

An Integrated Water Quality Modeling System with Dynamic Remote
Sensing Feedback

by

Yan Li

B.S. University of Science and Technology of China, 1997

M.S. Rochester Institute of Technology, 2001

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology

July 26, 2007

Signature of the Author _____

Accepted by _____
Coordinator, Ph.D. Degree Program Date

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

Ph.D. DEGREE DISSERTATION

The Ph.D. Degree Dissertation of Yan Li
has been examined and approved by the
dissertation committee as satisfactory for the
dissertation required for the
Ph.D. degree in Imaging Science

Dr. Anthony Vodacek, dissertation Advisor

Dr. Alfred Garrett

Dr. Carl Salvaggio

Dr. Frank Sciremammano

Date

DISSERTATION RELEASE PERMISSION
ROCHESTER INSTITUTE OF TECHNOLOGY
CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE

Title of Dissertation:

**An Integrated Water Quality Modeling System with Dynamic Remote Sensing
Feedback**

I, Yan Li, hereby grant permission to Wallace Memorial Library of R.I.T. to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or profit.

Signature _____ Date _____

An Integrated Water Quality Modeling System with Dynamic Remote Sensing Feedback

by
Yan Li

Submitted to the
Chester F. Carlson Center for Imaging Science
in partial fulfillment of the requirements
for the Doctor of Philosophy Degree
at the Rochester Institute of Technology

Abstract

A coupled hydrodynamic-optical water quality modeling system based on Dynamic Data Driven Applications Systems (DDDAS) concepts that assimilates remote sensing data into a hydrodynamic model was developed and tested. The modeling system includes the hydrodynamic model (ALGE), a radiative transfer model (Hydro-light), and remote imagery (MODIS) as a dynamic feedback. The DDDAS was implemented through an Ensemble Kalman Filter (EnKF) with a small ensemble space.

Large scale thermal structure and circulation patterns in Lake Ontario were simulated during the spring and summer seasons. High-resolution stream plume studies were performed in Conesus Lake and for the plume of the Niagara River in Lake Ontario. This work provided validation of the capabilities of the ALGE code to simulate the transport of sediment and passive tracer.

Although the ALGE model produces predictions of the distribution of the TSS constituents, visual examination of MODIS 250 m reflectance data clearly shows discrepancies between the model TSS output and the remote sensing data. These errors are due to the uncertainties in model physics, parameters, and forcing conditions. A Kalman

filter-based method was implemented in this research to provide a better estimate of the modeled TSS. MODIS 250 m reflectance data was used as a dynamic feedback in EnKF. A test was performed at the single simulation grid point at the Genesee River mouth to validate the performance of the EnKF method. The EnKF estimate and the ensemble mean had similar and lower RMSE than any single run. Further validation was undertaken to examine the effects of assimilating MODIS data for all grid points to estimate the plume dissipation. Results show that the spatial filtering via an EnKF is capable of capturing the episodic nature of storm events by using MODIS data as feedback. In this case the EnKF estimate RMSE is considerably smaller than the ensemble mean RMSE.

Acknowledgements

I would like to express my gratitude to all those who gave me the possibility to complete this thesis. Firstly, I would like to thank my Ph.D. advisor, Dr. Anthony Vodacek, for his generous advice, guidance and inspiration from the start to the very end. It has been a great pleasure to have Dr. Vodacek as my advisor. My thanks also go to the members of my committee for their guidance and review. In particular, I thank Dr. Alfred Garrett for providing and improving ALGE model and for always finding the time to help me and share his expertise in hydrodynamic modeling and meteorology. I would also like to thank Dr. Carl Salvaggio for his insightful comments and advice along the way. Thanks also go to Dr. Frank Sciremammano for his meaningful comments and being there at my committee meetings and defense.

I would also like to thank all the rest of the academic and support staff of the Center for Imaging Science at RIT.

Especially, I would like to give my special thanks to my husband Marvin for his support and patient love that enabled me to complete this work.

This work was funded under NSF grant CNS-0324989, USDA grant 2002-51130-01459, and NOAA grant NA06OAR4600217.

This work is dedicated to my parents and my husband, Marvin, for their love and support.

Contents

1	Introduction	22
1.1	Problem statement	22
1.2	Objectives	24
1.3	Overview of the study approach	26
2	Background	29
2.1	Lake hydrodynamics: physical processes	29
2.1.1	Water and heat budget	29
2.1.2	Thermal stratification and mixing	32
2.1.3	Thermal bar	33
2.1.4	Circulation structure of lakes	36
2.1.5	River plume effects on the lake hydrodynamics	36
2.2	Optical properties of natural water	38
2.2.1	Components of natural water	39
2.2.2	Spectral response of water constituents	40
2.2.3	Impacts of PSD on IOPs	43
2.3	Contributions to sensor reaching radiance over natural water	46
2.4	Data assimilation concepts and methods	49

2.4.1	The Kalman Filter	51
2.4.2	Prediction-correction state estimator	53
2.4.3	The EKF and EnKF	58
2.5	Study sites	59
2.5.1	Conesus Lake	59
2.5.2	Lake Ontario	61
2.5.3	Niagara River mouth	62
2.5.4	Rochester Embayment	63
3	Model Review	67
3.1	ALGE: Hydrodynamic model	67
3.1.1	Governing equations	67
3.1.2	Numerical scheme	73
3.1.3	Inputs	76
3.1.4	Model calibration	79
3.2	HYDROLIGHT: Radiative transfer model	79
3.2.1	Physical and mathematical design	80
3.2.2	Simulation parameters	82
4	Data Sources and Quality Control	83
4.1	Bathymetric grids	83
4.2	Adjustment of forcing functions	86
4.3	<i>In situ</i> optical data collection and water sampling in the Rochester Em- bayment	90
4.4	Use of remote sensing data	95
4.4.1	Atmospheric correction errors for MODIS	101

5	Ensemble Kalman Filter Implementation	105
5.1	System design	105
5.2	Generation of ensemble forcings	109
5.3	Generation of ensemble observations	115
5.3.1	MODIS data for data assimilation	115
5.3.2	MODIS observation error statistics	116
5.4	Open-loop simulations	119
6	Results and Discussion	120
6.1	Grid independence study and validation of ALGE using airborne thermal imagery	120
6.2	Thermal structure in Lake Ontario	125
6.3	Numerical simulations of large scale circulations	132
6.3.1	Conesus Lake	132
6.3.2	Lake Ontario	134
6.4	Numerical simulations of river plumes	140
6.4.1	Stream plumes in Conesus Lake	140
6.4.2	Niagara River plume in Lake Ontario	146
6.5	Impacts of a turbid plume on optical dynamics in Rochester Embayment	149
6.6	Updating ALGE with MODIS data via an EnKF	153
6.6.1	Grid point test at the Genesee River mouth	155
6.6.2	Spatial update of plume dissipation via 2D filtering	161
7	Conclusions and Recommendations	168
7.1	Conclusions	168
7.2	Recommendations	172

List of Figures

2.1	A schematic view of all the major energy inputs to a typical lake. (http://ceprofs.tamu.edu/kchang/ocen689/ocen689ch9.pdf)	31
2.2	Overview of the seasonal cycle in a dimictic lake. A density profile and the relationship between temperatures in the upper and lower regions are shown. (http://ceprofs.tamu.edu/kchang/ocen689/ocen689ch9.pdf)	33
2.3	Sampling locations and survey dates of Rodgers' field survey of Lake Ontario (1970) (Rodgers, 1971).	35
2.4	One temperature profile in Lake Ontario from Rodgers' field survey (1965) (Rodgers, 1971).	35
2.5	An aerial view showing the sediment plume from a river to a lake. . .	37
2.6	Absorption and scattering cross section spectra for chl-a, total sus- pended mineral, DOC, and water for Lake Ontario (Bukata et al., 1995).	42
2.7	Mineral specific absorption spectra used for the simulations in Hydro- light (Mobley and Sundman, 2000a).	43
2.8	Mineral specific scattering spectra used for the simulations in Hydro- light (Mobley and Sundman, 2000a).	44
2.9	Bottom irradiance reflectance spectra of different types of sea floors used for the simulations in Hydrolight (Mobley and Sundman, 2000a).	44

2.10 (a) A typical log-normal PSD plotted with respect to the radius (microns). (b) The same PSD as in (a) plotted with respect to log(radius) (microns).	45
2.11 Absorption and scattering coefficients calculated for calcite with a refractive index of 1.486.	46
2.12 Components of the radiance collecting by a remote sensor over a natural water body (Bukata et al., 1995)	48
2.13 The organization of the computations in KF assimilation algorithm. .	56
2.14 Bathymetry map (m) of Conesus Lake with stream gullies (Forest et al., 1978). The southern portion is deeper than the northern portion. Conesus Inlet, North McMillan Creek, Sutton Point Gully, Cottonwood Gully, Long Point Gully, Sand Point Gully, Greywood Gully, and McPherson Point were the eight tributary mass sources. Conesus Outlet was balanced to the mass sources.	60
2.15 Lake Ontario bathymetry (Chang, 2003).	62
2.16 Nearfield horizontal distribution of Niagara River Plume on June 23, 1982 (Murthy and Miners, 1992).	64
2.17 Farfield Lagrangian drifter trajectories from October 15 to November 20, 1984 showing the spatial extent of the Niagara River Plume (Murthy and Miners, 1992).	64
2.18 Turbid Genesee River plume flowing into Lake Ontario. Locations of Ontario beach and the USGS water quality site are shown as well. . .	66
3.1 Radiative transfer model regimes: air, air-water interface, water, water-air interface, and air again (Fairbanks, 1999).	80

4.1	Bathymetric maps developed by the Grid Translator and Hydroplot of GEODAS (GEOphysical DATA System). Depth (below lake level) and elevation (above sea level) are in meters. (a) Lake Ontario (b) Niagara River mouth (c) Genesee River mouth	85
4.2	Time series of simulation forcing factor conditions of April, May, and June, 2004 from the stream monitoring program and weather stations. A major storm event is shown in late May lasting seven days with a peak flow rate for North McMillan creek of about $100 \text{ m}^3/\text{s}$	88
4.3	Time series of simulation forcing factor conditions of April, May, June, and July, 2004 from the Buffalo weather station.	89
4.4	Overview map of the four stations. Background is an image of ALGE predicted TSS distribution showing the Genesee River plume flowing eastward.	91
4.5	Measured remote sensing reflectance at three stations for the August 9th collections.	92
4.6	Specific absorption coefficient for CHL and TSS from measurements taken on August 9, 2006 in the Rochester Embayment.	93
4.7	Specific scattering coefficient for CHL and TSS. Data is based on Bukata's published data (Bukata et al., 1995).	94
4.8	Absorption and scattering coefficients of pure water (Smith and Baker, 1981).	94
4.9	Absorption coefficient for CDOM from measurements taken on August 9, 2006 in the Rochester Embayment. CDOM has no scattering contribution.	95

4.10	MODIS 250 m radiance data (645 nm) of Lake Ontario showing the overview of the Genesee River plume (August 13, 2003). Right image is the zoomed image of the Rochester Embayment.	97
4.11	The geometry of the bidirectional reflectance. It characterizes the energy scattered into the hemisphere above a surface as a result of incident radiation.	98
4.12	Radiance indicatrix in polar coordinate is represented by a circle for a perfect Lambertian surface.	98
4.13	Image derived lake surface temperature distribution on Julian Day 96 (April 5 1820EST), Julian Day 132 (May 11 1800EST), Julian Day 151(May 30 1830EST), Julian Day 167 (June 15 1830EST), Julian Day 192 (July 10 1825EST) and Julian Day 206 (July 24 1835EST), 2004.	100
4.14	A true color rendering of the VNIR spectrometer bands of MISI of Genesee River plume. (a) May 11, 1999 (b) July 5, 2000 (c) March 31, 2003 (d) June 7, 2004.	102
4.15	The components affect the remote sensing signal at the top of the atmosphere in the 0.4 - 2.5 μm range (Vermote and Vermeulen, 1999). .	103
4.16	The impacts of uncertainty on the typical errors of surface reflectance retrieval: (a) absolute calibration (b) aerosol model (c) aerosol optical thickness (d) BRDF (Vermote and Vermeulen, 1999).	104
4.17	Total theoretical absolute and relative errors of the atmospheric correction algorithm (Vermote and Vermeulen, 1999).	104
5.1	Schematic representation of the coupled measurement-modeling system.	107

-
- 5.2 A complete operation cycle of the Kalman filter showing the equations of time update and measurement update (Welch and Bishop, 2006). 109
- 5.3 Genesee River discharge during the year of 2003 at streamflow gaging station 04232000 (Genesee River at Rochester). 111
- 5.4 TSS concentration during the year of 2003 at water quality site 431510077363501 (Genesee River at Charlotte Pump Station). 111
- 5.5 Discharge (a) and TSS concentration (b) of the Genesee River during the simulation period. The real observation is given in red, and the ensemble of forcing inputs with Gaussian distribution at each time step is given in black. 112
- 5.6 Wind direction (a) and wind speed (b) during the simulation period. The real observation is given in red, and the ensemble of forcing input with errors contributed is given in black. 113
- 5.7 MODIS 250 m surface reflectance data showing Genesee River plume dissipation in Lake Ontario during summer of 2003. The images are false-color RGB combination using bands 2-1-1. 117
- 5.8 Frequency of occurrence of deviations of MODIS derived remote-sensing reflectance. The red line shows Gaussian around the mean difference. 118
- 5.9 Spatial correlation of MODIS surface reflectance observation errors as a function of distance. 118

6.1	Thermal imagery of middle ((a) and (b)) and northern ((c) and (d)) portions of Conesus Lake: (a) 14:25 on May 19, 2003 (b) 13:05 on June 28, 2003. Three regions (Rectangle, Ellipse, and Polygon) were picked to convert from thermal radiance to surface temperature. The images were not corrected for aircraft roll and have been enhanced with a histogram stretch to show variation within the lake. The horizontal striping is an indication of line-to-line noise emphasized by the enhancement.	122
6.2	Simulated and observed average surface water temperature during April to June, 2004 at Long Point.	124
6.3	Observed and simulated temperature profiles on July 1, 2004 for stations: (a) Long Point Gully (b) Cottonwood Gully.	125
6.4	Surface temperature (color scale is in C as well as in F in Lake Ontario from the GLFS on May 20 and May 30, 1997 (Chang, 2003).	127
6.5	Surface temperature (color scale is in C as well as in F in Lake Ontario from the GLFS on June 10 and July 10, 1997 (Chang, 2003).	128
6.6	Observed temperature transects at 78.3W (color scale is in C as well as in F) in Lake Ontario from the GLFS on (a) May 20 (b) May 30 (c) June 10 (d) July 10, 1997 (Chang, 2003).	129
6.7	Simulated surface temperature distribution of Lake Ontario on (a) April 20 (b) April 30 (c) May 10 (d) May 20 (e) May 25 (f) May 30 (g) June 10 (h) June 20, 2004.	130
6.8	Simulated temperature transects at 77.8W in Lake Ontario on (a) April 20 (b) April 30 (c) May 10 (d) May 20 (e) May 25 (f) May 30 (g) June 10 (h) June 20, 2004.	131

- 6.9 Simulated surface water temperature along the sampling transects on May 25, 2004. Curve I, II, and III corresponding to the lines in (a) from north to south. 133
- 6.10 Time series of simulated surface currents (cm/s) for the entire lake during spring and summer conditions, 2004: (a) April 7 (b) April 26 (c) May 30 (d) June 29. Arrow length and direction represent the current strength and direction. Average wind direction was northwest, southeast, southwest, and south, respectively. One out of three nodes is plotted for clarity. 135
- 6.11 Prediction of circulation (cm/s) at different depths for the entire lake with wind blowing from the southeast (June 29, 2004): (a) surface (b) depth = 3m (c) depth = 6m. One out of three nodes is plotted for clarity. 136
- 6.12 Time series of simulated surface currents (cm/s) for the entire lake during spring and summer conditions, 2004: (a) April 7 (b) April 27 (c) June 10 (d) June 27 (e) July 10 (f) July 27. One out of three nodes is plotted for clarity. Lake surface temperature at each simulation grid is indicated by the color of the current vector. 138
- 6.13 Prediction of circulation (cm/s) at different depths for the entire lake with wind blowing from the southeast (June 27, 2004): (a) surface (b) depth = 9 m (c) depth = 17 m (d) predicted surface temperature distribution. One out of two nodes is plotted for clarity. 139
- 6.14 Plume at Sand Point Gully flowing southeast in June, 2004 during a storm event. 141

- 6.15 Water circulation and passive tracer transport at the mouth of Sand Point Gully during a major storm event during June 16-19, 2004. The figures show the results 18 hours after the beginning of simulation. The wind was blowing from northwest. Large vectors in (a) show nudging data (direction of flow currents, not to scale) from whole lake simulation. The circled area indicates the macrophyte bed location. (a) Surface (b) 2 m (c) Tracer transport at surface and contour of tracer distribution at 3 m. The flow current vectors at the stream mouth are not on the same scale in (a) and (b) as the current vectors away from the stream because of the extraordinarily larger velocity at stream mouth. One out of four nodes is plotted for clarity. 142
- 6.16 Water circulation and passive tracer transport at the mouth of Cottonwood Gully during a major storm event during May 20-25, 2004. The figures show the results 48 hours after the beginning of simulation. The wind was blowing from northwest. Large vectors in (a) show nudging data (direction of flow currents, not to scale) from whole lake simulation. The circled area indicates the macrophyte bed location. (a) Surface (b) 2 m (c) Tracer transport at surface and contour of tracer distribution at 3 m. The flow current vectors at the stream mouth are not on the same scale in (a) and (b) as the current vectors away from the stream because of the extraordinarily larger velocity at stream mouth. One out of four nodes is plotted for clarity. 143

6.17 Integrated sediment deposition pattern at Sand Point Gully after 53 hours and Cottonwood Gully after 117 hours with and without macrophyte bed. The circled area indicated macrophyte beds location. (a) and (b) Sand Point Gully, (c) and (d) Cottonwood Gully.	145
6.18 Simulated movement of the Niagara River Plume from a two-week simulation in summer 2004. Left column is the predicted movement of the plume and right column is the corresponding simulated surface flow currents. (a) & (b) June 5 (c) & (d) June 6.	147
6.19 Simulated movement of the Niagara River Plume from a two-week simulation in summer 2004. Left column is the predicted movement of the plume and right column is the corresponding simulated surface flow currents. (a) & (b) June 7 (c) & (d) June 10 (e) & (f) June 16. . .	148
6.20 ALGE predicted TSS distribution showing the Genesee River plume. Yellow represents the high concentration and blue represents the low concentration.	150
6.21 Relationship between measured and modeled TSS and CDOM concentrations.	150
6.22 An example of Chl profile obtained from Eq. 6.1.	151
6.23 Modeled remote sensing reflectance at three stations.	152
6.24 Measured and modeled remote sensing reflectance at stations 2, 3, and 4.	154
6.25 Relationship between TSS concentration (g/m^3) and remote-sensing reflectance R_{rs} generated by Hydrologht at the wavelength of 645 nm. Concentration of CHL and CDOM are assumed to be constants. . . .	156

6.26	TSS concentration (g/m^3) at mouth of the Genesee River during the study period from ALGE controlled simulations and truth. The ensemble of model simulations is given in black, the truth is given in red, and the filtering estimate is given in blue.	157
6.27	Kalman gain calculated by the EnKF during the study period. Values shown are for point-scale test.	159
6.28	Errors of TSS concentration (g/m^3) for the EnKF estimate with 11 ensemble members for the period of July 26 to August 18.	160
6.29	Errors in TSS concentration (g/m^3) for the mean of the ensemble (green), one ALGE open-loop run (magenta), and the EnKF estimate (blue) for the period of July 26 to August 18.	162
6.30	Cumulative errors of TSS concentration (g/m^3) for the mean of the ensemble (green), one ALGE controlled open-loop run (magenta), and the EnKF estimate (blue) for the period of July 26 to August 18.	163
6.31	TSS dissipation in the Rochester Embayment on (a) July 29 (b) July 30 (c) August 8 (d) August 13, 2003. First column: MODIS derived, second column: ensemble mean, third column: EnKF estimate with 12 ensemble members.	166
6.32	TSS dissipation in the Rochester Embayment on (a) August 14 (b) August 15 (c) August 16 (d) August 18, 2003. First column: MODIS derived, second column: ensemble mean, third column: EnKF estimate with 12 ensemble members.	167

List of Tables

3.1	Definitions and units for parameters and flags used in the ALGE model.	78
4.1	Water sampling estimates derived from laboratory measurements. . .	93
4.2	MODIS bands and the principal areas of application	96
5.1	EnKF definitions and notations	108
5.2	Definitions, values, units and data sources for parameters used in ALGE.	114
6.1	Comparison of lake surface temperature (LST) derived for MISI im- agery and simulation output	123
6.2	TSS and CDOM concentrations predicted by ALGE at four stations on August 9, 2006.	150

Chapter 1

Introduction

1.1 Problem statement

Computer-based models have been utilized to study and characterize the water quality in the Great Lakes since 1960 (Alwis, 1999). Recently these environmental simulation models, so far primarily physical or biogeochemical, but not linked, are increasingly recognized as a fundamental method to provide diagnostic and predictive outputs for spatially distributed, time-dependent environmental processes.

Physical processes such as the thermal cycle and water circulation in lakes have a pronounced influence on water quality conditions. In large temperate lakes stratification characteristics, such as the thermal bar, the thermocline depth, and the upwelling and downwelling dynamics are the major features of the thermal cycle that affect water quality. For example, the thermal bar is a downwelling plume of water which arises from the existence of a temperature of maximum density in fresh water. It slows the horizontal mixing of water, which results in the higher concentration of nutrients and pollutants coming off the land to remain in a small band of warmer water close to

the shore rather than distributing evenly over the entire body of water. Therefore, the presence of a thermal bar can affect carbon flow and lake ecology by controlling the seasonal distribution of dissolved or suspended water constituents in lakes.

The thermal characteristics of a lake together with wind patterns are responsible for the unique circulation patterns of each lake, that may be quite complicated within the nearshore zone. However, due to the limitation of numerical scales in time and space, relatively little research has been concerned with the physical properties of nearshore areas in lakes. Nevertheless, the large-scale circulations can drive many of the coastal hydrodynamic phenomena. Thus it is always necessary for local studies with a nearshore focus to extend into the open lake to determine the physical properties.

The transport and distribution of water constituents driven by the lake hydrodynamics can determine the optical dynamics in coastal areas. Light in the water is affected by both absorption and scattering processes, and these processes are tightly coupled to the type and amount of the components in the water, such as pigments, suspended sediments, and dissolved organic matter (Bukata et al., 1995). Optical properties of water constituents provide the link between the color of water which can be measured from space using satellite sensors and the constituents in the water body. The concentrations of dissolved and particulate substances are typically higher in coastal areas than deep-lake regions, causing coastal waters to appear typically greener and deep-lake regions to appear bluer. The particulate and dissolved substances in the water are also related to the reflectance spectrum at visible wavelengths through the radiative transfer equations. Therefore, it is important to understand the link between the remote sensing parameters of lakes and the vertical structure of the water constituents and their optical properties.

Modeling of lake physical properties is a powerful tool, however, due to the uncertainties in model physics, parameters, and initial conditions, hydrodynamic model estimates are imperfect and drift away from reality through simulation time. Data assimilation is capable of coupling the space-time continuity of a hydrodynamic model with intermittent remote observations in a framework that balances model and measurement uncertainties to provide an optimal solution. Recent research has shown that assimilation of remote sensing data can improve various types of simulations (Slater and Clark, 2006) (Crosson et al., 2002) (Reichle et al., 2002) (Andreadis, 2004). In this research, the Dynamic Data-Driven Applications Systems (DDDAS) concept (Darema, 2004) is used to implement the dynamic data assimilation modeling system. A DDDAS can feed data into an executing application either as the data is collected or from a data archive. The simulation makes predictions about the entity regarding how it will change and what its future state will be. The simulation is then continuously adjusted with data gathered from the entity. Simulations and measurements become a symbiotic feedback control system and the simulations can even steer future measurements to where they will be most effective. By using a DDDAS design, for a water quality monitoring and prediction system, the ability to incorporate real-time data from satellites or aircraft will greatly improve the overall reliability of the simulation results.

1.2 Objectives

The overall objective of this study is to develop and test a coupled water quality modeling system using the DDDAS concept that assimilates remote sensing data into the hydrodynamic model to understand the link between the lake hydrodynamics, including thermal structure and circulations, lake optical dynamics, and remote sensing pa-

rameters. The motivation of this work is based on the idea that the combination of individual models into a modeling system provides a better understanding of the lake hydrodynamics and water quality parameters that can be used for prediction and management. This coupled hydrodynamic-optical modeling system uses remote sensing data as a dynamic feedback for hydrodynamic modeling of materials transport and distribution that influence the optical dynamics in the lake.

Specific objectives of the hydrodynamic modeling are to provide information on how the thermal structure and circulation patterns are strongly affected by the thermal bar in lakes during spring and summer. The physical processes including the three-dimensional thermal structure and seasonal circulation pattern at space and time scales are studied. Additionally, this research also provides the transport and distribution of materials (dissolved and particulate) that enter the lakes from tributary streams. Since the input to hydrodynamic model is geo-referenced and site specific, the output can be directly overlaid with geo-referenced remotely sensed imagery for comparison and feature extraction of regions of interest.

The optical dynamics in the coastal area are dependent on the distribution of water constituents that can be affected by the thermal cycle and flow currents in lakes. The spectral shape of the reflectance of a water body depends on the state of air-water interface, bottom reflectance, and the absorption and scattering properties of water and other constituents. *In situ* measurements of optical properties, including the absorption and scattering properties as well as backscattering probability of water constituents, combined with distribution of the constituents predicted from the hydrodynamic model, enable the radiative transfer equations to calculate the expected remote sensing reflectance spectrum at various simulation grids. Model calculated reflectance is compared with remote observed reflectance and *in situ* measurements in the data

assimilation step.

A hydrodynamic model can predict the transport and distribution of water constituents discharged into lakes. However, hydrodynamic models are susceptible to errors and uncertainties in model physics, parameters, and initial conditions. Thus, hydrodynamic models work best if they are periodically re-initialized with observations. Data assimilation is used to produce an improved estimate of the model state by computing optimal weighting between a modeled and observed state if the estimates of the errors in the model and observation are given. Data assimilation provides a framework to merge satellite observations and hydrodynamic model predictions to estimate the model state in a way that accounts for both model and observation errors (McLaughlin, 1995). In this work a dynamic data-driven coupled modeling system is implemented through an Ensemble Kalman Filter (EnKF). High resolution hydrodynamic simulations are performed in a coastal area to study the transport and 3D distribution of sediments discharged into the lake. Satellite data is then dynamically fed back into the hydrodynamic model through an EnKF to obtain a better estimate of sediment distribution.

1.3 Overview of the study approach

A coupled measurement-modeling system incorporating a hydrodynamic model, a radiative transfer model, remote sensing observations, and data assimilation techniques is developed and tested.

The lake-wide thermal structure, water circulation, and mass transport are studied using the three-dimensional finite differencing hydrodynamic model ALGE. It solves the hydrostatic form of the partial differential equations that model conservation of

momentum, mass and thermal energy. ALGE is capable of producing high-resolution simulations for node-to-node matching with airborne and satellite imagery. Physical forcing factors, such as the initial water temperature, river discharge, river source concentration of water constituents, meteorological data, including wind speed/direction, air temperature, precipitable water, and cloud cover/height are required.

As we know, the nearshore waters tend to have complicated and unique thermal and circulation characteristics. However, only a few of the previous hydrodynamic studies of Lake Ontario focused on the near-field phenomena such as the flow circulation and spatial extent of the Niagara River plume (Murthy and Miners, 1992) (Masse and Murthy, 1990). In this research, high-resolution stream plume studies are performed in Conesus Lake and for the plume of the Niagara River in Lake Ontario. This work provided validation of the capabilities of the ALGE code to simulate the transport of sediment and passive tracer.

The primary goal of this research is to study the impacts of the turbid plume of the Genesee River on optical dynamics of the Rochester Embayment in Lake Ontario using the coupled modeling system. High resolution plume simulations are carried out to study the transport and distribution of water constituents during plume events by ALGE. Color Dissolved Organic Matter (CDOM) and Total Suspended Solids (TSS) are modeled as passive tracers and particles, respectively. *In situ* and laboratory Apparent optical Properties (AOPs) and Inherent Optical Properties (IOPs) measurements are made both in and outside the plume. Simulated CDOM and TSS profiles from the hydrodynamic simulations are entered into the radiative transfer model Hydrolight. Hydrolight uses the water constituents' distribution predicted from the hydrodynamic model to calculate the amount of light scattering and absorption within each layer of the water body and thereby predict the reflectance. The Hydrolight output is compared

with the measured AOPs and satellite data in a data assimilation step.

Many of the requirements for hydrodynamic modeling can be addressed by extensive use of remote sensing data. The water temperatures in lakes are retrieved from the satellite images as one of the initial parameters of hydrodynamic simulations. The surface water temperatures needed to compare with the model output are seldom available all over the lake except for the occasional stations on a buoy. Also, remotely sensed data from satellites can be used to monitor the occurrence of the thermal bar and other manifestations of the spatial patterns of water movement. A number of satellite products are currently available and having the potential to aid in development of thermal bar monitoring and modeling in Lake Ontario. However, many important environmental phenomena, such as stream plumes and the distribution of suspended sediments, are nearshore and cannot easily be examined in satellite imagery due to the large ground pixel size. However, the MODerate resolution Imaging Spectroradiometer (MODIS) sensor on board the Terra and Aqua satellites has two bands (645 and 865 nm) that acquires data at a 250 m ground spot diameter. While these two bands do not provide a rich optical signal, the 645 nm band does respond strongly to the sediment load in the plume and thus is an accurate measure of its distribution.

The models and the satellite data are brought together in the data assimilation step, which is based on the ensemble Kalman filter. An EnKF is implemented and applied in this research as a dynamic data assimilation tool to improve the hydrodynamic model performance and provide more accurate estimates of model state using remotely sensed observations. The model state is a probability distribution and represented by an ensemble of open-loop simulations. Then the model state is updated using the Bayes theorem by observations and the associated error estimates assuming the probability distributions are approximately normal.

Chapter 2

Background

2.1 Lake hydrodynamics: physical processes

2.1.1 Water and heat budget

The water budget is the hydrologic balance of a lake which provides useful information about the availability of water in lakes at any time. The basic parameters of the water budget are classified as recharge and discharge. These can be further divided as inflow from precipitation, surface runoff, and groundwater and outflow from river, evaporation and groundwater.

Temperature changes in lakes are the results of the variation of heat flux. The heat budget into and out of lakes includes several major factors: solar (short-wave) radiation, long-wave radiation, sensible heat transfer, latent heat transfer, and inflow and outflow. Radiation from sun and sky is the most important of the heating processes. Other heat sources are usually of much smaller importance, depending on local and seasonal circumstances. Evaporation is the main heat-loss process. Figure 2.1 shows

an overview of energy exchange in and out of a lake.

Solar radiation penetrates the atmosphere to reach the lake surface. Some of the solar radiation is reflected at the water surface and the remainder transmits into the lakes. Incoming solar energy varies seasonally and with the latitude and is greatly influenced by cloud cover. The fraction of the solar energy that is reflected away from the lake surface depends upon the solar angle, the turbidity of the atmosphere, the wave state of the lake, and the lake surface roughness.

Long-wave radiation originates from black-body radiation. Both the lake and the atmosphere emit black-body radiation. The amount of radiation emitted by the lake surface is proportional to the fourth power of the surface temperature, which results in a loss of thermal energy and cooling of the lake water. The radiation emitted by the clouds and atmosphere overlying the lake depends primarily upon the amount and height of the clouds and the temperature and moisture content of the atmosphere near the lake surface.

A sensible heat flux is a movement of heat between the lake surface and the atmosphere, which depends on the temperature difference between the lake surface and the adjacent air. Sensible heat transfer can be intensified by maintaining large temperature gradient at the air-water surface from the wind.

Latent heat is the energy required to change a substance to a higher state of matter (solid \rightarrow liquid \rightarrow gas). This same energy is released from the substance when the change of state is reversed. Evaporation from the lake surface extracts heat from the lake and results in cooling of the water surface. Condensation extracts heat from the atmosphere and adds it to the water surface, resulting in heating at the water surface. This process is determined by the relative humidity of the air and can only affect heat at the very surface of a lake. Evaporation and condensation are also accompanied by a

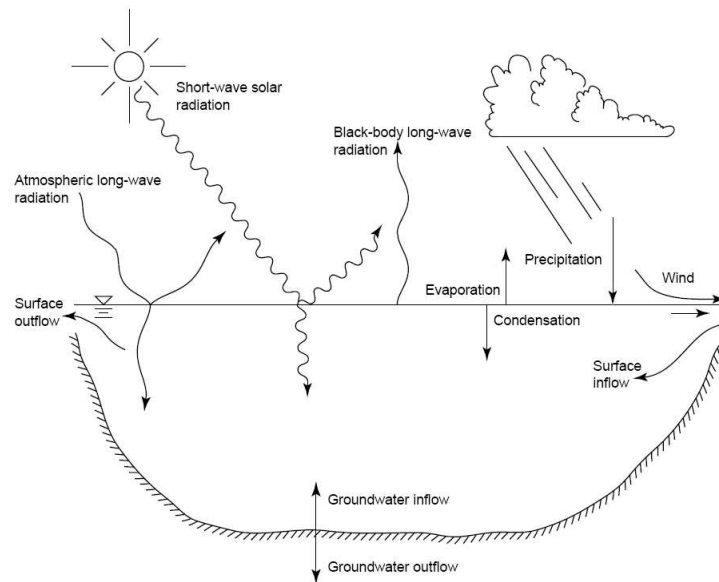


Figure 2.1: A schematic view of all the major energy inputs to a typical lake. (<http://ceprofs.tamu.edu/kchang/ocen689/ocen689ch9.pdf>)

flux of water. Thus, they can also affect the total water budget of a lake.

Hydrodynamics in large lakes, such as the stratification and mixing, water circulation, and material transport, depends on many factors including geometry, surrounding topography, hydrological and geochemical loadings, and meteorological conditions. Wind, waves, and thermal structure are primary determinants of water movements, mixing, and circulation in large lakes. The movement and mixing of water in natural systems affects water quality, biological community structure and productivity, and the transport of sediment and contaminant, especially in nearshore areas.

2.1.2 Thermal stratification and mixing

Thermal stratification and mixing are the changes in the temperature profile with depth within a lake system. Most important for stratification is the water temperature dependence of the water density, which reaches a maximum close to 4 °C. Figure 2.2 shows an overview of the seasonal changes in a dimictic temperate lake. During spring, the lake water is generally the same temperature from the surface to the bottom after the ice melts. Surface water can be pushed to the lake bottom and bottom water can rise to the surface by wind. The mixing of the lake water at this time of year is called spring turnover. As air temperatures rise in late spring, heat from the sun begins to warm the lake resulting in the warm water with less density floating over the cold water. The layers of warm water at the surface and the cold layer below it are called the epilimnion and hypolimnion, respectively. These two layers are separated by the thermocline which rapidly changes temperature with depth. During the summer the epilimnion reaches a maximum depth and the stratification is maintained for the remainder of the summer. Typically, the warm water of the epilimnion is unable to drive through the cold, dense water of the hypolimnion even if the surface water is circulated by wind. The epilimnion cools and becomes both denser and heavier as the fall approaches. The cooler and denser water sinks and fall winds stir the epilimnion which gradually erodes the thermocline. Eventually, the temperature of the entire lake is uniform. As the winter comes, surface water is in direct contact with the cold air and gets cooled faster than the water below. The cold, dense water sinks and further helps to stratify the lake. The surface water stops sinking when it is eventually cooled below 4 °C. Ice begins to cover the surface of the lake as surface water temperature reaches 0 °C, which prevents wind from mixing the lake water. A layer of low density water colder than 4 °C, but warmer than 0 °C forms just below the ice. The remainder of the

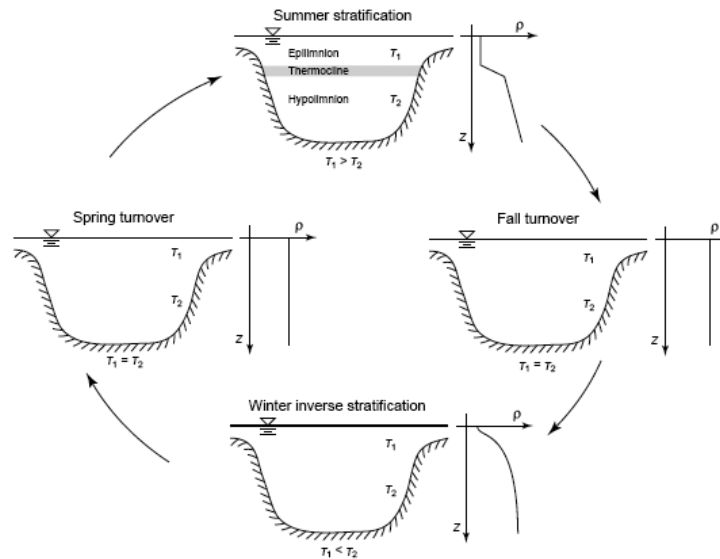


Figure 2.2: Overview of the seasonal cycle in a dimictic lake. A density profile and the relationship between temperatures in the upper and lower regions are shown. (<http://ceprofs.tamu.edu/kchang/ocen689/ocen689ch9.pdf>)

lake water is usually near 4 °C. This is called winter stratification.

2.1.3 Thermal bar

As a result of the increased surface heating in the early spring, the nearshore littoral zone begins warming more rapidly than deeper regions. Eventually the lake water is divided into two zones: offshore un-stratified water lower than 4 °C and nearshore weakly stratified water slightly higher than 4 °C. The convergence of 4 °C water between the two zones starts sinking, which is referred to as the thermal bar. The near-bottom currents moving from the bar and the near-surface currents converging onto the bar result in the region of strong downwelling currents. The thermal bar gradually moves offshore as the heating of inshore areas continues. The thermal bar is the main

characteristic of the coastal lake hydrodynamics during the spring in large lakes.

The thermal bar formation occurs during the spring when plants and animals in lakes are in the early stages of development. It can enhance algal growth by effectively trapping heat and nutrients because the sharp density front across the thermal bar effectively limits the exchange of materials between nearshore and offshore. This process is likely to encourage the growth of macrophytes in the nearshore area and have critical impacts on the ecosystem of the lake basin. On the other hand, the thermal bar also traps suspended sediments nearshore which may decrease plant growth due to light limitation. Measurements of the thermal bar have included satellite images and detailed observations, such as temperatures, current velocities, optical characteristics, and bio-chemical characteristics. The study of the thermal bar can be used to determine the physical dynamics and also to investigate its impact on water-quality conditions during the spring.

Field investigations of the thermal structure in Lake Ontario during the spring of 1970 were conducted by Rodgers (Rodgers, 1971). The survey program consisted of measuring temperature profiles at the stations in the deepest section of Lake Ontario. Sampling took place during nine surveys and seven of them were consecutive weekly surveys from May 11 to June 24 (Figure 2.3). The near-surface converging currents onto the bar were clearly observed. The thermal bar region was seen at over 12 km from the shore (Figure 2.4). Since the formation and propagation of the thermal bar can be affected by the slope of the lake bottom (Malm et al., 1993), the thermal bar develops quickly in the north side of Lake Ontario, while the south side of the lake has a steeper slope compared to the north side and is heated up much slower.

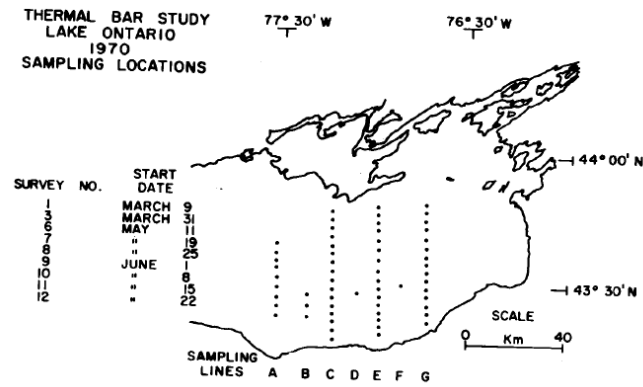


Figure 2.3: Sampling locations and survey dates of Rodgers' field survey of Lake Ontario (1970) (Rodgers, 1971).

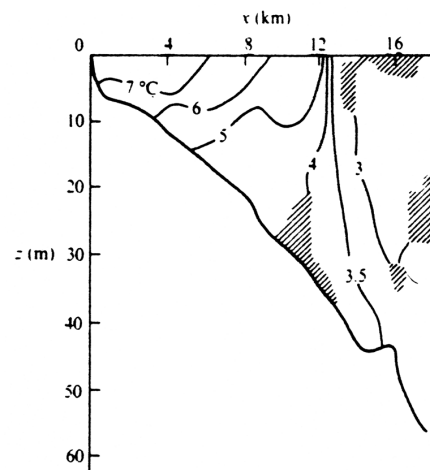


Figure 2.4: One temperature profile in Lake Ontario from Rodgers' field survey (1965) (Rodgers, 1971).

2.1.4 Circulation structure of lakes

Wind, solar radiation, gravity, and the earth's rotation are important forces causing water movements in lakes. Knowing water movements is important because currents and waves influence the distribution of water constituents, such as the dissolved substances, nutrients, microorganisms, and plankton.

When the wind blows over the lake surface, if the lake is un-stratified, then all of water is moving down-wind. During the stratification season, the vertical exchange is greatly reduced and the main motions are almost entirely horizontal. As momentum is transferred at the air-water surface, the lake receives less than 5% of the wind energy from the atmosphere. Surface waves transport and dissipate a portion of this energy, whereas the remaining energy forms large-scale wind-driven currents. Velocity of wind-driven currents is about 2% of wind velocity and is independent of the height of the surface waves. However, the wind-driven circulation can occasionally reach up to several tens of centimeters per second. In the Great Lakes, a cyclonic gyre is another type of large-scale horizontal current pattern driven by wind. Gyres form mainly as a result of the Coriolis effect and non-uniform wind forcing. Stratification also contributes to the formation of vortices.

2.1.5 River plume effects on the lake hydrodynamics

When a river enters into a lake, a distinct color difference between the river and lake water can commonly be observed. The area that appears to be an extension of river-colored water into a lake is called a plume (Figure 2.5). River water commonly appears browner in color than the lake because the river is ordinarily carrying more suspended particles such as silt and clay.

However, the plume is not just different in color. There is usually a density dif-

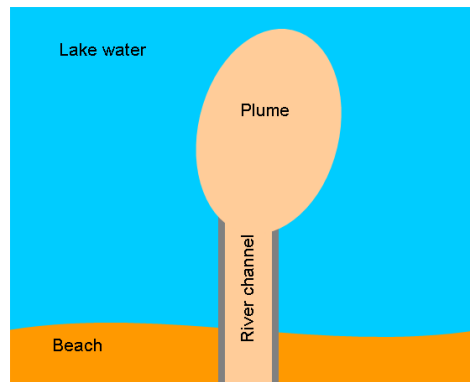


Figure 2.5: An aerial view showing the sediment plume from a river to a lake.

ference between the inflowing water and the ambient water body. A surface buoyant overflow is formed if the inflow is less dense than the surface water. If the inflow is heavier than the surface water, it forms an underflow along the bottom boundary. River flow is a main source of nutrients, suspended sediments and contaminants to lake. A large amount of rainfall in the watershed transfers a large amount of soil to the rivers and leads to high sediment loads. River sediments can also transfer man-made pollutants to the lake. As the soil is carried away by rain and into the rivers, it will be transferred to the river mouth and deposited in the river delta. An understanding of river plume dispersion is important for water quality management. The dispersion of the river plumes represents a redistribution mechanism. The incoming river plumes control the discharge of mass and constituents from the tributary area and the processes of redistribution and transformation in the water body of the lake.

The geometry of a river plume as well as the sharpness of its boundary may vary considerably depending on a number of factors, such as wind, wave, and current conditions as well as the amount of material carried in suspension. The tremendous amount of sediment and toxic materials that are carried to the Great Lakes daily by tributary

rivers decrease the water quality of the lakes. As a result, control of the erosion of sediment into streams is critical to maintaining high water quality in the Great Lakes.

2.2 Optical properties of natural water

The behavior of light within the water body determines the color and clarity of water. As a beam of light travels in a water body, it is attenuated by scattering and absorption. Scattering and absorption are called inherent optical properties (IOPs) because they do not depend on the amount or direction of sunlight (i.e. they depend solely upon the medium through which the light is traveling). IOPs include the absorption coefficient, the scattering coefficient, the volume scattering function, the index of refraction, the beam attenuation coefficient, and the single-scattering albedo. Apparent optical properties (AOPs) which describe the optical behavior of water bodies in a particular light field include the diffuse attenuation coefficients, the irradiance reflectance, and the remote sensing reflectance.

The absorption coefficient, $a(\lambda)$, defines the fraction of flux absorbed over an infinitesimal distance for a particular wavelength. Similarly, the scattering coefficient, $b(\lambda)$, evaluates the fraction of radiant flux that is scattered over a distance for a particular wavelength. In pure water, light is scattered in a manner consistent with Rayleigh scattering with approximately equal amounts being scattered backwards as forwards. However, the shape of the scattering (scattering phase function) changes drastically as the particles are introduced to the water, so that the majority of the radiant flux is scattered in the forward direction. The scattering phase function depends on the refractive index of the scattering particles in the water and the size distribution of the particles.

2.2.1 Components of natural water

The optical properties of water are dependent on the constituent composition of the water body. The absorption and scattering cross sections of each substance in a natural water body and its concentration level provide the absorption coefficient $a(\lambda)$ and scattering coefficient $b(\lambda)$ as:

$$\begin{aligned} a(\lambda) &= \sum_{i=1}^n C_i a_i(\lambda) = \text{Absorption coefficient}, \\ b(\lambda) &= \sum_{i=1}^n C_i b_i(\lambda) = \text{Scattering coefficient}, \end{aligned} \quad (2.1)$$

where C_i are the concentration levels, $a_i(\lambda)$ and $b_i(\lambda)$ are the absorption and scattering cross sections for each component.

Natural water bodies are comprised of an unlimited amount of substances. Yet, as a simplification, case II water constituents are divided up into pure water, colored dissolved organic matter (CDOM), chlorophyll-a (CHL), and total suspended solids (TSS).

CDOM is produced from the decay of plants and is usually brown or yellow in color. Surface run-off and river discharge introduce a wide variety of CDOM into the water. Both sources from within and outside the water can result in higher CDOM concentrations in inland and coastal waters than they are in the open ocean or middle of a large lake. For example, the range of concentration values of CDOM in a river can be several thousands of times higher than that in the deep ocean or large lake.

All phytoplankton in water bodies contain the photosynthetically active pigment chlorophyll-a. Chlorophyll-a is the most prominent chlorophyll since it exists in all green algae. Bukata suggested that chlorophyll-a is a reasonable surrogate for the

organic component of optically complex natural waters (Bukata et al., 1995). The knowledge of the concentration and location of chlorophyll-a is important to estimate the locations and concentrations of phytoplankton.

Suspended matter includes many types of fine organic and inorganic matter, clays, microorganisms, and plankton. The suspended solids are the most important of the constituents in the suspended matter because of their high scattering coefficients. They enter the water body from wind deposition, erosion and sediment transport in tributaries, and resuspension of bottom materials. These particles usually consist of finely ground sand, clay minerals, and metal oxides. The size of these particles can range from 3~4 microns in diameter to several hundreds of microns. Also, the concentration level can range from 0.02~0.17 g/m³ in the open ocean to as high as hundreds of g/m³ in some inland lakes. The literature reports the concentration level of these particles in Lake Ontario is about 0.2~8.9 g/m³ (Bukata et al., 1995).

2.2.2 Spectral response of water constituents

Variations in the spectral absorption of natural waters result directly from variations in the concentrations and chemical compositions of material substances distributed within the water volume. These absorbing and scattering materials may be present either in suspended particulates or as dissolved materials.

Bukata publishes the cross section curves for Lake Ontario waters in the range of 400 nm to 690 nm (Bukata et al., 1995). The cross section curves are given in Figure 2.6. Algal particles are strong absorbers in the blue and red portions of the light spectrum. By absorbing red and blue light, the chlorophyll and related pigments in algae make the water appear green. CDOM consists of humic substances and tannins from decayed plant matter, including algae and terrestrial plants. It strongly absorbs blue

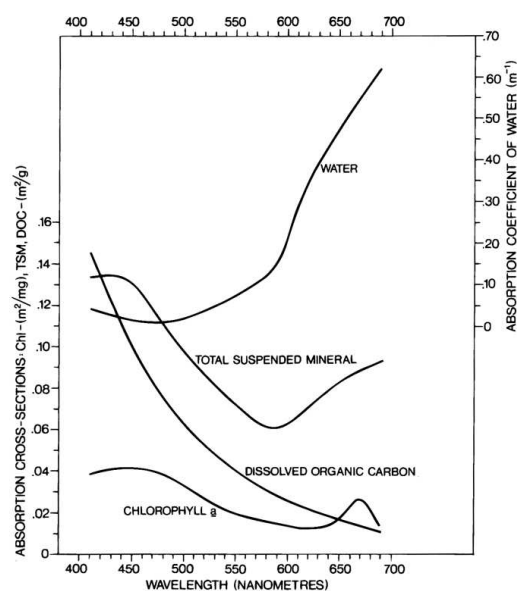
and ultraviolet light, so the absorption coefficient follows an exponential decay with increasing wavelength. In certain instances, there may be sufficient dissolved organic matter in the water to reduce the penetration of light in the water column. Pure water absorbs light in the long wavelength of the spectrum causing very clear water to have a bright blue appearance. Almost all of the incident near-infrared and shortwave-infrared (0.7 ~ 2.5 microns) radiant flux entering a pure water body is absorbed with negligible scattering taking place. Water bodies appear very dark in near infrared remote sensing images because they absorb almost all of the incident radiant flux, especially when the water is deep and pure and contains little suspended sediment or organic matter.

Scattering in the water column is important across all wavelengths of the visible-near infrared spectrum. The wave nature of light causes it to scatter more strongly from very small particles of sizes on the order of a few wavelengths. Inorganic sediments more strongly scatter light in the short wavelength region than long wavelength region. The scattering coefficient of CDOM is generally accepted to be small enough to ignore.

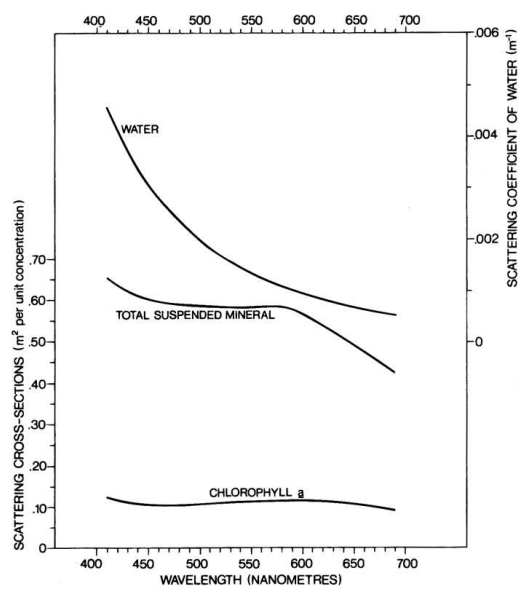
Therefore, Eq. (2.1) can be rewritten as the following:

$$\begin{aligned} a(\lambda) &= a_w(\lambda) + C_{CDOM}a_{CDOM}(\lambda) + C_{SM}a_{SM}(\lambda) + C_{chl}a_{chl}(\lambda), \\ b(\lambda) &= b_w(\lambda) + C_{SM}b_{SM}(\lambda) + C_{chl}b_{chl}(\lambda). \end{aligned} \quad (2.2)$$

Figure 2.7 shows the mass-specific absorption coefficient (m^2/mg) for brown earth, calcareous sand, yellow clay, red clay, and Bukata's data (Mobley and Sundman, 2000a). The absorption coefficient of the mineral particles ($1/m$) can be obtained by multiplying these values with the concentration of mineral particles in g/m^3 . The wavelength dependence of the mineral specific scattering data are shown in Figure 2.8 (Mobley and Sundman, 2000a). In shallow coastal waters, the water color is also influenced by the bottom reflectance and the water depth. Several irradiance reflectance



(a)



(b)

Figure 2.6: Absorption and scattering cross section spectra for chl-*a*, total suspended mineral, DOC, and water for Lake Ontario (Bukata et al., 1995).

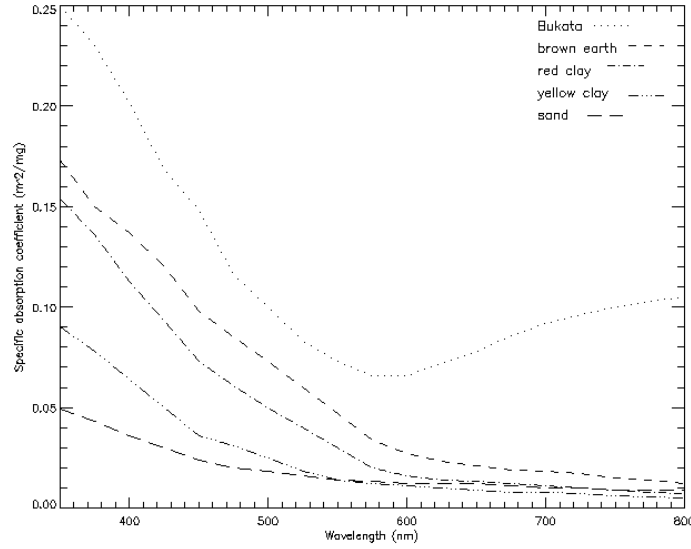


Figure 2.7: Mineral specific absorption spectra used for the simulations in Hydrolight (Mobley and Sundman, 2000a).

spectra for different types of sea floors, including brown, red, and green algae as well as coral sand and clean sea grass are shown in Figure 2.9 (Mobley and Sundman, 2000a).

2.2.3 Impacts of PSD on IOPs

Particles in a natural water body are seldom all of the same size. The size of particles may vary over quite a wide range. The range is normally broken up into a number of classes and the particles fall into each of the various size ranges. This range is called Particle Size Distribution (PSD) and is typically shown as a histogram chart with percentage-smaller-than on the y-axis and size ranges on the x-axis. PSD is an important parameter in modeling the fundamental IOPs of natural water body. Different PSDs can result in different absorption and scattering coefficients and the scattering phase function.

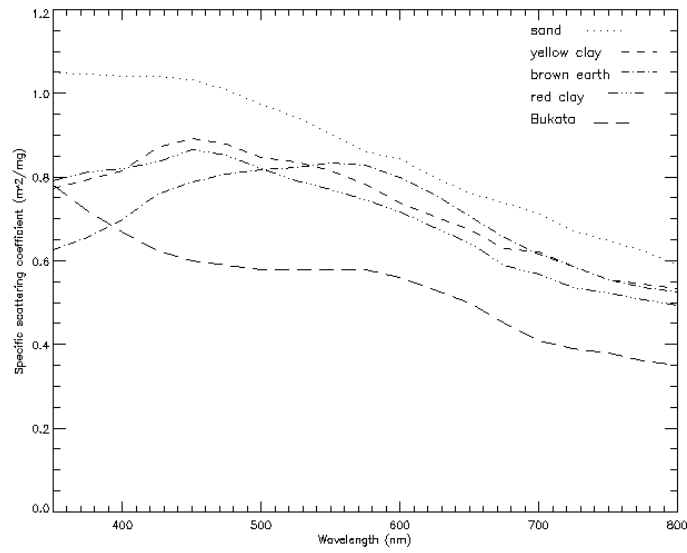


Figure 2.8: Mineral specific scattering spectra used for the simulations in Hydrolight (Mobley and Sundman, 2000a).

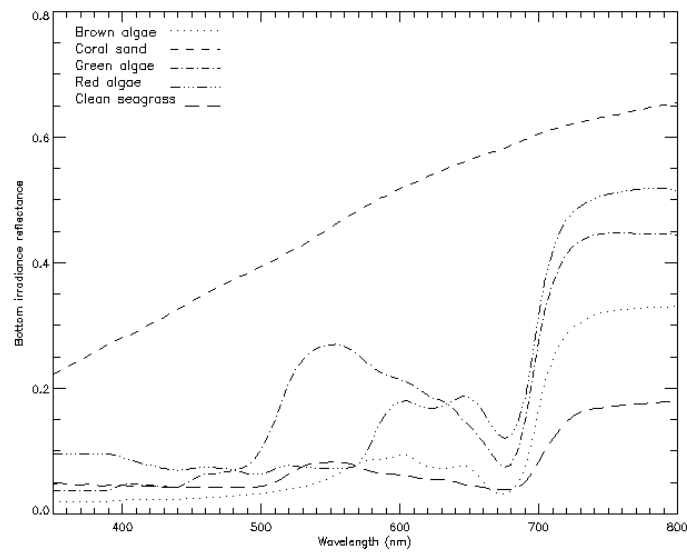


Figure 2.9: Bottom irradiance reflectance spectra of different types of sea floors used for the simulations in Hydrolight (Mobley and Sundman, 2000a).

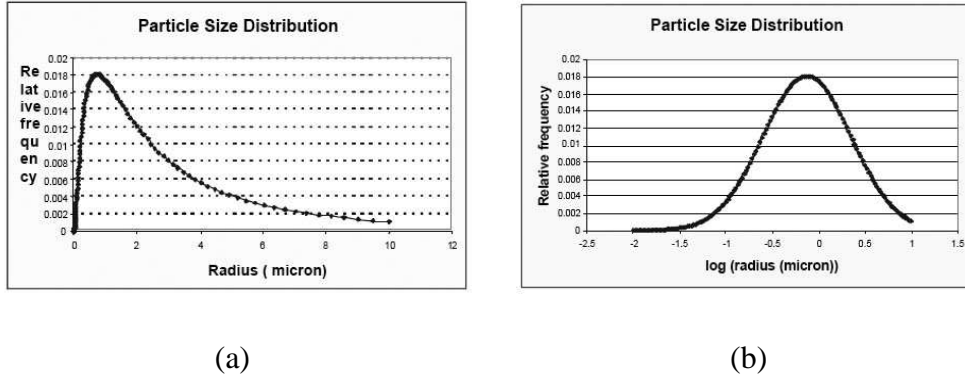


Figure 2.10: (a) A typical log-normal PSD plotted with respect to the radius (microns). (b) The same PSD as in (a) plotted with respect to $\log(\text{radius})$ (microns).

PSDs in the ocean are often found to be well approximated by a hyperbolic (Junge-like) distribution, which has the form:

$$N(D) = N_0(D/D_0)^{-m}, \quad (2.3)$$

where $N(D)$ is the number of particles per unit volume per unit bin width. D_0 is the maximum of the particle size. The exponent of the PSD typically varies between 3 and 5 so the number of smaller particles falls off with the inverse of a power of particle size.

The most common PSD is log-normal distribution which is based on the Gaussian distribution (Figure 2.10). If one specifies the median size (which in this case corresponds to the maximum frequency) and the spread of the distribution, the entire curve is fully specified. Log-normal distribution is the way that most PSDs are represented, and almost any real distribution can be approximated in this way.

Even though both particle types and refractive index affect IOPs, PSD is the main factor. Figure 2.11 shows the absorption and scattering coefficients calculated for cal-

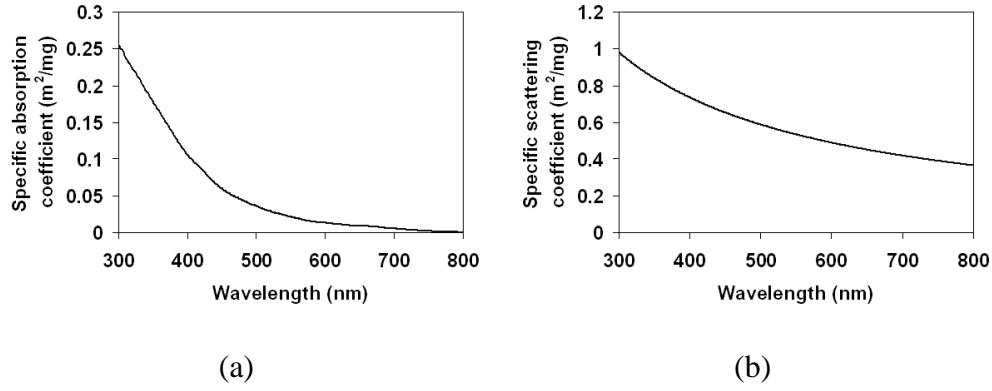


Figure 2.11: Absorption and scattering coefficients calculated for calcite with a refractive index of 1.486.

cite with a refractive index of 1.486 (J. Hamel, unpublished data, Rochester Institute of Technology).

2.3 Contributions to sensor reaching radiance over natural water

The total radiance $L(\theta, \phi, \lambda)$ recorded by a remote sensing system (aircraft or satellite) over a water body is a function of the electromagnetic energy from four sources (Figure 2.12) (Bukata et al., 1995):

$$L(\theta, \phi, \lambda) = L_a + L_s + L_v + L_b. \quad (2.4)$$

- L_a is the atmospheric path radiance resulting from the downwelling solar (E_{sun}) and sky (E_{sky}) radiation. This is an unwanted portion that never reaches the air-water interface.

- L_s is the radiance that reaches the air-water interface but only penetrates it a millimeter or so and is then reflected from the water surface. It therefore contains spectral information about the near-surface characteristics of the water.
- L_v is called subsurface volumetric radiance resulting from the downwelling radiation that penetrates the air-water interface, interacts with the organic/inorganic constituents in the water, and re-emerges from the water column without encountering the bottom. It provides information about the internal characteristics of the water column.
- L_b is the radiance that penetrates the air-water interface, reaches the bottom of the natural water body, propagates back through the water column, and exits the water column. It therefore represents the information about the bottom (e.g., depth, color).

Since it shows the direct result of aquatic absorption and scattering processes impacting on the photon flux in its subsurface propagation, L_v contains most of the information on the natural water body. L_v is a function of the concentration of each component in natural water, and the total amount of absorption and scattering attenuation that takes place in the water column due to these components:

$$L_v = f(a(\lambda), b(\lambda)). \quad (2.5)$$

Downwelling irradiance, E_d , is due to the stream of downwelling light and represents the irradiance attributable to the hemisphere above the horizontal plane. The upwelling irradiance, E_u , is due to the stream of upwelling light and represents the irradiance attributable to the hemisphere below the horizontal plane. Downwelling and

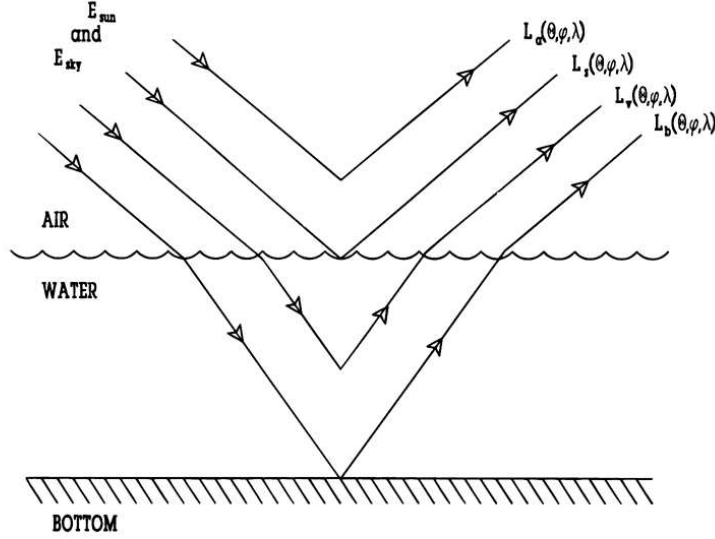


Figure 2.12: Components of the radiance collecting by a remote sensor over a natural water body (Bukata et al., 1995)

upwelling irradiance are obtained by integrating over the upper hemisphere and the lower hemisphere, respectively. They are defined as

$$E_d = \int_0^{2\pi} \int_0^{\pi/2} L(\theta, \phi) \cos\theta \sin\theta d\theta d\phi, \quad (2.6)$$

$$E_u = - \int_0^{2\pi} \int_{\pi/2}^{\pi} L(\theta, \phi) \cos\theta \sin\theta d\theta d\phi. \quad (2.7)$$

Irradiance reflectance, R , is the ratio of the upwelling irradiance at a point to the downwelling irradiance at that point. It can interpret the remotely sensed spectral data in terms of the compositions of natural water bodies. The irradiance reflectance is defined as

$$R(\theta, \phi, \lambda) = \frac{E_u}{E_d}. \quad (2.8)$$

The spectral remote-sensing reflectance (radiance reflectance), R_{rs} , measures how much of the downwelled irradiance that is incident on the water's surface is returned to the surface in a specific direction. The equation for the spectral remote-sensing reflectance is given below:

$$R_{rs}(\theta, \phi, \lambda) = \frac{L_u}{E_d}, \quad (2.9)$$

where L_u is the water leaving radiance as a function of the water leaving direction (θ , ϕ), the wavelength (λ), and the depth. The depth is set to the air just above the surface of the water. The downwelled irradiance evaluated just above the surface for each wavelength. Spectral remote-sensing reflectance $R_{rs}(\lambda)$ contains information about the properties of the water body surface layer whose thickness depends on the inherent and apparent optical properties.

Q -factor, ($Q = E_u/L_u$), converts irradiance reflectance, ($R = E_u/E_d$), to remote sensing reflectance, ($R_{rs} = L_u/E_d$). Q -factor is dependent on the solar zenith angle and the IOPs of the water column, specifically the volume scattering function and the single scattering albedo. Q is often considered to remain constant over wide geographical areas as well as over the visible spectrum.

2.4 Data assimilation concepts and methods

Data assimilation is a statistical technique to combine measurements of state variables with models that describe the time evolution of these variables to produce an optimal estimate of state variables. The error statistics of both the measurements and the model are used to estimate the best guess of the actual field. The observations of the variables (current and/or past) are assimilated into the model system by taking advantage of consistency constraints. When the measurements are very accurate, the best guess of

the state variables will be close to the observations. When the model produces a good forecast in comparison to the accuracy of the measurement, the field variables will be adjusted only slightly. The assimilation may identify the modeling problems and provide information about the quality of the measurements as well.

Data assimilation methods are divided into two classes: variational and sequential methods. Variational data assimilation is based on optimal control theory. 3-D Variational Assimilation (3DVA) and 4-D Variational Assimilation (4DVA) are the main methods. The major difference between 3DVA and 4DVA is that observations are only used when available at the time of analyses in 3DVA, whereas in 4DVA the past observations are included as well (*i.e.* time dimension is added). An important advantage of the variational approach is the ability to incorporate future measurements in the analysis. In this sense there are twice as many measurements available compared to the sequential technique, improving the accuracy of the assimilated field.

Most assimilation techniques analyze the measurements sequentially. The sequential data assimilation methods use a probabilistic framework and produce estimates of the whole system state sequentially by propagating information only forward in time. When a measurement is available, a new analyzed field value is calculated. After the analysis, the field is propagated forward in time until the next measurement and the procedure is repeated. An additional equation for the second-order statistical moment is integrated forward in time to predict error statistics for the model forecast as well. The error statistics are then used to calculate a variance-minimizing estimate whenever observations are available (Evensen, 1994) (Evensen, 2003).

2.4.1 The Kalman Filter

Kalman Filter (KF) is a recursive sequential technique which has been used widely in the areas of control and prediction of dynamic systems (Kalman, 1960) (Negenborn, 2003). The KF estimates the state of a linear dynamic system which may be corrupted by noise sources using its mean and the covariance matrix. There are two recursive operations involved: 1) Propagation of the state vector and covariance in time; 2) Analysis of the state vector and covariance based on the observations available at that time.

A linear dynamic system governed by the following equations consists of a linear system model and a linear measurement model. The system model, A , predicts the evolution of the state x_k over time. It relates the state at the previous time step $k-1$ to the state at the current step k . The measurement model, H , describes how measurements z_k are related to states. In practice, the system model and measurement model might change with each time step or measurement. The KF system is a hidden Markov chain since the state x_{k+1} depends only on the state x_k but not on previous ones $x_{0:k-1}$ and the measurement z_k depends only on the state x_k . The system model and the measurement model are defined as

$$x_k = Ax_{k-1} + w_{k-1}, \quad (2.10)$$

$$z_k = Hx_k + v_k. \quad (2.11)$$

The dynamics of systems and measurements are not perfect. Random variables w_k and v_k represent the process and measurement noise, respectively. They are assumed to be independent of each other, white, zero-mean, and with Gaussian distributions. The independence assumption implies the amount of noise in system and measurements have no effect on each other. The white noise assumption implies that errors are not

correlated through time. The Gaussian assumption deals with the amplitude of the noise and states that the system and measurement noise are often caused by multiple small noise sources. Since the noises are with the zero-mean and Gaussian distribution assumptions, they are defined as

$$p(w) \sim N(0, Q), \quad (2.12)$$

$$p(v) \sim N(0, R). \quad (2.13)$$

The process noise covariance Q and measurement noise covariance R are defined as matrices below and might change with each time step or measurement. The main diagonal of the covariance matrices Q and R contains the variance in the state and measurement variables respectively. The off-diagonal elements are zero due to the assumption of independent noise:

$$Q = \begin{pmatrix} \sigma_{11}^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_{22}^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma_{33}^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_{qq}^2 \end{pmatrix}$$

$$R = \begin{pmatrix} \sigma_{11}^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_{22}^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma_{33}^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_{rr}^2 \end{pmatrix}$$

where σ_{qq}^2 is the variance of q th state variable changes with responding to one time step. The process noise covariance cannot be usually measured because we cannot directly observe the process we measure. If Q is chosen with a large value (uncertainty), a poor model can produce acceptable results. Measurement noise can be determined before the filter operation.

2.4.2 Prediction-correction state estimator

The KF is a state estimator working on a prediction-correction basis. It estimates the true state of systems and gives a measure of how certain it is that the state estimate is the true state. The state of the system changes over time and is affected by noise. Since the variables of the state may not be directly observable, sensor observations are used even though they are subject to noise.

The Kalman filter estimates a state by obtaining feedback in the form of noisy measurement. The prediction-correction basis can also be thought of as time update and measurement update. The time update is used to project forward the current state and error covariance estimates to obtain the *a priori* estimates for the next time step. The measurement update incorporate a new measurement into the *a priori* estimate to obtain an improved *a posteriori* estimate.

In KF, \hat{x}_k^- is defined as *a priori* state estimate at time step k given knowledge of the system process prior to step k . \hat{x}_k^+ is defined as *a posteriori* state estimate at step k given measurement z_k . Then *a priori* and *a posteriori* estimate errors are defined as

$$e_k^- = x_k - \hat{x}_k^-, \quad (2.14)$$

$$e_k^+ = x_k - \hat{x}_k^+. \quad (2.15)$$

The *a priori* estimate error covariance and the *a posteriori* estimate error covariance are defined as

$$P_k^- = E[e_k^-(e_k^-)^T], \quad (2.16)$$

$$P_k^+ = E[e_k^+(e_k^+)^T]. \quad (2.17)$$

Steps

The KF algorithm consists of an initialization step and an alternatively performed prediction and correction steps. Before the KF performing prediction and correction steps, it is initialized by specifying the initial posterior state estimate \hat{x}_0^+ and the initial system uncertainty in the posterior state P_0^+ .

In the prediction phase, the prior state estimate and the prior system uncertainty are computed in Eqs. 2.18 and 2.19. These equations describe how the KF updates the belief when the system moves one step forward in time. The prior state is estimated by the last posterior state and the system model. The system uncertainty in the prior state includes the propagation of the system uncertainty from the last posterior state to the current prior state and the state process noise. The most likely state of the system is predicted by Eq. 2.18 and Eq. 2.19 represents how much uncertainty in the state estimate being the true state:

$$\hat{x}_k^- = A\hat{x}_{k-1}^+, \quad (2.18)$$

$$P_k^- = AP_{k-1}^+A^T + Q_{k-1}. \quad (2.19)$$

The correction phase consists of the correction of the prior state estimate and the prior system uncertainty. When a measurement z_k of the true state is available, a posteriori state estimate \hat{x}_k^+ is computed as a linear combination of an *a priori* estimate \hat{x}_k^- and a weighted difference between an actual measurement z_k and a measurement prediction $H(\hat{x}_k^-)$. The difference is called the measurement innovation, or the residual. A

residual of zero means that there is no discrepancy between the predicted measurement and the actual measurement. The uncertainty P_k^+ is updated through Eq. 2.21 at the same time as well. The correction phase is only performed when there is a measurement available. The new posterior belief then is used to compute the new prior belief for the next timestep:

$$\hat{x}_k^+ = \hat{x}_k^- + K_k(z_k - H(\hat{x}_k^-)), \quad (2.20)$$

$$P_k^+ = (I - K_k H) P_k^-. \quad (2.21)$$

In Eqs. 2.20 and 2.21, K_k is called the Kalman gain and defined as

$$K_k = P_k^- H^T (H P_k^- H^T + R_k)^{-1}. \quad (2.22)$$

The Kalman gain shows how much the measurement is taken into account.

Figure 2.13 shows the organization of the computations in KF assimilation. The input to the KF algorithms includes: the definition of the system and the measurement model, the initial condition for x and P , the sequence of measurements, and the sequence of model and observation error covariance matrices P and R . The output is the sequence of estimates of the model state and its error covariance matrix.

Kalman gain K_k

The Kalman gain is chosen to minimize the *a posteriori* covariance. This minimization can be accomplished by taking the derivative of the *a posteriori* estimate

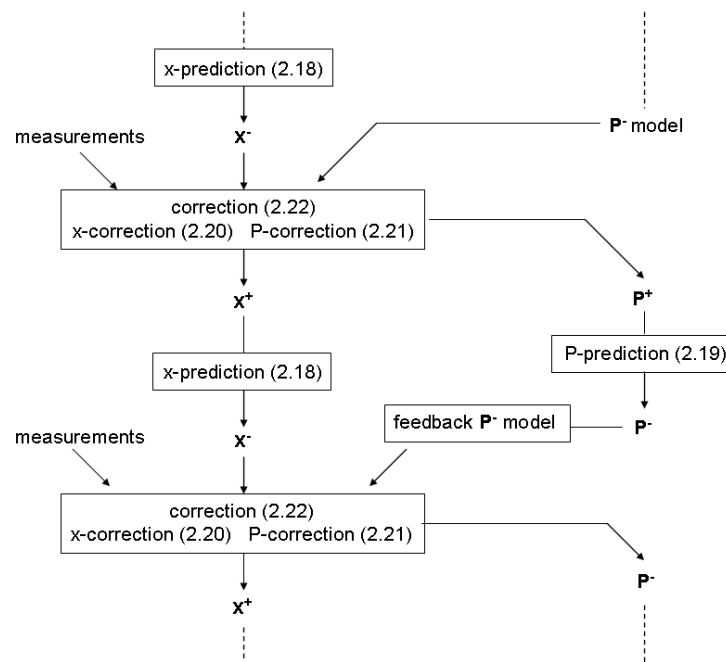


Figure 2.13: The organization of the computations in KF assimilation algorithm.

error covariance with respect to K_k , setting that to zero and then solving for K_k :

$$\begin{aligned}
P_k^+ &= E[e_k^+(e_k^+)^T] \\
&= E[(x_k - \hat{x}_k^- - K_k(z_k - H\hat{x}_k^-))(x_k - \hat{x}_k^- - K_k(z_k - H\hat{x}_k^-))^T] \\
&= E[(x_k - \hat{x}_k^- - K_k(Hx_k + v_k - H\hat{x}_k^-))(x_k - \hat{x}_k^- - K_k(Hx_k + v_k - H\hat{x}_k^-))^T] \\
&= E[((I - K_kH)(x_k - \hat{x}_k^-) + (K_kv_k))((I - K_kH)(x_k - \hat{x}_k^-) + (K_kv_k))^T] \\
&= (I - K_kH)E[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T](I - K_kH)^T + K_kE[v_kv_k^T]K_k^T \\
&\quad + 2(I - K_kH)E[(x_k - \hat{x}_k^-)v_k^T]K_k^T \\
&= (I - K_kH)P_k^-(I - K_kH)^T + K_kR_kK_k^T,
\end{aligned} \tag{2.23}$$

$$\begin{aligned}
\frac{\partial P_k^+}{\partial K_k} &= \frac{\partial}{\partial K_k}[(I - K_kH)P_k^-(I - K_kH)^T + K_kR_kK_k^T] \\
&= -2P_k^-H^T + 2K_k(HP_k^-H^T + R_k),
\end{aligned} \tag{2.24}$$

$$\begin{aligned}
K_k &= P_k^-H^T(HP_k^-H^T + R_k)^{-1} \\
&= \frac{P_k^-H^T}{(HP_k^-H^T + R_k)^{-1}}.
\end{aligned} \tag{2.25}$$

K_k determines how much of the information from the measurement should be taken into account when updating the state estimate. As the measurement error covariance R approaches zero, it implies the measurement will be taken into account (Eq. 2.26). If the *a priori* estimate error covariance P_k^- is close to zero, which means there is almost no uncertainty in the last state estimate, it results in that the measurement is not taken into great account (Eq. 2.27):

$$\lim_{R_k \rightarrow 0} K_k = H^{-1}, \tag{2.26}$$

$$\lim_{P_k^- \rightarrow 0} K_k = 0. \quad (2.27)$$

2.4.3 The EKF and EnKF

The traditional Kalman filter addresses the problem of estimating the state of a noisy system that can be described by a linear system model and a linear measurement model (Gelb, 1974). For nonlinear dynamic system consisting of a nonlinear system model and a nonlinear measurement model the Extended Kalman Filter (EKF) may be applied. The two Eqs. 2.10 and 2.11 can then be written as

$$x_k = f(x_{k-1}) + w_{k-1}, \quad (2.28)$$

$$z_k = h(x_k) + v_k, \quad (2.29)$$

where $f(\cdot)$ and $h(\cdot)$ are a nonlinear system function and a nonlinear measurement function. To predict the error statistics, an approximate linearized equation is used. A simple closure scheme is applied where third- and higher-order moments in the error covariance evolution equation are discarded. This closure technique results in an unbounded error variance growth caused by the linearization performed when higher-order moments are neglected (Evensen, 2003) (Evensen, 1994). Since the computational demand resulting from the error covariance integration limits the size of the problem, the EKF has been widely used in the application of modeling vertical soil moisture profiles (Crosson et al., 2002).

The Ensemble Kalman Filter (EnKF) is an alternative to the EKF. It has been examined and successfully applied in a number of study fields since it was first introduced by Evensen (Evensen, 2003) (Evensen, 1994). The EnKF takes a statistical approach to the solution of the basic KF equations for the covariance matrices of analysis and

background error. The EnKF avoids the expensive integration of the state error covariance matrix by propagating an ensemble of states from which the required covariance information is obtained at the time of the update. To generate a statistical sample of analysis, EnKF runs the analysis system several times for a given date, each time using backgrounds which differ by an amount of background error, and observations subject to observation error.

The major difference between EKF and EnKF is the EKF uses a linearized equation for the error covariance propagation while the EnKF propagates a finite ensemble of model trajectories. Other differences include: the EnKF can account for a wider range of model errors; the EnKF is easy to implement and more efficient in computation; the EKF cannot account for horizontal correlations in the system or measurement error for computational reasons.

2.5 Study sites

2.5.1 Conesus Lake

Conesus Lake, the westernmost Finger Lake, ($46^{\circ}54'$ N, $77^{\circ}43'$ W) of New York State, is 12.6 km long and 1.06 km wide, with the long axis oriented north and south. The lake has a surface area of about 13.4 km² and is 249 m above sea level (Bloomfield, 1978). Although the maximum depth of the lake is about 20.2 m, less than 6% of the lake's volume is deeper than 13.7 m, with a mean depth 11.5 m. The topography of the lake's watershed is characterized by gentle slopes in the north and steep slopes in the middle and southern portions of the lake (Figure 2.14). The outlet is at the north end, and the perimeter of the lake is surrounded by small streams that bring agricultural runoff to the lake from the surrounding landscape, including farms, woodlots, and vacation

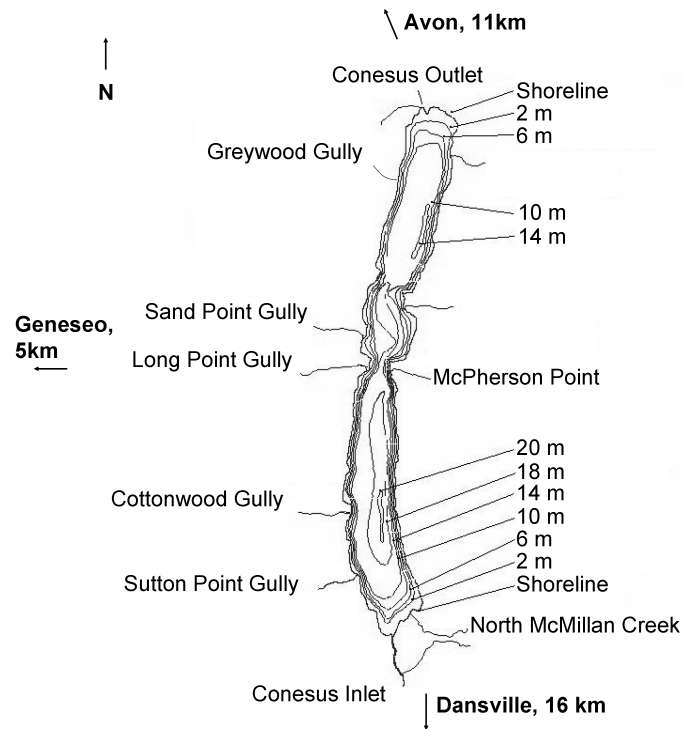


Figure 2.14: Bathymetry map (m) of Conesus Lake with stream gullies (Forest et al., 1978). The southern portion is deeper than the northern portion. Conesus Inlet, North McMillan Creek, Sutton Point Gully, Cottonwood Gully, Long Point Gully, Sand Point Gully, Greywood Gully, and McPherson Point were the eight tributary mass sources. Conesus Outlet was balanced to the mass sources.

homes that ring the lake. The major hydrologic inputs to the lake are Conesus Inlet and North McMillan Creek at the south end of the lake, which contribute as much as 70% of the water flowing into the lake. Conesus Lake is a dimictic, eutrophic lake covered by ice from late December until late March in most years (Makarewicz, 2001).

Studies conducted at the State University of New York (SUNY) at Brockport have demonstrated that nutrient loading during hydrometeorological events in Conesus Lake contributes massive amounts of phosphorus and nitrate to the lake in short periods of

time (Makarewicz et al., 2003). One consequence of the excess nutrients is a narrow but dense band of macrophytes and metaphyton (D'Auito et al., 2004) that grows along nearly the entire perimeter of the lake. In some shallow areas, particularly near streams, the macrophytes form expansive beds that cover most of the surface and bottom areas, and have become a distinct feature of Conesus Lake in recent years. Research at SUNY Geneseo suggests that nutrient input from agricultural runoff may be the single greatest contributor for macrophyte growth in Conesus Lake (D'Auito et al., 2004). However, the macrophyte beds are usually offset from the stream mouths. Areas nearest to streams typically have lower biomass, while the highest biomass is usually found near the middle of macrophyte bed which is often more than 100 m away from the stream source. Thus, studying the relationship between the macrophyte bed locations and stream sites and the effect of macrophytes on flow currents and material distribution in Conesus Lake is important for understanding the response of the lake to management effects in the watershed.

2.5.2 Lake Ontario

Lake Ontario is the smallest and easternmost of the Great Lakes with a surface area of 18,960 km². However, it has the highest ratio of watershed area to lake surface area. The lake is relatively deep, with an average depth of 86 m and a maximum depth of 244 m. Lake Ontario lies 99 m below Lake Erie, at the base of Niagara Falls. Approximately 80% of the water flowing into Lake Ontario comes from Lake Erie through the Niagara River. The remaining flow comes from Lake Ontario basin tributaries and precipitation. About 93 % of the water in Lake Ontario flows out to the St. Lawrence River and the remaining leaves through evaporation. Figure 2.15 shows the slope of the bottom topography is steep along the south shore and smooth along the

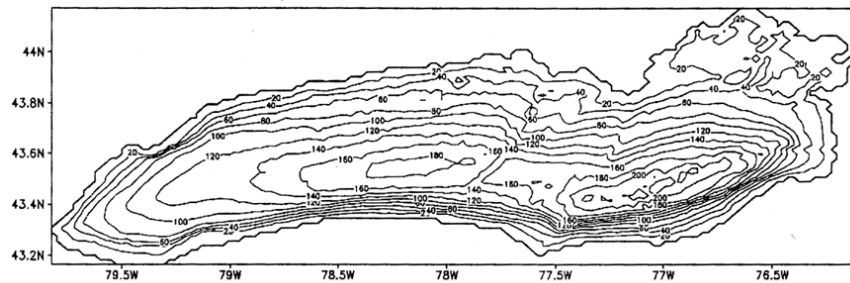


Figure 2.15: Lake Ontario bathymetry (Chang, 2003).

north shore. The freezing period of Lake Ontario is generally from middle January to early April with 15% ice coverage. The thermal bar forms in middle April to middle June, isolating the nearshore/offshore exchange of pollutants and thus affecting the nearshore water quality. The thermocline is usually formed from early June to middle October.

2.5.3 Niagara River mouth

The inflows play a significant role in introducing and redistributing substances in a large lake and also affect the lake's dynamic processes and water quality. Niagara River, estimated at $7000 \text{ m}^3/\text{sec}$, is the major inflow to Lake Ontario. Niagara River carries approximately 4.7 million tons of sediment into Lake Ontario annually, which represents 52% of the total sediment load to the lake. Water-quality analyses of the Niagara River have shown high concentrations of toxic chemicals that are introduced into Lake Ontario (Murthy and Miners, 1992).

The Niagara River is usually 3 to 4 °C warmer than the ambient water in Lake Ontario due to the high surface to volume ratio of Lake Erie. A buoyant thermal plume forms when the Niagara River discharges into Lake Ontario during certain times of

the year (Masse and Murthy, 1990). Much research has been conducted to investigate the mixing characteristics of the Niagara River plume in Lake Ontario (Murthy and Miners, 1992). In the work of Murthy and Miners, three physical experiments were performed to map the nearfield and farfield extent of the Niagara River plume in Lake Ontario. The mapping of the nearfield characteristics of the Niagara River plume was carried out by the conventional Lagrangian drifter experiments and electronic bathythermography thermal surveys. To map the farfield characteristics of the Niagara River plume, tracking Lagrangian drifting buoys by satellite over long periods of time was used. The data obtained from three different experiments suggest a three stage mixing process. In the initial stage, horizontal velocities from the Niagara River mouth are reduced significantly, and the river water is vertically well mixed over the shallow bar area. In the transition phase, a large clockwise eddy of between 10 km and 12 km in diameter is formed to the east of the Niagara River mouth. The eddy appears often and lasts for a few days (Figure 2.16). In the final stage, the river plume is deflected in response to lakewide circulation and the prevailing winds. In most circumstances, a plume develops from the Niagara River mouth and tends to extend eastward along the south shore of the lake (Figure 2.17). This nearshore area can be a zone in which fine particulate material is deposited.

2.5.4 Rochester Embayment

The Rochester Embayment is an area of Lake Ontario formed by the indentation of the Monroe County shoreline between Bogus Point (Parma, NY) and Nine Mile Point (Webster, NY). The northern boundary of the embayment is delineated by the straight line between these two points. The southern boundary includes approximately 9.6 km of the Genesee River. The warm turbid water of the Genesee River discharges into

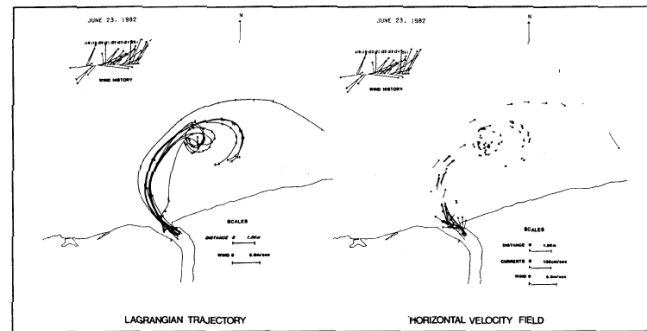


Figure 2.16: Nearfield horizontal distribution of Niagara River Plume on June 23, 1982 (Murthy and Miners, 1992).

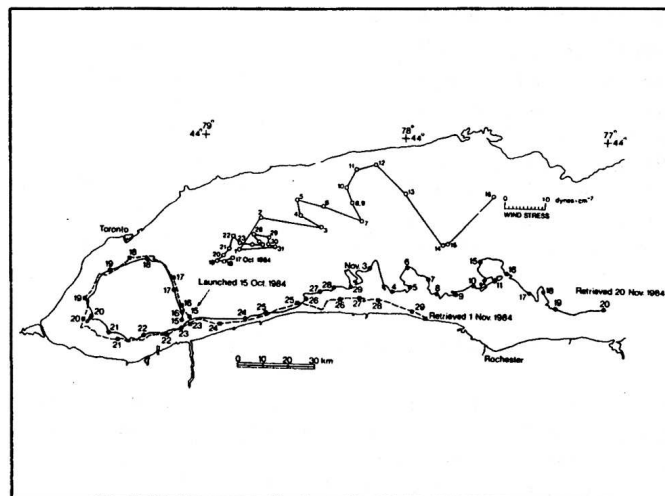


Figure 2.17: Farfield Lagrangian drifter trajectories from October 15 to November 20, 1984 showing the spatial extent of the Niagara River Plume (Murthy and Miners, 1992).

Lake Ontario at the Charlotte Pier (Figure 2.18). Rochester Embayment is one of the EPA areas of concern due to the point and non-point pollution. Ontario beach, just west of the mouth of the Genesee River, has been closed for one-quarter to one-half of the beach season due to the poor water clarity resulting from the Genesee river storm-water, the related bacteria level, and the excessive algae.

Suspended sediment concentrations in the Genesee River are quite high, although they remain less than 30 g/m^3 for at least 80% of a year, and exceed 200 g/m^3 for no more than 20 days due to highly energetic storms (Knauf, 2003). The process of advection and the dispersion of sediments and nutrients between the Genesee River system and Lake Ontario are complicated. The Genesee River plume usually swings either to the east or to the west after exiting from the port of Rochester jetty. The prevailing wind direction and subsurface circulation conditions are the two factors affecting the direction of the plume. Typically, the eastward track of the plume is usually seen during summer months due to the predominant winds from the west and pushes the river water away from the beach. However, easterly wind and westerly lake currents push potentially polluted river water onto the beach about 33% of the time (Knauf, 2003).



Figure 2.18: Turbid Genesee River plume flowing into Lake Ontario. Locations of Ontario beach and the USGS water quality site are shown as well.

Chapter 3

Model Review

3.1 ALGE: Hydrodynamic model

3.1.1 Governing equations

ALGE is a three-dimensional, time-dependent, hydrodynamic model that is capable of realistically predicting the movement and dissipation of stream plumes and the transport, diffusion and deposition of materials. ALGE has been extensively used to simulate and tested against flow, temperature, and tracer data from a variety of surface water systems (Garrett and Hayes, 1997) (Garrett et al., 2005). ALGE uses a finite differencing technique to solve the hydrostatic form of the partial differential equations that model conservation of momentum, mass and thermal energy. A module for predicting resuspension, transport, and deposition of materials is included in ALGE. Atmospheric energy exchange is modeled through turbulent sensible and latent heat transfer, including the effects of clouds. The amount of heat entering or leaving a body of water due to evaporation, convection, solar radiation absorption and thermal radiation flux

divergence is computed. ALGE is designed to produce high-resolution simulations for node-to-node matching with airborne and satellite imagery (Garrett, 2002) (Garrett and Hayes, 1997) (Garrett et al., 2000). The set of governing hydrostatic equations in a Cartesian coordinate system for ALGE model are given as (Garrett, 1995):

- Momentum conservation

$$\begin{aligned} \frac{\partial u}{\partial t} = & -\frac{\partial uu}{\partial x} - \frac{\partial vu}{\partial y} - \frac{\partial wu}{\partial z} - \frac{1}{\rho} \frac{\partial p}{\partial x} + fv \\ & + \frac{\partial}{\partial x} \left(K_H \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_H \frac{\partial u}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_H \frac{\partial u}{\partial z} \right) - C_d U u / \Delta z, \end{aligned} \quad (3.1)$$

$$\begin{aligned} \frac{\partial v}{\partial t} = & -\frac{\partial uv}{\partial x} - \frac{\partial vv}{\partial y} - \frac{\partial wv}{\partial z} - \frac{1}{\rho} \frac{\partial p}{\partial y} - fu \\ & + \frac{\partial}{\partial x} \left(K_H \frac{\partial v}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_H \frac{\partial v}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_H \frac{\partial v}{\partial z} \right) - C_d U v / \Delta z, \end{aligned} \quad (3.2)$$

$$\frac{\partial p}{\partial z} = -\rho g, \quad (3.3)$$

- Mass conservation

$$\frac{\partial w}{\partial z} = -\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}, \quad (3.4)$$

- Energy conservation

$$\begin{aligned} \frac{\partial T}{\partial t} = & -\frac{\partial uT}{\partial x} - \frac{\partial vT}{\partial y} - \frac{\partial wT}{\partial z} \\ & + \frac{\partial}{\partial x} \left(K_H \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_H \frac{\partial T}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_H \frac{\partial T}{\partial z} \right) + Q, \end{aligned} \quad (3.5)$$

where u , v and w are the velocity components, T is the temperature, Q represents any energy coming into or out of water body, p is the hydrostatic pressure, K_H is the

horizontal diffusion coefficient, f is the Coriolis parameter evaluated at the central latitude of the lake. The water temperature T (°C) is related to density ρ (kg/m³) by a quadratic fit (Garrett, 2002):

$$\rho = 0.0000161T^3 - 0.598T^2 + 0.0219T + 999.97. \quad (3.6)$$

Wind blowing at the surface constitutes a very important driving force for currents by transferring momentum from the wind to the water and pushes water downwind. Time-dependent wind stress is calculated as a function of the mean wind speed and direction. The bottom drag coefficient, C_D , is spatially variable. For areas of the lake bottom without macrophytes it is calculated as a function of roughness length by the logarithmic law $C_D = k^2[\ln(z_1/z_0)]^2$, where k is von Karman's constant and z_1 is the first node above the lake bottom. The roughness length z_0 is set to 0.001 m for all simulations in this research. Drag is also generated when water moves through vegetation, which removes kinetic energy and momentum from the flow. For areas of the lake bottom with macrophytes the drag term was modified and calculated as a function of leaf density $leafden$ (m² of macrophyte per m³ of water) by $C_D = a * leafden$ for dense macrophyte beds in the lake, where a is constant (Fischer-Antze et al., 2001). The leaf density values were derived from methods published by Gerber and Les (Gerber and Les, 1994). Eqs. (3.7) - (3.9) are vertically integrated forms of the momentum and mass conservation to predict the movement of a water body with a free surface.

- Vertically integrated momentum

$$\begin{aligned} \frac{\partial uh}{\partial t} = & -\frac{\partial u^2 h}{\partial x} - \frac{\partial uvh}{\partial y} + h \frac{\partial}{\partial x} (K_H \frac{\partial u}{\partial x}) + h \frac{\partial}{\partial y} (K_H \frac{\partial u}{\partial y}) + fvh \\ & - \frac{g}{\rho_0} \left(\frac{h^2}{2} \frac{\partial h}{\partial x} + \rho_0 h \frac{\partial h}{\partial x} + \rho_0 h \frac{\partial z_g}{\partial x} \right) \\ & - C_D U u + C_s \rho_a W^2 \cos \Psi / \rho_w, \end{aligned} \quad (3.7)$$

$$\begin{aligned} \frac{\partial vh}{\partial t} = & -\frac{\partial v^2 h}{\partial y} - \frac{\partial uvh}{\partial x} + h \frac{\partial}{\partial x} (K_H \frac{\partial v}{\partial x}) + h \frac{\partial}{\partial y} (K_H \frac{\partial v}{\partial y}) - fuh \\ & - \frac{g}{\rho_0} \left(\frac{h^2}{2} \frac{\partial h}{\partial y} + \rho_0 h \frac{\partial h}{\partial y} + \rho_0 h \frac{\partial z_g}{\partial y} \right) \\ & - C_D U v + C_s \rho_a W^2 \sin \Psi / \rho_w, \end{aligned} \quad (3.8)$$

- Vertically integrated mass

$$\frac{\partial h}{\partial t} = -\frac{\partial uh}{\partial x} - \frac{\partial vh}{\partial y} + M_s, \quad (3.9)$$

where h is water depth, $U = (u^2 + v^2)^{1/2}$, W is the wind speed, C_D and C_s are bottom drag coefficient and air-water surface drag coefficient, respectively. As noted above, drag is also generated when water moves through vegetation, which removes kinetic energy and momentum from the flow and results in reducing local flow velocities and enhancing the sediment deposition rate.

Mixing

The horizontal eddy diffusivities are related to three components which refer to turbulent mixing as a result of bottom roughness, horizontal velocity shear and buoyancy forces. The third component is important when the current velocities are low and the water movement mainly depends on the buoyancy forces. ALGE currently uses a

simplified version of the Yamada-Mellor closure model to compute vertical eddy diffusivity (Yamada, 1983) (Garrett, 1983) (Mellor and Yamada, 1982). A prognostic equation is solved for the turbulent kinetic energy:

$$\begin{aligned} \frac{\partial E}{\partial t} = & -\frac{\partial u E}{\partial x} - \frac{\partial v E}{\partial y} - \frac{\partial w E}{\partial z} + \frac{\partial}{\partial x} \left[K_H \frac{\partial E}{\partial x} \right] + \frac{\partial}{\partial y} \left[K_H \frac{\partial E}{\partial y} \right] + \frac{\partial}{\partial z} \left[K_Z \frac{\partial E}{\partial z} \right] \\ & + K_Z \left[\left(\frac{\partial u}{\partial z} \right)^2 + \left(\frac{\partial v}{\partial z} \right)^2 \right] + g K_Z \frac{1}{\rho} \frac{\partial \rho}{\partial z} - \frac{(2E)^{1.5}}{c_3 l}, \end{aligned} \quad (3.10)$$

where E is the turbulent kinetic energy, c_3 is a constant, and the turbulent length scale l defined by $l = \min(\Delta z, l_s)$. Δz is the vertical grid spacing, and l_s is turbulent length scale for a stably stratified condition $l_s = 0.76 E^{0.5} / \omega_b$, where ω_b is the Brunt-Vaisala frequency. The vertical eddy diffusivity is defined as $K_Z = l(2E)^{0.5} S_m$, where S_m is a function that accounts for the effects of stable density stratification on turbulence (Yamada, 1983).

Material transport

The transport equations for a generic scalar variable C_d , (C_p or C_ϕ), are obtained by considering a balance between convection and diffusion and defined as:

- Dissolved tracer

$$\begin{aligned} \frac{\partial C_d}{\partial t} = & -\frac{\partial u C_d}{\partial x} - \frac{\partial v C_d}{\partial y} - \frac{\partial w C_d}{\partial z} + \frac{\partial}{\partial x} \left[K_H \frac{\partial C_d}{\partial x} \right] + \frac{\partial}{\partial y} \left[K_H \frac{\partial C_d}{\partial y} \right] \\ & + \frac{\partial}{\partial z} \left[K_Z \frac{\partial C_d}{\partial z} \right] - \alpha (C K_D C_d - C_p) + S_d, \end{aligned} \quad (3.11)$$

- Particulate tracer

$$\begin{aligned} \frac{\partial C_p}{\partial t} = & -\frac{\partial u C_p}{\partial x} - \frac{\partial v C_p}{\partial y} - \frac{\partial w(w - w_s) C_p}{\partial z} + \frac{\partial}{\partial x} \left[K_H \frac{\partial C_p}{\partial x} \right] + \frac{\partial}{\partial y} \left[K_H \frac{\partial C_p}{\partial y} \right] \\ & + \frac{\partial}{\partial z} \left[K_Z \frac{\partial C_p}{\partial z} \right] - \alpha(C_p - CK_D C_d) + S_p, \end{aligned} \quad (3.12)$$

- Sediment

$$\begin{aligned} \frac{\partial C_\phi}{\partial t} = & -\frac{\partial u C_\phi}{\partial x} - \frac{\partial v C_\phi}{\partial y} - \frac{\partial w(w - w_s) C_\phi}{\partial z} + \frac{\partial}{\partial x} \left[K_H \frac{\partial C_\phi}{\partial x} \right] \\ & + \frac{\partial}{\partial y} \left[K_H \frac{\partial C_\phi}{\partial y} \right] + \frac{\partial}{\partial z} \left[K_Z \frac{\partial C_\phi}{\partial z} \right] - \delta_k(E_\phi - D_\phi)/\Delta z_b, \end{aligned} \quad (3.13)$$

where K_H and K_Z are diffusivity coefficients for the transported quantity in each coordinate direction, E_ϕ and D_ϕ are the resuspension flux and deposition flux, and w_s is the settling velocity.

Energy transfer

The heat transfer at the water surface in ALGE contains an explicit balance of short-wave solar radiation, long-wave atmospheric radiation, long-wave radiation emitted from the water surface, sensible heat transfer, and latent heat transfer between water and air. Short-wave radiation is reduced by cloud cover and adds heat at the surface and within a lake. Lakes lose heat via long-wave radiation, but also receive heat via long-wave radiation from the air and clouds. Also, the sensible heat transfer is driven by the temperature difference between lake water and adjacent air and latent heat transfer is driven by the relative humidity of the air. The energy exchange between the water body and atmosphere is specified as:

$$\frac{dT}{dt} = \frac{(H_s + H_L + S_w + L_w)}{(\Delta Z_s \rho_w c_{pw})} \quad (3.14)$$

where H_s , H_L , S_w , and L_w are the sensible heat, latent heat, solar radiation and thermal radiation fluxes, respectively. ΔZ_s is surface layer depth and c_{pw} is specific heat of water.

3.1.2 Numerical scheme

Finite differencing

In general, Partial Differential Equations (PDE) are much more difficult to solve analytically than Ordinary Differential Equations (ODE). To obtain analytical solutions to PDEs, numerical methods such as finite differences are needed. The governing equations presented above are solved numerically to obtain the velocity field and temperature distribution.

If a flow is described by a PDE of the form:

$$\frac{\partial p}{\partial t} = -v \frac{\partial p}{\partial x}, \quad (3.15)$$

the term on the right-hand-side is called the advection (or streaming, drift), depending upon the field of application. In early works for Computational Fluid Dynamics (CFD), a centered difference, which is accurate to second order, is used to discretize this as

$$\frac{\partial p}{\partial t}[i] = -v \frac{(p[i+1] - p[i-1]))}{\Delta x}. \quad (3.16)$$

Second-order finite difference scheme is applied in ALGE. The advection terms in ALGE are solved by the conservative and transportive upwind differencing scheme. Eqs. (3.1) - (3.5) were transformed directly into finite differencing equations, whereas

Eqs. (3.7) - (3.9) were rewritten as Eqs. (3.18) - (3.20) using the following definitions:

$$\begin{aligned} Q_x &= uh \\ Q_y &= vh, \end{aligned} \quad (3.17)$$

$$\begin{aligned} \frac{\partial Q_x}{\partial t} &= -\frac{\partial Q_x^2/h}{\partial x} - \frac{\partial Q_x Q_y/h}{\partial y} + h \frac{\partial}{\partial x} (K_H \frac{\partial Q_x/h}{\partial x}) + h \frac{\partial}{\partial y} (K_H \frac{\partial Q_x/h}{\partial y}) + f Q_y \\ &\quad - \frac{g}{\rho_0} \left(\frac{h^2}{2} \frac{\partial h}{\partial x} + \rho_0 h \frac{\partial h}{\partial x} + \rho_0 h \frac{\partial z_g}{\partial x} \right) \\ &\quad - C_D U Q_x/h + C_s \rho_a W^2 \cos \Psi / \rho_w, \end{aligned} \quad (3.18)$$

$$\begin{aligned} \frac{\partial Q_y}{\partial t} &= -\frac{\partial Q_x Q_y/h}{\partial x} - \frac{\partial Q_y^2/h}{\partial y} + h \frac{\partial}{\partial x} (K_H \frac{\partial Q_y/h}{\partial x}) + h \frac{\partial}{\partial y} (K_H \frac{\partial Q_y/h}{\partial y}) - f Q_x \\ &\quad - \frac{g}{\rho_0} \left(\frac{h^2}{2} \frac{\partial h}{\partial x} + \rho_0 h \frac{\partial h}{\partial x} + \rho_0 h \frac{\partial z_g}{\partial x} \right) \\ &\quad - C_D U Q_y/h + C_s \rho_a W^2 \sin \Psi / \rho_w, \end{aligned} \quad (3.19)$$

$$\frac{\partial h}{\partial t} = -\frac{\partial Q_x}{\partial x} - \frac{\partial Q_y}{\partial y} + M_s. \quad (3.20)$$

After Q_x , Q_y and h are solved, u and v are computed from the updated Q through the updated values of h .

One dimensional advection equation and its finite difference form are defined as

$$\frac{\partial Q_x}{\partial t} = -\frac{\partial Q_x Q_x/h}{\partial x}, \quad (3.21)$$

$$\frac{\Delta Q_i}{\Delta t} = -\frac{(u_r Q_r - u_l Q_l)}{\Delta x}, \quad (3.22)$$

where

$$\begin{aligned} u_r &= \frac{(Q_{i+1}/h_{i+1} + Q_i/h_i)}{2} \\ u_l &= \frac{(Q_i/h_i + Q_{i-1}/h_{i-1})}{2}, \end{aligned} \quad (3.23)$$

and

$$\begin{aligned} Q_r &= Q_i, \text{ if } u_r > 0 \\ Q_r &= Q_{i+1}, \text{ if } u_r < 0 \\ Q_l &= Q_{i-1}, \text{ if } u_l > 0 \\ Q_l &= Q_i, \text{ if } u_l < 0. \end{aligned} \quad (3.24)$$

Also, to force the conservation, $u_r = 0$ if the $i+1$ node is land and $u_l = 0$ if the $i-1$ node is land. Instead of upwind differences, centered differences are used for solution of the mass conservation Eqs. (3.4) and (3.20).

The upwind differencing scheme is stable provided that the Courant-Friedrichs-Lewy (CFL) stability criterion is satisfied. The CFL condition on the vertically integrated hydrodynamic equations limits the time step to:

$$\Delta t = [(gH)(\Delta x^{-2} + \Delta y^{-2})]^{-\frac{1}{2}}. \quad (3.25)$$

Boundary conditions

Applying appropriate boundary conditions is the principle difficulty in hydrodynamic modeling of the open water body. There are three different options for lateral boundary conditions in ALGE. The boundary could have a zero-gradient for temperature and the normal velocity component that is directed inward from the boundary.

Secondly, the zero-gradient boundary conditions could be used everywhere. In addition, the temperature and velocity could be fixed at the boundary throughout the simulation at their initial values. To create a land-water distribution array, all three different types of boundary conditions can be mixed as long as the mixture is physically realistic.

Nudging

To accurately track a stream plume in a limited area and generate realistic results, the procedure known as nudging is applied in this study (Garrett, 1995). Nudging uses the large scale solution to drive a high resolution, limited area simulation of the plume. The simulation of the entire lake generates a time series of current velocities which are applied as a forcing term near the boundary of the limited area simulation. The additional nudging terms include a weighting function that decreases the magnitude of the nudging term at nodes close to shore. The weights are zero close to shore and then increase to 1.0 about halfway between the shoreline and the offshore boundary.

3.1.3 Inputs

The inputs to the ALGE include surface meteorological data, upper air meteorological data, land-water distribution, lake bathymetry, and miscellaneous data, such as grid size, number of nodes, etc. The tidal forcing is needed when one models estuarine. The nudging inputs and the corresponding weight function are required for limited-area simulation.

Surface meteorological data

ALGE uses hourly surface meteorological data including: wind direction (deg), wind speed (m/s), air temperature (K), dewpoint temperature (K), cloud cover (0.0 to 1.0), cloud height (km), and pressure (mb).

Upper air meteorological data

The hourly upper air data at specific heights above the lake surface includes temperature (K) and precipitable water (cm). There are 11 levels for upper air data with the first observational height assumed to be 100 m and then 1 km, 2 km and so on.

Land-water distribution

A 2-D array is used to identify the nodes that represent land (0's) and nodes that represent water (1's). All non-zero numbers are assumed to be water in ALGE. However, three types of special locations over water are: mass and energy sources (7's), mass and energy sinks (6's), and inflow/outflow boundaries (8's and 9's, depending on various boundary conditions).

Lake bathymetry

Water depths are read into ALGE via a 2-D array that contains one integer per node. Digits 1 through 9 and alphanumeric characters a through z define the number of levels at a horizontal location (x, y). The product of this number and the vertical grid spacing defines the real depths at location (x, y).

Other input parameters

The other input parameters include resolution of the bathymetric data, starting time (in local hours and Julian days), total simulation time (in hours), latitude, longitude, roughness, leaf density (macrophyte leaf surface area/unit volume, m^2/m^3), temperatures of boundary inflow/outflow and mass source, particle density and diameter, number of nodes in each direction, tracer distribution coefficient, and flags for time and parameter usage. Hourly discharge and hourly river source sediments and tracer concentration are used as dynamic variables to simulate the river plume dissipation in lakes. A full list of definitions and units for parameters and flags used in ALGE is given in Table 3.1.

Table 3.1: Definitions and units for parameters and flags used in the ALGE model.

Parameter definition	Unit
<i>Static variable</i>	
Horizontal grid spacing	m
Vertical grid spacing	m
Simulation time	hour
Latitude	degree
Longitude	degree
Julian day	unitless
Roughness	m
Water grass surface area/unit volume	m^2/m^3
Multiplier for Courant limit	unitless
Initial timestep	second
Inflow/Outflow temperature	$^{\circ}\text{C}$
Mass source temperature	$^{\circ}\text{C}$
Particle density	g/cm^3
Particle diameter	m
Distribution coefficient for tracer	unitless
Number of grids in x direction	unitless
Number of grids in y direction	unitless
<i>Dynamic variable</i>	
River source TSS concentration	g/m^3
River discharge	m^3/s
<i>Flag</i>	
Use of nudging data	0 = off, 1 = on
Heat transfer functions	0 = off, 1 = on
Outfall temperature	0 = fixed, 1 = delta-t
Restart option	0 = off, 1 = on
Turbulence model	0 = Yamada turbulence, 1=molecular viscosity
Fresh or salt water	0=fresh, 1=salt
Steady or time-varying mass source	0=steady-state, 1=time-varying

3.1.4 Model calibration

To assess the impact of grid resolution of ALGE simulations, grid independence study was carried out in Conesus Lake by running identical simulations with the only difference being the number of nodes in the grid. Validation of the hydrodynamic model was performed through the comparisons of the simulated temperature structure in Conesus Lake with the airborne thermal images and field data. ALGE was also employed to determine the physical dynamics of the spring thermal bar in Lake Ontario and investigate its impact on water-quality conditions during the spring season as well. Large scale circulations were studied in Conesus Lake and Lake Ontario. The impacts of river plumes on the lake hydrodynamics were studied at mouth of the Niagara River and two stream gullies in Conesus Lake. It was done by simulating the dispersion and redistribution mechanism of the river (or stream gullies) discharged mass and constituents.

3.2 HYDROLIGHT: Radiative transfer model

HYDROLIGHT 4.1 is a radiative transfer numerical model that obtains the radiance distribution within and leaving any plane-parallel water body by solving the time-independent radiative transfer equations (Mobley and Sundman, 2000b). The model is designed to solve a wide range of problems in optical oceanography and limnology. The radiative transfer regimes for the HYDROLIGHT model shown in Figure 3.1 include five modules: radiation in the air, transition from air to water, reflection under water, transition from water to air, and propagation to the sensor. Input to the model consists of the absorbing and scattering properties of the water body which can vary with depth and wavelength, the concentration profiles of water constituents, the

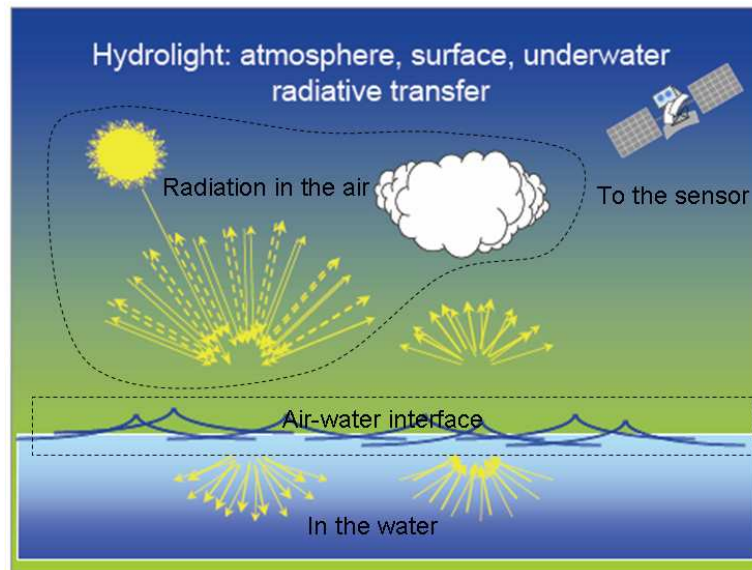


Figure 3.1: Radiative transfer model regimes: air, air-water interface, water, water-air interface, and air again (Fairbanks, 1999).

nature of the wind-blown water surface and of the bottom of the water column, and the sun and sky radiance incident on the water surface. The model also includes the effects of inelastic scatter by chlorophyll fluorescence, CDOM fluorescence, and Raman scattering by the water itself. HYDROLIGHT can also simulate internal layers of bioluminescing microorganisms.

3.2.1 Physical and mathematical design

Although the natural water body is horizontally inhomogeneous, the horizontal scales of significant optical variability are typically about tens of meters to kilometers that are much greater than the vertical scales (typically centimeters to tens of meters). In Hydrolight, the water body is treated as consisting of optically independent patches of water, so that each patch can be modeled as a horizontally homogeneous water volume

whose optical properties vary only with depth. To simulate the entire, horizontally inhomogeneous water body, a 1D radiative transfer model is applied independently at the center of each patch. Hydrolight can also be applied to the water patch associated with each pixel in the remote sensing imagery.

One physical consideration is that the horizontal size of each water patch needs to be at least several photon mean free paths. Photon mean free paths are defined as the inverse of the beam attenuation coefficient. The horizontal variability in the open ocean is usually on scales of several kilometers while the photon mean free paths are usually less than 50 m. The optical properties and boundary conditions can change horizontally on scales of meters to tens of meters in coastal water due to river runoff, sediment resuspension, and variable shallow bottom topography. However, the photon mean free paths of this type water are only about tens of centimeters to a few meters due to the increasing turbidity. In both cases the use of a 1-D radiative transfer model is still justified. In addition, the technique of time-independent radiative transfer is valid if the time required for the light field to reach a steady state in the water body after the optical properties or boundary conditions changing are less than the time scales for changes in environmental conditions.

The spectral radiance $L(z, \theta, \phi, \lambda)$ ($\text{W m}^{-2} \text{sr}^{-1} \text{nm}^{-1}$) describes the time-independent, one-dimensional light field within the water body. The spectral radiance is completely determined by the depth (z), direction (θ, ϕ), and wavelength (λ) behavior of the light field. Other variables, including various radiance, diffuse attenuation functions (K-functions), and remote sensing reflectance, can be computed through the spectral radiance based on the definitions.

The depth z is defined as positive downward from zero at the mean sea surface. The polar angle θ ($0 \leq \theta \leq 180^\circ$) is measured from zero in the nadir direction, and the

azimuthal angle ϕ ($0 \leq \phi \leq 360^\circ$) is measured from zero in the downwind direction. The sensor-reading radiance is the average of $L(z, \theta, \phi, \lambda)$ over the finite solid angle $\Delta\Omega$ and the bandwidth $\Delta\lambda$. The solid angle and bandwidth are determined by the field of view of the instrument and the wavelength response of the instrument. The directional averaging is performed by dividing all directions into small separate regions bounded by lines of constant θ and ϕ . The individual regions Q_{uv} (quads), are labeled by discrete indices $u = 1, 2, \dots, 20$ and $v = 1, 2, \dots, 24$ to show the θ and ϕ positions. Thus the resolutions of θ and ϕ are 10° and 15° , respectively. In order to solve the RTE numerically in Hydrolight model, we discretize it by averaging over direction and wavelength. The direction- and band-averaged radiances at any selected set of depth z_k , $k = 1, 2, \dots, K$, are computed in the Hydrolight:

$$L(k, u, v, j) = \frac{1}{\Delta\Omega_{uv}\Delta\lambda_j} \int_{Q_{uv}} \int_{\Delta\lambda_j} L(z_k, \theta, \phi, \lambda) \sin\theta d\theta d\phi d\lambda. \quad (3.26)$$

3.2.2 Simulation parameters

Numerical simulations of the remote sensing reflectance were carried out with Hydrolight (version 4.1) radiative transfer model. The spectrum of light from 400 nm to 700 nm in 10 nm intervals was applied to compute the results. The bioluminescence, chlorophyll fluorescence, CDOM fluorescence, and Raman scattering were included. The wind speed was assumed to be a constant 5 m/s to simulate calm but not completely smooth wave conditions. The illumination condition was based on the illumination in Rochester area during the summer near noon. The condition of the lake bottom was assumed to be covered with sand. It is feasible to compare the simulated remote sensing reflectance with MODIS derived remote sensing reflectance, as well as *in situ* measurements.

Chapter 4

Data Sources and Quality Control

4.1 Bathymetric grids

Conesus Lake

A coarse geo-referenced bathymetry data set of Conesus Lake has been integrated into the hydrodynamic model to study the lake surface temperature distribution and the circulation pattern. The whole lake simulations divide the lake into 70×232 control volumes. Horizontal grid spacing is uniform in this research. The grid spacing is 46.9 m in longitude and in latitude. In the z-direction (depth) there are a maximum of six levels with a resolution of 3.1 m. Figure 2.14 shows the lake bathymetry map and the tributary streams used as mass sources and mass sink for hydrodynamic simulations. Locations of weather stations providing meteorological data for simulations and distance away from the lake are given in Figure 2.14 as well.

A low cost bathymeter suitable for generation of bottom profile maps of small to mid-size lakes has been developed in the Digital Imaging and Remote Sensing (DIRS) Laboratory at the Rochester Institute of Technology (RIT). The device uses commer-

cial components for depth finding (Garmin Fishfinder 100 depth sounder) and positioning (Garmin GPS eTrex) combined with a data storage and management unit developed at DIRS. To obtain high-resolution bathymetry near a stream mouth, the unit was deployed from a canoe. In this research, the high-resolution bathymetry was re-sampled to a grid with a 2 m horizontal spacing and a 1 m vertical spacing for input to ALGE. Cottonwood Gully and Sand Point Gully were chosen as the study sites to carry out stream plume simulations (Figure 2.14).

Lake Ontario

The water depth of Lake Ontario incorporated with ALGE in this research is from the new bathymetry released by the National Geophysical Data Center (NGDC). The custom grids shown in Figure 4.1 were created using the GEOphysical DATA System (GEODAS) Grid Translator Design-a-Grid system. The whole lake was divided into 34×194 horizontal lattice points. The horizontal grid spacing is uniform, about 1.5 km in longitude and in latitude. There are 35 levels with a resolution of 4.4 m in the z-direction (depth).

The higher resolution bathymetric maps of the Niagara River mouth and the Rochester Embayment (Figure 4.1 (b) and (c)) were created with the same method. The horizontal grid spacing are 325 m and 135 m for the Niagara River mouth and the Rochester Embayment, respectively. The vertical spacing is 3 m for both cases. Figure 4.1 (b) shows the bathymetry at the river mouth includes an inner shallow shelf region with a depth of 20 m extending approximately 5 to 6 km from the shore followed by a shelf edge from 20 to 100 m depth at a distance of 10 km offshore. The topographic features affect the transport and mixing of the Niagara River plume in the lake.

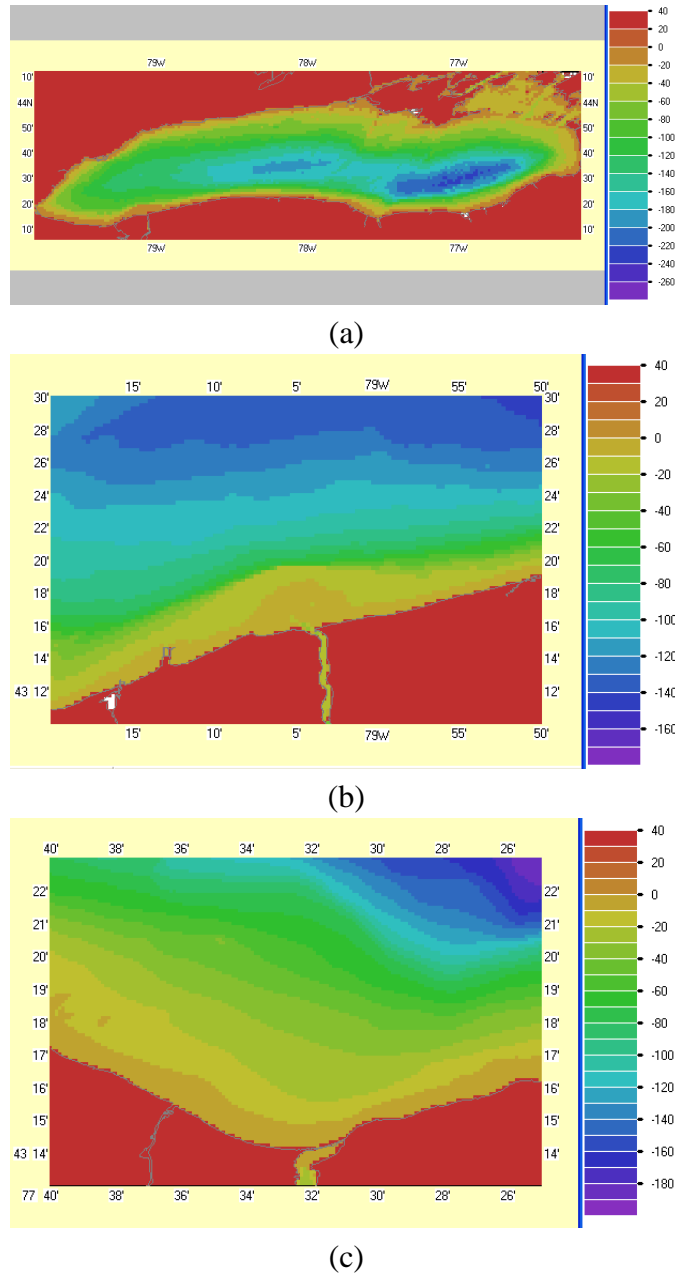


Figure 4.1: Bathymetric maps developed by the Grid Translator and Hydroplot of GEODAS (GEOphysical Data System). Depth (below lake level) and elevation (above sea level) are in meters. (a) Lake Ontario (b) Niagara River mouth (c) Genesee River mouth

4.2 Adjustment of forcing functions

To perform more accurate hydrodynamic simulations of the lake, forcing variables such as wind speed, dew point temperature, and air temperature are adjusted to include the effect of over land stations. Due to the much higher aerodynamic roughness overland compared to the much lower aerodynamic roughness over water, the wind speeds from overland stations (airports and other surface stations) were adjusted by an empirical method to obtain a fairly good agreement with overwater wind speeds. In this study, we applied the empirical overland-overwater wind speed adjustment given by (Resio and Vincent, 1977):

$$u_w = u_l F_1(u_l) F_2(\Delta T), \quad (4.1)$$

where $F_1(u_l) = 1.2 + 1.85/u_l (ms^{-1})$, $F_2 = 1 - (\Delta T/|\Delta T|)(|\Delta T|/1920)^{1/3}$, u_w is the overwater wind speed, u_l is overland wind speed, and $\Delta T = T_a - T_w (^\circ C)$.

To calculate the latent heat flux in the energy transfer model more realistically, the overwater dew point temperature was estimated from overland values by a simple empirical formula (Phillips and Irbe, 1978):

$$T_{dw} = T_{dl} - c_1(T_{dl} - T_w), \quad (4.2)$$

where $c_1 = 0.35$, T_{dw} and T_{dl} are the dewpoint temperatures of overwater and overland, respectively. Air temperature was also adjusted with an empirical formula (Beletsky and Schwab, 2001):

$$T_a = 0.4T_{al} + 0.6T_w, \quad (4.3)$$

where T_{al} and T_w are the air temperature overland and averaged water surface temper-

ature, and T_a is the air temperature over water.

Hourly meteorological data are available from the National Weather System (NWS). After the meteorological data were downloaded and adjusted to over water values, for the missing hours, the data from the very closest hours before and after were extracted and linearly interpolated. Triangular interpolation was applied to provide lake-wide mean meteorological forcing conditions.

Hourly meteorological data is available from three stations located in Livingston County around the Conesus Lake: Dansville Municipal Airport (DSV), Avon, and Geneseo. Conesus Lake is a long, narrow lake and the hill sides adjacent to the lake are generally forested with the agricultural land generally on the hill tops. To include the effect of trees on the wind speed over water, wind speed from overland stations was reduced to the square root of its original value. A time series of the daily forcing, including wind speed, wind direction, and inflow discharge, for April, May, and June 2004 is presented in Figure 4.2. Wind is stronger in early spring, while it is relatively weak in May and June, with an average speed of no more than 4 m/s. However, there are some occasions where wind speeds exceed 5 m/s, including May 20 and June 13, which occurred before storm events. Northwest and southeast winds appear to be the dominant wind directions during April to June 2004.

Meteorological data from the station at the Buffalo Airport (WBAN 14733) was applied to hydrodynamic simulations of Lake Ontario. Figure 4.3 shows a time series of daily wind direction and speed for April, May, June, and July, 2004. Wind is stronger in early spring and July, while it is relatively weak in May and June. The prevailing winds blow from the west over Lake Ontario. With the major axis in the west-east direction, the eastward wind can have a strong effect on lake circulation pattern.

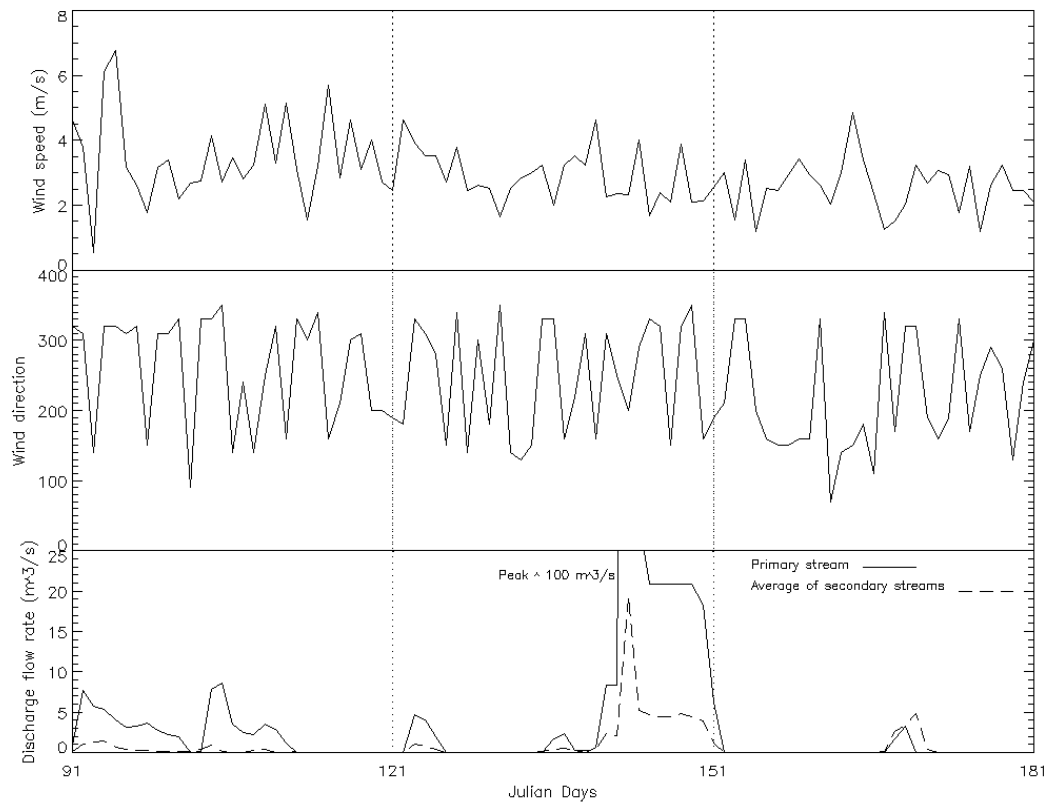


Figure 4.2: Time series of simulation forcing factor conditions of April, May, and June, 2004 from the stream monitoring program and weather stations. A major storm event is shown in late May lasting seven days with a peak flow rate for North McMillan creek of about $100 \text{ m}^3/\text{s}$.

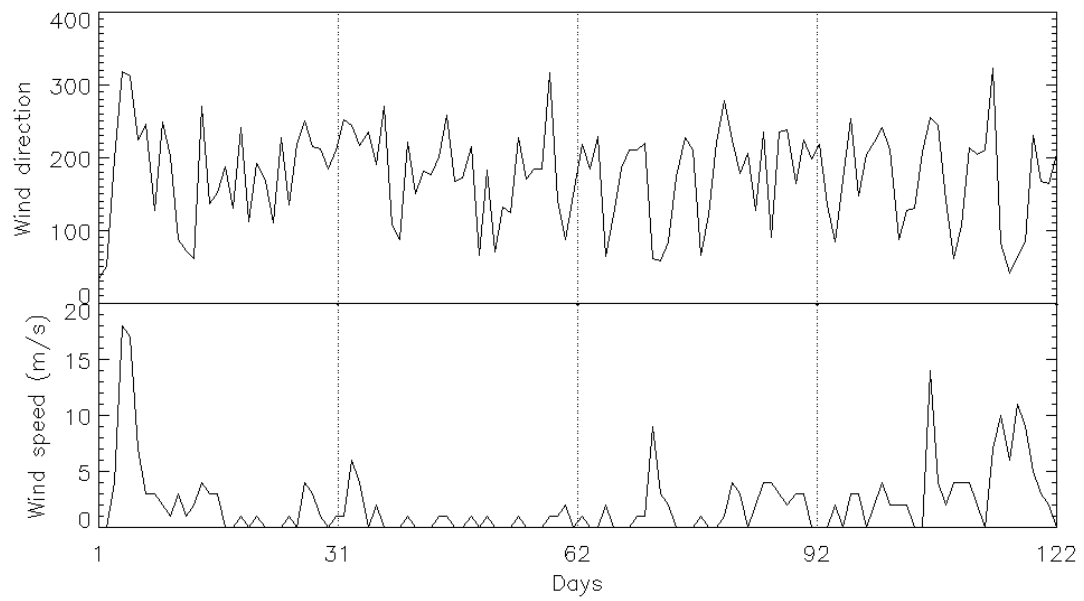


Figure 4.3: Time series of simulation forcing factor conditions of April, May, June, and July, 2004 from the Buffalo weather station.

4.3 *In situ* optical data collection and water sampling in the Rochester Embayment

The field optical measurements are performed to characterize the light field and the spectral IOPs of the aquatic medium. After an estimate of the overall absorption and scattering coefficients is given by field measurements, additional laboratory analysis is carried out to decompose these properties into the optically active components. The absorption and scattering spectra of each component determined in the lab will be the basic IOPs input to the radiative transfer model.

Field data were collected at several sites in the Rochester Embayment including both inside and outside of the plume on August 9, 2006. Figure 4.4 shows a summary of the sampling stations around the embayment. Stations 1 and 2 are plume-dominated and stations 3 and 4 are offshore. Vertical profiles of optical parameters were measured at each location and supplemented with water samples that were taken to the laboratory for more analysis.

A HydroRad-4 spectroradiometer was deployed in this study to measure the surface and sub-surface AOPs. Irradiance sensors are available and are usually placed in tandem. The HydroRad is an in-water profiling instrument measuring spectral downwelling and upwelling irradiance in the wavelength range of 400 ~ 700 nm. It was deployed away from the boat to reduce the background influences of the platforms on the measured light field. Remote sensing reflectance (R_{rs}) values were calculated as

$$R_{rs} = \frac{L_u}{E_d} = \frac{E_u}{QE_d}, \quad (4.4)$$

where E_u and E_d are the HydroRad measured upwelling and downwelling irradiance evaluated just above the surface for each wavelength. Q -factor is dependent on the

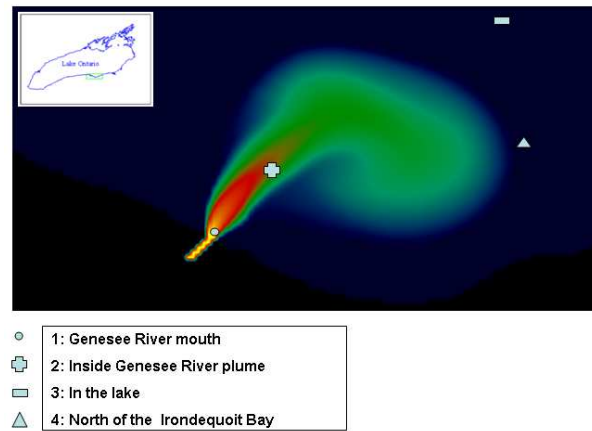


Figure 4.4: Overview map of the four stations. Background is an image of ALGE predicted TSS distribution showing the Genesee River plume flowing eastward.

solar zenith angle and the IOPs of the water column, specifically the volume scattering function and the single scattering albedo. Q is often considered to remain constant over wide geographical areas as well as over the visible spectrum. In this study Q was assumed to be a constant 5.5. Spectral remote-sensing reflectance contains information about the properties of the water body surface layer whose thickness depends on the inherent and apparent optical properties. Figure 4.5 shows the values of R_{rs} calculated by Eq. 4.4 for stations 2, 3, and 4.

Samples from the water bottles were taken for determination of TSS and CDOM concentrations and the absorption spectrum for TSS, CDOM, and CHL. Table 4.1 lists the average values of measured water quality parameters estimated through lab analysis on the water samples obtained from each of the stations. Suspended solid concentrations were determined by filtration through 0.4 micron pre-weighed filters which were dried and reweighed. The laboratory measured TSS concentrations minimum and maximum were 0.5 g/m^3 at station 3 and 16.7 g/m^3 at station 1. Since CDOM has

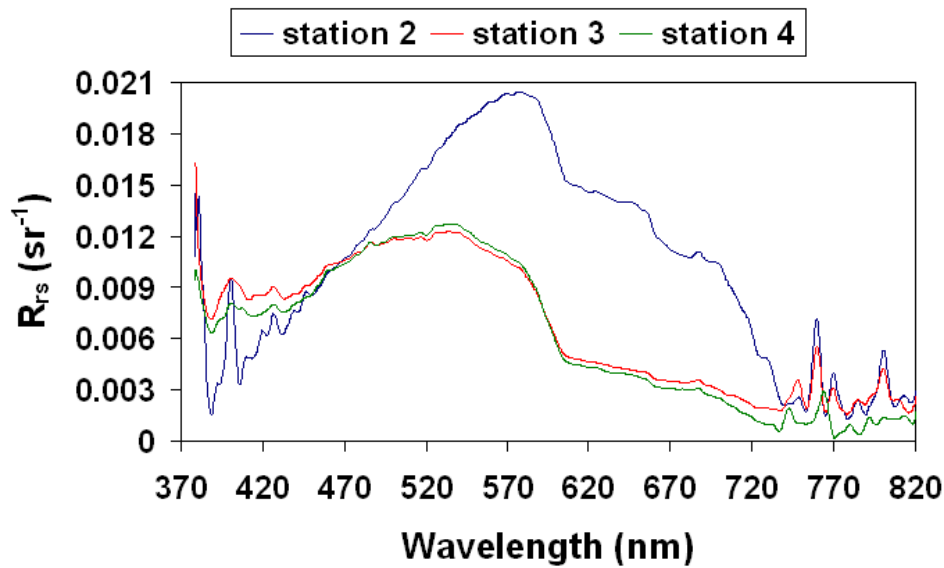


Figure 4.5: Measured remote sensing reflectance at three stations for the August 9th collections.

typically exponential absorption spectrum and an increase in the true CDOM concentration is likely to produce a proportional increase in the absorption spectrum, a value based on the mean CDOM absorption coefficient at 350 nm was used as an indirect measure of concentration. The laboratory measured minimum and maximum absorption coefficients were 0.163 m^{-1} at station 3 and 5.935 m^{-1} at station 1. Chlorophyll concentrations applied in this study were estimated from the water samples in 1999. The minimum and maximum of CHL concentration were 0.594 mg/m^3 for offshore Lake Ontario and 2.0 mg/m^3 for Genesee River mouth.

Figure 4.6 shows the measured absorption coefficient for CHL and TSS which is a mean of the measurements from all the stations. The chlorophyll absorption curve was scaled to 0.05 at 450 nm with the two absorption peaks occurring at 440 nm and 670 nm. TSS has maximum absorption in the blue region. CHL absorbance causes

Table 4.1: Water sampling estimates derived from laboratory measurements.

Location	Station 1	Station 2	Station 3	Station 4
TSS (g/m^3)	16.7	5.8	0.5	0.8
CDOM (scolor)	5.935	1.556	0.163	0.213

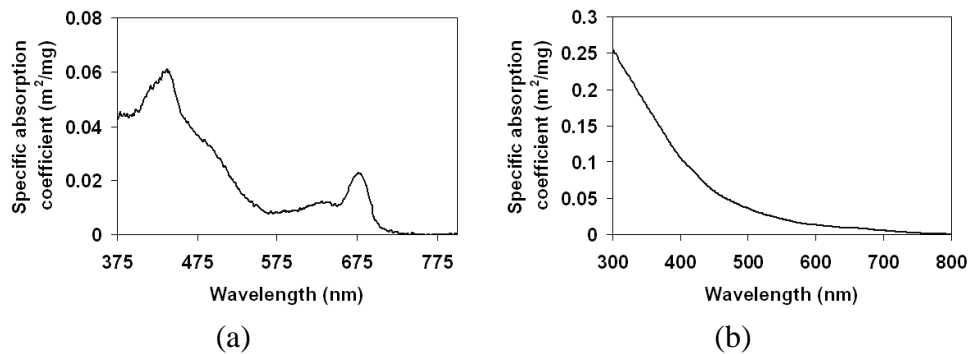


Figure 4.6: Specific absorption coefficient for CHL and TSS from measurements taken on August 9, 2006 in the Rochester Embayment.

a reflectance minimum near 440 nm. However, the reflectance in blue region has low sensitivity to CHL concentration because absorbance by CDOM and scattering by particles dominate the reflectance in this region. The absorption spectrum is an exponential function of wavelength, and most absorption occurs before 400 nm.

The specific scattering coefficient (m^2/mg) for CHL in Lake Ontario is based on Bukata's published data of spectral scattering coefficients. The last measurement shows the specific scattering coefficient of chlorophyll is 0.095 at 695 nm (Figure 4.7). The specific scattering coefficient (m^2/mg) for TSS in Lake Ontario is based on Bukata's published data as well (Figure 4.7).

The absorption spectrum of pure water and the scattering spectrum of pure sea water from 380 to 800 nm are illustrated in Figure 4.8 (Smith and Baker, 1981). Pure

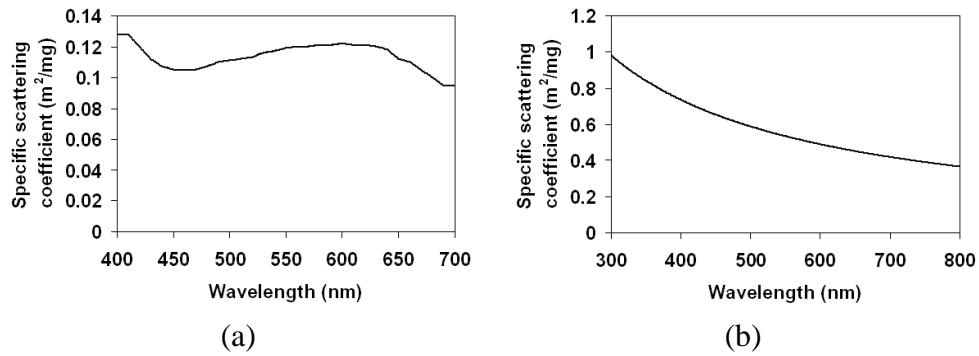


Figure 4.7: Specific scattering coefficient for CHL and TSS. Data is based on Bukata's published data (Bukata et al., 1995).

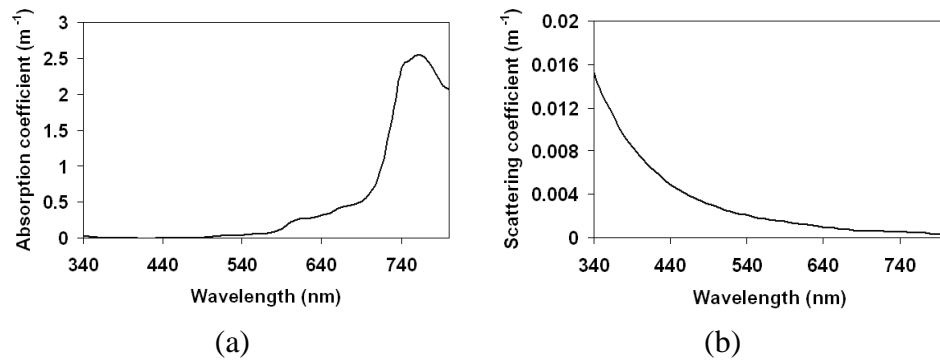


Figure 4.8: Absorption and scattering coefficients of pure water (Smith and Baker, 1981).

water has most of its absorption at wavelengths longer than 700 nm. The scattering coefficient of pure water is less than 0.02 m^{-1} . Figure 4.9 shows the typical exponential absorption spectrum of CDOM and the absorption coefficient was set to 1.0 at 350 nm.

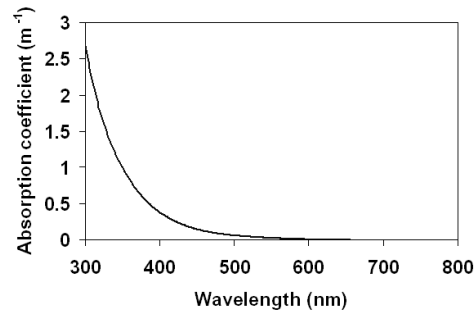


Figure 4.9: Absorption coefficient for CDOM from measurements taken on August 9, 2006 in the Rochester Embayment. CDOM has no scattering contribution.

4.4 Use of remote sensing data

The MODerate resolution Imaging Spectroradiometer (MODIS) is an advanced multipurpose NASA sensor, which is a key instrument aboard the Terra (EOS AM) and Aqua (EOS PM) satellites. Terra MODIS and Aqua MODIS view the entire earth surface every 1 to 2 days, acquiring data in 36 spectral bands. MODIS bands have spatial resolutions of 250 m (bands 1-2), 500 m (bands 3-7) and 1 km (all the rest) (Table 4.2). For most applications in coastal areas the 1 km resolution of typical ocean color sensors is not sufficient and the high-resolution sensors (e.g. Landsat ETM) have drawbacks such as a narrow swath, very infrequent overpasses, and high cost, MODIS 250 m bands are very promising for coastal monitoring applications particularly related to sediment monitoring. Bands (1-2) with 250 m spatial resolution are applied in this research to detect features in coastal areas.

The MODIS surface reflectance data products (bands 1-7) provide an estimate of the surface spectral reflectance for each band as it would be measured at ground level in the absence of atmospheric scattering or absorption. In this study, surface reflectance data for the two MODIS bands with 250 m spatial resolution (bands 1-2) were used to

Table 4.2: MODIS bands and the principal areas of application

Reflective Solar Bands		
<i>250 m</i>	<i>500 m</i>	<i>1 km</i>
Land/Cloud/Aerosols Boundaries	Land/Cloud/Aerosols Properties	Ocean Color/Phytoplankton/ Biogeochemistry
Band 1 (620-670 nm)	Band 3 (459-479 nm)	Band 8 (405-420 nm)
Band 2 (841-876 nm)	Band 4 (545-565 nm)	Band 9 (438-448 nm)
	Band 5 (1230-1250 nm)	Band 10 (483-493 nm)
	Band 6 (1628-1652 nm)	Band 11 (526-536 nm)
	Band 7 (2105-2155 nm)	Band 12 (546-556 nm)
		Band 13L (662-672 nm)
		Band 13H (662-672 nm)
		Band 14L (673-683 nm)
		Band 14H (673-683 nm)
		Band 15 (743-753 nm)
		Band 16 (862-877 nm)
		Band 17 (890-920 nm)
		Band 18 (931-941 nm)
		Band 19 (915-965 nm)
		Band 26 (1.360-1.390 μm)
Emissive Bands		
		Band 20 (3.660-3.840 μm)
		Band 21 (3.929-3.989 μm)
		Band 22 (3.939-3.989 μm)
		Band 23 (4.020-4.080 μm)
		Band 24 (4.433-4.498 μm)
		Band 25 (4.482-4.549 μm)
		Band 27 (6.535-6.895 μm)
		Band 28 (7.175-7.475 μm)
		Band 29 (8.400-8.700 μm)
		Band 30 (9.580-9.880 μm)
		Band 31 (10.780-11.280 μm)
		Band 32 (11.770-12.270 μm)
		Band 33 (13.185-13.485 μm)
		Band 34 (13.485-13.785 μm)
		Band 35 (13.785-14.085 μm)
		Band 36 (14.085-14.385 μm)



Figure 4.10: MODIS 250 m radiance data (645 nm) of Lake Ontario showing the overview of the Genesee River plume (August 13, 2003). Right image is the zoomed image of the Rochester Embayment.

detect the sediment dominated plume in the Rochester Embayment. The best observations during a 24-hour period, as determined by overall pixel quality and observational coverage, are matched geographically. Figure 4.10 is a MODIS 250 m surface reflectance image showing the Genesee River plume on August 15, 2003.

Since reflectance is dependent on both sun and sensor's directions, natural surfaces create scattered reflection in all directions (Figure 4.11). If the energy reflected is exactly the same along all directions, the surface is said to be Lambertian (Figure 4.12). In this case radiance and reflectance are both independent of the viewing angles. As a result of the fact that radiance reflected back to the sensor is defined directionally whereas irradiance received by the surface comes from the entire upper hemisphere, a numerator π , is used as a normalization factor. By assuming that the water surface is Lambertian, the remote-sensing reflectance R_{rs} can be calculated from the surface reflectance R based on:

$$\frac{R}{R_{rs}} = \frac{E_u/E_d}{L_u/E_d} = \frac{E_u}{L_u} = \pi. \quad (4.5)$$

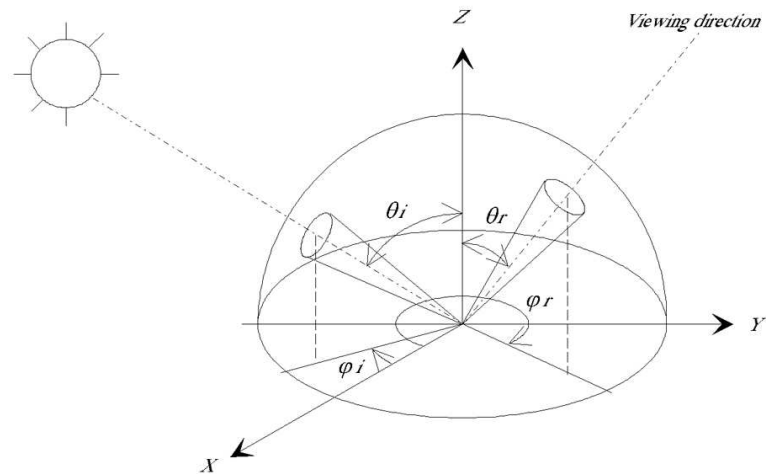


Figure 4.11: The geometry of the bidirectional reflectance. It characterizes the energy scattered into the hemisphere above a surface as a result of incident radiation.

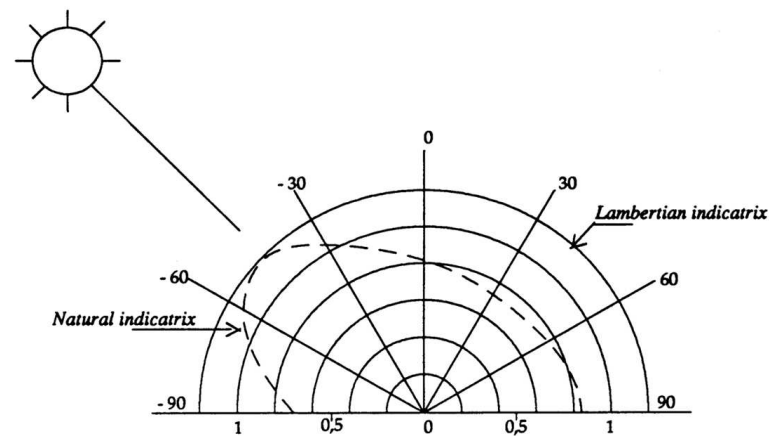


Figure 4.12: Radiance indicatrix in polar coordinate is represented by a circle for a perfect Lambertian surface.

The MODIS thermal bands provide sea surface temperature with 1 km resolution over the global oceans. Figure 4.13 shows the spatial pattern of seasonal warming and cooling for Lake Ontario in 2004 for a series of MODIS images acquired from early April to late July. Blue represents the cooler temperature and red represents the warmer temperature. The minimum surface temperature range of 0.3-2.5 °C occurs from early April to late May. In early April 2004, the entire lake is still below 4 °C even though the temperatures along the northern shoreline exceed the average temperature. In mid-May, the nearshore waters heat faster than offshore waters because of the shallow bathymetry and warmer discharge from the rivers. This phenomenon is seen in the figure of Julian Day 132 along the north shore and around the Niagara River mouth. Figure 4.13 (b) also shows a temperature gradient (thermal bar) developing along the north shore. The spring thermal bar period lasts about 4 weeks. The lake surface warms above 4 °C after the disappearance of the thermal bar. A relatively cool upwelling region is seen in the figure of Julian Day 206 along the north shore and in the western basin. During this period, surface temperatures in the upwelling zone are well below surface temperatures outside the upwelling region.

Satellite remote sensing of the bio-optical properties of coastal waters can be important to monitor the coastal environment. Coastal waters with high sediment load usually have a higher reflectance in the visible region than deep water. However, the spatial resolution of the currently operating ocean color instruments is too coarse for coastal applications. Smaller scale phenomena, such as river discharges and dynamics of coastal sediment are resolved by high resolution imagery. The DIRS Laboratory at RIT developed and operates an airborne spectrometer called the Modular Imaging Spectrometer Instrument (MISI). MISI is an imaging spectrometer designed to provide radiometrically accurate data sets (Schott et al., 2001). It is a line scanner system op-

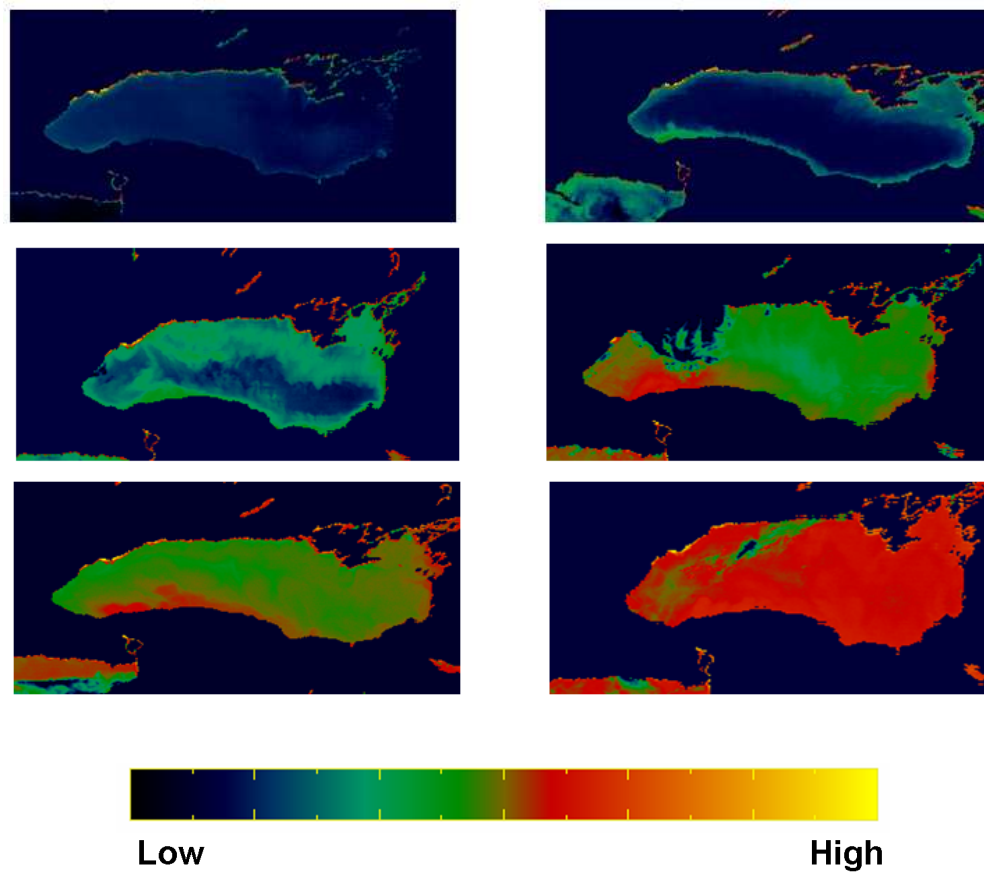


Figure 4.13: Image derived lake surface temperature distribution on Julian Day 96 (April 5 1820EST), Julian Day 132 (May 11 1800EST), Julian Day 151(May 30 1830EST), Julian Day 167 (June 15 1830EST), Julian Day 192 (July 10 1825EST) and Julian Day 206 (July 24 1835EST), 2004.

erating in 84 spectral channels covering the spectral range from 0.4 to 14 micrometers. Figure 4.14 shows example RGB renderings of the MISI images of the Genesee River plume in the Rochester Embayment on different dates. Examples of thermal band data used in this study are shown in section 6.1. Images from the thermal infrared detectors are converted to apparent temperature using onboard blackbodies that are imaged during each scan (Schott et al., 2001). With a flight altitude of 3,000 m, the spatial resolution of the thermal images is about 3 m. RIT's MISI has underflown LANDSAT 5 and 7 to provide calibration of the LANDSAT thermal channel (Schott et al., 2001).

4.4.1 Atmospheric correction errors for MODIS

This study uses the MODIS reflectance product. Atmospheric correction is a necessary step in converting the top of the atmosphere signal to a surface signal. In this case, MODIS level 1B radiances are used to generate the surface reflectance product through an algorithm that compensates for the atmospheric contribution to the level 1B radiance. MODIS land products including surface albedo, snow cover, land cover, land cover change, vegetation indices and biophysical variables rely on the solar reflective portion of the spectrum. The quality of the atmospheric correction algorithm and the accuracy of the surface reflectance are important to the quality of MODIS land products (Vermote and Vermeulen, 1999). The current atmospheric correction algorithm corrects for the effects of gaseous and aerosol absorption and scattering, adjacency effects caused by variation of land cover, Bidirectional Reflectance Distribution Function (BRDF), and atmosphere coupling effects (Vermote and Vermeulen, 1999). Figure 4.15 shows the variable constituents influencing the remote sensing signal at the top of the atmosphere.

Uncertainty arises in the atmosphere correction procedure from many sources. The

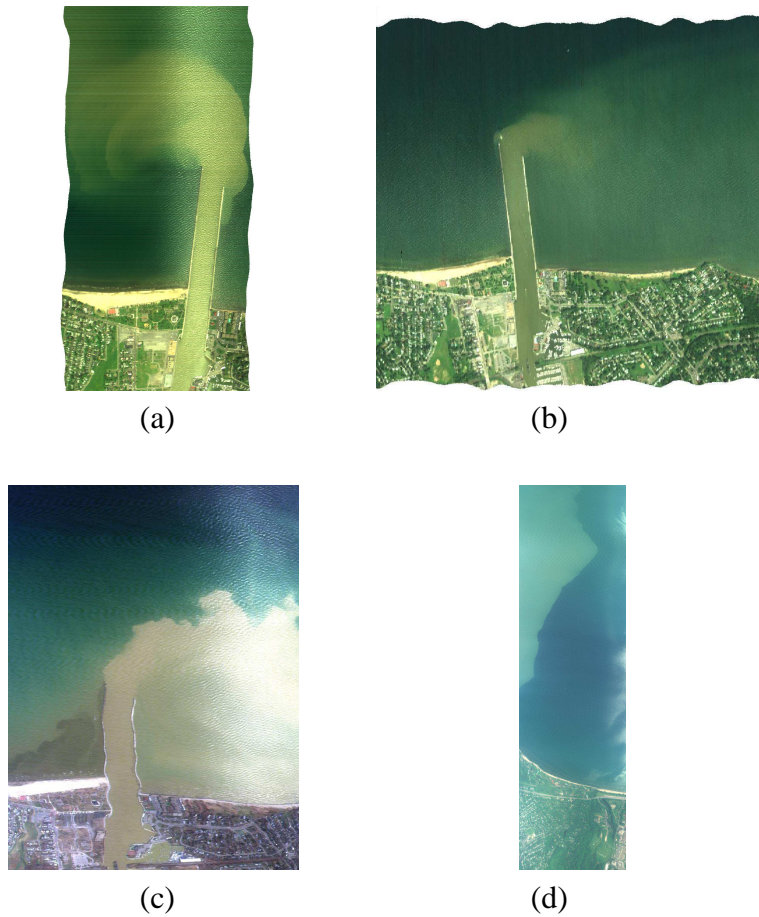


Figure 4.14: A true color rendering of the VNIR spectrometer bands of MISI of Genesee River plume. (a) May 11, 1999 (b) July 5, 2000 (c) March 31, 2003 (d) June 7, 2004.

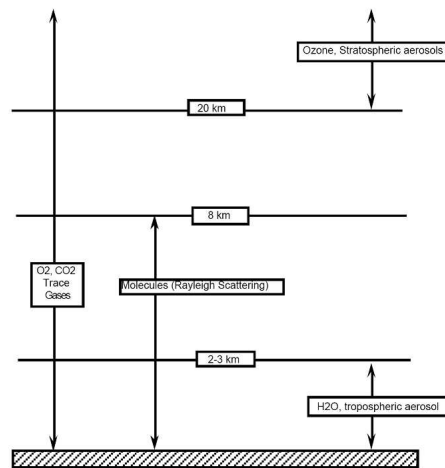


Figure 4.15: The components affect the remote sensing signal at the top of the atmosphere in the 0.4 - 2.5 μm range (Vermote and Vermeulen, 1999).

reflectance at the top of the atmosphere is affected by the uncertainty on the absolute calibration. Research shows an error of $\pm 2\%$ on the top of the atmosphere reflectance can translate to higher error of surface reflectance (Vermote and Vermeulen, 1999). In addition, uncertainty of aerosol optical thickness can affect the surface reflectance retrieval. Errors in the aerosol model and the Lambertian approximation add to the uncertainty of the surface reflectance as well. Figure 4.16 shows the various impacts of uncertainty on the errors of surface reflectance retrieval. Total typical errors of surface reflectance for MODIS bands 1 - 7 is given in Figure 4.17.

band	rel. error %	abs. error
1	4	0.003
2	2	0.015
3	14	0.004
4	3.5	0.004
5	2	0.015
6	2	0.008
7	2	0.003

(a)

band	abs. error
1	0.002
2	0.017
3	0.002
4	0.005
5	0.016
6	0.009
7	0.006

(b)

band	abs. error of SR for $\delta_A(.55\mu\text{m})=0.1$	abs. error of SR for $\delta_A(.55\mu\text{m})=0.5$
1	0.003	0.008
2	0.008	0.018
3	0.006	0.013
4	0.003	0.007
5	0.007	0.016
6	0.003	0.007
7	0.002	0.004

(c)

band	rel. error (%) and abs. error backscattering direction
1	9 0.003
2	2 0.011
3	8.5 0.002
4	4.5 0.003
5	1.5 0.007
6	1 0.003
7	1 0.001

(d)

Figure 4.16: The impacts of uncertainty on the typical errors of surface reflectance retrieval: (a) absolute calibration (b) aerosol model (c) aerosol optical thickness (d) BRDF (Vermote and Vermeulen, 1999).

band	abs. error	rel. error % (range)
1	0.005	10-33
2	0.014	3-6
3	0.008	50-80
4	0.005	5-12
5	0.012	3-7
6	0.006	2-8
7	0.003	2-8

Figure 4.17: Total theoretical absolute and relative errors of the atmospheric correction algorithm (Vermote and Vermeulen, 1999).

Chapter 5

Ensemble Kalman Filter Implementation

5.1 System design

The flow circulation around the river mouth is essential to understanding the transport and the final distribution of the TSS and CDOM discharged into lakes. River plumes are typically dominated by optically complex water and experience significant seasonal and episodic variation of optical properties. The concentrations of the constituents in waters dominate the optical properties, as well as the water quality. Although the ALGE model produces predictions of the distribution of the TSS constituents, visual examination of MODIS 250 m reflectance data clearly shows discrepancies between the model TSS output and the remote sensing data. In this context the modeled TSS is called a state variable. Like any model, these errors are due to the uncertainties in model physics, parameters, and initial conditions. A Kalman filter-based method is used in this research to provide a better estimate of the state variables. State variables

are modeled through system equations between updates based on the dynamics of the system. When the remote observations are available, an EnKF is applied on the model grid to update the predictions of the state variables. The resulting state variables are then used to drive the system model until the next set of remotely data is available. Clearly, only state variables that are remotely observable can be updated.

The framework used in this research to implement the coupled measurement-modeling system with remote sensing as dynamic feedback is represented in Figure 5.1. The process begins with the initialization of the hydrodynamic model ALGE with dynamic and static variables, which is indicated on the left side of Figure 5.1. The dynamic variables are as a function of time and are assumed to have a uniform spatial distribution over the simulation domain. They include the air/dew point temperature, wind speed/direction, cloud cover/height, pressure, upper air temperature, precipitable water, river discharge, and river TSS/CDOM concentration. The static variables are applied as the initialization of ALGE model. To compare the observed reflectance from MODIS scene and the calculated reflectance from radiative transfer model, the distribution of TSS in the river plume is the primary output considered here even though ALGE simulates other physical variables (temperature, flow current, *etc.*). When the remotely sensed reflectance spectra are available, modeled TSS concentration profiles, as well as lab measured IOPs and the concentration of CHL and CDOM, are entered into the forward radiative transfer model Hydrolight to produce remote sensing reflectance spectra on the model grid. Using modeled and remotely sensed reflectance values, an EnKF is applied to update the previous prediction of TSS concentration values and produce a better estimate of them. The updated TSS concentration profiles are then fed back to the hydrodynamic model. ALGE runs forward in time until the next remote observation is available. An important advantage of this system is that there is no need for an inverse

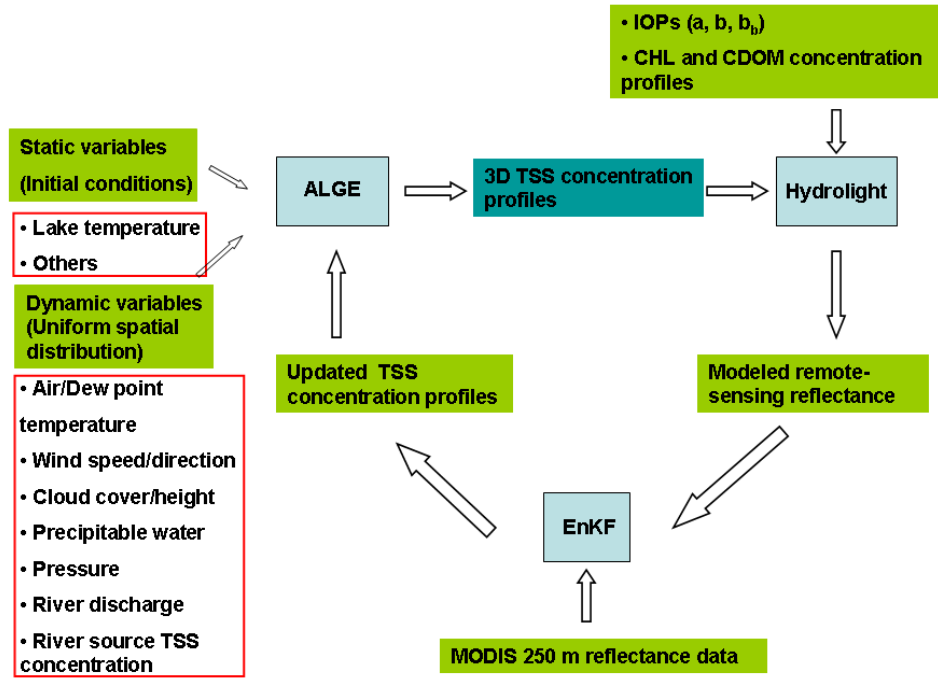


Figure 5.1: Schematic representation of the coupled measurement-modeling system.

radiative transfer model to predict the TSS from the remotely measured reflectance spectra.

Table 5.1 lists the definitions of the variables in EnKF used in this research. The TSS concentrations at the river mouth are defined as a state variable. Measurements contain the remote-sensing reflectance derived from MODIS 250 m data. The system model is the hydrodynamic model ALGE which formulates the evolution of TSS concentration profiles. The measurement model is the non-linear representation of the radiative transfer model Hydrolight.

The complete operation cycle including time update and measurement update used in EnKF in this research is illustrated in Figure 5.2. Figure 5.2 also shows the equations for each update as well. The basic steps for updating TSS concentration using the

Table 5.1: EnKF definitions and notations

Variable	Name	Definition in this study
x_k	state variable	TSS concentration
A	non-linear system model	ALGE
Q	process noise	Estimated
P_k	system uncertainty	Covariance between ensemble model states
H	non-linear measurement model	Hydrolight
z_k	observation	MODIS derived R_{rs}
R_k	measurement noise	Derived from MODIS

EnKF are:

- Generate an ensemble of forcing variables and run the ALGE model forward in time to obtain a forecast estimate of state variables at time k for all ensemble members.
- Compute the error covariance of the state variables and the measurements.
- Calculate the Kalman gain. It is a weighting factor that determines how much of the information from the measurement should be taken into account when updating the state estimate.
- Update the predicted state variables and obtain a better estimate of the state variable when the measurement is available. The state variables are updated at time step k by applying the Kalman gain to the difference between the measurement and the predicted measurement.

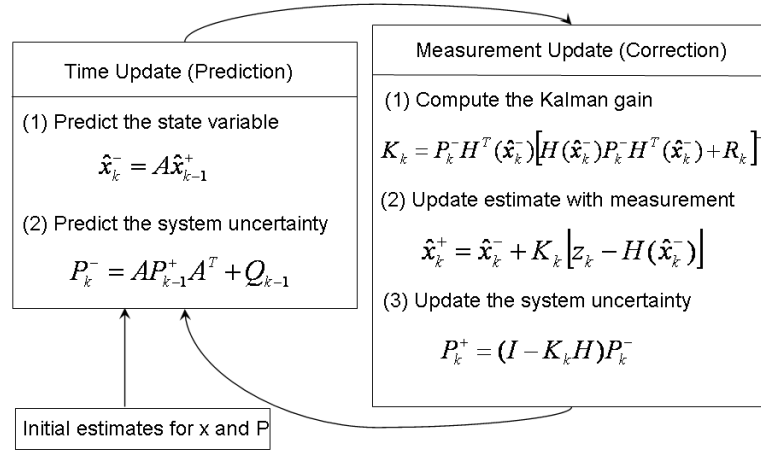


Figure 5.2: A complete operation cycle of the Kalman filter showing the equations of time update and measurement update (Welch and Bishop, 2006).

5.2 Generation of ensemble forcings

To implement the ensemble Kalman filter, system model uncertainty is estimated by running ALGE with an ensemble of forcing variables. Generally, model uncertainty can be created from errors in forcing variables and model parameters. In this research, only errors in the forcing data are included. The ensemble of forcing variables includes hourly data for the river source TSS concentration (g/m^3), the river discharge (m^3/s), and the average wind speed/direction. An ensemble of model states at each time step is generated by running ALGE with the forcing inputs which are treated as stochastic terms. The covariance between the ensemble of states provides an estimate of system model uncertainty. The open-loop simulations represent model uncertainties through an ensemble of model states. Each ensemble corresponds to a different realization of the forcing data that drives ALGE to produce the TSS distribution. The error terms of both the system model and the measurement model are propagated forward in time.

This is an important distinction that separates Kalman filter based assimilation methods from other simple methods. For EnKF, it's important to make sure the ensemble spreads enough to cover the modeling system uncertainty.

Discharge and TSS concentration data of the Genesee River were collected at the USGS streamflow-gaging station (Genesee River at Rochester, 04232000) and the USGS water quality site (Charlotte Pump station, 431510077363501) during the year of 2003 (Figure 5.3 and 5.4). Heavy rainfall in late July and early August lead to the streamflow with a peak of about $300 \text{ m}^3/\text{s}$ and TSS concentration with a peak of about 190 g/m^3 . The time series of discharge, river source TSS concentration, and wind speed and wind direction are given in Figure 5.5 and 5.6 for a time period of July 26 to August 18, 2003. Discharge and river TSS concentration shown in Figure 5.5 are from the same data source as Figure 5.3 and 5.4. Wind speed and direction are observations taken at the Rochester Airport (WBAN 14768). In Figure 5.5 and 5.6, the real observations from the stations are shown in red curve (or red dots for wind direction). Using the observations as mean values and 10% to 15% of that as standard deviations, Gaussian distributions at each time step for each forcing variable are created. These errors which are artificially introduced to the forcing data are assumed to be large enough to capture the uncertainty in the modeling system. The individual ensemble member for each forcing input is obtained by randomly choosing one value from the Gaussian distributions at each time step and is shown as the blue curve (blue dots for wind direction).

Other than the four forcing variables mentioned above, ALGE is also driven by other parameters. Values for these static and dynamic variables and the methods to obtain them are given in Table 5.2.

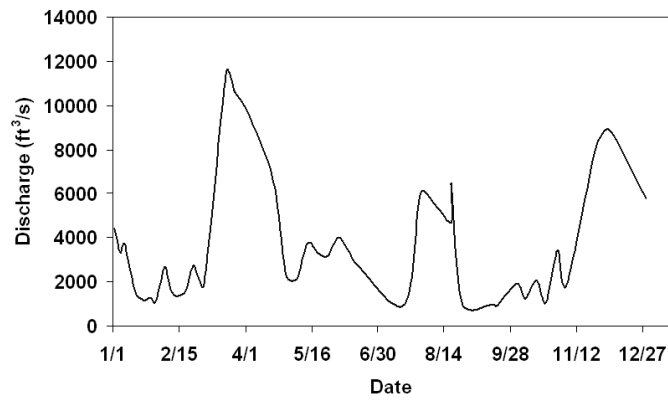


Figure 5.3: Genesee River discharge during the year of 2003 at streamflow gaging station 04232000 (Genesee River at Rochester).

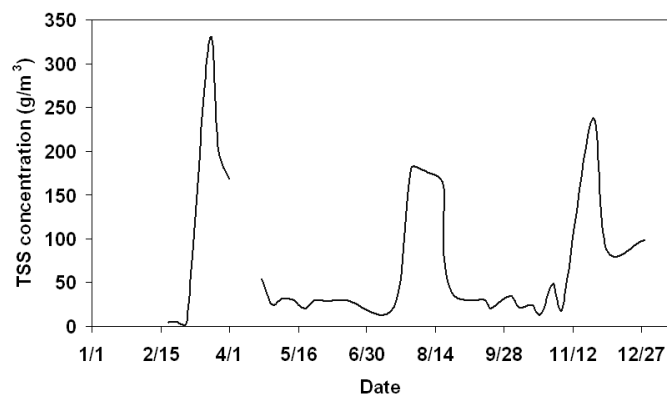
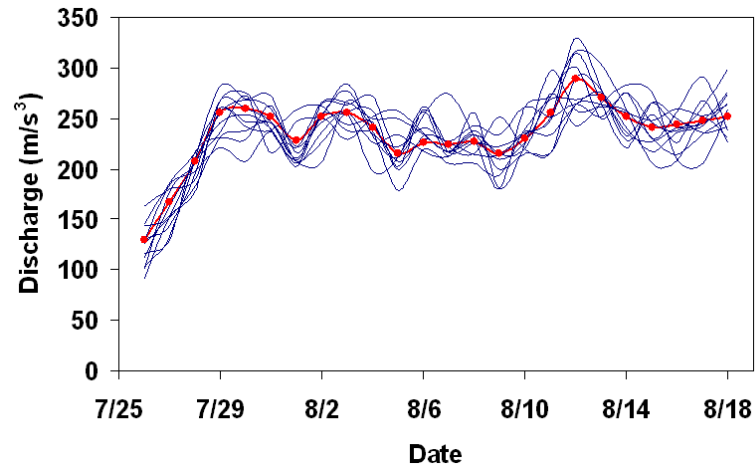
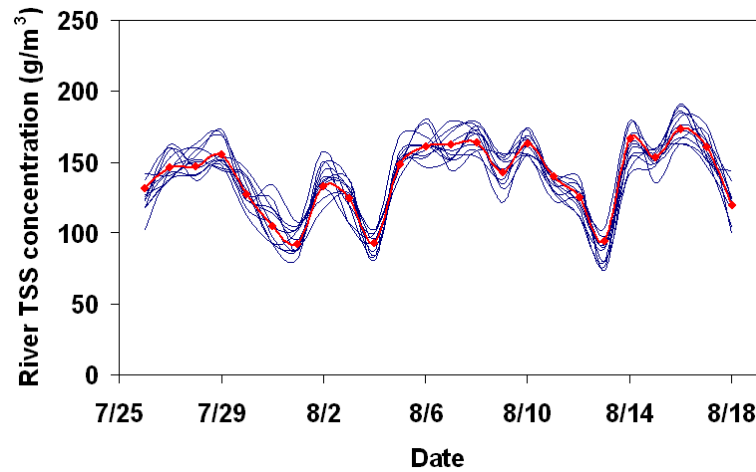


Figure 5.4: TSS concentration during the year of 2003 at water quality site 431510077363501 (Genesee River at Charlotte Pump Station).

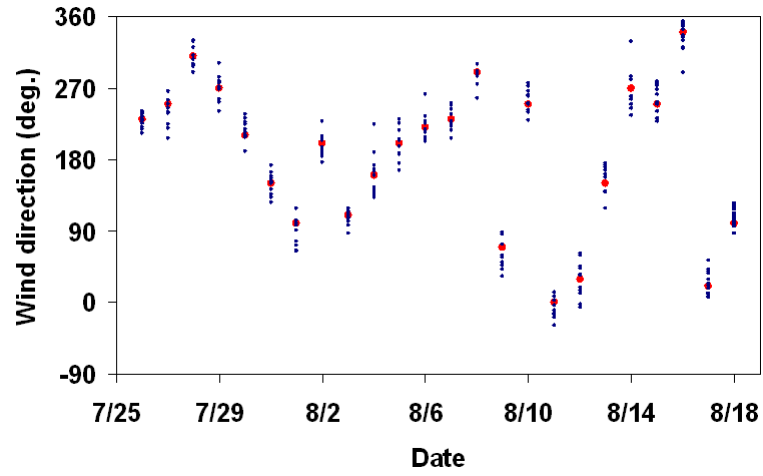


(a)

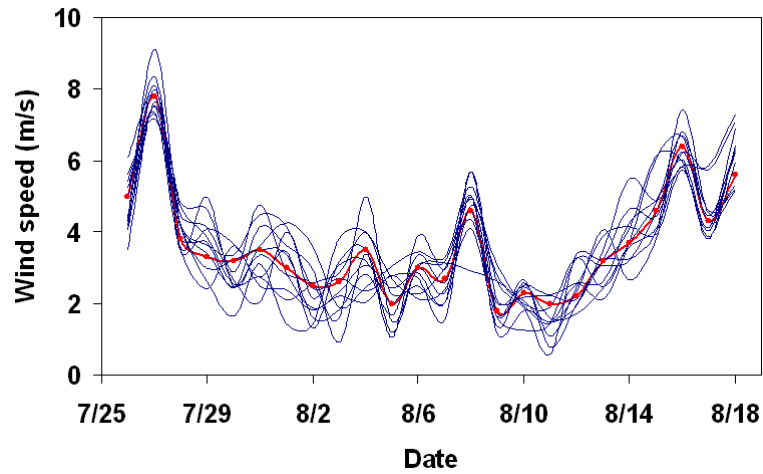


(b)

Figure 5.5: Discharge (a) and TSS concentration (b) of the Genesee River during the simulation period. The real observation is given in red, and the ensemble of forcing inputs with Gaussian distribution at each time step is given in black.



(a)



(b)

Figure 5.6: Wind direction (a) and wind speed (b) during the simulation period. The real observation is given in red, and the ensemble of forcing input with errors contributed is given in black.

Table 5.2: Definitions, values, units and data sources for parameters used in ALGE.

Parameter definition	Value	Unit	Data source
<i>Static variable</i>			
Horizontal grid spacing	135	m	National Geophysical Data Center
Vertical grid spacing	3	m	National Geophysical Data Center
Simulation time	576	hour	
Latitude	43.3	degree	
Longitude	77.55	degree	
Julian day	193	unitless	
Roughness	0.001	m	Estimated
Inflow/Outflow temperature	11.0	°C	Estimated
Mass source temperature	13.0	°C	Estimated
Particle density	2.5	g/cm ³	Estimated
Particle diameter	2.0e-6	m	Estimated
Number of grids in x direction	115	unitless	
Number of grids in y direction	62	unitless	
<i>Dynamic variable</i>			
River source TSS concentration		g/m ³	USGS water quality site 431510077363501
River discharge		m ³ /s	USGS streamflow-gaging station 04232000
Air temperature		K	Rochester airport (WBAN 14768)
Dew point temperature		K	
Cloud height		km	
Cloud cover		unitless	
Pressure		mb	
Wind direction		degree	
Wind speed		m/s	
Upper air temperature		K	Buffalo airport (WBAN 14733)
Precipitable water		cm	

5.3 Generation of ensemble observations

5.3.1 MODIS data for data assimilation

The ensemble Kalman filter applied in this research requires observations of remote sensing reflectance with error estimates. Daily L2G MODIS surface reflectance data was obtained from Earth Observing System Data Gateway. The spatial resolution is 250 m (bands 1-2) and the data is an estimate of the reflectance as it would be measured at lake surface in the absence of atmospheric scattering or absorption. The spectral range of MODIS band 1 is 620-670 nm and the range of band 2 is 841-867 nm. By assuming the water surface is a Lambertian surface, the remote-sensing reflectance values were calculated as

$$R_{rs} = \frac{R}{\pi}, \quad (5.1)$$

where R is the surface reflectance retrieved from MODIS.

Daily MODIS data was examined and eight relatively cloud-free images were downloaded during the simulation period of July 26 to August 18, 2003. A land-mask and cloud-mask were applied to the MODIS data based on band 2 reflectance. A Region Of Interest (ROI) was chosen around the Rochester Embayment for each MODIS scene and the image map was converted from UTM to Latitude/Longitude. All pixels within the study area were inspected visually to insure they were free of sun glint. Figure 5.7 is a false-color RGB combination using the two 250 m bands of MODIS. Vegetation appears red, clouds appear white, and water appears black. The sediment-dominated plume water appears blue since band 2 does not include plume information. Average wind speed and direction are shown in the figure as well. Ontario beach was closed on July 29, August 13, August 14, August 15, and August 18, 2003 due to the extreme high flow of the Genesee River (Knauf, 2003). However, on July 29

and August 15, beach was closed although the plume was flowing eastward. Almost no plume is shown on August 16 compared with the previous day, which indicates there may be an error in atmosphere correction.

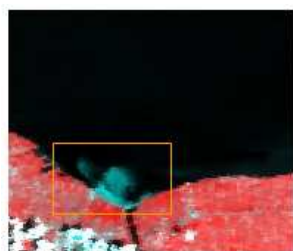
5.3.2 MODIS observation error statistics

With regard to observation error statistics within the assimilation theory, there are three major requirements and any violation can result in bad assimilation results (Albert, 2004). The three requirements are: 1) the measurements should be bias-free; 2) the observation error should be close to a Gaussian distribution; 3) the observation error should have a small spatial correlation. In this research, all three requirements were tested with the summer MODIS scene shown in Figure 4.10.

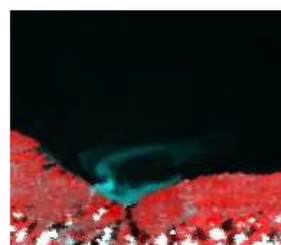
A small region in the center of deep lake was chosen as an ROI to calculate the measurement error. The mean value of surface reflectance in ROI was used as the true value to calculate the observation bias. A small overall bias of less than 3.85% is obtained by subtracting the mean.

The histogram of the MODIS observation error in that ROI by subtracting the mean is shown in Figure 5.8. The regression curve in red shows the error is approximately a Gaussian distribution. The second requirement of the data assimilation can be expected to be satisfied by the MODIS measurements.

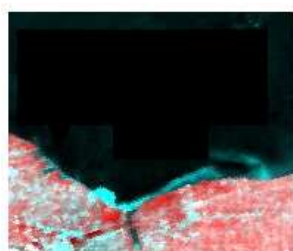
The spatial correlation of observation errors was calculated and binned into 250 m intervals. In Figure 5.9, the correlation between the first column and the rest of the image-ROI is shown as a function of distance. The spatial correlation values range from -0.1 to 0.15. Thus, the third requirement of the data assimilation can be expected to be fulfilled by the MODIS observations.



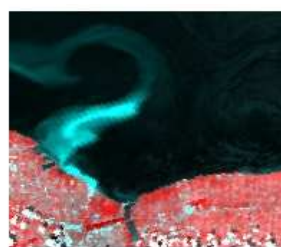
(a) July 29 (west, 7.5 mph)



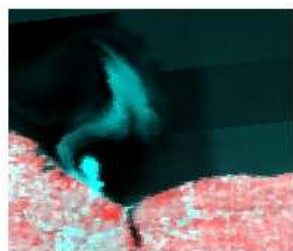
(b) July 30 (southwest, 4 mph)



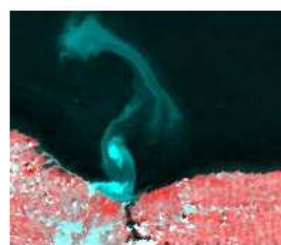
(c) August 8 (northwest, 9 mph)



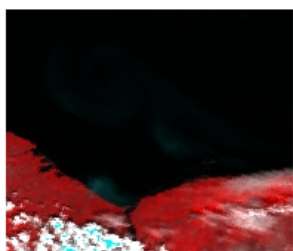
(d) August 13 (northeast, 3 mph)



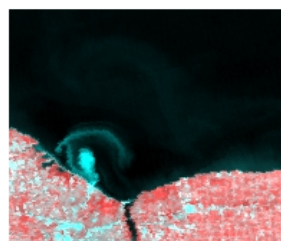
(e) August 14 (west, 9 mph)



(f) August 15 (southwest, 8 mph)



(g) August 16 (northwest, 12 mph)



(h) August 18 (east, 10 mph)

Figure 5.7: MODIS 250 m surface reflectance data showing Genesee River plume dissipation in Lake Ontario during summer of 2003. The images are false-color RGB combination using bands 2-1-1.

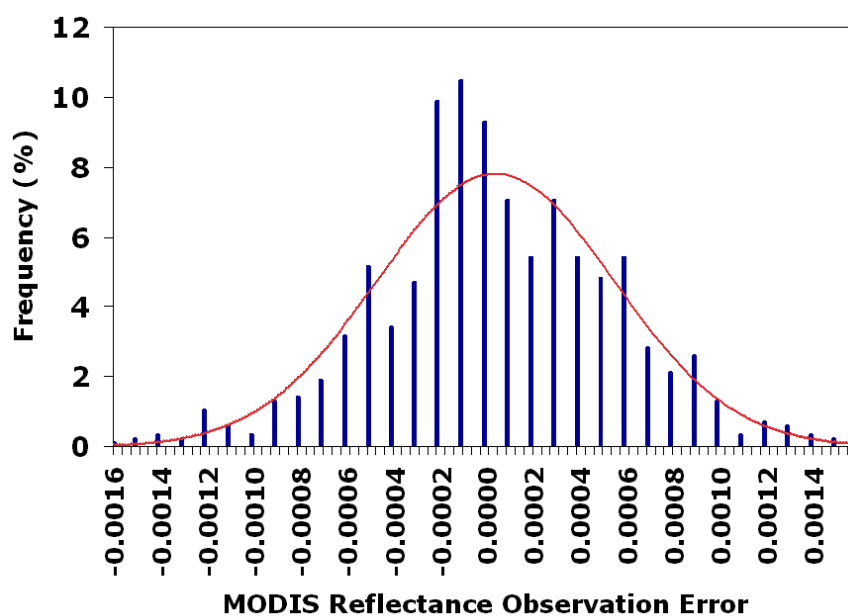


Figure 5.8: Frequency of occurrence of deviations of MODIS derived remote-sensing reflectance. The red line shows Gaussian around the mean difference.

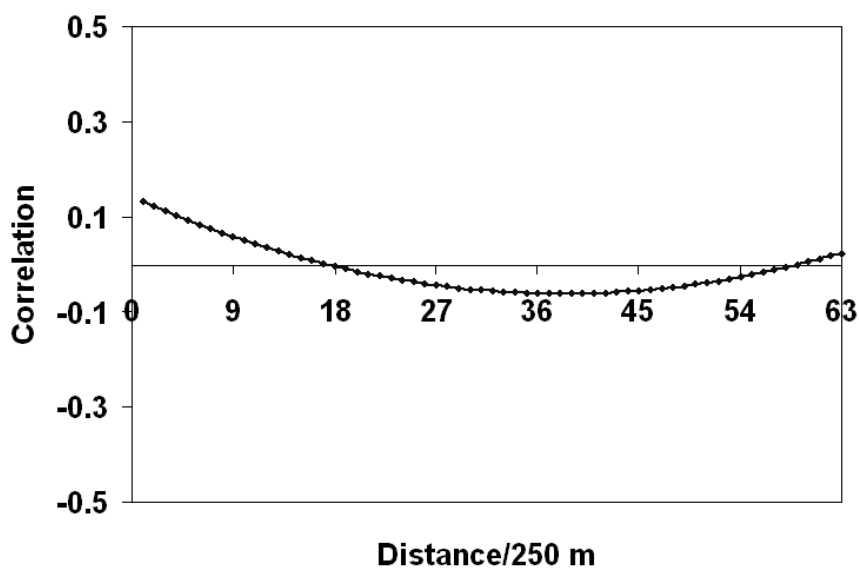


Figure 5.9: Spatial correlation of MODIS surface reflectance observation errors as a function of distance.

5.4 Open-loop simulations

The Rochester Embayment study site was divided into a grid for the ALGE simulation with a horizontal grid spacing of 135 m and a vertical grid spacing of 3 m. The study period includes 24 days from July 26 to August 18, 2003. A small ensemble space with 12 members is generated from the open-loop runs with different forcing variables. The estimated plume dissipation at the study site for each ensemble member is nudged from a whole lake simulation to obtain a reasonable boundary condition. Model uncertainty is calculated as the error covariance of the state variables among the ensemble members. Update is performed only when the MODIS data is available. By equating the model variables with the assimilated quantities, the matrices H and R are re-binned to match the proper dimension of the MODIS data to perform matrix multiplication.

Chapter 6

Results and Discussion

6.1 Grid independence study and validation of ALGE using airborne thermal imagery

To assess the impact of grid resolution on the ALGE simulations, a grid sensitivity study was performed on Conesus Lake using identical simulations with the only difference being the number of nodes in the grid (Li et al., 2007). For the whole lake simulations, three grid sizes were tested, $232 \times 70 \times 6$ nodes with 40 m horizontal spacing, $308 \times 90 \times 6$ nodes with 30 m horizontal spacing, and $537 \times 158 \times 6$ nodes with 20 m horizontal spacing. The results show that the surface flow circulation patterns are similar for each grid size and that the nudging factors derived from the three simulations are the same magnitude and direction.

The effect of horizontal grid spacing on high resolution plume simulations was also examined. A grid with twice as many horizontal nodes (1.5 m spacing) as the nominal grid (2 m spacing) was created for Cottonwood Gully. The comparison of the surface flow field shows that the simulation with a 1.5 m grid reveals some small scale

features such as eddies near the surface which are not present in the simulation that used 2 m spacing. The integrated sediment deposition pattern modeled with the 1.5 m grid is slightly wider spread to the south of the stream mouth, which may be a result of the higher resolution mesh allowing ALGE to generate a recirculating cell that is not resolved with the coarser mesh. It is concluded that increasing the grid resolution produces minor differences in the results but the basic conclusions are the same.

The validation of ALGE began with comparison of surface temperatures of Conesus Lake from the 2003 simulation with temperatures derived from remote thermal images. Figure 6.1 shows the thermal imagery (8 - 10 micrometer) of the middle and northern portions of Conesus Lake collected during MISI overflights in May and June 2003. Three regions were chosen from the images to retrieve water surface temperature from thermal radiance measured by MISI. The results showed the distribution of water surface temperature was nearly uniform. Since the lake surface temperature converted from thermal radiance did not have noticeable gradients, we took the average of the three regions for each image to compare with the average surface temperature from the simulation. The comparison of lake surface temperature between the thermal imagery and the simulation is provided in Table 6.1. The agreement has an RMSE of 0.9 °C. With the errors associated with various blackbody sources and calibration assumptions, MISI's on-board blackbody and transducer system can be calibrated to within 0.3 K. When we add uncertainties from approximations in ALGE and the input data used by ALGE, we consider these results for matching remote sensing data with a two or three month simulation to be very good.

Further validation was provided by field measurements and a simulation for spring and summer 2004. Direct temperature measurements were taken at Long Point and Cottonwood of Conesus Lake from a boat at noon on the day the simulations ended.

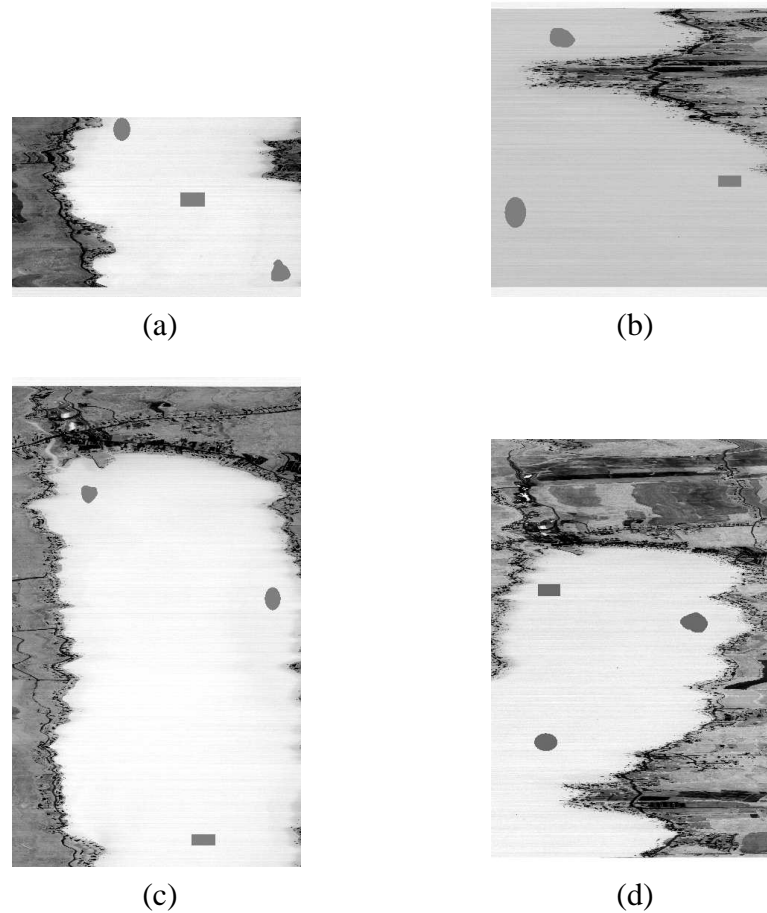


Figure 6.1: Thermal imagery of middle ((a) and (b)) and northern ((c) and (d)) portions of Conesus Lake: (a) 14:25 on May 19, 2003 (b) 13:05 on June 28, 2003. Three regions (Rectangle, Ellipse, and Polygon) were picked to convert from thermal radiance to surface temperature. The images were not corrected for aircraft roll and have been enhanced with a histogram stretch to show variation within the lake. The horizontal striping is an indication of line-to-line noise emphasized by the enhancement.

6.1 Grid independence study and validation of ALGE using airborne thermal imagery 123

Table 6.1: Comparison of lake surface temperature (LST) derived for MISI imagery and simulation output

	May 19 <i>Middle portion</i>	June 28	May 19 <i>Northern portion</i>	June 28
Region 1 (Rectangle)	17.1	20.2	14.5	22.6
Region 2 (Ellipse)	15.1	21.0	15.3	20.7
Region 3 (Polygon)	15.7	21.4	15.8	21.1
Average LST from above (°C)	16.0	20.9	15.2	21.2
Average LST from simulation (°C)	15.1	20.0	15.6	20.2

The water temperatures were measured at various depths by using calibrated thermistors attached to a weighted line. In addition, a chain of recording temperature sensors (Onset Stowaway TidBit sensor) at different depths was placed near Long Point Park from April through October 2004. The Onset Stowaway TidBit sensor is a completely sealed underwater temperature logger and is used to provide hourly temperature data at several depths in the water column.

In Figure 6.2, a comparison of the whole lake simulation from April to June 2004 found overall good agreement between the model output of average surface temperature as a function of time and the measured temperatures from the near surface TidBit sensor at Long Point. The accuracy of the TidBit sensor is about ± 0.2 °C over 0 to 50 °C. The model output is smoother than the temperature measurement during the study period. Although the predicted surface temperature is warmer than the observed temperature in spring and colder than the observed temperature in summer, the overall comparison is very good with an RMSE of 0.8 °C.

Figure 6.3 shows the comparison of observed and simulated temperature profiles at two locations where subsurface temperatures in Conesus Lake were measured at the end of June, 2004. Overall, the agreement between observed and computed tempera-

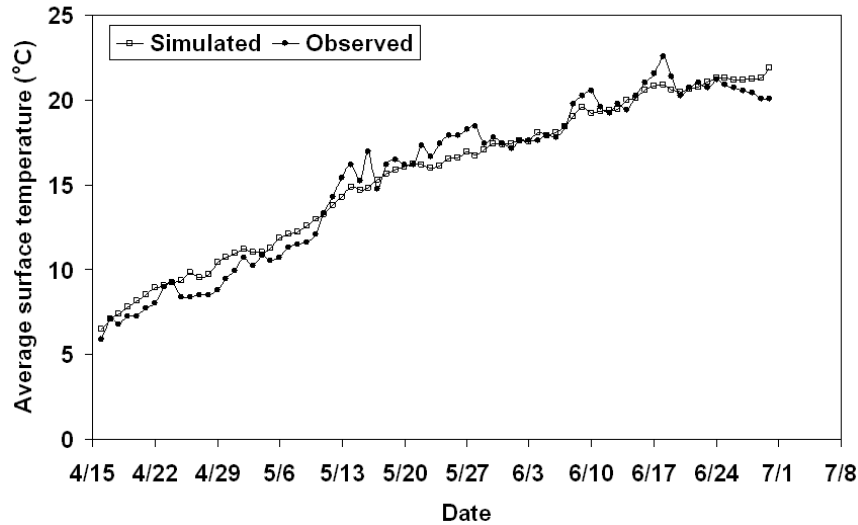


Figure 6.2: Simulated and observed average surface water temperature during April to June, 2004 at Long Point.

ture profiles is again good. The RMSE temperature difference between the observed and simulated values above a depth of 9 m is less than 0.7 °C. Below a depth of 10 m, the simulated temperatures are significantly warmer than the observed temperature. The discrepancy may be partly caused by the lack of vertical resolution in the simulation ($dz = 3.1$ m). The simulation with 1 m vertical resolution was run and a better temperature match at the surface and the bottom was found. However, the intermediate layers still did not show the same well developed thermocline as the measurements. Due to forest cover, the steep hillsides of the lake, and the interpolation into a single wind speed, the estimate of the wind speed and direction over the lake has a large uncertainty and therefore the effect of turbulent mixing may not be adequately simulated. In any case the simulated currents are very similar for each vertical resolution so the nudging factors are also very similar.

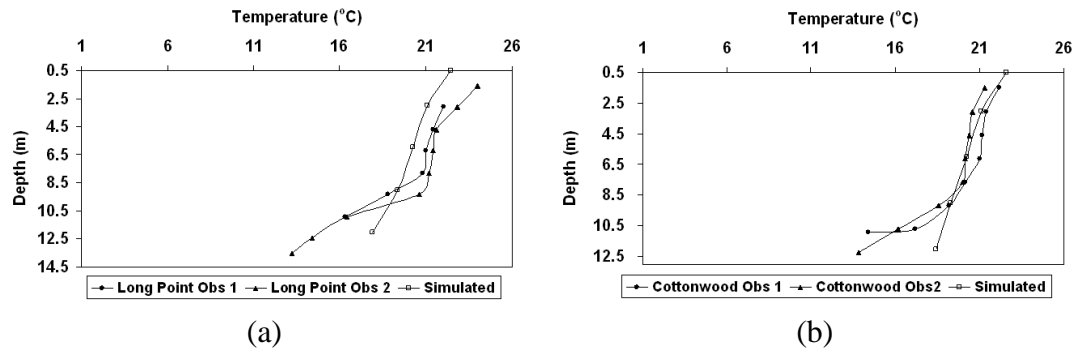


Figure 6.3: Observed and simulated temperature profiles on July 1, 2004 for stations: (a) Long Point Gully (b) Cottonwood Gully.

6.2 Thermal structure in Lake Ontario

Characteristics of the temperature structure in Lake Ontario are important for understanding both the entire lake and nearshore water circulation processes and the consequent impacts on lake water quality. The previous observations and the satellite images have clearly showed the spatial and temporal characteristics of temperature distributions during the thermal bar period. The surface temperature field exhibits a warm coastal ring and a cold interior core, whereas the vertical temperature structure shows a stratified coastal region and a well-mixed interior region. Satellite images from the Great Lakes Forecasting System (GLFS) (Figure 6.4 and 6.5) illustrate the expansion of the thermal bar in Lake Ontario in 1997. The thermal bar begins with an incomplete warm ring along the coast and expands lakeward during the spring warming. After the incomplete warm ring extends along the entire coast, the coastal band is formed and confined to the coastal region throughout the thermal bar period. Figure 6.6 shows several observed temperature transects of Lake Ontario from GLFS in 1997. The water temperatures in the midlake are less than 4 °C and the development of the thermal

bar on the north shore is seen in Figure 6.6 (a). Due to the depth of the water and the configuration of the bottom, the thermal bar is more developed along the shallower north shore compared to the deeper south shore (Figure 6.6 (b) and (c)). As the heating intensifies towards the midsummer period, the progression of the thermal bar moves towards the center of the lake. Figure 6.6 (d) shows typical summer stratification occurs when the surface water temperature warms over the entire lake and the thermal bar disappears.

Figure 6.7 depicts a time series of the simulated surface temperature obtained for four months of simulation for Lake Ontario in 2004. The warm coastal ring starting from the northern shoreline and Niagara River area is seen in Figure 6.7 (a) and (b) due to the characteristics of the lake bathymetry. The bottom slope is steep along the south shore and smooth along the north shore. Since the inflowing water of Niagara River is usually 3 to 4 °C warmer than Lake Ontario, the warmest surface region occurs near the mouth of Niagara River during the early thermal bar period (late April and early half May). Then the warm ring starts to complete the circle around the lake's perimeter (Figure 6.7 (d)). The surface temperature contours approximately are parallel to the depth contour. A much broader warm band starts to form north to the cold core in late May which moves the cold core towards the south. Figure 6.7 (e) shows the cold core breaks into two small pieces in the deepest region at the end of the thermal bar period (late May and early June). Depending on meteorological conditions, the thermal bar may last for a period of from 1 to 3 months, varying from year to year.

A time series of the simulated water temperature transect at 77.8W in Lake Ontario is presented in Figure 6.8. It shows the lake changing from entirely thermally mixed in early spring to strongly stratified in late summer. Figure 6.8 (a) and (b) illustrate well mixing from top to bottom at temperatures near or below 4 °C in early spring. Then

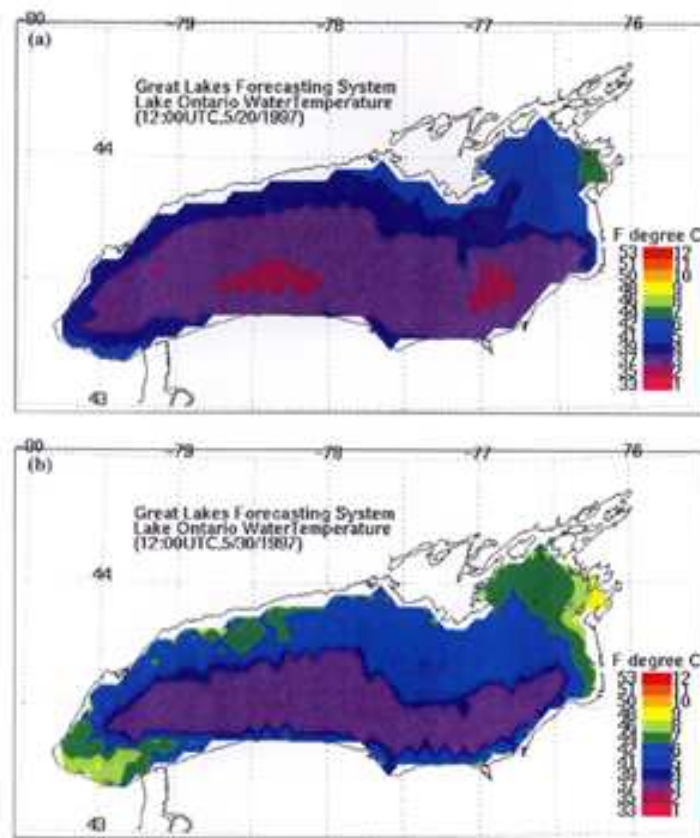


Figure 6.4: Surface temperature (color scale is in C as well as in F in Lake Ontario from the GLFS on May 20 and May 30, 1997 (Chang, 2003).

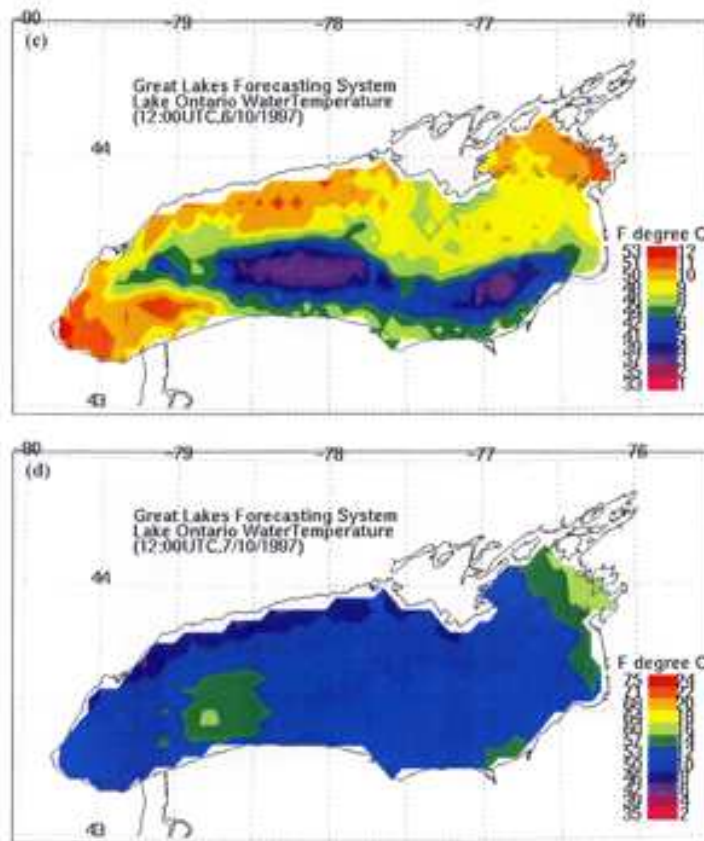


Figure 6.5: Surface temperature (color scale is in C as well as in F in Lake Ontario from the GLFS on June 10 and July 10, 1997 (Chang, 2003).

the spring-time warming tends to heat and stratify the near-shore shallow areas first. Figure 6.8 (c), (d) and (e) show that the thermal bar that separates the stratified and homogeneous areas of the lake begins its offshore movement and propagates across almost the entire lake. Stratification eventually covers the whole lake, and a well-developed thermocline generally persists throughout the summer.

Significant wind events can cause upwelling and downwelling along the shoreline

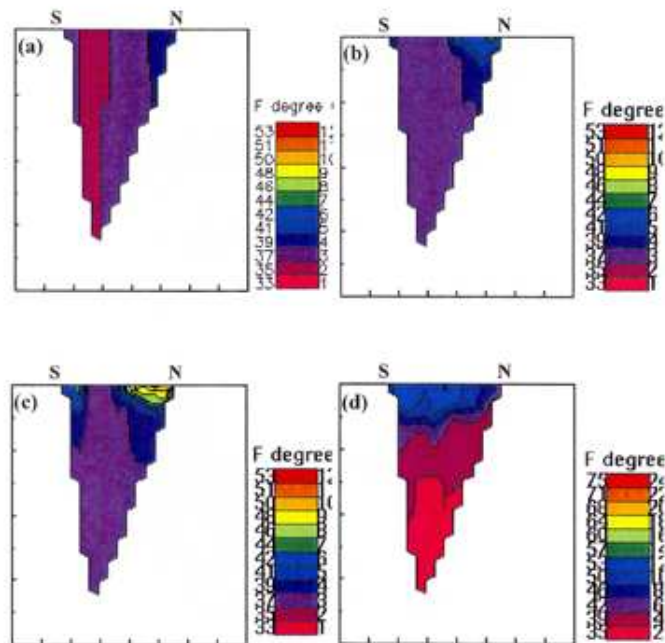


Figure 6.6: Observed temperature transects at 78.3W (color scale is in C as well as in F) in Lake Ontario from the GLFS on (a) May 20 (b) May 30 (c) June 10 (d) July 10, 1997 (Chang, 2003).

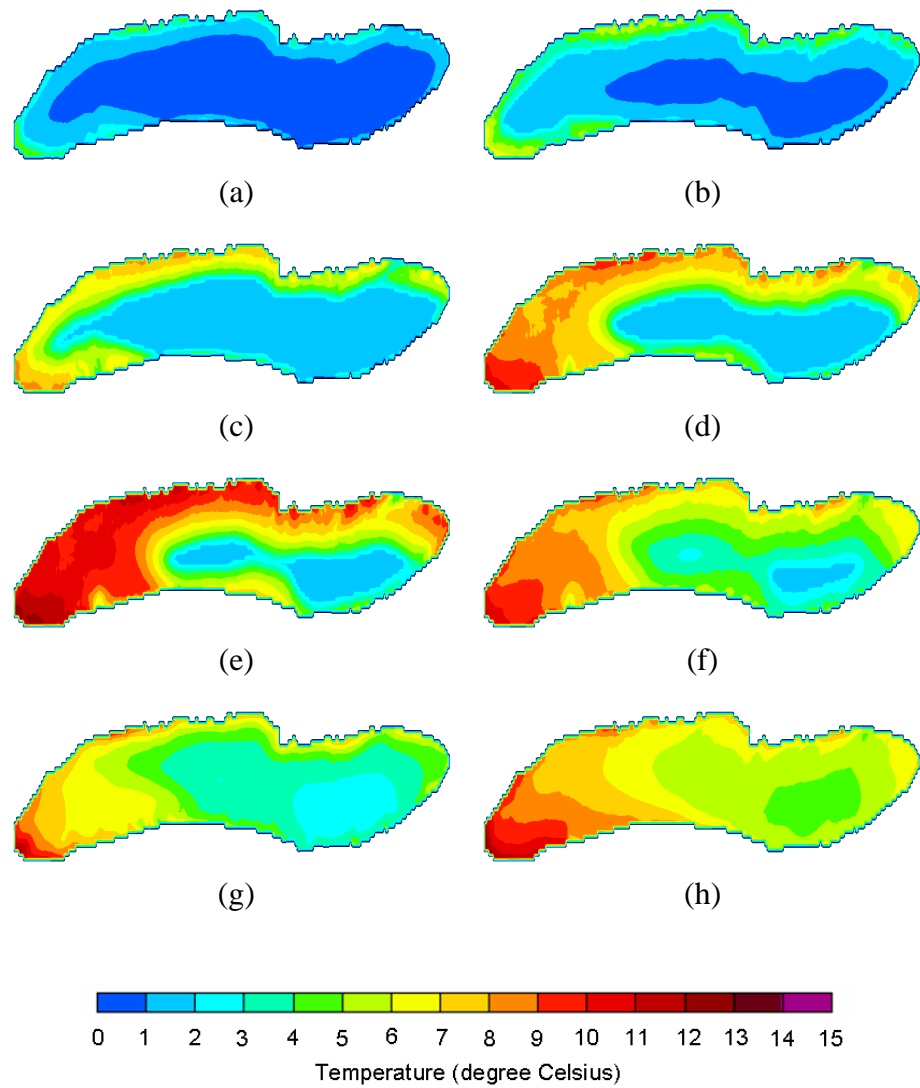


Figure 6.7: Simulated surface temperature distribution of Lake Ontario on (a) April 20 (b) April 30 (c) May 10 (d) May 20 (e) May 25 (f) May 30 (g) June 10 (h) June 20, 2004.

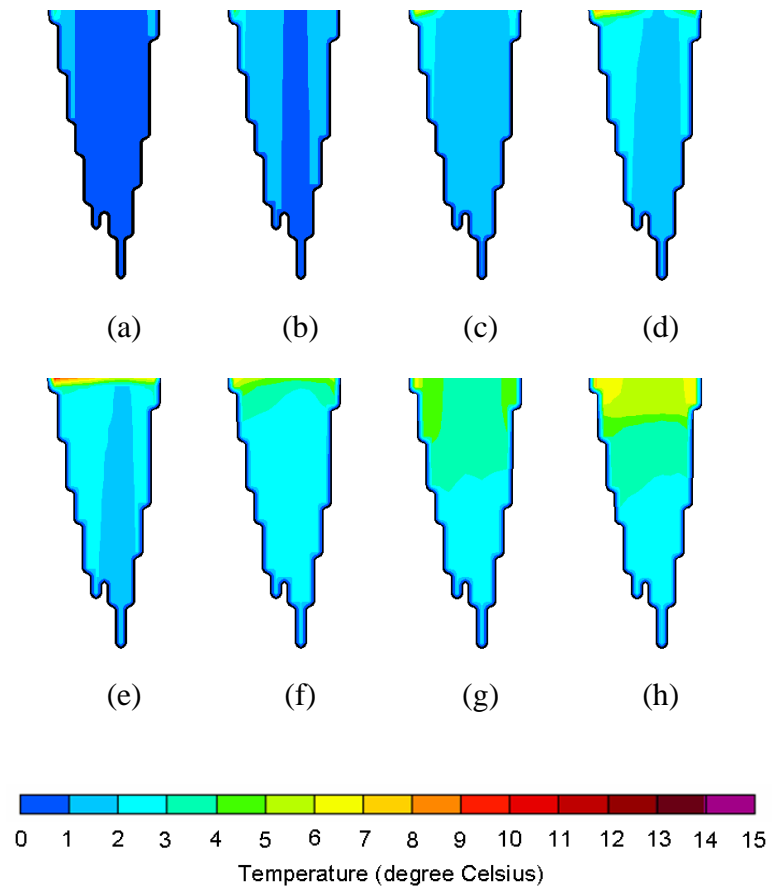


Figure 6.8: Simulated temperature transects at 77.8W in Lake Ontario on (a) April 20 (b) April 30 (c) May 10 (d) May 20 (e) May 25 (f) May 30 (g) June 10 (h) June 20, 2004.

during the summer stratified season. It is known that the phenomenon of upwelling along the north shore and downwelling along the south shore is resulting from the eastward component of the wind stress in Lake Ontario. The upwelling events are characterized by relatively weak eastward flow, and downwelling events with strong westward currents (Murthy and Miners, 1992). Research showed that upwelling events generally associate with cooling down the region, whereas downwelling events associate with warming up the region. A relatively cool coastal upwelling region at the northeast corner on May 25, 2004 is shown in Figure 6.9. The nearshore temperature is about 6 - 9 °C whereas water temperatures of the rest part of the lake are 10 - 13 °C. The upwelling events result from the prevailing eastward wind on that day.

6.3 Numerical simulations of large scale circulations

Large-scale circulation patterns in lakes can be affected by wind stress curl, topography, and stratification. During the stratification season, the vertical exchange is greatly reduced and the main motions are almost entirely horizontal. The large scale circulations produced by the whole lake ALGE simulations are required to develop the nudging for stream plume simulations.

6.3.1 Conesus Lake

Figure 6.10 depicts a time series of the simulated surface currents obtained for three months of simulation of 2004 for the Conesus Lake. The arrows indicate the speed and direction of the predicted motion at a node located at the origin of each arrow. The most striking features are the convergence zones and overall complicated circulation pattern in the north and south basins during a northwest wind (Figure 6.10 (a)). The

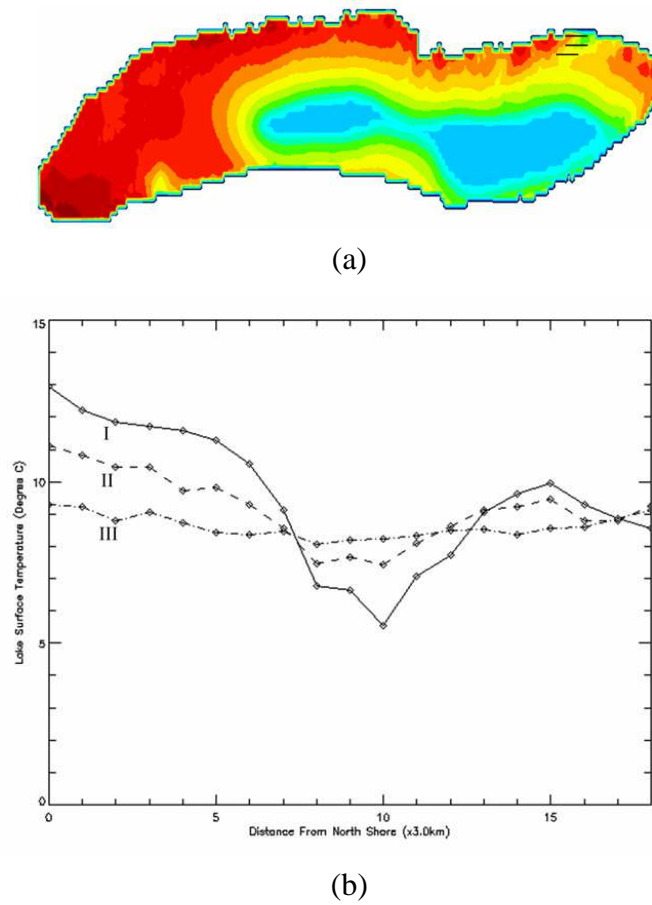


Figure 6.9: Simulated surface water temperature along the sampling transects on May 25, 2004. Curve I, II, and III corresponding to the lines in (a) from north to south.

average current speed over the entire lake was 1.0-1.1 cm/s with a maximum speed reaching 4.4 cm/s during April. The average current speed over the entire lake was 0.7-0.8 cm/s with maximum speed up to 4.2 cm/s during May and June. The storm-induced flow currents can be strong in Conesus Lake in summer with maximum speed up to several tens of cm/s. The maximum flow current during a major storm in May was 19.8 cm/s. The circulation and flow currents were stronger in April than in May and June, when wind speeds were higher and the winds were more northerly (Figure 4.2).

The predicted water circulation patterns at different depths in Conesus Lake on June 29, 2004 are presented in Figure 6.11. Circulation at the surface, 3 m, and 6 m is shown in Figure 6.11 (a), (b), and (c), respectively. On June 29 wind was blowing from the southeast (about 150°). For the surface layer, the currents are generally in the direction of the wind. For examples, the velocity vectors at the surface are south to north and parallel to the shoreline along the western shore. The overall effect of the southerly wind is a conveyor belt behavior with downwelling at the north end and upwelling at the south end. The currents are opposite the direction of the wind at a depth of 3 m in Figure 6.11 (b). Currents at a depth of 3 m are stronger than currents at a depth of 6 m. Generally, the wind-induced currents are in the direction of the wind near the surface and are in the opposite direction in lower layers. The conservation of water volume in the enclosed lake basin is guaranteed by the reversal in current direction.

6.3.2 Lake Ontario

Large scale water movement in Lake Ontario during spring and summer of 2004 was simulated by ALGE. The earlier observations showed a thermally-driven coastal cur-

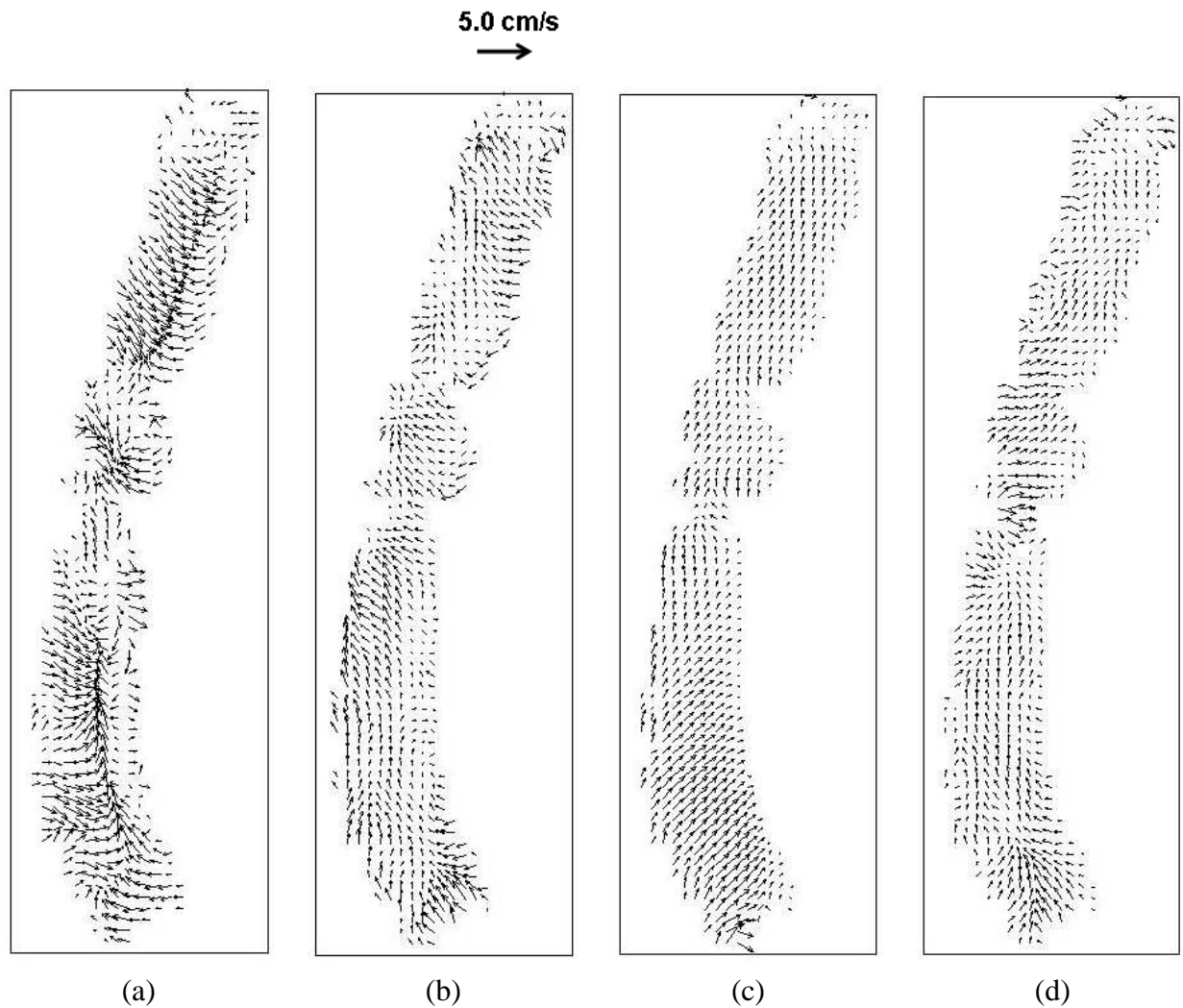


Figure 6.10: Time series of simulated surface currents (cm/s) for the entire lake during spring and summer conditions, 2004: (a) April 7 (b) April 26 (c) May 30 (d) June 29. Arrow length and direction represent the current strength and direction. Average wind direction was northwest, southeast, southwest, and south, respectively. One out of three nodes is plotted for clarity.

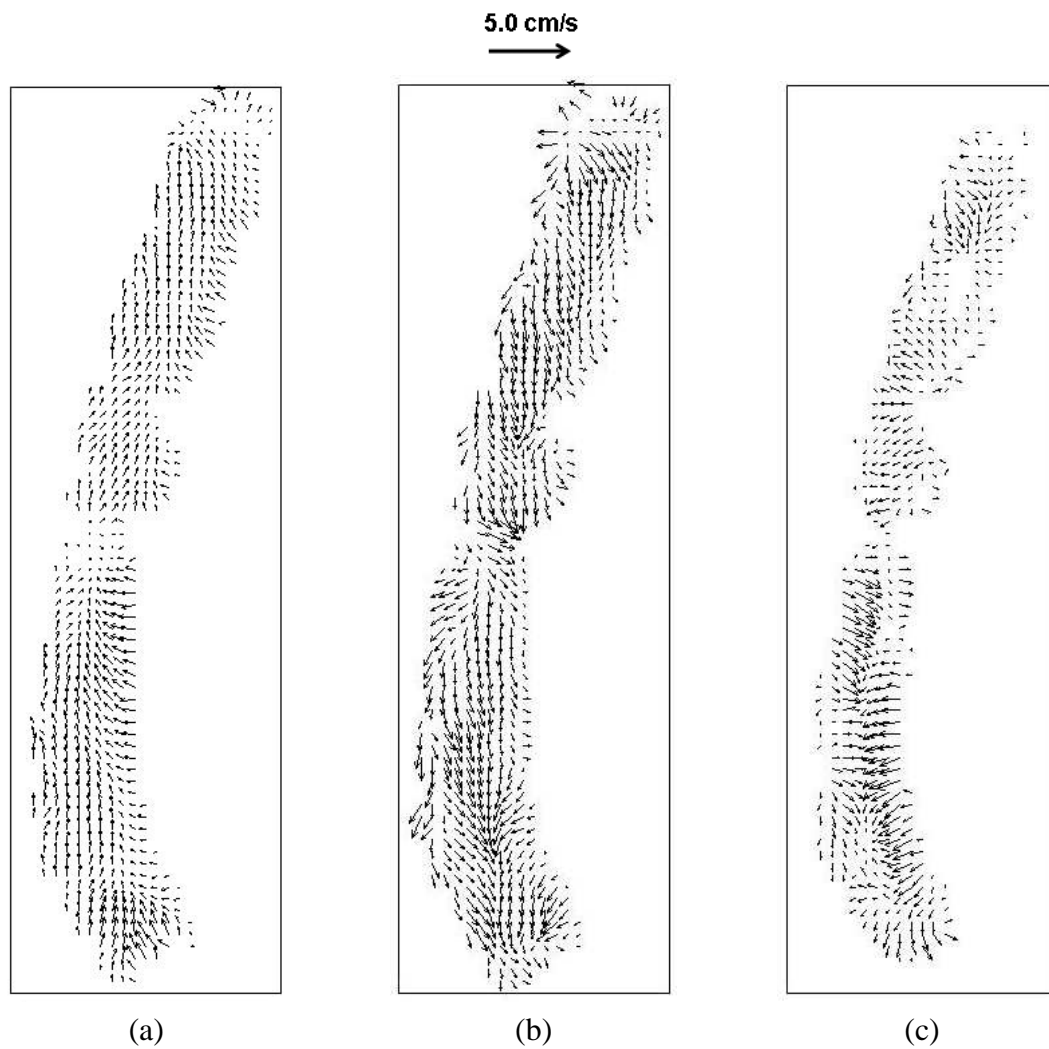


Figure 6.11: Prediction of circulation (cm/s) at different depths for the entire lake with wind blowing from the southeast (June 29, 2004): (a) surface (b) depth = 3m (c) depth = 6m. One out of three nodes is plotted for clarity.

rent in the counterclockwise direction under the light wind and a wind-driven coastal current in the wind direction under strong winds (Chang, 2003). Figure 6.12 depicts a time series of the simulated surface currents and the water temperature at each grid point for Lake Ontario in 2004. Nearshore currents are generally stronger than offshore currents. The average surface currents in Lake Ontario are relatively weak through spring and summer of the year (on the order of several cm/s). However, under the strong wind the flow currents can be strong in Lake Ontario with maximum speed up to several tens of cm/s. The average monthly simulated surface currents from April through July are 4.0 cm/s, 3.3 cm/s, 7.8 cm/s, and 7.7 cm/s with maximum speeds reaching 16.9 cm/s, 13.7 cm/s, 19.2 cm/s, and 15.0 cm/s.

Figure 6.12 (a), (b), (d), and (e) show the surface wind-driven circulation in the whole lake under the average wind direction of southwest, southwest, northwest, and south, respectively. Comparing to Figure 6.12 (a), Figure 6.12 (b) shows a similar flow circulation pattern but slightly weaker currents because the resulting wind speed was slower on that day. Figure 6.12 (c) and (f) show a strong westward current along the north shore of Lake Ontario, which formed part of the counter-clockwise circulation in the whole lake.

The predicted water circulation patterns at different depths in Lake Ontario on July 27, 2004 are presented in Figure 6.13. Circulation at the surface, 9 m, and 17 m is shown in Figure 6.13 (a), (b), and (c), respectively. Figure 6.13 (d) shows the simulated surface temperature distribution on July 27, 2004. The minimum surface temperature is observed at the south-east corner and along the middle southern shoreline on that particular day. The cold surface temperatures are resulting from the coastal upwelling events which bring up the colder water from the deep layers.

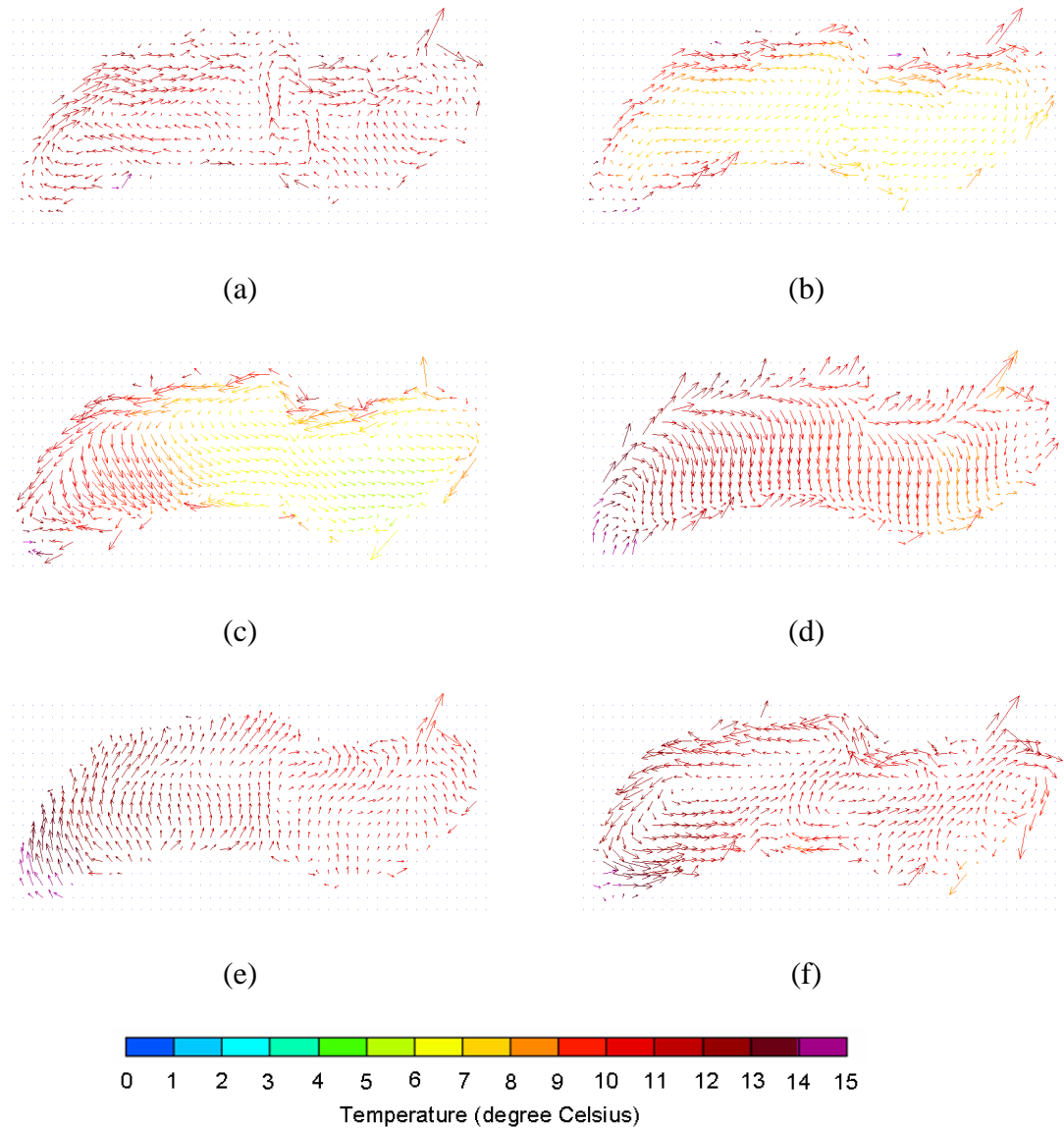


Figure 6.12: Time series of simulated surface currents (cm/s) for the entire lake during spring and summer conditions, 2004: (a) April 7 (b) April 27 (c) June 10 (d) June 27 (e) July 10 (f) July 27. One out of three nodes is plotted for clarity. Lake surface temperature at each simulation grid is indicated by the color of the current vector.

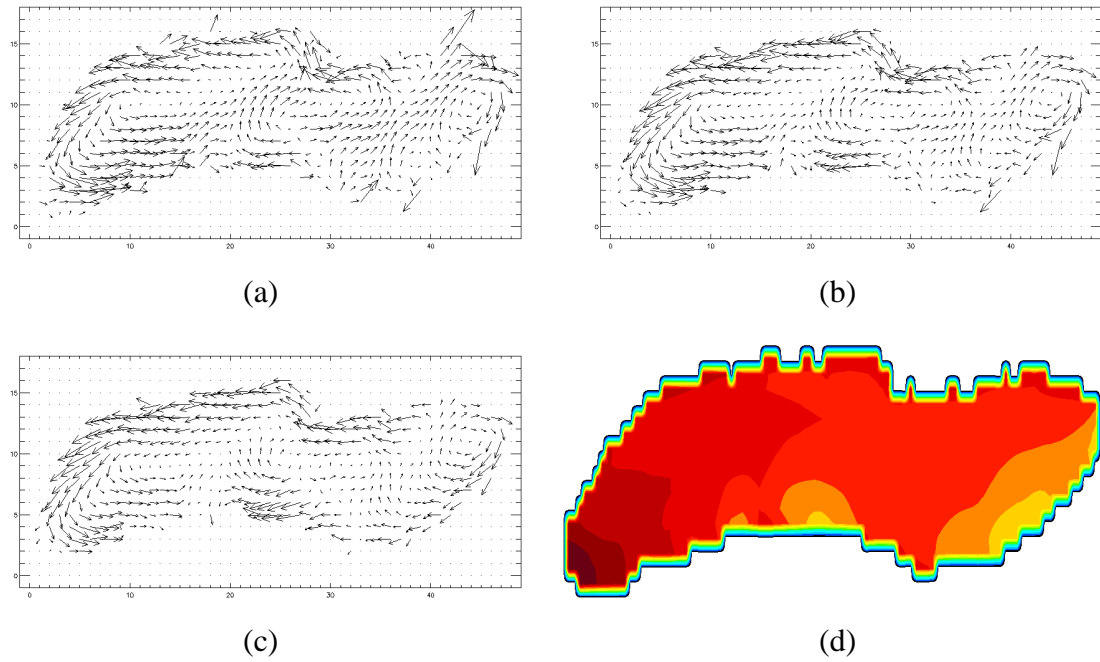


Figure 6.13: Prediction of circulation (cm/s) at different depths for the entire lake with wind blowing from the southeast (June 27, 2004): (a) surface (b) depth = 9 m (c) depth = 17 m (d) predicted surface temperature distribution. One out of two nodes is plotted for clarity.

6.4 Numerical simulations of river plumes

6.4.1 Stream plumes in Conesus Lake

The hydrodynamic model results for the entire lake provides the overall wind-driven circulation in Conesus Lake during the spring and summer. However, during hydrometeorological events, water circulation can be locally modified by stream plumes. Nudging data from the whole lake simulations were applied to investigate the fate and transport of tracer and sediment at stream mouths during storm events in the late spring and early summer.

Sand Point Gully

Figure 6.14 is a photograph of the Sand Point Gully stream plume during a storm event of June 16-19, 2004. At the peak stream flow rate of $1.35 \text{ m}^3/\text{s}$, the maximum current velocity of the simulation was 0.8 m/s at the stream mouth of Sand Point Gully, which is over ten times the average velocity at deeper areas (Figure 6.15 (a)). In Figure 6.15 (a) and (b), a convergence zone is seen southeast of the stream mouth at Sand Point Gully. During the storm event, the incoming water has a higher density than that of the lake due to its suspended load. Because of the higher density of sediment laden water flowing into the lake, the plume sinks creating the convergence zone and sediments are more likely to deposit on the shallow nearshore zones. Figure 6.15 (c) shows tracer moved toward the southeast at the surface and at 3 m during storm event. This tracer movement is consistent with the plume momentum and is coincident with moderate but steady winds from the northwest. The nudging currents do not appear to strongly affect the plume.

Cottonwood Gully

The storm event that occurred during May 20-25, 2004 at Cottonwood Gully had a



Figure 6.14: Plume at Sand Point Gully flowing southeast in June, 2004 during a storm event.

maximum flow of $0.6 \text{ m}^3/\text{s}$, which produced a current velocity at the stream mouth of about 0.45 m/s (Figure 6.16 (a)). Compared to Sand Point Gully, a steep slope along the nearshore is found at Cottonwood Gully. For much of the Cottonwood Gully area depths of 3 m could occur within 15 m of the shore. Due to the relatively deep bottom near the stream mouth at Cottonwood Gully, the stream plume sediments need a longer residence time before depositing on the bottom of the lake. A surface and near surface gyre is shown immediately adjacent to stream mouth in Figure 6.16 (a) and (b) which increases the residence time. In Figure 6.16 (c), the plume is observed to stay near the shoreline and then move north toward the macrophyte bed even though the wind is from the northwest. This plume appears to be strongly driven by the nudging currents.

Site comparison

The nudging data from the whole lake simulation at Cottonwood Gully is pointing from south to north under all wind conditions even when the wind was blowing from north-northwest as indicated by the stream plume simulation (see Figure 6.10

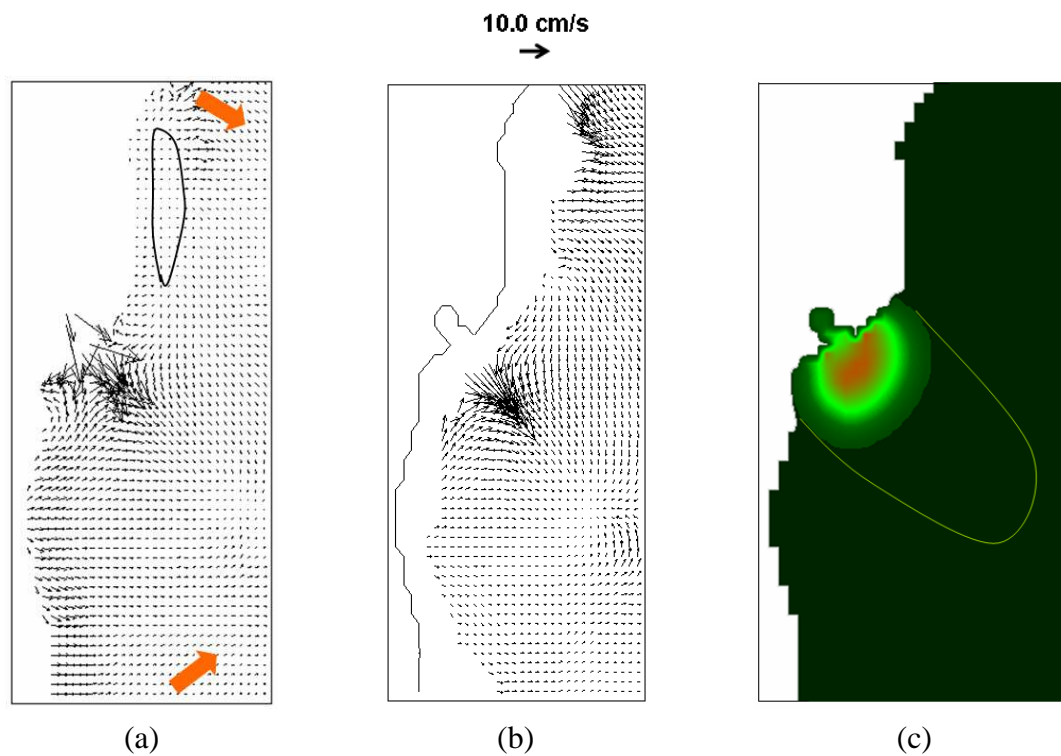


Figure 6.15: Water circulation and passive tracer transport at the mouth of Sand Point Gully during a major storm event during June 16-19, 2004. The figures show the results 18 hours after the beginning of simulation. The wind was blowing from northwest. Large vectors in (a) show nudging data (direction of flow currents, not to scale) from whole lake simulation. The circled area indicates the macrophyte bed location. (a) Surface (b) 2 m (c) Tracer transport at surface and contour of tracer distribution at 3 m. The flow current vectors at the stream mouth are not on the same scale in (a) and (b) as the current vectors away from the stream because of the extraordinarily larger velocity at stream mouth. One out of four nodes is plotted for clarity.

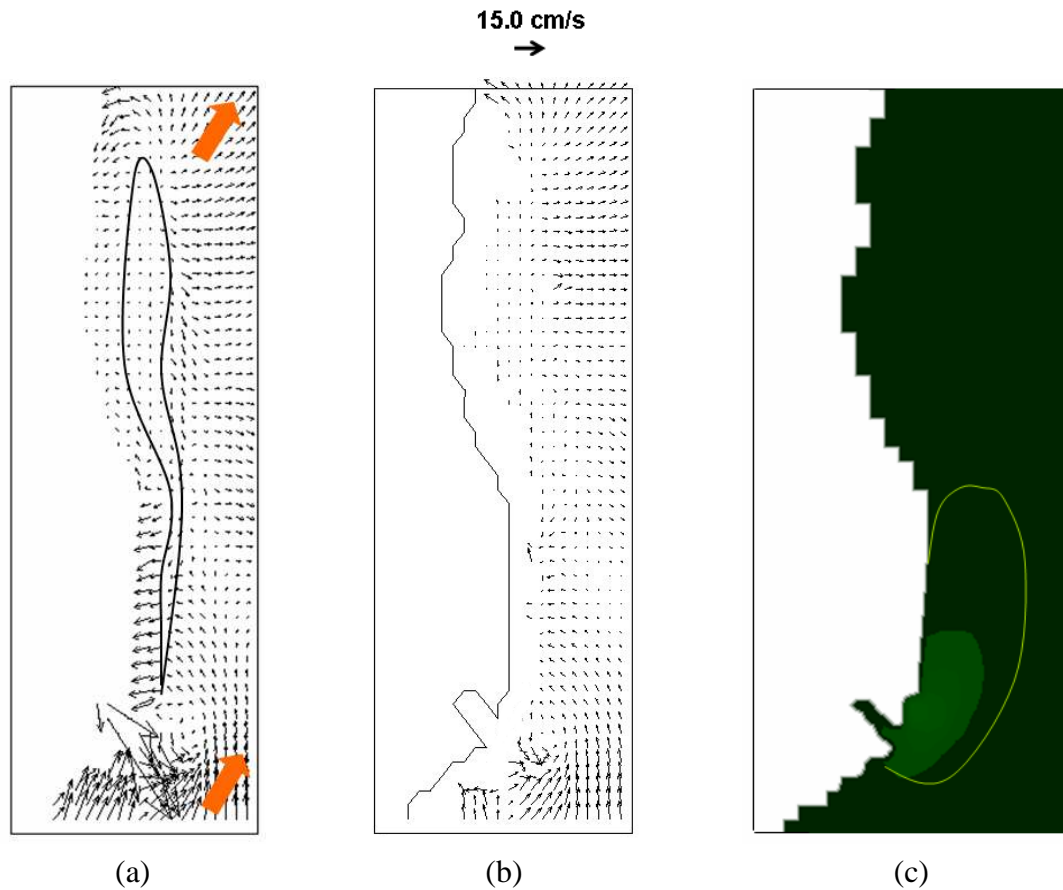


Figure 6.16: Water circulation and passive tracer transport at the mouth of Cottonwood Gully during a major storm event during May 20-25, 2004. The figures show the results 48 hours after the beginning of simulation. The wind was blowing from northwest. Large vectors in (a) show nudging data (direction of flow currents, not to scale) from whole lake simulation. The circled area indicates the macrophyte bed location. (a) Surface (b) 2 m (c) Tracer transport at surface and contour of tracer distribution at 3 m. The flow current vectors at the stream mouth are not on the same scale in (a) and (b) as the current vectors away from the stream because of the extraordinarily larger velocity at stream mouth. One out of four nodes is plotted for clarity.

and 6.16). As a result of the steep bathymetry at Cottonwood and maximum depth of 19 m, the stream plume is strongly affected by the main current from the entire lake. However, the nudging data do not appear to have such a strong impact on the plume at Sand Point Gully. We also found out that the magnitude of nudging data at Sand Point Gully is slightly smaller than that at Cottonwood Gully.

Slow flow currents are seen to form at the macrophyte beds at both Sand Point Gully and Cottonwood Gully (Figures 6.15 and 6.16). The average velocity in the middle of macrophyte beds are 2.9 cm/s at Sand Point Gully and 1.8 cm/s at Cottonwood Gully. These velocities are much smaller than those outside the macrophyte beds and offshore in deeper water. The average velocities in the deeper waters at those two sites can reach 4.5 cm/s to 5 cm/s. The friction of the macrophytes enhances tracer and sediment residence time within the macrophyte beds.

The sediment deposition patterns show qualitative information about where small particles with small settling velocities will be transported and deposited at stream mouths and create potential macrophyte habitat. The integrated sediment deposition patterns at the two study sites with and without macrophyte drag are given in Figure 6.17. We expect submerged macrophytes to slow down the flow and enhance sedimentation and our results do show that process occurring at both sites. The region of sediment deposition at Cottonwood Gully is much more narrowly constrained than at Sand Point Gully because of the strong control by the northward flowing nudging currents even though the simulation was about twice as long. At Sand Point the deposition is more wide spread which may be a result of more complex interplay between the weaker nudging currents which are more variable at this shallower site and the control of the plume by the wind direction and plume density and momentum. This is further supported by the size of the macrophyte beds in relation to the size of the watershed for

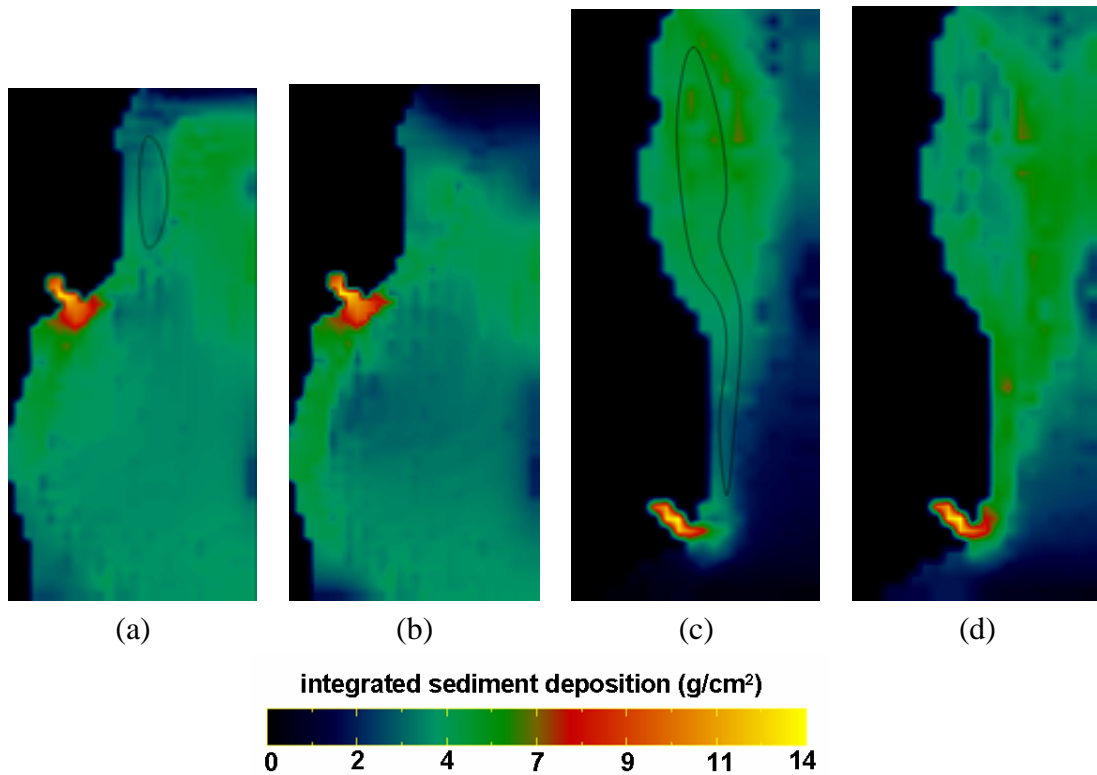


Figure 6.17: Integrated sediment deposition pattern at Sand Point Gully after 53 hours and Cottonwood Gully after 117 hours with and without macrophyte bed. The circled area indicated macrophyte beds location. (a) and (b) Sand Point Gully, (c) and (d) Cottonwood Gully.

each stream. The macrophyte beds at both stream mouths are about 8,000 m² although the Sand Point watershed (325 ha) is about 4.3 times larger than the Cottonwood watershed (76 ha). If we assume that the supply of sediment scales with the size of the watershed, then Cottonwood has much more focused and consistent sediment deposition pattern than Sand Point where sediment is deposited over wider area and into water too deep for macrophyte growth.

6.4.2 Niagara River plume in Lake Ontario

In this research, high-resolution hydrodynamic simulations were carried out during the period of two weeks in summer 2004 (June 5 - June 16) to demonstrate the mixing characteristics of Niagara River thermal plume in Lake Ontario. Time series flow currents from the large scale whole lake simulations were used as nudging for plume simulations. Eastward winds appear to be the dominant wind direction during early-mid June. Wind is relatively weak in most of the simulation time with an average speed of less than 1.5 m/s. However, there are some occasions where wind speeds reach up to 7 m/s.

Figures 6.18 and 6.19 show the predicted movement of the Niagara River plume during two weeks in summer 2004. Figure 6.18 shows the stage of fast extension of the Niagara River plume within the first two days of the simulation time. Figure 6.18 (c) illustrates the inflowing water flows out of the mouth to a distance of approximately 8.2 km. The flow is hydraulically controlled within the first 5 to 6 km from the mouth of the river. The magnitude of horizontal flow currents is strong during this momentum-dominated period with maximum speeds up to several tens of cm/s. Figure 6.19 shows the stages of transition and slow deflection of the Niagara River plume during the next ten days of the simulation time. A large clockwise eddy forming to the east of the Niagara River mouth is shown in Figure 6.19 (b). The eddy lasts for a few days (Figure 6.19 (b) and (d)). The spread of the Niagara River plume in Lake Ontario is determined by the prevailing wind conditions and lake circulation patterns. In Figure 6.19 (e), the river plume is diverted to the east, and the weakly buoyant plume responds to the prevailing winds and lakewide circulation forces.

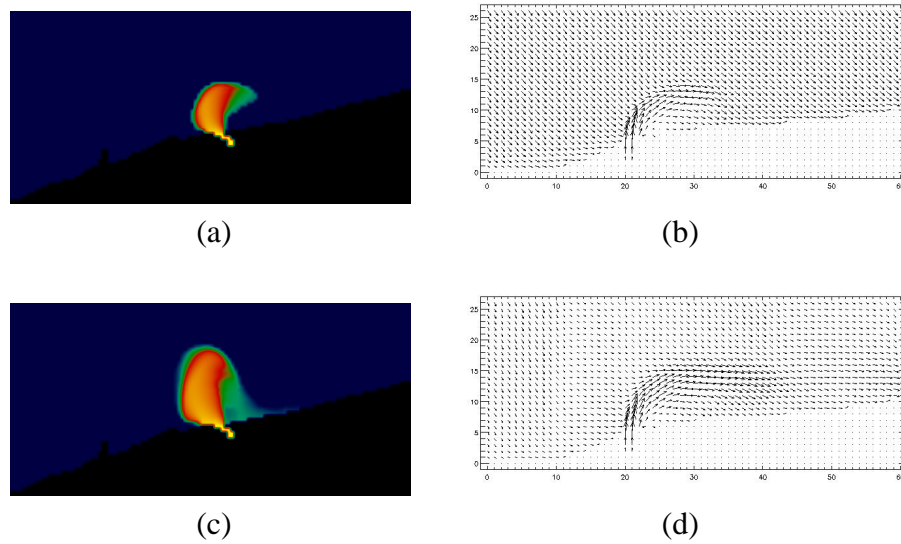


Figure 6.18: Simulated movement of the Niagara River Plume from a two-week simulation in summer 2004. Left column is the predicted movement of the plume and right column is the corresponding simulated surface flow currents. (a) & (b) June 5 (c) & (d) June 6.

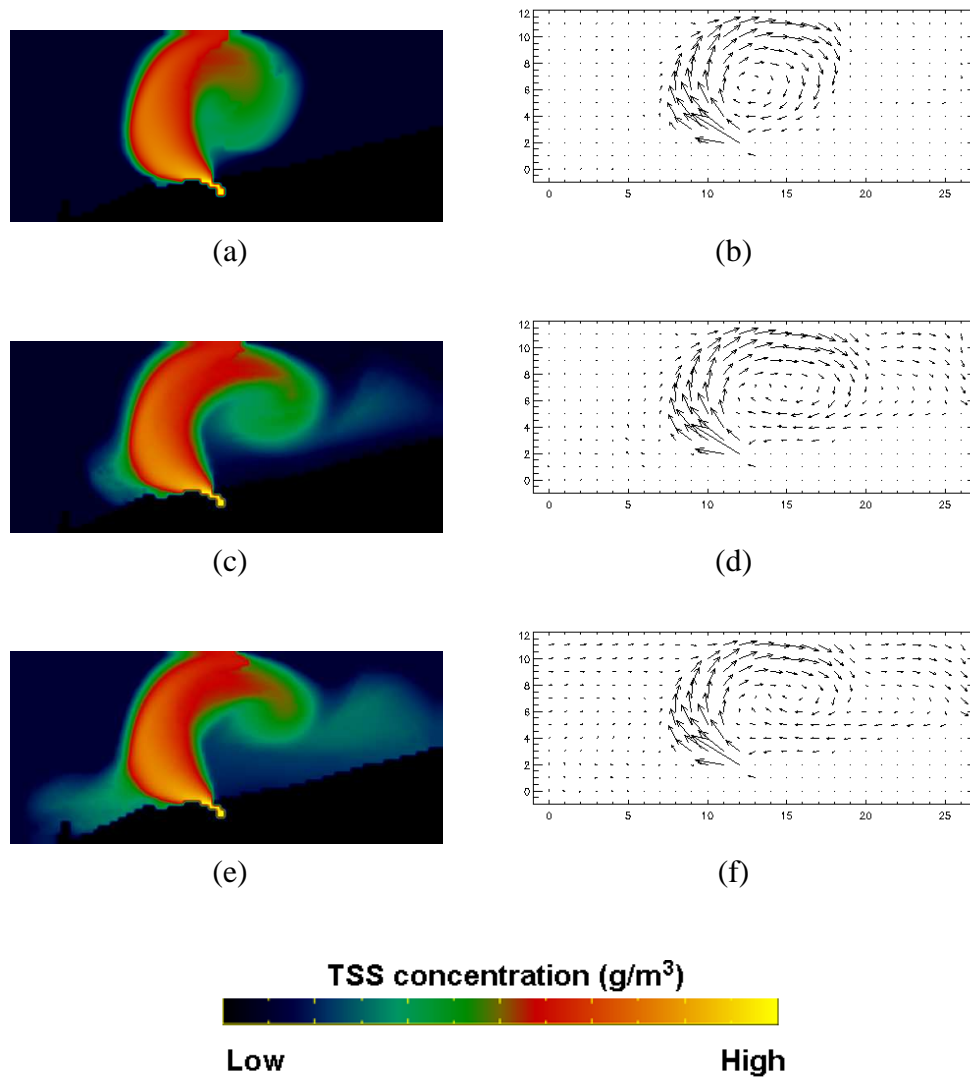


Figure 6.19: Simulated movement of the Niagara River Plume from a two-week simulation in summer 2004. Left column is the predicted movement of the plume and right column is the corresponding simulated surface flow currents. (a) & (b) June 7 (c) & (d) June 10 (e) & (f) June 16.

6.5 Impacts of a turbid plume on optical dynamics in Rochester Embayment

To evaluate the impacts of the Genesee River plume on the optical dynamics of the Rochester Embayment in Lake Ontario, high resolution plume simulations were carried out to study the transport and distribution of water constituents from the Genesee River plume. TSS was simulated as single-sized particles with diameter of 2 microns and density of 2.5 g/cm^3 and CDOM simulated as dissolved tracer in ALGE. The initial concentrations of TSS and CDOM in the lake were set to zero. The distribution of TSS and CDOM built up as particles and tracer enter the lake from the river. In this study, the TSS discharged into Lake Ontario from Genesee River is assumed to be fine particles with small settling velocities.

Figure 6.20 shows the distribution of TSS predicted by ALGE after two weeks of simulation for August 2006. The inflowing water flows out of the Genesee River mouth and spreads to the east, and the weakly buoyant plume responds to the prevailing winds and lake-wide circulation forces. Table 6.2 shows the simulated concentrations of TSS and CDOM at four stations on August 9, 2006. To validate our results from ALGE, modeled values were compared to concurrently measured concentrations of TSS and CDOM at four stations in our sampling area (Figure 6.21). The concentrations of TSS and CDOM varied significantly from near the river mouth to further offshore. The correlations of TSS concentration indicate that modeled values were slightly underestimating measured values except for station 1. This is consistent with setting the initial values in lake to zero. CDOM concentration correlations were also strong, illustrating the strong control the Genesee River can exert on the optical properties of the embayment.

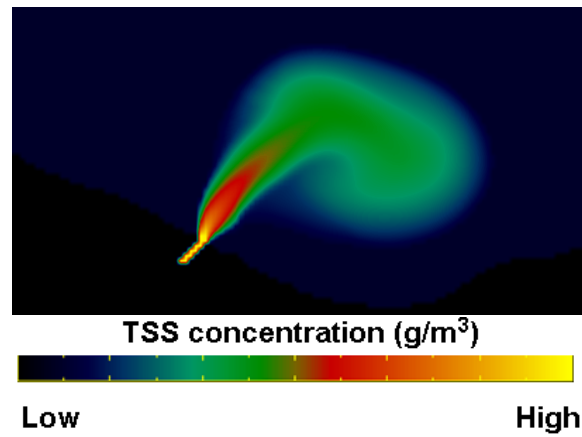


Figure 6.20: ALGE predicted TSS distribution showing the Genesee River plume. Yellow represents the high concentration and blue represents the low concentration.

Table 6.2: TSS and CDOM concentrations predicted by ALGE at four stations on August 9, 2006.

Location	Station 1	Station 2	Station 3	Station 4
TSS (g/m^3)	16.96	3.769	0.403	0.689
CDOM (scalar $a(350)$)	5.996	1.356	0.152	0.184

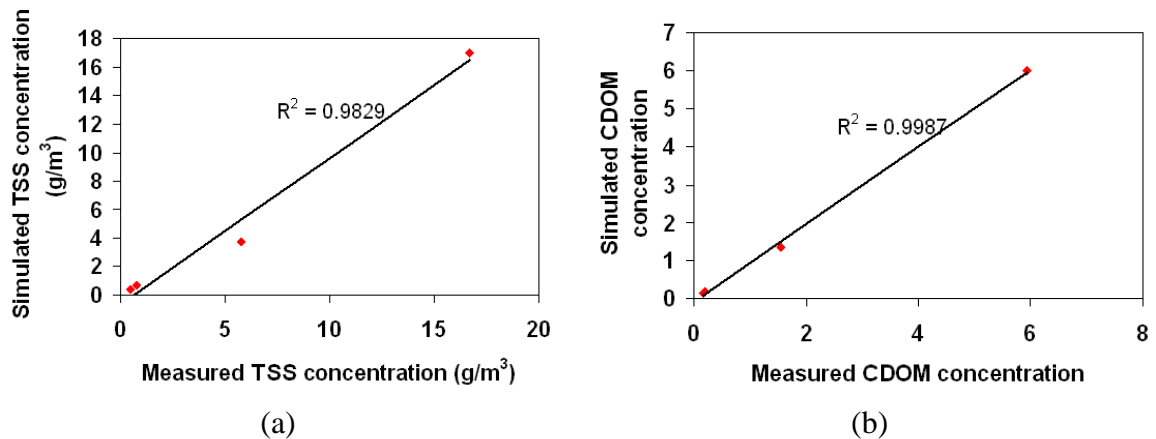


Figure 6.21: Relationship between measured and modeled TSS and CDOM concentrations.

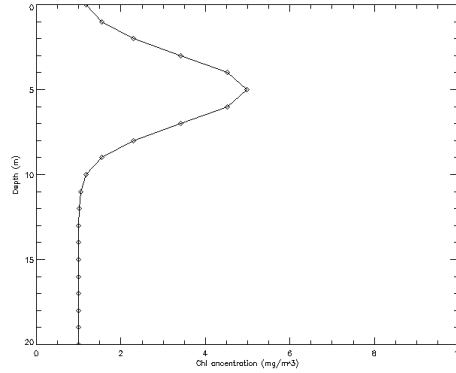


Figure 6.22: An example of Chl profile obtained from Eq. 6.1.

In absence of the algal growth model, the vertical profiles of CHL is approximated by a Gaussian function plus a constant background (Stramska and Stramski, 2005):

$$Chl(z) = Chl_0 + h/[\sigma(2\pi)^{0.5}]exp[-(z - z_{max})^2/2\sigma^2], \quad (6.1)$$

where $Chl(z)$ is the concentration profile of CHL, Chl_0 is the background value of CHL (mg/m^3), z_{max} is the depth of CHL maximum, and $Chl_{max}=h/[\sigma(2\pi)^{0.5}]$ determines the amplitude of CHL maximum above the value of Chl_0 . In this study, $h=20$, $\sigma=2$, $Chl_0=1.0$, and $z_{max}=5.0$. Figure 6.22 shows an example of CHL profile with the maximum CHL concentration at a depth of 5 m.

Figure 6.23 shows the Hydrolight predicted remote sensing reflectance values for three stations using the in situ measured IOPs in Figure 4.6, the concentration values in Table 6.2, and the estimated CHL profile. The Lake Ontario surface water reflectance of station 3 and 4 exhibit the characteristic spectrum of waters with low constituent concentrations. The sediment plume is characterized by high reflectance values in the green region. The reflectance of the plume-dominated station shows a decrease in the

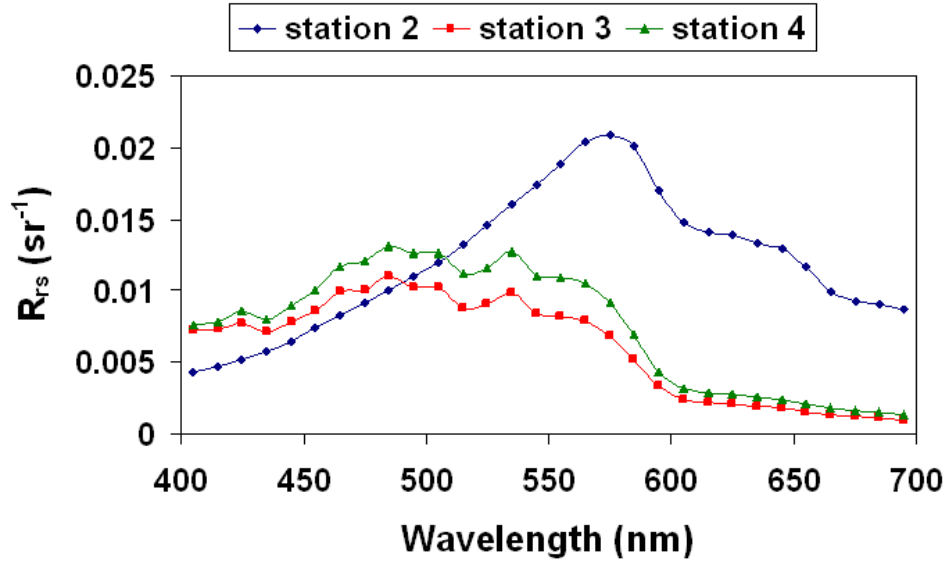


Figure 6.23: Modeled remote sensing reflectance at three stations.

blue region consistent with absorption due to large amounts of TSS and CDOM.

Figures 6.24 show the comparison of remote-sensing reflectance between the model prediction and the measured values for three stations around the embayment. The modeled and measured R_{rs} values tend to follow the same spectral trend for all three stations. The high concentrations of TSS and CDOM in areas near river mouth results in the characteristic low blue reflectance. Reflectance of station 3 and 4 is smaller in the red wavelengths from absorption by CHL and water. Range of the reflectance values in Figure 6.24 (c) is slightly greater than that in Figure 6.24 (b) due to the higher constituent concentrations. Figure 6.24 (c) shows some of the modeled R_{rs} are larger than the observed values at station 4. It is probably due to the estimate of CHL concentration is higher than the real values since station 4 is offshore. Reflectance values derived from MODIS 1 km scene for stations 3 and 4 are also shown in Figures 6.24 (b) and (c). MODIS derived reflectance values are smaller than the modeled values in blue and

green regions. Comparison to MODIS 1 km data is not good for either the measured or modeled reflectance both in magnitude and spectral shape. This may be a result of a misregistration between MODIS scene pixels picked to calculate the reflectance and the locations of stations 3 and 4 due to the low spatial resolution. In addition, the discrepancy may be caused by the lack of accurate correction of atmospheric effect for coastal areas. The modeled reflectance values are slightly smaller than the measured values for all stations. One reason could be the scattering coefficients and phase functions of TSS used to calculate the reflectance values are not proper for Lake Ontario due to the lack of consideration for the effects of various particle size distributions (PSD) on IOPs.

6.6 Updating ALGE with MODIS data via an EnKF

In this section, a test is performed at the simulation grid point at the mouth of the Genesee River to validate the performance of the EnKF method. Results from the model controlled runs are compared with the assimilation runs. Since the ensemble member simulations contain only the errors associated with the model inputs and assumptions, they provide a baseline of model accuracy without any satellite observations. The filtering estimate of the TSS concentration is compared with real observations which were obtained from one USGS water quality site (Charlotte Pump station, 431510077363501). Errors of estimates from a single ALGE run (called an open-loop run), the ensemble mean, and the EnKF estimate are compared. The difference between individual ensemble members and the mean of the assimilation results demonstrates the value of incorporating satellite measurements into the model estimation of TSS concentration.

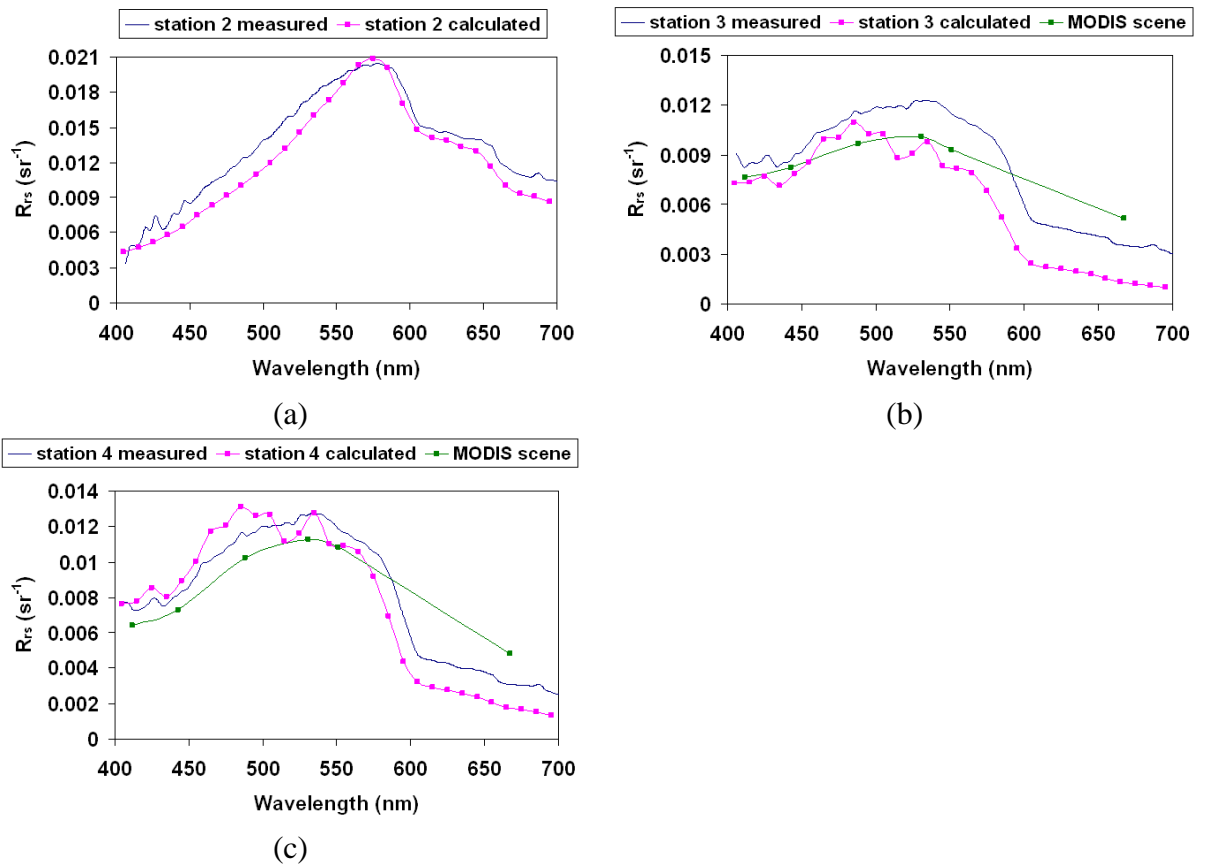


Figure 6.24: Measured and modeled remote sensing reflectance at stations 2, 3, and 4.

The estimate of plume dissipation at the mouth of the Genesee River for all grid points by 2D spatial EnKF filtering is shown in this section as well. Due to the lack of ground truth in the lake, it is difficult to validate the spatial filtering quantitatively. The estimate of the plume dissipation from one open-loop run, MODIS data, and 2D EnKF filtering are compared.

6.6.1 Grid point test at the Genesee River mouth

Ontario beach has been a focus of research for unsafe conditions for swimmers due to high flow of Genesee River and poor water quality (Knauf, 2003). Streamflow has been measured at the Genesee River at Rochester two or three times per week. Water samples are also collected from the Genesee River at Charlotte Pump station a few times per week by automatic sample to provide water-quality data, such as TSS concentration (Sherwood, 2005). During the swimming season of 2003, heavy rainfall causing Genesee river storm water in combination with turbid conditions in the near shore resulted in beach closures on 25 days (Knauf, 2003).

Figure 6.25 shows the relation between the state variable and the measurement. TSS concentration is converted to remote-sensing reflectance using the measurement model Hydrolight. Higher TSS concentration results in higher reflectance values. Reflectance increases significantly when the TSS concentration is low and approaches saturation when TSS concentration is greater than 150 g/m^3 . The EnKF was applied using simulated and observed remote-sensing reflectance to update the TSS concentration profile estimates.

To validate the benefit of remote updates, a series of ALGE controlled simulations in which errors were intentionally introduced into model forcing variables are performed. The model controlled simulations (or open-loop runs) are defined as the

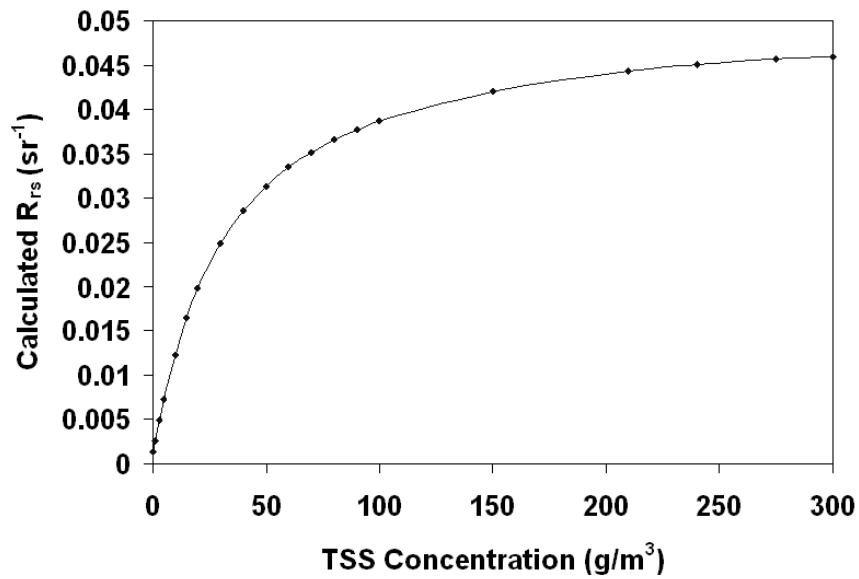


Figure 6.25: Relationship between TSS concentration (g/m^3) and remote-sensing reflectance R_{rs} generated by Hydrologht at the wavelength of 645 nm. Concentration of CHL and CDOM are assumed to be constants.

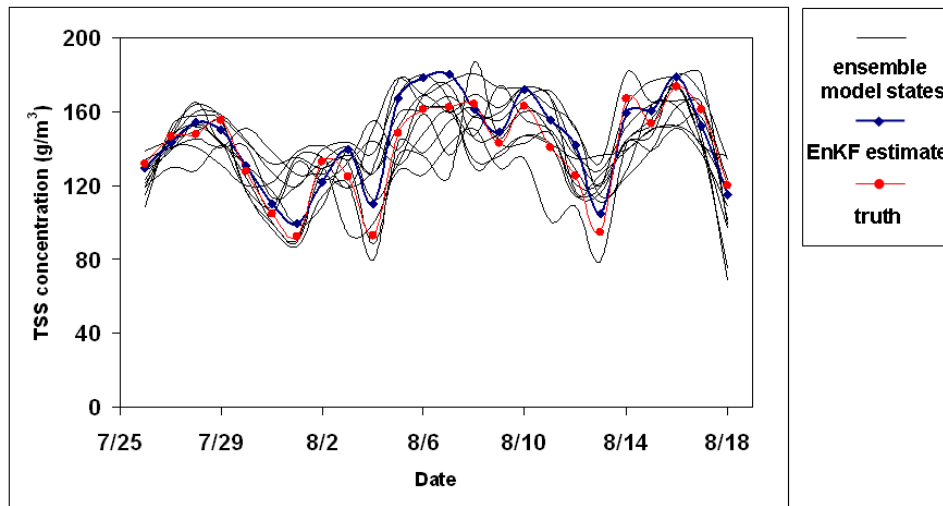


Figure 6.26: TSS concentration (g/m^3) at mouth of the Genesee River during the study period from ALGE controlled simulations and truth. The ensemble of model simulations is given in black, the truth is given in red, and the filtering estimate is given in blue.

runs which are driven by various combinations of forcing inputs. In each simulation, the model was run at one hour time steps. Figure 6.26 shows the TSS concentration (g/m^3) at the grid point at the mouth of Genesee River during the study period for 12 ensemble members in black. Also shown in Figure 6.26 for comparison is the observations of TSS concentration obtained from the USGS water quality site. Results from the ensemble runs match the basic trend of the real observations. The ensemble largely envelopes the truth and over- or underestimate the TSS concentration. Since the truth falls within the ensemble all the times, it is evidence to show the previous choices for forcing variables are very reasonable. Figure 6.26 also shows the EnKF TSS concentration estimate at the grid point at the mouth of Genesee River during the study period in blue.

Kalman gain determines the inclusion of measurements into state estimates and is

influenced by the uncertainty in the prior state estimate and the uncertainty in the measurement estimate. According to the definition of Kalman gain, as the measurement error covariance R approaches zero, the Kalman gain moves towards $\frac{1}{H}$. The higher the prior uncertainty, the less confidence the EnKF has in the state estimate and thus the more it includes information from the measurements. As the *a priori* estimate covariance P_k^- approaches zero, the Kalman gain moves towards zero. Higher uncertainty in the measurements indicates less of the information from the measurements should be taken into account.

In Figure 6.27, the Kalman gain value at the simulation grid point at the mouth of the Genesee River is plotted. During the simulation period, eight MODIS scenes are able to be used as a feedback via EnKF into ALGE. The values of the Kalman gain reflect how much information of MODIS data can be included into TSS concentration estimate by ALGE. On the days of July 29, August 8, August 14, and August 15, Kalman gain is large with the absolute values from 750 to 1400. For those days the EnKF will emphasize the MODIS data more than the prior state estimate. On the other hand, on the days of July 30, August 13, and August 16, absolute values of the Kalman gain are less than 100. For those days the information from the MODIS data is not that trustworthy. For the last day of simulation period, the Kalman gain has the value of about 250 which indicates MODIS data is partly taken into account.

One of the most straightforward ways to validate the performance of the EnKF is to compare the EnKF estimate to the true state. In this case these are daily TSS values from the Charlotte pumping station that provides the true state values. Figure 6.28 shows the errors of the EnKF estimate on each day during the study period at the river mouth. Errors shown in the figure are the absolute values of the difference between the real observations and the EnKF estimate. The red rectangles indicate a day

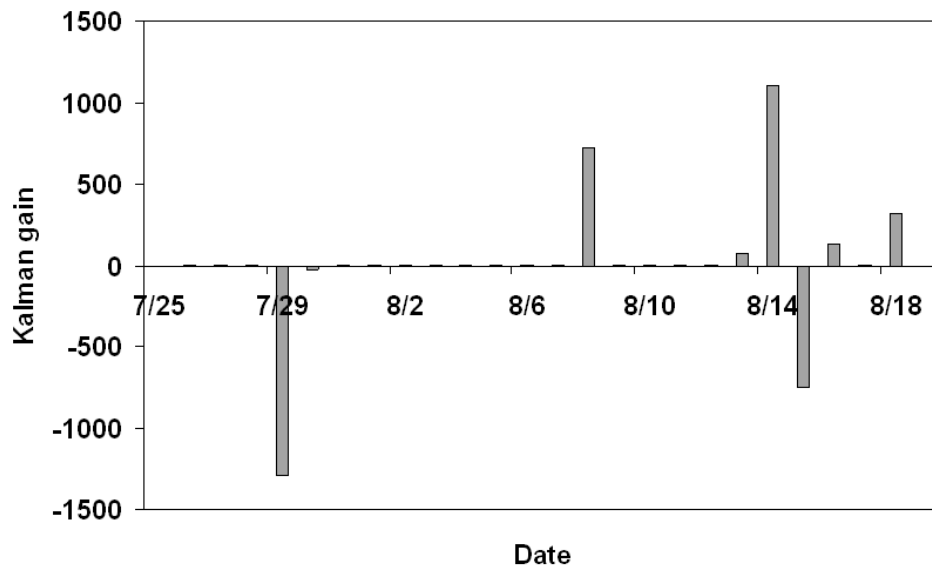


Figure 6.27: Kalman gain calculated by the EnKF during the study period. Values shown are for point-scale test.

MODIS is available for measurement updating. As expected, at times when MODIS data is available, the EnKF is very effective at correcting the TSS concentration to the truth and reducing the errors. Figure 6.28 shows the increase in uncertainty due to prediction and the decrease in uncertainty due to correction. The error of the EnKF estimate typically increases after the correction. In other words, the EnKF prediction deteriorates until a new observation becomes available. Kalman gain determining the balance of uncertainty between the system and measurement results in the error always reaches the local minimum when the MODIS data is available. The time between updates also affects the magnitude of the error change.

One technique used to improve model forecast is to calculate the mean of the ensemble prediction (Houtekamer and Derome, 1995) (Murphy, 1988) (Szunyogh and Toth, 2002) (Whitaker and Lough, 1998). Figure 6.29 shows estimate errors in TSS

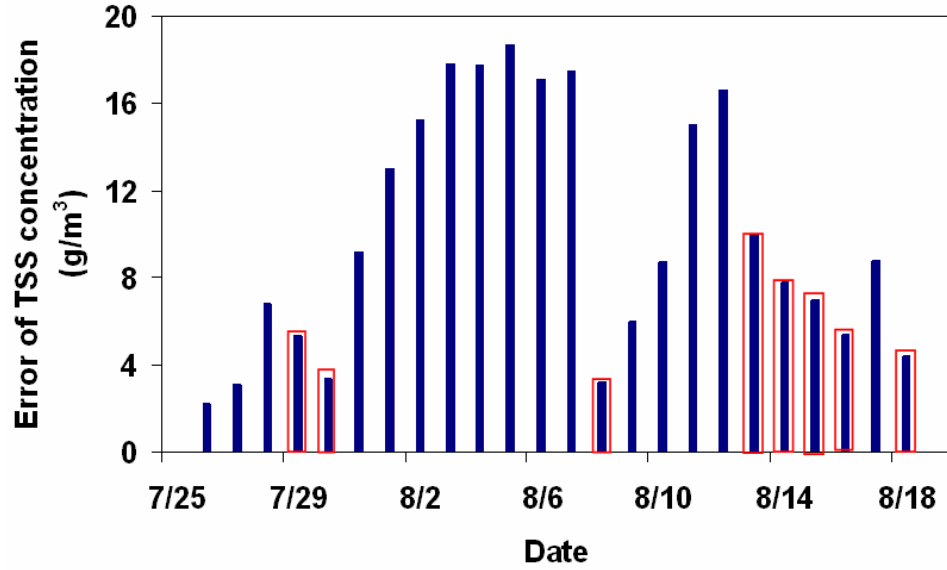


Figure 6.28: Errors of TSS concentration (g/m^3) for the EnKF estimate with 11 ensemble members for the period of July 26 to August 18.

concentration for the ensemble mean, one open-loop run, and the EnKF on each day during the study period. The open-loop run was randomly chosen from 12 ensemble members. The root-mean-square error (rmse) was calculated for each of them based on:

$$RMSE = \frac{1}{24} \sqrt{\sum_{i=1}^{24} (x_i - x_o)^2}, \quad (6.2)$$

where x_i is either the ensemble mean, the singel ALGE run, or the EnKF estimate, and x_o is the real observations of TSS concentration. The RMSE of TSS concentration for the open-loop simulation is 4.32 m/s^3 compared with for the EnKF estimate 2.33 m/s^3 . It shows MODIS assimilation by EnKF improves the RMSE of TSS concentration estimate for the grid point at the river mouth. The result also shows the mean improves the model estimate nearly as well as EnKF. Figure 6.30 shows the cumulative errors

of TSS concentration from the ensemble mean, one open-loop run, and the EnKF estimate. The open-loop run of ALGE has the largest error towards the end of study period. Results from the mean of the state estimate from all the ensemble members show slightly smaller errors than the EnKF estimate. However, the EnKF estimation accuracy can be easily traded off with computational effort by simply adjusting the number of ensemble members. Thus it is possible for the EnKF to produce better estimation if the ensemble space is increased. This examination of the performance of the ensemble mean versus the EnKF also illustrates two predictable results. First, it is reasonable to assume that the ensemble mean should be a good prediction and this is apparent from Figure 6.26. Since the ensemble bounds the truth the mean will lie closer to the truth than any single run. Second the mean has a smoothing action so that it does not capture real excursions in the TSS that occur day to day as a result of the episodic nature of storms. On the other hand, when data is available, the EnKF adjustment from one day to the next can be large if the Kalman gain is large and the maximum EnKF estimate error is considerably smaller than the maximum ensemble mean error.

6.6.2 Spatial update of plume dissipation via 2D filtering

Spatially demonstrating the improvements gained by using remote-sensing data as feedback can be difficult. Uniform spatial distribution of the dynamic inputs to ALGE has a large impact on the ability to model plume dissipation accurately. Remotely sensed data are also subject to various types of error sources. In addition, lack of observations from station networks as truth makes it difficult to quantify the benefits of using MODIS data in the EnKF method.

Figures 6.31 and 6.32 show the comparison of TSS dominated plume dissipation at

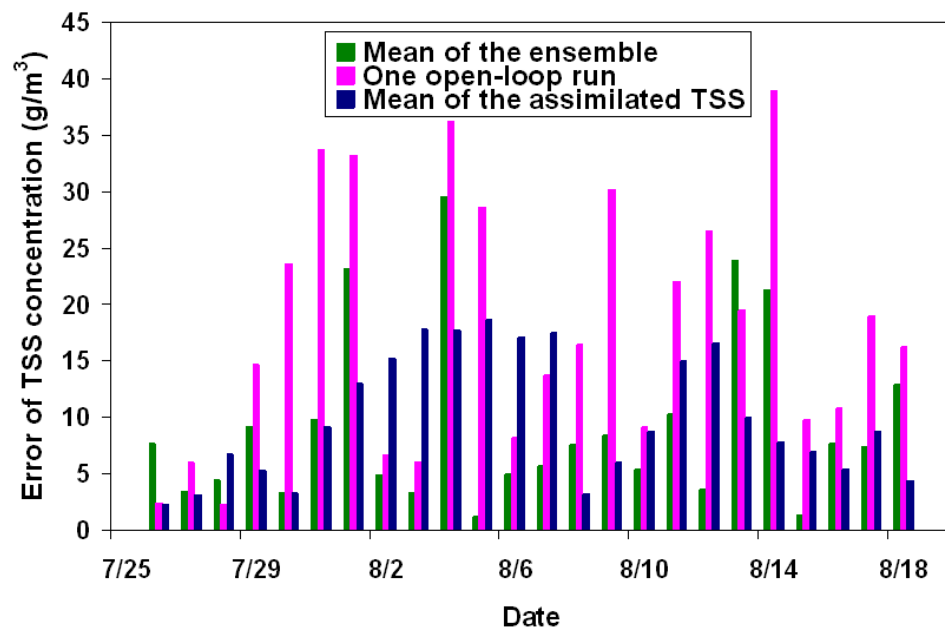


Figure 6.29: Errors in TSS concentration (g/m^3) for the mean of the ensemble (green), one ALGE open-loop run (magenta), and the EnKF estimate (blue) for the period of July 26 to August 18.

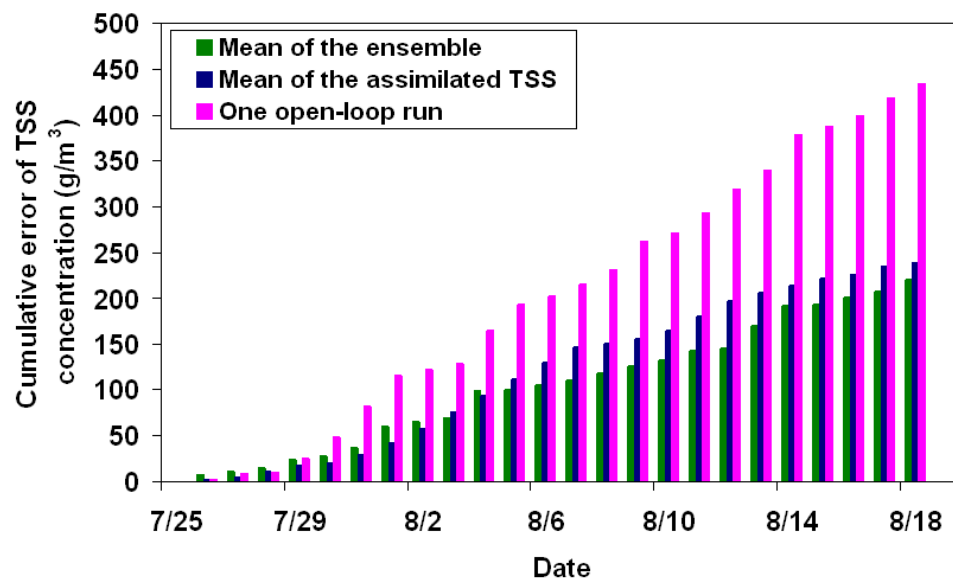


Figure 6.30: Cumulative errors of TSS concentration (g/m^3) for the mean of the ensemble (green), one ALGE controlled open-loop run (magenta), and the EnKF estimate (blue) for the period of July 26 to August 18.

the Genesee River mouth during the study period of July 26 to August 18. Sediments are simulated to show where the turbid plume flows into Lake Ontario. The colors represent sediment concentration in the water. To illustrate the impacts of using MODIS data as a feedback, only results for eight days in which MODIS data is available are shown in the figures. The first column in Figures 6.31 and 6.32 demonstrates MODIS derived surface reflectance showing the plume dissipation. MODIS scenes show that most of the time the plume is flowing eastward, but on some days wind and lake currents can push the plume westward onto the beach. Estimated plume dissipation from the mean of 12 ensemble members is shown in the second column. The EnKF estimate is given in the third column.

On July 29, the result of assimilating the MODIS data can be seen in the EnKF estimate. On next day, the ensemble mean shows the plume flowing slightly westward while MODIS shows the plume flowing eastward. The filtering estimate updates the direction from westward to eastward, but without a tail. The long tail is possibly due to finer particles in water which is not included in the ALGE simulation. On August 8, the results of the MODIS scene, the ensemble mean, and the filtering estimate are similar. On August 13th, the EnKF updates the direction of plume dissipation in a westward direction, but the area on the upper-left corner with high reflectance value is missing. That part of the plume shown in MODIS scene is possibly due to resuspension originating further west. On the 14th, 15th, and 18th, the difference between the filtering estimate and the ensemble mean clearly shows the result of using MODIS as feedback. However, on the 16th, the EnKF estimate takes more of the prior state estimate from the ALGE open-loop runs into account. This matches the fact that there might be an atmosphere correction error on that day for MODIS. Overall, the results show that with more observations being assimilated into ALGE, the EnKF works better towards the

end of the study period.

Since the ensemble mean has a smoothing action, it does not capture real excursions in the TSS that occur day to day as a result of the episodic nature of storms. On August 14 and 18, the EnKF adjustment is large enough to show the long tails as a result of the storm events by using MODIS data as feedback. The measure of estimation error is the difference between the true state and the estimate at any given time and location. To validate the performance of the EnKF spatial filtering, MODIS data is applied as the true state and compared with the ensemble mean and the EnKF estimate because of the lack of the ground truth in the lake. The TSS concentration is converted to the remote-sensing reflectance by a look-up table shown in Figure 6.25. On August 18, the MODIS scene has a maximum remote-sensing reflectance of 0.0452, while the maximum of the EnKF estimation is 0.0441. The RMSE averaged over all pixels in the study area is a convenient method to evaluate the EnKF method. The reflectance is assumed to be zero on the land. The number of the pixels in the study area is 115×62 . The spatial RMSE of remote-sensing reflectance between the MODIS data and the ensemble mean is 0.0083. It reduces to 0.0035 when comparing the 2D filtering result and the MODIS data. Results show that the EnKF estimate RMSE is considerably smaller than the ensemble mean RMSE. Results also suggest that the EnKF improves the hydrodynamic simulation not only in point scale, but also in spatial domain.

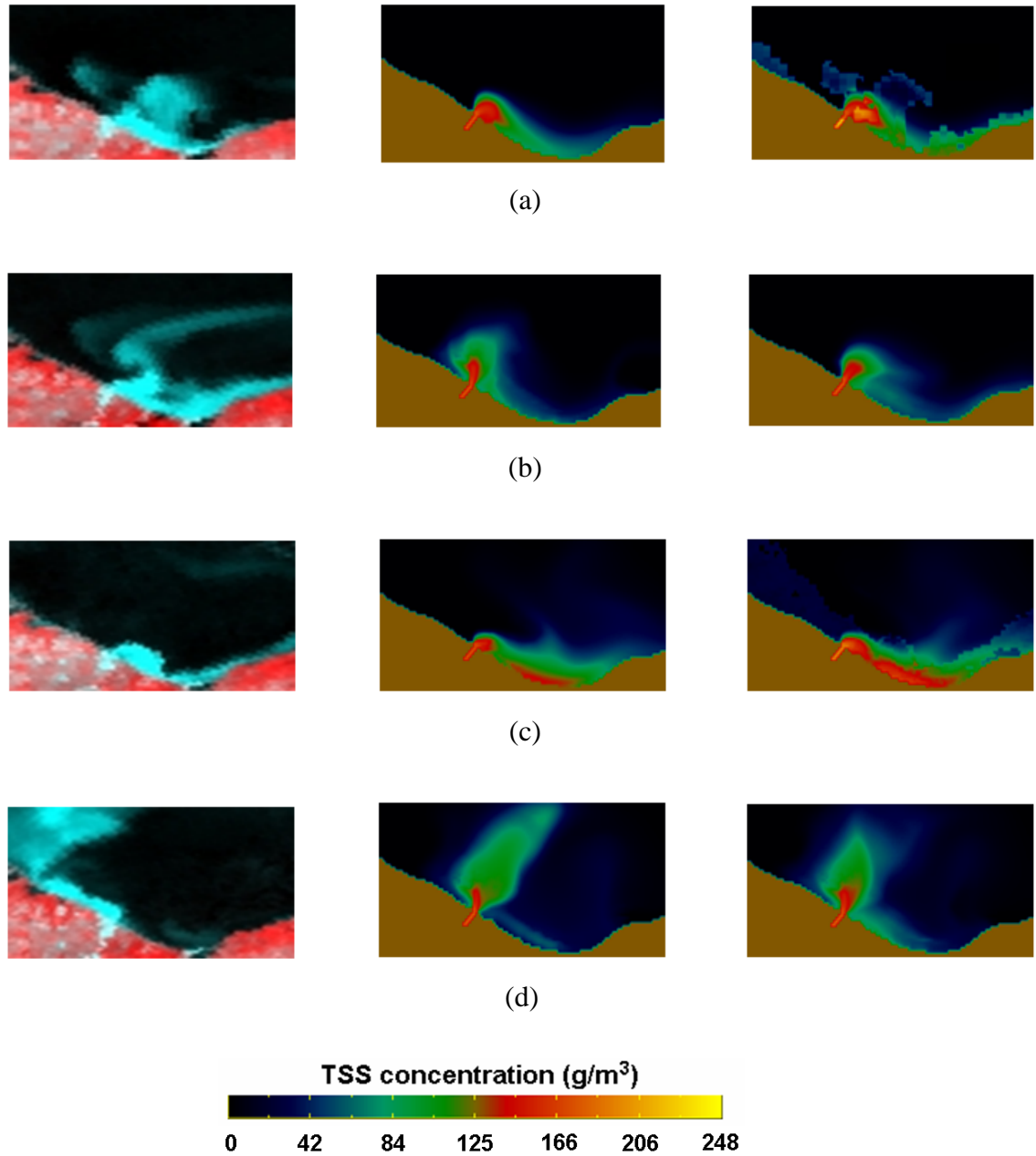


Figure 6.31: TSS dissipation in the Rochester Embayment on (a) July 29 (b) July 30 (c) August 8 (d) August 13, 2003. First column: MODIS derived, second column: ensemble mean, third column: EnKF estimate with 12 ensemble members.

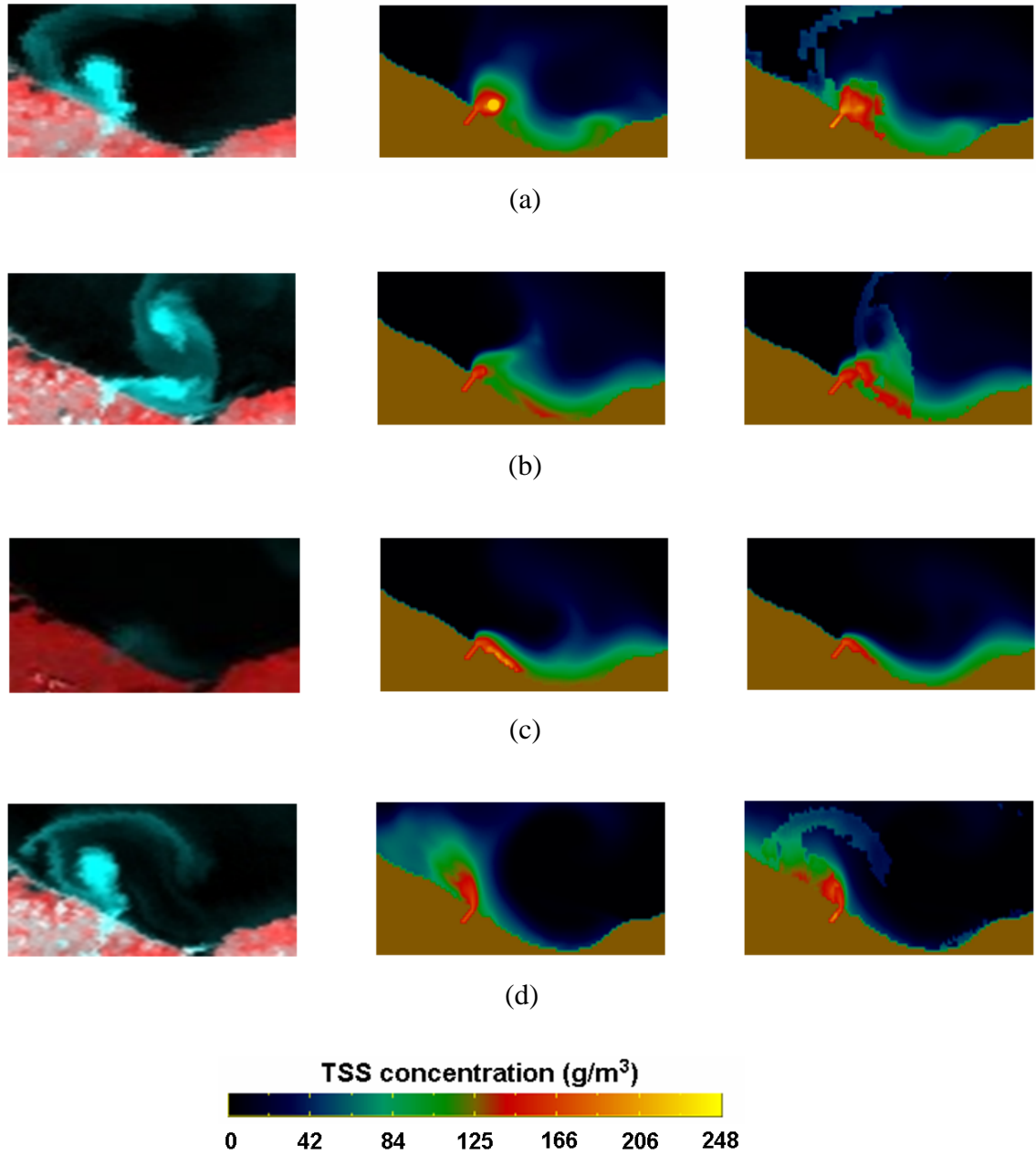


Figure 6.32: TSS dissipation in the Rochester Embayment on (a) August 14 (b) August 15 (c) August 16 (d) August 18, 2003. First column: MODIS derived, second column: ensemble mean, third column: EnKF estimate with 12 ensemble members.

Chapter 7

Conclusions and Recommendations

7.1 Conclusions

A coupled water quality modeling system using DDDAS concepts that assimilates remote sensing data into a hydrodynamic model was developed and tested. The modeling system includes the hydrodynamic model (ALGE), a radiative transfer model (Hydrolight), and remote imagery (MODIS) as a dynamic feedback. The DDDAS is implemented through an Ensemble Kalman Filter (EnKF) with a small ensemble space.

Validation of the ALGE simulations used airborne thermal images from RIT's MISI to assess the pattern of surface temperature across the Conesus Lake in mid May and late June 2003. The comparison of lake surface temperature between the thermal imagery and the ALGE simulation shows an RMSE of 0.9 °C. MISI's on-board blackbody and transducer system can be calibrated to within 0.3 K. Due to the uncertainties from approximations in ALGE and the input data used by ALGE, these results for matching remote sensing data with a two or three month simulation are very good.

The study of thermal structure in Conesus Lake in general shows a good agreement between measurements and simulations. The RMSE of averaged surface temperature between measurements and simulations is 0.8 °C. The simulated subsurface temperatures show a better agreement with the observed temperature profiles at a depth above 9 m since the inaccurate meteorological conditions might introduce excess vertical mixing. In general ALGE is capable of simulating the vertical thermal structure in Conesus Lake.

The maps of modeled seasonal water circulation in Conesus Lake show strong effects of lake shape and orientation on the circulation pattern. When the wind is blowing from west-northwest perpendicularly to the lake, two strong convergences are seen easily in both the north and the south basins. In particular, the circulation is stronger in early spring than in summer. The water circulation at different depths for the entire lake also shows the effect of wind for a typical summer stratified lake.

With boundary conditions for currents generated from a whole lake simulation, the results of specific stream plume simulations using a nested grid approach show the local fate of model-generated passive tracer and the pattern of sediment deposition at stream mouths. Tracer via streams and creeks are controlled by the vertical mixing, the net effect of local event-driven circulation, and main flow in Conesus Lake. Macrophytes generate most of the resistance to flow because of the drag force they exert when water moves through vegetation. The sediment deposition patterns show the effect of macrophytes on enhancing sedimentation near stream mouths. Results show that water circulation at stream mouths during storm events is highly current-driven at Cottonwood Gully and can explain the location of the macrophyte beds with respect to the stream mouth. However, at Sand Point Gully it appears the combination of wind direction and nudging currents affect the plume resulting in a less focused plume.

These results suggest this type of detailed study of individual stream plumes is necessary to understand the linkage between tributary inputs of nutrients and macrophyte distribution in the lake.

Further testing of ALGE used a four-month hydrodynamic simulation for Lake Ontario during April to July in 2004 to study the thermal structure and water movement affected by the thermal bar in spring and summer. The simulated thermal cycles show the spatial and temporal characteristics of temperature distributions during the thermal bar period. The entire lake simulation also shows the upwelling and downwelling along the shoreline during the summer stratified season under the significant wind events. The simulated large-scale current circulations show the thermally-driven coastal currents in the counterclockwise direction under the light wind and the wind-driven coastal currents in the wind direction under the strong wind. A study of nearshore plume behavior was also carried out at the Niagara River mouth to investigate the plume extension. The simulated plume dissipation shows a three stage mixing process similar to that shown for published results, and providing further validation of ALGE.

The effects of the turbid Genesee River plume on the optical dynamics in the Rochester Embayment are studied using the DDDAS coupled modeling system. The comparisons of the concentrations of TSS and CDOM from field stations show overall good agreements between the estimates from lab analysis and the model output. The simulated reflectance values of the plume-dominated station show the characteristic spectrum of turbid waters with absorption and scattering due to large amounts of TSS and CDOM. The Lake Ontario surface water reflectance of offshore stations exhibits the characteristic spectrum of waters with low constituent concentrations. Overall, the agreement between the 250 m MODIS reflectance and the modeled reflectance values is good although the data set is small. MODIS data with a spatial resolution of 1 km is

not sufficient for most applications in coastal areas. With the higher spatial resolution, MODIS 250 m band is sufficient to detect features in coastal areas and very promising for coastal monitoring applications.

A TSS data assimilation study was performed in which MODIS 250 m reflectance data was used as a dynamic feedback to update the hydrodynamic model ALGE. EnKF was implemented to improve ALGE's estimate of TSS concentration by balancing the uncertainties associated with meteorological and hydrological forcing variables and MODIS observations. The modeling system was tested at the single simulation grid point at the mouth of the Genesee River to validate the performance of the EnKF method. The evaluation of the EnKF method is based on the comparison between the open-loop estimate, the ensemble mean, the EnKF estimate, and the truth for the grid point at the Genesee River mouth. The RMSE of TSS concentration for the open-loop simulation is 4.32 g/m^3 vs. 2.33 g/m^3 for the EnKF estimate. The difference between individual ensemble members and the assimilation results shows the improvement gained by assimilating MODIS measurements into the hydrodynamic model via EnKF. However, simply taking an average of all the ensemble model states reduces errors as well, thus, the mean of the ensemble is also a choice to produce better estimates of state variables if not enough observations are available as feedback. However, since the mean has a smoothing action, it does not capture real excursions in the TSS that occur day to day as a result of the episodic nature of storms.

Further validation was undertaken to examine the effects of assimilating MODIS data on the estimate of spatial plume dissipation. This time the plume dissipation observed in the MODIS scene was compared to the ensemble mean and the EnKF estimate with 12 ensemble members. The EnKF is able to recover TSS concentration over an 24-day simulation period with observations available on eight days. The spatial

filtering for all simulation grid points via an EnKF is capable of capturing the episodic nature of storm event by using MODIS data as feedback. The spatial RMSE (considering all grid points) of remote-sensing reflectance between the MODIS data and the ensemble mean is 0.0083. The RMSE reduces to 0.0035 when comparing the spatial EnKF result and the MODIS data. These results suggest that the EnKF improves the hydrodynamic simulation not only at the river mouth, but also across the spatial domain of the Rochester Embayment.

7.2 Recommendations

The approach described in this research may be enhanced to potentially provide better accuracy of the estimate from hydrodynamic simulations. This section describes the main areas for further research.

To assimilate MODIS data, the measurement model, Hydrolight, is applied to convert model state estimate (TSS concentration) to measurement (R_{rs}). However, the effects of sediment particle size distribution on scattering coefficients and phase function was not considered in the calculation of R_{rs} via Hydrolight. The scattering coefficients and phase functions have a strong effect on the reflectance data. The proper particle size distributions for Lake Ontario are needed to be taken into account to obtain more accurate scattering properties.

In this research, only remote observations from day time visible channels MODIS are used as dynamic feedback into the hydrodynamic model to provide better estimate of model states. Recent research shows that the error of the EnKF estimates decreases with time as more and more observations are assimilated (Reichle et al., 2002). One of the recommendations is that other satellite data and field observations be used as

feedback whenever they are available. One choice for satellite data assimilation is Landsat. The Landsat series of instruments [Thematic Mapper (TM), Enhanced Thematic Mapper (ETM+)] have a spatial resolution of 30 m, which is good for coastal water monitoring. TM and ETM+ have three bands to represent blue, green, and red in visible region and two infrared bands as well. However, due to a revisit time of about 16 days, Landsat data is not assimilated into hydrodynamic modeling in this research. But combining Landsat data and MODIS data, the EnKF will have more chances to update the hydrodynamic modeling and provide a more accurate estimate of state variables. Since MODIS thermal data with 1 km spatial resolution shows the plume dissipation at night time, those data can be assimilated into the hydrodynamic model, providing more opportunities for obtaining cloud free data. In addition, field observations from sensors in the river or on buoys can be used as a dynamic feedback to the hydrodynamic model as well.

The ensemble of model states is generated by treating model forcings as stochastic variables. Bias errors were not included but could be added to future simulations. One possible source of forcing bias could be wind direction at the lake versus wind direction at the Rochester Airport which is about 15 km from the lake. Other than forcing variables, model parameters contribute to system uncertainty as well. In the future, an artificial error can be introduced to some model parameters to generate a wider spread of the ensemble to capture the uncertainty inherent in the system model.

Reichle et al. (2002) indicate that the errors of the EnKF estimate decrease and converge with increasing ensemble size. However, increasing the ensemble space may easily increase the computational complexity. Thus, adjusting the number of ensemble members to obtain better estimation accuracy should consider if computational time will be significantly increased.

Bibliography

- Albert, P. (2004). *Remote sensing of atmospheric water vapor for numerical weather prediction*. PhD thesis, University of Berlin.
- Alwis, D. A. (1999). Simulation of the formation and propagation of the thermal bar on Lake Ontario. Master's thesis, Rochester Institute of Technology.
- Andreadis, K. M. (2004). Assimilating remotely sensed snow observation into a macroscale hydrologic model. Master's thesis, University of Washington.
- Beletsky, D. and Schwab, D. J. (2001). Modeling circulation and thermal structure in Lake Michigan: Annual cycle and interannual variability. *J. Geophys. Res.*, 106:19,745 – 19,771.
- Bloomfield, J. (1978). Lakes of New York State: Ecology of the Finger Lakes.
- Bukata, R. P., Jerome, J. H., Kondratyev, K. Y., and Pozdnyakov, D. V. (1995). *Optical properties and remote sensing of inland and coastal waters*. CRC Press.
- Chang, C. (2003). Thermal bar dynamics in Lake Ontario. Master's thesis, University of New York at Buffalo.

- Crosson, W. L., Laymon, C. A., Inguva, R., and Schamschula, M. P. (2002). Assimilating remote sensing data in a surface flux-soil moisture model. *Hydrological Processes*, 16:1645 – 1662.
- Darema, F. (2004). Dynamic Data Driven Applications Systems: A new paradigm for application simulations and measurements. In *Computational Science - ICCS 2004: 4th International Conference, Kraków, Poland, June 6 - 9, 2004, Proceedings, Part III*, pages 662 – 669. Springer Berlin/Heidelberg.
- D'Auito, P., Makarewicz, J. C., and Bosch, I. (2004). The impact of stream nutrient loading on macrophytes and metaphyton in Conesus lake, USA. *Verh. Internat. Verein Limnol.*, 29:1373 – 1377.
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *J. Geophys. Res.*, 599:10143 – 10162.
- Evensen, G. (2003). The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53:343 – 367.
- Fairbanks, R. R. (1999). *A characterization of the impact of clouds on remotely sensed water quality*. PhD thesis, Rochester Institute of Technology.
- Fischer-Antze, T., Stoesser, T., Bates, P. D., and Olsen, N. R. B. (2001). 3d numerical modeling of open-channel flow with submerged vegetation. *Journal of Hydraulic Research*, 39:3903 – 3910.
- Forest, H. S., Wade, J. Q., and Maxwell, T. F. (1978). *Lakes of New York State*, chapter The limnology of Conesus Lake, pages 147 – 150, 163 – 167. Academic Press, Inc.

- Garrett, A. J. (1983). Drainage flow predictions with a one-dimensional model including, canopy, soil and radiation parameterizations. *J. Climate and Applied Meteorology*, 22:79 – 90.
- Garrett, A. J. (1995). ALGE: A 3-D thermal plume prediction code for lakes, rivers and estuaries. Savannah River Technology Center, Aiken, South Carolina.
- Garrett, A. J. (2002). Analyses of MTI imagery of power plant thermal discharge. In *Imaging Spectrometry VII, Proceedings of SPIE*, page 12.
- Garrett, A. J. and Hayes, D. W. (1997). Cooling lake simulations compared to thermal imagery and dye tracers. *J. Hydraul. Eng.*, 123:885 – 894.
- Garrett, A. J., Hayes, D. W., and Bollinger, J. S. (2005). Hydrodynamic modeling of tritium transport in flooded Savannah River swamp. In *American Nuclear Society 2005 Winter Annual Meeting*.
- Garrett, A. J., Irvine, J. M., and King, A. D. (2000). Application of multispectral imagery to assessment of a hydrodynamic simulation of an effluent stream entering the Clinch River. *Photogrammetric Engineering & Remote Sensing*, 66:329 – 335.
- Gelb, A., editor (1974). *Applied optimal estimation*. MIT press, Cambridge, MA.
- Gerber, D. T. and Les, D. H. (1994). Comparison of leaf morphology among submersed species of myriophyllum (Haloragaceae) from different habitats and geographical distributions. *American Journal of Botany*, 81:973 – 979.
- Houtekamer, P. L. and Derome, J. (1995). Methods for ensemble prediction. *Monthly Weather Review*, 123:2181 – 2196.

- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82:35 – 45.
- Knauf, C. L. (2003). 2003 Ontario beach monitoring report. Monroe County Health Department, Bureau of Environmental Quality.
- Li, Y., Vodacek, A., Raqueno, N., Kremens, R., Garrett, A. J., Bosch, I., Makarewicz, J. C., and Lewis, T. W. (2007). Circulation and stream plume modeling in Conesus lake. *Environmental Modeling and Assessment*. Accepted.
- Makarewicz, J. C. (2001). Trophic interactions: Changes in phytoplankton community structure coinciding with alewife introduction. *Verh. Internat. Verein Limnol.*, 27:7.
- Makarewicz, J. C., Bosch, I., and Lewis, T. W. (2003). Update of soil and nutrient loss from subwatersheds of Conesus Lake: 2001. Technical report, SUNY Brockport, Brockport, New York.
- Malm, J., Grahn, L., Mironov, D., and Terzhevik, A. (1993). Field investigation of thermal bar in Lake Ladoga, spring 1991. *Nordic Hydrology*, 24:339 – 358.
- Masse, A. K. and Murthy, C. R. (1990). Observations of the Niagara River thermal plume. *J. Geophys. Res.*, 95:16,097 – 16,110.
- McLaughlin, D. (1995). Recent advances in hydrologic data assimilation. *Rev. Geophys.*, 37:977 – 984.
- Mellor, G. L. and Yamada, T. (1982). Development of turbulence closure model for geophysical fluid problems. *Rev. of Geophys. and Space Phys.*, 20:25.

- Mobley, C. D. and Sundman, L. K. (2000a). Hydrolight 4.1 Technical Documentation. Sequoia Scientific, Inc.
- Mobley, C. D. and Sundman, L. K. (2000b). Hydrolight 4.1 User Guide. Sequoia Scientific, Inc.
- Murphy, J. M. (1988). The impact of ensemble forecasts on predictability. *Quarterly Journal of the Royal Meteorological Society*, 114:463 – 493.
- Murthy, C. R. and Miners, K. C. (1992). Mixing characteristics of the Niagara River plume in Lake Ontario. *Water Pollution Research Journal of Canada*, 24:143 – 162.
- Negenborn, R. (2003). Robot localization and kalman filters on finding your position in a noisy world. Master's thesis, Utrecht University.
- Phillips, D. W. and Irbe, J. G. (1978). Lake to land comparison of wind, temperature, and humidity on Lake Ontario during the International Field Year for the Great Lakes (IFYGL). Technical Report CLI-2-77, Atmospheric Environment Service Report, Downsview, Ontario.
- Reichle, R. H., Mclaughlin, D. B., and Entekhabi, D. (2002). Hydrological data assimilation with the ensemble kalman filter. *Monthly Weather Review*, pages 103 – 114.
- Resio, D. T. and Vincent, C. L. (1977). Estimation of winds over the Great Lakes. *Journal Waterway Port Coastal Ocean Division, ASCE*, 102:265 – 283.
- Rodgers, G. K. (1971). Field investigation of the thermal bar in Lake Ontario: precision temperature measurements. In *14th Conf. Great Lakes Res.*, pages 618 – 624.

- Schott, J. R., Barsi, J. A., Nordgren, B. L., Raqueno, N. G., and Alwis, D. (2001). Calibration of Landsat thermal data and application to water resource studies. *Remote Sensing of Environment*, 78:108 – 117.
- Sherwood, D. A. (2005). Water resources of Monroe County, New York, water years 2000-02 - atmospheric deposition, ground water, streamflow, trends in water quality, and chemical loads in streams. Technical Report 2005-5107, Monroe County Department of Health.
- Slater, A. G. and Clark, M. P. (2006). Snow data assimilation via an ensemble kalman filter. *Journal of Hydrometeorology*, 7:478 – 493.
- Smith, R. C. and Baker, K. S. (1981). Optical properties of the clearest natural waters (200-800 nm). *Applied Optics*, 20:277 – 184.
- Stramska, M. and Stramski, D. (2005). Effects of a nonuniform vertical profile of chlorophyll concentration on remote-sensing reflectance of the ocean. *Applied Optics*, 44:1735 – 1747.
- Szunyogh, I. and Toth, Z. (2002). The effect of increased horizontal resolution on the NCEP global ensemble mean forecasts. *Monthly Weather Review*, 130:1125 – 1143.
- Vermote, E. F. and Vermeulen, A. (1999). Atmosphere correction algorithm: Spectral reflectances (MOD09). Technical Report NASA Contract NAS5-96062, University of Maryland, Department of geography.
- Welch, G. and Bishop, G. (2006). An introduction to the Kalman Filter. Department of Computer Science, University of North Carolina at Chapel Hill.

- Whitaker, J. S. and Lough, A. F. (1998). The relationship between ensemble spread and ensemble mean skill. *Monthly Weather Review*, 126:3292 – 3302.
- Yamada, T. (1983). Simulations of nocturnal drainage flows by a q^2l turbulence closure model. *J. Atmos. Sci.*, 40:91 – 106.