Retrieval of Lake Erie Water Quality Parameters from Satellite Remote Sensing and Impact on Simulations with a 1-D Lake Model

by

Kiana Zolfaghari

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Geography

Waterloo, Ontario, Canada, 2016

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Lake Erie is a freshwater lake, and the most southern of the Laurentian Great Lakes in North America. It is the smallest by volume, the fourth largest in surface area (25,700 km^2), and the shallowest of the Laurentian Great Lakes. The lake's high productivity and warm weather in its watershed has attracted one-third of the total human population of the Great Lake's basin. The industrial and agricultural activities of this huge population has caused serious environmental problems for Lake Erie namely harmful algal blooms, dissolved organic/inorganic matters from river inputs, and sediment loadings. If these sorts of water contaminations exceed a certain level, it can seriously influence the lake ecosystem. Hence, an effective and continuous water quality monitoring program is of outmost importance for Lake Erie.

The use of Earth observation satellites to improve monitoring of environmental changes in water bodies has been receiving increased attention in recent years. Satellite observations can provide long term spatial and temporal trends of water quality indicators which cannot be achieved through discontinuous conventional point-wise in situ sampling. Different regression-based empirical models have been developed in the literature to derive the water optical properties from a single (or band ratio of) remote sensing reflectance (radiance). In situ measurements are used to build these regressions. The repeated in situ measurements in space and/or time causes clustered and correlated data that violates the assumption of regression models. Considering this correlation in developing regression models was one of the topics examined in this thesis. More complicated semi-analytical models are applied in Case II waters, aiming to distinguish several constituents confounding water-leaving signals more effectively. The MERIS neural network (NN) algorithms are the most widely used among semi-analytical models. The applicability of these algorithms to derive chl-a concentration and Secchi Disk Depth (SDD) in Lake Erie was assessed for the first time in this thesis. Satellite-observations of water turbidity were then coupled with a 1-D lake model to improve its performance on Lake Erie, where the common practice is to use a constant value for water turbidity in the model due to insufficient in situ measurements of water turbidity for lakes globally.

In the first chapter, four well-established MERIS NN algorithms to derive chl-a concentration as well as two band-ratio chl-a related indices were evaluated against in situ measurements. The investigated products are those produced by NN algorithms, including Case 2 Regional (C2R), Eutrophic (EU), Free University of Berlin WeW WATER processor (FUB/WeW), and CoastColour (CC) processors, as well

as from band-ratio algorithms of fluorescence line height (FLH) and maximum chlorophyll index (MCI). Two approaches were taken to compare and evaluate the performance of these algorithms to predict chl-a concentration after lake-specific calibration of the algorithms. First, all available chl-a matchups, which were collected from different locations on the lake, were evaluated at once. In the second approach, a classification of three optical water types was applied, and the algorithms' performance was assessed for each type, individually. The results of this chapter show that the FUB/WeW processor outperforms other algorithms when the full matchup data of the lake was used (root mean square error (RMSE) = 1.99 mg m^{-3}, index-of-agreement (I_a) = 0.67). However, the best performing algorithm was different when each water optical type was investigated individually. The findings of this study provide practical and valuable information on the effectiveness of the already existing MERIS-based algorithms to derive the trophic state of Lake Erie, an optically complex lake.

Unlike the first chapter, where physically-based and already trained algorithms were implemented to evaluate satellite derived chl-a concentration, in the next chapter, two lake-specific, robust semi-empirical algorithms were developed to derive chl-a and SDD using Linear Mixed Effect (LME) models. LME considers the correlation that exists in the field measurements which have been repeatedly performed in space and time. Each developed algorithm was then employed to investigate the monthly-averaged spatial and temporal trends of chl-a concentration and water turbidity during the period of 2005-2011. SDD was used as the indicator of water turbidity. LME models were developed between the logarithmic scale of the parameters and the band ratio of B7:665 nm to B9:708.75 nm for $\log_{10}$chl-a, and the band ratio of B6:620 nm to B4:510 nm for $\log_{10}$SDD. The models resulted in RMSE of 0.30 for $\log_{10}$chl-a and 0.19 for $\log_{10}$SDD. Maps produced with the two LME models revealed distinct monthly patterns for different regions of the lake that are in agreement with the biogeochemical properties of Lake Erie.

Lastly the water turbidity (extinction coefficient; K_d) of Lake Erie was estimated using the globally available satellite-based CC product. The CC-derived K_d product was in a good agreement with the SDD field observations (RMSE=0.74 m^{-1}, mean bias error (MBE)=0.53 m^{-1}, I_a=0.53). CC-derived K_d was then used as input for simulations with the 1-D Freshwater Lake (FLake) model. An annual average constant K_d value calculated from the CC product improved simulation results of lake surface water temperature (LSWT) compared to a "generic" constant value (0.2 m^{-1}) used in previous studies (CC lake-specific yearly average K_d value: RMSE=1.54 ºC, MBE= -0.08 ºC; generic constant K_d value: RMSE=1.76 ºC, MBE= -1.26 ºC). Results suggest that a time-independent, lake-specific, and constant

K_d value from CC can improve FLake LSWT simulations with sufficient accuracy. A sensitivity analysis was also conducted to assess the performance of FLake to simulate LSWT, mean water column temperature (MWCT) and mixed layer depth (MLD) using different values of K_d. Results showed that the model is very sensitive to the variations of K_d, particularly when K_d value is below 0.5 m^{-1}. The sensitivity of FLake to K_d variations was more pronounced in simulations of MWCT and MLD. This study shows that a global mapping of the extinction coefficient can be created using satellite-based observations of lakes optical properties to improve the 1-D FLake model.

Overall, results from this thesis clearly demonstrate the benefits of remote sensing measurements of water quality parameters (such as chl-a concentration and water turbidity) for lake monitoring. Also, this research shows that the integration of space-borne water clarity (extinction coefficient) measurements into the 1-D FLake model improves simulations of LSWT.

Acknowledgements

I would like to express my deepest appreciation and gratitude to my supervisor, Prof. Claude Duguay, for his understanding, all continuous support, encouragement, patience and expert guidance throughout my PhD experience. Without his mentoring and incredible counsel it would have never been possible to conduct this research.

I also would like to express my appreciation to the members of my PhD committee, Dr. Tiit Kutser, Dr. Richard Kelly, Dr. Ellsworth LeDrew, and Dr. Ralph E.H. Smith for serving as my committee members and their constructive comments. I also want to thank Dr. Caren Binding from Environment Canada, for providing me with the unique opportunity of joining her research group on Lake Erie *Limnos* cruise.

I would like to acknowledge my colleagues in "*Duguay Research Group*" for providing support and friendship. I also want to thank all the staff and many faculty members in the Department of Geography and Environmental management for providing me with the administrative and technical support during the course of my PhD.

Last but not least, I would like to express my heart-felt gratitude to my parents for supporting me throughout my entire life, in particular, for being a constant source of love, concern, and strength during this endeavor. I doubt that I will ever be able to convey my appreciation fully, but I owe them everything. I am so thankful to my brother, Kaveh, who is always with me, not only spiritually, but also physically. My great gratitude also goes to my other siblings, Kimia and Qmars, for their perpetual love and heart-warming words. Finally, my sincere gratefulness to all my close friends for their whole hearted support and care.

This work was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to Prof. Claude Duguay.

Dedication

To my beloved mother, Tahereh Dabidian and amazing father, Mohammad Ali Zolfaghari, for their all endless efforts and dedications to help me make my dreams come into reality.

Table of Contents

AUTHOR'S DECLARATION ii
Abstract iii
Acknowledgements vi
Dedication vii
Table of Contents viii
List of Figures xi
List of Tables xiv
List of Abbreviations xvi
Chapter 1 General Introduction 1
1.1 Preface 1
1.2 Motivation 2
1.3 Objectives 3
1.4 Thesis Contribution 3
1.5 Thesis Structure 4
Chapter 2 Background 6
2.1 Introduction 6
2.2 Lake Water Quality 6
2.2.1 Physical and Chemical Water Properties 7
2.2.2 Biological Water Properties 11
2.3 Field Measurements of Water Quality 13
2.3.1 Chlorophyll-a Measurement 13
2.3.2 Turbidity Measurement 16
2.4 Satellite Remote Sensing of Water Quality 18
2.4.1 Remote Sensing Algorithms 19
2.4.2 Past and Current Remote Sensing Satellites 28
2.5 FLake Model 31
2.6 Study Site and Water Sampling Protocol 33
2.7 Summary 35
Chapter 3 Evaluation and Blending of MERIS Chlorophyll-a Retrieval Algorithms for Optically Complex Lake Erie 37
3.1 Introduction 37

3.2 Data and Methods 39
3.2.1 Study Area 39
3.2.2 Algorithms Evaluated 41
3.2.3 Processing of MERIS data 44
3.2.4 Clustering Method Using Remote Sensing Reflectance 45
3.2.5 Chlorophyll-a Retrieval Algorithms Calibration and Error Definitions 46
3.3 Results and Discussion 47
3.3.1 In situ Measurements 47
3.3.2 Matchup Data Result 48
3.3.3 Chlorophyll-a Products Evaluation 49
3.3.4 Clustering Analysis Results 53
3.3.5 Chlorophyll-a Product Evaluation in Different Water Types and Blending Algorithm Results 57
3.4 Summary and Conclusion 59
Chapter 4 Estimation of Water Quality Parameters in Lake Erie from MERIS Using a Linear Mixed Effect Model 62
4.1 Introduction 62
4.2 Material and Methods 64
4.2.1 Study Site 64
4.2.2 Field Measurements of Water Quality Parameters 66
4.2.3 Satellite Data and Processing 66
4.2.4 Water Quality Parameters Algorithms 67
4.2.5 Accuracy Assessment 68
4.3 Results and Discussion 69
4.3.1 Lake Erie as an Optically Complex Water Body 69
4.3.2 Linear Mixed Effect Model Calibration 71
4.3.3 Evaluation of Linear Mixed Effect Models 73
4.3.4 Spatial and Temporal Variability of Chl-a and SDD 74
4.3.5 Uncertainties of the Applied Linear Mixed Effect Model on MERIS 79
4.4 Conclusion 80
Chapter 5 Satellite-Derived Light Extinction Coefficient and its Impact on Water Temperature Simulations in a 1-D Lake Model 82

5.1 Introduction 82
5.2 Data and Methods 84
5.2.1 Study Site and Station Observations 84
5.2.2 Satellite Solar Irradiance Model 86
5.2.3 Satellite-Derived Extinction Coefficient 87
5.2.4 FLake Model and Configuration 88
5.2.5 Accuracy Assessment 89
5.3 Results and Discussion 89
5.3.1 Evaluation of Modeled Solar Irradiance Data 89
5.3.2 Evaluation of Satellite-Derived K_d 91
5.3.3 FLake Model Results 96
5.4 Summary and Conclusion 104
Chapter 6 General Conclusions 106
6.1 Overall Summary 106
6.2 Limitations 107
6.3 Future Research Directions 110
References 113

List of Figures

Figure 2-1 Density of water (and ice) as a function of temperature. Note maximum density of water at 4ºC. (Data from Pauling 1953 and Hutchinson 1957). 8
Figure 2-2 Lake stratification and turnover (Redrawn after Wetzel, 1975) 9
Figure 2-3 From left: Approximate absorption spectrum of colored dissolved organic matter (a_{CDOM}); Absolute values for sea water absorption (a) and backscattering (b_b); Relative values (normalized at 440 nm) of absorption (a) and scattering (b) for chl-a in phytoplankton (Robinson, 2004). 19
Figure 2-4 Overview of recently (2007-2014) published papers on remote sensing of chl-a for optically complex water bodies. The algorithms retrieve chl-a from satellite imagery using matchup-validated empirical and semi-analytical algorithms. Arrows with more than one color are indicative of studies that utilized more than one sensor. Starred references indicate those studies that include Lake Erie. C2R, EU, BL, FUB/WeW are the NN-based semi-analytical algorithms available in the BEAM software and are explained in more details in section 3.2.2.1. 27
Figure 2-5 Overview of recently (2007-2015) published papers on remote sensing of SDD for optically complex water bodies. The algorithms retrieve SDD from satellite imagery using matchup-validated empirical and semi-analytical algorithms. Arrows with more than one color are indicative of studies that utilized more than one sensor. Starred references indicate those studies that include Lake Erie. Studies in C2R, EU, BL are the NN-based semi-analytical algorithms available in the BEAM software and are explained in more details in section 3.2.2.1. 28
Figure 2-6 Location map and shape of the Great Lakes 34
Figure 3-1 Location of Lake Erie in the Laurentian Great Lakes (inset), and the sampled points of the Canadian Coast Guard ship *Limnos* during 2004 to 2012, in situ data used in this study. 41
Figure 3-2 Chl-a retrieved from each algorithm after tuning for the full lake relative to the matchup in situ data. Solid line corresponds to 1:1 relationship. 53
Figure 3-3 Chl-a concentration mapping by locally tuned FUB/WeW algorithm during a bloom event in September 3, 2011. 53
Figure 3-4 Matchup in situ data classified into three clusters from the two-step clustering approach. Black lines are the individual reflectance data; colorful lines are the reflectance means of the three optical water types as illustrated by the legend. (Rrs: Remote Sensing Reflectance). 55
Figure 3-5 Comparing matchup in situ data characteristics across the three optical water types. 56
Figure 3-6 Distribution of matchup in situ data for different optical water types. 57

Figure 3-7 Chl-a concentration derived from blending algorithm versus in situ measurements for all matchup in situ data over Lake Erie. Solid line corresponds to a 1:1 relation. 58
Figure 4-1 Location of Lake Erie and its boundary (Canada and US). In situ sampling stations from cruises that took place in September 2004, May, July, and September 2005, May and June 2008, July and September 2011, and February 2012 are illustrated by empty triangles. 65
Figure 4-2 Relationships between in situ SDD and three bio-optical parameters of the water: chl-a (A), TSM (B), and $a_{CDOM}(440)$(C). 70
Figure 4-3 Correlation coefficients between MERIS red/NIR water-leaving reflectance ratio and in situ chl-a (left) and between MERIS water-leaving reflectance ratio and in situ SDD (right). R1 and R2 represent nominator and denominator, respectively. Values along the diagonal line from lower left to top right indicate correlation with reflectance of a single wavelength. 72
Figure 4-4 Comparison between MERIS estimates of chl-a (left) and SDD (right) using LME models and in situ measurements for Lake Erie. The solid diagonal line is the 1:1 line. 74
Figure 4-5 Chl-a average (Avg, left) and standard deviation (St.Dev., right) (log10 scale) from March to October for the study period (2005-2011). 77
Figure 4-6 SDD average (Avg, left) and standard deviation (St.Dev., right) (log10 scale) from March to October for the study period (2005-2011). 78
Figure 4-7 Bathymetry of Lake Erie (source: NOAA). 79
Figure 5-1 Maps showing Lake Erie in Laurentian Great Lakes and the location of stations where different parameters were measured. 85
Figure 5-2 Comparison of daily variations of NWRI-EC observations versus SUNY solar irradiance model: a) April - September 2004, b) April - October 2008. 90
Figure 5-3 Scatter plot of NWRI-EC and SUNY mean daily solar irradiance (data from 2004 and 2008). The obtained statistical indices are included. The dashed line shows the best-fit line. Solid line corresponds to 1:1 relationship. 91
Figure 5-4 Relation between satellite-derived K_d and in situ SDD matchups 93
Figure 5-5 Spatial variation of satellite-derived K_d in Lake Erie, on 3 September 2011. Location of NDBC station is shown on the map as a solid dot. 94
Figure 5-6 Temporal and spatial variation of satellite-derived K_d in Lake Erie for different months of a year: May- August 2010. Location of NDBC station is shown on the map as a solid dot. 94
Figure 5-7 Temporal and spatial variation of K_d in Lake Erie during May of two consecutive years: 2008 and 2009. Location of NDBC station is shown on the map as a solid dot. 95

Figure 5-8 Variations of CC-derived K_d for the selected location during the study period (2003-2012). 96
Figure 5-9 LSWT simulation results for 2005 - 2007; from: CRCM-12.6, CRCM-20, CC-derived average for K_d during selected month of each year (0.81, 0.71, and 0.73 m^{-1}; respectively), and the merged simulations based on each month average K_d. The corresponding observations for LSWT, and CC-derived K_d values are also plotted. 98
Figure 5-10 Modeled (y-axis) versus observed (x-axis) LSWT for yearly average, merged, CRCM-12.6, and CRCM-20 simulations during the ice-free seasons in 2005-2007. A linear fit (dashed line) and its coefficients are shown on the plot. The statistics related to the regression of parameters, and a 1:1 relationship (solid line) are also shown. 100
Figure 5-11 LSWT and MWCT simulation results in 2008, when using the lowest (Min), average (Avg), and the highest (Max) K_d values. Results from the CRCM-12.6 simulation is also plotted. 103
Figure 5-12 MLD simulation results for the lowest (Min), average (Avg), and the highest (Max) K_d values in 2008. CRCM-12.6 results are also plotted. 104

List of Tables

Table 2-1 Summary of the key variables and their inter-relationships (Protocol Manual for Water Quality Sampling in Canada, 2011) 7
Table 2-2 Examples of point, diffuse, and in-stream sources of nutrients (Shaw et al., 2009) 11
Table 2-3 Most frequently used satellite data products for water quality 30
Table 3-1 Chl-a, TSM and CDOM training ranges of the NN processors (source: Palmer et al. (2014); Ruescas et al. (2014)). 43
Table 3-2 Flags of excluded pixels 45
Table 3-3 Descriptive statistics of full range in situ data *Limnos* collected from 2004 to 2012 on Lake Erie. 47
Table 3-4 Descriptive statistics of in situ matchup subset data after removing those flagged in each algorithm. Chl-a and TSM concentrations are reported in mg m^{-3} and g m^{-3}. CDOM absorption in m^{-1}. 49
Table 3-5 Locally tuned equations for the full lake derived in training step. N is the number of training data which is 70% of the matchup in situ data after removing those flagged by the algorithm. Chl_aprodcut is the product of each algorithm related to the chl-a concentration, whereas Chl_apredicted is the one derived from applying the locally tuned equations. 50
Table 3-6 Chl-a retrieval performance statistical indices for each algorithm derived in testing step for the full lake. Chl_a predicted derived from tuning equation is compared to the in situ chl-a concentration. N is 30% of matchup in situ data after removing those flagged by the processor. 50
Table 3-7 Descriptive statistics of in situ matchup data optical properties across three different water types. N is the number of matchup in situ data that is assigned to each water type. 56
Table 3-8 Performance of the models calibrated for each single water type. 58
Table 3-9 Performance of the calibrated models for the full lake, and the blending algorithm. N is the number of matchup in situ data after removing those flagged ones in each algorithm. 59
Table 4-1 Lake Erie Basins Information (source: (Painter et al., 2000)) 65
Table 4-2 Flags of excluded pixels 67
Table 4-3. Descriptive statistics of in situ measurements for Lake Erie (2004-2012). N is the number of times samples were collected at stations. St. dev. is standard deviation. Chl-a and TSM are in mg m^{-3}, and g m^{-3}, respectively; a_{CDOM} in m^{-1}, and SDD in m. 69
Table 5-1 Flags of excluded pixels 88

Table 5-2 CC-derived average values of K_d for each month (2005-2007). The values correspond to the time of year when water LSWT observations, as well as the CC derived K_d values, are available. 96
Table 5-3 Simulated LSWT compared to in situ observations (2005 – 2007). Period corresponds to the time of year when LSWT and K_d values were available. 99

List of Abbreviations

AIC	Akaike Information Criterion
AOP	Apparent Optical Property
BEAM	Basic ERS & ENVISAT (A)ATSR MERIS
BIC	Bayesian Information Criterion
CC	CoastColour
CDOM	Colored Dissolved Organic Matters
Chl-a	Chlorophyll-a
CPA	Color-Producing Agent
CRCM	Canadian Regional Climate Model
ESA	European Space Agency
FLake	Freshwater Lake model
FLH	Fluorescence Line Height
GOES	Geostationary Operational Environmental Satellites
HIRLAM	High Resolution Limited Area Model
HyspIRI	Hyperspectral Infrared Imager
I_a	Index of Agreement
IOCCG	International Ocean Colour Coordinating Group
IOP	Inherent Optical Property
LDCM	Landsat Data Continuity Mission
LME	Linear Mixed Effect
LSWT	Lake Surface Water Temperature
MBE	Mean Bias Error
MCI	Maximum Chlorophyll index
MERIS	MEdium Resolution Imaging Spectrometer
MLD	Mixed Layer Depth
MODIS	Moderate-resolution Imaging Spectrometer
MSI	Multispectral Imager
MWCT	Mean Water Column Temperature
NDBC	National Data Buoy Center
NEMO	Nucleus for European Modelling of the Ocean
NLET	National Laboratory for Environmental Testing
NN	Neural Network
NOAA	National Oceanic and Atmospheric Administration
NREL	National Renewable Energy Laboratory
NSRDB	National Solar Radiation Database
NWP	Numerical Weather Prediction
NWRI	National Water Research Institute
OCC	Ontario Climate Center
OLCI	Ocean and Land Colour Instrument
OLI	Operational Land Imager
QAA	Quasi Analytical Algorithm
RCM	Regional Climate Model
RMSE	Root Mean Square Error
SDD	Secchi Disk Depth
SeaWiFS	Sea-Viewing Wide Field-of-View Sensor
TSM	Total Suspended Matters

Chapter 1
General Introduction

1.1 Preface

In addition to a general introduction, a background chapter and a general conclusion, the thesis contains three journal articles that examine the applicability of remote sensing observations to retrieve the bio-optical properties of water bodies and how these observations can improve water temperature simulations in a lake model. The first paper, submitted to the international journal *Remote Sensing of Environment*, presents an extensive evaluation and comparison of different existing chlorophyll-a (chl-a) retrieval algorithms that are intended for use with MEdium Resolution Imaging Spectrometer (MERIS) data, and made available in the Basic ERS & ENVISAT (A)ATSR MERIS (BEAM) toolbox. A two-step clustering method was also applied to derive different existing optical classes in the lake. The applicability of the algorithms for each water type was evaluated. The second paper, submitted to the international journal *Remote Sensing*, presents the development and assessment of a new algorithm based on a Linear Mixed Effect (LME) model to derive chl-a concentration and Secchi Disk Depth (SDD). The third paper, submitted to the international journal *Hydrology and Earth System Sciences*, evaluated a globally available MERIS product for water turbidity. The satellite-derived water turbidity was applied in a lake model, the 1-D Freshwater Lake (Flake) model, to improve the results of simulated lake surface water temperature (LSWT). The sensitivity of the FLake model to water turbidity was also investigated based on simulated LSWT, mean water column temperature (MWCT), and mixed layer depth (MLD).

The work presented in this thesis was conducted under the supervision of Professor Claude Duguay who aided a great deal with the initial proposal of each article, and also provided funding for the research. Dr. Caren Binding (Environment Canada) provided the in situ data of Lake Erie and also the opportunity to join her research group on the Lake Erie *Limnos* Cruise. Dr. Ram Yerubandi, also from Environment Canada, provided meteorological data for Lake Erie. All materials in composition of the original articles provided in the thesis are the sole production of the primary investigator listed as first author in the journal publications. The research presented in this thesis is the direct result of a collaboration with the listed co-authors. Dr. Homa Kheyrollah Pour, from the "Duguay Research Group", supported this research through comments and advice related to running the FLake model. Dr. Daniel Odermatt, from "Odermatt & Brockmann GmbH", provided comments and advice about the algorithms available in the Basic Envisat and ERS (A)ATSR and MERIS toolbox (BEAM). The manuscripts were edited for content and composition by the co-authors.

1.2 Motivation

The Laurentian Great Lakes are a vital natural resource that contains 18% of the Earth's potable freshwater (Reynold, 1996). Among these lakes, Lake Erie is the smallest by volume but yet the most productive of them due to its shallowness and warm temperature (Munawar and Weisse, 1989; IJC. International Joint Commission Canada and United States, 2013). Due to the high productivity of the lake, a large population density is living in its watershed (11.6 million people (IJC. International Joint Commission Canada and United States, 2013)); and a large percentage of agricultural lands exists in its drainage basin (63% of the lake's watershed (IJC. International Joint Commission Canada and United States, 2013)). Therefore, the local population on Lake Erie basin are in the substantial dependency on the lake for providing their economical, recreational and industrial needs (Gobler and Wilhelm, 2015). These human activities severely threaten the water quality of Lake Erie. Municipal sewage treatment plants contribute to nutrient loadings, particularly phosphorus. There are other anthropogenic sources as well, but diffuse runoff from rural and urban lands is today a leading factor among these sources (IJC. International Joint Commission Canada and United States, 2013). Excess nutrients enrichment, known as eutrophication, is particularly evident in Lake Erie. Lake Erie is experiencing profound changes due to eutrophication including severe and frequent algal blooms (IJC. International Joint Commission Canada and United States, 2013). It is the most subjected to sediment loading among the Great Lakes. The fine sediments covering the lake bottom can be easily distributed when the shallow lake is stirred up by winds, therefore degrading water clarity (Lake Erie LaMP Work Group, 2004).

Aquatic ecosystem management can control and decrease the environmental stresses with an effective monitoring of water quality parameters on Lake Erie. Water temperature, suspended and dissolved materials, algae growth, and turbidity are among water quality parameters that can impact submerged ecosystem health. In situ surveys allow for measurement of these parameters along depth profiles. However, the most important drawback of relying only on the field work in the water quality monitoring programs is the technical and financial supports to operate these surveys in adequate spatial and temporal coverage. Remote sensing methods can complement in situ sampling. Over the last few decades, remote sensing observations have revolutionized research especially in remote high latitude regions. Satellite imagery permits water quality monitoring from local to global scales, while field measurements cannot operate the in situ data collection in the required temporal and spatial scales. However, field measurements are still vital for developing remote sensing algorithms, and also for a precise interpretation of satellite-derived optical data. There is an urgent need to develop retrieval algorithms of lake optical properties from recent and current satellite missions, and evaluate them against available in situ measurements for both lake monitoring and modeling purposes.

The MEdium Resolution Imaging Spectrometer (MERIS) on board the European Space Agency (ESA)'s Envisat platform collected data from March 2002 until April 2012. The observations were provided in 15 spectral bands (spanning the reflective solar spectral range from 390 to 1040 nm), 16 bit radiometric resolution, 300 m full spatial resolution, and with a three-day revisit cycle at the equator. MERIS resolutions is more appealing than other current satellites sensors for monitoring aquatic ecosystems. The band centers and widths were designed for an improved insight into water optical properties. The 10-year MERIS archived data is an imperative source for studies on lake systems dynamics at regional and global scales. The in-water algorithms are developed and validated to explore the capacity of MERIS observation to accurately retrieve the biogeochemical properties of lakes. Although MERIS is not active anymore, the forthcoming Sentinel-3 Ocean and Land Colour Instrument (OLCI) will provide continuity for global monitoring of lakes. The validation of available MERIS products is of key importance in order to develop algorithms for the upcoming Sentinel-3 OLCI sensor which has MERIS heritage and improves upon the data collected by MERIS (Palmer et al., 2014). The motivation of this thesis is to examine the applicability of remote sensing methods for the identification of important water quality parameters that can assist a great deal with the monitoring and preserving the important ecological system of Lake Erie. This work also aims to improve the results of a lake model through integration of space-borne water clarity measurements.

1.3 Objectives

The overall goal of this research is to evaluate and improve chlorophyll-a (chl-a) concentration and water turbidity retrieval algorithms from MERIS satellite data on Lake Erie. It also aims to investigate how the lake-specific water turbidity derived from satellite observations can improve simulations of LSWT from a lake model (Freshwater Lake or FLake model), where the common practice is to use a constant value for water turbidity in FLake. The specific objectives of this thesis are to: 1) evaluate the applicability of different existing MERIS algorithms to derive water optical parameters of interest in different optical classes of lake, both individually and all together at once; 2) develop a semi-empirical algorithm using regression methods based on a linear mixed effect (LME) model approach to estimate chl-a concentration and water turbidity from MERIS satellite observations and evaluate the results against the field measurements; and 3) investigate the improvement of the one-dimensional (1-D) FLake model to simulate lake water surface temperature through integration of satellite-derived turbidity from the globally available satellite-based CoastColour product.

1.4 Thesis Contribution

In this thesis, the performance of four well-established MERIS semi-analytical NN-based algorithms to derive chl-a concentration in Lake Erie is evaluated. These algorithms are the most widely used of the semi-

analytical techniques due to their availability within the publically available BEAM toolbox. The lessons learned from this evaluation (i.e. strengths and limitations of the algorithms) will help improve the processors being developed for the future Sentinel-3 OLCI sensor, which will provide continuity to MERIS data.

A LME model approach is also employed to develop regression-based empirical equations for deriving chl-a concentration and SDD in Lake Erie. There are different band ratio (or single band) algorithms used in the literature for deriving water optical properties. These models are based on developing a statistical regression between satellite band ratios (or single bands) and the optical properties. The in situ data required for calibrating these regression algorithms are collected repeatedly in space and/or time, hence representing clustered or dependent data which violates the assumption of regression models. Therefore, in this thesis, a LME model approach is employed as the regression method. This approach is more robust statistically than the common regression-based equations developed in previous investigations on remote sensing of water quality parameters.

The last significant contribution of this thesis is in the improvement of simulations with the 1-D FLake model. Water turbidity is one of the input parameters of the model to describe the optical properties of lakes. However, due to the scarcity of in situ measurements of this parameter in field monitoring programs, a "generic" constant value is often used to run the model. In this thesis, the potential of remote sensing observations to fill this data gap and improve the performance of the model is demonstrated. Satellite-derived water turbidity is used as input in FLake model simulations. Model output shows clear improvements in water temperature simulations for Lake Erie.

1.5 Thesis Structure

This manuscript-based thesis consists of six chapters. The current chapter presents the rationale and objectives of the thesis, outlining the need for monitoring chl-a concentration and water turbidity in Lake Erie, and coupling the satellite-based observations of lake water turbidity in a lake model.

Chapter 2 reviews a list of lake water quality parameters, and how their variations can affect lake ecosystems. Field measurement methods for parameters with optical properties, including chl-a concentration and turbidity, are covered. Different satellite-based algorithms developed to measure these variables over water bodies are then described. The most important available satellites that are commonly used to derive optical properties of water bodies are discussed. Background information about the requirement for a global mapping of lake turbidity in lake models and the applicability of integrating satellite-derived lake turbidity within 1-D lake models are provided. This chapter also reviews the

characteristics of the study site, Lake Erie, and how chl-a concentration and water turbidity are measured in the field for the lake.

Chapter 3 addresses the first objective of the thesis by evaluating existing MERIS algorithms to derive chl-a concentration and also a blended algorithm using in situ measurements collected by Environment Canada for Lake Erie.

Zolfaghari, K., Duguay, C.R., Odermatt, D. Evaluation of MERIS Chlorophyll-a Retrieval Algorithms for Optically Complex Lake Erie.

Chapter 4 addresses the second objective by developing a Linear Mixed Effect (LME) model to derive chl-a concentration and water turbidity from MERIS observations and evaluate them against the field measurements provided by Environment Canada for Lake Erie.

Zolfaghari, K., Duguay, C.R. Estimation of Water Quality Parameters in Lake Erie from MERIS Using a Linear Mixed Effect Model.

Chapter 5 addresses the improvement of the 1-D FLake model to simulate lake surface water temperature by integrating satellite-based measurements of water clarity through the extinction coefficient. This optical property is provided from the globally available CoastColour product and was evaluated with the field observations provided by Environment Canada on Lake Erie. The model results were compared against in situ measurements of surface water temperature collected by the National Data Buoy Center (NDBC) of NOAA.

Zolfaghari, K., Duguay, C.R. Kheyrollah Pour, H. Satellite-Derived Light Extinction Coefficient and its Impact on Water Temperature Simulations in a 1-D Lake Model.

The final chapter provides a summary of key findings of the research. It also presents some of the limitations of the thesis and future research directions.

Chapter 2
Background

2.1 Introduction

This chapter includes 5 sections, each section covering the pertinent background to the research in this study. Section 2.2 describes critically important physical, chemical, and biological water quality parameters that can be estimated using remote sensing methods. Section 2.3 is focused on field measurement methods of water quality parameters of interest in this thesis: chlorophyll-a (chl-a) concentration and turbidity. Different remote sensing satellites, and algorithms currently available to derive bio-optical properties of waterbodies are described in section 2.4. Section 2.5 discusses the necessity of considering the integration of satellite-derived optical parameter (extinction coefficient) for lake model simulations. A description of the general characteristics of Lake Erie and also the employed water quality in situ measurement methods for this research are provided in section 2.6. A summary of the key elements of this chapter is then given in section 2.7.

2.2 Lake Water Quality

Lakes play an important role in different applications including: urbanization (drinking water supplies, agriculture, transportation), industry (fishery, power generation), and recreation (Guan et al., 2012). The water quality of lakes is vital to human activities and needs. It also plays a critical role in sustaining regional ecosystems (Wang et al., 2011). The physical, chemical, and biological properties of lake water and the aquatic ecosystem can be used to describe the water quality and its suitability for a particular purpose such as: drinking, protection and maintenance of aquatic life, agricultural irrigation, and recreation. The physical and chemical properties are considered as a "snapshot" of water condition at the moment of sampling; while the biological factors show the combined impact of diverse water quality factors (Water Quality Task Group, 2006). The selection of these physical, chemical, and biological characteristics used to measure water quality depends on the water quality issues and water uses of interest (Water Quality Task Group, 2006). Dissolved organic and inorganic matters, total suspended solids, chl-a, dissolved oxygen, conductivity, pH, and temperature are among the parameters typically being measured (Salama et al., 2012).

Degradation in water quality is a result of environmental stresses from natural and anthropogenic sources. Naturally-sourced stresses cause variations in water quality resulting from changes in seasons, climate, type of soils and rocks the water flows through. However, human activities such as urban and industrial development, farming, mining, combustion of fossil fuels, animal-feeding operations can also change the quality of natural waters. There are many lake water quality monitoring programs designed to keep records

of changes in water quality indicators. They are based on the use of different standards and guidelines to classify water quality for designated uses. Summary of the physical, chemical, and biological water quality indicators, that the Canadian Council of Ministers of the Environment (CCME) referred them as the key parameters, are listed in Table 2-1. In this table, *variables* are a surrogate measurement for the *second variables*. Some of these parameters are measurable using satellite observations (as it will be covered in Section 2.4). More details of these variables are provided below. Other variables such as nitrogen and phosphorus concentrations, which are not regularly measured by satellite imagery directly, are also outlined here as the critically important water quality factors.

Table 2-1 Summary of the key variables and their inter-relationships (Protocol Manual for Water Quality Sampling in Canada, 2011)

Variable	Second Variable	Relationship
Specific Conductivity (SC)	Dissolved Solids (TDS)	SC and TDS are usually related for each waterbody. TDS is the sum of constituents such as chloride, sulphate, etc.
Turbidity	Suspended Solids (TSS)	Turbidity can be related to the amount of solids in suspension
Temperature	Dissoved Oxygen	The amount of oxygen in water increases with cooler temperatures
pH and Temperature	Ammonia	The toxicity of ammonia in water increases with higher pH and temperatures
Secchi Disk Reading	Turbidity, color, algae	Measures light penetration in a lake that is reduced by these factors
Chloride	Nitrite	The Toxicity of nitrite in water decreases with increased chloride
Hardness	Alkalinity	Alkalinity and hardness often have similar concentrations in water
Hardness, Dissolved Organic Carbon (DOC)	Metals	The toxicity of some metals (*e.g.*, copper, zinc) decreases with increasing hardness and DOC

2.2.1 Physical and Chemical Water Properties

2.2.1.1 Temperature

Effect on aquatic ecosystem health: Water temperature is an important factor in the aquatic ecosystem as it has a major influence on biological activities and growth of aquatic organisms. Different species such as zooplankton, phytoplankton, fish, and insects have a certain tolerance to temperature changes. Furthermore, temperature has a leading role in the chemical properties of water, as the chemical reactions accelerate at higher temperatures. Changes in water chemistry can accordingly have effects on submerged biology. As one of the chemical consequences of water temperature is its impact on oxygen levels. Warm water can hold less oxygen compared to cold water. So it may be saturated but not holding enough oxygen for living species. Another impact of water temperature on submerged ecosystems is that some of compounds existing in water, such as ammonia, are more toxic for living species at higher temperatures (Wetzel, 2001). Water

temperature has become an important indicator of climate change that can impact various habitats in or out of aquatic systems (Peters et al., 2011).

Reasons for variations: Water temperature can change naturally in different days, and seasons. Daily changes are more noticeable at the surface which are warmer during the day and cooler at night. Layers of water with different temperatures will form in deep lakes in summer and winter, because of the unique relationship between water temperature and its density (water reaches its maximum density at 4^{o}C (Figure 2-1)). This process is called thermal stratification. In the fall, the surface water begins to cool and sink. Also, during spring, the increasing density of the warming water causes this surface layer to sink. Accordingly, when the layers reach similar temperature in spring and fall, wind forces are able to turbulently mix the whole layers from top to bottom. This process is referred to as lake turnover (Figure 2-2). Stratification and layer mixing associated with a turnover will affect the physical, and chemical characteristics developed in different layers, that eventually will have an effect on biological communities (Golterman, 1975; Illinois Environmental Protection Agency, 2014).

Light can influence water temperature as it decreases exponentially with depth. Furthermore, lake temperature is affected by the lake size and temperature of inflows into the lake.

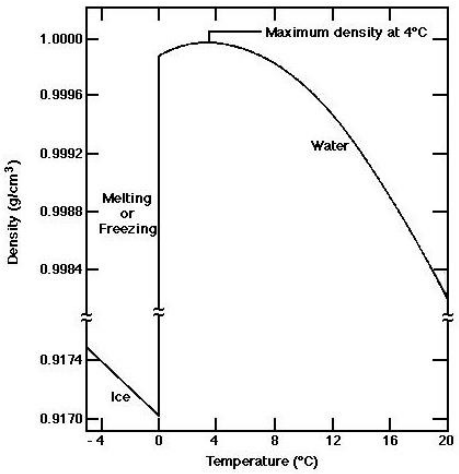


Figure 2-1 Density of water (and ice) as a function of temperature. Note maximum density of water at 4^{o}C. (Data from Pauling 1953 and Hutchinson 1957).

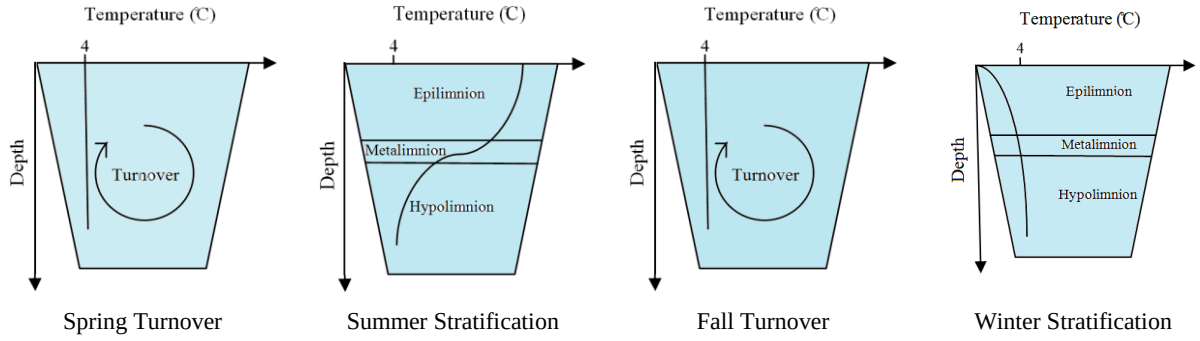


Figure 2-2 Lake stratification and turnover (Redrawn after Wetzel, 1975)

2.2.1.2 Total suspended material

Effect on aquatic ecosystem health: Suspended materials (organic or inorganic) including silt, plankton, industrial materials, and sewage enter into the lake water. High concentration of total suspended materials reduces water clarity and prevents light penetration. As a result, photosynthesis rate and its by-product, oxygen, will decrease. Furthermore, as total suspended matters (TSM) absorbs heat, increasing the temperature is another cause for a lake to hold low levels of dissolved oxygen. Therefore, the waterbody will lose its ability to support living organisms. Suspended solids can disrupt natural movements and migration of species in water. Eventually, deposited solids will alter the bottom of the lake which can have an impact on fish habitat. Reduced growth rates, lowered resistance to disease, and clogged gills are direct harms of total suspended materials on fish (Vincent, 2008).

Reasons for variations: Factors affecting the concentration of suspended solids in lake include: waste water discharges, industrial wastes, erosion from urban and agricultural lands, increased growth of phytoplankton, and high population of bottom feeders that can stir up the deposited suspended solids at the bottom of the lake. Also, turbulence caused by wind in shallow lakes or increased stream discharges can re-suspend them (Vincent, 2008).

2.2.1.3 Total dissolved material

Effect on aquatic ecosystem health: Total dissolved solids refer to any organic and inorganic material dissolved in water. Inorganic salts include principally calcium (ca^{2+}), magnesium (mg^{2+}), potassium (K^{+}), sodium (Na^{+}), bicarbonates (HCO_3^{-}), chlorides (Cl^{-}), and sulfates (SO_4^{2-}). High concentration of dissolved solids in water can cause changes in turbidity, taste, and odor of lake water. Turbidity will cause increase in temperature and decrease of photosynthesis rate in water as a result of less light penetration. Thus less dissolved oxygen levels in water, and high temperature will lead to death of many aquatic organisms. Furthermore, possible changes in pH caused by dissolved ions in water can affect the overall health of

species. High concentration of dissolved salts can dehydrate the animal skins. Cloudy ice cubes, softer and faster melting ice can be other consequences of dissolved solids in water (Bronmark and Hansson, 2005).

Reasons for variations: These ions and organic matters can originate from sedimentary rocks that water is flowing over, acidic rainfall, silt, sewage, urban and agricultural runoff, and industrial wastewater (Bronmark and Hansson, 2005).

2.2.1.4 Turbidity

Effect on aquatic ecosystem health: The clarity of water is defined using turbidity. Phytoplankton is the major source of turbidity in open water zones, whereas in areas closer to shore, detritus, and silt can also have an influence on water clarity. Phytoplankton is the microscopic plants living in water. The detritus (organic material) is from dead algae, higher plants, zooplankton, bacteria, fungi, etc. that can be produced within the water column, or watershed and washed into the lake. Silt (inorganic or mineral sediments) comes mostly from shoreline erosion and from the re-suspension of bottom sediments due to wind mixing. High turbidity can modify light penetration into the lake, and decrease the photosynthesis. The reduced photosynthesis will result in a lower food production, and oxygen release which are required for other living organisms. Turbidity effect on phytoplankton growth is complex and depends on the period of a specified turbidity status as well. For instance, very high levels of turbidity in a short period of time can be less of a problem compared to a lower level of turbidity that remains longer. Moreover, high concentration of particulate in water can fill in shallow lakes and the spaces between rocks that can be used by aquatic organisms as habitat (Wetzel, 2001; Canadian Council of Ministers of the Environment, 2003).

Reasons for variations: Phytoplankton growth changes in different seasons. It starts to increase in spring with warmer temperature until growth peaks in summer for shallow lakes. As a result, less water clarity is expected in these seasons compared to winter that cooler weather will decrease phytoplankton growth. However, the water clarity increase in fall and winter will be decelerated due to more winds and water column sediments mixes in shallow lakes. Also, in deep lakes experiencing stratification in summer, turbidity may increase in fall turnover. As the surface water starts to cool down in fall, lake layers mix and the nutrients release from the lower layer of the lake may cause an algal bloom which results in decreasing water clarity. Furthermore, various organic and inorganic materials existing in the lake vary depending on the hydrological events such as storms and snowmelt to wash them into the lake, and also wind speed to stir up the lake bottom (Wetzel, 2001; Canadian Council of Ministers of the Environment, 2003).

2.2.1.5 Phosphorus, Nitrogen

Effect on aquatic ecosystem health: A limiting nutrient is a chemical necessary for plant growth, however, it is available in smaller amount that is required for algae to increase their abundance. Phosphorus and

nitrogen are major nutrients that can limit biological productivity in a waterbody; whereas other factors such as availability of light can also be a limiting environmental factor. These nutrients in aquatic ecosystem can occur in the form of organic or inorganic, dissolved or suspended in the water (Florida Lakewatch, 2000).

Eutrophication is a process, both natural and anthropogenic, of additional nutrients in waterbodies including: lakes, rivers, estuaries, and oceans. It will result in changes of primary production and species composition in the waterbody. The natural process of aging in lakes will result in a eutrophic status over a long (geological) time scale. On the other hand, industrial activities can be a reason of anthropogenic nutrient input to waterbodies as to raise the trophic status promptly, which is referred to as "cultural eutrophication". The eutrophication process starts from poor nutrient state (oligotrophic), through mesotrophic state of having additional nutrients, to a final eutrophic status of waters where water quality decreases as a result of excess nutrients available in both water and sediments. One important consequence of eutrophication is the algae bloom that can oxygenate the water during the assimilation of available nutrients through photosynthesis. On the other hand, large quantities of dead algae sinking to the bottom of the lake will be decomposed by aerobic bacteria consuming oxygen. Low levels of available oxygen will lead to the death of fish and other aquatic organism. Moreover, oxygen depletion will establish the activity of anaerobic bacteria in bottom layers that can produce toxic gases harmful for lake ecosystem (Kondratyev, 1999; Shaw et al., 2009).

Reasons for variations: The contribution of nutrients can be from either point or diffuse as the external sources or from internal inputs (Table 2-2) (Shaw et al., 2009). Agriculture, domestic and industrial wastes, and phosphorus-rich lake bottom sediments are the most important sources of nutrient contribution into the lake.

Table 2-2 Examples of point, diffuse, and in-stream sources of nutrients (Shaw et al., 2009)

Point Sources	Diffuse Sources	In-stream Sources
Sewage treatment plants	Storm runoff from rural land	Release of nutrients from bottom sediments
Piggeries, Feedlots, Dairies	Urban runoff	Stream bank collapse
Industrial effluents	Groundwater discharge	The seasonal mixing of phosphorus- enriched water from deeper layers in lakes or reservoirs with surface waters
Irrigation drains	Atmospheric fall-out	

2.2.2 Biological Water Properties

2.2.2.1 Phytoplankton

Effect on aquatic ecosystem health: Algae are photosynthetic aquatic organisms and the range of their size is from micro unicellular to giant multicellular. Lack of stems, roots, and leaves differentiate them from

plants. There are three main accessory pigments available in different types of algae that play a major role in photosynthesis: Chlorophylls, Carotenoids, and Biliproteins. The presence of each of them in algae will add the ability of sunlight absorption in a certain optical window. Pigment composition in different algae reflects the adaption to the natural light environment to make it suitable for algae to grow. Different types of algae are recognized based on the presence or absence of various photosynthetic pigments (Ston et al., 2002; Greisberger and Teubner, 2007).

Phytoplankton are a subset of algae and are the suspended aquatic microorganisms. Their biomass is one of the important bio-optical characteristics of waterbodies. Their distribution and population define the waterbody's health, composition, and ecological status of the lake. Phytoplankton are unicellular algae living in both salty and fresh waters. These microorganisms use photosynthetic pigments to capture sunlight, carbon dioxide and nutrients available in water to produce organic food, and release oxygen. Phytoplankton can be further classified into three categories of green algae, diatoms, and blue-green algae (also called cyanobacteria). The phytoplankton community structure, and the dominance of a specific group, depends on the environmental factors including nutrient concentrations (Huszar and Caraco, 1998; Florida Lakewatch, 2000).

Phytoplankton biomass has a critical role in bacterial life, and in the life chain of water organisms. They can provide the organic food for other living organisms in water. Accordingly fish production flourish in areas with massive phytoplankton biomass. Moreover phytoplankton population can be an indicator of climate change as they need a specific environmental life condition. Water temperature, salinity, and pollution are among factors affecting phytoplankton life. In addition, phytoplankton acts as a basis in the global carbon-cycle. They can absorb carbon dioxide from water to do photosynthesis using sunlight and release oxygen and store carbon dioxide in the form of organic material as by-product. Oceans absorb additional carbon dioxide available from the atmosphere to replace the reduced carbon dioxide concentration. The dead phytoplankton biomass sinks to the ocean floor carrying the stored carbon in their cells which will be covered by other sinking materials. Accordingly, ocean will create a "carbon sink" to store world's excess atmospheric carbon (Golterman, 1975).

On the other hand, a high population of phytoplankton, which is a consequence of having high concentration of nutrients needed for photosynthesis, is harmful for water ecosystems. The enhanced growth of phytoplankton, which is referred to as algal bloom, can result in production of toxins harmful for ecology and human health. Also, bacteria, decomposing the dead body of phytoplankton, deplete the oxygen in water causing low levels of oxygen required for other organisms to survive. Accordingly, the result will be a dead zone (Bresciani and Giardino, 2012).

Reasons for variations: As discussed above, various parameters including light, nutrients availability, temperature, salinity, turbulence, and stratification can vary in different seasons and locations of the lake leading to a change of lake's environment. Consequently, the suitability of ecosystems for phytoplankton growth will change. Phytoplankton growth is dependent on the availability of enough nutrients, and light to do the photosynthesis (Deus, 2013). Temperature is an important environmental factor that can regulate phytoplankton growth (Binding et al., 2011a). Harris et al. (2006) conclude that a 4ºC increase in water temperature can result in a 20% increase of primary production and a 43% increase in respiration (Harris et al., 2006).

2.3 Field Measurements of Water Quality

In situ measurements remain the most accurate solution for water quality monitoring programs (Moore et al., 2014). Mueller et al. (2003) provided a complete set of protocols for the measurement of biogeochemical, bio-optical, and optical parameters in waterbodies. In this chapter, the focus is on field measurement methods for chl-a and turbidity, as these are the two water quality parameters of interest in this thesis. Uncertainties associated with each of the methods are also summarized.

2.3.1 Chlorophyll-a Measurement

Chlorophyll is the green pigment found abundantly in nearly all algae. Chlorophyll allows plants and algae to use sunlight in the process of photosynthesis for growth (Florida Lakewatch, 2000). It has different forms; coded a, b, c, and d. These forms have various concentrations in different photosynthetic organisms. Chl-a is the most abundant form (Ston et al., 2002; Greisberger and Teubner, 2007). Thus, the monitoring of chl-a is an indirect measurement of phytoplankton biomass and its temporal variations; and it is a critical factor to measure for water quality monitoring (Moses et al., 2009). Furthermore, chl-a is empirically correlated to nutrients concentrations, in particular, phosphorus (Bresciani and Giardino, 2012).

The use of chl-a for phytoplankton biomass is criticized, because chl-a content can change due to species composition, and the cells physiological state (DosSantos et al., 2003). In addition, the type of phytoplankton cannot be distinguished when chl-a concentration is used as indicator of phytoplankton biomass, as all of them have this photosynthetic pigment. Applying pigment analysis or using the microscope are the approaches can be taken to distinguish phytoplankton composition (US EPA, 1983). Alternatively, carbon, which is a basic currency to measure food production of phytoplankton, can be used as an indicator of phytoplankton biomass. It can be measured by an elemental analyzer. The disadvantage of this method is measuring all suspended carbon in the water sample that can be from detritus and inorganic sources.

Fine meshed phytoplankton nets (often 10 μm) can be used to collect and settle phytoplankton biomass. Sedimentation chambers can also be used to concentrate them. The chl-a concentration in the concentrated biomass is measured. This is possible due to chl-a unique spectral properties and using different methods with various levels of sophistication and accuracy. These methods range from simple and widely used spectrophotometric and fluorometric methods for chl-a, b, and c to more sophisticated chromatographic methods.

It has be noted that, because chlorophyll pigments are sensitive to heat and light, the sampling have to be done in a dark room. This is to prevent them from growing and to avoid degradation during extraction and applying measurement techniques. The preferred method of keeping samples cold is storage in nitrogen with longer period of storage (e.g. one year) and the least degradation. It is also helpful for pigment extraction, as the cell walls and membrane are weakened during rapid temperature changes. Samples can also be kept in ultra-cold freezers (-90˚C) for less than 60 days. Conventional deep freezers can hold the samples for less than 20 hours before transferring them to liquid nitrogen or ultra-cold freezers (Mueller et al., 2003).

2.3.1.1 Spectrophotometric Measurements of Chl-a

The first approach for chl-a measurement was spectrophotometric methods which were established 40 years ago and are still used today (Hydrology Project Training Module, 2000). In these set of methods, using a membrane filter, a known volume of water is filtered to concentrate the algal cells in sample (filter 50 ml to 2000 ml of water using a 0.5 μm filter pore size). An aqueous solution of acetone is used to extract pigments from the concentrated algal cells. The extract is then clarified by centrifugation. Absorption of the extract in different wavelengths is determined spectrophotometrically. The measurements are applied in a set of standard equations to derive chl-a concentrations (Salinas, 1988; University of South Florida, 2010).

Monochromatic equation: this method requires absorption at the wavelength of 665 nm before and after a 90-second acidification step. Chl-a concentration, which is corrected for phaeophytin (a common degradation product of chl-a), and the amount of phaeophytin in the sample are measured. Chl-b and c concentration cannot be derived using this equation.

Trichromatic equation: this method requires absorption measurements at the wavelengths of 665, 645, and 630 nm to calculate the amount of chl-a, b, and c in the sample. The chl-a is uncorrected and phaeophytin concentration cannot be measured.

2.3.1.1.1 Interferences

Over- or underestimation of chl-a in solutions containing all of the pigments is inevitable, due to spectral absorption overlaps (Arar, 1997a, 1997b; Arar and Collins, 1997; DosSantos et al., 2003). The spectral bandpass of the specterometer has to be adjusted to a band narrow enough. This adjustment avoids the overlap of absorption for different pigments to measure chl-a accurately at a given wavelength. The American Society for Testing and Materials (ASTM, 1984) reports a 40% error in estimation of chl-a concentration while using a spectrometer with 20 nm spectral bandpass (American Society for Testing and Materials, 2009). A spectral bandpass of 2 nm has been shown to work adequately to measure chl-a concentrations (Salinas, 1988).

Chl-a is overestimated in the presence of phaeophytin using trichromatic equations. This is because phaeophytin absorbs light at the same wavelength as chl-a (665 nm); however, correction for the presence of it will not be considered in this method (Florida Departement of Protection Plan, 2011). However, the monochromatic equation with the acidification step corrects for phaeophytin interference. During the acidification, chl-b is converted to phaeophytin-b, which contributes in phaeophytin-a absorption. Therefore, the correction for phaeophytin can be potentially overestimated in the monochromatic equation (Florida Departement of Protection Plan, 2011). The use of the monochromatic method is preferred over the trichromatic equation, since the later poorly estimates chl-b and –c concentrations that leads to unrealistic (sometimes negative) results for chl-a. Chl-a is the most abundant component of chlorophylls in phytoplankton biomass, therefore the monochromatic method of measuring chl-a concentration has been shown to estimate most accurately the phytoplankton biomass (Salinas, 1988).

Absorption reading at 750 nm is conducted to remove the effect of turbidity for both equations (Hydrology Project Training Module, 2000). High absorption at 750 nm indicates a poorly clarified solution which needs centrifugation or filtration before analysis.

2.3.1.2 Fluorometric Measurements of Chlorophyll-a

Pigments are first extracted from the sample using the same process as the spectrophotometric method. This method is based on the fact that chl-a fluoresces at 665 nm, when radiation at a wavelength of 430 nm excites it. The excitation and emission wavelengths are determined using optical filters (Hydrology Project Training Module, 2000). Any material that fluoresces in the red region can interfere in the chl-a measurement. The acidification procedure can be added to correct for the presence of phaeophytin; and the chl-a concentration is measured considering the decrease of fluorescence after acidification (Arar and Collins, 1997).

A modified fluorometric method with special narrow bandpass filter can be used to nearly eliminate the interference of phaeophytin, and chl-b. In this case, fluorometry without performing acidification is applied; it causes a minimal overestimation of chl-a (Florida Departement of Protection Plan, 2011).

Fluorometric measurements are cost-effective, easy, and fast for determining chl-a. A smaller sample volume is required with this method. Also, the analysis can be conducted on the ship during lengthy cruises. However, fluorometric methods are more appropriate for total pigments estimations due to possible interferers in case of having various kinds of pigments (DosSantos et al., 2003; Mueller et al., 2003).

2.3.1.2.1 Interferences

Depending on the type of algae, and therefore the pigments amount present in the sample source, there are spectral interferences using the fluorometric method that may under- or over-estimate the concentration of chl-a. If chl-b is present, during the acidification step, it will be converted to phaeophytin-b. Fluorescence contribution of phaeophytin-b to the one of phaeophytin-a will result in overestimation of phaeophytin-a correction and underestimation of chl-a (Arar and Collins, 1997; Florida Departement of Protection Plan, 2011). The presence of chl-c in the sample will cause overestimation of chl-a (Arar and Collins, 1997; Florida Departement of Protection Plan, 2011).

2.3.1.3 HPLC measurement of Chl-a

The most recent method is the chromatography which separates various pigments and their degradation products before determining their concentration. The concentration of chl-a, b, c, and other species of interest can be determined after separation in the chromatography column and in a single instrument run using a fluorescence or absorption detector (Hydrology Project Training Module, 2000).

Results from High-Performance Liquid Chromatography (HPLC) may be lower than those from spectrophotometric and fluorometric methods. This is because of the ability of chromatography methods to separate pigments and avoid the positive interferences happening in spectrophotometry and fluorimetry (Florida Departement of Protection Plan, 2011).

Although this method is slow and requires expensive equipment to be operated by specialized technicians, the information gained from this method is important especially when pigment separation in the sample is necessary. Due to separation of diverse chlorophyll types and degradation products, this method is precise and is subject to low interference (DosSantos et al., 2003).

2.3.2 Turbidity Measurement

Turbidity is a general indicator of environmental health of waterbodies. Suspended and dissolved matters make water appear cloudy and muddy. Particulates such as clay, silt, finely divided organic matter,

plankton, and other microscopic organisms, organic acids, and dyes affect turbidity measurements (Chauncey, 2005). Water clarity or optical depth is measured widely using a Secchi disk. The disk is attached to a rope that is lowered into the lake until it disappears from sight. Greater Secchi Disk Depth (SDD) indicates clearer water that lets the light penetrate more deeply into the lake (Ingram and Young, 2010).

Another method of measurement uses optical probes, and is based on "comparison of the intensity of light scattered in the sample under defined conditions with the intensity of light scattered in a reference suspension". The turbidity is higher if the intensity of light is high (EPA Guidance Manual, 1999; US epa, 2003). A wide variety of probes are available for turbidity measurement. These probes do not use the same wavelength, and angles of measurement to detect the scattered light. In addition, the signal processing strategies are different. Based on the differences in probes, various units of measurements are established and assigned considering the probe design (Chauncey, 2005).

2.3.2.1 Interference

The SDD measure is a worldwide-accepted method, easy and economic; however it still has some limitations. Limitations include the subjectivity of the measure due to its dependency on the ability of the operator's eye. Also, the measure is influenced by environmental factors during the day such as the inclination of the sun, the state of the water, the time of day, and weather. SDD measurements are more reliable under certain conditions. Measurements should be made during calm days between 9 a.m. and 3 p.m. on the leeward side of the boat, and with the sun at the back of the person who is measuring. Secchi measurements have limitations in shallow waters, where it may reach the bottom without having disappeared. These limitations of the Secchi disk method can be improved by using optical devices such as turbidimeter (Civera et al., 2013).

In the presence of floating debris and coarse sediments that settle rapidly, the turbidity is low. Samples should be gently agitated by swirling to reduce particle settling. Otherwise the particles will be broken apart or air will be entitled into the sample when violently agitated (EPA Guidance Manual, 1999; US epa, 2003). Water samples always contain bubbles. They act like particles and scatter light resulting in incorrect measurements of turbidity (positive bias). The bubbles are generated because of chemical or biological processes, during filling of a sample container, or temperature fluctuations that result in a reduced solubility of gas in liquid. Also, temperature changes can alter particles' behavior or create new particles. Particle settlement and temperature changes can lead to incorrect measurements. Therefore, measurements have to be done expeditiously (EPA Guidance Manual, 1999; US epa, 2003).

Instrument fouling, that is the result of biological growth or scratches, can also cause incorrect measurements. In case of scratches that block the sensor light beam, there will be a negative bias. Also, scratches that result in scattering of the sensor light beam will have a positive bias on turbidity measurements (Chauncey, 2005). Stray light will also increase the apparent light scatter and therefore will cause the positive bias on measurements (Chauncey, 2005).

2.4 Satellite Remote Sensing of Water Quality

Effective management of water resources requires monitoring water quality parameters from past to present. One way of achieving this is through field sampling and laboratory measurements. This necessitates technical and financial support to operate cruise surveys in adequate spatial and temporal coverage. However these efforts can still include multiple challenges, such as logistical issues, and breaks in sampling due to changes in funding priorities. Furthermore, field measurements are difficult to obtain in high-latitude remote regions due to accessibility. Given all of the above, complementary methods for water quality measurements are needed. Satellite remote sensing has proven to be a viable method for deriving bio-optical parameters, especially over lake regions where direct field measurements are difficult to obtain (Le et al., 2013).

Optically significant water constituents, such as chl-a, suspended particulate matters, and colored dissolved organic matters, can be derived from satellite remotely sensed measurements of the reflected sunlight at given wavelengths in the visible and near-infrared spectrum (Cococcioni et al., 2007). This is due to the fact that these constituents absorb and scatter various wavelengths of the incident light in water (Figure 2-3). Hence, satellite remotely sensed water-leaving radiance can provide valuable information about changes in the aquatic environment such as algal blooms. There are two different optical properties associated with in-water constituents and their concentrations that is derived from satellite measurements: inherent optical properties (IOPs) and apparent optical properties (AOPs). AOPs are related to water constituents and light field, and are measured by satellite as water-leaving reflectance, whereas IOPs are solely a function of water constituents and their distribution and optical properties. Although IOPs are difficult to measure from optical studies, they can be related to AOPs. The derived IOPs can then be converted to constituent concentrations. Accordingly, the goal of optical studies is to measure AOPs and then derive IOPs from them to evaluate physical, chemical, and biological properties of water (Lee et al., 1994).

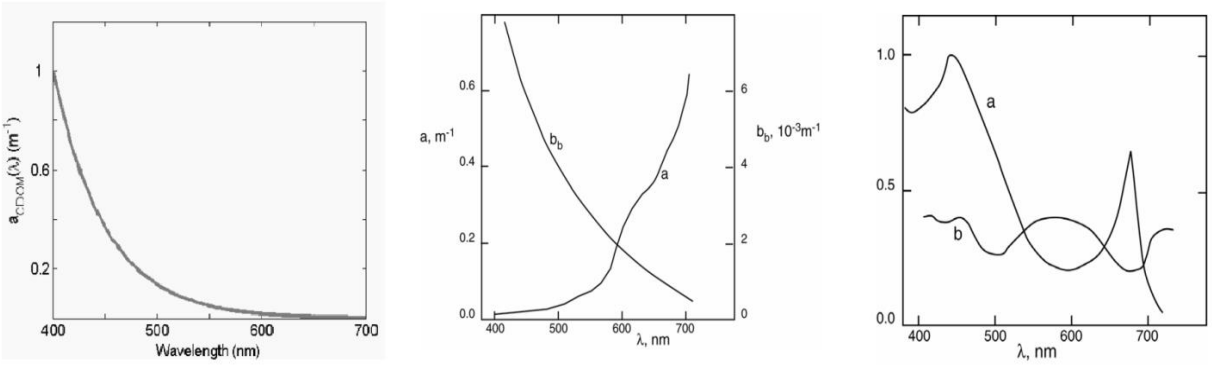


Figure 2-3 From left: Approximate absorption spectrum of colored dissolved organic matter (a_{CDOM}); Absolute values for sea water absorption (a) and backscattering (b_b); Relative values (normalized at 440 nm) of absorption (a) and scattering (b) for chl-a in phytoplankton (Robinson, 2004).

According to the optical classification by Morel and Prieur, there are two different types of natural waters, namely Case I and Case II (Morel and Prieur, 1977). In deep ocean waters, the predominant constituent is phytoplankton and other constituents that co-vary with chl-a concentration. Such waters are referred to as Case I waters. Their optical properties are dominated by phytoplankton and the reflected light can be related to chl-a concentration. Case II waters contain other constituents such as suspended materials and dissolved materials which do not co-vary with phytoplankton biomass. Thus, the optical properties of Case II waters are more complex by producing overlapping, uncorrelated absorption and scattering through its components (Moses et al., 2009). Satellite remote sensing has been increasingly used to monitor the water quality parameters of Case I and II waters. The general parameters being mentioned in literature as being detectable from ocean color satellites are chl-a (Gower and King, 2012; Sugumaran and Thomas, 2012; Lesht et al., 2013; Sá et al., 2015; Watanabe et al., 2015); TSM (Liu et al., 2010; L. Wang et al., 2012; Song et al., 2012; Tilstone et al., 2012); colored dissolved organic matter (CDOM) (López et al., 2012; Tian et al., 2012; Matsuoka et al., 2013); diffuse attenuation (water transparency and SDD) (Doron et al., 2011; Thayapurath and Talaulikar, 2012; Binding et al., 2015); cynobacterial pigment phycocyanin (Randolph et al., 2008; Ruiz-Verdú and Koponen, 2008; Kutser, 2012); and water surface temperature (Petrusevics et al., 2011; Chen et al., 2013; Zhu et al., 2013).

In this section, different remote sensing approaches for water quality monitoring are described. The most common satellites used to derive optical properties of waterbodies are also mentioned. The rationale for employing MERIS in this thesis is also discussed.

2.4.1 Remote Sensing Algorithms

Morel and Gordon (1980) describe three approaches for the interpretation of water color data in order to derive the optical constituents' properties: empirical, semi-analytical, and analytical. Empirical models are

based on statistical regressions to imply effective data optimization but they have limited transferability (e.g. Kabbara et al., 2008). In analytical models, radiative transfer theory is used to provide the relationship of upwelling irradiance (or radiance) with the water constituents. An inversion technique is applied on the resultant system of equations to derive constituents' concentrations from irradiance (or radiance) values measured at several wavelengths (Morel, 1980). The semi-analytical approach is explained when a simplified solution of the radiative transfer equation is adopted which requires approximation or calibration of the bio-optical pieces of the model with empirical equations (e.g. Carder et al., 1999; Binding et al., 2012). Statistical regression can be based on properties known from physics (e.g. investigating the nature of radiance spectral peak for chl-a concentration; Gitelson, 1992) which describes the epithet of "semi" from side of empirical approaches (Odermatt et al., 2012a). In this thesis, semi-empirical and empirical terms are used interchangeably.

Several empirical models have been developed to establish a regression between the water quality parameter of interest and water-leaving reflectance or radiance in certain spectral bands, the ratio of these bands, or the combination of individual bands and their ratio, using different types of equations (linear or nonlinear) (Ruiz-Verdú and Koponen, 2008; Wang et al., 2012; López et al., 2012; Potes et al., 2012; Le et al., 2013). Empirical algorithms are robust, and simple to implement, especially for cases where in situ data is regularly collected. Regression-based algorithms generally work well for the specific lake of interest and for the time when the in situ data was collected; however, the applicability of the algorithms to other lakes and different seasons has to be verified (Gons et al., 2008).

The empirical sensor-specific standard algorithms, which are based on using the primary chl-a absorption band (blue/green), exist for all medium resolution ocean color spectrometers. These algorithms are applicable to open ocean waters, and include the OC2, OC3, OC4, depending on the number of bands used (O'Reilly et al., 1998, 2000). However, O'Reilly et al. (1998) and other studies have clearly indicated that blue-green models do not perform well in Case 2 waters. In turbid waters, the spectral peak shifts from blue to green to red with increasing sediment loads. However, the blue/green band ratio algorithms interpret this shift solely as an increase of chl-a, leading to significant overestimation of chl-a concentration. Also, CDOM absorption in the blue region of spectrum is greatest. The effect of CDOM absorption decreases exponentially with increasing wavelengths, being near negligible in the red/NIR for the majority of waters (e.g. in the Great Lakes) (Binding et al., 2008). Therefore, these wavelengths offer minimal sensitivity to other CPAs. In optically complex Case II waters, the secondary absorption peak of chl-a in 665 nm is promoted due to the least variations from other water constituents apart from increasing water absorption. However, the major limitation of these methods is the absence of such feature in oligotrophic and some mesotrophic Case II waters (Odermatt et al., 2012a). Several studies have investigated the potential of using

red/NIR bands for estimating chl-a concentration by exploiting the absorption and fluorescence signals of chl-a at these wavelengths. A few recent ones are mentioned in the following.

The solar-simulated fluorescence signal from chl-a in surface waters can be used to measure near-surface phytoplankton biomass. Fluorescence Line Height (FLH) is defined as the increase in radiance near 685 nm relative to the measured radiance in a baseline band. However, some studies document that lake turbidity is a critical factor in the effective use of fluorescence. Gilerson et al. (2007) show non-algal and mineral particulates to have an elevated scattering at fluorescence emission, such that the effect of fluorescence on reflectance spectra in many turbid waters is almost negligible (Gilerson et al., 2007; Binding et al., 2011; Binding et al., 2012). Also, in intense bloom conditions, the absorption of chl-a at 685 nm dominates the fluorescence radiance. This leads to an absorption depth measurement instead of FLH (Gower and King, 2012). Gons et al. (2008) report the successful use of MERIS FLH in oligotrophic waters of the Laurentian Great Lakes, and failure of this method (with diminishing or negative FLH) in mesotrophic or eutrophic lakes (Gons et al., 2008).

The MCI is also a band-difference index similar to FLH, but it measures the height of water leaving reflectance peak at 709 nm relative to a baseline between bands at 681 and 753 nm (i.e. for MERIS bands). Binding et al. (2011) indicate that MCI is more effective solution than FLH for monitoring algal blooms in Lake of the Woods (a lake occupying parts of the Canadian provinces of Ontario and Manitoba, and the U.S. state of Minnesota with chl-a concentration ranging from 1.9 to 70.5 mg m^{-3}) (MCI: R^2= 0.72, FLH: R^2=0.57). This study found that MCI for Lake of the Woods was fairly independent of the effect of high dissolved organic and suspended matters, and also the failure of atmospheric correction procedures that is common for MERIS in eutrophic waters (Binding et al., 2011). Binding et al. (2013) compared the potential of MERIS MCI product to detect algal blooms under a variety conditions for eutrophic, oligotrophic, and turbid waters of Lake of the Woods, Lake Ontario, and Lake Erie, respectively. The in situ measured MCI in Lake of the Woods showed strong logarithmic relationship with the in situ chl-a concentration ranging from 8.4 to 289 mg m^{-3} (R^2=0.91). Despite the contamination of water-leaving reflectance due to other CPAs (TSM and CDOM), the red/NIR portion of spectrum was dominated by the effect of algal cells. The study suggests that MCI is fairly insensitive to CDOM, but significant sensitivity to mineral scattering. However, for Lake Erie, where mineral sediments often dominate the optical signal rather than chl-a, MCI showed a strong linear relationship with chl-a concentration in the range of 0.7-20 mg m^{-3} (R^2=0.70). The apparent success of MCI in this study was attributed to the dominance remote sensing reflectance due to the thick cyanobacteria bloom, eliminating the contribution from the mineral turbidity below. For oligotrophic waters of Lake Ontario, with chl-a concentration less than 10 mg m^{-3}, the peak near 709 nm shows very little sensitivity (Binding et al., 2013).

Different blue/green band ratio regression models were tested in Witter et al. (2009) to derive chl-a concentration in Lake Erie using the SeaWiFS observations. The authors considered the evaluation of twelve globally-calibrated ocean-derived color algorithms, one regional algorithm derived for Case II waters of Baltic Sea, and a set of regional algorithms developed for the western, central, eastern basins of Lake Erie. Their results showed that none of the ocean-derived algorithms performed adequately for Lake Erie ($1.72<RMSE<3.41$ mg m^{-3}; $0.65<$correlation coefficient $(R)<0.73$). However, the regionally calibrated quadratic algorithms were promising for the eastern basin (RMSE=0.74 mg m^{-3}, R=0.79), and possibly the central basin (RMSE=1.07 mg m^{-3}, R=0.82). All the algorithms performed poorly in the western basin. They suggested that for the eastern and central basins of Lake Erie, a regional approach using "ocean-like" algorithms represents a significant improvement over direct use of globally-calibrated ocean algorithms. However, in the western basin, a different approach to algorithm development is required, when the regionally calibrated algorithms considered in that study did not produce satisfactory results. Moore et al. (2014) considered applying red/NIR bands in Lake Erie in addition to the blue/green bands to derive chl-a concentration from MERIS observations. A clustering approach was presented to manage the selection between the two algorithms: NASA-OC4 and MERIS 3-band algorithms. NASA's OC4 algorithm is based on using the blue/green region of spectrum and is suitable for deriving low chl-a concentrations, while the MERIS 3-band algorithm that employs red/NIR bands becomes more important in turbid and/or eutrophic waters The retrieval accuracy improved from RMSE (MBE) values of 0.416 (-0.194) and 0.437 (0.115) for NASA OC4 and MERIS 3-band, respectively, to 0.320 (0.023) in the blending algorithm (the accuracies were reported in the logarithmic units).

The western basin of Lake Erie is optically complex due to the existence of multiple CPAs, such as phytoplankton, suspended matters, and dissolved organic carbon. To improve the understanding of mixed spectral signatures, hyperspectral remote sensing technology was used in Ali and Ortiz (2014). The study estimated the concentration of chl-a and TSM in the optically complex waters of the western basin of Lake Erie. Ali and Ortiz (2014) collected upwelling radiances in visible and NIR range (400-900 nm) using a field-based spectroradiometer. A partial least squares (PLS) method was applied on the continuous narrow spectral bands. The method models the covariance between hyperspectral data and CPAs and identifies the optimal bands that characterize most of the variances in the CPAs. The method produced R^2 (RMSE) values of 0.84 (1.18 mg m^{-3}) and 0.90 (1.26 mg.L^{-1}) for estimating chl-a and TSM concentrations. Their study illustrated the potential of hyperspectral data and the PSL method for extracting the absorption features of various CPAs in optically complex Case II waters (Ali and Ortiz, 2014).

Other water constituents can also be estimated using satellite observations. Brezonik et al. (2015) applied different published band ratio models on the simulated Landsat 8, Sentinel-2 and Sentinel-3 bands from

field-measured reflectance spectra in several sampling sites in Minnesota and Wisconsin to estimate CDOM absorption at 440 nm. High R^2 values were obtained (0.83-0.86). The study concluded that the broader Landsat 8 bands performed as well as the narrower Sentinel bands and hyperspectral bands. This is probably because CDOM absorption decreases exponentially with increasing wavelengths; and there is no specific peaks and troughs in the absorbance or reflectance spectra (Brezonik et al., 2015). Lobo et al. (2015) employed a non-linear empirical regression algorithm between measured TSM and reflectance of Landsat satellite family sensors (MSS, TM, and OLI) at red band for Tapajós River Basin in Brazil, and reported RMSE= 1.39 and R^2 = 0.94.

For monitoring the variations of SDD in the Laurentian Great Lakes using empirical models, Binding et al. (2007) merged data from the Coastal Zone Color Scanner (CZCS) and the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) over Lake Erie and Ontario. A regression equation was developed between coincident measurements of water transparency and remotely-sensed water-leaving radiance at ~ 550 nm using a power fit (Lake Erie: RMSE=24.8%, R^2=0.71, N=400; Lake Ontario: RMSE= 23%, R^2= 0.65, N=420) (Binding et al., 2007). In Binding et al. (2015), a third-order polynomial relationship was fitted between 1/SDD and remote sensing reflectance at ~ 550 nm, after merging the observations from the CZCS, SeaWiFS, and MODIS-Aqua sensors to monitor SDD variations in the Laurentian Great Lakes. An acceptable retrieval accuracy was derived across the three sensors (N = 1328, RMSE = 29.79 %, MBE = 1.13 %, R^2 = 0.74) (Binding et al., 2015).

The empirical algorithms used in the investigations described above are based on using in situ data for training regression models. However, the correlation that stems from the repeated in situ measurements in space and/or time is not considered for developing the regressions. Therefore, the second objective of this thesis is motivated from this gap in the literature, to further improve the regression-based remote sensing algorithms that will enable successful retrieval of chl-a and SDD on the studied lake.

In turbid waters, there is still significant contribution from suspended matters in the remote sensing signal in the red/NIR region; and the signals cannot be attributed to only chl-a absorption and fluorescence and water alone. In these techniques, the spatial distribution of chl-a concentration can still mirror that of known suspended sediments. Because of the complexity of Case II waters, approaches that employ semi-analytical algorithms seem more effective at distinguishing several constituents confounding water-leaving signals. Semi-analytical methods use the maximum information that can be gained from different bands in satellite images, to solve for several different parameters related to the water constituents, simultaneously. Hence, there is a high probability of differentiating signals coming from different in-water constituents (Matthews et al., 2010). More sophisticated physically based inversion methods have been used to estimate in-water

IOPs from water-leaving reflectance and related them to the concentrations of CPAs employing empirical models.

Binding et al. (2012) used inverse modeling of water-leaving radiance to simultaneously extract chl-a and mineral suspended particulate matters (MSPM) concentrations for Lake Erie from red/NIR bands of the MODIS-Aqua sensor. The estimated concentrations were in close agreement with the in situ observations (chl-a: N=98, RMSE= 2.21 mg m^{-3}, R^2 = 0.95; MSPM: N=98, RMSE = 1.04 mg m^{-3}, R^2 = 0.91) (Binding et al., 2012). Giardino et al. (2015) applied a physically based method that relies on spectral inversion procedures to simultaneously derive the concentrations of water constituents in Lake Trasimeno (Italy) from airborne imaging spectrometry. The bio-optical inversion procedure was implemented in a tool named BOMBER. The studied lake was characterized by both optically deep and shallow turbid waters and dense submerged aquatic vegetation (Giardino et al., 2015). Salama and Verhoef (2015) developed the 2SeaColor model which is a semi-analytical model with an inversion scheme of radiative transfer equation, to retrieve the downwelling attenuation coefficient for Lake Naivasha in Kenya. In this study, MERIS-derived K_d values were calculated from the bulk IOPs which were derived from inverting the observed atmospherically corrected remote sensing reflectance spectra. The retrieved values were compared to the measured ones and showed relatively high degree of accuracy with R^2 value of 0.72.

Between spectral inversion algorithms, the NN inversion techniques are more dominant. This is probably because of their improved availability as MERIS level 2 products, and also the lake processor within the publically available BEAM toolbox. The BEAM plug-ins have been demonstrated to be transferrable to some lakes from coastal zone settings. However, the evaluation of spectral inversion techniques has been conducted to a lesser extent compared to the band ratio algorithms (Odermatt et al., 2012a). The evaluation of the MERIS processors is still important in order to improve the existing time series, even though MERIS data is no longer available since April 2012. Also, the forthcoming Sentinel-3 OLCI sensor will provide continuity to the MERIS mission. Therefore, the learned lessons from the NN-derived products from MERIS in the BEAM toolbox can help improve the processors being developed for OLCI (Kallio et al., 2015).

Palmer et al. (2014) evaluated the performance of different MERIS spectral NN-inversion processors and also band–difference indices for Lake Balaton, Hungary, to derive chl-a concentration. The retrieved values from different processors were calibrated and validated using in situ data. Highly variable accuracy results from different NN-based algorithms and the robust FLH and MCI algorithms were shown. In general, both band-difference algorithms tested performed well (R^2 values equal to 0.78 and 0.62, respectively), whereas the neural network algorithms were found to much less accurately retrieve chl-a concentration ($0.36<R^2<0.48$). Kallio et al. (2015) also evaluated MERIS spectral inversion processors for the estimation

of chl-a, TSM concentrations, a_{CDOM} (443), and SDD in four lakes in southern Finland. The lakes represent oligotrophic, mesotrophic, and humic (high a_{CDOM} and low TSM) lake types. Remote sensing reflectance, mass-specific IOPs (SIOP), and water quality parameters were collected in the field for evaluation of the algorithms. The tested algorithms included Boreal, and FUB/WeW processors. FUB/WeW estimated water constituents with lower R^2 (<0.17) than the Boreal processor, since a_{CDOM} for most of the studied lakes was outside the definition range of FUB/WeW. Boreal lake processor retrieved chl-a concentrations with R^2 values of 0.64 and 0.83, and SDD with R^2 values of 0.39 and 0.85 depending on the approach used to estimate them. Matthews et al. (2010) utilized the standard Level 2 Case II waters product and Eutrophic lake processors in BEAM as the semi-analytical methods to derive chl-a, TSM, SDD, and a_{CDOM} for Zeekoevlei Lake in South Africa. The standard product resulted in negative reflectance values in the blue region, indicative of atmospheric correction failure. This is due to breakdown of the assumption of negligible water-leaving radiance in the NIR, as a result of high sedimentation rates in Zeekoevlei. The spectra produced from the Eutrophic Lake processor showed not negative but uncharacteristic spectral shape relative to the in situ measurement. Therefore, the NN-based products failed to produce reasonable comparisons with in situ data. Failure of atmospheric correction can also be explained by the discrepancies between the IOPs of Zeekoevlei and those used to simulate data (using radiative transfer equation) to train the atmospheric correction module of the NN algorithm (Matthews et al., 2010; Tilstone et al., 2012).

The MERIS NN-based inversion processors, available as plug-in in the BEAM software, are trained using in situ data collected mostly from European waters. These algorithms were designed aiming to cover the optical properties of lakes of various types. This motivated the first objective of this thesis to test the applicability of MERIS processors available in the BEAM toolbox to derive water quality parameters for Lake Erie. However, the NN processor should be preferably trained with the conditions (water optics, water quality ranges, atmospheric properties, etc.) of the region of interest. Evaluation of semi-analytical spectral inversion algorithms is complicated as it requires in situ measurements of bio-optical parameters, in particular IOPs and SIOPs which have rarely been measured for water bodies, and also spectral remote sensing reflectance for atmospheric correction evaluation. Kallio et al. (2015) summarized the factors causing divergence between semi-analytically derived optical properties and the in situ measurements as follow:

1) Inaccuracies in SIOPs used in the bio-optical models,

2) Atmospheric correction accuracy and adjacency effect,

3) IOPs estimated by the inversion algorithm,

4) Applied conversion factors to estimate water quality concentrations from IOPs, and

5) Unsuitable training ranges for NNs.

Thus, the first paper of this thesis can be considered as a pilot study to evaluate the performance of the MERIS NN-inversion based processors to derive chl-a concentration and SDD. However, in order to fully investigate the reasons for discrepancies between the in situ measurements and satellite observations, a more complete in situ data collection of bio-optical properties is necessary.

In general, semi-analytical methods are advantageous over empirical methods because of the reduced requirement for concurrent in situ data and generalizability over spatial and temporal scales, and different satellite sensors. Therefore, the advantage of semi-analytical models over empirical approaches, is more related to their adjustability, and also the justification that arises from their physically sound procedure, rather than accuracy (Odermatt et al., 2012a). On the other hand, the empirical approaches are simple, providing rapid access to data and their correlation; however at a cost of not expressing the cause-effect relationship. The semi-analytical inversion approaches are more complex. Their use of many spectral bands, if not all of them, makes the approach very sensitive to errors in atmospheric correction (Matthews et al., 2010; Doron et al., 2011). Some of the semi-analytical algorithms require training (re-parameterization) using the regional IOPs of the study area. Also, an error in any stage of procedure (e.g. atmospheric correction), propagate errors in the estimation of other parameters. This will complicate the validation of these methods, as mentioned above.

An overview of Chl-a and SDD retrieval algorithms presented in recent studies (2007-2015) is depicted in Figure 2-4 and Figure 2-5 with corresponding sensors and in situ measurement ranges. The bands employed in the empirical models are also shown.

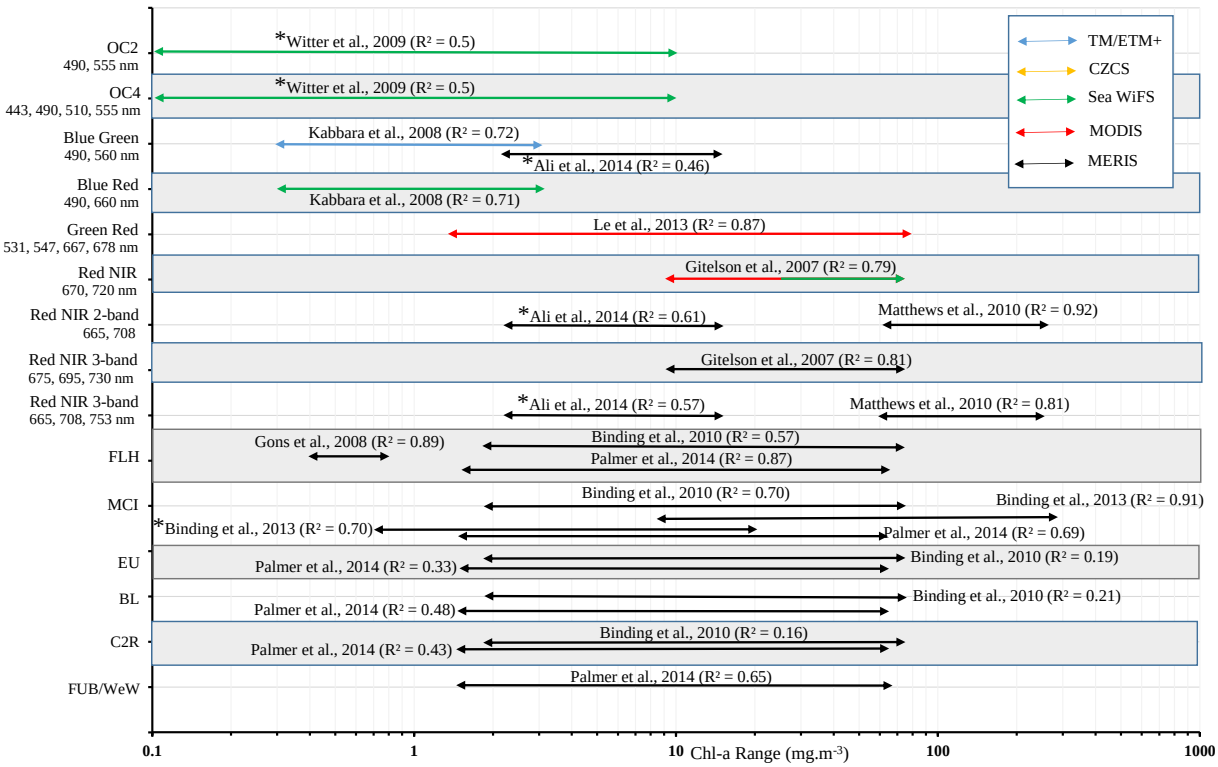


Figure 2-4 Overview of recently (2007-2014) published papers on remote sensing of chl-a for optically complex water bodies. The algorithms retrieve chl-a from satellite imagery using matchup-validated empirical and semi-analytical algorithms. Arrows with more than one color are indicative of studies that utilized more than one sensor. Starred references indicate those studies that include Lake Erie. C2R, EU, BL, FUB/WeW are the NN-based semi-analytical algorithms available in the BEAM software and are explained in more details in section 3.2.2.1.

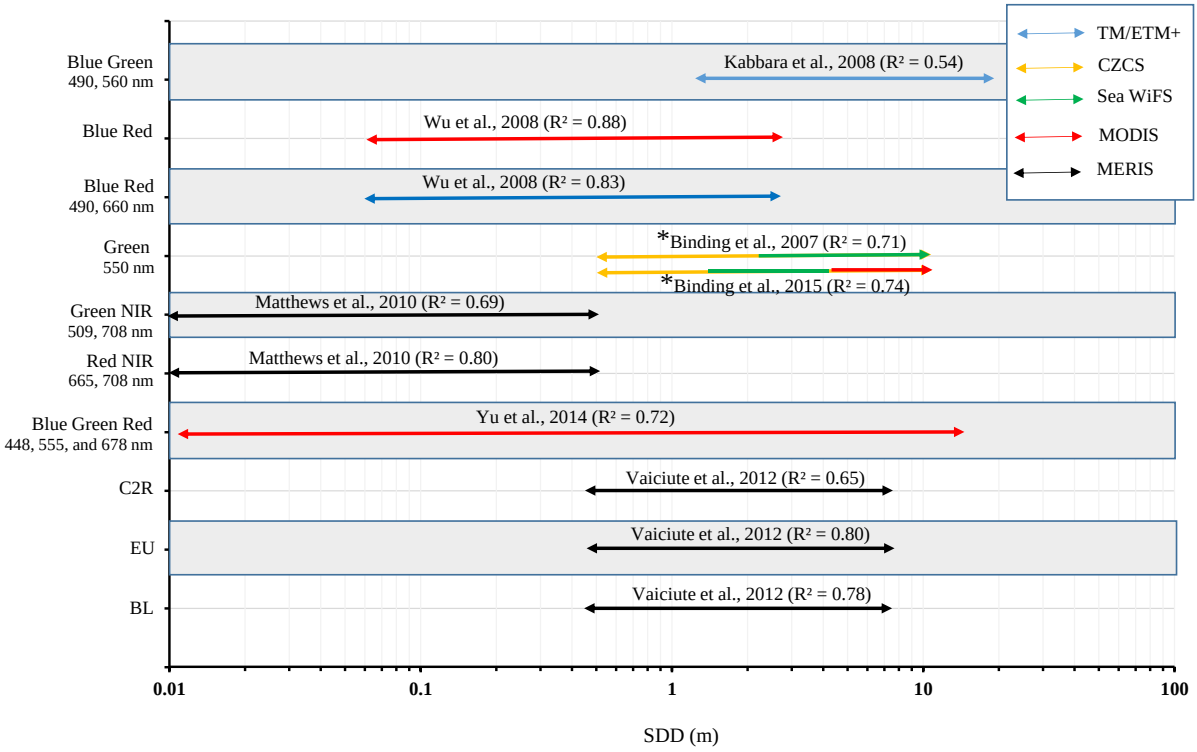


Figure 2-5 Overview of recently (2007-2015) published papers on remote sensing of SDD for optically complex water bodies. The algorithms retrieve SDD from satellite imagery using matchup-validated empirical and semi-analytical algorithms. Arrows with more than one color are indicative of studies that utilized more than one sensor. Starred references indicate those studies that include Lake Erie. Studies in C2R, EU, BL are the NN-based semi-analytical algorithms available in the BEAM software and are explained in more details in section 3.2.2.1.

2.4.2 Past and Current Remote Sensing Satellites

Historical, current, and scheduled ocean color satellites are listed on the International Ocean Colour Coordinating Group (IOCCG) website (http://www.ioccg.org). Satellite sensors such as Sea-Viewing Wide Field-of-View Sensor (SeaWiFS), MERIS, and Moderate-resolution Imaging Spectrometer (MODIS) are currently being used to deliver past and current ocean color data. The characteristics of SeaWiFS, MERIS, and MODIS, as well as their products specific to water quality, are summarized in Table 2-3. The archive of SeaWiFS image data is available from 1997 to 2010. MERIS is available for a 10-year period from 2002 to 2012. MERIS, although it stopped operating in April 2012, is still contributing significantly to the remote sensing of inland water color due to its high spectral, temporal, and radiometric resolutions. The large number of research papers published in recent years on MERIS satellite imagery, compared to other satellites to retrieve different optically active water constituents and other optical properties, is a testament of this. MODIS onboard of the Terra and Aqua satellites is still in operation; however, the red and NIR spectral resolution of MERIS is better suited to derive a more accurate secondary chl-a absorption

maximum compared to MODIS, which is especially important in Case II waters (Odermatt, Pomati, et al., 2012). This is one of the main reasons why MERIS was retained as the remote sensing data source in this thesis.

Although the Landsat Data Continuity Mission (LDCM), a collaboration between NASA and USGS, is primarily designed for land applications, pre-launch simulations showed its potential to provide useful optical properties measurements of aquatic environments (Pahlevan and Schott, 2013). It is the eighth satellite of the Landsat series, which was launched in February 2013. The mission carries the Operational Land Imager (OLI) sensor. OLI has two additional wavebands compared to Landsat 7 that can improve ocean color measurements and high-altitude cirrus cloud observations. The added blue band results in higher accuracy of water quality monitoring and the shortwave-infrared band enables detection of variations in atmospheric aerosols, thus eliminating the confusing signals produced by them and making more accurate observations of water quality parameters (Irons and Masek, 2012).

A significant advantage of OLI over other existing ocean color capable missions is its 30 m spatial resolution. This resolution is at least an order of magnitude higher than ocean color capable missions including SeaWiFS, MERIS, and MODIS. A higher spatial resolution is beneficial for studying the heterogeneous coastal and inland waters to resolve the fine spatial structure of the water constituents. It is also important to distinguish land from water, especially in narrow rivers (Franz et al., 2014). Therefore, OLI on Landsat has the potential to make a valuable contribution to ocean color studies, specifically over coastal environments (Franz et al., 2014). However, the 16-day revisit time is still a factor that make Landsat 8 less appealing than MERIS for water quality monitoring over Case II waters and lakes with turbid waters; where water quality parameters can change in order of a few days to hours.

Table 2-3 Most frequently used satellite data products for water quality

Platform	Sensor	Operation	Product	Spatial Resolution	Spectral Resolution	Temporal Resolution			
						Daily	3-Day Composite	8-Day Composite	Monthly
OrbView-2 (USA)	SeaWiFS	Sep 1997 - Feb 2011	chlor_a (OC4)	1.1 km, 4.5 km, 9 km	8 Spectral Bands (402 nm - 885 nm)				
			nLw_412						
			nLw_443						
			nLw_490						
			nLw_510						
			nLw_555						
			nLw_670						
			Kd (490)						
			PIC						
			POC						
			CDOM Index						
			PAR						
Envisat (Europe)	MERIS	June 2002 - Apr 2012	algal_1	300 m, 1 km	15 Spectral Bands - (390 nm - 1040 nm)				
			algal_2						
			nLw_413						
			nLw_443						
			nLw_490						
			nLw_510						
			nLw_560						
			nLw_620						
			nLw_665						
			nLw_681						
			nLw_709						
			nLw_754						
			nLw_775						
			nLw_865						
			nLw_890						
			TSM						
			CDOM Index						
			PAR						

Platform	Sensor	Operation	Product	Spatial Resolution	Spectral Resolution	Temporal Resolution			
						Daily	3-Day Composite	8-Day Composite	Monthly
Terra (Aqua) (USA)	MODIS	Jan 2000 (May 2002) - Current	chlor_a (OC3)	1 km, 4 km, 9 km	36 Spectral Bands (400 nm - 14.4 µm)				
			nLw_412						
			nLw_443						
			nLw_469						
			nLw_488						
			nLw_531						
			nLw_547						
			nLw_555						
			nLw_645						
			nLw_667						
			K_d (490)						
			PIC						
			POC						
			CDOM Index						
			PAR						
			IPAR						
			nFLH						
			SST						

Legend

chlor_a, algal: Sub-surface chl-a concentration
nLw: Remote Sensing Reflectance
PIC: Particulate Inorganic Carbon
POC: Particulate Organic Carbon
K_d: Diffuse Attenuation Coefficient
PAR: Photosynthetically Available Radiation
IPAR: Instantaneous PAR
nFLH: normalized FLH
SST: Sea Surface Temperature

2.5 FLake Model

Lakes are an important component of the climatic system influencing the regional and local climate (Martynov et al., 2010). They affect the local weather by modifying the air temperature, wind and precipitation in their surrounding (Eerola et al., 2010). Their influence in regions with an abundance small lakes or with large lakes is considerable (Martynov et al., 2012). Therefore, accounting for lakes in weather forecasting and climate modeling is essential for many lake-rich regions, such as Northern Europe and North America (e.g. Canadian Shield and the Great Lakes area) (Martynov et al., 2012). Samuelsson et al. (2010), for example, assessed the impact of considering the presence of lakes in climate model simulations in contrast to land only. Results showed that lakes induce a warming on the European climate for the fall and winter seasons (Samuelsson et al., 2010). By increasing the spatial resolution of Numerical Weather Prediction (NWP) models, it is possible to improve the handling of the lake influence on weather (Eerola et al., 2010).

To account for lakes, a variable representing the exchange of water and energy between lakes and the atmosphere is needed in atmospheric models. The lake surface temperature, as well as ice-coverage and other lake surface properties influence the atmospheric boundary layer; and are used as inputs to NWP models to handle lake-atmosphere interactions and describe the energy exchange between the two mediums. These properties are obtained based on either observations or reliable simulations. However, real-time observations of these variables are not available in practice for operational NWP models (Eerola et al., 2010). Eerola and Rontu (2010) investigated three different ways to derive lake surface temperature: climatic information, by assimilating lake surface temperature observations, and by applying lake parameterization schemes for prediction of lake surface temperature. They tested these different approaches in the High Resolution Limited Area Model (HIRLAM) to show the differences in the atmospheric temperature forecast. They used Freshwater Lake (FLake) model as the lake parameterization scheme and produced promising results (Eerola et al., 2010). Thus, model-based information on lake surface properties can be used when there are limited sources of observations.

The choice of the lake model scheme to couple with the NWP models is a challenge. There are lake models with different levels of complexity from simple mixed-layer models to computationally expensive three-dimensional models (Martynov et al., 2012). The appropriate model has to reproduce the behavior of surface conditions of the lakes present within the simulation domain. Also, the model has to compromise between adequately reproducing the lake surface properties and using the reasonable computational time and memory sources (Martynov et al., 2012). Another factor affecting the choice

of the optimized lake model is the type of lakes in the simulation domain. The type of lakes span from tropical lakes to temperate-zone lakes and tundra ponds (Martynov et al., 2012). One-dimensional (1-D) lake models are often used to couple with NWP and climate models to parameterize small lakes due to being cheap, computationally fast and simple. For temperate or warm climatic regions, the 1-D FLake model has been found to perform adequately for small lakes and also for large lakes with shallow depths (Martynov et al., 2010, 2012).

The FLake model (Mironov, 2005, 2008; Mironov et al., 2010), the model used in this thesis, is a computationally efficient model. Flake is used as a lake parameterization scheme in weather forecasting, climate modeling, and other numerical prediction systems for environmental purposes. This model is also used as a stand-alone lake model. The model is based on a two-layer water temperature profile. Mixed layer is at the surface and above the thermocline, which starts from the mixed layer's outer edge extending to the bottom of the lake. FLake is able to predict the vertical temperature profile and mixing conditions in a lake. The model is applicable to lakes with various depths at different time scales, from a few hours to a year. The water temperature structure in the thermocline is described based on the concept of self-similarity (assumed shape) of the temperature-depth curve. This concept is also used to describe the temperature profile of the thermally active layer of bottom sediments and the ice and snow cover. The adopted two-layer parameterization of the vertical temperature profile is:

$$\theta = \begin{cases} \theta_s, & 0 \le z \le h \\ \theta_s - (\theta_s - \theta_b)\Phi_\theta(\xi), & h \le z \le D \end{cases} \qquad \text{Equation 2.1}$$

whereas θ_s is the temperature of the upper mixed layer of depth $z = h$, θ_b is the temperature of the basin bottom at $z = D$. $\Phi_\theta(\xi) \equiv \frac{\theta_s - \theta}{\theta_s - \theta_b}$ is a dimensionless function of dimensionless depth $\xi \equiv \frac{z-h}{D-h}$.

The main external parameters in the model are the lake depth and optical characteristics of the lake water which is parameterized via the light extinction coefficient. For deep lakes, a virtual bottom depth of 40-60 m is typically used in simulations, instead of the real lake depth (Mironov, 2008). A sensitivity study was conducted by Potes et al. (2012) to investigate the importance of the extinction coefficient in the development of the lake surface temperature using FLake. The results showed the significant sensitivity of the model to different extinction coefficient values (Potes et al., 2012). Heiskanen et al. (2015) also investigated the sensitivity of FLake to changes in the extinction coefficient. Their results revealed that the model is very sensitive to variations in the extinction coefficient when it is lower than 0.5 m^{-1}. Formation of the thermal stratification layer depends strongly on the extinction coefficient.

This study concluded that a global mapping of extinction coefficients would be most beneficial in regions with relatively clear waters (e.g. high altitude lakes) (Heiskanen et al., 2015). Turbidity is measured frequently in lakes as part of water quality monitoring programs. However, these measurements are spatially and temporally limited. Thus, there is no dataset with global or regional coverage currently available on lake extinction coefficients. This is a challenging issue for the FLake model to consider the optical parameter of water, extinction coefficient, in the physics of the model. At this time, the extinction coefficient is considered as a constant value (Mironov et al., 2010). On the other hand, such a global mapping of extinction coefficients could be created using satellite remote sensing to provide the global characterization of lake optical properties. This is one of the topics examined in this thesis.

2.6 Study Site and Water Sampling Protocol

The Laurentian Great Lakes - Superior, Huron, Michigan, Erie and Ontario (Figure 2-6) - form the largest system of fresh, surface water lakes on earth, that contains 18% of the world supply of surface freshwater (Wetzel, 2001). The system provides drinking water, food, recreation, and transportation for a growing population of about forty million (Shuchman et al., 2013). The ecological state of the lakes, on both sides of the US-Canadian border, reflects the environmental health in the heart of the continent (Gons et al., 2008). In recognition of the importance of the lakes as a significant resource of fresh surface water and the need to restore and maintain their water quality, the Great Lakes Water Quality Agreement was signed between the two governments of Canada and the United States in 1972 to enhance the water quality in the Great Lakes basin ecosystem (Herdendorf, 1984).

Lake Erie, the study site for this thesis, is the smallest in volume (484 km^3), the shallowest (mean depth: 19 m, maximum depth: 64 m), and the most southern of the Great Lakes (Herdendorf, 1984). Due to the shallowness, Lake Erie warms quickly in the spring and summer and cools quickly in the fall. In long, cold winters, a large percentage of lake is covered with ice, and occasionally freezes over completely. However, in warm winters, there may be no ice at all. The warmer temperature and shallowness of Lake Erie basin make this lake the most biologically productive among the other Great Lakes (Lake Erie LaMP Work Group, 2000). About one-third of the total population of the Great Lakes basin reside within Lake Erie watershed, in total 11.6 million people (Daher, 1999). Being the most highly populated basin of the Great Lakes, Lake Erie surpasses all other lakes in the amount of effluent received from sewage treatment plants. It is also the most subjected to the sediment loadings, due to intensive agricultural development, in particular in the southwest Ontario and northwest Ohio. The

shallow nature of the lake basin makes it particularly vulnerable to loadings (Lake Erie LaMP Work Group, 2000).

Therefore, of all the Great Lakes, Lake Erie is exposed to the greatest stress from urbanization, agriculture, and industrialization. It was the first lake to demonstrate a serious eutrophication problem due to increased nutrient loadings beginning in the 1950s. Lake Erie has undergone considerable environmental changes over time due to different source of stresses, and the physical characteristics of the lake have a direct bearing on how the lake ecosystem react to all these various stressors (Lake Erie LaMP Work Group, 2000).

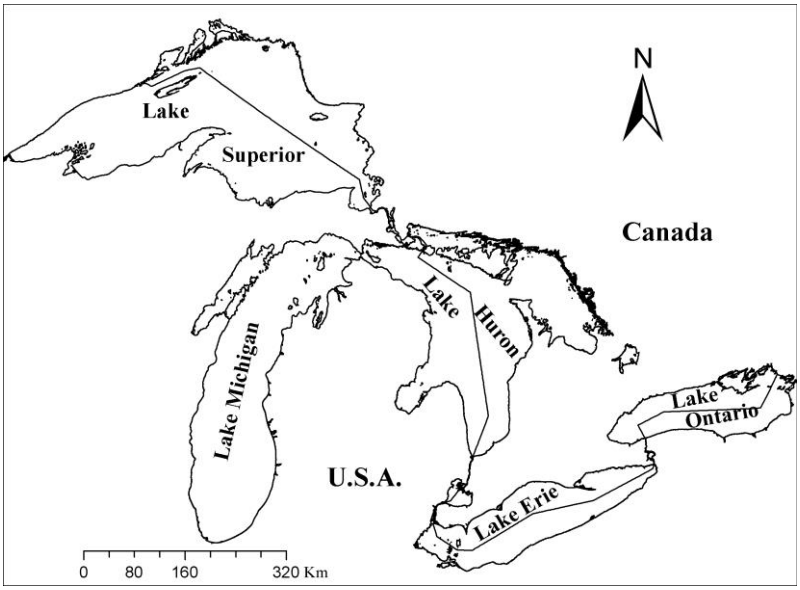


Figure 2-6 Location map and shape of the Great Lakes

Environment Canada Research cruises on board of the Canadian Coast Guard ship *Limnos*, the cruises that provided the in situ measurements of water quality parameters used in this thesis, collects water samples in different locations on Lake Erie, to provide measurements of a wide range of optical properties as well as concentrations of the main color-producing agents (CPA) for developing and validation of remote sensing algorithms as a complementary method for water quality monitoring programs.

Composite water samples are collected at all stations from the surface mixed layer of the lake using Niskin bottles. The samples are filtered through a Whatman GF/F fiber filter (0.7 μm) in the field. The filtered samples are then frozen and sent to the laboratory (the National Laboratory for Environmental Testing (NLET)) for extraction of the CPAs concentrations including chl-a. The chl-a measurement

method is based on the trichromatic spectrophotometry following fixation using a 90% acetone solution and centrifugation. Absorption of the residue at specified wavelengths of 663 nm, 645 nm, and 630 nm are determined. Chl-a, b and c values are calculated using SCOR/UNESCO equations in the analytical range of 0.1-100 mg m^{-3}. The following trichromatic Equations 2.2 are recommended by SCOR/UNESCO to measure chl concentrations:

$$chl\_a = 11.64\, e_{663} - 2.16\, e_{645} + 0.10\, e_{630} \qquad \text{Equations 2.2}$$

$$chl\_b = -3.94\, e_{663} + 20.97\, e_{645} - 3.66\, e_{630}$$

$$chl\_c = -5.53\, e_{663} - 14.81\, e_{645} + 54.22\, e_{630}$$

where chl-a (b or c) is in $\mu g.cm^{-3}$; and e_{663}, e_{645}, e_{630} are the absorbances (cm^{-1}) of light path at 663, 645, 630 µm after subtracting the 750 nm reading (UNESCO, 1966). The reported chl-a concentration also contains phaeopigments, which are degradation products of chl-a: phaeophytin and pheophorbide (Environment Canada, 1997). SDD measurement, a worldwide accepted method to estimate turbidity in water bodies, is regularly conducted during the cruises. Chl-a and SDD measurement methods follow the Ocean Optics Protocols for Satellite Ocean Color Sensor Validation (Mueller et al., 2003; Pegau et al., 2002).

2.7 Summary

In order to increase our knowledge and observing capabilities of water quality, satellite remote sensing is being used increasingly as a tool for monitoring waterbodies. The strength of remote sensing methods lies in their ability to provide both spatial and temporal views of surface water quality parameters that is typically not possible from in situ measurements alone. Remote sensing makes it possible to monitor water quality more effectively and efficiently, identifying waterbodies with significant water quality problems, so that it can support developing lake management strategies. MERIS satellite imagery measured the light emanating from the water surface at visible and near-infrared wavelengths and provided frequent acquisitions with sufficient spatial coverage for large lakes, such as Lake Erie. In particular, MERIS-type imagery is suitable for monitoring Case II waters with a sufficient (300 m) spatial resolution at nadir, with a 2-3 day revisit cycle, and with 15 spectral bands designed to allow for a more accurate retrieval of the secondary chl-a absorption maximum (e.g. at around 675 nm, more suitable for eutrophic waters).

Therefore, MERIS data was retained for this thesis to measure water quality parameters for Lake Erie, a complex water body with properties ranging from oligotrophic to mesotrophic/eutrophic (productive) and from clear to turbid in its different basins (from eastern to western basin, respectively). The parameters of interest include chl-a concentration and water turbidity. Chl-a is a photosynthetic pigment available in all kinds of phytoplankton. It is a trophic status indicator, and is used to measure phytoplankton biomass in waterbodies. Turbidity indicates the general environmental health of waterbodies, and is a measure of light penetration depth in lakes. The optical characteristics of lakes in lake models is parameterized using turbidity. However, no dataset with global or regional coverage is currently available for lake turbidity. Therefore, a constant value of this parameter is used in lake models. Satellite remote sensing has the potential to map lake optical properties globally, including extinction coefficient, and integrate it for simulations in lake models.

Chapter 3
Evaluation and Blending of MERIS Chlorophyll-a Retrieval Algorithms for Optically Complex Lake Erie

3.1 Introduction

Historically, the main application of ocean color satellites and bio-optical algorithms have been managed for the open ocean waters (Moore et al., 2014), where the optical properties are governed by the concentration of algal and other co-varying materials in water (referred to as Case I waters (Morel and Prieur, 1977)). Optical models developed for open ocean waters fail when applied for environments where the optical properties are dictated by materials additional to algal particles that do not co-vary with each other (the so called Case II waters). Hence, algorithms designed specifically for use in turbid inland and coastal waters, which are highly susceptible to Case II conditions (Moore et al., 2014), have been developed. Some of these algorithms are intended for use with MEdium Resolution Imaging Spectrometer (MERIS) data and are accessible in the Basic ERS & ENVISAT (A)ATSR MERIS (BEAM) toolbox available from the European Space Agency (ESA). These include artificial neural network (NN) processors and band-ratio approaches. The NN based processors include Case 2 Regional (C2R) (Doerffer and Schiller, 2007), Eutrophic (EU) (Doerffer and Schiller, 2008), FUB/WeW (Schroeder et al., 2007), and CoastColour (CC) processors (Ruescas et al., 2014). Two linear baseline algorithms including maximum chlorophyll index (MCI) and fluorescence line height (FLH) are the band ratio approaches (Gower et al., 1999, 2005; Gower and King, 2012).

Several studies have reported on the performance of algorithms implemented in BEAM for a variety of water bodies (Alikas et al., 2008; Giardino et al., 2010; Matthews et al., 2010; Odermatt et al., 2010; Vaiciute, 2012; Attila et al., 2013; Palmer et al., 2014; Kiefer et al., 2015; Sá et al., 2015). Palmer et al. (2014) applied six algorithms for the retrieval of chlorophyll-a (chl-a) over optically complex Case II waters of Lake Balaton, including NN and band differences algorithms. The NN processors (e.g. Boreal, EU, C2R, FUB/WeW processors) were found to much less accurately retrieve chl-a concentrations, whereas band difference algorithms of MCI and FLH performed best overall. In particular, FLH performed best especially for chl-a concentrations greater than 10 mg m^{-3} (Palmer et al., 2014). Sá et al. (2015) evaluated selected chl-a satellite products, including those from the CC processor (OC4, NN, merged). For 3-hour time intervals between in situ sampling and satellite overpass, CC products displayed different performance results to derive chl-a concentration for the

Western Iberian coast ($0.249<RMSE<0.278$, $0.139<MBE<0.200$, $0.78<R^2<0.85$, uncertainties were reported on the logarithmic scale). The CC_NN algorithm, suitable for coastal waters, provided better results than CC's other chl-a products for Iberian West coast (Sá et al., 2015). Kiefer et al. (2015) applied the FUB/WeW processor to retrieve chl-a concentrations of Lake Geneva. Their study confirmed that this algorithm is more suitable for use over oligo- to mesotrophic lakes, as opposed to eutrophic lakes (Kiefer et al., 2015). A highly variable accuracy and performance is found in these studiess applying the processors on different lakes. The variable results underline that the processor performance is highly related to the optical properties of the water body, and the data range used to train the NNs. Although, the difference in methods used to collect in situ observations of chl-a concentration can also cause divergence in the results.

Moore et al. (2014) proposed an optical water type framework for optically complex waters. In their study, two algorithms were considered: (1) the NASA OC4 algorithm based on blue/green bands, and (2) a MERIS 3-band algorithm based on red/near-infrared (NIR) bands. Higher uncertainty was found for the OC4 algorithm when chl-a concentration exceeded 10 mg m^{-3}, whereas better performance was obtained for the MERIS 3-band algorithm for higher concentrations. The authors proposed a blending approach based on optical water types to manage a selection between these two chl-a algorithms. In this approach, different optical classes of water body are derived from a fuzzy logic classification scheme applied to satellite-derived water leaving reflectance (Moore et al., 2001, 2014).

More recently, Lyu et al. (2015) analyzed different optical characteristics of five inland lakes, located in different basins in China, based on MERIS-derived remote sensing reflectance and applying an automatic two-step clustering method to derive optical water classes. The inland waters were clustered into three optical classes. The optical characteristics of water types I and III were dominated by chl-a concentration and inorganic suspended matters, respectively. Water type II had a combined optical characteristics of water types I and III. Four different chl-a retrieval algorithms including: two-band, three-band, four-band, and Synthetic Chlorophyll Index (developed by Shen et al. (2010) exclusively for estuaries with high total suspended matter (TSM) concentrations, and without a requirement for recalibration of band locations) were evaluated for the three water types. Different levels of accuracy were reached for each algorithm applied to different optical types, indicating that no single algorithm could be successfully applied to all water types.

The present study assesses the performance of six chl-a concentration retrieval algorithms implemented in the BEAM software: C2R, EU, FUB/WeW, CC NN processors as well as FLH and

MCI indices. These algorithms represent two fundamentally different approaches to derive chl-a concentration in optically complex waters. The NN approaches are trained to varying ranges of parameter concentration and optical property; whereas the indices make use of the height of a peak above a baseline, related to chl-a concentration. The algorithms are evaluated for Lake Erie, an optically complex water body with properties ranging from oligotrophic to mesotrophic/eutrophic (productive) and from clear to turbid in its different basins (from eastern to western basin, respectively). The lake is subject to sediment re-suspension. River inputs carry colored dissolved organic matters (CDOM) into the lake. Floating harmful algal blooms are also characteristic of Lake Erie. Thus, there is high potential for a variety of optical water types to exist in the lake (Moore et al., 2014). In this paper, an extensive comparison of different chl-a retrieval algorithms applied to MERIS data and their performance is conducted for Lake Erie. A classification approach is then applied to derive clusters with different optical characteristics using remote sensing reflectance. Finally, a blending approach is used to combine the results of chl-a products derived from different algorithms assigned to each water class.

3.2 Data and Methods

3.2.1 Study Area

The lake of interest in this study is Lake Erie, one of the Great Lakes of North America (Figure 3-1). Lake Erie has a surface area of 25,700 km^2 with an average depth of 19 m. It is the fourth largest in surface area, the most southern, the shallowest and the warmest of the Laurentian Great Lakes (Anderson et al., 2015). The lake's high productivity as well as the warm weather in its watershed has attracted one-third of the total human population of the Great Lakes basin. Human activities are responsible for several environmental problems such as floating harmful algal blooms, and chemical contaminations that are carried into the lake from river inputs, considered as a source of dissolved organic/inorganic matters (Djoumna et al., 2014). In addition, Lake Erie is subject to sediment re-suspension due to its shallow depth, and consequently suspended particulate matters play a large role in Lake Erie's low water clarity (Binding et al., 2010).

The various dissolved and suspended water constituents of Lake Erie determine the variability of its water leaving reflectance. Hence, the lake is considered as optically complex and classified as a Case II water body. Lake Erie trophic status ranges from oligotrophic to mesotrophic/eutrophic in its different basins (from eastern to western basin, respectively). Total phosphorus (TP) loads from point and nonpoint sources result in algal blooms and poor water clarity in the lake. The contribution of TP

in changing the trophic state of the lake and its turbidity is different through a year. Based on forty years of historical data collected before 2010, Lake Erie Nutrient Science Task Group (2009) reported a higher correlation between the three parameters of TP, chl-a concentration (a trophic status indicator), and Secchi disk depth (SDD, a water clarity indicator) in summer rather than spring. More frequent sediment loads and re-suspension of lake sediments (due to the shallow depth of Lake Erie) in spring delivers higher particulate phosphorus from soil erosion, which is not readily available for algae uptake. Hence, although there is low correlation between TP and chl-a concentration in Lake Erie during spring (R^2=0.20), there is still a good overall relationship between TP concentration and SDD (R^2=0.63). Conversely, in summer, TP concentration contains more bioavailable phosphorus, which result in higher correlation between all three parameters compared to spring (chl-a and TP: R^2=0.71; SDD and TP: R^2=0.75; SDD and chl-a: R^2=0.65) (Lake Erie Nutrient Science Task Group, 2009).

Phosphorus loads are highly variable in the different basins of Erie. The International Joint Commission, signed between Canada and the United States to protect shared water bodies such as Lake Erie, reports that 64% of the 2003-2011 average loads are received in the western basin, while 26 and 11% of the loads are received in the central and eastern basins, respectively. A general trend of decreasing phosphorus concentration from west to east and from near-shore to the offshore is observed in the lake (IJC. International Joint Commission Canada and United States, 2013).

Recent increases in total phosphorus and chl-a concentrations have been reported, especially in the western basin of Lake Erie (Michalak et al., 2013). This highlights the importance of ongoing monitoring programs that track parameters indicative of trophic state with the intent of reducing phosphorus loads into the lake. Research cruises on board the Canadian Coast Guard ship *Limnos* provided measurements of chl-a concentration and optical properties for this study from a total of 89 distributed stations in Lake Erie during September 2004, May, July, and September 2005, May and June 2008, July and September 2011 and February 2012 (Figure 3-1). In situ chl-a measurements ranged between 0.2 and 70.1 mg m^{-3}, with an average of 4.26 ± 7.84 mg m^{-3}.

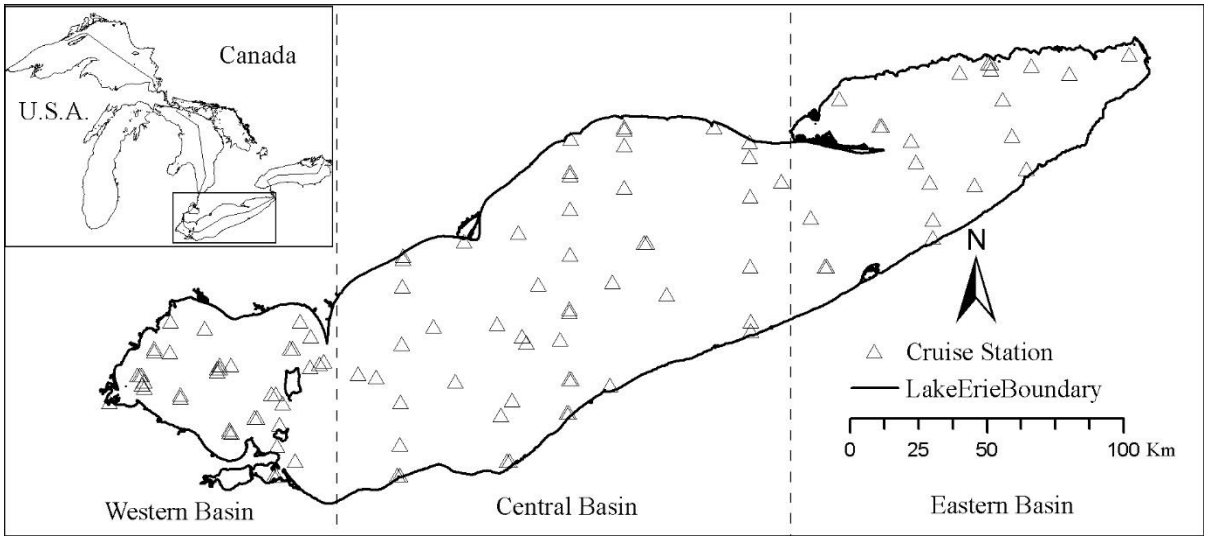


Figure 3-1 Location of Lake Erie in the Laurentian Great Lakes (inset), and the sampled points of the Canadian Coast Guard ship *Limnos* during 2004 to 2012, in situ data used in this study.

3.2.2 Algorithms Evaluated

In the present study, the performance of four NN processors including C2R, EU, FUB/WeW, CC are evaluated to estimate chl-a concentration. Band difference algorithms developed using FLH and MCI indices are also tested. The details of the algorithms are provided in the following sections.

3.2.2.1 Neural Network Algorithms

The relationship between the water-leaving reflectance and inherent optical properties (IOP) of water constituents is known from radiative transfer theory (Carder et al., 2004; A. A. Gitelson et al., 2008; Tilstone et al., 2012). The inversion of a radiative transfer model can be approximated with a NN approach. A NN approach is adopted to achieve the requirements of high accuracy and, at the same time, efficiency in the operational processing of MERIS satellite imagery to derive the optical properties (Doerffer and Schiller, 2007). The processing toolbox for MERIS (BEAM software) incorporates many of these NN-based methods for interpretation of remote sensing signals over Case II waters including: C2R, EU, FUB/WeW and CC modules.

Doerffer and Schiller (2007) described the C2R processor, designed for more optically complex conditions in Case II waters. This processor is based on a coupled forward and backward NN architecture. A separate NN first performs atmospheric correction. Eight atmospherically corrected MERIS spectral bands, as well as geometric information, are then used as input in the coupled NN to

perform the in-water algorithm. The special feature of combining the forward and backward models allows to test if the measured spectrum is within the scope of the training set. The networks are trained using the IOPs and water constituents' concentration data that are collected from European waters, mainly in the North Sea as well as the North Atlantic, Baltic Sea, and Mediterranean Sea. The established bio-optical dataset was followed by Hydrolight radiative transfer simulation of a large dataset of reflectance to use as training data. This processor results in the generation of three IOPs at MERIS band 2 (443 nm) including: (1) absorption coefficient of phytoplankton pigments, (2) absorption coefficient of yellow substance and (3) scattering coefficient of particles. The derived IOPs are converted to the concentration of chl-a and TSM (Doerffer and Schiller, 2007).

The EU lake processor in BEAM replicates the atmospheric correction and the coupled NN architecture of the C2R processor. It also uses the same input of the water leaving reflectance of 8 MERIS bands derived from atmospheric correction. However, it is trained using a different range of bio-optical data collected from eutrophic Spanish lakes (Doerffer and Schiller, 2008). The EU and C2R lake processors also provide the atmospherically corrected water-leaving reflectance for MERIS bands 1-10, 12 and 13.

The FUB/WeW processor is also based on inversion of a radiative transfer model applying NNs. However, unlike the C2R and EU lake processors it consists of four NNs. FUB/WeW is a one-step coupled atmospheric correction and in-water constituent retrieval algorithm (Kratzer et al., 2008), where MERIS measured radiances are used as input to all four networks. Atmospheric correction is performed in one network separately. Each water constituent is retrieved individually from the other three networks and directly from the top-of-atmosphere measured radiances. Concurrent in situ data, collected from North Sea, Gulf of Cadiz in the North Atlantic, Baltics Sea, and Mediterranean Sea close to Israel border, was used for validation of the water retrieval algorithms. This processor was specifically designed for European coastal waters to derive the atmospherically corrected water-leaving reflectance for MERIS bands 1-7 and 9; as well as the concentrations of water constituents including chl-a and TSM; and the absorption of yellow substance at 443 nm (Schroeder et al., 2007).

The CC processor employs a wider range of optical properties as the training data compared to the other NN processors in BEAM. The BEAM NN processors' specific training ranges for chl-a, TSM concentration and CDOM absorption are compared in Table 3-1. Also, the CC algorithms benefit from a newly developed atmospheric correction algorithm optimized for coastal areas (Sá et al., 2015) that represent optically-complex waters present in inland lakes. There are different chl-a products generated

by the CC module implemented in BEAM. One product uses NN as the inversion approach and is an advanced version of MERIS algal_2 product (CC_NN). Another one is based on empirical NASA-OC4 band ratio algorithm (CC_OC4) (Sá et al., 2015). In addition, a blending of the OC4 (for clear waters) and NN (for turbid waters) algorithms is available as another product (CC_merged) of CC processor. CC_OC4 product is used when the TSM concentration is below 5 g m^{-3}. For TSM concentration above 10 g m^{-3} CC_NN is calculated. A weighting average between chl-a products from CC_OC4 and CC_NN is used to estimate chl-a for TSM concentrations between 5 and 10 g m^{-3}. Thus CC_merged is a smooth transition between two algorithms (Ruescas et al., 2014).

Table 3-1 Chl-a, TSM and CDOM training ranges of the NN processors (source: Palmer et al. (2014); Ruescas et al. (2014)).

Processor	Chl-a (mg m^{-3})	TSM (g m^{-3})	CDOM (a_{440} m^{-1})
EU	1-120	0.42-50.9	0.1-3
C2R	0.016-43.18	0.0086-51.6	0.005-5
FUB/WeW	0.05-50	0.05-50	0.005-1
CC	0.001-100	0.2-2000	-

3.2.2.2 Band-Ratio Algorithms

Band-ratio algorithms incorporating red and NIR water-leaving radiances, such as FLH and MCI, have been applied in coastal, oceanic and inland waters with varying degrees of success (Binding et al., 2010; Gons et al., 2008; Matthews et al., 2010; Palmer et al., 2015; Stumpf et al., 2012). These algorithms consider the shape of MERIS spectral radiances at red and NIR to measure the enhancement of radiance at a certain wavelength above a baseline which is drawn between the radiance values at two suitable wavelengths (Gons et al., 2008). This is especially useful for Case II waters in order to reduce the masking of pigment spectral features by concentrations of CDOM and TSM, which is stronger in blue and green spectral regions (Gower et al., 1999).

FLH measures the relative spectral peak at 685 nm which is caused by solar-stimulated fluorescence from chl-a (Gower et al., 1999; Gower and King, 2012). FLH can be used effectively to measure near-surface phytoplankton biomass up to a concentration of 20 mg m^{-3} (Gower and King, 2012). In higher concentrations, any further FLH increase with the concentration is prevented by absorption of stimulated and emitted radiation. However, a radiance relative peak will form in the wavelength range of 700-710 nm, measured as MERIS MCI (Gower and King, 2012). Similar to FLH, MCI measures a peak but near 708 nm which is a characteristics of intense surface algal blooms (Gower et al., 2005; Gower and King, 2012). The FLH and MCI algorithms are formulated as:

$$FLH/MCI = L_2 - L_1 - (L_3 - L_1) \times \frac{(\lambda_2 - \lambda_1)}{(\lambda_3 - \lambda_1)}$$ Equation 3.1

whereby L is the radiance or reflectance, depending on the use of MERIS Level 1b or atmospherically corrected Level 2 data (Palmer et al., 2014). FLH measures the peak at MERIS band 8 (680.5 nm) above the baseline between bands 7 (664 nm) and 9 (708 nm). MCI measures the relative peak of MERIS band 9 (708 nm) above band 8 (680.5 nm) and band 10 (753 nm) (Palmer et al., 2014). The height-above-baseline algorithms can be applied with or without atmospheric correction (Matthews et al., 2010; Binding et al., 2011). In this study, we investigated Level 1b (which is a product level prior to atmospheric correction) FLH and MCI.

3.2.3 Processing of MERIS data

MERIS full resolution images of Lake Erie for the period 2004-2012 were inspected to visually select images that are at least partially cloud free. The selected images corresponded to dates within a 2-day time window of in situ data collection. Thus, nine of the MERIS Level 1 images downloaded from the European Space Agency (ESA)'s archive were analyzed and processed within the BEAM v5.0 (Brockmann Consult GmbH) toolbox. Quality control flags were applied on pixels prior to extract and use them in matchup analysis to compare chl-a products with field measurements.

In all algorithms, Level 1 flagged pixels were excluded from matchup analysis. For C2R, EU, and FUB/WeW processors, the NN Level 2 flagged pixels were also excluded. Pixels flagged in Level 1P and Level 2 CC processing (atmospheric correction (L2R) and in-water algorithms (L2W) are applied in Level 2) were excluded from further processing (Table 3-2). The remaining pixels were used in the matchup dataset for training and evaluating the algorithms (see section 3.2.5 for details).

Table 3-2 Flags of excluded pixels

Level of Data	Flags
Level 1	“invalid”, “coastline”, “land_ocean” (pixel is over land), “bright”, “suspect”, “glint_risk” (at risk of glint)
Level 2-C2R/EU	“land”, “cloud_ice” (cloud or ice), “atc_oor” (atmospheric correction out of range), “tosa_oor” (top-of-standard-atmosphere out of range), “toa_oor” (top-of-atmospheric-reflectance out of range), “solzen” (large solar zenith angle), “wlr_oor” (water leaving reflectance out of scope), “conc_oor” (concentration out of training range), “ootr” (spectrum out of training range), “whitecaps”, “invalid”
Level 2-FUB	“Level 1b_masked”, “Chl_in” (input retrieval failure of chl-a), “Chl_out” (output retrieval failure of chl-a), “Yel_in” (input retrieval failure of yellow substance), “Yel_out” (ouput retrieval failure of yellow substance), “TSM_in” (input retrieval failure of total suspended matter), “TSM_out” (output retrieval failure of total suspended matter), “atm_in” (input retrieval failure of atmospheric correction), “atm_out” (output retrieval failure of atmospheric correction)
Level 1P-CC	“CC_land”, “CC_cloud”, “CC_cloud_ambigious”, “CC_cloud_buffer”, “CC_cloud_shadow”, “snow_ice”, “CC_mixedpixel”
Level 2-CC	“solzen” (large solar zenith angle), “c2r_whitecaps”, “aot560_oor” (atmospheric correction out of range), “toa_oor” (top-of-atmosphere reflectance out of range), “tosa_oor” (top-of-standard-atmosphere out of range), “nn-wlr_oor” (water-leaving-reflectance out of training range), “nn_conc_oor” (water constituents out of training range), “nn_ootr” (spectrum out of training range with a default threshold set on chi-square)

3.2.4 Clustering Method Using Remote Sensing Reflectance

The two-step clustering scheme, available in SPSS V20, is an unsupervised classifier which does not require a foreknowledge of the classes characteristics. The two-step clustering approach is a combination of the two hierarchical and relocation clustering methods. In the first step of clustering analysis in this method, the data is pre-clustered into many small sub-clusters. Then in the next step, the sub-clusters from the first step are combined together or divided into smaller clusters to produce the desired number of clusters. The similarity between two clusters is calculated using the log-likelihood distance. The distance between two clusters decreases as they are combined into one cluster. The predictor importance, which is the relative importance of each predictor in estimating the model, is provided as an output of the two-step clustering approach. The predictor importance in SPSS is based on variable importance statistics. It indicates how well each variable can differentiate different clusters. Since the values are relative, the sum of all predictors’ importance values equals to one.

The two-step clustering is capable of handling large datasets, both continuous and categorical variables. Also, in case of unknown desired number of clusters, the two-step clustering can determine

the proper number of clusters automatically (SPSS, 2001) using a combination of both the distance changes and either Akaike Information Criterion (AIC) (Akaike et al., 1998) or Bayesian Criterion (BIC) (Schwarz, 1978) methods.

Here, we selected remote sensing reflectance in MERIS bands 1-10, 12, and 13 in the matchup data as the predictors for clustering. AIC was used as the method of finding the optimal number of clusters. The optical feature of the derived optical classes were studied. Furthermore, the performance of different chl-a retrieval algorithms were assessed for each single optical type. A blending approach was then used to choose the most appropriate algorithm for each given water optical condition.

3.2.5 Chlorophyll-a Retrieval Algorithms Calibration and Error Definitions

The resulted matchup dataset was randomly divided into training and testing data in 100 iterations. In each iteration, the data sets used for algorithms training and evaluation were kept independent, where 70% of the matchups were used for algorithm calibration and 30% for evaluation. Ordinary least square regression equations were used in the calibration step of each iteration to relate the in situ measurements of chl-a to the derived chl-a products from C2R, EU, FUB/WeW, and CC processors and also MCI/FLH indices. Locally tuned equations were derived from this step and were applied on the algorithms' products to predict chl-a concentration in testing matchup data. The statistical parameters of each model performance were derived between the predicted chl-a concentration and measured ones in the field for testing data. These steps were repeated for 100 iterations; and the final statistical indices, slope and intercept of the locally tuned algorithms were reported as the average of the ones over all iterations.

Satellite-derived chl-a concentration estimates were evaluated against in situ data using statistical indices, including the index-of-agreement (I_a) and two dimensioned statistics of average model performance error: the root-mean-square-error (RMSE) and the mean bias error (MBE) (Willmott and Matsuura, 2005). The dimensionless I_a is used to report the model performance and compare it with the other ones. The models in this study are the chl-a retrieval processors. A refined version of I_a, the one adopted in this study, has recently been proposed by Willmott et al. (2012). The index is bounded between -1 (the worst model performance) and +1 (the best model performance) (Willmott, 1981; Willmott et al., 2012). RMSE is a comprehensive metric that combines the mean and variance of model errors into a single statistic (Moore et al., 2014). MBE is predicted (modeled) values minus the observed in situ data. It therefore intends to report the average bias of the model which is the average over- or under-prediction (Willmott and Matsuura, 2005).

3.3 Results and Discussion

3.3.1 In situ Measurements

Table 3-3 summarizes the water quality indicators collected on Lake Erie during *Limnos* cruises in 2004 to 2012 for all basins combined, as well as on a per basin basis. A chl-a concentration gradient from the east to the west basin is apparent with an average value of 1.08 ± 0.4, 2.41 ± 2.07, and 8 ± 11.82 mg m^{-3} in the east, central, and west basins respectively. SDD decreases from 6.5 ± 2.19 m in the east basin to 4.97 ± 2.53 m in the central basin, and 1.74 ± 1.28 m in the west basin. Therefore, there is a water clarity gradient from west to east that is consistent with the gradients of in situ chl-a concentration employed in this study, and also the TP loads reported in the International Joint Commission Canada and United States (2013).

Table 3-3 Descriptive statistics of full range in situ data *Limnos* collected from 2004 to 2012 on Lake Erie.

Basin	Stat	Chl-a (mg m^{-3})	TSM (g m^{-3})	CDOM (m^{-1})	SDD (m)
All	N	186	186	156	113
	Min	0.20	0.18	0.04	0.20
	Mean	4.31	5.70	0.32	3.73
	Max	70.10	50.50	2.36	11.00
	S.D.	7.90	7.97	0.34	2.71
Central	N	88.00	88.00	72.00	55.00
	Min	0.30	0.29	0.05	0.30
	Mean	2.41	3.33	0.22	4.97
	Max	12.00	38.24	0.74	11.00
	S.D.	2.07	5.50	0.14	2.53
Eastern	N	28.00	27.00	19.00	10.00
	Min	0.20	0.18	0.04	4.00
	Mean	1.08	1.84	0.15	6.50
	Max	1.90	11.10	0.28	10.00
	S.D.	0.40	2.57	0.08	2.19
Western	N	70.00	71.00	65.00	48.00
	Min	0.30	1.10	0.13	0.20
	Mean	8.00	10.09	0.48	1.74
	Max	70.10	50.50	2.36	6.00
	S.D.	11.82	9.80	0.45	1.28

The west and central basins of Lake Erie are shallow enough to have wave-driven sediment re-suspension (Lake Erie Nutrient Science Task Group, 2009). Table 3-3 shows higher amounts of

suspended matters in the west and central basins with averages of 10.09 ± 9.8 and 3.33 ± 5.5 g m^{-3}, respectively, as measured during the *Limnos* cruises

3.3.2 Matchup Data Result

Table 3-4 shows the descriptive statistics of the matchup data after removal of flags corresponding to different algorithms, for all basins lumped together. The total number of matchup data derived for each algorithm was variable due to the specific flags applied in different algorithms (described in section 3.2.3). All algorithms were flagged similarly in level 1. Level 2 flags were the same for the EU and C2R lake processors. The FUB/WeW NN processor excluded more pixels compared to other NN processors (N=96). FLH and MCI band ratio algorithms resulted in a larger number of matchups (N=117) as only level 1 flags are raised by these algorithms. According to the in situ data, chl-a concentration varies between 0.2 mg m^{-3} and 70.1 mg m^{-3} with the average of 4.31 mg m^{-3} in the study period (Table 3-3). Therefore, FLH and MCI processors covered a larger variation of chl-a considering the remaining matchup data in these algorithms (chl-a: min=0.2 mg m^{-3}, max=61.7 mg m^{-3}). Whereas FUB/WeW processor, with the lowest number of matchup data, covered only chl-a concentration between 0.2 mg m^{-3} and 12 mg m^{-3}. However, the average chl-a concentration was close in all matchup data for different chl-a retrieval algorithms.

Table 3-4 Descriptive statistics of in situ matchup subset data after removing those flagged in each algorithm. Chl-a and TSM concentrations are reported in mg m^{-3} and g m^{-3}. CDOM absorption in m^{-1}.

Processor	Parameter	N	Min	Max	Avg.	St. Dev.
EU-C2R	Chl-a	109	0.20	23.10	3.35	3.53
	TSM	108	0.18	50.50	5.46	7.94
	CDOM	78	0.04	1.75	0.28	0.31
FUB/WeW	Chl-a	96	0.20	12.00	3.16	2.63
	TSM	96	0.29	35.73	5.12	6.47
	CDOM	67	0.04	1.75	0.28	0.32
CC	Chl-a	103	0.20	23.10	3.46	3.59
	TSM	102	0.18	50.50	6.29	8.84
	CDOM	78	0.04	1.75	0.29	0.30
MCI	Chl-a	117	0.20	61.70	3.78	6.39
	TSM	116	0.18	50.50	5.90	8.54
	CDOM	86	0.04	1.75	0.29	0.30
FLH	Chl-a	117	0.20	61.70	3.77	6.38
	TSM	116	0.18	50.50	5.71	8.41
	CDOM	86	0.04	1.75	0.29	0.30

3.3.3 Chlorophyll-a Products Evaluation

The locally tuned equations relating NN chl-a products and FLH/MCI indices to in situ measurements of chl-a concentration were retrieved by applying a linear regression on randomly selected 70% of the matchup in situ data. These equations are shown in Table 3-5. The remaining 30% of matchup in situ data were used to derive model performance statistical indices and compare the accuracy of different algorithms to derive chl-a concentration. Retrieval performance of each is presented in Table 3-6. The C2R processor was found to perform the best in terms of I_a (0.69), with an improvement of about 6% over FUB/WeW processor (I_a=0.65); although the C2R processor underperformed FUB/WeW by about 11% based on RMSE results (C2R: RMSE = 2.21 mg m^{-3}, FUB/WeW: RMSE = 1.99 mg m^{-3}).

Overall, the NN processors show a marginal difference in chl-a retrieval accuracy relative to each other. Variation in the performance of models can be attributed to the algorithms' training ranges. A comparison of Table 3-2 and Table 3-3 reveals that the optical properties measured in Lake Erie are represented in the training range of the C2R processor, whereas low levels of chl-a, CDOM and TSM concentrations in Lake Erie are not covered in the training range of the EU processor. Also, the training range of FUB/WeW does not cover high concentrations of CDOM matchup in situ data.

Table 3-5 Locally tuned equations for the full lake derived in training step. N is the number of training data which is 70% of the matchup in situ data after removing those flagged by the algorithm. Chl_aprodcut is the product of each algorithm related to the chl-a concentration, whereas Chl_apredicted is the one derived from applying the locally tuned equations.

Basin	Processor	N	Locally tuned equation
Full Lake	EU	75	$Chl\_a_{predicted}=0.69\times Chl\_a_{prodcut} + 0.10$
	C2R	75	$Chl\_a_{predicted}=0.70\times Chl\_a_{prodcut} - 0.16$
	FUB/WeW	68	$Chl\_a_{predicted}=3.03\times Chl\_a_{prodcut} + 1.68$
	CC_OC4	71	$Chl\_a_{predicted}=0.99\times Chl\_a_{prodcut} - 0.41$
	CC_NN	71	$Chl\_a_{predicted}=0.28\times Chl\_a_{prodcut} + 2.01$
	CC_merged	71	$Chl\_a_{predicted}=0.28\times Chl\_a_{prodcut} + 1.82$
	FLH	80	$Chl\_a_{predicted}= -6.95\times FLH + 3.93$
	MCI	80	$Chl\_a_{predicted}= 2.97\times MCI + 4.62$

Table 3-6 Chl-a retrieval performance statistical indices for each algorithm derived in testing step for the full lake. Chl_a predicted derived from tuning equation is compared to the in situ chl-a concentration. N is 30% of matchup in situ data after removing those flagged by the processor.

Basin	Processor	N	RMSE (mg m^{-3})	MBE (g m^{-3})	I_a
Full Lake	EU	34	2.40	-0.01	0.67
	C2R	34	2.21	-0.02	0.69
	FUB/WeW	31	1.99	-0.01	0.65
	CC_OC4	32	2.50	0.03	0.64
	CC_NN	32	2.28	0.02	0.64
	CC_merged	32	2.32	0.03	0.63
	FLH	37	5.45	-0.15	0.44
	MCI	37	5.29	0.00	0.52

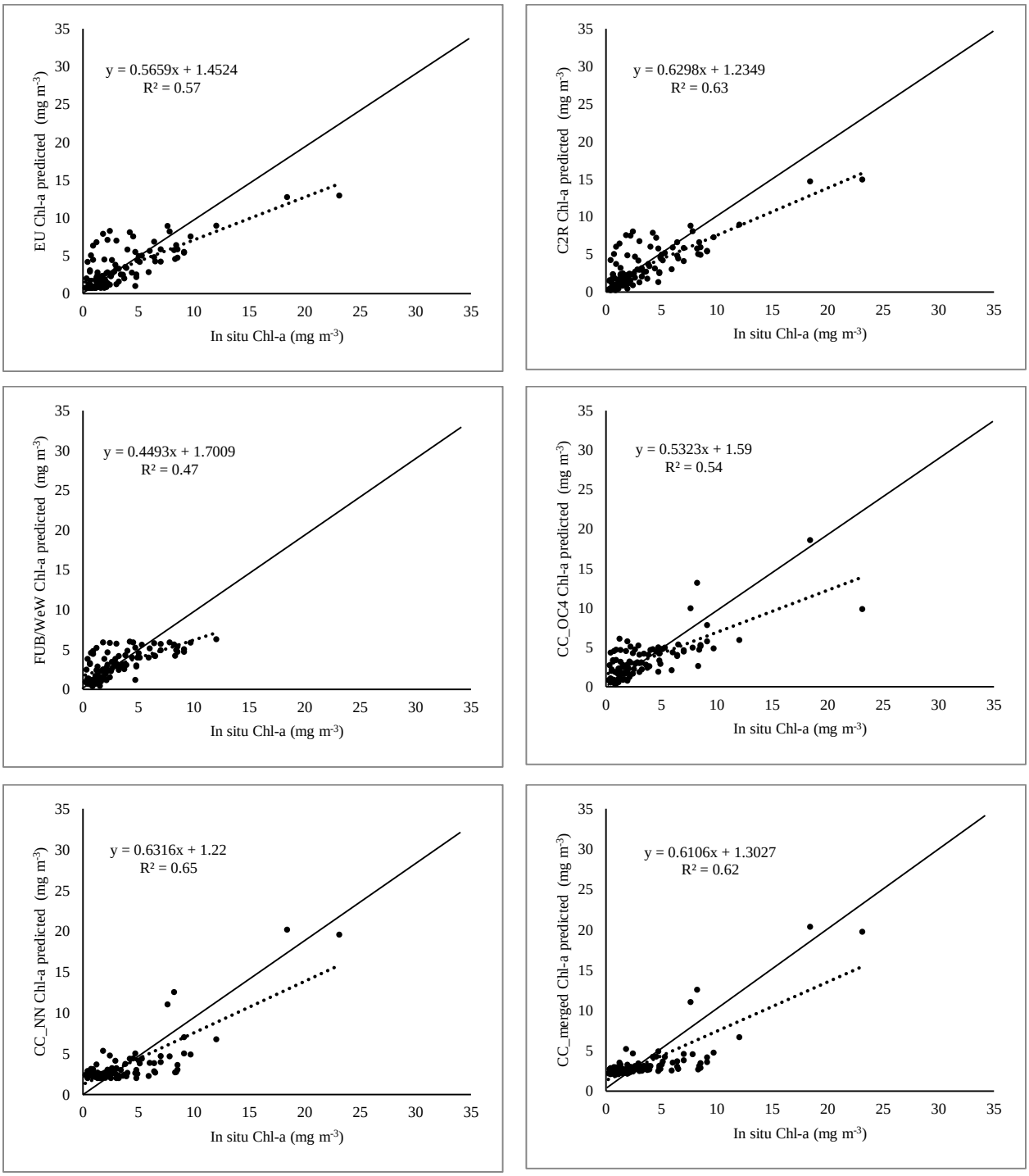
FLH derived chl-a has the lowest accuracy (I_a = 0.44, RMSE = 5.45 mg m^{-3}) with a negative slope from the derived tuning equation. Gons et al. (2008) found a strong linear relationship between chl-a concentration and FLH for oligotrophic waters of Keweenaw Bay (Lake Superior). However, FLH poorly predicted chl-a concentration for eutrophic to hyper-eutrophic waters of Green Bay (Lake Michigan). In mesotrophic to eutrophic and hyper-eutrophic waters, FLH diminishes to become negative when chl-a increases. This is because the absorption of chl-a signal becomes dominant at 681 nm for high concentrations (Gons et al., 2008). Binding et al. (2011b) and Palmer et al. (2014) also found a negative correlation between in situ chl-a measurements and the FLH index in their studies. At high chl-a concentration, FLH is more representative of an absorption line depth instead of being a FLH (Binding et al., 2011b; Palmer et al., 2014). The other possibility for the reduced fluorescence is the contribution of cyanobacteria chl-a molecules. Most of these molecules belong to non-fluorescing

photosystems. This results in the absence of a distinct FLH signal (Binding et al., 2011b). MCI values estimated from MERIS Level 1b on Lake Erie tend to be negative. This happens when there is a radiance peak near 685 nm and FLH is high. But MCI has positive values when the peak shifts to longer wavelengths around 705 nm (Gower et al., 2005). The model results of Gower et al. (2005) show that MCI positive values (a peak in the wavelength range of 700-710 nm) occur in the case of an intense algal bloom with a chl-a concentration of above ca. 100 mg m^{-3} (Gower et al., 2005). The MCI-derived chl-a product results in a low accuracy in the present study (I_a = 0.52, RMSE = 5.29 mg m^{-3}). This is due to the fact that there is not enough high concentration of chl-a to form the MCI peak.

The predicted chl-a concentrations obtained by applying the locally tuned equations are shown in Figure 3-2 against the in situ measurements for all matchup data. The CC_merged product shows similar results to the CC_NN product. This indicates that the merging was mostly based on the CC_NN product, which is developed for turbid waters (e.g. Lake Erie).

Between the C2R and the FUB/WeW processors, the former underperformed the latter (11% failure in RMSE values versus 6% improvement in I_a values). Thus, the FUB/WeW processor was selected to map chl-a concentration for a selected day in September 2011, when Lake Erie experienced the largest algal bloom in its recorded history before the 2015 bloom (NOAA, 2015) (Figure 3-3). The peak intensity was over three times greater than any observed bloom in the past in 2011 (Michalak et al., 2013). The FUB/WeW processor captures the west-to-east gradient trend of trophic condition in Lake Erie. Also, the map reveals the bloom extent and spatial feature details during the bloom event in September 2011. It shows the bloom expanding into the central basin from the western basin. It also displays the second phase of bloom that was forming along the northern shorelines of the central basin (Michalak et al., 2013). Michalak et al. (2013) reported that the bloom in early September 2011 was at least 2.4 times greater than the previous largest bloom of 2008. The bloom was four times larger than the average bloom from 2002 to 2010. The bloom continued to grow until mid-October (Michalak et al., 2013). Michalak et al. (2013) stated that the record-breaking nutrient loads in Lake Erie was associated with the long-term trends in agricultural practices. The nutrient loads resulted in increasing phosphorus loading into the western basin. The high concentration of phosphorus was also coupled with special meteorological conditions (severe precipitations) in spring 2011. On the other hand, the unusual strong re-suspension immediately after the bloom onset, as well as the uncommonly warm condition in late spring, are hypothesized in Michalak et al. (2013) as the factors providing the ideal

conditions for further algal growth and bloom development on Lake Erie in that year (Michalak et al., 2013).



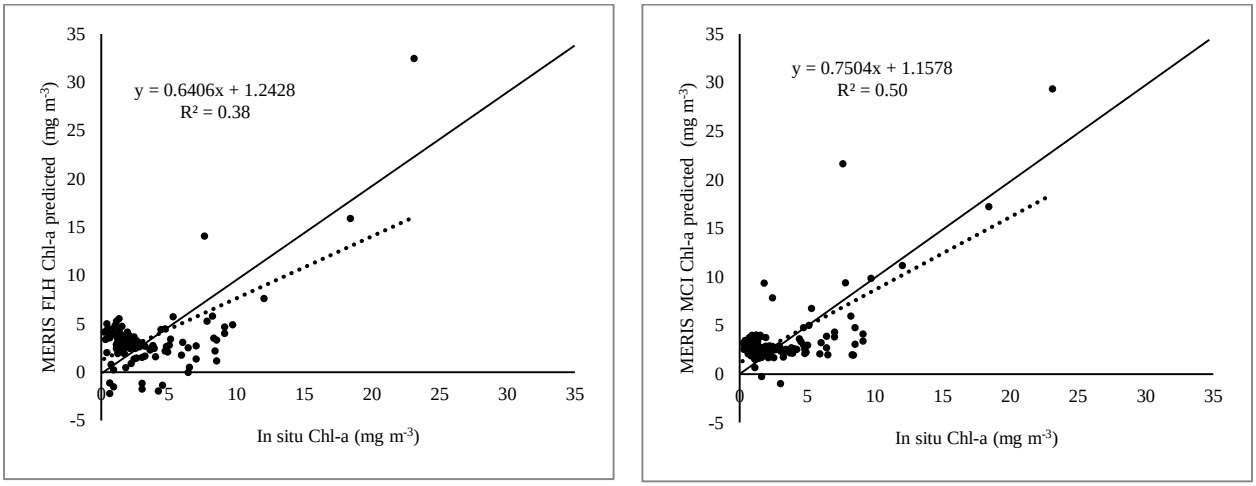


Figure 3-2 Chl-a retrieved from each algorithm after tuning for the full lake relative to the matchup in situ data. Solid line corresponds to 1:1 relationship.

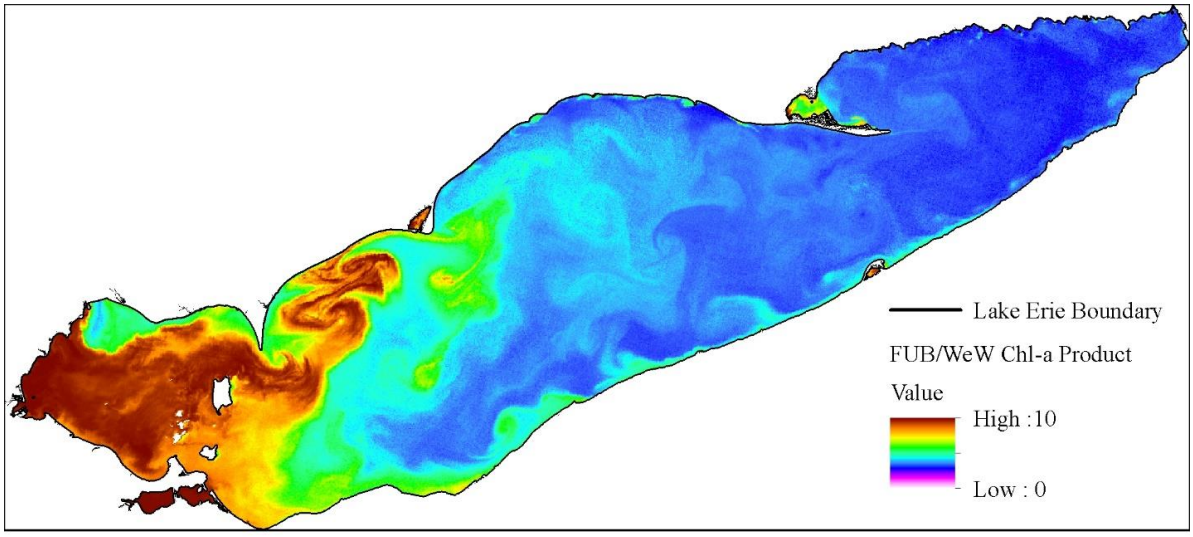


Figure 3-3 Chl-a concentration mapping by locally tuned FUB/WeW algorithm during a bloom event in September 3, 2011.

3.3.4 Clustering Analysis Results

Because there were no in situ measurements of reflectance available in our study for evaluation purpose, the NN processor resulting in the best relationship with field measurements of water quality parameter (chl-a concentration) was used as the best estimator of remote sensing reflectance. This NN was selected between EU, C2R, and CC processors which are the only NN processors among the algorithms used in this study that produce the remote sensing reflectance for MERIS bands 1-10, 12, and 13. The C2R processor applies the same atmospheric correction method as the EU lake processor in BEAM,

resulting in the same remote sensing reflectance data. Also, the C2R processor resulted in the best agreement with in situ data (comparable performance with the FUB/WeW processor). Therefore, remote sensing reflectance derived from the C2R processor was used as input to cluster the matchup data by applying the two-step clustering method in the SPSS software. The relative importance of MERIS bands used as the predictors in the two-step clustering approach were reported above 0.5.

Three spectrally distinct optical water types were automatically classified from applying the clustering algorithm on all matchup in situ data. The difference between clusters' optical properties can be more readily distinguished when the reflectance mean of each cluster are plotted together (Figure 3-4). Therefore, the optical water types (type 1, 2, and 3) are represented based on the form of the remote sensing reflectance means. Water types 3 through 1 show a pattern of increasing peak magnitude at 560 nm which is associated with the enhanced backscattering from living and nonliving sources such as phytoplankton and sediments, respectively (Moore et al., 2014). Water types 1 and 2 also show peaks around 708 nm with different magnitudes. This peak is characteristic of algal blooms with high chl-a concentrations (Gower et al., 2005). Type 3 has an overall low spectral magnitude with relatively flat features after 680 nm compared to two other types. Water types 1 and 2 have a higher spectral magnitude, but different from each other. A higher number of matchup data are associated with type 3 waters. Thus, in the matchup of situ data used in this study, optical properties of type 3 waters are dominant. Each type is descriptive of the average optical properties that is governed in that class, whereas the optical properties depend on the concentration of optically active water constituents such as phytoplankton, suspended and dissolved matters.

Table 3-7 shows the descriptive statistics of water quality parameters measured across the optical water types. These parameters include concentration of chl-a, TSM, absorption of CDOMs (a_{CDOM}), and SDD, that were derived from matchup in situ data. Figure 3-5 compares the characteristics of each optical water type in terms of the measured matchup in situ data. Water type 3 is showing the highest average SDD (5.5 ± 2.37 m) compared to the two other water types (type 1: 0.9 ± 0.49 m; type 2: 1.12 ± 0.63 m). In addition, the lowest concentration of TSM, chl-a and CDOMs are measured in type 3 (Chl-a: 2.20± 1.99 mg m^{-3}, TSM: 1.84 ± 1.65 g m^{-3}, a_{CDOM}: 0.16 ± 0.1 m^{-1}). Thus, this water type represents the less turbid of the two other types, which explains the relatively flat spectra after 680 nm in Figure 3-4. Water type 1 has the highest concentration of TSM (12.48 ± 7.51 g m^{-3}), whereas the highest chl-a concentration is measured in water type 2 (5.67 ± 5.1 mg m^{-3}). Water type 1 represents the most turbid among water types with a SDD average of 0.9 ± 0.49 m.

Figure 3-6 presents the distribution of different optical types across the lake. Turbid type 1 waters, with highest concentrations of suspended matters from the three types, are mostly located in the shallow western basin of the lake and also around Turkey/Long Point and Port Colborne. The deepest parts of the eastern basin, east of Long Point is identified as having the highest sedimentation rates in the lake, probably because of shoreline erosion (Kemp et al., 1977). Type 2 waters are still turbid but with highest levels of chl-a concentrations. This water type is focused mostly close to the shorelines, including Port Stanley which has watershed dominated by agriculture (Hiriart-Baer et al., 2013) and particularly in western basin. The Detroit River and the Maumee River are sources of sediment loads and nutrients to the western basin; they influence lake turbidity and algal biomass. West Erie is also affected by inflows from Sandusky Bay carrying nutrients into the lake.

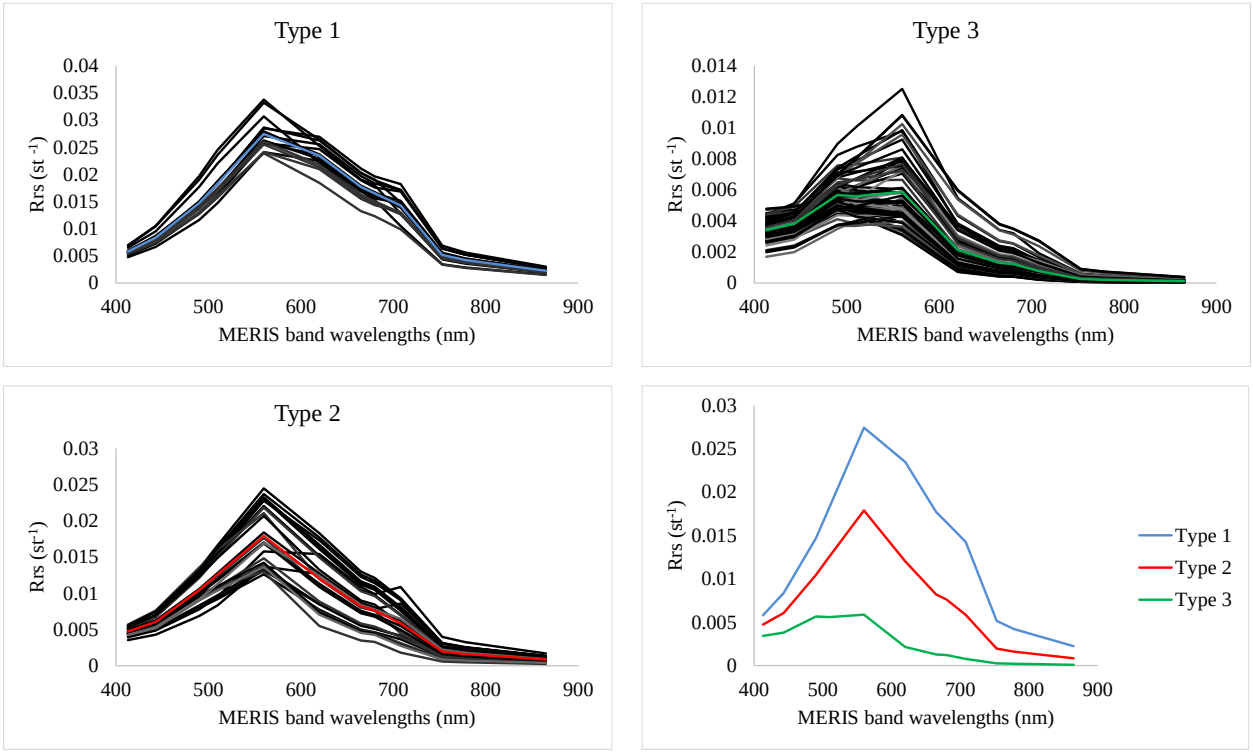


Figure 3-4 Matchup in situ data classified into three clusters from the two-step clustering approach. Black lines are the individual reflectance data; colorful lines are the reflectance means of the three optical water types as illustrated by the legend. (Rrs: Remote Sensing Reflectance).

Table 3-7 Descriptive statistics of in situ matchup data optical properties across three different water types. N is the number of matchup in situ data that is assigned to each water type.

Type	Water Quality Parameter	N	Max	Min	Avg.	St. Dev.
Type 1	Chl-a (mg m^{-3})	16	12.00	0.60	4.41	3.60
	TSM (g m^{-3})	15	30.42	4.49	12.48	7.51
	a_{CDOM} (m^{-1})	13	1.75	0.16	0.44	0.44
	SDD (m)	7	1.50	0.30	0.90	0.49
Type 2	Chl-a (mg m^{-3})	26	23.10	0.40	5.67	5.10
	TSM (g m^{-3})	26	50.50	1.17	10.73	11.79
	a_{CDOM} (m^{-1})	20	1.49	0.13	0.46	0.39
	SDD (m)	14	2.50	0.20	1.12	0.63
Type 3	Chl-a (mg m^{-3})	67	9.10	0.20	2.20	1.99
	TSM (g m^{-3})	67	10.74	0.18	1.84	1.65
	a_{CDOM} (m^{-1})	45	0.52	0.04	0.16	0.10
	SDD (m)	30	11.00	1.50	5.50	2.37

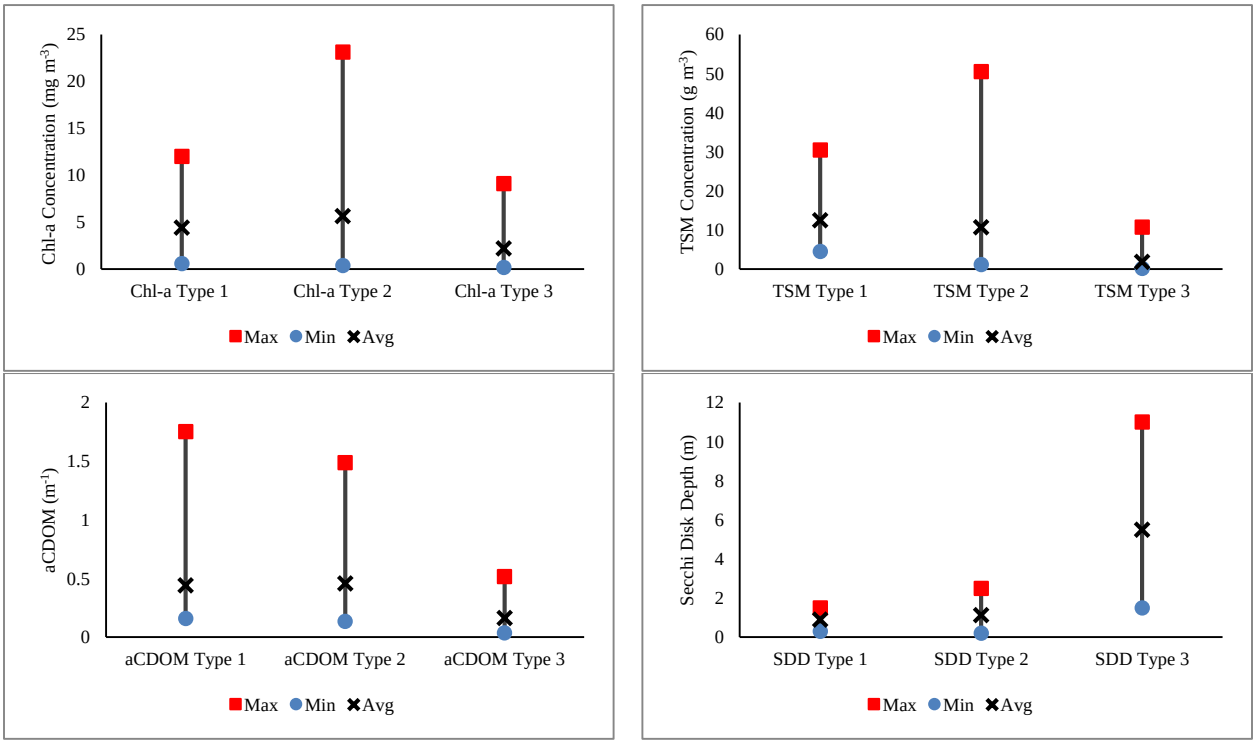


Figure 3-5 Comparing matchup in situ data characteristics across the three optical water types.

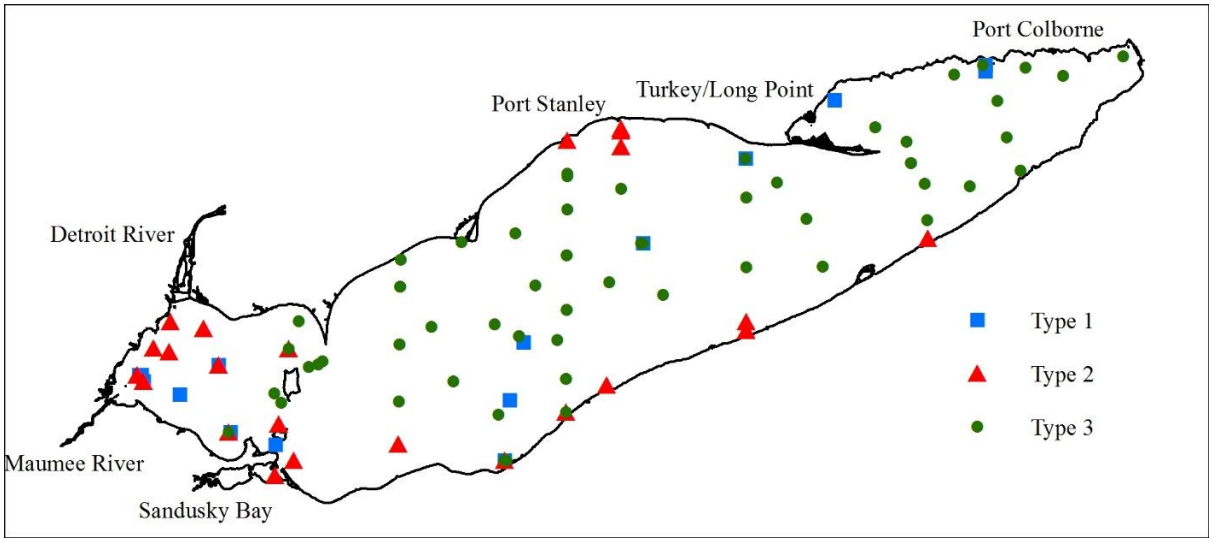


Figure 3-6 Distribution of matchup in situ data for different optical water types.

3.3.5 Chlorophyll-a Product Evaluation in Different Water Types and Blending Algorithm Results

A narrower range of optical properties and environmental conditions that each algorithm can perform well is revealed by testing the algorithms' performance for each individual water type (Moore et al., 2014). The most variation in the performance of different algorithms for each single water type was observed in I_a; where it ranged from 0.41 to 0.61 for type 1, 0.59 to 0.70 for type 2, and 0.13 to 0.72 for type 3 (Table 3-8). Therefore in terms of I_a variations, type 3 was the most sensitive to applying different algorithms, with chl-a concentration ranging from 0.2 to 9.1 mg m^{-3}, and with an average of 2.20 ± 1.99 mg m^{-3}. In this water type, FLH index had a poor performance in predicting chl-a concentration, probably because the concentration of chl-a in this type was not high enough to produce a fluorescence peak, whereas the best performing algorithm in that type was obtained from applying the C2R processor (C2R: RMSE=1.20 mg m^{-3}, MBE= -0.39 mg m^{-3}, I_a=0.72; FLH: RMSE=2.70 mg m^{-3}, MBE=1.38 mg m^{-3}, I_a=0.13). FLH had its best performance in water type 2 with highest chl-a concentrations compared to other types. The high concentration of TSM in water type 1 can eliminate the contribution of fluorescence in the water-leaving reflectance and result in a poor performance of FLH in estimating chl-a concentration for this water type.

The best performing algorithm was assigned to a particular water type based on I_a values (due to the most range of variations observed in this model performance indictor for each water type). Type 1 water was best predicted by the CC_merged product, type 2 by the CC_ NN product, and type 3 waters

by the C2R processor. The best performing algorithms were tuned to the assigned water types; and resulted in the blended product. The blended product is shown in Figure 3-7 against the in situ measurements. The overall RMSE, and I_a improved minimally for the blended chl-a product compared to applying the algorithms separately (RMSE: blended=1.91 mg m^{-3} versus FUB/WeW=1.99 mg m^{-3}; I_a: blended=0.73 versus C2R=0.71) (Table 3-9).

Table 3-8 Performance of the models calibrated for each single water type.

Statistical Indices	Water Type	EU	C2R	FUB/WeW	CC_OC4	CC_NN	CC_merged	FLH	MCI
RMSE	Type 1	3.69	3.43	3.19	3.13	2.79	2.79	4.12	4.59
	Type 2	3.16	2.84	1.99	3.59	2.65	2.67	3.61	2.94
	Type 3	1.23	1.20	1.34	1.42	1.66	1.80	2.70	2.12
MBE	Type 1	2.43	2.11	1.15	0.44	0.16	0.02	-2.65	2.00
	Type 2	-0.63	-0.31	-0.18	-0.25	-0.91	-0.81	-1.64	-1.05
	Type 3	-0.34	-0.39	-0.29	-0.11	0.21	0.22	1.38	0.45
I_a	Type 1	0.47	0.51	0.51	0.55	0.60	0.61	0.41	0.48
	Type 2	0.65	0.68	0.61	0.65	0.70	0.69	0.59	0.66
	Type 3	0.71	0.72	0.70	0.67	0.56	0.54	0.13	0.40

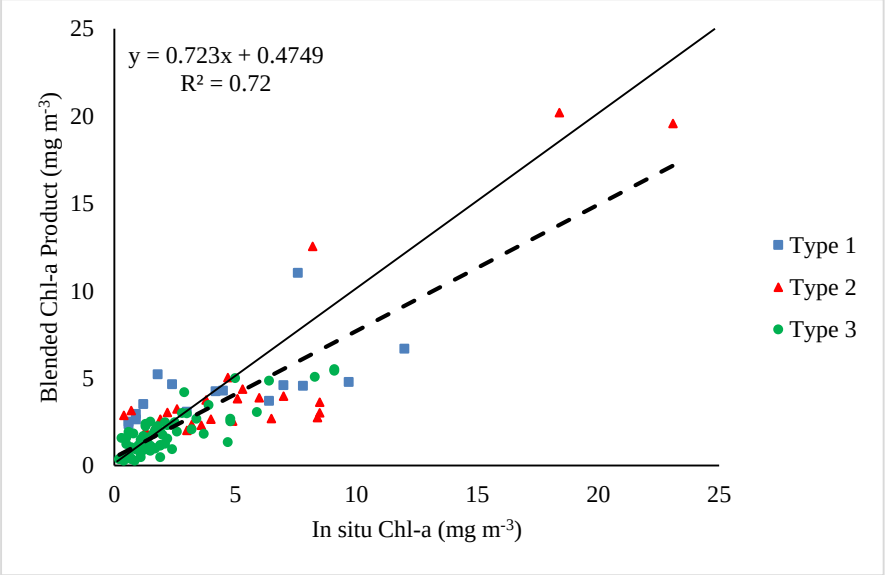


Figure 3-7 Chl-a concentration derived from blending algorithm versus in situ measurements for all matchup in situ data over Lake Erie. Solid line corresponds to a 1:1 relation.

Table 3-9 Performance of the calibrated models for the full lake, and the blending algorithm. N is the number of matchup in situ data after removing those flagged ones in each algorithm.

Full Lake	EU	C2R	FUB/WeW	CC_OC4	CC_NN	CC_merged	FLH	MCI	Blended
N	109	109	99	103	103	103	117	117	108
RMSE	2.31	2.13	1.99	2.44	2.15	2.22	3.19	2.81	1.91
MBE	0.00	-0.01	-0.04	-0.06	-0.08	-0.07	0.04	0.32	-0.45
I_a	0.69	0.71	0.67	0.68	0.67	0.66	0.47	0.59	0.73

3.4 Summary and Conclusion

Different NN and band ratio algorithms available in the BEAM software were evaluated against in situ chl-a measurements obtained during several ship cruises on Lake Erie. Matchup in situ data was divided into two independent subsets for calibration and evaluation purposes. Evaluation results using randomly selected 30% matchup in situ data shows that the C2R processor can perform marginally better than FUB/WeW processor in terms of I_a with an improvement of 6%. Whereas it underperformed FUB/WeW processor with 11% increase in RMSE values. The locally tuned equation was applied on FUB/WeW derived chl-a product to predict chl-a concentration on a selected day in September 2011. Lake Erie had experienced the largest algal bloom in its recorded history in that year before 2015. The National Oceanic and Atmospheric Administration (NOAA) reported that the 2015 bloom in western Erie was the most severe one in this century. The severity index indicates the amount of biomass in the bloom. The index value in 2015 is 10.5 as compared to 10 in 2011 bloom. The 2015 bloom covered a large area in the central basin; however, the bloom growth was disrupted and weakened by strong winds. The winds declined the bloom much faster than previous record-breaking blooms (NOAA, 2015).

The remote sensing reflectance data derived from the C2R processor was used as the input in a two-step clustering algorithm. The clustering approach was applied to differentiate different optical classes in Lake Erie; and then select the best performing algorithm for each water type. The clustering algorithm resulted in three separate water type classes in Lake Erie, each showing distinct optical properties based on different reflectance characteristics. Type 1 waters were associated with high suspended matter concentrations and the most turbid of the three types. Type 2 was still turbid but with the highest chl-a concentration values. Type 3 represented the least turbid water type with the lowest concentrations of chl-a and TSM. Also, the highest SDD was measured in type water 3. The performance of locally tuned algorithms for each water type was evaluated individually. There was no single algorithm performing best for all water types. Results showed that, although the FUB/WeW

processor is the best overall, its performance was only the best for type 3 waters in terms of RMSE and I_a. The best performing algorithm in type 1 waters based on I_a was CC_merged product. In type 2 waters, however, the CC_NN product resulted in a higher I_a. The C2R processor outperformed other algorithms for type 3 waters in terms of I_a values.

A blending ("ensemble") approach was applied on the results of best performing algorithms that were tuned for the assigned water type. The blended algorithm resulted in a slight improvement in RMSE and I_a compared to all other algorithms applied on all matchup in situ data collected from the full lake. Although there was a marginal improvement in I_a (RMSE) of about 4% (9%) versus applying the single C2R processor (FUB/WeW) on all basins.

Different water bodies have specific optical properties, therefore the calibration of semi-analytical algorithms to the characteristics of the lake is still required. However, various concentrations of optically active water constituents including algae, and suspended/dissolved matters in different parts of a lake can result in different optical properties to exist in a single water body (i.e., optically complex lakes). Thus, different algorithms have varying performance for the identified optical water classes. The clustering approach can assign each pixel to a specific optical water type, and find the best performing algorithm for that water type. Therefore, a higher estimation accuracy of satellite-derived chl-a concentration can be obtained by developing algorithms specific to the optical properties of the optically complex inland and coastal waters.

Clearly, there are still limitations associated with the satellite remote sensing of chl-a concentration including the time differences between in situ data collection and satellite overpass, and also different spatial coverage (satellite footprint versus in situ point-sampling). These factors are especially important for inland waters characterized by high temporal and spatial optical variability. The forthcoming Sentinel-2 and Sentinel-3 satellites will provide information about inland waters and coastal zones with improvement in satellite imagery applicability to monitor algal blooms. The future Sentinel-2 MSI sensor is complementary to the Landsat-8 mission with high spatial resolutions of 10, 20, and 30 m, added channels, and also a 5-day revisit frequency. Sentinel-3 OLCI sensor has MERIS heritage and improves upon it with an additional 6 spectral bands. It has a more frequent revisit cycle compared to MSI (every 1-2 days); however a coarser spatial resolution of 300 m at nadir.

Finally, the in situ data sampling has to be extensive enough to statistically represent the spatial and temporal variability in the optical properties of the inland lake. Therefore, a dataset conveying the dynamic range of optical variability observed in the water body is preferable for training and evaluating

different algorithms for the water body; however the matchup data resulted from applying flags in the present study did not cover the full variations captured in the field data.

Chapter 4
Estimation of Water Quality Parameters in Lake Erie from MERIS Using a Linear Mixed Effect Model

4.1 Introduction

Lake Erie is the most southern and shallowest of the Laurentian Great Lakes. Total suspended matters (TSM) are a major contributor to the lake's low water clarity (Binding et al., 2012). The problem of excess nutrients and resulting algal blooms are also threatening the ecosystem of the lake and the economic activities of the surrounding regions. The ecological state of Lake Erie significantly affects its role as a natural, social, and economic resource, considering that the lake is as an essential drinking water source that also offers many opportunities for recreational activities, fisheries and tourism. As a result, the Lakewide Management Plan was signed in 1972 to restore and maintain the ecological health of the lake (Daher, 1999). Ongoing efforts to support this plan require high-resolution measurements of the water quality parameters on a variety of spatial and temporal scales. Conventional field-based measurements of these parameters can be expensive and are often sparse in either space or time or both. Remote sensing has the potential to infer the lake bio-optical/water quality parameters overcoming these concerns.

The emerging water reflectance measured by remote sensing instruments depends on the water itself and its constituents. The water constituents interact with the photons of light and modify the incoming and outgoing radiation at various wavelengths. Therefore, remote sensing measurements of water leaving reflectance can be related to the composition and concentration of water constituents. In Case I waters, chl-a concentration (phytoplankton population) and its co-varying particles are dominating the optical properties, which is the case in nearly all open ocean waters. However, optically complex inland waters and coastal waters are referred to as Case II waters, where chl-a alone is a poor predictor of light attenuation, and variations in TSM and colored dissolved organic matter (CDOM) appear to also be important in light scattering or absorption, and therefore the water leaving reflectance (Morel and Prieur, 1977).

The delivery of data on water color has been explored using satellite sensors such as Landsat Thematic Mapper (TM)/Enhanced Thematic Mapper (ETM+) due to their relatively high spatial resolution of 30 m (Zhao et al., 2011; McCullough et al., 2012a; Tebbs et al., 2013). However, its shortcomings in other capacities such as its relatively low temporal resolution (i.e. 16 days) make the

use of MODIS (Moderate-resolution Imaging Spectroradiometer) and MERIS (Medium Resolution Imaging Spectrometer) data more attractive for water quality monitoring (Binding et al., 2011a; McCullough et al., 2012b; Saulquin et al., 2013). Images from these satellite sensors compensate for the limitations of Landsat at the expense of a lower spatial resolution (ca. 250-500 m). MERIS was originally designed for water quality monitoring applications. Therefore, compared to MODIS, it has a more suitable spectral resolution in the red and near-infrared (NIR) to derive the secondary chlorophyll-a (chl-a) absorption maximum (Odermatt et al., 2012a). This is essential for Case II waters, as is the case for Lake Erie, where chl-a is not the predominant color-producing agent (CPA) and multiple CPAs may also confound the reflectance signal, particularly at shorter wavelengths. Hence, the traditional and empirical blue/green band ratio algorithms result in large uncertainties due to the limited ability to distinguish signals coming from the covariant water constituents.

As a consequence of the ambiguities related to the shorter wavelengths, several authors have investigated the applicability of red and NIR wavelengths for estimating chl-a concentration in turbid optically complex waters to aim for a minimal sensitivity to other water-coloring parameters. Red-NIR band ratio algorithms have been found to work well in Lake Chagan (chl-a concentration: 6.4 to 58.21 mg m^{-3}) (Duan et al., 2010), as well as Curonian Lagoon (chl-a concentration: 44.1 to 85.3 mg m^{-3}) (Bresciani and Giardino, 2012) and Zeekoevlei Lake (chl-a concentration: 61 to 247.4 mg m^{-3}) (Matthews et al., 2010) that have chl-a concentration ranges typical of mesotrophic lakes and eutrophic waters in Lake Erie (Binding et al., 2012). Band ratio algorithms developed to derive Secchi disk depth (SDD) variations are making use of bands in the visible range of the spectrum. Two multiple linear regression models have been developed separately, based on blue and red bands of Landsat TM and MODIS, to predict the logarithm of SDD in Poyang Lake National Nature Reserve in China (Wu et al., 2008). A linear regression model has also been proposed based on the logarithmic transformation of MERIS band ratio (490 nm to 620 nm) to estimate the natural logarithm of SDD in the Baltic Sea (Kratzer et al., 2008). However, the correlated errors resulting from repeated measurements in space and time are not considered in the regression models developed in these studies. Multiple measurements per variable will result in non-independency, which violates the assumptions of regression methods. The linear mixed effect (LME) model (Pinheiro et al., 2015) approach developed herein is appropriate for cases where observations are collected in time and/or space for the same parameter, and therefore represent clustered or dependent data.

The applicability of LME models is tested in this study to estimate chl-a concentration and SDD from the CoastColour (CC) atmospherically corrected MERIS reflectance product (Ruescas et al., 2014) in support of water quality monitoring in Lake Erie. Although in situ measurements remain the most accurate solution for water quality monitoring programs, satellite remote sensing can be added for routine and synoptic measurements (Moore et al., 2014). Chl-a is widely measured as an indicator of eutrophication and primary production. SDD is another environmental descriptor that is indicative of water clarity. It also provides a highly relevant measure of the extent of the euphotic layer where primary production is possible (Kratzer et al., 2003; Fleming-Lehtinen and Laamanen, 2012). Therefore, both parameters are of interest in this study. Section 4.2 of this paper provides an overview of the location and general chemical, physical and biological properties of Lake Erie. It also outlines the field methods and measurement campaigns conducted to acquire in situ water quality observations. The processing scheme and LME models applied to the MERIS data are described at the end of this section. Section 4.3 presents results from the evaluation of satellite-derived estimates of chl-a and SDD against in situ measurements and discusses the temporal and spatial variations in the water quality parameters. Finally, Section 4.4 provides concluding remarks about the study.

4.2 Material and Methods

4.2.1 Study Site

Lake Erie (42° 11′N, 81° 15′W; Figure 4-1) is the smallest (by volume), the shallowest, and the warmest of the Great Lakes (Michalak et al., 2013). It is a monomictic lake (with occasional dimictic years). The shallow western basin is polymictic, since stratification is destroyed by wind-driven mixing. However, the central and eastern basins develop stable thermoclines in summer (Gobler and Wilhelm, 2015). The lake is covering an area of 25,700 km^2, with average and maximum depths of 19 m and 64 m, respectively (Bootsma and Hecky, 2003). The lake is naturally divided into three basins of different depths: the shallow western basin, the central basin, and the deep eastern basin (Table 4-1). The basins are separated approximately based on the Lake Erie Islands (~ 82^{O} 49′ W) and the Long Point-Erie Ridge (~ 80^{O} 25′W) (Figure 4-1) (Binding et al., 2012). River discharge into Lake Erie originates mostly from the St. Clair River and Lake St. Clair through the Detroit River. Other smaller rivers and streams in the territory of Lake Erie also contribute to water inflow into the lake. Lake Erie drains into Lake Ontario through the Niagara River and shipping canals (Daher, 1999; Painter et al., 2000).

Table 4-1 Lake Erie Basins Information (source: (Painter et al., 2000))

Lake Erie	Mean Depth (m)	Maximum Depth (m)
West Basin	7.4	19
Central Basin	18.3	25
East Basin	24	64

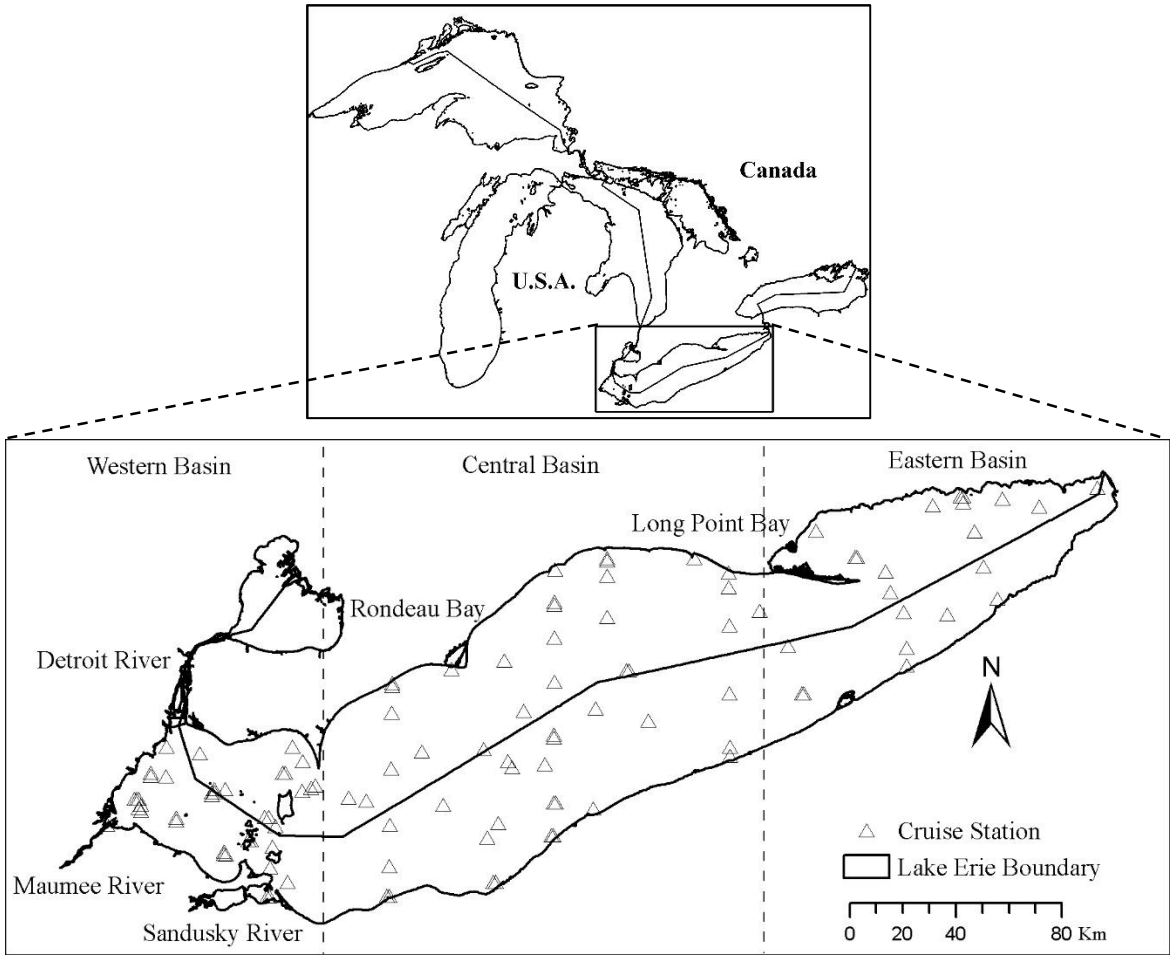


Figure 4-1 Location of Lake Erie and its boundary (Canada and US). In situ sampling stations from cruises that took place in September 2004, May, July, and September 2005, May and June 2008, July and September 2011, and February 2012 are illustrated by empty triangles.

Lake Erie is exposed to greater stress than any other of the Great Lakes due to agricultural practices and urbanization in its surroundings. Chemically enriched runoff from agricultural lands in the basin flows into the lake. In addition, the lake receives the most effluents from wastewater treatment works (Daher, 1999; Painter et al., 2000). Lake Erie has experienced substantial eutrophication over the past half century due to excess phosphorus loads from point and nonpoint sources producing algal blooms

(Michalak et al., 2013). In general, phosphorus concentration in Lake Erie decreases from west to east and from near-shore to the offshore (IJC. International Joint Commission Canada and United States, 2013).

4.2.2 Field Measurements of Water Quality Parameters

Sample collection in Lake Erie was conducted on board of the Canadian Coast Guard ship *Limnos* during September 2004, May, July, and September 2005, May and June 2008, July and September 2011, and February 2012. A total of 89 distributed stations were visited to provide measurements of a wide range of optical properties as well as concentrations of the main CPAs in different locations of the lake (Figure 4-1).

Composite water samples were collected at all stations, during 2004 to 2012, from the surface mixed layer of the lake using Niskin bottles. The samples were filtered through a Whatman GF/F fiber filter (0.7 μm) in the field. The filtered samples were then frozen and sent to the laboratory for extraction of chl-a concentrations following the method of the National Laboratory for Environmental Testing (NLET). This method is based on spectrophotometry following fixation using a 90% acetone solution and centrifugation (Environment Canada, 1997). SDD measurement is a worldwide accepted method to estimate turbidity in water bodies. Chl-a and SDD measurement methods followed the Ocean Optics Protocols for Satellite Ocean Color Sensor Validation (Mueller et al., 2003; Pegau et al., 2002).

4.2.3 Satellite Data and Processing

Launched by the European Space Agency (ESA) on 1 March 2002, the MERIS sensor was one of the instruments operating on the Envisat polar-orbiting satellite platform. Contact was lost with Envisat on 8 April 2012, which marked the end of the mission. MERIS was primarily dedicated to ocean color studies. MERIS was a push-broom imaging spectrometer that could measure the solar radiation reflected from the Earth's surface in a high spectral and radiometric resolution (15 spectral bands across the range 390 nm to 1040 nm) with a dual spatial resolution (300 and 1200 m). MERIS scanned the Earth with global coverage every 2-3 days.

In this study, CC L2R (Version 2) MERIS reflectance full resolution images with full or partial coverage of Lake Erie between September 2004 and February 2012 were acquired through the Calvalus on-demand processing portal. The CC MERIS Level 2R product is generated using an atmospheric correction algorithm applied to the Level 1P product, which is a refined top of atmosphere radiance product with improved geolocation, calibration, equalization, smile correction, in addition to precise

coastline and additional pixel characterization information (e.g. cloud, snow). The atmospheric correction procedure is based on two processors implemented in the Basic ERS & ENVISAT (A)ATSR MERIS (BEAM) software (Version 5.0): the Case II Regional (C2R) lake processor and also glint correction processor (Ruescas et al., 2014). A detailed description of C2R can be found in Doerffer and Schiller (2007).

The MERIS images were selected to be within a 2-day time window of in situ water quality measurements for the study period (2004-2012). This criterion was set to maximize the number of possible satellite and in situ measurements match-ups; at the same time reducing the effect of time heterogeneity over the lake, assuming that the water quality parameters would not change significantly in this time frame. Atmospherically corrected MERIS L2R reflectance values were extracted from pixels covering the geographic location of the stations. A valid pixel expression was defined that excluded all pixels with properties listed in Table 4-2. Spatial averaging of pixels surrounding the station could be a technical solution to increase the number of resulting match-ups, when the considered pixel is excluded due to flags (Heim et al., 2014). However, the horizontal spatial heterogeneity of parameters over the lakes prevents the averaging analysis.

Table 4-2 Flags of excluded pixels

Level 1	Level 1P	Level 2
Glint_risk	Land	AOT560_OOR (Aerosol optical thickness at 550 nm out of the training range)
Suspect	Cloud	TOA_OOR (Top of atmosphere reflectance in band 13 out of the training range)
Land_ocean	Cloud_ambigious	TOSA_OOR (Top of standard atmosphere reflectance in band 13 out of the training range)
Bright	Cloud_buffer	Solzen (Large solar zenith angle)
Coastline	Cloud_shadow	
Invalid	Snow_ice	
	MixedPixel	

4.2.4 Water Quality Parameters Algorithms

Semi-empirical algorithms are based on the regression between individual bands or band ratios, and the dependent variables, which are chl-a and SDD in this study. Different combinations can be considered based on the 15 MERIS spectral bands. Lakes with various optical and biological properties can produce different levels of correlation with these band combinations. This makes the use of semi-empirical algorithms a robust approach that can work on the lake of interest and within the time period that data samples were collected. The best band combinations were determined from the highest

calculated Pearson correlation coefficients (R) against the logarithmic scale of in situ measurements of water quality parameters. It has long been known that the chl-a concentration distribution in the ocean is lognormal (Campbell, 1995). Also, the logarithmic function linearizes the relationship of in situ observations to band ratios and makes the distribution more symmetric (normal). Considering the optical complexity of Lake Erie (see Section 4.3.1), red and NIR bands are required to derive chl-a concentration. Therefore, the selected band ratio for estimating chl-a was chosen among MERIS bands centered at B05: 560 nm, B06: 620 nm, B07: 665 nm, B008: 681.25 nm, B09: 708.75 nm, B10: 753.75 nm, B12: 778.75 nm, and B13: 865 nm. The band ratio for deriving SDD was selected among the visible bands: B01: 412.5 nm, B02: 442.5 nm, B03: 490 nm, B04: 510 nm, B05: 560 nm, B06: 620 nm, B07: 665 nm, B08: 681.25 nm.

Sampling-wise, the same in situ measurements are repeated in time (month) over Lake Erie during the study period. Multiple measurements per variable in space or time will generally result in correlated errors and clustered data, which violate the assumptions of regression methods. Accordingly, random effect of time has to be added to the error term of the general regression models to account for measurements being made in clusters of time. Also, the measurements for different stations are inter-dependent. Different locations can affect each other's measurements, depending on their distance. Therefore, there is spatial dependency in in situ observations. To consider both random and fixed effects in the regression, a LME model approach was selected to handle the repeated measurements and also the spatial autocorrelation of in situ observations.

Two separate LME models were developed between the logarithmic scale of in situ chl-a and SDD with selected individual bands or band ratios of MERIS atmospherically corrected reflectance. The models were then used to predict chl-a and SDD over Lake Erie at different times. A LME model allows the prediction to be made at the outermost level (level=0: predictions are only based on fixed effects, as it would be in a standard regression model), and innermost level (Level≠0: predictions are based on estimated random and fixed effects).

4.2.5 Accuracy Assessment

Cross validation was performed to assess the accuracy of derived chl-a and SDD estimates. Repeated random sub-sampling was used to reduce variability due to random sampling effect. Ten rounds were repeated by splitting the in situ measurements into training (70%) and testing (30%) datasets. The model uncertainty indicators were reported as the average over the iterations. The mean bias error (MBE: the average difference), and the root mean square error (RMSE) were used as the model uncertainty

indicators to describe chl-a and SDD retrieval accuracies. The index-of-agreement (I_a) was also calculated as it is a statistics indicative of model performance. The refined version of I_a, the one adopted in this study, has recently been proposed by Willmott et al. (2012). The index is bounded between -1 (the worst model performance) and +1 (the best model performance) (Willmott, 1981; Willmott et al., 2012). The RMSE, and MBE statistics are defined as follows:

$$RMSE = \sqrt{mean(log10x_{pr} - log10x_{obs})^2} \qquad \text{Equations 4.1}$$

$$MBE = mean(log10x_{pr} - log10x_{obs})$$

where x_{pr} is the predicted value of the chl-a concentration or SDD from the algorithm, x_{obs} is the observed value of the quantity which is measured in the field. The RMSE is a comprehensive metric as it combines the mean and variance of the error distribution into a single term (Szeto et al., 2011). MBE also reveals the systematic errors (Moore et al., 2014).

Statistical analyses including: (1) finding the best MERIS atmospherically corrected reflectance band ratios, based on correlation with in situ observations; (2) development of LME models (regression method) based on selected band ratios and in situ observations; and (3) cross validation were performed in R programming language (Version 3.2.1) (R Core Team, 2015). All significant levels are reported at $p < 0.005$.

4.3 Results and Discussion

4.3.1 Lake Erie as an Optically Complex Water Body

Descriptive statistics of various bio-optical parameters measured in Lake Erie over the 2004-2012 period are summarized in Table 4-3.

Table 4-3. Descriptive statistics of in situ measurements for Lake Erie (2004-2012). N is the number of times samples were collected at stations. St. dev. is standard deviation. Chl-a and TSM are in mg m^{-3}, and g m^{-3}, respectively; a_{CDOM} in m^{-1}, and SDD in m.

	N	Min	Max	Mean	St. dev.
Chl-a	190	0.20	70.10	4.27	7.82
TSM	190	0.18	50.50	5.75	8.00
a_{CDOM}	160	0.04	2.36	0.31	0.33
SDD	117	0.20	11.00	3.69	2.68

Figure 4-2 shows the contribution of different water constituents in the attenuation of light in Lake Erie by investigating correlations between the concentrations of chl-a and TSM, and also absorption of CDOM in 440 nm ($a_{CDOM}(440)$) with the measured SDD. The graphs reveal that the concentrations of chl-a, TSM, and $a_{CDOM}(440)$ are correlated with SDD over the period of measurements. TSM and to a lesser extent CDOM are important contributors to water turbidity observed in Lake Erie with coefficient of determination (R^2) values of 0.67 and 0.54, respectively.

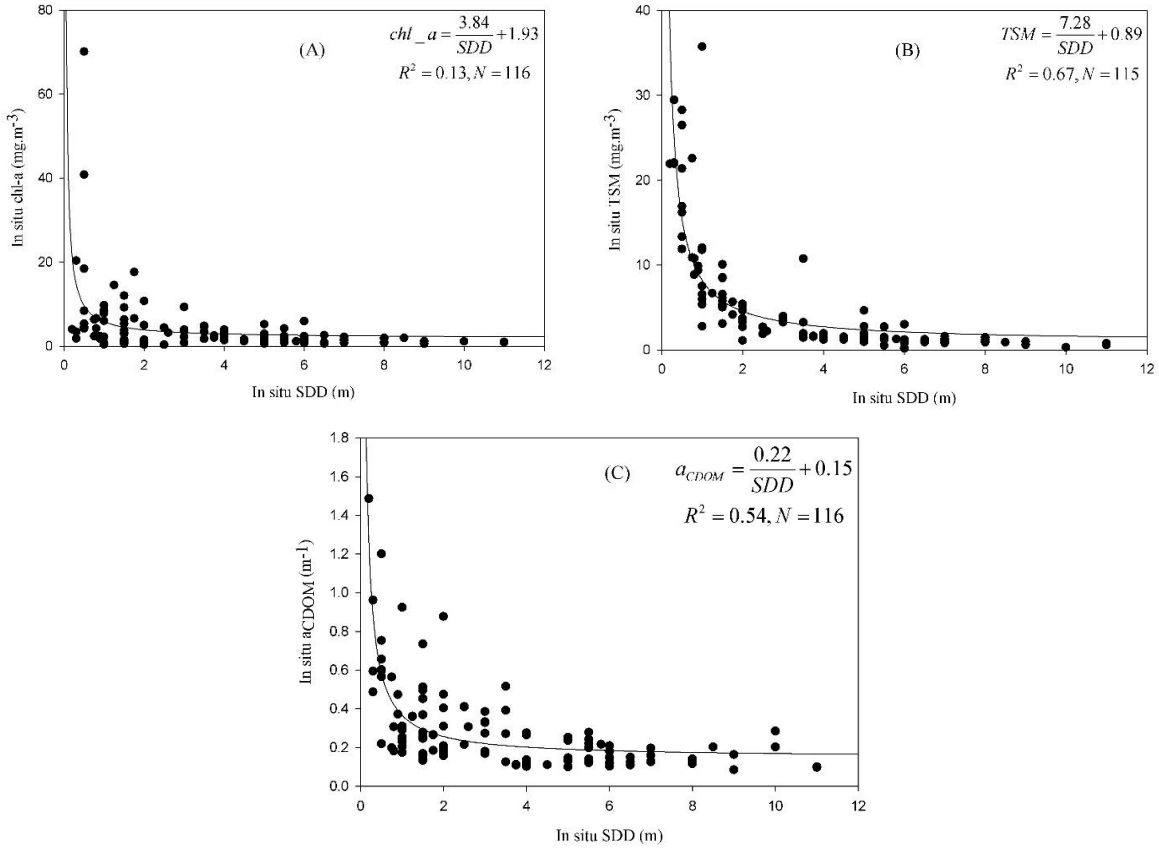


Figure 4-2 Relationships between in situ SDD and three bio-optical parameters of the water: chl-a (A), TSM (B), and $a_{CDOM}(440)$(C).

The relative contribution of TSM compared to CDOM can be reduced in microtidal estuaries, or depending on the bathymetry of the lake (Branco and Kremer 2005). In shallow Lake Erie, re-suspension of bottom sediments leads to higher water turbidity. Also, the Detroit and Maumee rivers contribute large sediment loads into the western basin of the lake. Kempt et al. (1997) identified the regions off Long Point and the mouths of these two rivers as the points of highest sedimentation rates

in the lake. The study by Binding et al. (2012) identified these zones as the highest turbid areas in the lake, confirming that TSM plays a major role in optically complex Lake Erie (Binding et al., 2012).

Based on the results presented in Table 4-3 and in Figure 4-2, Lake Erie can be classified as a typical Case II water system where other water constituents play a major and independent role in its low water clarity, besides chl-a concentration. Therefore, red and NIR reflectance are consistently the most reliable and expedient remote sensing variables in predictive algorithms for chl-a concentrations assessments in Lake Erie (Duan et al., 2010).

4.3.2 Linear Mixed Effect Model Calibration

The natural variation of the water quality parameter being measured determines the required period of concurrency between satellite overpass and in situ observation (Shi et al. 2014). In this study, satellite images were selected in a 2-day time window of in situ data collection. Using this criterion resulted in 16 MERIS CC L2R images being available for analysis. Applying defined flags produced 117 (60) pairs of atmospherically corrected reflectance and in situ chl-a (SDD) observations.

Pearson correlation (R) coefficients were calculated from MERIS bands or band ratios against the logarithmic scale of chl-a concentration and SDD measurements to select the best band or band ratios for the regression analysis of chl-a and SDD, separately. Figure 4-3 shows the range of correlation coefficients between the parameters of interest (chl-a and SDD) in logarithmic scale and both individual bands and band ratios of water-leaving reflectance. The ratio of B07:665 nm to B09:708.75 nm has the highest correlation with chl-a concentration (R=-0.68) and the ratio of B13:865 nm to B10:753.75 nm the weakest correlation (R=-0.14). The highest and lowest correlation coefficients between SDD measurements and individual spectral bands or the ratio of them are observed for the band ratio of B06:620 nm to B04:510 nm (R=-0.90) and ratio of B04:510 nm to B02:442 nm (R=-0.16), respectively. From this analysis, band ratio B07:665 nm to B09:708.75 nm and band ratio of B06:620 nm to B04:510 nm were selected to investigate their predictive capability in estimating chl-a concentration and SDD, respectively, using LME regression models.

Gitelson et al. (2007) applied a two-band model, as the special case of a conceptual three-band model (Dall'Olmo and Gitelson, 2006), to the turbid (Case II) waters of Chesapeake Bay to estimate chl-a concentration. The tuning process found the ratio of 720 nm to 670 nm as the optimal spectral band ratio, with the maximal R^2 of 0.79 in a positive correlation. Water samples collected from Chesapeake Bay contained widely variable chl-a concentration (9 to 77.4 mg m^{-3}), when SDD ranged from 0.28 to

1.5 m. Duan et al. (2010) also found the band ratio of 710/670 nm positively correlated to the chl-a concentration in eutrophic Lake Chagan with R^2=0.70. Chl-a concentration in this lake was between 6.40 and 58.21 mg m^{-3} and SDD rarely exceeded 0.50 m (Duan et al., 2010). Hicks et al. (2013) reported that the logarithmic scale of SDD measurements and logarithmic scale of Landsat 7 ETM+ band ratio of B01(0.450–0.515 nm)/B03(0.630–0.690 nm) were positively correlated with a high Pearson correlation (R=0.82). This study was conducted for shallow lakes (ranging from 1.8 to 8.7 m depth) in the Waikato region in New Zealand, with SDD in situ observations varying between 0.005 and 3.78 m (Hicks et al., 2013).

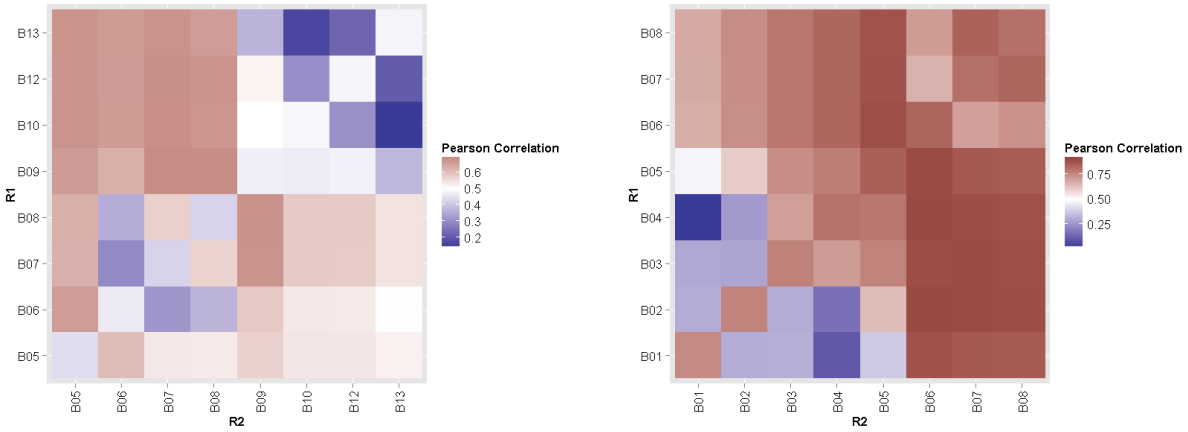


Figure 4-3 Correlation coefficients between MERIS red/NIR water-leaving reflectance ratio and in situ chl-a (left) and between MERIS water-leaving reflectance ratio and in situ SDD (right). R1 and R2 represent nominator and denominator, respectively. Values along the diagonal line from lower left to top right indicate correlation with reflectance of a single wavelength.

The LME models were developed from the relationships between selected band ratios and the logarithmic scale of chl-a and SDD. The month of in situ data collection represented the random effect due to repeated measurements in time. Random slopes and intercepts were considered for each group of measurements in a single month. Goodness of fit (r-squared; R^2) for the outermost and innermost levels of prediction were 0.49 and 0.56 for chl-a (versus R^2 value of 0.46 for a regular regression) and 0.78 and 0.83 for SDD (versus R^2 value of 0.81 for a regular regression), respectively. Therefore, the innermost level of predictions were used to predict chl-a concentration and SDD, with different values of slope and intercept for each month. The model developed for chl-a has average values of -1.16 (standard deviation: 3e-5, standard error: 0.1) and 2.44 (standard deviation: 0.12, standard error: 0.2) for slope and intercept, and the model developed to predict SDD has values of -1.04 (standard deviation: 0.09, standard error: 0.08) and 0.99 (standard deviation: 0.05, standard error: 0.08) for slope and

intercept. The estimated slope and intercept values for both models are significant ($p < 0.005$). The derived models are summarized using only the fixed effects as follows:

$$\log_{10}(chl\_a) = \frac{B07(665\ nm)}{B09\ (708.75\ nm)} \times (-1.16) + 2.44 \qquad \text{Equations 4.2}$$

$$\log_{10}(SDD) = \frac{B06\ (620\ nm)}{B04\ (510\ nm)} \times (-1.04) + 0.99$$

4.3.3 Evaluation of Linear Mixed Effect Models

The model uncertainty indicators were derived for two models ($\log_{10}$chl-a: N = 117, RMSE = 0.31, MBE = 0.018, I_a = 0.66; $\log_{10}$SDD: N = 60, RMSE = 0.19, MBE = 0.006, I_a = 0.79). Comparisons between the measured and predicted $\log_{10}$ (chl-a) and $\log_{10}$ (SDD) using the LME models show that the values are in close agreement with paired observations, mostly evenly distributed along the 1:1 line (Figure 4-4). The chl-a model is, however, not sensitive enough to detect changes in concentrations below 1 mg m^{-3} and, as a result, the predicted values are not showing the variations corresponding to the small amount of in situ chl-a concentration measurements. Also, there is a larger scatter for the predicted values corresponding to the smaller amount of SDD in situ observations.

Moore et al. (2014) applied a blending approach, to manage the selection between two band ratio algorithms in blue/green and red/NIR regions, based on the optical water type classification on Lake Erie. RMSE and MBE values were 0.32, and 0.023 in logarithmic chl-a units (Moore et al., 2014). Sá et al. (2015) evaluated CC chl-a products including: OC4, NN, and merged products, for the Western Iberian coast. The uncertainty estimation analysis was presented on the logarithmic scale of chl-a ($0.249<RMSE<0.278$, $0.139<MBE<0.200$; for 3-hour time intervals) (Sá et al., 2015). Wu et al. (2008) estimated SDD in Poyang Lake in China from two multiple regression models. The models were developed using spectral bands of Landsat TM and MODIS, separately. In both models the blue and red bands were used in the regression. The logarithmic scale of SDD was predicted with RMSE values of 0.20 and 0.37 for the models, respectively (Wu et al., 2008). Results of the present study indicate that the LME models can be used to derive the bio-optical quantities; the models provide accuracies comparable to that of other studies. A good agreement between the selected band ratios (B07/B09 for chl-a and B06/B04 for SDD) of atmospherically corrected CC L2R MERIS data and in situ

measurements of chl-a and SDD in logarithmic scale were derived for Lake Erie for the 2004-2012 study period.

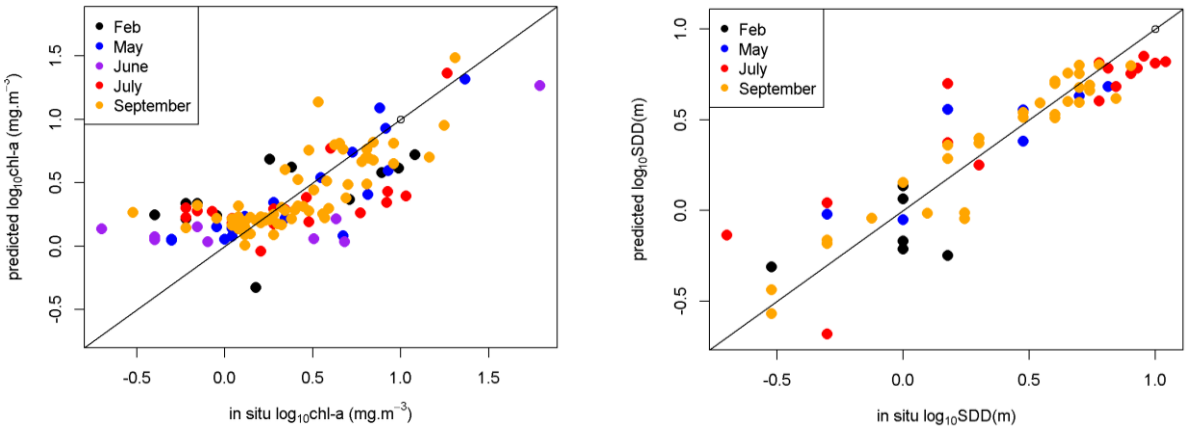


Figure 4-4 Comparison between MERIS estimates of chl-a (left) and SDD (right) using LME models and in situ measurements for Lake Erie. The solid diagonal line is the 1:1 line.

4.3.4 Spatial and Temporal Variability of Chl-a and SDD

The average chl-a concentration and SDD for each month between 2005 and 2011 are shown in Figures 5 (left) and 6 (left). In situ data in 2004 were only collected in September, and the values were not estimated for the months before September. Also, the Envisat satellite stopped operating in April 2012, hence there were no full year time series estimated for 2004 and 2012. These years were therefore disregarded in the time series analysis below. The statistics related to the number of available pixels for each month is included in Figure 4-5 (left), which are the same for SDD measurements.

The three basins of Lake Erie are characterized by distinct physical, chemical, biological, and optical properties. The highest chl-a concentrations and turbidity are experienced in different times of the year for each basin. The western basin always experiences a more intense algal bloom compared to the two other basins, with the most and least concentrations in September (6.62 ± 4.67 mg m^{-3})-October (4.83 ± 3.67 mg m^{-3}) and June (2.39 ± 3.68 mg m^{-3}), respectively. Lake Erie's central basin experiences spring bloom in April (2.08 ± 0.67 mg m^{-3}) and a more intense bloom in fall (October: 2.80 ± 0.79 mg m^{-3}). The eastern basin shows the least chl-a concentrations of the three basins, and its highest algal intensity occurs in summer (August, 1.78 ± 1.27 mg m^{-3}). Some more specific areas of the lake are affected by prolonged intense algal bloom, including Maumee Bay, Sandusky Bay, Rondeau Bay and Long Point Bay. These specific areas are known to experience cyanobacteria blooms due to constant nutrient

enrichment (Binding et al., 2012). Maumee River drains a large watershed which is dominated by agricultural fields, and also is a tributary of the largest storm runoff within the Lake Erie basin (Bolsenga and Herdendorf, 1993; Morang et al., 2011).

There is also a north-south gradient noticeable in both chl-a concentration and SDD in western Lake Erie. This gradient can be explained by inflows from Detroit River. The Detroit River is a major source of flows from the upper Great Lakes into Lake Erie, which carries contaminated sediments and nutrients from a highly urbanized and industrialized watershed into western Lake Erie (Marvin et al., 2002; Binding et al., 2012). However, the comparatively clearer water that is carried through this river from the upper Great Lakes can create the north-south gradient in the Lake Erie (Binding et al., 2012). Also, Dolan (1993) reported that municipal phosphorus loads from US sources have a higher magnitude compared to the Canadian ones during the period 1986-1990. Therefore, if the same trend of phosphorus loads in those years occurs during the time period of this study, the observed differences between north and south near-shore algal productivity can be enlightened (Binding et al., 2012).

The highest SDD values for the full Lake are estimated in July (5.38 ± 1.16 m), whereas the lowest SDD estimates are observed in March (2.44 ± 1.20 m) and October (2.52 ± 1.13 m). Re-suspension, shoreline erosion and loading from different sources such as rivers are among the most important factors influencing SDD estimates. Wind, as the primary source of kinetic energy, affects the sediment redistribution in the water column in Lake Erie (Binding et al., 2010). The high-energy and short-lived winter storms are a characteristic of Lake Erie wave climate that interrupts a long period of relative calm weather (Morang et al., 2011). These strong storms usually occur before the lake freezes (in October, November, and December) and also in spring after ice break-up (March and April) (Binding et al., 2012). However, it should be noted that the depth of the lake directly affects the amount of kinetic energy generated by wind. In other words, the re-suspension of sediment loads generated by wind in the shallower areas can be more pronounced than in the deeper areas of Erie. Comparing SDD estimates and depths of the lake in Figure 4-7, it can clearly be seen that the deeper areas are relatively clearer, while the shallow areas are more turbid. The maximum depth of the western basin is only 11 m (Ortiz et al., 2013). Hence, being the shallowest area, the western basin is the most vulnerable to physical processes such as re-suspension. Therefore, re-suspension of TSM can result in a prolonged constant turbidity in West Erie basin, besides inflows from rivers that can carry sediments into the lake and reduce in SDD reduction. Fine-grained sediments come from the Detroit River (1.6 million tons/yr) and the Maumee River (1.2-1.3 tons/yr) (Carter, 1977).

Standard deviations of chl-a (Figure 4-5-right) and SDD (Figure 4-6-right) were also calculated for each month (March to October) to show variations of log_{10}chl-a and log_{10}SDD from the average values. Figures 5 and 6 show that the greatest variability occurs in the western basin for both log_{10}Chl-a and log_{10}SDD. This area of the lake is the estuary of the Detroit River and close to Maumee Bay and Sandusky Bay. The largest variability in west Erie is estimated in March (Figure 4-5-right and Figure 4-6-right). Precipitation and runoff during this time of year, after the ice break-up period on the lake, cause more variations in nutrient availability and water column re-suspension effect on algal biomass and lake turbidity. The offshore areas and eastern basin have the least variations in chl-a concentration and SDD patterns. These lake sections appear to experience low fluctuations in the availability of required resources for algal bloom such as nutrients. Also, eastern basin of Lake Erie is the deepest with an average depth of 24 m (max depth = 64 m). Physical processes such as re-suspension have the least effect on the turbidity and its variations in the deep parts of the lake, as opposed to the shallow western basin.

The meteorological forcings can also have impact on the magnitude and timing of blooms. In general, a temperature increase leads to higher rates of photosynthesis and therefore to a greater phytoplankton growth rate under adequate resource supplies such as nutrients and light. Light-limited photosynthesis rate is insensitive to temperature, whereas a light-saturated one increases with temperature (Winder and Sommer, 2012). The resource availability of light and nutrients can be accompanied by vertical mixing. Therefore, the seasonal cycles of stratification and wind-induced vertical mixing are the key variables that condition the growth rate of phytoplankton in the water column (Winder and Sommer, 2012). Stratification results in a nutrient-depleted condition at the water surface, when the upward flux of nutrients from the deep water layers is suppressed. Also, the overall impact of windiness decreases light availability in the lower depth due to re-suspension of sediments (Winder and Sommer, 2012). As a result, the balance found between meteorological forcings, which sometimes can have opposite effects, is one of the driving factors determining the bloom condition. Phytoplankton production is a complex function and can be controlled by resources dynamics, species composition, and predator-prey interactions in the ecosystem (Winder and Sommer, 2012).

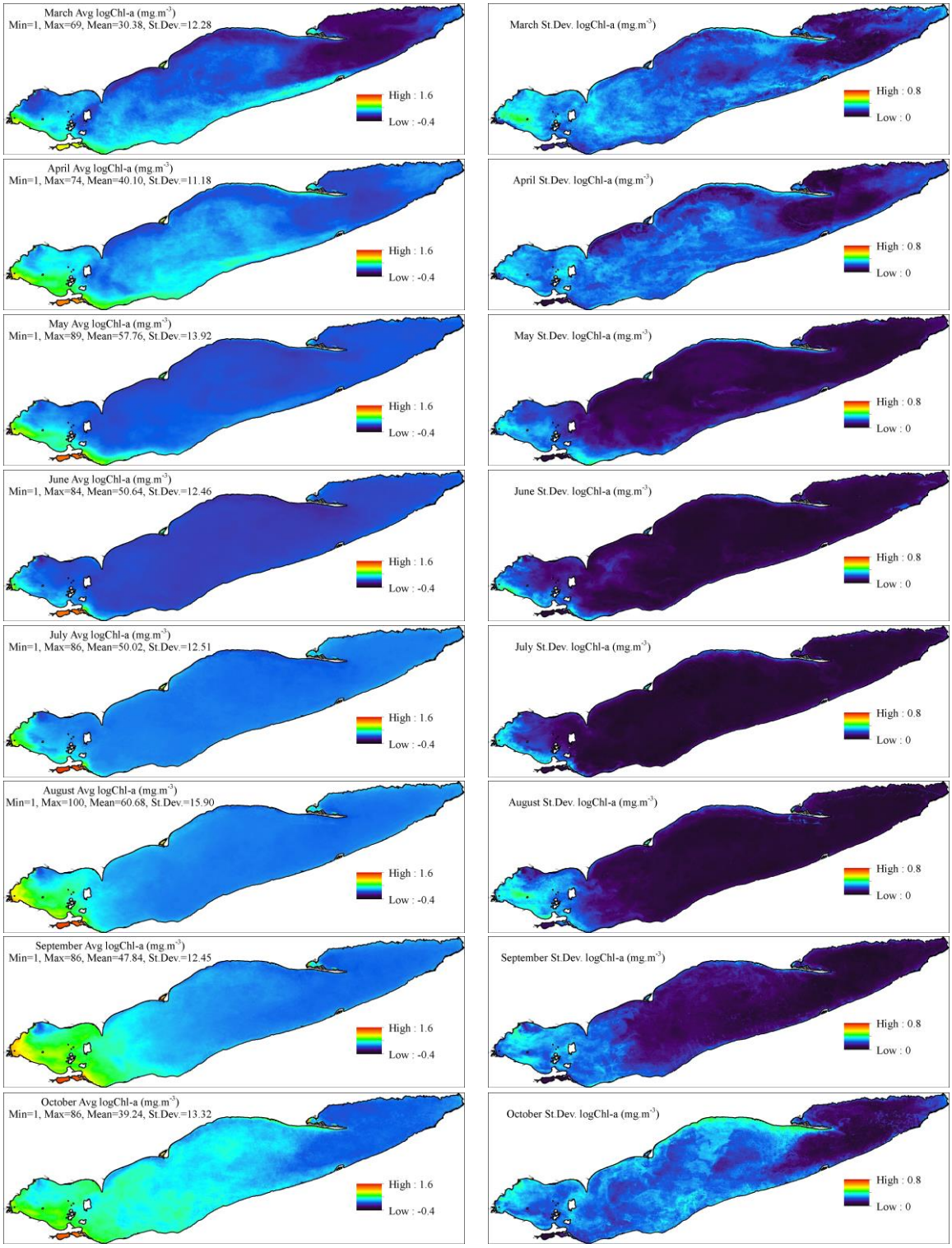


Figure 4-5 Chl-a average (Avg, left) and standard deviation (St.Dev., right) (log10 scale) from March to October for the study period (2005-2011).

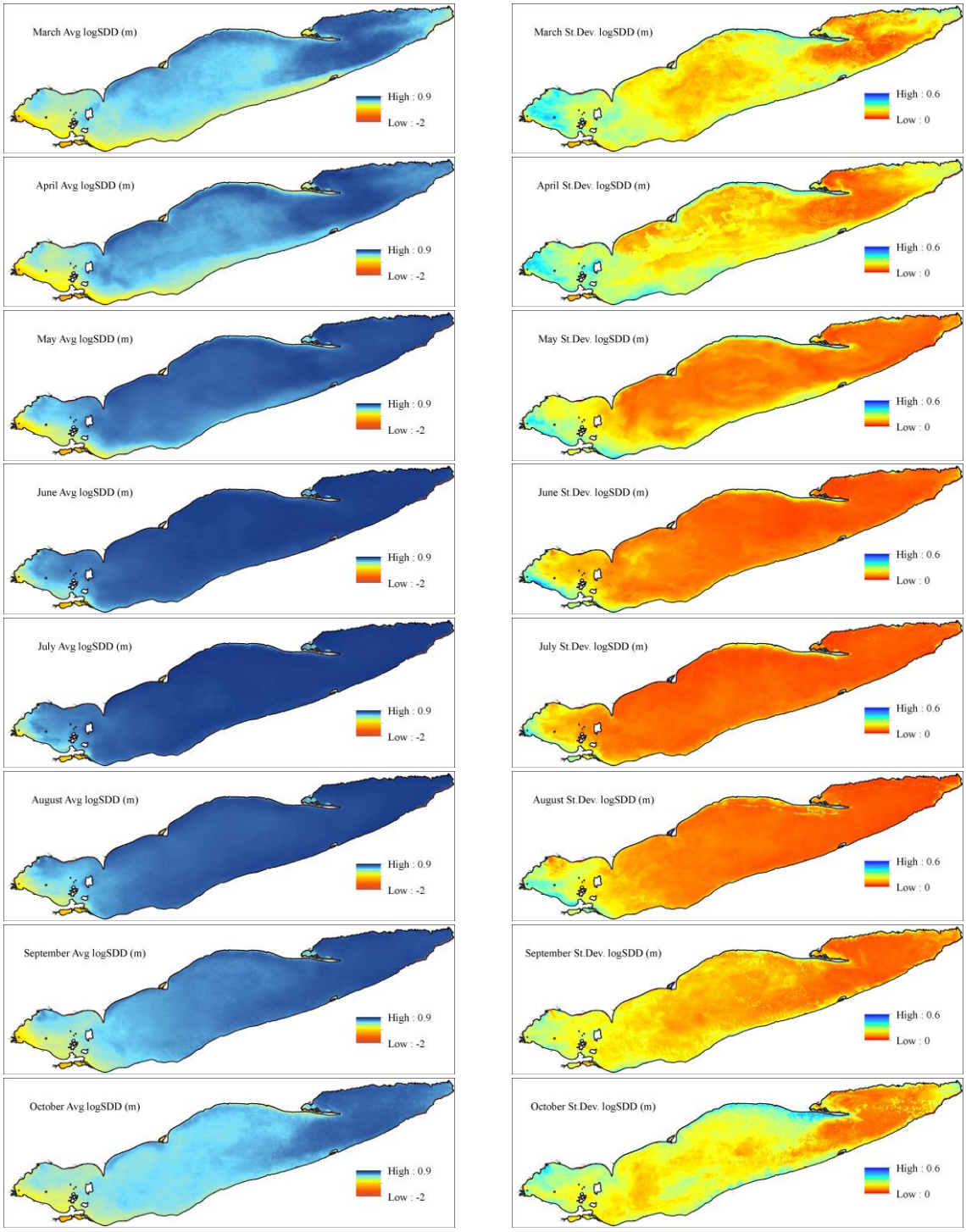


Figure 4-6 SDD average (Avg, left) and standard deviation (St.Dev., right) (log10 scale) from March to October for the study period (2005-2011).

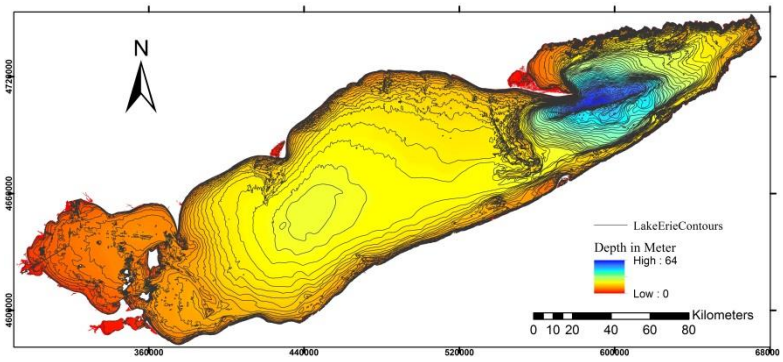


Figure 4-7 Bathymetry of Lake Erie (source: NOAA).

4.3.5 Uncertainties of the Applied Linear Mixed Effect Model on MERIS

The influence of other existing particulates in a Case II water, such as CDOM and TSM, will be significantly decreased employing the selected wavelengths to develop the chl-a LME model, as opposed to empirical blue-to-green band ratio algorithms. The absorption of CDOM is greatest in the blue region and certainly decreases exponentially with increasing wavelength, being near negligible in the NIR for the majority of the Great Lakes waters (Binding et al., 2012). The wavelengths also have a minimal sensitivity to TSM, but the absorption and scattering of suspended matters can still interfere within the chl-a algorithm selected wavebands. Increasing sediment loads result in the reflectance peak to move from blue to green to red in turbid waters (Bukata et al., 1995). Therefore, the semi-empirical models need to be tuned for each water body of interest characterized by different optical properties, in order to obtain the optimized wavelengths that can discriminate algal from suspended matters, and result in improved retrieval accuracy.

In addition, one has to consider that there is a relatively higher level of errors in computing water-leaving reflectance at longer wavelengths. Water absorption in red-NIR is strong and produces less water-leaving reflectance and, accordingly, a lower signal-to-noise ratio. This error is even higher in the case of clear waters where there is a low concentration of CDOM and TSM to produce water-leaving reflectance (Binding et al., 2012). In Lake Erie, however, the contribution of suspended and dissolved matters in water-leaving reflectance is high enough to allow the use of the proposed wavelengths in this study and produce a strong agreement between the modeled and observed values.

In the northern part of the western basin of Lake Erie, benthic algae can be seen at the surface when the water is clear enough. Consequently, in some remote sensing methods, benthic algae can contribute to the water leaving reflectance (Binding et al., 2012). In the present study, however, there is no need

to distinguish benthic algae from surface algae. The rapid in-water attenuation of the wavelengths selected in this study for chl-a model means that the remote sensing reflectance originates mostly from the upper 30 cm of water column in the lake (depends on the diffuse attenuation coefficient) (Binding et al., 2010). Therefore, there is no contribution of reflectance from algae at the bottom of the lake or subsurface. The estimated chl-a concentration is attributed to the surface or near surface algae even in the shallow areas or sections of the lake with clear water. The in situ samples to measure chl-a concentration in Lake Erie were collected from the surface mixed layer. There is a constant relationship between chl-a concentration at the surface and the one averaged over the mixing layer, as Lake Erie is shallow and exposed to strong wind-driven mixing to create a mainly mixed water column condition (Binding et al., 2010).

In situ data are required for algorithm evaluation purposes and also for parameterizing the LME models. The water quality parameters measured in the field can change at a scale smaller than that of the satellite image pixel resolution (300 m for MERIS), especially in Lake Erie due to different river inputs and wind effect. Thus, multiple measurements around stations are necessary to consider spatial heterogeneity. Also, the time lapse between satellite overpasses and in situ data collection may characterize a large change in the water quality parameter magnitude. The extent of these variations depends on the particular condition in the water body and defines the time window to be considered between satellite and in situ measurements. There are also some uncertainties associated with in situ data collections. Over- or underestimation of chl-a concentration measurements in the field is inevitable when the collected samples contain all of the pigments, due to spectral absorption overlaps (Arar, 1997a, 1997b; Arar and Collins, 1997; DosSantos et al., 2003). SDD measurements are subjective and may vary depending on the operator's ability. In shallow water bodies, disk contrast disappears at a shorter depth due to bottom reflections. Also, the disk can reach the bottom of the shallow parts of the lake without disappearing (Civera et al., 2013), which is not the case in Lake Erie as the depth measured in the survey was always larger than SDD (Binding et al., 2012).

4.4 Conclusion

This paper presented and assessed a remote sensing approach that utilizes spectral bands in the red and NIR portions of the spectrum to estimate chl-a concentration, and visible bands to determine SDD from MERIS images obtained over Lake Erie for the 2004-2012 period. LME models were developed based on the selected bands and in situ measurements. This method presents advantages over the traditional regression models that are only based on fixed effects, and that do not consider the correlation that

stems from repeated measurements in space and time. Also, the LME models for chl-a and SDD are semi-empirical models that, unlike the semi-analytical models, do not require detailed knowledge of the inherent optical properties (IOPs) of the CPAs in water.

Despite the limitations of remote sensing methods, they can still be considered as providing a complementary approach for the estimation of parameters related to water optical properties for many lakes over large areas and with frequent temporal coverage. In situ measurements of water quality parameters at sufficient temporal and spatial resolutions are, on the other hand, also problematic due to field logistics and extended periods without sampling as a result of changes in funding priorities by agencies. Environmental managers and policy makers can use satellite-derived information in support of decision-making programs. These programs require synoptic measurements of water quality parameters that can vary significantly in time and space. Measurements at an acceptable frequency are required in order to discern potential water quality problems associated with the lake. Upcoming Sentinel-3a and b satellite missions of ESA, which will each carry the OLCI (Ocean and Land Colour Instrument) sensor (heritage of MERIS), will mark a new era in the measurement of lake water quality parameters from space. OLCI has an optimized design to minimize sun-glint and will provide 21 spectral bands compared to the 15 bands available from MERIS.

Chapter 5
Satellite-Derived Light Extinction Coefficient and its Impact on Water Temperature Simulations in a 1-D Lake Model

5.1 Introduction

There has been significant progress made in recent years in the representation of lakes in regional climate models (RCM) and numerical weather prediction (NWP) models. Lakes are known to be as important as land surface component affecting weather and climate, especially in lake-rich regions of the northern hemisphere (Eerola et al., 2010; Samuelsson et al., 2010; Martynov et al., 2012). They can influence the atmospheric boundary layer by modifying the air temperature, wind and precipitation. Therefore, consideration of lakes in NWP/RCM is essential (Martynov et al., 2010; Kheyrollah Pour et al., 2012). In order to account for lakes in NWP/RCM, a description of energy exchanges between lakes and the atmosphere is required (Eerola et al., 2010). Lake Surface Water Temperature (LSWT) is one of the key variables when investigating lake-atmosphere energy exchanges (Kheyrollah Pour et al., 2012). There are various approaches to obtaining LSWT and integrating it in NWP models, such as through climatic observations, assimilation, and lake parameterization schemes (Eerola et al., 2010; Kheyrollah Pour et al., 2014a; Kheyrollah Pour et al., 2014b). Currently, LSWT is broadly modeled in NWP models using one-dimensional (1-D) lake models as lake parameterization schemes (Martynov et al., 2012). For instance, the 1-D Freshwater Lake (FLake) model performs adequately in various lake sizes, shallow to relatively deep (artificially limited to 40-60 m depth (Kourzeneva et al., 2012a)), located in both temperate and warm climate regions (Kourzeneva, 2010; Martynov et al., 2010, 2012; Mironov et al., 2010, 2012; Samuelsson et al., 2010; Kourzeneva et al., 2012a; Kourzeneva et al., 2012b).

One of the optical parameters required as input in the FLake model is water clarity. This variable is considered as an apparent optical property and is parameterized using the light extinction coefficient (K_d) to describe the absorption of shortwave radiation within the water body as a function of depth (Heiskanen et al., 2015). A global constant value of K_d is usually used to run lake models, including FLake. For example, Martynov et al. (2012) coupled FLake in the Canadian Regional Climate Model (CRCM) by specifying a K_d value equal to 0.2 m^{-1} (Martynov, pers. comm., 2015) for all North American Lakes, including Lake Erie for years 2005-2007. Heiskanen et al. (2015) evaluated the sensitivity of two 1-D lake models, LAKE and FLake, to the seasonal variations and the general level

of K_d for simulating water temperature profiles and turbulent fluxes of heat and momentum in a small boreal Finnish lake. Modeled values were compared to those measured for the lake during the ice-free period of 2013. The study found a critical threshold for K_d (0.5 m^{-1}) in 1-D lake models. They conclude that for too clear waters ($K_d < 0.5\ m^{-1}$), the model is much more sensitive to K_d. The study recommends a global mapping of K_d to run FLake model for regions with clear waters ($K_d < 0.5\ m^{-1}$) for future use in NWP models. The authors suggest that this global mapping can be time-independent (i.e. with a constant value per lake) (Heiskanen et al., 2015), and this can be derived from satellite imagery. Potes et al. (2012) used empirically derived water clarity from Medium Resolution Imaging Spectrometer (MERIS) satellite measurements to test the sensitivity of FLake to this parameter. The sensitivity analysis was conducted using two K_d values, representing the expected extreme water turbidity cases for their study (1.0 m^{-1} for clear water and 6.1 m^{-1} for turbid water). The importance of lake optical properties was evaluated based on the evolution of LSWT and heat fluxes. Their results show that the water clarity is an essential parameter affecting the simulated LSWT. The daily mean LSWT range increased from 1.2 ºC in clear water to 2.4 ºC in turbid water (Potes et al., 2012). Water clarity measurements are included in water quality monitoring programs; however, global measurements of turbidity are not yet available. Satellite remote sensing can provide lake models with turbidity observations at higher spatial and temporal resolutions, to fill the gap of field measurements.

In recent years, a number of algorithms have been devised to retrieve different water optical parameters, including water turbidity, from satellite observations for coastal (ocean) and lake waters (Binding et al., 2007, 2015; Wu et al., 2009; Zhao et al., 2011; Potes et al., 2012; Attila et al., 2013; Olmanson et al., 2013). Coastal waters are optically more complex compared to open ocean, and large optical gradients exist. There is more than only one component (phytoplankton species, various dissolved and suspended matters with non-covarying concentrations) in coastal waters and lakes that determines the variability of water-leaving reflectance. Considering this complexity, the development of algorithms for coastal waters is more challenging. MERIS, which operated from March 2002 to April 2012, collected data from the European Space Agency's (ESA) Envisat satellite. The spatial resolution and spectral bands settings were carefully selected in order to meet the primary objectives of the mission; addressing coastal monitoring from space. The best possible signal-to-noise ratio, additional channels to measure optical signatures as well as the relatively high spatial resolution of 300 m are some of the specific instrument characteristics (Ruescas et al., 2014). In 2010, ESA launched the CoastColour project to fully exploit the potential of MERIS instrument for remote sensing of coastal zone waters. CoastColour (CC) is providing a global dataset of MERIS full resolution data of coastal

zones which are processed with best possible regional algorithms to produce water-leaving reflectance and optical properties (Ruescas et al., 2014).

The objectives of this study are to: 1) evaluate satellite-derived K_d values for a large lake in the Great Lakes region; 2) apply the evaluated satellite-derived K_d in FLake model to investigate the improvement of model performance to reproduce LSWTs. Three different values of K_d were used in the simulations: yearly average, monthly average, and a constant value to demonstrate the impact of a time-independent, lake-specific K_d value in simulating LSWT; and 3) understand the sensitivity of the FLake model to K_d values based on simulated LSWT, mean water column temperature (MWCT), and mixed layer depth (MLD) during the ice-free seasons of Lake Erie (from April to November). In this study, solar irradiance and K_d values were derived from satellite-based observations. Lake-specific water clarity maps were produced for Lake Erie. The value of integrating satellite-derived solar irradiance and K_d in the 1-D FLake model is determined.

5.2 Data and Methods

5.2.1 Study Site and Station Observations

Lake Erie (42° 11′N, 81° 15′W) is a large shallow temperate freshwater lake covering a surface area of 25,700 km^2. The lake is characterized by three basins: shallow western, central, and deep eastern basins with maximum depths of 19 m, 25 m, and 64 m, respectively. Lake Erie is monomictic with occasional dimictic years. The shallow western basin is polymictic, since stratification is destroyed by wind-driven mixing. However, the central and eastern basins develop stable thermoclines in summer (Gobler and Wilhelm, 2015). It is the shallowest and smallest by volume of the Laurentian Great Lakes (Daher, 1999). These characteristics make Lake Erie unique from the other Great Lakes.

The meteorological forcing variables required to run the model includes solar radiation, air temperature, air humidity, wind speed, and cloudiness. Daily meteorological data including air temperature, wind speed, and water temperature were supplied by the National Data Buoy Center (NDBC) of NOAA, station 45005 for 2003-2012. The location of this station is shown in Figure 5-1 (41°40' N, 82°23' W, and depth: 12.6 m). Air temperature is measured from 4 m above the water surface, and anemometer height is 5 m above the water surface to measure the wind speed, whereas the water surface has an elevation of 173.9 m above mean sea level. Water temperature is measured from 0.6 m below the water line. The NDBC station is selected to run the model, since water temperature observations collected by the buoy station can be used to evaluate the model output. Other

meteorological data required for model simulations at the NDBC station were obtained from additional nearby stations.

Solar radiation data was supplied by the National Water Research Institute (NWRI), Environment Canada (EC), from a station located in the western basin of Lake Erie (see Figure 5-1). Daily measurements were available only for 2004 and 2008. Therefore, a daily time series of solar radiation data used to run FLake for the entire study period (2003-2012) at the NDBC station was provided by a solar irradiance model. Details on the solar irradiance model are given in section 5.2.2.

Air humidity and cloudiness was available in a daily format, and provided by EC-Ontario Climate Center (OCC) from the Windsor station (climate ID: 6139525) for 2003-2012. The location of this station is shown in Figure 5-1, which is a near-shore station close to the NDBC station. The distance between OCC and NDBC stations is less than 81 km.

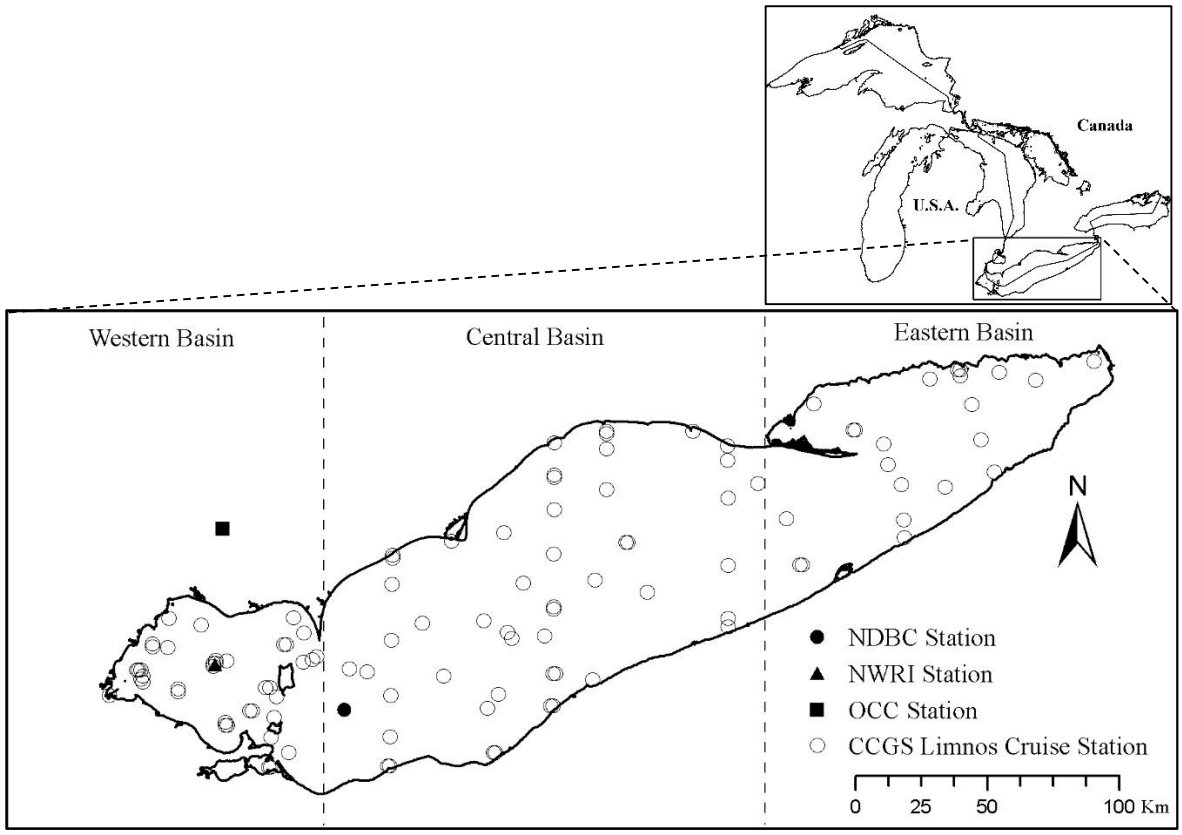


Figure 5-1 Maps showing Lake Erie in Laurentian Great Lakes and the location of stations where different parameters were measured.

FLake requires information on water transparency (downward light K_d). MERIS satellite imagery is used to derive K_d for the NDBC station during the study period. The method is described in details in section 5.2.3. Available Secchi disk depth (SDD) field measurements were used to estimate lake water turbidity. The data was provided by EC and utilized in this study to evaluate the satellite-derived K_d. Research cruises on board the Canadian Coast Guard Ship *Limnos* visited Lake Erie at a total of 89 distributed stations in five different years (September 2004; May, July, and September 2005; May and June 2008; July and September 2011; and February 2012). Locations of these stations are shown in Figure 5-1.

5.2.2 Satellite Solar Irradiance Model

The SUNY model, a satellite solar irradiance model, has been developed to exploit Geostationary Operational Environmental Satellites (GOES) for deriving solar irradiance using cloud, albedo, elevation, temperature, and wind speed observations (Kleissl et al., 2013). The state of the atmosphere and the Earth's cloud cover can be monitored from geostationary satellites. The basic principles of solar-irradiance modeling based on inputs from geostationary satellites and atmospheric models are described in Kleissl et al. (2013). Data from these sources are used to generate site and time specific high-resolution maps of solar irradiance with the SUNY model. This model has been regularly updated. The first version of SUNY model was based on using the visible images from GOES. It was introduced in 2002 based on earlier works of Cano et al. (1986) and Zelenka et al. (1999). To overcome the weaknesses of first version related to arid regions, version 2 of the model was introduced in 2004. The model was thoroughly validated against 10 US locations representing a wide range of climatic environments (Perez et al., 2002). The second version is used by the National Renewable Energy Laboratory (NREL) in the National Solar Radiation Database (NSRDB) (Wilcox, 2012). The model has been continuously improved over years. Versions 3 and 4 are also developed to improve the performance of the model to resolve the snow and cloud cover using both visible and infrared channels imagery, and across a diversity of climatological conditions (Gueymard et al., 2015).

The second version of the SUNY model (Version 2.4) is available in SolarAnywhere® (https://www.solaranywhere.com), providing the daily mean solar irradiance data for this study. The model provides a gridded data set with a spatial resolution of one tenth of a degree (ca. 10 km). The solar irradiance data was extracted from a tile corresponding to the NWRI station (see Figure 5-1) for 2004 and 2008, when observations were available for evaluation, and also for FLake model run on Lake Erie for the full study period (2003-2012).

5.2.3 Satellite-Derived Extinction Coefficient

MERIS operated on-board the ESA Envisat polar-orbiting satellite until April 2012. The sensor was a push-broom imaging spectrometer which measured solar radiation reflected from the Earth's surface high spectral and radiometric resolutions with a dual spatial resolution (300 m and 1200 m). Measurements were obtained in the visible and near-infrared part of the electromagnetic spectrum (across the 390 nm to 1040 nm range) in 15 spectral bands during daytime, whenever illumination conditions were suitable, and with a full spatial resolution of 300 m at nadir, with a 68.5° field-of-view. MERIS scanned the Earth with a global coverage of every 2-3 days.

In this study, a total of 326 full resolution archived MERIS images encompassing the NDBC station in Lake Erie (see Figure 5-1) were acquired from CC (Version 2) products through the Calvalus on-demand processing service for the period of 2003-2012. CC Level2W products are the result of in-water processing algorithms to derive optical parameters from the water leaving reflectance. These parameters include inherent optical properties (IOPs), concentrations of water constituents, and other optical water parameters such as spectral vertical K_d. The IOP parameters are first derived applying two different inversion algorithms: neural network (NN) and Quasi Analytical Algorithm (QAA). The derived IOPs are then converted to estimate constituents' concentrations and apparent optical properties (AOP), including diffuse K_d for different spectral bands applying Hydrolight simulations (Ruescas et al., 2014).

Diffuse K_d product in CC MERIS L2W data were evaluated against SDD in situ data collected during *Limnos* cruises. The CC-derived diffuse K_d values were extracted for pixels on the same day and location as the *Limnos* cruise stations. The satellite-derived K_d values were then extracted for the pixel at the geographic location of the NDBC station. A valid pixel expression was defined in all pixel extraction steps that excluded pixels with properties listed in Table 5-1.

Table 5-1 Flags of excluded pixels

Level 1	Level 1P	Level 2
Glint_risk	Land	AOT560_OOR (Aerosol optical thickness at 550 nm out of the training range)
Suspect	Cloud	TOA_OOR (Top of atmosphere reflectance in band 13 out of the training range)
Land_ocean	Cloud_ambigious	TOSA_OOR (Top of standard atmosphere reflectance in band 13 out of the training range)
Bright	Cloud_buffer	Solzen (Large solar zenith angle)
Coastline	Cloud_shadow	NN_WLR_OOR (Water leaving reflectance out of training range)
Invalid	Snow_ice	NN_CONC_OOR (Water constituents out of training range)
	MixedPixel	NN_OOTR (Spectrum out of training range)
		C2R_WHITECAPS (Risk of white caps)

5.2.4 FLake Model and Configuration

The FLake model is a self-similar parametric representation (assumed shape) of the temperature structure in the four media of the lake including water column, bottom sediments, and in the ice and snow. The water column temperature profile is assumed to have two layers: a mixed layer with constant temperature and a thermocline that extends from the base of mixed layer to the lake depth. The shape of thermocline temperature is parameterized using a fourth-order polynomial function of depth that also depends on a shape coefficient C_T. The value of C_T lies between 0.5 and 0.8 so that the thermocline can neither be very concave nor very convex. FLake has an optional scheme for the representation of bottom sediments layer, which is based on the same parametric concept (Martynov et al., 2012; De Bruijn et al., 2014). The system of prognostic equations for parameters is described in Mironov (2008).

The prognostics ordinary differential equations are solved to estimate the thermocline shape coefficient, the mixed layer depth, bottom, mean and surface water column temperatures, and also parameters related to the bottom sediment layers (Mironov, 2008; Mironov et al., 2010; Martynov et al., 2012). The mixed layer depth is calculated considering the effects of both convective and mechanical mixing, also accounting for volumetric heating which is through the absorption of net shortwave radiation (Thiery et al., 2014). The non-reflected shortwave radiation is absorbed after penetrating the water column in accordance with the Beer-Lambert law (Mironov, 2008; Mironov et al., 2010; Martynov et al., 2012).

Stand-alone FLake simulations were conducted for the NDBC station. The setup conditions of NDBC buoy station (height of the wind measurement: 5 m, height of the air temperature measurements: 4 m, depth of the water temperature measurements: 0.6 m), the measured meteorological parameters,

as well as the geographic location and depth of this site (41°40' N, 82°23' W, and depth: 12.6 m) were used to configure the FLake model. A fetch value of 100 km was used to run all simulations. It was found that there is only little sensitivity to modifications in this parameter for Lake Erie. The same result found for Lake Kivu in Thiery et al. (2014). The bottom sediments module was switched off in all simulations and the zero bottom heat flux condition is adopted. The ability of FLake to reproduce the observed temperature variations using different K_d values was tested by comparing the simulated LSWT to the corresponding in situ observations in the NDBC station. Also, the model sensitivity to variations in water turbidity was assessed studying the LSWT, MWCT, and MLD.

5.2.5 Accuracy Assessment

To assess the model outputs, three statistical indices were calculated: the root mean square error (RMSE), the mean bias error (MBE), and the index-of-agreement (I_a). RMSE is a comprehensive metric that combines the mean and variance of model errors into a single statistic (Moore et al., 2014). The MBE is calculated as the modeled values minus the in situ observations. Therefore, a positive (negative) value of this error shows an overestimation (underestimation) of the parameter of interest. I_a is a descriptive measure of model performance. It is used to compare different models and also modeled against observed parameters. I_a was originally developed by Willmott in the 1980s (Willmott, 1981) and a refined version of it was presented by Willmott et al. (2012). The refined version, which was adopted in this study, is dimensionless and bounded by -1.0 (worst performance) and 1.0 (the best possible performance). These statistical indices are considered as robust measures of model performance (e.g. Hinzman et al., 1998; Kheyrollah Pour et al., 2012; Willmott and Wicks, 1980).

5.3 Results and Discussion

5.3.1 Evaluation of Modeled Solar Irradiance Data

The solar irradiance data derived from SUNY model were evaluated using the corresponding available NWRI-EC observations for 2004 and 2008. Figure 5-2 shows daily variations of solar irradiance derived from NWRI-EC observations and the paired SUNY model derived values for the two years. These two solar radiation sources are highly correlated with coefficient of determination value (R^2) equal to 0.93. The SUNY model has a high agreement with observed solar irradiance data with a slight underestimation of 2.18 Wm^{-2} (N = 362, RMSE = 21.58 Wm^{-2}, MBE = -2.18 Wm^{-2}, I_a = 0.88) (Figure 5-3).

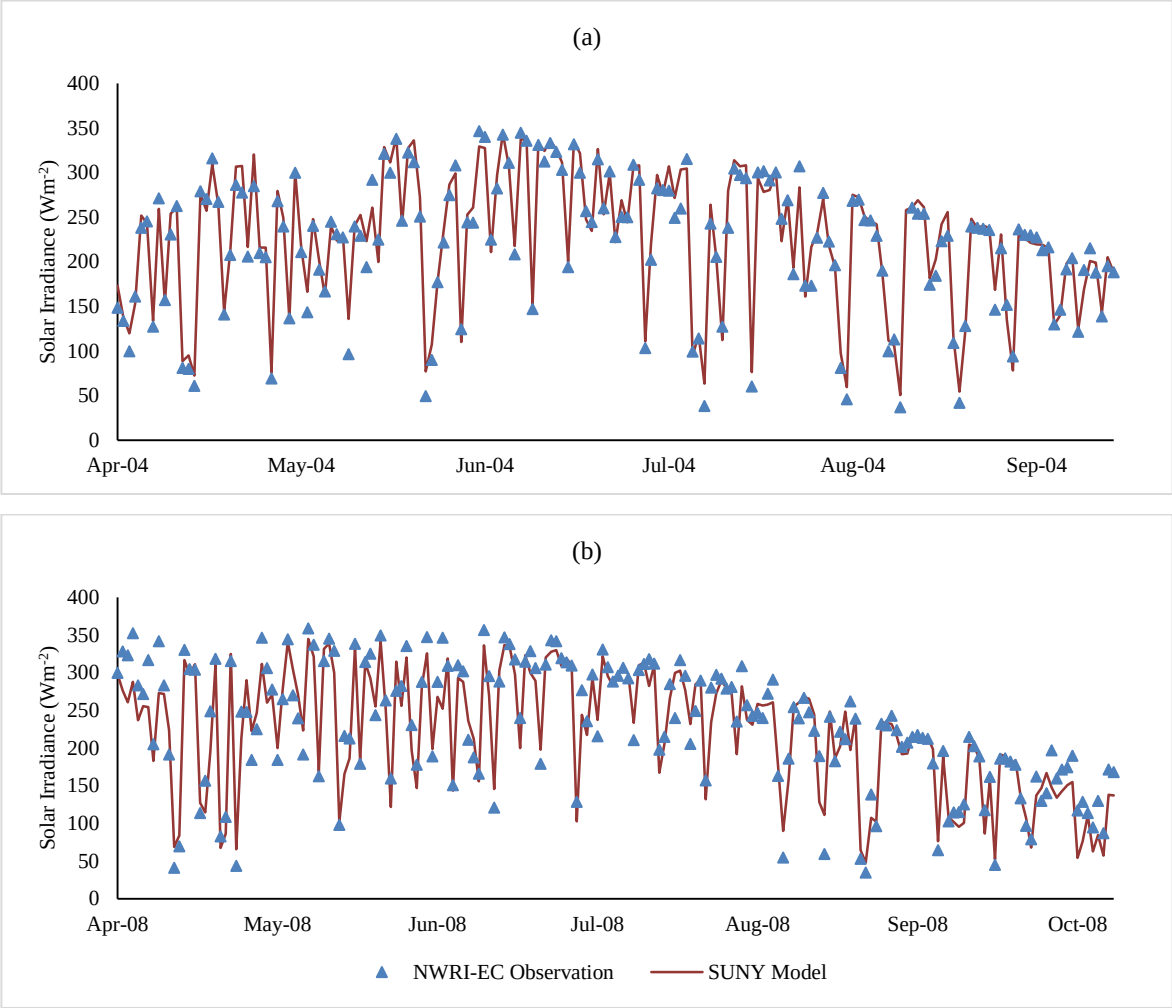


Figure 5-2 Comparison of daily variations of NWRI-EC observations versus SUNY solar irradiance model: a) April - September 2004, b) April - October 2008.

The accuracy of the SUNY model significantly depends on its ability to recognize the difference between a cloudy and clear sky background conditions (Dise et al., 2013). A proper differentiation between cloud and ground reduces when the ground is highly reflective, resulting in an increase in the model uncertainty. For example, larger errors are produced in regions with prolonged persistent cloud cover, low-lying clouds, and snow cover (Vignola et al., 2007; Nottrott and Kleissl, 2010; Perez et al., 2010). Nottrott and Kleissl (2010) compared the satellite-derived solar irradiance data from the SUNY modeled dataset in the NSRDB to the measurements for 27 weather stations in California during the years 1998-2005. The spatial and temporal differences between these two datasets were analyzed and related to the meteorological phenomena. Generally, for the study period, the SUNY model produced

accurate and high quality irradiance data with MBE of 5%, which was smaller than the sensor accuracy in ground stations. However, the study illustrates that the errors in SUNY-modeled solar irradiance data is larger near coastal stations than inland stations, particularly in the morning during summer. At coastal stations, year-round positive MBEs up to 18%, and monthly MBEs up to 54% in the summer months during morning were observed (Nottrott and Kleissl, 2010). Errors were attributed to problems with the model parameterization of dense cloud cover.

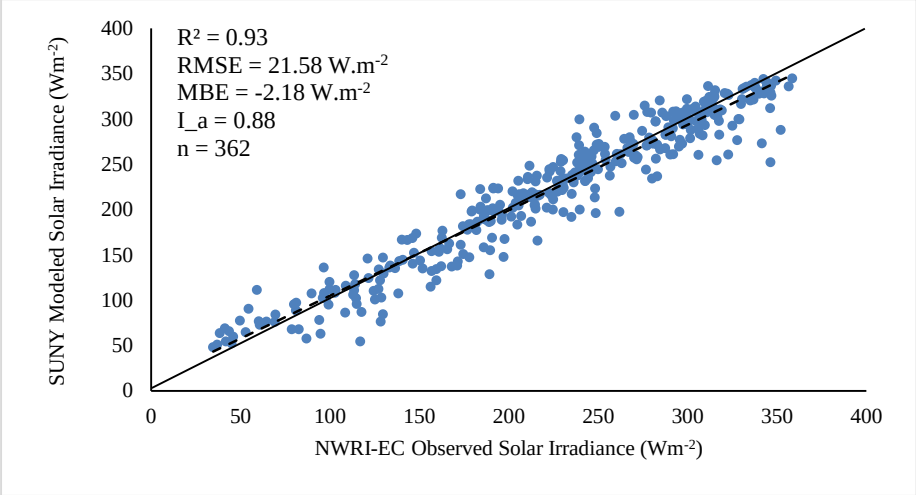


Figure 5-3 Scatter plot of NWRI-EC and SUNY mean daily solar irradiance (data from 2004 and 2008). The obtained statistical indices are included. The dashed line shows the best-fit line. Solid line corresponds to 1:1 relationship.

5.3.2 Evaluation of Satellite-Derived K_d

5.3.2.1 Evaluation of CoastColour K_d

The assessment of the satellite-derived K_d retrieval reliability highly depends on the comparison with independent in situ SDD measurements. The relation between K_d and SDD was established by the pioneer study of Poole and Atkins (1929):

$$SDD \times K_d = K \qquad \text{Equation 5.1}$$

where K is a constant value of 1.7 (Poole and Atkins, 1929). Studies have found a high variability of this constant depending on the type of the lake considered (Koenings and Edmundson, 1991). Armengol et al. (2003) also show that K_d and SDD are negatively correlated and they developed an empirical relation between these two parameters using Equation 5.1.

In this study, applying a cross validation approach, an empirical relation based on Equation 5.1 was developed between in situ measured SDD and CC-derived K_d. SDD measurements were conducted 117 times during cruises on Lake Erie from 2004 to 2012. These spatially-distributed measurements have minimum, maximum, mean, and standard deviation values of 0.2, 11, 3.69, and 2.68 m, respectively. CC L2W satellite products were acquired on the same day as the in situ measurements. Applying defined flags produced 49 data pairs (matchup dataset) of CC observations of K_d and SDD in situ data that were collected on the same day and location.

The matchup dataset was divided into training and testing data in 100 iterations. In each iteration, the data used for the equation's training and evaluation were kept independent, where 70% of the sample was used for equation calibration and 30% for evaluation. Ordinary least square regression was used in the calibration step of each iteration to relate the in situ measurements of SDD to the CC-derived K_d. Locally tuned equations were derived from this step and applied on SDD observations to predict K_d in testing matchup data. The statistical parameters of the model performance were derived between the estimated K_d from SDD observations and the paired CC-derived values. These steps were repeated for 100 iterations; and the final statistical indices, slope and power of the locally tuned equation was reported as the average of the ones derived over all iterations.

Results from the above procedure show that K_d and SDD are highly correlated with $R^2 = 0.78$. The extinction coefficient can be derived from equation $K_d = 1.64 \times SDD^{-0.76}$. There is a good agreement between the satellite-derived K_d and the corresponding ones estimated from in situ measured SDD (N = 49, RMSE = 0.63 m^{-1}, MBE = -0.09 m^{-1}, I_a = 0.65) (Figure 5-4). Arst et al. (2008) obtained a similar regression formula between SDD and K_d for the boreal lakes in Finland and Estonia representing different types of water, from oligotrophic to hypertrophic. SDD is a suitable characteristics to describe water transparency for small values of K_d. However, for high values of K_d (ranging above 4 m^{-1}), Arst et al. (2008) and Heiskanen et al. (2015) suggested that SDD is unable to describe any changes in K_d. Figure 5-4 also shows that SDD cannot describe the scatter of K_d for values above 4 m^{-1}. Therefore, the estimation of K_d from SDD should be used with caution, motivating the investigation on the potential of integrating satellite-based estimations of K_d into lake models.

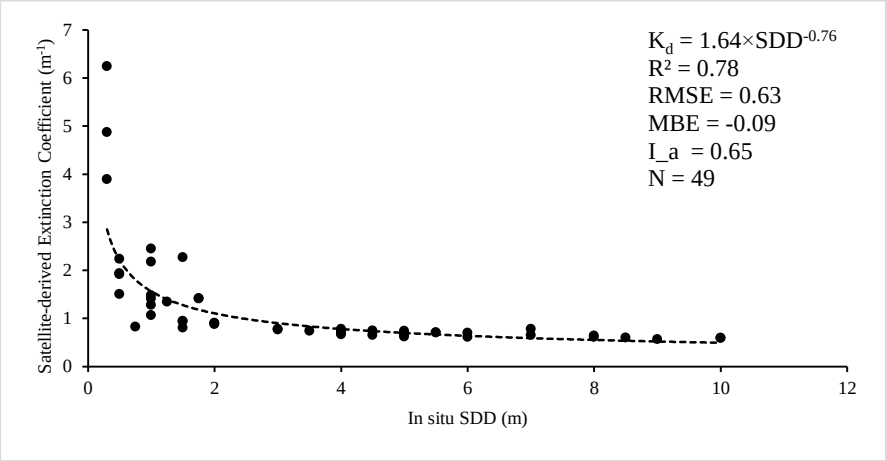


Figure 5-4 Relation between satellite-derived K_d and in situ SDD matchups

5.3.2.2 Spatial and Temporal Variations in K_d

This section described how the CC satellite observations can explain the spatial and temporal variations of K_d. Spatial variations of K_d derived from the CC algorithm are shown in Figure 5-5 for a selected day (3 September 2011). This particular day of 2011 is selected as the lake experienced its largest algal bloom in its recorded history in that year, before the new recent record of 2015 (Michalak et al., 2013; NOAA, 2015). The bloom was expanding from the western basin into the central basin. Algal bloom are one of the factors affecting the water clarity of Lake Erie (NOAA, 2015). Other parameters include the concentrations of suspended and dissolved matters in the lake. The western basin is the shallowest region of the lake; and therefore is the most vulnerable to sediment re-suspension that also results in reducing water clarity. The map shows that Lake Erie experienced different levels of turbidity in various locations with an average of 0.90±0.80 m^{-1} for all basins. The NDBC station is also shown on the map as a reference (with K_d = 0.87 m^{-1} on 3 September 2011).

Since fully cloud-free MERIS satellite images for consecutive months were only available in 2010, four months (May-August 2010) are selected to illustrate variations in K_d on a monthly-basis for one year (Figure 5-6). K_d over the full lake during May, June, July, and August has average values of 0.82±0.85 m^{-1}, 0.72±1.10 m^{-1}, 0.73±1.20 m^{-1}, 0.78±0.55 m^{-1}, respectively. The western basin is always experiencing the lowest levels of water clarity in comparison to other regions of the lake, with a maximum K_d in May. This can be the result of a spring algal bloom, and also wind-driven re-suspension of sediments. K_d at the NDBC station varies between 0.68 m^{-1}, 0.62 m^{-1}, 0.66 m^{-1}, and 0.85 m^{-1} from May to August 2010, respectively.

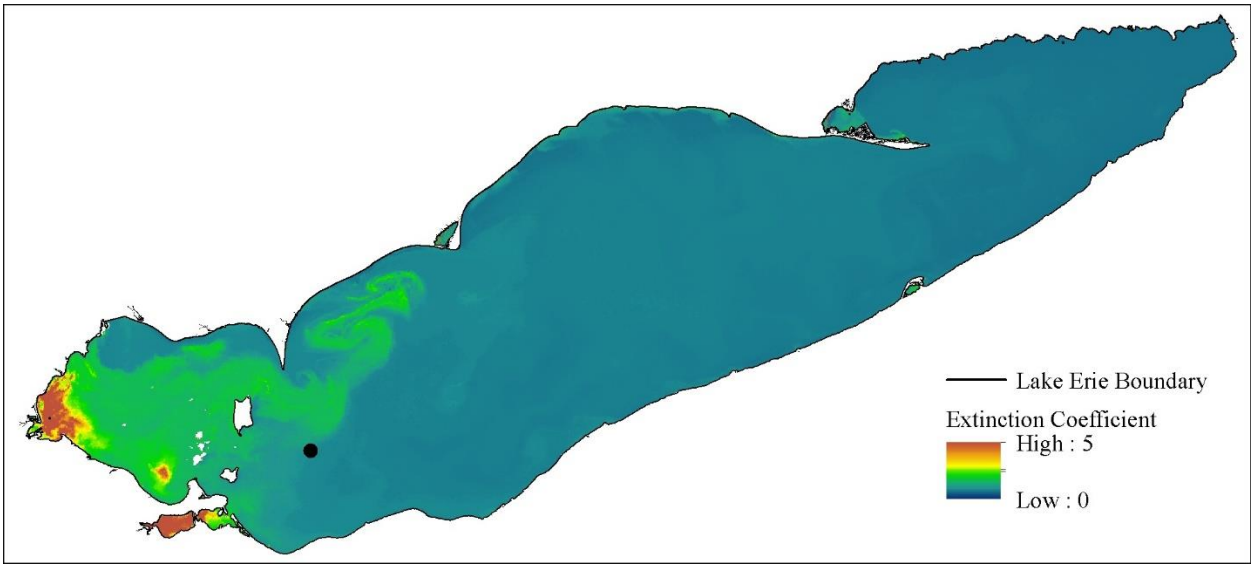


Figure 5-5 Spatial variation of satellite-derived K_d in Lake Erie, on 3 September 2011. Location of NDBC station is shown on the map as a solid dot.

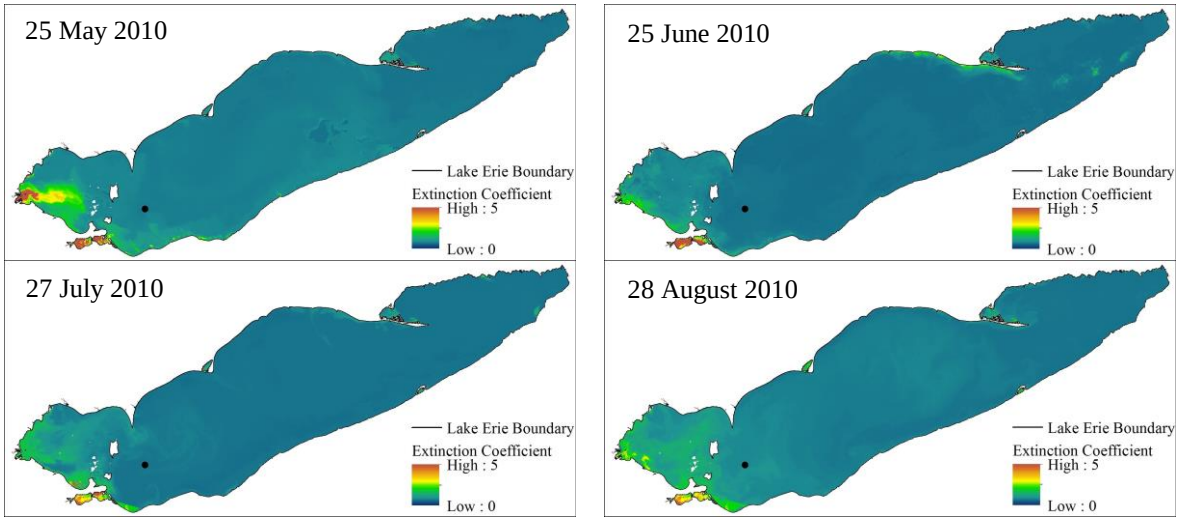


Figure 5-6 Temporal and spatial variation of satellite-derived K_d in Lake Erie for different months of a year: May- August 2010. Location of NDBC station is shown on the map as a solid dot.

Two MERIS images with full coverage of Lake Erie were only available in May of two consecutive years (2008 and 2009). Hence, the MERIS image of May 2008 and May 2009 were selected to show variations in K_d between two years. Although the images are for the same month of the year, K_d still varies across the lake (Figure 5-7). In May 2008 an average value of 0.77 ± 0.49 m^{-1} is estimated for the entire lake, while in May 2009 the average value is 0.90 ± 0.93 m^{-1}. Comparing the estimated maps for the two years suggests that the spring bloom in 2009 was stronger than the one in 2008 for the western

basin. However, algal bloom in all basins of Lake Erie for the complete year of 2008, was recorded as the third largest that the lake experienced before the occurrence of the breaking record blooms in 2011 and 2015 (Michalak et al., 2013; NOAA, 2015). K_d value estimated for the NDBC station is 0.69 and 0.62 m^{-1} in May 2008 and 2009, respectively.

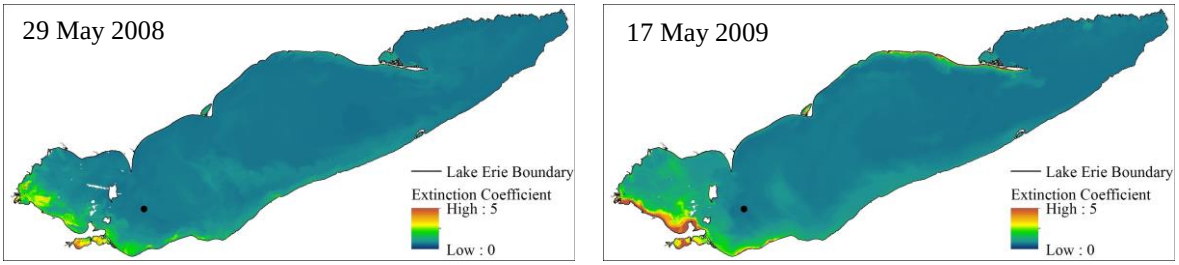


Figure 5-7 Temporal and spatial variation of K_d in Lake Erie during May of two consecutive years: 2008 and 2009. Location of NDBC station is shown on the map as a solid dot.

Figure 5-8 depicts variations of K_d for the NDBC station during the study period (2003-2012). In the shallow section of Lake Erie, re-suspension of bottom sediments is the most important factor that leads to higher water turbidity. Therefore the highest K_d values are related to the turn-over times in spring and fall. Smith et al. (1999) measured diffuse vertical attenuation coefficient for different stations in Lake Erie during May to August 1997. Infrequent high values of attenuation coefficient (>16 m^{-1}) was observed in very turbid waters, whereas 75% of the measured values were ≤2 m^{-1}. Hiriart et al. (2002) sampled two stations in turbid western basin during months of June, July and August 1997. Values of 1.61, 0.36, and 1.16 m^{-1} were observed in the months, respectively (Hiriart-Baer et al., 2002). The results from applying the CC algorithm on MERIS satellite imagery show that the maximum value of K_d is 3.54 m^{-1}, estimated in April 2003. A minimum value of 0.58 m^{-1} is estimated in June 2007. The average value of K_d during the study period is 0.90 m^{-1} with a standard deviation of 0.38 m^{-1}. Hence, these values, identified as the average, the lower, and the upper limits of turbidity at the NDBC station were used to carry out a sensitivity analysis with FLake (see section 5.3.3.2).

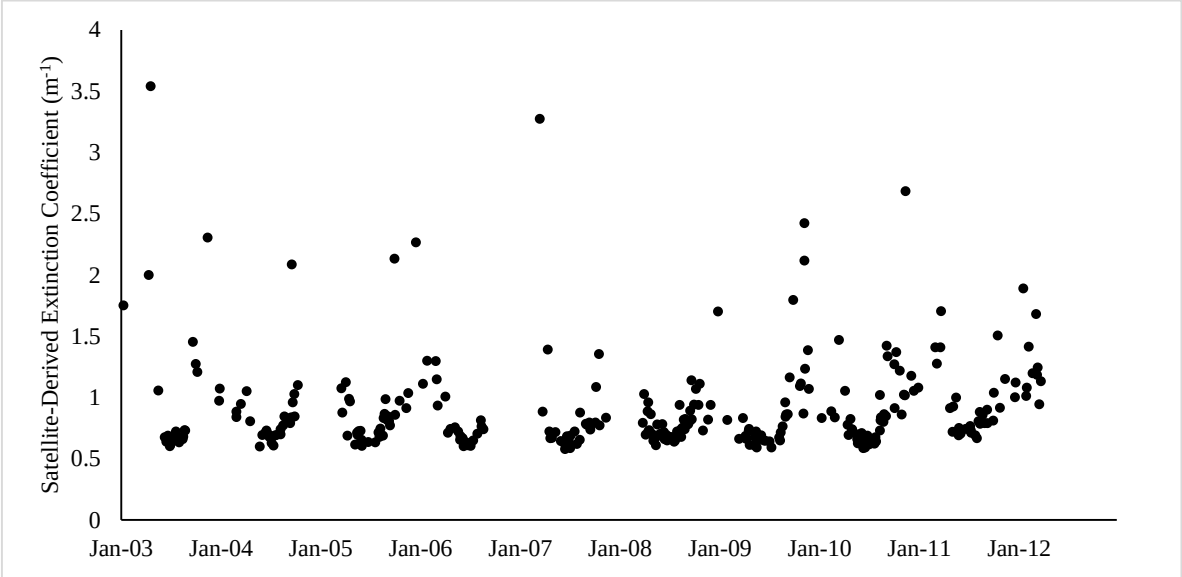


Figure 5-8 Variations of CC-derived K_d for the selected location during the study period (2003-2012).

5.3.3 FLake Model Results

5.3.3.1 Improvement of LSWT Simulations with Satellite-Derived K_d

Martynov et al. (2012) focused on 2005 to 2007 to run FLake at the NDBC station using a constant value of 0.2 m^{-1} for K_d. They simulated the lake using both realistic and excessive depths of 20 and 60 m, respectively, for a grid tile corresponding to the NDBC station. Applying a more realistic lake depth parameterization improved the performance of the model to reproduce the observed surface temperature. In this section, K_d values were derived from the CC algorithm for different months during the same years as in Martynov et al. (2012) (2005-2007). Table 5-2 displays the average K_d values for each month of these years. The monthly averaged values are only focused on the months of the year when both LSWT observations and CC–derived K_d values were available. The average value of K_d in these months in each year is considered as the average value of K_d for that year.

Table 5-2 CC-derived average values of K_d for each month (2005-2007). The values correspond to the time of year when water LSWT observations, as well as the CC derived K_d values, are available.

Year	Apr.	May	June	July	Aug.	Sep.	Oct.	Nov.	Avg.
2005		0.69	0.62	0.63	0.79	1.07	0.92	0.97	0.81
2006	0.82	0.70	0.62	0.65	0.77				0.71
2007	0.86	0.72	0.64	0.65	0.76				0.73

Figure 5-9 shows the results of different FLake simulations of LSWT at the NDBC station. The model was run first applying $K_d = 0.2\ m^{-1}$ from Martynov et al. (2012) using both the real lake depth at the station (12.6 m: CRCM-12.6) and also a tile depth corresponding to the station in their study (20 m: CRCM-20). Then, simulations using the yearly average K_d for each year of study are plotted (Avg). The K_d values derived from the monthly average of each year were used to simulate the surface water temperature and produce a merged LSWT product. Results of the merged product are also plotted (Merged).

Comparing LSWT in situ observations with the modeled values in Figure 5-9 demonstrate that in Avg and Merged simulations for 2005-2007, surface temperature in spring is modeled warmer and in summer-fall colder than in situ observations (spring: $MBE_{Avg} = 1.31$ ºC, $MBE_{Merged} = 1.25$ ºC; summer-fall: $MBE_{Avg\ Kd} = -1.27$ ºC; $MBE_{Merged\ Kd} = -1.37$ ºC). CRCM-12.6 and CRCM-20 are always producing a colder LSWT with maximum under-prediction in July-August (for 2005-2007: -2.93 ºC $< MBE_{July\text{-}August} <$ -0.99 ºC). Simulation with a larger depth (CRCM-20) tends to more slowly gain (lose) heat in spring (fall), compared to all other simulations.

The performance of each simulation is summarized in Table 5-3 during the period of data availability. For all years, the average and merged simulations perform better than simulations using K_d ($0.2\ m^{-1}$) as in Martynov et al. (2012), with improvement in RMSE and MBE for both real depth and tile depth. In all three years, LSWT simulated from the K_d value employed in Martynov et al. (2012) results in an underestimation (CRCM-12.6: MBE= -1.52 ºC, -0.98 ºC, -1.08 ºC; CRCM-20:-1.54 ºC, -1.09 ºC, -1.35 ºC; during years 2005, 2006, and 2007, respectively). In 2005, the average of for the year demonstrates a better performance compared to the merged results; contrary to the results of 2007. However, the merged results in 2006, the MBE is improved compared to the simulation using the average K_d; whereas its performance decreases in terms of RMSE. The extent of K_d variations in each month might not be captured by the available MERIS images. Therefore, the merged results cannot always perform better than the year average, which can be more representative of K_d variations.

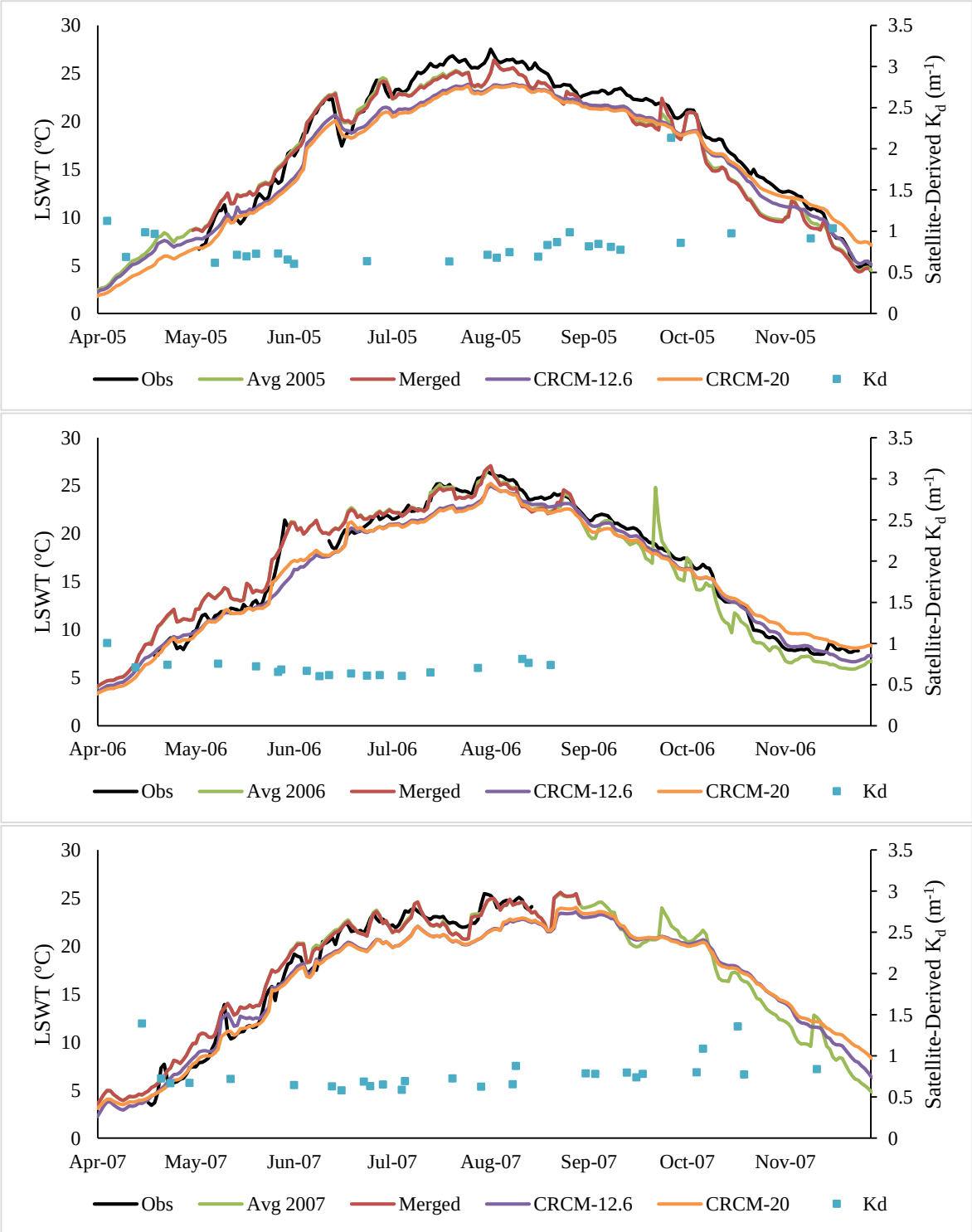


Figure 5-9 LSWT simulation results for 2005 - 2007; from: CRCM-12.6, CRCM-20, CC-derived average for K_d during selected month of each year (0.81, 0.71, and 0.73 m^{-1}; respectively), and the merged simulations based on each month average K_d. The corresponding observations for LSWT, and CC-derived K_d values are also plotted.

Table 5-3 Simulated LSWT compared to in situ observations (2005 – 2007). Period corresponds to the time of year when LSWT and K_d values were available.

Period	K_d	RMSE	MBE	I_a
2005 May-Nov	Avg2005	1.69	-0.86	0.87
	Merged	1.76	-0.95	0.86
	CRCM-12.6	1.88	-1.52	0.85
	CRCM-20	2.12	-1.54	0.83
2006 Apr-Aug	Avg2006	1.40	0.59	0.89
	Merged	1.42	0.54	0.89
	CRCM-12.6	1.50	-0.98	0.89
	CRCM-20	1.47	-1.09	0.89
2007 Apr-Aug	Avg2007	1.37	0.62	0.90
	Merged	1.35	0.57	0.91
	CRCM-12.6	1.78	-1.08	0.86
	CRCM-20	1.80	-1.35	0.87

Figure 5-10 illustrates the evaluation of simulation results for LSWT using the yearly average, and monthly average values of K_d for three years 2005-2007, altogether. CRCM-12.6 and CRCM-20 LSWT simulation results are also evaluated versus the corresponding in situ observations of LSWT. All simulations result in a high agreement with in situ measurements while underestimating the LSWT. However, the CRCM-20 simulation tends to produce the coldest LSWT (the most under-prediction) with MBE = -1.37 ºC. Simulations with the yearly average K_d results in the least underestimation of LSWT (MBE=-0.08 ºC).

Figure 5-10-a and -b show that the resulting LSWT from yearly average and monthly average K_d are not significantly different, whereas simulations with yearly average K_d reproduces LSWT with improved RMSE and MBE values compared to monthly average (Avg: RMSE=1.54 ºC, MBE=-0.08 ºC; Merged: RMSE=1.57 ºC, MBE=-0.14 ºC). It is possible that the extent of K_d variations is best represented by the yearly average value. Therefore, using a constant annual open water season value for K_d seems sufficient to reproduce LSWT in 1-D lake simulations with high accuracy. The time-dependent (monthly average) K_d does not improve simulation results for Lake Erie (K_d ranging from 0.58 to 3.54 m^{-1} with average value of 0.90 m^{-1} during open water seasons of 2002-2012). However, comparing results from Figure 5-10-a and –c shows improvement in LSWT simulations when a lake-specific value of K_d is used (Avg: RMSE=1.54 ºC, MBE=-0.08 ºC; CRCM-12.6: RMSE=1.76 ºC, MBE=-1.26 ºC). Under-prediction of LSWT decreases when the yearly-average CC-derived K_d values

are used, rather than a generic constant value (0.2 m^{-1}). Heiskanen et al. (2015) suggest that the effect of K_d seasonal variations on LSWT simulations are not significant for lakes with K_d values higher than 0.5 m^{-1} (e.g. Lake Erie) . Therefore, a lake-specific, time-independent, and constant value of K_d can be used in 1-D lake models.

Martynov et al. (2012) conclude that applying a more realistic lake depth parameterization improves the FLake model performance. Using the realistic lake depth at the NDBC station slightly improves the model performance in reproducing LSWT compared to simulations employing the corresponding tile depth (20 m) (CRCM-12.6: RMSE=1.76 ºC, MBE=-1.26 ºC; CRCM-20: RMSE=1.88 ºC, MBE=-1.37 ºC) (Figure 5-10-c and –d).

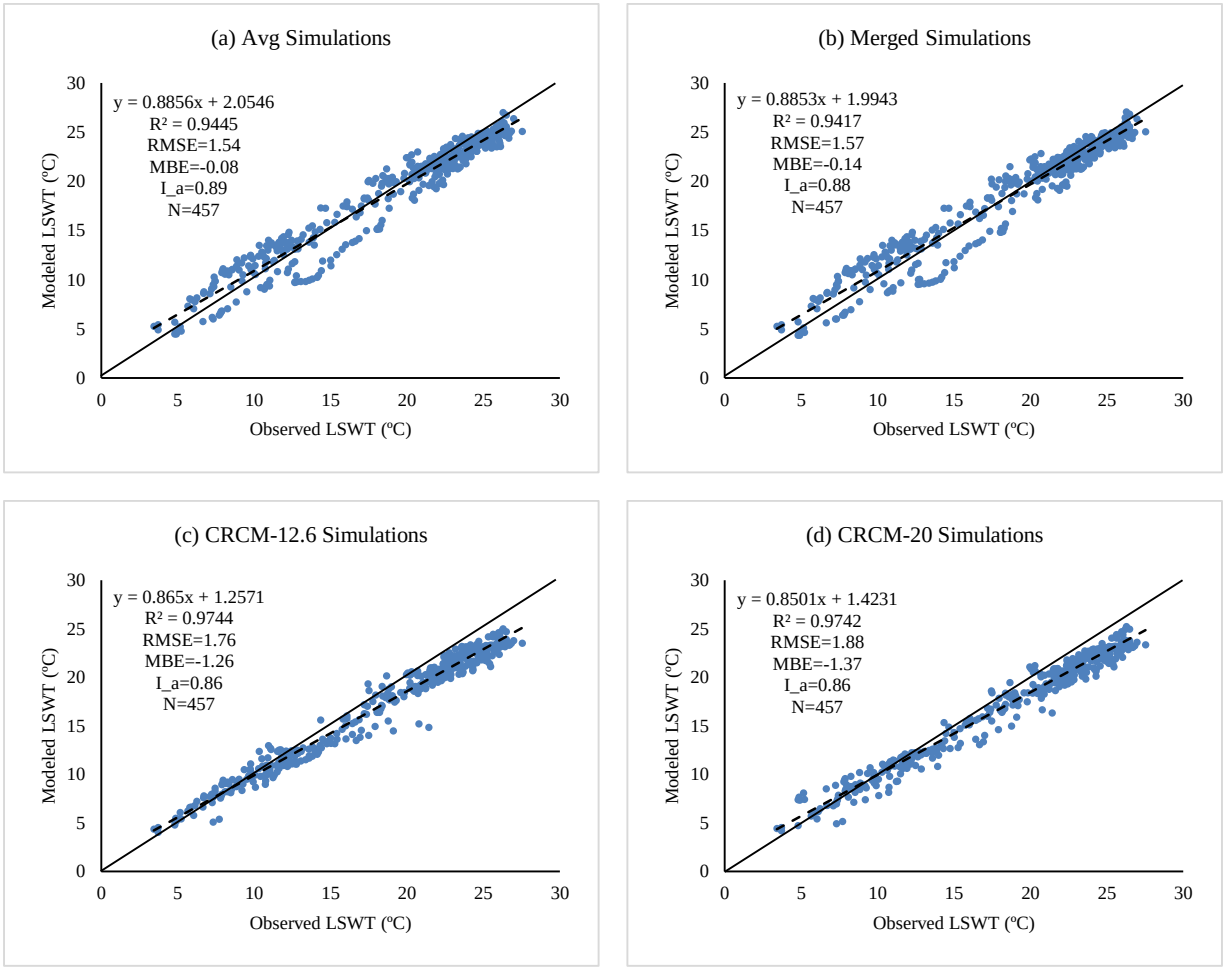


Figure 5-10 Modeled (y-axis) versus observed (x-axis) LSWT for yearly average, merged, CRCM-12.6, and CRCM-20 simulations during the ice-free seasons in 2005-2007. A linear fit (dashed line) and its coefficients are shown on the plot. The statistics related to the regression of parameters, and a 1:1 relationship (solid line) are also shown.

5.3.3.2 Sensitivity of FLake to K_d Variations

The sensitivity of FLake to reproduce LSWT, MWCT, and MLD using different values of K_d is investigated for 2008 in this section. The in situ data collected in 2008 provided a more complete dataset to run the sensitivity analysis due to the availability of incoming longwave radiation measurements during that year, which are necessary as FLake forcing data. Figure 5-11 presents simulation results for LSWT and MWCT in 2008 using the lowest, average, and highest values of K_d observed in the study period (minimum K_d=0.58 m^{-1}, average K_d =0.90 m^{-1}, maximum K_d =3.54 m^{-1}). The water temperature simulation from CRCM-12.6 (realistic depth at station) simulation is also plotted.

LSWT in the extreme clear water (K_d = 0.2 m^{-1}; CRCM-12.6) shows smoother variations during open water season in 2008 as opposed to the extreme turbid water simulation (maximum or Max) which displays more abrupt LSWT variations (Figure 5-11). This is because solar radiation is absorbed faster in turbid waters. It penetrates less deeply and warms up the shallow surface layer. This shallow layer exchanges heat faster with the atmosphere, resulting in sudden surface water temperature variations as opposed to clear waters. Also, the extreme turbid water simulation shows warmer LSWT in spring and colder LSWT in fall compared to the extreme clear water simulation results. In spring (the start of heating season), darker surface waters absorb heat faster than clear water because of existing particles in water. However, in fall the loss of energy to the atmosphere is also faster due to the shallow penetration of solar radiation into dark waters. On average, Max simulation resulted in 0.09 ºC higher LSWT compared to the average (Avg) simulation, whereas the minimum (Min) simulation produced on average 0.02 ºC colder LSWT during 2008. CRCM-12.6 simulation with K_d value of 0.2 resulted in a larger difference compared to Avg simulation, 0.55 ºC colder LSWT. Therefore, when K_d changes from its minimum to its average value (or from its maximum to its average value) the model is not significantly sensitive to reproduce LSWT, whereas a change in K_d value from its average value to 0.2 m^{-1} results in a larger variation in the modeled LSWT.

For both clear and dark waters, LSWT is warmer than the MWCT, due to being exposed to more intense solar radiation. Shortwave radiation is attenuated as it reaches a greater depth, particularly in turbid waters. In the extreme clear water simulation, the MWCT is on average 0.99º C colder than LSWT, whereas for the extreme dark water the average difference is 4.82º C.

The MWCT in the extreme turbid condition (Max) is less than for all other clear water simulations. This is because the lower layers in dark waters accumulate less heat during the heating season as opposed to clear waters which results in less heat storage and lower water column temperature in turbid

waters (Potes et al., 2012; Heiskanen et al., 2015). The solar radiation penetrates less deeply and is absorbed by the surface layer, thereby heating it; where the surface layer transfers the energy faster to the atmosphere, resulting in a colder water column in turbid waters. The MWCT decreases by 0.94 ^{0}C (increases by 0.63 ^{0}C) when K_d changes from its average to its maximum (minimum) value during the study period. The increase in MWCT is even larger when K_d changes from its average to 0.2 m^{-1} (2.25 ^{0}C). Therefore, K_d variations has a larger impact on MWCT than on LSWT, and the largest difference is when K_d is estimated to be extremely clear.

Rinke et al. (2010) conclude that the thermal structure of lakes is particularly sensitive to changes in K_d when its value is below 0.5 m^{-1}. More recently, Heiskanen et al. (2015) confirmed the critical threshold of K_d (ca. 0.5 m^{-1}). They suggest that the response of 1-D lake models to K_d variations is nonlinear. The models are much more sensitive if the water is estimated to be too clear. Heiskanen et al. (2015) recommend to use a value of K_d that is too high rather than too low in lake simulations, if the clarity of lake is not known exactly.

Figure 5-12 shows variations of the MLD in 2008 derived from simulations using the minimum, average, and maximum K_d values, and CRCM-12.6 simulation. In the extreme dark water simulation, the MLD is shallower than the other simulations with clear water (an average difference of 4.94 m in 2008 between two simulations with extreme K_d values). In turbid waters, solar radiation does not penetrate as far beyond the water surface as opposed to clear waters; and it will get absorbed by the particles in water. Therefore, clear waters have a deeper mixed layer when the solar radiation can penetrate further and distribute to a larger volume in the water column. Deepening of the thermocline layer in dark waters is slower compared to more clear waters, as there is a less penetration of solar radiation into the water column and also heat is transferred more slowly between water layers to stabilize the temperatures in different layers.

In all simulations, the maximum mixing depth occurs in both spring and fall. The maximum depth of spring mixing is at the same time for all simulations; however, fall mixing occurs at different times for each simulation. CRCM-12.6 reaches its maximum (fall) MLD at the end of summer, while it is reached in late fall, before an ice sheet forms on the water body, for all other simulations. The Min simulation reaches its maximum MLD in early November, before the Avg and Max MLD simulations (at the end of November). Therefore, the maximum fall MLD occurs earlier when the water was clearer. Due to the fact that the MLD is influenced by the water column thermal structure, the effect of K_d on the MLD is also larger when the K_d value is estimated to be less than 0.5 m^{-1}. CRCM-12.6 produces a

MLD 3.47 m deeper compared to Avg simulations, whereas the Min (Max) simulations result in MLD 1.15 m (1.47 m) deeper (shallower) compared to the Avg simulation.

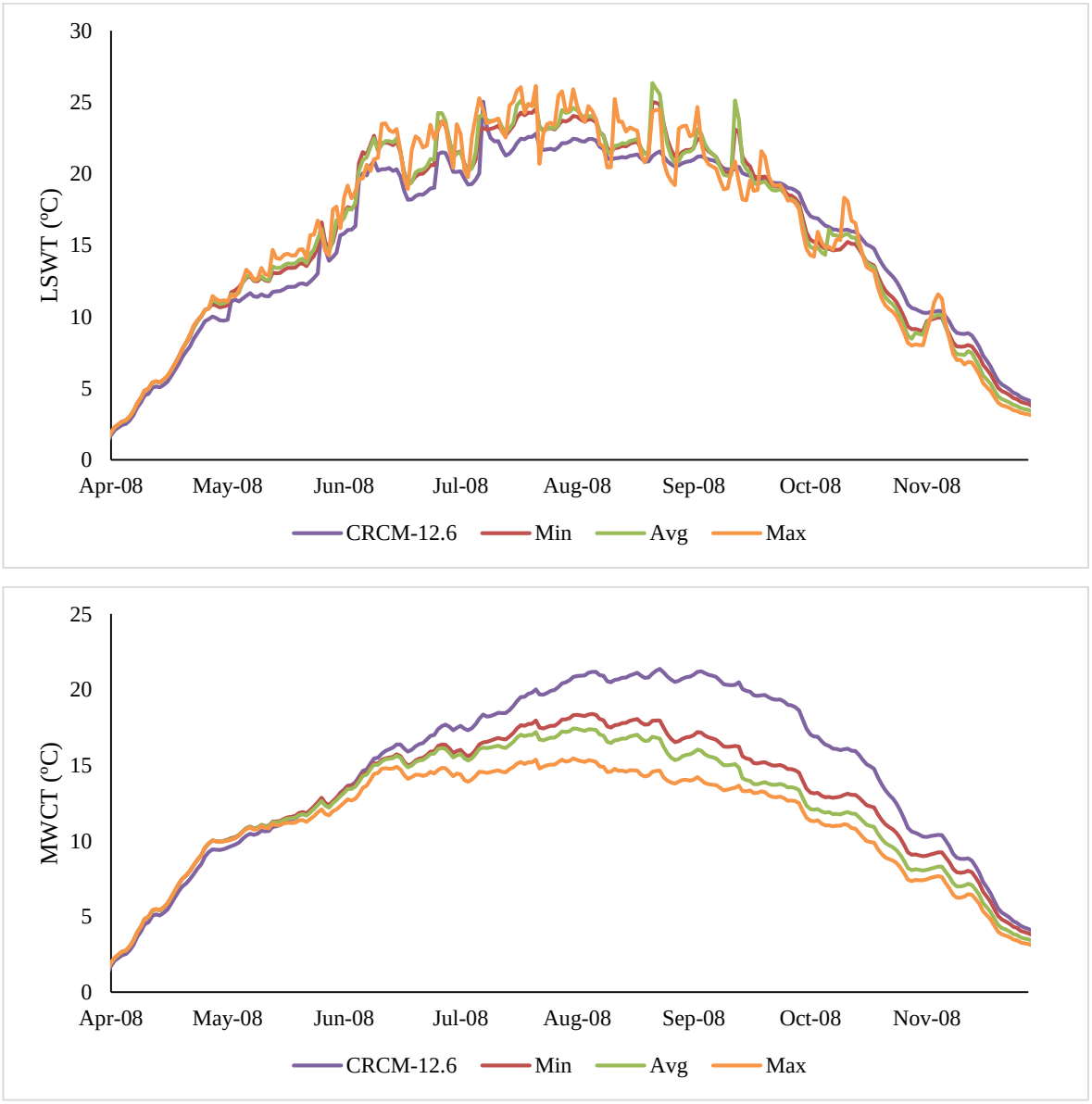


Figure 5-11 LSWT and MWCT simulation results in 2008, when using the lowest (Min), average (Avg), and the highest (Max) K_d values. Results from the CRCM-12.6 simulation is also plotted.

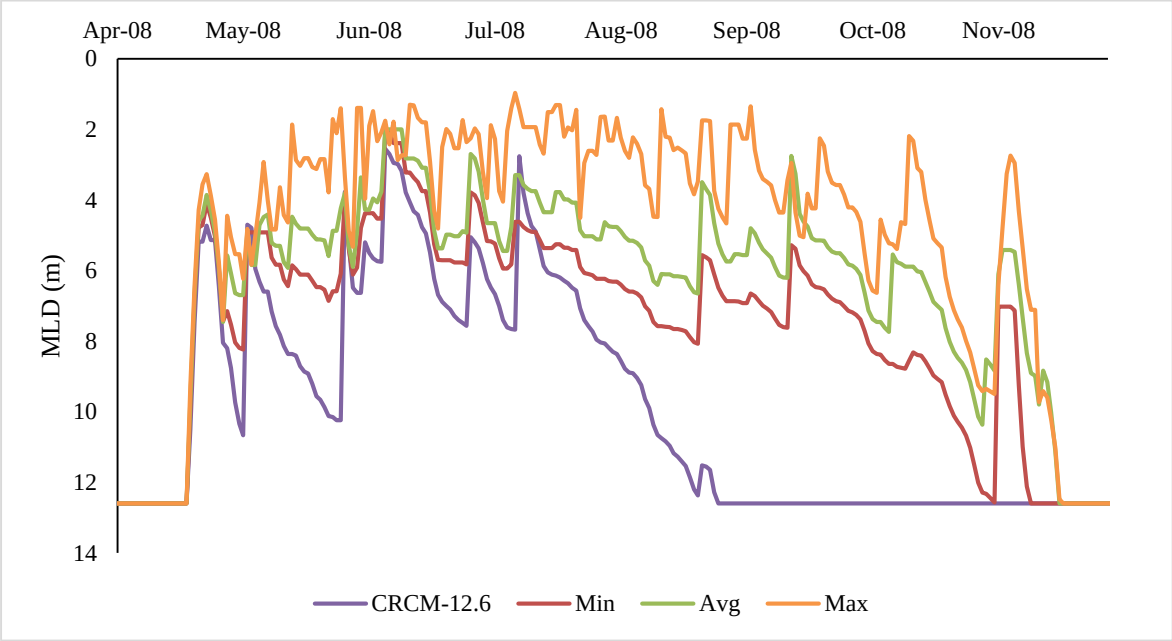


Figure 5-12 MLD simulation results for the lowest (Min), average (Avg), and the highest (Max) K_d values in 2008. CRCM-12.6 results are also plotted.

5.4 Summary and Conclusion

Spatial and temporal variations of K_d in Lake Erie were derived from the globally available satellite-based CC product during open water seasons 2003-2012. The CC product was evaluated against SDD in situ measurements. The solar irradiance data were obtained from a satellite-based solar-to-irradiance model and evaluated with corresponding in situ observations. There was a high agreement between the modeled and in situ measurements. The estimated K_d values and solar irradiance data, in addition to complementary meteorological observations during the study period, were used to force the 1-D FLake model. The model was run for a selected site (NDBC buoy station) on Lake Erie, a large shallow temperate freshwater lake.

The FLake model was run for 2005-2007, the same years as in Martynov et al. (2012) who conducted model simulations with a constant value of K_d (0.2 m^{-1}) at NDBC station on Lake Erie. Results presented herein clearly showed that employing satellite-derived K_d values improve the FLake model simulations. An annual lake-specific K_d value, derived from the CC satellite-based algorithm, reproduced LSWT with a better agreement to in situ LSWT observations compared to using the constant value of K_d employed in previous studies. Results also showed that, although K_d varies in time, a time-invariant (constant) annual value is sufficient to obtain reliable estimates of LSWT with FLake

for Lake Erie. This finding is in agreement with the recent study of Heiskanen et al. (2015) who determined that the impact of seasonal variations of K_d on the simulated thermal structure is small enough, for a lake with high K_d values (> 0.5 m^{-1}), to allow use of a constant value in time.

Sensitivity of the FLake model to the range of potential K_d variations was tested for the NDBC station. The model performance to reproduce LSWT, MWCT, and MLD was assessed. Different K_d values were used to run the model: minimum, average, maximum values of K_d that were derived from the satellite-based CC algorithm during the study period (2003-2012), and also the constant value used in Martynov et al. (2012) study. A warmer spring and colder fall was predicted running the model for dark water compared to clear water. The MWCT for clear water is warmer than the dark water. This is because solar radiation can penetrate deeper in clear water. This study also investigated the MLD for clear and dark waters. A shallower MLD is simulated for dark water. Therefore, the results showed that the 1-D FLake model is very sensitive to the variations in K_d. However, the sensitivity was even more significant when a K_d value less than 0.5 m^{-1} was used. Rinke et al. (2010) and Heiskanen et al. (2015) also found the critical threshold value of K_d to be 0.5 m^{-1}. These studies suggest that simulating the thermal structure of lakes using 1-D lake models is particularly sensitive to K_d changes when its value is less than 0.5 m^{-1}. The sensitivity of K_d variations was more pronounced in simulation results for MWCT and MLD compared to LSWT.

Integrating lake specific K_d values can improve the performance of 1-D lake models. However, field measurements of K_d are not widely available. This study demonstrates that satellite observations are a reliable data source to provide lake models with global estimates of K_d with high spatial and temporal resolutions. The globally available CC product can be used as a source to fill gaps in K_d in situ observations, and improve the performance of lake models such as FLake. Although MERIS is no longer active, the Ocean and Land Colour Instrument (OLCI) to be operated on the ESA Sentinel-3 satellite will provide continuity of MERIS-like data. OLCI has MERIS heritages and improves upon it with an additional 6 spectral bands.

Chapter 6
General Conclusions

6.1 Overall Summary

The overall aim of this research was to test the potential of different MEdium Resolution Imaging Spectrometer (MERIS) algorithms to derive chlorophyll-a (chl-a) concentration and water turbidity in the optically complex Lake Erie. The satellite-derived water turbidity, represented as the light extinction coefficient, was then used to test the potential of coupling the globally available MERIS product of water clarity with a lake model, where usually only a constant value of this variable is used. Different algorithms including the neural network (NN) and band ratio methods which are available in the Basic ERS & ENVISAT (A)ATSR MERIS (BEAM) software were validated using the in situ data collected on Lake Erie in Chapter 3. The NN algorithms include Case 2 Regional (C2R), Eutrophic (EU), FUB/WeW, and CoastColour (CC) algorithms, and the band-ratio algorithms used in the research were fluorescence line height (FLH), and maximum chlorophyll index (MCI). The FUB/WeW and C2R algorithms outperformed other algorithms with improvement in the index-of-agreement (I_a) and root mean square error (RMSE). C2R algorithm underperformed FUB/WeW algorithm with 11% increase in RMSE value. The variable results underlines that the processors performance is highly related to the optical properties of the water body and also the data range used to train the NNs. Thus, a final product was produced as the result of blending best performing algorithm for each specific optical class in Lake Erie. The optical classes were derived from applying a two-step automatic clustering method. The blended product has an improvement of about 3% in I_a compared to C2R algorithm when applied on all Lake Erie basins (blended: I_a= 0.73; C2R: I_a=0.71).

Besides the NN and band ratio algorithms investigated in Chapter 3, a semi-empirical regression model was also developed for Lake Erie in Chapter 4. The empirical relationships may not be transferrable to other locations and times. Thus, the algorithms are only applicable for Lake Erie and for the time when the in situ data were collected. The field measurements of the optical parameters were used to train the regression. A Linear Mixed Effect (LME) model was used as the regression method between the band ratio of MERIS-derived reflectance and in situ data of chl-a concentration and water turbidity (Secchi Disk Depth, SDD) in two separate models. This regression method has advantages over the traditional regression models which are based only on fixed effects, and not considering the correlation that stems from repeated measurements in space and time. The model was

developed between the logarithmic scale of the parameters and band ratio of B7:665 nm to B9:708.75 nm for log_{10}chl-a and band ratio of B6:620 nm to B4:510 nm for log_{10}SDD. The predicted log_{10}chl-a and log_{10}SDD were in close agreement with the ones measured in the field (log_{10}chl-a: I_a=0.67, RMSE=0.3; log_{10}SDD: I_a=0.88, RMSE=0.19). The monthly spatial and temporal patterns of chl-a concentration and SDD were also estimated applying the derived model on the images acquired over Lake Erie from 2004 to 2012.

There is also the globally available CC MERIS product for water turbidity (K_d). The archived globally mapped and reliable CC products for water ecosystem studies is open access and free to end-users. In Chapter 5 the product was evaluated against the in situ measurements of SDD to estimate the extinction coefficient (I_a=0.65). Extinction coefficient is an essential optical input in Freshwater Lake (FLake) model as a lake representing scheme. However, because of the limited in situ measurements of water turbidity, a constant value is usually used in the lake model. Therefore there is a need to produce the extinction coefficient maps that are derived specifically for each lake. According to this research, the globally available CC product can be used as a good estimator of lake water turbidity to provide the lake models with high spatial and temporal resolution satellite-based measurements of this optical parameter. Results clearly showed that the model is sensitive to variations of extinction coefficient. A time-independent, yearly average, and constant extinction coefficient improved Mean Bias Error (MBE) and RMSE to reproduce lake surface temperature, compared to employing a constant K_d value (0.2 m^{-1}) (lake-specific yearly average K_d value: RMSE=1.54 ºC, MBE= -0.08 ºC; constant K_d value: RMSE=1.76 ºC, MBE= -1.26 ºC).

Overall this research has shown the potential of satellite sensors to derive water optical properties that are essential parameters in lake management programs to monitor lake water quality and also to improve lake models that are used as lake parameterization schemes in high-resolution numerical weather prediction models. The major advantage of employing satellite-based measurements of water optical parameters is to derive high spatial and temporal resolution maps which cannot be produced relying only on in situ measurements.

6.2 Limitations

The distribution of chl-a concentration range in the matchup data did not represent the full variations in the lake, especially for values above 10 mg m^{-3}. There were only two measurements above 25 mg m^{-3} amongst the in situ data. However, there was no corresponding satellite observations for those

measurements after applying flags in C2R, EU, FUB/WeW, and CC algorithms. These two measurements, with chl-a concentrations above 25 mg m^{-3}, were taken at the same location in the western basin during July 2005 and June 2008. For chl-a concentrations between 10 and 25 mg m^{-3}, there were seven in situ observations. Four of them were removed due to cloud cover detection by all algorithms. FUB/WeW applies a different atmospheric correction algorithm compared to other NN algorithms. In Chapter 3, the FUB/WeW algorithm dropped the number of matchups for chl-a concentrations between 10 and 25 mg m^{-3} to only one. SDD matchup data derived from the CC algorithm, which were used in both Chapters 4 and 5, produced a good distribution of SDD values in the data range. Both matchup data sets resulted for chl-a and SDD were collected from different location on the lake, with a good spatial distribution. To improve the accuracy of the developed models, especially for estimating high chl-a concentrations, the matchup data had to represent the magnitude and spatial distribution of the parameter of interest for the lake. This was required since the data was to be used later to calibrate the output of different algorithms used in Chapter 3 or to develop the regression method in Chapter 4. Also the models' performance can be assessed better when the actual range of the parameter of interest in lake is used for the accuracy analysis. Therefore, additional measurements with high chl-a concentrations during algal blooms in Lake Erie would be useful.

The water samples were collected from the surface mixed layer to measure chl-a concentrations in Lake Erie. On the other hand, André (1992) stated that as a general rule the satellite observations are associated with an average concentration within a surface layer of a certain depth (André, 1992). There is no unique method of averaging the stratified layers in a heterogeneous water column within the specified surface layer thickness. However, in a homogenous (well-mixed) water column, which is hardly the case in nature, chl-a concentrations obtained from satellites can be proportional to chl-a concentrations at any depth. Different studies evaluated the information that satellite observations can provide concerning the vertical distribution of constituents within a defined thickness of the surface layer (Gordon and Clark, 1980; André, 1992; Zaneveld et al., 2005; Sokoletsky and Yacobi, 2011). According to the study of Gordon and McCluney (1975), water-leaving reflectance measured by satellite is originating from a depth known as the penetration depth which can be detected by optical sensors (Gordon and McCluney 1975). Gordon and Clark (1980) proposed the optically weighted pigment concentration in the penetration depth for the purpose of interpreting satellite information and comparing them with in situ data (Gordon and Clark, 1980). However André (1992) showed that where the concentration varies strongly within the first meters of a water body, the optically weighted concentration is inadequate to interpret the satellite observation. Therefore, to resolve the ambiguity of

the vertical representativeness of remotely sensed constituent concentrations, a vertically resolved in situ data is required on Lake Erie (Odermatt, Pomati, et al., 2012).

Water-leaving reflectance derived from applying atmospheric corrections in different NN algorithms was not validated. Atmospheric corrections are a critical step over water bodies, since the radiance signal emerging from the water column is much less than that of land. Atmospheric corrections become even more challenging in highly turbid inland and coastal waters where the 'black pixel' assumption of negligible water-leaving radiance in the near-infrared (NIR) is no longer valid due to scattering from suspended matters (Binding et al., 2010). Thus, typical atmospheric corrections fail and other schemes based on radiative transfer models or other approaches are required (Matthews et al., 2010). The accuracy of atmospheric correction algorithms used in different models is very important to evaluate the satellite-derived water quality products. In a band ratio algorithm with bands near each other, atmospheric effects are normalized (Matthews et al., 2010). However, for optimization methods such as NN algorithms, accurate atmospheric corrections are especially important. These methods use the water-leaving reflectance as an input to bio-optical models to derive the water constituents concentrations, and are based on iterations to reduce the difference between observed and predicted water-leaving reflectance (Matthews et al., 2010). Hence the usefulness of semi-analytical methods in turbid waters may be limited because of the accuracy of atmospheric corrections. The remote sensing measured water-leaving reflectance can be validated with concurrent field reflectance measurements.

There are also limitations associated with the design of satellite remote sensing instruments. Some requirements are necessary for ocean (lake) color remote sensing instruments. Because the sea (lake) is a much darker object than land surfaces, the radiometric resolution associated with sensors has to be sensitive to low levels while maintaining a sufficient range to not saturate in case of bright pixels happening due to sun glint. Furthermore, in the case of spectral band features, ocean color sensors need to have narrowly defined wavebands. The spectral peak related to the pigment absorption or scattering has a bandwidth of less than 20 nm (Chen et al., 2013). Both of these features, high radiometric resolution with narrow spectral bands, prevent having too narrow instantaneous-field-of-view (IFOV). The flux of photons to detectors has to be possible with an acceptable signal-to-noise ratio (SNR). To compensate for this, the spatial or temporal (or both) resolutions of satellite sensors need to be degraded.

A high spatial resolution demands a narrower swath and, consequently, the revisit interval is not sufficient for monitoring changes in water properties, which can occur in the matter of hours (Robinson, 2004). Small-sized waterbodies with Case II waters are more challenging for applying water quality

algorithms with remote sensing. They are more temporally dynamic with more heterogeneous areas. Consequently, medium to coarse spatial resolution of some satellite images (e.g. MERIS and MODIS) is not sufficient for monitoring small lakes and rivers, except for the very large ones. Also, measurements have to be performed with sufficient spectral details to detect and distinguish different pigments, and water components' concentrations (e.g. colored dissolved organic matter) in Case II waters. Spectral bands in the NIR are required to enhance the atmospheric information, and improve the results of atmospheric corrections.

Another limitation associated with optical sensors is when the lakes are covered by clouds. As a result, the images may not be useful for several days due to cloud coverage. Therefore, there is a need to upgrade cloud detection algorithms which may not detect clouds and result in anomalous errors. Some of these limitations are expected to be resolved following the launch of future satellites.

6.3 Future Research Directions

The applicability of NN algorithms over the water body of interest has to be examined in terms of suitability of the inherent optical properties (IOPs) and concentrations ranges. The bio-optical models used in the MERIS NN algorithms convert the derived IOPs to constituent concentrations based on measured SIOP coefficients. Because IOPs are also provided as algorithm outputs, this relationship can be re-parameterized. Thus, SIOP can be specified to suite the particular region of interest when in situ measurements of IOPs are available (Palmer et al., 2014).

The phytoplankton species composition in Lake Erie, according to the study of Binding et al. (2008), range from diatoms and dinoflagellates in the spring to cyanobacteria, including significant microcystis blooms, in summer (Binding et al., 2008). Cyanobacteria blooms tend to be intensified at the surface. Remote sensing can provide a significant benefit to detect, monitor, and forecast these events (Kudela et al., 2015). Kudela et al. (2015) demonstrate that cyanobacteria blooms can be detected with the availability of high spectral resolution satellite images, and appropriate atmospheric corrections using spectral-shape algorithms (Kudela et al., 2015). Vincent et al. (2004) employed Landsat TM sensor data, despite its relatively wide range of spectral bandwidths, to detect phycocyanin as an accessory pigment in Lake Erie. This pigment can be used as an indicator of cyanobacterial blooms (Vincent et al., 2004).

Considering that MERIS is no longer active, there are other aquatic color satellite missions that can provide information about inland waters and coastal zones with a tradeoff in the measurement

resolutions. Their applicability to monitor algal blooms on Lake Erie can be investigated. For example, the Landsat 8 OLI sensor offers a spatial resolution of 30 m suitable for monitoring small freshwater systems, with refined heritage bands, along with three additional bands (including a deep blue band at 443 nm which is chl-a absorption peak). However, the temporal resolution of 16-days is a drawback for monitoring highly dynamic processes.

The Sentinel-2a MSI (Multispectral Imager) sensor is complementary to Landsat 8 with added channels (excluding the thermal bands of Landsat 8's Thermal Infrared Sensor), with spatial resolutions of 10, 20, and 30 m, and a 10-day revisit frequency (initially with the one satellite launched; a 5-day revisit once both Sentinel-2a and 2b are in orbit). MSI covers the visible, near, and shortwave infrared regions of spectrum in 13 spectral bands. Landsat 8 and MSI only have four main visible and near-infrared bands with a spatial resolution of 30 and 10 m, respectively. Thus, their spatial resolutions are adequate for most lakes, while it is likely for them to have suitable sensitivity for estimating chl-a concentration utilizing a band at approximately 665 nm. The combination of MSI on Sentinel-2a,b with Landsat 8 will eventually provide a coverage of better than 5 days at relatively high spatial resolution.

The first Sentinel-3 is planned for launch in early 2016. Sentinel-3's OLCI (Ocean and Land Color Instrument) sensor is designed to harvest biological information from the ocean and land surfaces. It has a spatial resolution of 300 m (1 km in open ocean and sea ice, 300 m over coastal zones) to provide an equivalent precision and level of accuracy to MERIS satellite images with 1-2 days repeat coverage and high spectral resolution (spanning from visible to infrared). OLCI possesses 6 additional spectral bands (21 bands in visible and infrared) compared to MERIS. The spectral bands are used to provide better cloud screening, better atmospheric corrections, and discrimination of water components, land, and atmospheric changes in temporal series of data. A narrow channel at 673 nm (7.5 nm width) has been added to improve chl-a fluorescence measurements. There is an improvement over MERIS observations by mitigating the impact of sun glint over the ocean with low-noise-equivalent radiances in all channels. Its radiometric dynamic range is compatible to both situations of low oceanic signals in the case of clear atmosphere and higher signals in the presence of high aerosol loading (optimized for ocean color measurements) (Berger et al., 2012; ESA, 2012). OLCI's revisit cycle will be more frequent compared to OLI and MSI (every 1-2 days), although with a coarser spatial resolution of 300 m at nadir.

Multi-spectral global missions have targeted band positions. Thus, selected biophysical variables of water column can be derived from a given sensor. Hyperspectral data provide measurements in several narrow bands to highlight the specific absorption features of several water column properties (Hestir et

al., 2015). The underlined features can be used to derive the empirically calibrated band indices for the water region of interest. Also, the additional spectral information can overcome the underdetermined problems to simultaneously deriving a large number of optical properties in the radiative transfer inversion algorithms used in semi-analytical methods. The Hyperspectral Infrared Imager (HyspIRI) is NASA's mission for the coming decade. It has uniquely well designed spatial (60 m) and spectral resolutions designed to capture multiple biophysical variables including chl-a concentration in freshwater ecosystems. However, the temporal resolution of 19 days equatorial revisit cycle (shorter) can provide measurements at the seasonal time scale, but may not be able to track algal bloom dynamics (Hestir et al., 2015).

Finally, the sensitivity of the 1-D FLake model to the variations of K_d for simulating other lake properties such as water temperature profiles, ice thickness, ice on/off days should be evaluated. More complex 3-D lake models are now starting to be used to reproduce large lake properties. For example, Environment Canada has recently implemented a fully coupled 3-D atmosphere-lake modelling system to represent the complex air-lake interaction over the Great Lakes region. In this system, the hydrodynamic lake component is based on NEMO (Nucleus for European Modelling of the Ocean; Madec et al. 1998) (Dupont et al., 2012). The contribution of satellite-derived water clarity in improving simulations with more complex 3-D lake models such as NEMO is another interesting research avenue to explore in the near future.

References

Akaike, H., Parzen, E., Tanabe, K., and Kitagawa, G. 1998. Selected Papers of Hirotugu Akaike. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Springer Series in Statistics* (p. 448). Springer New York.

Ali, K. A., and Ortiz, J. D. 2014. Multivariate approach for chlorophyll-a and suspended matter retrievals in Case II type waters using hyperspectral data. *Hydrological Sciences Journal*, *6667*(February), null–null.

Alikas, K., Ansko, I., Reinart, A., Lill, E., and Valdmets, K. 2008. Testing Available Meris Image Processors for Lakes. In *the 2nd MERIS (A)ATSR User Workshop*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?rep=rep1&type=pdf&doi=10.1.1.227.2909

Anderson, K. R., Chapman, D. C., Wynne, T. T., Masagounder, K., and Paukert, C. P. 2015. Suitability of Lake Erie for bigheaded carps based on bioenergetic models and remote sensing. *Journal of Great Lakes Research*, *41*(2), 358–366.

André, J.-M. 1992. Ocean color remote-sensing and the subsurface vertical structure of phytoplankton pigments. *Deep Sea Research Part A. Oceanographic Research Papers*, *39*(5), 763–779.

Arar, J. E. 1997a. *Determination of Chlorophylls a and b and Identification of Other Pigments of Interest in Marine and Freshwater Algae Using High Performance Liquid Chromatography with Visible Wavelength Detection*. *EPA, Cincinnati. (EPA method 447.0).* Retrieved from http://greentechdev.awardspace.info/it/docs/epa/EPA-Method-447.pdf

Arar, J. E. 1997b. *Determination of chlorophylls a, b, c 1c and pheopigments in marine and freshwater algae by visible spectrophotometry*. *EPA, Cincinnati. (EPA method 446.0).*

Arar, J. E., and Collins, G. B. 1997. *Determination of Chlorophyll a and Pheophytin a in Marine and Freshwater Algae by Fluorescence*. *EPA, Cincinnati. (EPA method 445.0).*

Arst, H., Erm, A., Herlevi, A., Kutser, T., Leppäranta, M., Reinart, A., and Virta, J. 2008. Optical properties of boreal lake waters in Finland and Estonia. *Boreal Environment Research*, *13*(April), 133–158.

Attila, J., Koponen, S., Kallio, K., Lindfors, A., Kaitala, S., and Ylostalo, P. 2013. MERIS Case II water processor comparison on coastal sites of the northern Baltic Sea. *Remote Sensing of Environment*, *128*, 138–149.

Berger, M., Moreno, J., Johannessen, J. a., Levelt, P. F., and Hanssen, R. F. 2012. ESA's sentinel missions in support of Earth system science. *Remote Sensing of Environment*, *120*, 84–90.

Binding, C. E., Greenberg, T. A., and Bukata, R. P. 2011a. Time series analysis of algal blooms in Lake of the Woods using the MERIS maximum chlorophyll index. *Journal of Plankton Research*, *33*(12), 1847–1852.

Binding, C. E., Greenberg, T. A., and Bukata, R. P. 2012. An analysis of MODIS-derived algal and mineral turbidity in Lake Erie. *Journal of Great Lakes Research*, *38*(1), 107–116.

Binding, C. E., Greenberg, T. a., and Bukata, R. P. 2013. The MERIS Maximum Chlorophyll Index; its merits and limitations for inland water algal bloom monitoring. *Journal of Great Lakes Research*, *39*, 100–107.

Binding, C. E., Greenberg, T. A., Jerome, J. H., Bukata, R. P., and Letourneau, G. 2011b. An assessment of MERIS algal products during an intense bloom in Lake of the Woods. *Journal of Plankton Research*, *33*(5), 793–806.

Binding, C. E., Greenberg, T. a., Watson, S. B., Rastin, S., and Gould, J. 2015. Long term water clarity changes in North America's Great Lakes from multi-sensor satellite observations. *Limnology and Oceanography*, *60*(6), 1–20.

Binding, C. E., Jerome, J., Bukata, R., and Booty, W. 2008. Spectral absorption properties of dissolved and particulate matter in Lake Erie. *Remote Sensing of Environment*, *112*(4), 1702–1711.

Binding, C. E., Jerome, J. H., Bukata, R. P., and Booty, W. G. 2007. Trends in Water Clarity of the Lower Great Lakes from Remotely Sensed Aquatic Color. *Journal of Great Lakes Research*, *33*(4), 828–841.

Binding, C. E., Jerome, J. H., Bukata, R. P., and Booty, W. G. 2010. Suspended particulate matter in Lake Erie derived from MODIS aquatic colour imagery. *International Journal of Remote Sensing*, *31*(19), 5239–5255.

Bolsenga, S. J., and Herdendorf, C. E. 1993. *Lake Erie and Lake St. Clair Handbook*. Detroit, Michigan: Wayne State University Press.

Bootsma, H. A., and Hecky, R. E. 2003. A Comparative Introduction to the Biology and Limnology of the African Great Lakes. *Journal of Great Lakes Research, 29*(Supplement 2), 3–18.

Bresciani, M., and Giardino, C. 2012. Retrospective analysis of spatial and temporal variability of

chlorophyll-a in the Curonian Lagoon. *Journal of Coastal Conservation*, *16*(4), 511–519.

Brezonik, P. L., Olmanson, L. G., Finlay, J. C., and Bauer, M. E. 2015. Factors affecting the measurement of CDOM by remote sensing of optically complex inland waters. *Remote Sensing of Environment*, *157*, 199–215.

Bronmark, C., and Hansson, L.-A. 2005. *The Biology of Lakes and Ponds* (Second Edi). NewYork: Oxford University Press.

Bukata, R. P., Jerome, J. H., Kondratyev, A. S., and Pozdnyakov, D. V. 1995. *Optical Properties and Remote Sensing of Inland and Coastal Waters*. Boca Raton, FL.: CRC Press.

Campbell, J. W. 1995. The lognormal distribution as a model for bio-optical variability in the sea. *Journal of Geophysical Research*, *100*(C7), 13237–13254. Retrieved from http://onlinelibrary.wiley.com/doi/10.1029/95JC00458/full

Canadian Council of Ministers of the Environment. 2003. *Canadian water quality guidelines for the protection of aquatic life*. *Canadian Environmental Quality Guidelines*.

Cano, D., Monget, J. M., Albuisson, M., Guillard, H., Regas, N., and Wald, L. 1986. A method for the determination of the global solar radiation from meteorological satellite data. *Solar Energy*, *37*(1), 31–39.

Carder, K. L., Chen, F. R., Cannizzaro, J. P., Campbell, J. W., and Mitchell, B. G. 2004. Performance of the MODIS semi-analytical ocean color algorithm for chlorophyll-a. *Advances in Space Research*, *33*(7), 1152–1159.

Carder, K. L., Chen, F. R., Lee, Z. P., Hawes, S. K., and Kamykowski, D. 1999. Semianalytic Moderate-Resolution Imaging Spectromete algorithms for chlorophyll a and absorption with bio-optical domains based on nitrate-depletion temperatures, *104*, 5403–5421.

Carter, C. H. 1977. Sediment–Load Measurements along the United States Shore of Lake Erie. *Columbus Ohio: Ohio Division of Geological Survey Report of Investigations*, *102*, 24.

Chauncey, W. A. 2005. Field Measurement. In *Handbooks for Water-Resources Investigations* (Vol. 1, pp. 1–55). U.S. Geological Survey. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:U.S.+Geological+Survey+TWRI+Book#3

Chen, J., Zhang, M., Cui, T., and Wen, Z. 2013. A Review of Some Important Technical Problems in Respect of Satellite Remote Sensing of Chlorophyll-a Concentration in Coastal Waters. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1–15.

Civera, J. I., Miró, N. L., Breijo, E. G., and Peris, R. M. 2013. Secchi depth and water quality control: Measurement of sunlight extinction. In T. B. Hughes (Ed.), *Mediterranean Sea: Ecosystems, Economic Importance and Environmental Threats* (pp. 91–114). Nova Science.

Cococcioni, M., Lazzerini, B., and Marcelloni, F. 2007. Estimating the concentration of optically active constituents of sea water by Takagi–Sugeno models with quadratic rule consequents. *Pattern Recognition*, *40*(10), 2846–2860.

Daher, S. 1999. *Lake Erie LaMP Status Report*.

Dall'Olmo, G., and 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*, *44*(3), 412–422.

De Bruijn, E. I. F., Bosveld, F. C., and Van Der Plas, E. V. 2014. An intercomparison study of ice thickness models in the Netherlands. *Tellus A*, *66*, 21244–21255.

Deus, R. 2013. Three-dimensional model for analysis of spatial and temporal patterns of phytoplankton in Tucuruí reservoir, Pará, Brazil.

Dise, J., Kankiewicz, A., Schlemmer, J., Hemker, K., Kivalov, S., Hoff, T., and Perez, R. 2013. Operational improvements in the performance of the SUNY satellite-to-solar irradiance model using satellite infrared channels. *Conference Record of the IEEE Photovoltaic Specialists Conference*, 960–963.

Djoumna, G., Lamb, K. G., and Rao, Y. R. 2014. Sensitivity of the Parameterizations of Vertical Mixing and Radiative Heat Fluxes on the Seasonal Evolution of the Thermal Structure of Lake Erie. *Atmosphere-Ocean*, *52*(4), 294–313.

Doerffer, R., and Schiller, H. 2007. The MERIS Case 2 water algorithm. *International Journal of Remote Sensing*, *28*(3-4), 517–535.

Doerffer, R., and Schiller, H. 2008. *Algorithm Theoretical Basis Document (ATBD) Lake Water Algorithm for BEAM*. *GKSS*.

Doron, M., Babin, M., and Hembise, O. 2011. Ocean transparency from space: Validation of algorithms

estimating Secchi depth using MERIS, MODIS and SeaWiFS data. *Remote Sensing of Environment*, *115*(12), 2986–3001.

DosSantos, A. C. A., Calijuri, M. C., Moraes, E. M., Adorno, M. A. T., Falco, P. B., Carvalho, D. P., … Benassi, S. F. 2003. Comparison of three methods for Chlorophyll determination: Spectrophotometry and Fluorimetry in samples containing pigment mixtures and spectrophotometry in samples with separate pigments through High Performance Liquid Chromatography. *Acta Limnologica Brasiliensia*, *15*(3), 7–18. Retrieved from http://ecologia.icb.ufmg.br/~rpcoelho/Energetica_Aulas/projetos_2006/compara%E7%E3o de tres metod. para determ. de clorofila.pdf

Duan, H., Ma, R., Xu, J., Zhang, Y., and Zhang, B. 2010. Comparison of different semi-empirical algorithms to estimate chlorophyll-a concentration in inland lake water. *Environmental Monitoring and Assessment*, *170*(1), 231–44.

Dupont, F., Chittibabu, P., Fortin, V., Rao, Y. R., and Lu, Y. 2012. Assessment of a NEMO-based hydrodynamic modelling system for the Great Lakes. *Water Quality Research Journal of Canada*, *47*, 198–214.

Eerola, K., Rontu, L., Kourzeneva, E., and Shcherbak, E. 2010. A study on effects of lake temperature and ice cover in HIRLAM. *Boreal Environment Research*, *15*, 130–142. Retrieved from http://cat.inist.fr/?aModele=afficheN&cpsidt=22782352

EPA Guidance Manual. 1999. *TURBIDITY METHODS & MEASUREMENT*.

ESA, E. S. A. 2012. *Sentinel-3: ESA's Global Land and Ocean Mission for GMES Operational Services*. the Netherlands.

Fleming-Lehtinen, V., and Laamanen, M. 2012. Long-term changes in Secchi depth and the role of phytoplankton in explaining light attenuation in the Baltic Sea. *Estuarine, Coastal and Shelf Science*, *102-103*, 1–10.

Florida Departement of Protection Plan. 2011. *Applicability of Chlorophyll a Methods*.

Florida Lakewatch. 2000. *A Beginner's Guide to Water Management — Nutrients*. Gainesville, Florida.

Franz, B. a, Bailey, S. W., Kuring, N., and Werdell, P. J. 2014. Ocean Color Measurements from Landsat-8 OLI using SeaDAS. *Proceedings of the Ocean Optics XXII*, 26–31.

Giardino, C., Bresciani, M., Valentini, E., Gasperini, L., Bolpagni, R., and Brando, V. E. 2015.

Airborne hyperspectral data to assess suspended particulate matter and aquatic vegetation in a shallow and turbid lake. *Remote Sensing of Environment*, *157*, 48–57.

Giardino, C., Bresciani, M., Villa, P., and Martinelli, A. 2010. Application of Remote Sensing in Water Resource Management: The Case Study of Lake Trasimeno, Italy. *Water Resources Management*, *24*(14), 3885–3899.

Gilerson, A., Zhou, J., Hlaing, S., Ioannou, I., Amin, R., Gross, B., … Ahmed, S. 2007. Fluorescence contribution to reflectance spectra for a variety of coastal waters. *SPIE*, *6680*, 66800C–66800C–12.

Gitelson, A. a. 1992. The peak near 700 nm on radiance spetral of algae and water: relationships of its magnitude and position with chlorophyll concentration. *International Journal of Remote Sensing*, *13*(17), 3367–3373.

Gitelson, A. A., Dall'Olmo, G., Moses, W., Rundquist, D. C., Barrow, T., Fisher, T. R., … Holz, J. 2008. A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: Validation. *Remote Sensing of Environment*, *112*(9), 3582–3593.

Gobler, C. J., and Wilhelm, S. W. 2015. *Estimations and implications of plankton mortality in Lake Erie*.

Golterman, H. L. 1975. *Physiological limnology: An approach to the physiology of lake ecosystems*. Amsterdam, The Netherland: Elsevier Science Ltd.

Gons, H. J., Auer, M. T., and Effler, S. W. 2008. MERIS satellite chlorophyll mapping of oligotrophic and eutrophic waters in the Laurentian Great Lakes. *Remote Sensing of Environment*, *112*(11), 4098–4106.

Gordon, H. R., and Clark, D. K. 1980. Remote sensing optical properties of a stratified ocean: an improved interpretation. *Applied Optics*, *19*(20), 3428–3430.

Gower, J. F. R., Doerffer, R., and Borstad, G. A. 1999. Interpretation of the 685nm peak in water-leaving radiance spectra in terms of fluorescence, absorption and scattering, and its observation by MERIS. *International Journal of Remote Sensing*, *20*(9), 1771–1786.

Gower, J. F. R., and King, S. 2012. Use of satellite images of chlorophyll fluorescence to monitor the spring bloom in coastal waters. *International Journal of Remote Sensing*, *33*(23), 7469–7481.

Gower, J. F. R., King, S., Borstad, G. A., and Brown, L. 2005. Detection of intense plankton blooms

using the 709 nm band of the MERIS imaging spectrometer. *International Journal of Remote Sensing*, *26*(9), 2005–2012.

Greisberger, S., and Teubner, K. 2007. Does pigment composition reflect phytoplankton community structure in differing temperature and light conditions in a deep alpine lake? An approach using HPLC and delayed fluorescence techniques 1. *Journal of Phycology*, *43*(6), 1108–1119.

Guan, Q., Guo, Z., Liu, C., and Lei, X. 2012. A tentative study of water quality retrieval in low-level-polluted Case II waters using analytical model. *Geoscience and Remote Sensing Symposium (IGARSS)*, 805–807. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6350872

Gueymard, C., Perez, R., Schlemmer, J., Hemker, K., Kivalov, S., and Kankiewicz, A. 2015. Satellite-to-Irradiance Modeling – A New Version of the SUNY Model. *42nd IEEE PV Specialists Conf*, (JUNE).

Harris, L. a., Duarte, C. M., and Nixon, S. W. 2006. Allometric Laws and Prediction in Esturine and Costal Ecology. *Estuaries and Coasts*, *29*(2), 340–344.

Heim, B., Abramova, E., Doerffer, R., Günther, F., Hölemann, J., Kraberg, a., … Wegner, C. 2014. Ocean colour remote sensing in the southern laptev sea: Evaluation and applications. *Biogeosciences*, *11*(15), 4191–4210.

Heiskanen, J. J., Mammarella, I., Ojala, A., Stepanenko, V., Erkkilä, K.-M., Miettinen, H., … Nordbo, A. 2015. Effects of water clarity on lake stratification and lake-atmosphere heat exchange. *Journal of Geophysical Research: Atmospheres*, *120*(15), 7412–7428.

Herdendorf, C. E. 1984. *Lake Erie Water Quality 1970-1982: A Management Assessment*.

Hestir, E. L., Brando, V. E., Bresciani, M., Giardino, C., Matta, E., Villa, P., and Dekker, A. G. 2015. Measuring freshwater aquatic ecosystems: The need for a hyperspectral global mapping satellite mission. *Remote Sensing of Environment*, *167*, 181–195.

Hicks, B. J., Stichbury, G. a, Brabyn, L. K., Allan, M. G., and Ashraf, S. 2013. Hindcasting water clarity from Landsat satellite images of unmonitored shallow lakes in the Waikato region, New Zealand. *Environmental Monitoring and Assessment*.

Hinzman, L. D., Goering, D. J., and Kane, D. L. 1998. A distributed thermal model for calculating soil temperature profiles and depth of thaw in permafrost regions. *Journal of Geophysical Research*,

103(D22), 28975–28991.

Hiriart-Baer, V. P., Binding, C. E., and Howell, T. E. 2013. Dissolved organic matter quantity and quality in Lake Simcoe compared to two other large lakes in southern Ontario. *Inland Waters*, *3*(2), 139–152.

Hiriart-Baer, V. P., Greenberg, B. M., Guildford, S. J., and Smith, R. E. 2002. Effects of ultraviolet radiation on rates and size distribution of primary production by Lake Erie phytoplankton. *Canadian Journal of Fisheries and Aquatic Sciences*, *59*, 317–328.

Huszar, V. L. D. M., and Caraco, N. F. 1998. The relationship between phytoplankton composition and physical–chemical variables: A comparison of taxonomic and morphological–functional descriptors in six temperate lakes. *Freshwater Biology*, *40*, 679–696.

Hydrology Project Training Module. 2000. *How to measure Chlorophyll-a*.

IJC. International Joint Commission Canada and United States. 2013. *Lake Erie Ecosystems Priority, Scientific Findings and Policy: Recommendations to Reduce Nutrient Loadings and Harmful Algal Blooms*.

Illinois Environmental Protection Agency. 2014. *Lake Stratification and Mixing*. Illinois. Retrieved from http://www.epa.state.il.us/water/conservation/lake-notes/lake-stratification-and-mixing/lake-stratification.pdf

Ingram, R., and Young, J. 2010. *Lake Simcoe Ice-Free Sampling Manual 2010*.

Irons, J., and Masek, J. 2012. *Landsat data continuity mission*. *National Aeronautics and Space Administration-USGS*. Retrieved from http://modarch.gsfc.nasa.gov/sci_team/meetings/200610/presentations/land/irons.pdf

Kabbara, N., Benkhelil, J., Awad, M., and Barale, V. 2008. Monitoring water quality in the coastal area of Tripoli (Lebanon) using high-resolution satellite data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *63*(5), 488–495.

Kallio, K., Koponen, S., Ylöstalo, P., Kervinen, M., Pyhälahti, T., and Attila, J. 2015. Validation of MERIS spectral inversion processors using reflectance, IOP and water quality measurements in boreal lakes. *Remote Sensing of Environment*, *157*, 147–157.

Kemp, a. L. W., MacInnis, G. a., and Harper, N. S. 1977. Sedimentation Rates and a Revised Sediment Budget for Lake Erie. *Journal of Great Lakes Research*, *3*(3-4), 221–233.

Kheyrollah Pour, H., Duguay, C. R., Martynov, A., and Brown, L. C. 2012. Simulation of surface temperature and ice cover of large northern lakes with 1-D models: A comparison with MODIS satellite data and in situ measurements. *Tellus, Series A: Dynamic Meteorology and Oceanography*, *64*(1), 17614–17633.

Kheyrollah Pour, H., Duguay, C., Solberg, R., and Rudjord, Ø. 2014a. Impact of satellite-based lake surface observations on the initial state of HIRLAM. Part I: evaluation of remotely-sensed lake surface water temperature observations. *Tellus A*, *66*, 21534–21546.

Kheyrollah Pour, H., Rontu, L., and Duguay, C. 2014b. Impact of satellite-based lake surface observations on the initial state of HIRLAM. Part II: Analysis of lake surface temperature and ice cover. *Tellus A*, *66*, 21395–21413.

Kiefer, I., Odermatt, D., Anneville, O., Wüest, A., and Bouffard, D. 2015. Application of remote sensing for the optimization of in-situ sampling for monitoring of phytoplankton abundance in a large lake. *Science of The Total Environment*, *527-528*, 493–506.

Kleissl, J., Perez, R., Cebecauer, T., and Šúri, M. 2013. *Solar Energy Forecasting and Resource Assessment*. *Solar Energy Forecasting and Resource Assessment*. Elsevier.

Koenings, J. P., and Edmundson, J. a. 1991. Secchi disk and photometer estimates of light regimes in Alaskan lakes: Effects of yellow color and turbidity. *Limnology and Oceanography*, *36*(1), 91–105.

Kondratyev, K. Y. 1999. *Limnology and Remote Sensing: A Contemporary Approach*.

Kourzeneva, E. 2010. External data for lake parameterization in Numerical Weather Prediction and climate modeling. *Boreal Environment Research*, *15*(2), 165–177.

Kourzeneva, E., Asensio, H., Martin, E., and Faroux, S. 2012a. Global gridded dataset of lake coverage and lake depth for use in numerical weather prediction and climate modelling. *Tellus Series a-Dynamic Meteorology and Oceanography*, *64*, 15640–15654.

Kourzeneva, E., Martin, E., Batrak, Y., and Moigne, P. Le. 2012b. Climate data for parameterisation of lakes in Numerical Weather Prediction models. *Tellus A*, *64*, 17226–17243.

Kratzer, S., Brockmann, C., and Moore, G. 2008. Using MERIS full resolution data to monitor coastal waters — A case study from Himmerfjärden, a fjord-like bay in the northwestern Baltic Sea. *Remote Sensing of Environment*, *112*(5), 2284–2300.

Kratzer, S., Håkansson, B., and Sahlin, C. 2003. Assessing Secchi and Photic Zone Depth in the Baltic Sea from Satellite Data. *Remote Sensing of Environment*, *32*(8), 577–585. Retrieved from http://www.jstor.org

Kudela, R. M., Palacios, S. L., Austerberry, D. C., Accorsi, E. K., Guild, L. S., and Torres-Perez, J. 2015. Application of hyperspectral remote sensing to cyanobacterial blooms in inland waters. *Remote Sensing of Environment*, *167*, 196–205.

Kutser, T. 2012. The possibility of using the Landsat image archive for monitoring long time trends in coloured dissolved organic matter concentration in lake waters. *Remote Sensing of Environment*, *123*, 334–338.

Lake Erie LaMP Work Group. 2000. *Lake Erie LaMP Status Report*.

Lake Erie LaMP Work Group. 2004. *Lake Erie LaMP Status Report*.

Lake Erie Nutrient Science Task Group. 2009. *Status of Nutrients in the Lake Erie Basin*. *Lake Erie Millennim Network 5th Biennial Conference*. Retrieved from http://www.epa.gov/greatlakes/lakeerie/erie_nutrient_2010.pdf

Le, C., Hu, C., Cannizzaro, J., English, D., Muller-Karger, F., and Lee, Z. 2013. Evaluation of chlorophyll-a remote sensing algorithms for an optically complex estuary. *Remote Sensing of Environment*, *129*, 75–89.

Lee, Z., Carder, K. L., Hawes, S. K., Steward, R. G., Peacock, T. G., and Davis, C. O. 1994. Model for the interpretation of hyperspectral remote-sensing reflectance. *Applied Optics*, *33*(24), 5721–32. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20935974

Lesht, B. M., Barbiero, R. P., and Warren, G. J. 2013. A band-ratio algorithm for retrieving open-lake chlorophyll values from satellite observations of the Great Lakes. *Journal of Great Lakes Research*, *39*(1), 138–152.

Liu, H., Xu, L., Ding, J., Zhuoma, B., Deng, X., and Liu, Z. 2010. Atmospheric correction and land surface temperature retrieval method for FY-3 IR observations. *2010 Second IITA International Conference on Geoscience and Remote Sensing*, 131–134.

López, R., Del Castillo, C. E., Miller, R. L., Salisbury, J., and Wisser, D. 2012. Examining organic carbon transport by the Orinoco River using SeaWiFS imagery. *Journal of Geophysical Research*, *117*(G3), G03022.

Madec, G., Delecluse, P., Imbard, M., and Lévy, C. 1998. *Opa 8 ocean general circulation model - reference manual*. *Tech. rep.*

Martynov, A., Sushama, L., and Laprise, R. 2010. Simulation of temperate freezing lakes by one-dimensional lake models: Performance assessment for interactive coupling with regional climate models. *Boreal Environment Research*, *15*(2), 143–164.

Martynov, A., Sushama, L., Laprise, R., Winger, K., and Dugas, B. 2012. Interactive lakes in the Canadian Regional Climate Model, version 5: The role of lakes in the regional climate of North America. *Tellus, Series A: Dynamic Meteorology and Oceanography*, *64*(1), 16226–16248.

Marvin, C. H., Charlton, M. N., Reiner, E. J., Kolic, T., MacPherson, K., Stern, G. a., … Painter, S. 2002. Surficial Sediment Contamination in Lakes Erie and Ontario: A Comparative Analysis. *Journal of Great Lakes Research*, *28*(3), 437–450.

Matsuoka, a., Hooker, S. B., Bricaud, a., Gentili, B., and Babin, M. 2013. Estimating absorption coefficients of colored dissolved organic matter (CDOM) using a semi-analytical algorithm for southern Beaufort Sea waters: application to deriving concentrations of dissolved organic carbon from space. *Biogeosciences*, *10*(2), 917–927.

Matthews, M. W., Bernard, S., and Winter, K. 2010. Remote sensing of cyanobacteria-dominant algal blooms and water quality parameters in Zeekoevlei, a small hypertrophic lake, using MERIS. *Remote Sensing of Environment*, *114*(9), 2070–2087.

McCullough, I. M., Loftin, C. S., and Sader, S. a. 2012a. Combining lake and watershed characteristics with Landsat TM data for remote estimation of regional lake clarity. *Remote Sensing of Environment*, *123*, 109–115.

McCullough, I. M., Loftin, C. S., and Sader, S. a. 2012b. High-frequency remote monitoring of large lakes with MODIS 500m imagery. *Remote Sensing of Environment*, *124*, 234–241.

Michalak, A. M., Anderson, E. J., Beletsky, D., Boland, S., Bosch, N. S., Bridgeman, T. B., … Zagorski, M. A. 2013. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proceedings of the National Academy of Sciences*, *110*(16), 6448–6452.

Mironov, D. 2005. *Parameterization of lakes in numerical weather prediction. Part 1: Description of a lake model*. *German Weather Service*. Germany. Retrieved from http://cosmo-

model.cscs.ch/content/model/modules/flake/docs/ParLak_Part1_a.pdf

Mironov, D. 2008. Parameterization of Lakes in Numerical Weather Prediction. Description of a Lake Model. *Reports - COSMO*, 47.

Mironov, D., Heise, E., Kourzeneva, E., Ritter, B., and Schneider, N. 2010. Implementation of the lake parameterisation scheme FLake into the numerical weather prediction model COSMO. *Boreal Environment Research*, *15*, 218–230. Retrieved from http://www.borenv.net/BER/pdfs/ber15/ber15-218.pdf

Mironov, D., Ritter, B., Schulz, J.-P., Buchhold, M., Lange, M., and Machulskaya, E. 2012. Parameterisation of sea and lake ice in numerical weather prediction models of the German Weather Service. *Tellus A*, *64*, 17330–17346.

Moore, T. S., Campbell, J. W., and Feng, H. F. H. 2001. A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(8), 1764–1776.

Moore, T. S., Dowell, M. D., Bradt, S., and Ruiz-Verdu, A. 2014. An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters. *Remote Sensing of Environment*, *143*, 97–111.

Morang, A., Mohr, M. C., and Forgette, C. M. 2011. Longshore Sediment Movement and Supply along the U.S. Shoreline of Lake Erie. *Journal of Coastal Research*, *27*(4), 619–635.

Morel, A. 1980. In-water and remote measurements of ocean color. *Boundary-Layer Meteorology*, *18*(2), 177–201.

Morel, A., and Gordon, H. R. 1980. Report of the working group on water color. *Boundary-Layer Meteorology*, *18*(3), 343–355.

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

Moses, W. J., Gitelson, a a, Berdnikov, S., and Povazhnyy, V. 2009. Estimation of chlorophyll- a concentration in case II waters using MODIS and MERIS data—successes and challenges. *Environmental Research Letters*, *4*(4), 1–8.

Mueller, J. L., Bidigare, R. R., Trees, C., Balch, W. M., Dore, J., Drapeau, D. T., … Perl, J. 2003. Ocean Optics Protocols For Satellite Ocean Color Sensor Validation, Revision 5, Volume V:

Biogeochemical and Bio-Optical Measurements and Data Analysis Protocols. In J. L. Mueller, G. S. Fargion, & C. R. Mcclain (Eds.), *Ocean Optics Protocols For Satellite Ocean Color Sensor Validation* (Vol. V).

Munawar, M., and Weisse, T. 1989. Is the "microbial loop" an early warning indicator of anthropogenic stress? *Hydrobiologia*, *188-189*(1), 163–174.

NOAA. 2015. *Experimental Lake Erie Harmful Algal Bloom Bulletin*.

Nottrott, A., and Kleissl, J. 2010. Validation of the NSRDB–SUNY global horizontal irradiance in California. *Solar Energy*, *84*(10), 1816–1827.

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. a., Carder, K. L., Garver, S. a., … McClain, C. 1998. Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research*, *103*(C11), 24937.

O'Reilly, J. E., Maritorena, S., O'Brien, M., Siegel, D., Toole, D., Menzies, D., … Culver, M. 2000. SeaWiFS Postlaunch Technical Report Series. *Nasa*, *11*, 1–49. Retrieved from http://www.csa.com/partners/viewrecord.php?requester=gs&collection=TRD&recid=N0225155AH

Odermatt, D., Giardino, C., and Heege, T. 2010. Chlorophyll retrieval with MERIS Case-2-Regional in perialpine lakes. *Remote Sensing of Environment*, *114*(3), 607–617.

Odermatt, D., Gitelson, A. A., Brando, V. E., and Schaepman, M. 2012a. Review of constituent retrieval in optically deep and complex waters from satellite imagery. *Remote Sensing of Environment*, *118*, 116–126.

Odermatt, D., Pomati, F., Pitarch, J., Carpenter, J., Kawka, M., Schaepman, M., and Wüest, A. 2012b. MERIS observations of phytoplankton blooms in a stratified eutrophic lake. *Remote Sensing of Environment*, *126*, 232–239.

Olmanson, L., Brezonik, P., and Bauer, M. 2013. hyperspectral remote sensing to assess spatial distribution of water quality characteristics in large rivers: The Mississippi River and its tributaries in Minnesota. *Remote Sensing of Environment*, *130*, 254–265.

Ortiz, J. D., Witter, D. L., Ali, K. A., Fela, N., Duff, M., and Mills, L. 2013. Evaluating multiple colour-producing agents in Case II waters from Lake Erie. *International Journal of Remote Sensing*, *34*(24), 8854–8880.

Pahlevan, N., and Schott, J. R. 2013. Leveraging EO-1 to evaluate capability of new generation of landsat sensors for coastal/inland water studies. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *6*(2), 360–374.

Painter, S., Myers, D. N., and Letterhos, J. 2000. *Lake Erie Lakewide Management Plan (LaMP) Technical Report Series, Characterization of Data and Data Collection Programs for Assessing Pollutants of Concern to Lake Erie*.

Palmer, S. C. J., Hunter, P. D., Lankester, T., Hubbard, S., Spyrakos, E., N. Tyler, A., … Tóth, V. R. 2014. Validation of Envisat MERIS algorithms for chlorophyll retrieval in a large, turbid and optically-complex shallow lake. *Remote Sensing of Environment*, *157*, 158–169.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., and Vignola, F. 2002. A new operational model for satellite-derived irradiances: Description and validation. *Solar Energy*, *73*(5), 307–317.

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K. J., and Zelenka, A. 2010. Improving the performance of satellite-to-irradiance models using the satellite's infrared sensors. In *American Solar Energy Society Annual Conference*. Retrieved from http://www.asrc.albany.edu/people/faculty/perez/2010/ir.pdf

Peters, N. E., Krysanova, V., Lepisto, A., Prasad, R., Thoms, M., Wilby, R., and Zandaryaa, S. 2011. *Water Quality: Current Trends and Expected Climate Change Impacts (186)*. Wallingford, United Kingdom: IAHS Press. Retrieved from https://e-publications.une.edu.au/vital/access/manager/Repository/une:8538

Petrusevics, P., Bye, J., Luick, J., and Teixeira, C. E. P. 2011. Summer sea surface temperature fronts and elevated chlorophyll-a in the entrance to Spencer Gulf, South Australia. *Continental Shelf Research*, *31*(7-8), 849–856.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team. 2015. {nlme}: Linear and Nonlinear Mixed Effects Models. Retrieved from http://cran.r-project.org/package=nlme

Poole, H. H., and Atkins, W. R. G. 1929. Photo-electric measurements of submarine illumination throughout the year.

Potes, M., Costa, M. J., and Salgado, R. 2012. Satellite remote sensing of water turbidity in Alqueva reservoir and implications on lake modelling. *Hydrology and Earth System Sciences*, *16*(6), 1623–

1633.

R Core Team. 2015. R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from https://www.r-project.org/

Randolph, K., Wilson, J., Tedesco, L., Li, L., Pascual, D. L., and Soyeux, E. 2008. Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll a and phycocyanin. *Remote Sensing of Environment*, *112*(11), 4009–4019.

Rinke, K., Yeates, P., and Rothhaupt, K. O. 2010. A simulation study of the feedback of phytoplankton on thermal structure via light extinction. *Freshwater Biology*, *55*(8), 1674–1693.

Robinson, I. S. 2004. *Measuring the Oceans from Space: The principles and methods of satellite oceanography*. Springer.

Ruescas, A., Brockmann, C., Stelzer, K., Tilstone, G. H., and Beltrán-Abaunza, J. M. 2014. *DUE Coastcolour Final Report , version 1*.

Ruiz-Verdú, A., and Koponen, S. 2008. Development of MERIS lake water algorithms: validation results from Europe. In *Proceeding of the '2nd MERIS / (A)ATSR User Workshop*. Frascati, Italy. Retrieved from https://earth.esa.int/pub/ESA_DOC/meris_workshop_2008/papers /o73_ruiz.pdf

Sá, C., D'Alimonte, D., Brito, A. C., Kajiyama, T., Mendes, C. R., Vitorino, J., … Brotas, V. 2015. Validation of standard and alternative satellite ocean-color chlorophyll products off Western Iberia. *Remote Sensing of Environment*, *168*, 403–419.

Salama, M. S., Radwan, M., and van der Velde, R. 2012. A hydro-optical model for deriving water quality variables from satellite images (HydroSat): A case study of the Nile River demonstrating the future Sentinel-2 capabilities. *Physics and Chemistry of the Earth, Parts A/B/C*, *50-52*, 224–232.

Salama, M. S., and Verhoef, W. 2015. Two-stream remote sensing model for water quality mapping: 2SeaColor. *Remote Sensing of Environment*, *157*, 111–122.

Salinas, J. T. 1988. *A CRITICAL COMPARISON OF METHODS FOR THE DETERMINATION OF PHYTOPLANKION CHLOROPHYLL*. Oregonstate University.

Samuelsson, P., Kourzeneva, E., and Mironov, D. 2010. The impact of lakes on the European climate as simulated by a regional climate model. *Boreal Environment Research*, *15*, 113–129. Retrieved from http://cat.inist.fr/?aModele=afficheN&cpsidt=22782351

Saulquin, B., Hamdi, A., Gohin, F., Populus, J., Mangin, A., and D'Andon, O. F. 2013. Estimation of the diffuse attenuation coefficient K dPAR using MERIS and application to seabed habitat mapping. *Remote Sensing of Environment*, *128*, 224–233.

Schroeder, T., Schaale, M., and Fischer, J. 2007. Retrieval of atmospheric and oceanic properties from MERIS measurements: A new Case-2 water processor for BEAM. *International Journal of Remote Sensing*, *28*(24), 5627–5632.

Schwarz, G. 1978. Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.

Shaw, G. R., Moore, D. P., and Garnet, C. 2009. *Environmental and Ecological Chemistry, Volume 2*. (A. Sabljic, Ed.). EOLSS Publishers Company Limited.

Shen, F., Zhou, Y.-X., Li, D.-J., Zhu, W.-J., and Salama, M. S. 2010. Medium Resolution Imaging Spectrometer (MERIS) Estimation of Chlorophyll-a Concentration in the Turbid Sediment-laden Waters of the Changjiang (Yangtze) Estuary. *International Journal of Remote Sensing*, *31*(17-18), 4635–4650.

Shuchman, R. a., Leshkevich, G., Sayers, M. J., Johengen, T. H., Brooks, C. N., and Pozdnyakov, D. 2013. An algorithm to retrieve chlorophyll, dissolved organic carbon, and suspended minerals from Great Lakes satellite data. *Journal of Great Lakes Research*, *39*, 14–33.

Smith, R. E. H., Furgal, J. A., Charlton, M. N., Greenberg, B. M., Hiriart-Baer, V. P., and Marwood, C. 1999. Attenuation of ultraviolet radiation in a large lake with low dissolved organic matter concentrations. *Canadian Journal of Fisheries and Aquatic Sciences*, *56*(8), 1351–1361.

Sokoletsky, L. G., and Yacobi, Y. Z. 2011. Comparison of chlorophyll a concentration detected by remote sensors and other chlorophyll indices in inhomogeneous turbid waters. *Applied Optics*.

Song, K., Li, L., Wang, Z., Liu, D., Zhang, B., Xu, J., … Wang, Y. 2012. Retrieval of total suspended matter (TSM) and chlorophyll-a (Chl-a) concentration from remote-sensing data for drinking water resources. *Environmental Monitoring and Assessment*, *184*(3), 1449–70.

SPSS. 2001. *The SPSS TwoStep Cluster Component: A scalable component enabling more efficient customer segmentation*. SPSS Inc. Chicago. Retrieved from http://www.spss.ch/upload/1122644952_The SPSS TwoStep Cluster Component.pdf

Ston, J., Kosakowska, A., and Lotocka, M. 2002. Pigment composition in relation to phytoplankton community structure and nutrient content in the Baltic Sea. *Oceanologia*, *44*(4), 419–437.

Retrieved from http://agro.icm.edu.pl/agro/element/bwmeta1.element.agro-article-190160a5-a0a9-4092-8f28-dda8582dec6f

Stumpf, R. P., Wynne, T. T., Baker, D. B., and Fahnenstiel, G. L. 2012. Interannual variability of cyanobacterial blooms in Lake Erie. *PloS One*, *7*(8), e42444.

Sugumaran, R., and Thomas, J. 2012. A Semi-Analytical Model for the Multitemporal Prediction of Chlorophyll-a in an Iowa Lake Using Hyperion Data. *Photogrammetric Engineering and Remote Sensing*, *78*(12), 7812. Retrieved from http://cat.inist.fr/?aModele=afficheN&cpsidt=26720928

Szeto, M., Werdell, P. J., Moore, T. S., and Campbell, J. W. 2011. Are the world's oceans optically different? *Journal of Geophysical Research*, *116*(C7).

Tebbs, E. J., Remedios, J. J., and Harper, D. M. 2013. Remote sensing of chlorophyll-a as a measure of cyanobacterial biomass in Lake Bogoria, a hypertrophic, saline–alkaline, flamingo lake, using Landsat ETM+. *Remote Sensing of Environment*, *135*, 92–106.

Thayapurath, S., and Talaulikar, M. 2012. Comparison of measured and satellite-derived spectral diffuse attenuation coefficients for the Arabian Sea. *International Journal of Remote Sensing*, *33*(2), 570–585. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/01431161.2010.543435

Thiery, W., Martynov, A., Darchambeau, F., Descy, J. P., Plisnier, P. D., Sushama, L., and Van Lipzig, N. P. M. 2014. Understanding the performance of the FLake model over two African Great Lakes. *Geoscientific Model Development*, *7*(1), 317–337.

Tian, Y., Yu, Q., and Zhu, W. 2012. Estimating of chromophoric dissolved organic matter (CDOM) with in-situ and satellite hyperspectral remote sensing technology. In *IGARSS* (pp. 2040–2042). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6350975

Tilstone, G. H., Peters, S. W. M., van der Woerd, H. J., Eleveld, M. A., Ruddick, K., Schönfeld, W., … Shutler, J. D. 2012. Variability in specific-absorption properties and their use in a semi-analytical ocean colour algorithm for MERIS in North Sea and Western English Channel Coastal Waters. *Remote Sensing of Environment*, *118*, 320–338.

UNESCO. 1966. Determination of photosynthetic pigments in sea-water. *Monographs on Oceanographic Methodology*, *1*, 1–69.

University of South Florida. 2010. *Standard Operating Procedure for: Spectrophotometric*

Determination of Corrected and Uncorrected Chlorophyll-a and Pheophytin.

US epa. 2003. *Analytical Method for Turbidity Measurement*.

Vaiciute, D. 2012. Using MERIS/Envisat data to assess the Secchi depth—A case study from Lithuanian Baltic Sea waters. *Baltic International Symposium (BALTIC), 2012 IEEE/OES*, 1–5. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6249175

Vignola, F., Harlan, P., Perez, R., and Kmiecik, M. 2007. Analysis of satellite derived beam and global solar radiation data. *Solar Energy*, *81*(6), 768–772.

Vincent, R. K., Qin, X., McKay, R. M. L., Miner, J., Czajkowski, K., Savino, J., and Bridgeman, T. 2004. Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. *Remote Sensing of Environment*, *89*(3), 381–392.

Vincent, W. F. 2008. *Polar lakes and rivers limnology of Arctic and Antarctic aquatic ecosystems*.

Wang, L., Zhao, D.-Z., Yang, J., and Chen, Y. 2012. Retrieval of total suspended matter from MODIS 250 m imagery in the Bohai Sea of China. *Journal of Oceanography*, *68*(5), 719–725.

Wang, M., Shi, W., and Tang, J. 2011. Water property monitoring and assessment for China's inland Lake Taihu from MODIS-Aqua measurements. *Remote Sensing of Environment*, *115*(3), 841–854.

Watanabe, F., Alcântara, E., Rodrigues, T., Imai, N., Barbosa, C., and Rotta, L. 2015. Estimation of Chlorophyll-a Concentration and the Trophic State of the Barra Bonita Hydroelectric Reservoir Using OLI/Landsat-8 Images. *International Journal of Environmental Research and Public Health*, *12*(9), 10391–10417.

Water Quality Task Group. 2006. *A Canada-wide Framework for Water Quality Monitoring*.

Wetzel, R. G. 2001. *Limnology: Lake and River Ecosystems* (3rd ed.). Academic Press.

Wilcox, S. 2012. *National Solar Radiation Database 1991 – 2010 Update : User ' s Manual*. Retrieved from http://www.nrel.gov/docs/fy12osti/54824.pdf

Willmott, C. J. 1981. On the validation of models. *Physical Geography*, *2*(2), 184–194.

Willmott, C. J., and Matsuura, K. 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*(1), 79–82.

Willmott, C. J., Robeson, S. M., and Matsuura, K. 2012. A refined index of model performance. *International Journal of Climatology*, *32*(13), 2088–2094.

Willmott, C. J., and Wicks, D. E. 1980. An Empirical Method for the Spatial Interpolation of Monthly Precipitation within California. *Physical Geography*, *1*(1), 59–73.

Winder, M., and Sommer, U. 2012. Phytoplankton response to a changing climate. *Hydrobiologia*, *698*(1), 5–16.

Witter, D. L., Ortiz, J. D., Palm, S., Heath, R. T., and Budd, J. W. 2009. Assessing the application of SeaWiFS ocean color algorithms to Lake Erie. *Journal of Great Lakes Research*, *35*(3), 361–370.

Wu, G., Leeuw, J. De, and Liu, Y. 2009. Understanding Seasonal Water Clarity Dynamics of Lake Dahuchi from In Situ and Remote Sensing Data. *Water Resources Management*, *23*(9), 1849–1861.

Wu, G., Leeuw, J. De, Skidmore, A. K., Prins, H. H. T., and Liu, Y. 2008. Comparison of MODIS and Landsat TM5 images for mapping tempo–spatial dynamics of Secchi disk depths in Poyang Lake National Nature Reserve, China. *International Journal of Remote Sensing*, *29*(8), 2183–2198. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/01431160701422254

Zaneveld, J. R., Barnard, A., and Boss, E. 2005. Theoretical derivation of the depth average of remotely sensed optical parameters. *Optics Express*, *13*(22), 9052–9061.

Zelenka, A., Perez, R., Seals, R., and Renné, D. 1999. Effective accuracy of satellite-derived hourly irradiances. *Theoretical and Applied Climatology*, *62*, 199–207.

Zhao, D., Cai, Y., Jiang, H., Xu, D., Zhang, W., and An, S. 2011. Estimation of water clarity in Taihu Lake and surrounding rivers using Landsat imagery. *Advances in Water Resources*, *34*(2), 165–173.

Zhu, J., Hu, J., and Liu, Z. 2013. On summer stratification and tidal mixing in the Taiwan Strait. *Frontiers of Earth Science*, 1–10.