

EVALUATING THE ROLE OF GROUNDWATER IN  
CIRCULATION, THERMAL STRUCTURE AND  
NUTRIENT-ALGAL DYNAMICS WITHIN  
A DEEP INLAND LAKE

By

Ammar Safaie Nematollahi

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements for  
the degree of

Civil Engineering - Doctor of Philosophy

2017

## **ABSTRACT**

### **EVALUATING THE ROLE OF GROUNDWATER IN CIRCULATION, THERMAL STRUCTURE AND NUTRIENT–ALGAL DYNAMICS WITHIN A DEEP INLAND LAKE**

By

Ammar Safaie Nematollahi

Inland lakes readily respond to changes in external forcing and therefore serve as sentinels of climate change. Many parts of the world continue to experience declining groundwater levels due to anthropogenic activities such as high-capacity pumping for agriculture or decreases in natural recharge rates of aquifers while lake surface temperatures continue to rise and show a clear warming trend. The responses of individual lakes to these stressors could vary depending upon the positions of the lakes within the landscape and the nature of lake-groundwater interactions. Since temperature is a key driver that affects the structure and function of ecosystems including biological productivity, nutrient cycling and hypoxia, groundwater-fed lakes could be altered drastically due to declining groundwater contribution. Thus, it is crucial to understand the role of groundwater in biophysical processes and to determine what regime shifts may occur in the absence of lake-groundwater interactions. To address this question, extensive field datasets were collected in the Gull Lake, a deep, dimictic, groundwater-fed, inland lake in Michigan, with bottom cooling and strong stratification during summer. The lake supports diverse warm and cold water fisheries. Detailed three-dimensional hydrodynamic and temperature models of Gull Lake coupled to nutrient and algal dynamics were developed to study the effect of groundwater on physical, chemical, and biological processes in the lake. Coupled biophysical processes in the water column are closely linked to meteorological forcing. Therefore, meteorological forcing fields were carefully reconstructed from a network of weather station data, and were assessed using outputs

from a mesoscale numerical weather forecasting model (WRF). A novel manifold method of reconstructing dynamically evolving spatial fields is presented for assimilating data from sensor networks in lake and watershed models. The manifold method has been developed based on the assumption that geophysical and meteorological data can be mapped onto an underlying differential manifold. A comparative evaluation of turbulence models was also conducted to improve descriptions of vertical mixing and thermal structure of the lake. The performance of the biophysical model was first evaluated against high-resolution *in situ* observations, including currents, lake levels, temperature, nutrients, dissolved oxygen, and chlorophyll data. After successfully applying the model to describe current conditions, the developed models were used to understand the responses of the lake ecosystem when the groundwater contribution is absent. Results suggest that groundwater-fed lakes have the ability to buffer seasonal water temperature variations in the hypolimnion, which helps them to withstand disturbances from surface-induced changes. However, groundwater depletion was accompanied by changes in the structure and function of lake ecosystems including lake level changes, rising water temperatures, increased growth rates of algae, oxygen depletion, early anoxia, reduction of light availability, and eutrophication. These results highlight the significant role played by groundwater in inland lakes and indicate that groundwater-dependent ecosystems tend to show greater resilience. In addition to providing insights into key biophysical processes in inland lakes, this study is expected to help strengthen management efforts to improve or maintain the resilience of lake ecosystems.

Suggested citation:

Safaie, A. (2017), “*Evaluating the Role of Groundwater in Circulation, Thermal Structure, and Nutrient-Algal Dynamics within a Deep Inland Lake*”, Ph.D. Dissertation, Michigan State University, East Lansing, Michigan, United States.

Copyright by  
AMMAR SAFAIE NEMATOLLAHI  
2017

*To my family and my love*

## ACKNOWLEDGEMENTS

This research was supported by a grant from the National Science Foundation, CyberSEES program (Award # 1331852). I would like to gratefully acknowledge my academic advisor, Prof. Mantha S. Phanikumar, for his enlightening guidance, enthusiastic encouragement, and continuous support throughout my doctoral program. I am thankful to my committee members, Prof. Farhad Jaber, Prof. Elena Litchman, and Dr. Yadu Pokhrel for their time, suggestions, and feedback on my doctoral dissertation. Fieldwork and data collection would have not been possible without the assistance of Prof. Litchman and her crew, Pam Woodruff, Allyson Hutchens, and Kirill Shchapov, and I truly appreciate them all. I also thank Prof. Stephen K. Hamilton for sharing the inflow and outflow data for Gull Lake. Additionally, many thanks go to Mike Gallagher, Jim Allen, Andrew Fogiel, and Michael Nichols for their assistance with deployments of our instruments in Gull Lake. I am grateful to Prof. Hayder Radha and Dr. Chinh Dang for their collaboration and for introducing me to the manifold-based approach. I extend my appreciation to Prof. Shu-Guang Li and his research team, Dr. Huasheng Liao and my good friend Zachary K. Curtis, for their helpful discussions and great collaboration. My sincere thanks also go to my colleagues Tuan D. Nguyen, Guoting Kang, and Han Qiu for their fruitful discussions and assistance with field data collection. Most importantly, I would like to express my heartfelt thanks and gratitude to my family for their continuous support, dedication, encouragement, and prayers in every step of my life. Above all, I would like to express my sincerest gratitude to God for everything I have been given.

# TABLE OF CONTENTS

LIST OF TABLES .....	ix
LIST OF FIGURES .....	x
CHAPTER 1 .....	1
1 Introduction.....	1
1.1 Problem description .....	1
1.2 The study area .....	4
1.3 Summary of specific aims.....	7
1.4 Expected benefits and significance .....	8
1.5 Dissertation structure .....	9
CHAPTER 2 .....	11
2 Reconstruction of Geophysical and Meteorological Data in Integrated Earth System Models .....	11
2.1 Introduction.....	11
2.2 Materials and methods .....	15
2.2.1 Application of manifold methods for data assimilation in integrated Earth system models .....	15
2.2.1.1 Manifold approach.....	15
2.2.1.2 Test case: Analytical function.....	25
2.2.1.3 Assimilating meteorological data for improved lake circulation modeling: Lake Michigan .....	26
2.2.1.4 Assimilating geophysical data for improved lake circulation modeling: Gull Lake .....	29
2.2.2 Reconstruction of meteorological data in Gull Lake.....	31
2.2.2.1 Land-based weather station network .....	31
2.2.2.2 Mesoscale weather prediction model.....	33
2.3 Results and Discussion .....	35
2.3.1 Manifold method .....	35
2.3.2 Observed and predicted meteorological data in Gull Lake .....	41
2.4 Conclusions.....	46
CHAPTER 3 .....	48
3 Evaluating the Role of Groundwater in Circulation and Thermal Structure .....	48
3.1 Introduction.....	48
3.2 Materials and methods .....	52
3.2.1 Observational data.....	52
3.2.2 Numerical hydrodynamic model .....	54
3.2.3 Turbulent closure models .....	61
3.2.3.1 Mellor-Yamada 2.5 level turbulence model .....	61
3.2.3.2 The $k - \varepsilon$ turbulence model.....	63
3.2.4 Light attenuation lengths .....	65

3.2.5 Precipitation/Evaporation .....	66
3.2.6 Groundwater .....	67
3.3 Results and discussion .....	77
3.4 Conclusions.....	99
 CHAPTER 4 .....	 102
4 Evaluating the Role of Groundwater in Nutrient - Algal Dynamics .....	102
4.1 Introduction.....	102
4.2 Materials and methods .....	105
4.2.1 Field sampling and laboratory analyses .....	105
4.2.2 Biophysical model.....	106
4.3 Results and discussion .....	117
4.4 Conclusions.....	126
 CHAPTER 5 .....	 128
5 Conclusions.....	128
 APPENDIX.....	 131
 BIBLIOGRAPHY.....	 134

## LIST OF TABLES

<b>Table 2-1</b> Properties of the numerical grids used for the hydrodynamic models. ....	28
<b>Table 2-2</b> Cross-validation results for the analytical function based on different sampling points selected randomly. ....	36
<b>Table 2-3</b> RMSE values (m/s) of alongshore and cross-shore velocities in Lake Michigan for comparison of the manifold method with other standard methods used in limnology and oceanography. ....	37
<b>Table 2-4</b> Cross-validation results for wind field over Lake Michigan. ....	38
<b>Table 2-5</b> RMSE values (m/s) of eastward and northward velocities in Gull Lake for comparison of the manifold method with other standard methods used in limnology and oceanography. ....	41
<b>Table 2-6</b> Cross-validation results for Gull Lake bathymetry.....	41
<b>Table 3-1</b> Details of instruments deployed in Gull Lake. ....	53
<b>Table 3-2.</b> Description of different models evaluated.....	60
<b>Table 3-3.</b> Averaged heat budget components of Gull Lake in the summers of 2014 and 2015. ....	80
<b>Table 3-4.</b> RMSE values (m/s) of eastward and northward velocities at ADCP locations.....	81
<b>Table 3-5.</b> RMSE values (m/s) of the observed and simulated vertical velocity profiles at ADCP locations for some selected time steps. ....	86
<b>Table 3-6.</b> Set of parameters used for the groundwater module. ....	88
<b>Table 4-1.</b> Nutrient concentration ranges in Prairieville Creek and a pond lab reservoir well near Gull Lake from 2009-2014. ....	106
<b>Table 4-2</b> Parameter values used in the water quality model.....	116

## LIST OF FIGURES

<b>Figure 1-1</b> Map of Gull Lake .....	5
<b>Figure 1-2</b> Gull Lake's watershed boundaries and digital elevation model (DEM) map.....	6
<b>Figure 2-1</b> Some examples of manifolds (a) torus, (b) sphere, and (c) a two-dimensional cross-section of a six-dimensional Calabi-Yau manifold.....	16
<b>Figure 2-2</b> Illustration of the proposed manifold approach for estimation of missing data at point $P_0$ .....	17
<b>Figure 2-3</b> A tangent space created from the set of nearest points using (a) coordinates of selected neighborhoods or (b) Kernel regression. ....	19
<b>Figure 2-4</b> Manifolds representing (a) bathymetry of Gull Lake and (b) wind components over Lake Michigan in three dimensional space.....	24
<b>Figure 2-5</b> (a) Analytical function used to test the manifold method for interpolation of scattered data. Random sampling was used to generate scatter points as shown in figures (b, 30 points), (c, 60 points) and (d, 90 points) to reconstruct the function. ....	25
<b>Figure 2-6</b> Locations of the ADCPs deployed during summer 2008 and weather stations surrounding Lake Michigan.....	27
<b>Figure 2-7</b> (a) Bathymetry of Gull Lake. (b) Boat tracks generated during the sampling survey in Gull Lake. ....	30
<b>Figure 2-8</b> Selected weather stations surrounding Gull Lake.....	32
<b>Figure 2-9</b> Circular grid with a radius of 42 km used for the reconstruction of meteorological fields from the Land-based weather station network.....	33
<b>Figure 2-10</b> Three nested simulation domains with resolutions of 30 km, 10 km, and 2 km used for the WRF model. ....	34
<b>Figure 2-11</b> Performance of the manifold method evaluated using observed and simulated currents at different stations in Lake Michigan. Different number of nearest neighbors were used to reconstruct the wind field using the manifold method with kernel regression. ....	38
<b>Figure 2-12</b> Comparison of simulated and observed vertically averaged currents at the location M in Lake Michigan. (a) Alongshore velocity (b) Cross-shore velocity .....	39
<b>Figure 2-13</b> Comparison of simulated (black lines) and observed (red lines) vertically averaged currents at the ADCP location in Gull Lake. (a) Eastward velocity and (b) Northward velocity 40	
<b>Figure 2-14</b> Comparisons of observed meteorological data at KBS LTER station with simulated results based on NCEP/NCAR data during April to mid-September of 2014 (a) air pressure (kPa),	

(b) air temperature ( $^{\circ}\text{C}$ ), (c) downward shortwave radiation ( $\text{W}/\text{m}^2$ ), (d) downward longwave radiation ( $\text{W}/\text{m}^2$ ), (e) relative humidity (%), (f) wind speed at 10-meter height (m/s), (g) eastward wind velocity (m/s), (h) northward wind velocity (m/s)..... 42

**Figure 2-15** Scatter plots of WRF-simulated results versus observed meteorological forcing at KBS LTER during April to mid-September of 2014 (black dot) and 2015 (red dot): (a) air pressure (kPa), (b) air temperature ( $^{\circ}\text{C}$ ), (c) downward shortwave radiation ( $\text{W}/\text{m}^2$ ), (d) downward longwave radiation ( $\text{W}/\text{m}^2$ ), (e) relative humidity(%), (f) wind speed at 10-meter height (m/s), (g) eastward wind velocity (m/s), (h) northward wind velocity (m/s)..... 43

**Figure 2-16** Comparisons of averaged simulated forcing (WRF) and observed meteorological forcing (Stations) over Gull Lake during April to mid-September of 2014. .... 44

**Figure 2-17** Comparisons of averaged simulated forcing and meteorological forcing over Gull Lake during April to mid-September of 2015..... 45

**Figure 2-18** Comparisons of the simulated and observed wind speed in 2014 and 2015 at the KBS LTER station. .... 45

**Figure 3-1** Map of Gull Lake showing locations of the *in situ* measurements. .... 53

**Figure 3-2** Schematic of the thermistor chain..... 54

**Figure 3-3.** The unstructured computational mesh created based on the bathymetry-based refinement algorithm..... 59

**Figure 3-4.** Quaternary geology of Lower Peninsula and Gull Lake. Figure adapted by author from *Farrand* [1982] ..... 68

**Figure 3-5.** Bedrock Geology of the Lower Peninsula and the Gull Lake area. Figure adapted by author from *Wilson* [1987]..... 70

**Figure 3-6.** (a) Groundwater flow map in the study area. (b) A cross-section of the aquifer along the longitudinal transect of the lake..... 71

**Figure 3-7.** Time series of relatively constant bottom water temperatures (blue symbols) measured at a depth of 32 m between 2011 and 2015 compared to the surface water temperature and air temperature for the period 2011-2015 (black line). .... 76

**Figure 3-8.** Water budget components in Gull Lake in (a) 2014 and (b) 2015..... 77

**Figure 3-9.** Comparisons of water level fluctuations recorded by ADCPs (red line), simulated without groundwater (blue line), and simulated with groundwater (black and yellow lines) in (a) 2014 and (b) 2015. .... 78

**Figure 3-10.** Comparison of Gull Lake evaporation in 2014 computed by the mass transfer approach (method1) and the COARE algorithm (method2)..... 79

<b>Figure 3-11.</b> Comparisons of simulated water level fluctuations in Gull Lake using evaporation computed by the mass transfer approach (method1) and the COARE algorithm (method2). .....	79
<b>Figure 3-12.</b> Comparison of simulated and observed vertically averaged salinity at HL location for year 2014. ....	80
<b>Figure 3-13.</b> Comparisons of observed and simulated vertically-averaged velocity profiles at ADCP locations in the summer of 2014. ....	82
<b>Figure 3-14.</b> Scatter plots of velocities (observed versus simulated for different models) for year 2014. ....	83
<b>Figure 3-15.</b> Probability distributions of RMSE values based on observed and simulated vertical velocity profiles .....	84
<b>Figure 3-16.</b> Examples of comparisons of the observed and simulated vertical velocity profiles at ADCP locations in the summer of 2014. ....	85
<b>Figure 3-17.</b> Observed PAR data used to determine the parameters of the shortwave penetration model. ....	87
<b>Figure 3-18.</b> Contour plots of the observed and simulated water temperature at the TC location in the summer of 2014. ....	88
<b>Figure 3-19.</b> Comparisons of simulated and observed time series of (a) surface water temperature ( $T_0$ ), and bottom temperature at (b) TC, (c) M14, and (d) S14 locations in 2014. ....	89
<b>Figure 3-20.</b> Taylor diagram which shows statistical comparisons of observed water temperature with simulated results of model 1 to 4. The statistics are normalized by the standard deviation of the observations .....	90
<b>Figure 3-21.</b> Comparisons of simulated vertical profiles of temperature with weekly Hydrolab data in 2014. ....	91
<b>Figure 3-22.</b> Comparisons of the simulated vertically averaged velocities and observations at ADCP locations in the summer of 2015. ....	92
<b>Figure 3-23.</b> Time series of wind, current and significant wave height at SS15 location. (a) Eastward wind speed (m/s), (b) comparison of observed and simulated eastward current speeds (m/s), (c) northward wind speed (m/s), (d) comparison of observed and simulated northward current speeds (m/s), (e) observed significant wave height (m). Gray areas represent examples of periods with significant wave activity during which the model could not adequately capture the observed currents. ....	93
<b>Figure 3-24.</b> Comparisons between observed (red lines) and simulated (black lines) power spectral densities at ADCP locations in 2014 and 2015. ....	94
<b>Figure 3-25.</b> Examples of comparisons of the observed and simulated vertical velocity profiles at ADCP locations in the summer of 2015. ....	96

**Figure 3-26.** Contour plots of the observed and simulated water temperature at the thermistor chain location (TC) in the summer of 2015. .... 97

**Figure 3-27.** Comparisons of simulated and observed time series of (a) surface water temperature at TC, and bottom temperature at (b) TC, (c) SS15, and (d) SD15 locations in 2015. .... 98

**Figure 3-28.** Comparisons of simulated vertical profiles of temperature with weekly Hydrolab data in 2015. .... 99

**Figure 4-1** Comparisons of simulated vertical profiles of dissolved oxygen (mg O<sub>2</sub>/L) with Hydrolab data in 2015. Simulated results of Model 1 (blue line) are in the absence of groundwater effects, and results of Model 2 (black line) are with groundwater effects. .... 118

**Figure 4-2.** Contour plots of (a) observed DO (mg O<sub>2</sub>/L), (b) simulated DO with groundwater effects (Model 2) and (c) simulated DO in the absence of groundwater effects (Model 1) in Gull Lake in 2015. .... 119

**Figure 4-3.** Comparisons of simulated vertical profiles of Chl-a (µg C/L) with Hydrolab data in 2015. Simulated results of Model 1 (blue line) are in the absence of groundwater effects, and results of Model 2 (black line) are with groundwater effects. .... 120

**Figure 4-4.** Contour plots of (a) observed Chl-a (µg C/L), (b) simulated Chl-a with groundwater effects (Model 2) and (c) simulated Chl-a in the absence of groundwater effects (Model 1) in Gull Lake in 2015. .... 121

**Figure 4-5.** Contour plots of (a) observed DO saturation (%), (b) simulated DO saturation with groundwater effects and (c) simulated DO saturations in the absence of groundwater effects in Gull Lake in 2015. .... 122

**Figure 4-6.** Comparisons of observed and simulated vertical profiles of dissolved nitrate (mg N/L) in 2015. Simulated results of Model 1 (blue line) are in the absence of groundwater effects, and results of Model 2 (black line) are with groundwater effects. .... 123

**Figure 4-7.** Comparisons of observed and simulated vertical profiles of SRP (µg P/L) in 2015. Simulated results of Model 1 (blue line) are in the absence of groundwater effects, and results of Model 2 (black line) are with groundwater effects. .... 124

# CHAPTER 1

## 1 Introduction

### 1.1 Problem description

Climate-induced changes in hydrodynamics of inland lakes are known to have physical, chemical, and biological effects on lake ecosystems. Wind speed, current shear, stratification and vertical mixing have significant effects on the cycling of nutrients, phytoplankton distribution, as well as on the growth rates of phytoplankton and benthic filter-feeders important for benthic-pelagic trophic coupling [Denman and Gargett, 1983; Edwards *et al.*, 2005; Rowe *et al.*, 2015]. In addition, wind-mixing and thermal stratification have implications for regulating the supply of phytoplankton to zebra mussels [Boegman *et al.*, 2008]. Changes in water temperature, thermal and mixing dynamics, strength and duration of stratification have impacts on primary production and associated food web, predation and growth rates of zooplankton and fish, nutrient supply, and deepwater anoxia [MacKay *et al.*, 2009]. Hence, understanding coupled aquatic biophysical processes in lakes is critical to an improved understanding of lake ecosystems and their responses to climate change.

Lake surface temperature is a key response indicator of climate change. Long-term air temperature and ice cover records suggest that freshwater ecosystems have experienced

temperature warming and shortening ice duration in response to the global warming over the past 150 years [Magnuson, 2000]. Remote sensing observations of inland water bodies since 1985 also provide evidence of significant warming trends in their surface temperatures to the extent that in some regions, such as the Great Lakes and Northern Europe, water temperature is rising more rapidly than regional air temperature [Schneider and Hook, 2010]. The lake warming has been reported even in the largest freshwater lakes around the world such as Lake Baikal [Hampton et al., 2008], Lake Superior [Austin, 2013], Lake Michigan [Brooks and Zastrow, 2002], Lake Erie [Burns et al., 2005], and Lake Tanganyika [O'Reilly et al., 2003]. These warming trends have resulted in longer summer stratifications, combined with higher summer water temperatures and a shorter winter-spring mixing period. Such changes in thermal stratifications have consequential impact upon the ecology of lakes. For instance, the changes could reduce primary productivity by limiting upwelling of nutrients from deep water [Brooks and Zastrow, 2002; O'Reilly et al., 2003; Chang et al., 2015], and change planktonic processes and community dynamics of other aquatic organisms further up in the food chain [Lane et al., 2008; MacKay et al., 2009; Rühland et al., 2015].

Surface fluxes are usually assumed to be the only important driving forces responsible for the aforementioned climate-induced changes. From a system dynamics point of view, if a system is composed of only a reinforcing loop, the closed loop of cause and effect would lead to instability via accelerating growth [Sterman, 2000]. Balancing loops, on the other hand, serve to resist attempted changes in order to maintain a balance and keep the system at a desired state. Identifying these reinforcing and balancing processes in a complex dynamic system like a freshwater ecosystem is crucial to better understand thermal structure, circulation, nutrients, and algal dynamics in inland lakes. It is known that primarily groundwater-fed rivers are buffered against

increasing seasonal temperature variation [Combes, 2003]. A shallow temperate lake fed by groundwater might also have a significant bottom heating in winter and bottom cooling in summer [Kettle *et al.*, 2012]. This groundwater effect may act as a balancing loop within a lake system. If this hypothesis is true, then groundwater can be introduced as a key contributor to offset warming and anoxia in the groundwater-fed lakes. However, long-term groundwater depletion has been reported in many different regions of the world [Rodell *et al.*, 2009; Wada *et al.*, 2010; Konikow, 2013; Joodaki *et al.*, 2014; Pokhrel *et al.*, 2015; Dalin *et al.*, 2017]. Therefore, groundwater dependent ecosystems are likely to be under dual pressure from surface temperature warming and groundwater depletion in response to natural climate variability and human activities. Thus, in order to protect ecosystems of groundwater-fed lakes, we need to understand the role of groundwater in their biophysical processes, and determine what regime shifts may occur in their ecosystems in the absence of lake-groundwater interactions.

There are a variety of numerical ocean models, such as POM [Blumberg, and Mellor, 1987] ROMS [Haidvogel and Beckmann, 1999], and FVCOM [Chen *et al.*, 2003b, 2006] that can be used to predict and obtain a better understanding of aquatic biophysical processes. Performance of these mechanistic models is highly dependent on accuracy of surface forcing fields (i.e. wind stress and heat flux). These forcing fields can be obtained from observational data, the output of a weather forecast model, or a combination thereof [Xue *et al.*, 2015]. In the Great Lakes region, real-time and historic meteorological data are available from buoys and monitoring stations across the lakes. However, realistic meteorological forcing fields (reconstructed using point observations) are still needed and the hydrodynamic results could be enhanced by using an improved algorithm for reconstruction of meteorological fields based on data from a network of weather stations [Safaie *et al.*, 2016] or utilizing a weather forecast model [Xue *et al.*, 2015]. In contrast, for

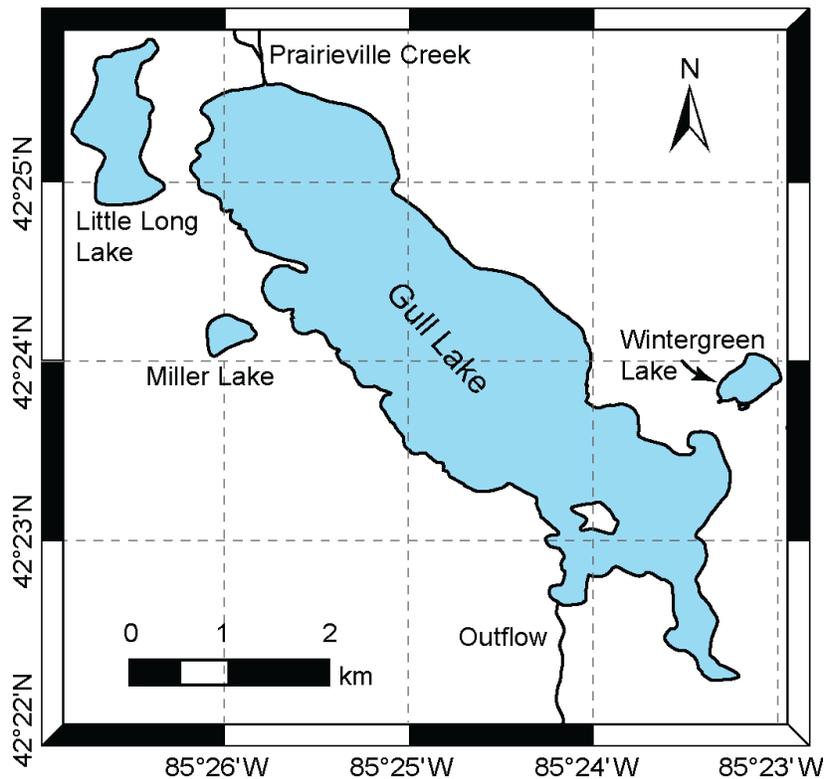
relatively small inland lakes nestled within the landscape, meteorological forcing data are not always available and intraregional data have to be compiled from local weather stations (e.g., from airports) several kilometers away from the lake [*Hondzo and Stefan, 1993; Markfort et al., 2010*]. Moreover, meteorological data from a single station near the lake may not be representative of the spatial distribution of meteorological fields around the lake and may produce unreasonable results [*Rueda et al., 2005*].

## 1.2 The study area

To evaluate the role of groundwater, we used field datasets and a coupled biophysical model of Gull Lake, a relatively small (8 km<sup>2</sup> surface area) but deep (34 m maximum depth) clear water lake in the in Kalamazoo County in southwestern Michigan. Gull Lake is a hardwater, groundwater-fed lake, representative of many other inland lakes in the Great Lakes region. This lake is an attractive recreational site that supports diverse warm and cold water fisheries. Over 70 species of fish with a seasonal pattern in their predator dynamics are reported in Gull Lake [*Lane, 1979*]. The presence of a Michigan State University research field station (the W.K. Kellogg Biological station), located on shores of Gull Lake, has resulted in long term biological and limnological research. In 1941, a bathymetric map of the lake was created to aid fisheries research. The earliest limnological study of Gull Lake by *Perry and Brown [1942]* reported the thermal stratification in the lake. The lake is usually stratified from May to early November. Gull Lake was of great scientific interest following the zebra mussel invasion in 1994, and the subsequent rapid increase of *Microcystis* biomass after the invasion [*Wilson and Sarnelle, 2002; Sarnelle et al., 2005; Bruesewitz et al., 2009; Horst et al., 2014; White et al., 2015*]. Gull Lake nitrification and denitrification rates in littoral sediments are relatively high likely due to the presence of zebra mussels [*Bruesewitz et al., 2009*]. Observations from 1965 to 1975 showed a high growth rate of

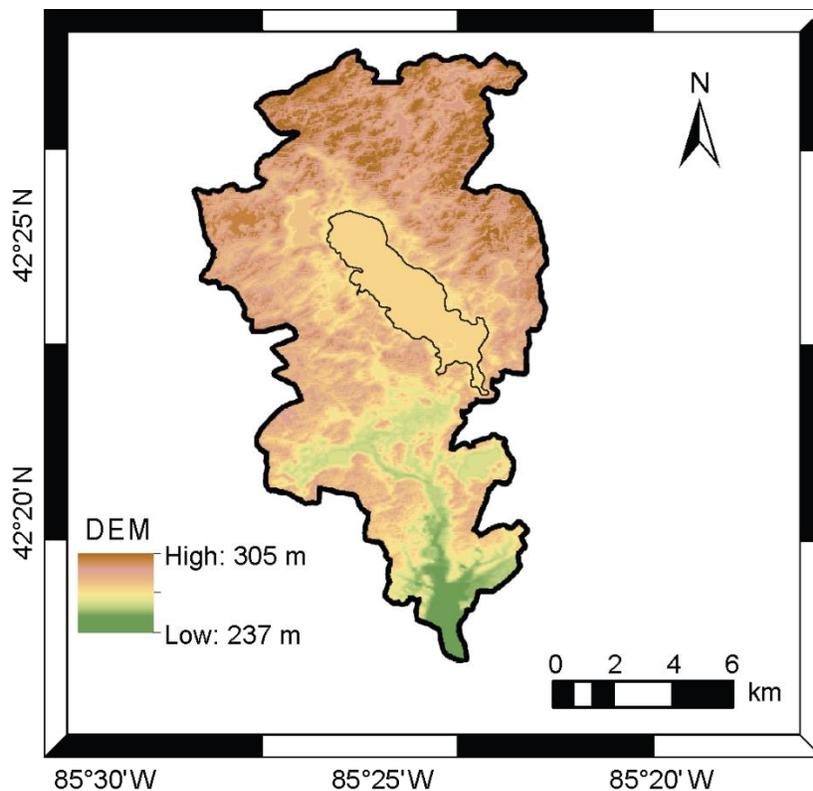
algae in Gull Lake which was controlled by phosphorus loading into the lake. Long term phosphorus concentration of Gull Lake has been stable since 1994, and currently the water quality in the lake is reported to be in good condition. In summers, however, some concerns have been raised about blue-green algal blooms [White and Hamilton, 2014].

Prairieville Creek provides the main surface inflow into Gull Lake from the north. This tributary had a small annual flow rate of 0.19 (m<sup>3</sup>/s) in 2014. The lake has a single outflow southward to the Gull Creek (Figure 1-1). The flow rate of the outflow and the lake level are controlled by a sluice-gate dam. Gull Lake also receives water from three small lakes: Little Long, Wintergreen, and Miller Lakes. All these lakes have glacial kettle type basins, and the soil in this area is classified as glacial outwash sand and gravel, and postglacial alluvium.



**Figure 1-1** Map of Gull Lake

Bedrock type divides Gull Lake into two bedrock aquifer systems. The northern half of the lake is underlain by a permeable Marshall formation, and the remaining part consists of Coldwater shale with a relatively low permeability. The topographic elevation around the lake ranges between 237 to 305 meters with a downward slope toward the south/southeast (Figure 1-2). It has been reported that, other than small tributaries, inflow of Gull Lake is provided by groundwater discharge through springs from the bottom of the lake [Perry and Brown, 1942; Moss, 1972]. Tague [1977] estimated the water budget of Gull Lake in 1974, with 40% of its water from groundwater inflow, 25% from direct precipitation onto the lake surface, and 35% from stream inflows.



**Figure 1-2** Gull Lake's watershed boundaries and digital elevation model (DEM) map.

During the summer stratification period, Gull Lake is alkaline with average pH values of 9 and 8.2 in the epilimnion and hypolimnion, respectively. The level of alkalinity reflects the

interaction of the lake with groundwater. *Kinsman-Costello et al.* [2015] reported that the groundwater contribution to the water budget of the lake was as high as 90% in late summer of 2010. This percentage of the groundwater influence was calculated from magnesium ion ( $Mg^{2+}$ ) concentrations in shallow (<2m deep) waters as a conservative tracer for groundwater, assuming that groundwater and precipitation are the only source of dissolved  $Mg^{2+}$  in the lake. In inland shallow lakes, heat flux exchange between water and lake sediment needs to be taken into consideration for vertical thermal diffusivity analyses [*Hondzo et al.*, 1991]. It was suggested that heat exchange between water and lake sediments in inland shallow lakes needs to be taken into consideration for vertical thermal diffusivity analyses [*Hondzo and Stefan*, 1993]. In most hydrodynamic models, however, the role of groundwater exchange in the energy budget and seasonal stratification are assumed to be negligible.

### **1.3 Summary of specific aims**

The major objective of this study was to evaluate the role of groundwater in physical, chemical, and biological processes in Gull Lake which is a typical representation of many other groundwater-fed inland lakes in Michigan. Despite decades of water quality monitoring and biological research in Gull Lake, a numerical model of the lake has not been developed to date. A major motivation of this work is to develop a hydrodynamic model of Gull Lake coupled to nutrient (N, P) and algal dynamics. The working hypothesis here is that groundwater flow through Gull Lake causes significant bottom cooling in summer, to an extent that it has an important role in offsetting hypolimnetic warming and anoxia in the lake. Groundwater also has an effect on temperature, and it is well-known that phytoplankton and algae growth rates are dependent on temperature. This work is an attempt to quantify contributions of groundwater in a lake ecosystem by addressing the following research questions:

1. What is the role of groundwater in circulation and thermal structure within Gull Lake?

Aim 1: Evaluate the role of groundwater in circulation and thermal structure.

Aim 2: Quantify the impact of ignoring groundwater contribution on temperature.

Aim 3: Quantify the percentage contribution of groundwater to the water budget of the lake.

2. What is the role of groundwater in water quality of Gull Lake?

Aim 1: Understand the effect of groundwater on dissolved oxygen.

Aim 2: Understand the effect of groundwater on nutrient-algal dynamics.

Aim 3: Understand the impact of ignoring groundwater contribution on nutrients and algae.

## **1.4 Expected benefits and significance**

The aim of this research is to combine a coupled biophysical model and field observations to provide improved understanding of physical and biological behaviors of lakes. This research highlights the importance of algal blooms in the context of climate change and how groundwater tends to slow down the effects of the climate change. Insights gained from the developed complex mechanistic model will advance the understanding of the response of dissolved oxygen in inland lakes to the warming trend of air temperature throughout the world. Another aspect of this study is to document a systematic approach to modeling hydrodynamics and water quality of inland lakes that can be used for other fresh water bodies. These outcomes could strengthen management efforts to decrease nutrient loads, control hypolimnetic oxygenation, and monitor human activities that affect groundwater – surface water interactions.

## 1.5 Dissertation structure

A coupled three-dimensional hydrodynamic and water quality model of Gull Lake was developed to investigate the effect of groundwater on physical, chemical, and biological processes in the lake. The model was refined and assessed through the following steps:

1. Collecting accurate bathymetry of the lake in order to improve the performance of the numerical model.
2. Developing a novel manifold-based method to assimilate geophysical and meteorological data in integrated Earth system models for better representations of the meteorological fields and bathymetry.
3. Evaluating the performance of the model for different forcings including a mesoscale weather forecast model.
4. Refinement of the unstructured numerical mesh horizontally and vertically using a bathymetry-based refinement algorithm to improve the accuracy of simulated circulation and thermal structure.
5. A comparative study of turbulence models to identify superior formulations and to further improve descriptions of vertical mixing and thermal structure of the lake.
6. Utilizing high-resolution *in situ* observations to evaluate the performance of the models in describing coupled biophysical processes.

The developed models were then used to predict the responses of the lake ecosystem caused by disconnection of the lake from groundwater. This doctoral dissertation is divided into the following five main chapters:

### **Chapter 1.** Introduction

**Chapter 2.** Reconstruction of geophysical and meteorological data in integrated Earth system models

**Chapter 3.** Evaluating the role of groundwater in circulation and thermal structure

**Chapter 4.** Evaluating the role of groundwater in nutrients, and algal dynamics

**Chapter 5.** Conclusions

## CHAPTER 2

# 2 Reconstruction of Geophysical and Meteorological Data in Integrated Earth System Models

### 2.1 Introduction

Performance of integrated Earth system models relies upon accuracy of meteorological and geophysical data. *In situ* observations generally have sparse and inhomogeneous distribution in space and time, and it is often infeasible to accurately reconstruct the true field from the data. However, more information about the structure of the field and its evolution, allows for better approximations [Barth *et al.*, 2008]. For instance, currents in large lakes such as the Laurentian Great Lakes are mostly controlled by wind. Therefore, by improving the representation of wind fields in models of lake circulation, we expect to describe coupled biophysical processes in lakes more accurately. For example, Safaie *et al.* [2016] demonstrated that improved representation of meteorological fields based on natural neighbor interpolation of weather station data produced superior results for currents and bacterial concentrations relative to similar results based on a nearest neighbor interpolation of the same data. Accurate representation of geophysical features such as topography and bathymetry is also important in Earth system models and their components,

and model performance depends on the interpolation method used to assign the topographic information over a numerical mesh in processed-based models.

In order to assimilate observations into the models, and estimate variables at unsampled locations and/or times, it is crucial to use a suitable interpolation method. Various deterministic [e.g., nearest neighbor, natural neighbor, inverse distance weighting (IDW), spline, polynomial] and geostatistical (e.g. kriging) interpolation methods have been developed to generate spatial fields. There have been numerous efforts to compare different spatial interpolation methods in order to identify the best method for a given model application. *Yan et al.* [2014] compared different interpolation methods, including IDW, global polynomial interpolation, local polynomial interpolation, radial basis functions, ordinary kriging (OK), simple kriging (SK), universal kriging (UK), and co-kriging (CK) to determine the water/land boundary point elevation based on *in situ* water level data from 14 control stations in Dongting Lake. They used a cross-validation method to select the optimal method, which was found to be the OK method. *Merwade* [2009] studied the effect of spatial trend on interpolation of river bathymetry, and compared the performance of different interpolation methods. The number of measurements and their spatial arrangement, as well as channel morphology and geology were found to influence the accuracy of the interpolation results [*Merwade*, 2009]. Due to the effects of these and other factors on the performance of various methods, comparisons of different spatial interpolation methods could not point out the best universal interpolation method [*Li and Heap*, 2008; *Šiljeg et al.*, 2015].

Many studies, such as aforementioned works, have used cross-validation for assessing the performance of the interpolation methods. In this method, a subset of the original dataset is withheld to be used later for validating the interpolated field constructed from the rest of the observational data. Mean error (ME), root mean square error (RMSE) and the coefficient of

determination ( $R^2$ ) are commonly used to evaluate the performance of each interpolation method [Suparta and Rahman, 2016]. However, every problem has a unique method of interpolation that works best for a given distribution of observations and the intended use of the interpolated data. Density of a sensor network, spatial variability of the variable of interest and its distribution, and observational errors, all influence the accuracy of the interpolated field [MacEachren and Davidson, 1987]. For example, Luo et al. [2008] compared seven spatial interpolation techniques to identify which method produced the best estimation of the wind speed data recorded across England and Wales. Their study showed that kriging is the best method, and that the thin plate spline method had higher ME and RMSE values. However, in Suparta and Rahman [2016] the performance of the thin plate spline interpolation based on the RMSE and  $R^2$  values was found to be better than kriging for less dense data points over the selected interpolation surface. Therefore, comparing interpolation methods using the cross-validation method without considering the data structure and the purpose of interpolation is not guaranteed to produce the best representation of the underlying data.

In this study, a novel manifold method is proposed to assimilate different types of spatiotemporal data in integrated Earth system models based on the hypothesis that an environmental dataset (including independent variables such as longitude, latitude, and time, and the measured variables of interest) can be mapped onto an underlying differential manifold. Working directly in the high dimensional space generally involves dealing with complex algorithms. Modeling the high dimensional data using manifolds with fewer degrees of freedom has captured a great deal of attention recently [Zhang et al., 2016a]. The use of low-dimensional manifolds not only reduces computational load for further processing, but also helps visualize the

entire dataset, which is an important initial step to make sense of the data before proceeding with more goal-directed modeling and analyses [Ma *et al.*, 2011].

The problem of non-linear dimensionality reduction for a set of high dimensional data points is known as manifold learning. Examples of early works for non-linear dimensionality reduction include Isomap [Tenenbaum *et al.*, 2000], local linear embedding (LLE) [Roweis and Saul, 2000], and Eigenmaps [Belkin and Niyogi, 2003], which have been used to learn the true underlying low-dimensional manifold structure of the data. Since then, the manifold model has been exploited extensively in numerous applications such as face recognition, action classification, segmentation, image denoising, image/video super-resolution, and multi-scale image analysis [Carin *et al.*, 2011; Dang *et al.*, 2013; Dang and Radha, 2014]. Most of these manifold learning methods have been inspired by linear techniques, mainly based on the assumption that non-linear manifolds can be approximated by locally linear parts [Mordohai and Medioni, 2010]. Two pioneering works in this area are the Isomap approach [Tenenbaum *et al.*, 2000] and the LLE algorithm [Roweis and Saul, 2000]. The Isomap algorithm aims to preserve the geodesic distance among points from the input dataset. On the other hand, the LLE algorithm targets the local linear geometry of neighbors in a manifold. Numerous works on manifold learning have been further developed. A comprehensive review of prior works can be found in van der Maaten *et al.* [2009].

In this chapter, the effectiveness of the presented manifold algorithm is evaluated through assimilation of geophysical and meteorological data in lake models (section 2.1), although the methods described are general and can be used in many other areas of computational geosciences. We first apply the proposed method to reconstruct wind fields (time-varying vector fields) over Lake Michigan. Since currents in Lake Michigan are primarily driven by wind, we expect to improve the simulation of hydrodynamic and biophysical variables of interest by improving the

wind fields. Instead of relying on the cross-validation of interpolated wind data, however, we use a well-tested hydrodynamic model of Lake Michigan and compare current measurements with simulated currents to test the interpolation methods. The manifold methods are used to reconstruct meteorological data, which are available from buoys and monitoring stations across the lakes, to improved simulation of circulation in Lake Michigan. Then the method is applied to assimilate bathymetry data as a scalar field for use in a hydrodynamic model of Gull Lake.

For small inland lakes nestled within the landscape, a nearby network of meteorological stations are not always available. In addition, meteorological data from a single station near the lake or from local weather stations several kilometers away from the lake may not be representative of the spatial distribution of meteorological fields around the lake. Therefore, we used a mesoscale weather prediction model along with data from a network of land-based weather stations to assess the accuracy of reconstructed forcing over Gull Lake (section 2.2). The predicted and observed weather data will be used in Chapter 3 to run the hydrodynamic model of Gull Lake.

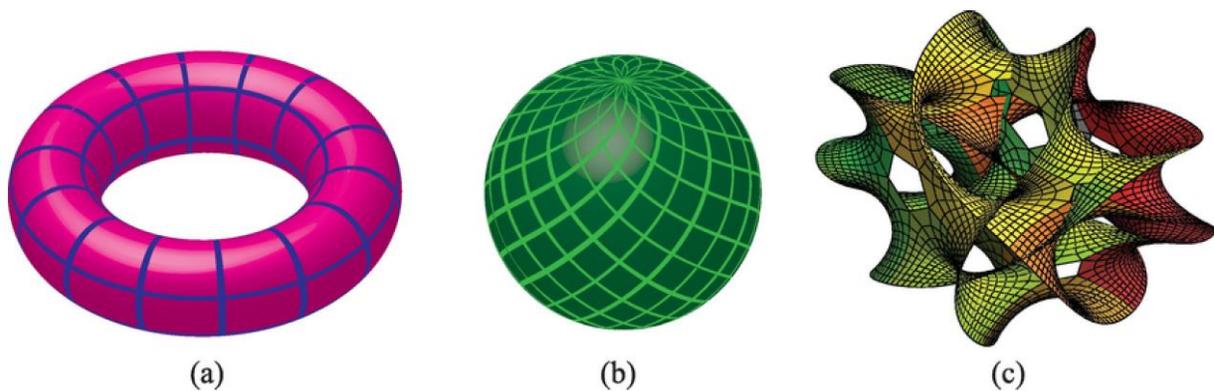
## **2.2 Materials and methods**

### **2.2.1 Application of manifold methods for data assimilation in integrated Earth system models**

#### **2.2.1.1 Manifold approach**

A manifold ( $M$ ) is an  $n$ -dimensional topological space such that each point of  $M$  and its neighborhood can be approximated by a small flat piece in the Euclidean space,  $\mathbb{R}^n$ . We can think of a manifold as a set of low-dimensional curves and surfaces within higher dimension Euclidean spaces [Guillemin and Pollack, 2010]. Some typical examples of manifolds are smooth surfaces, such as a torus (Figure 2-1a) or a sphere (Figure 2-1b), where each point and its neighborhood can

be approximated by a small flat linear-subspace within the three-dimensional Euclidean space. Another example of a manifold in a high dimensional space is a Calabi-Yau manifold which has found important applications in theoretical physics (e.g. superstring theory). Figure 2-1c shows a two-dimensional cross-section of a six-dimensional Calabi-Yau manifold. Surfaces of all these three manifolds are not a Euclidean space. The laws of the Euclidean geometry, however, are valid locally.

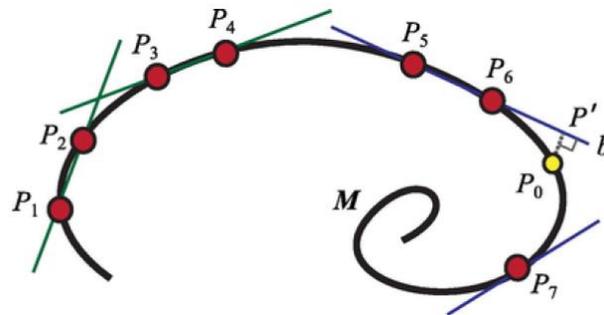


**Figure 2-1** Some examples of manifolds (a) torus, (b) sphere, and (c) a two-dimensional cross-section of a six-dimensional Calabi-Yau manifold.

Based on Einstein's theory of relativity, physical events are located on the continuum (manifold) of space-time. Therefore, station locations and times of observations form a space-time manifold viewed as a four-dimensional vector space. One way to handle spatiotemporal interpolation problems, inspired by this concept, is to integrate space and time simultaneously [*Li and Revesz, 2004*]. An underlying assumption behind this approach is that time and space dimensions can be treated as equally important [*Li et al., 2014a*]. In order to add time as another dimension of space, time values are needed to be scaled for a spatiotemporal dataset by a scaling speed [*Schwab and Beletsky, 1998; Li et al., 2014a*]. For a point measurement, we can then define a four-vector  $P^\mu = (ct, \vec{x})$  where  $c$  is a time scale,  $t$  is the time coordinate and  $\vec{x}$  is a three-

dimensional vector space. We assume that the set of high-dimensional data points  $P$  (and the estimated data points  $P_0$ ) belongs to a differential manifold  $M$ , which may be curved and have a complicated topology, but the neighborhood of each point is approximately similar to a small piece of Euclidean space (resembles  $\mathbb{R}^D$ ). Since a traditional distance measure is built upon the geometry of Euclidean space, we adapt the calculation to a neighborhood or a small region of the assumed manifold.

An example of a one-dimensional curve in Figure 2-2 illustrates the general idea of the manifold estimation approach. The set of points  $P$  in Figure 2-2 includes sample data points where we have measured data as well as a point  $P_0$  where data are missing. For example, in the context of the wind field data, one full measurement (or data point) includes five components: time, longitude, latitude, wind speed, and wind direction. The partially missing data point may contain known components (time, longitude, latitude) and unknown or missing components (wind speed and wind direction).



**Figure 2-2** Illustration of the proposed manifold approach for estimation of missing data at point  $P_0$ .

Suppose that it is desired to estimate the wind field for a data point  $P_0 \in \mathbb{R}^n$  ( $n =$  dimension of vector space) from a set of training data points that belong to a manifold  $M$ . The space/time coordinates of the point (the independent variables) are known. However, the data (the dependent

variable) are missing. We denote  $P_0 = \begin{bmatrix} P_0^\mu \\ P_0^\nu \end{bmatrix} \in \mathbb{R}^n$  as the data point using the superscript  $\mu$  to

denote the independent variables and the superscript  $\nu$  to denote the dependent variable which is

the missing component of interest here.  $P_0^\mu \in \mathbb{R}^{n_\mu}$  is the sub-vector of the known components, and

$P_0^\nu \in \mathbb{R}^{n_\nu}$  ( $n_\mu + n_\nu = n$ ) is the corresponding sub-vector field (e.g., wind vector) for the missing

component where  $P_0^\nu = \vec{V} = (u, v)$  and  $u$  and  $v$  are the orthogonal components of the wind ( $\vec{V}$ ).

The training data points, for example  $P = \{P_1, P_2, \dots, P_7\}$  in Figure 2-2, also include the two

components  $P_i = \begin{bmatrix} P_i^\mu \\ P_i^\nu \end{bmatrix} \in \mathbb{R}^n$ , but there is no missing component here since both dependent and

independent variables are assumed to be known at the nearby stations. Given a point  $P_0^\mu \in \mathbb{R}^{n_\mu}$ , the

algorithm locates a set of nearest points to  $P_0^\mu$  based on the distances  $d(P_i^\mu, P_0^\mu)$  between pairs of

points  $P_i^\mu$  and  $P_0^\mu$ . In order to determine local neighbors of  $P_0^\mu$ , we can calculate the distances

between  $P_0^\mu$  and either all other points within a fixed radius  $\varepsilon$ , or all of its  $k$  nearest neighbors

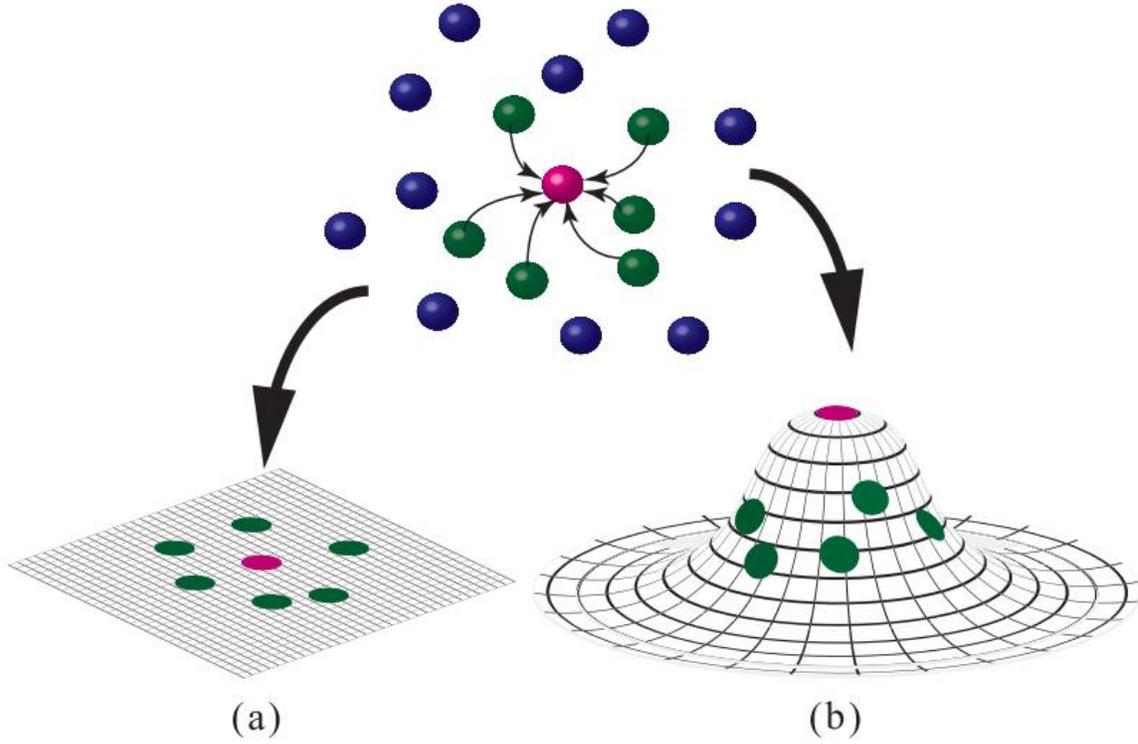
[Tenenbaum et al., 2000]. Then, a tangent space (linear subspace) of the manifold  $M$  at the point

$P_0$  is created from the set of nearest points (Figure 2-3a), denoted by  $T_{P_0}(M) = \begin{bmatrix} T^\mu \\ T^\nu \end{bmatrix}$  where  $T^\mu, T^\nu$

denote the tangent spaces for the independent and dependent variables in the data at  $P_0^\mu$  and  $P_0^\nu$ .

Finally, the point  $P_0 = \begin{bmatrix} P_0^\mu \\ P_0^\nu \end{bmatrix} \in \mathbb{R}^n$  will be located as the closest point that belongs to that tangent

space.



**Figure 2-3** A tangent space created from the set of nearest points using (a) coordinates of selected neighborhoods or (b) Kernel regression.

The tangent space approximation is describe as follow. Consider a smooth  $n$ -dimensional manifold  $M$  embedded in a  $D$ -dimensional Euclidean space. To understand the local geometry of the surface  $f(x)$  near a point  $x \in \mathbb{R}^n$ , we consider the first-order Taylor series expansion of the surface:

$$f(\bar{x}) = f(x) + \frac{\partial f(x)}{\partial x} (\bar{x} - x) + O(\|\bar{x} - x\|^2) = f(x) + J_f (\bar{x} - x) + O(\|\bar{x} - x\|^2) \quad (1)$$

where  $J_f(x) \in \mathbb{R}^{D \times n}$  is the Jacobian matrix of  $f$  at the point  $x$ . If the components of  $f(x)$  can be defined as:  $f(x) = [f_1(x), f_2(x), f_3(x) \cdots f_D(x)]^T$  and  $x = [x_1, x_2, x_3 \cdots x_n]^T$ , then the Jacobian can be written as:

$$J_f(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_D}{\partial x_1} & \dots & \frac{\partial f_D}{\partial x_n} \end{bmatrix} \quad (2)$$

To understand the local shape of the surface in Eq. (1), we seek to determine the space  $(\bar{x} - x)$ , such that as we move away from  $x$ , the value of the function does not change to within first order. This is equivalent to finding the space  $T$  such that:

$$T = \{(\bar{x} - x) \mid J_f(x)(\bar{x} - x) = 0\} \quad (3)$$

This space is the tangent space to the surface at point  $x$  and is the right null space of the Jacobian matrix  $J_f(x)$ . The space orthogonal to the tangent space is the row space of the Jacobian and orthogonal representations of these spaces can be obtained from SVD:

$$J_f = \begin{bmatrix} U_{\parallel} & U_{\perp} \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_{\parallel}^T \\ V_{\perp}^T \end{bmatrix} \quad (4)$$

where  $V_{\perp}$  spans the tangent space (right null space) and  $V_{\parallel}$  spans the row space. Alternatively, in Eq. (4), the right null space of  $J_f$  is the columns of  $V$  corresponding to zero singular values.

Therefore, the tangent space of the manifold  $M$  at  $y = f(x)$  is:

$$T(M) = \text{span}(J_f(x)) \quad (5)$$

From a practical computation point of view, given a set of sample points  $y = \{y_1, y_2, y_3, \dots, y_m\}$ , the tangent space can be directly estimated using SVD. If  $C^m$  denotes the local covariance matrix:

$$C^m = \frac{1}{m} \sum_{i=1}^m y_i y_i^T = U \Sigma U^T \quad (6)$$

where  $U = [u_1, u_2, u_3, \dots, u_D]$  and  $\Sigma = \text{diag}[\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_D]$  denote the eigenvector and eigenvalue matrices respectively, then the optimal (in a least-squares sense)  $n$ -dimensional linear subspace is the span of the  $n$ -largest eigenvectors in  $U$ :

$$T(M) \cong \text{span}\{u_1, u_2, u_3, \dots, u_n\} \quad (7)$$

Additional details and other methods of estimating the tangent space are described in *Dang et al.* [2014]. Now by having the tangent space, we can use the Euclidean distance of an orthogonal projection from a point to the tangent space to represent the closest distance between that point and the tangent space. Since a tangent space is a linear space (or affine space in a more general case), one point can orthogonally project into that space. The question is how to define neighbors for each data point? The underlying idea is how to define similarity distance among the training data points, and then the overall similarity matrix. Several methods have been considered in the past, such as k-nearest neighbors [*Press*, 2007],  $\epsilon$ -ball method [*Allard et al.*, 2012] or the use of sparse representation theory [*Dang et al.*, 2014; *Dang and Radha*, 2015].

The estimation of  $P_0^\nu$  is performed using the following steps:

1. Given a set of neighboring points, estimate the tangent space  $T^\mu$  at the point of interest,  $P_0$  :

Details of the method for creating a tangent space from a set of data points were described above. One simple method is to create a tangent space using singular value decomposition (SVD, *Press* [2007]). By way of an example in Figure 2-2, a tangent space (line  $b$ ) is created for  $P_0^\mu$  from a set of its neighboring points ( $P_5^\mu$  and  $P_6^\mu$ ). This tangent space at  $P_0^\mu \in M$  is denoted by  $T^\mu$ .

2. Find the orthogonal projection of  $P_0^\mu$  onto the tangent space:

The closest point  $P' \in T^\mu$  to the given point  $P_0^\mu$  is located at the intersection of the line  $b$  and the line perpendicular to it which passes through the point  $P_0^\mu$ .  $P'$  which is a projection of  $P_0^\mu$  onto the subspace  $T^\mu$  can be represented as an approximation of point  $P_0^\mu$ . The orthogonal projection of vector point  $P_0$  in a high-dimensional space onto a low-dimensional vector subspace is given by:

$$\prod_{T^\mu}(P_0^\mu) = A (A^T A)^{-1} A^T P_0^\mu = AA^+ P_0^\mu \quad (8)$$

where  $A = T^\mu \in \mathbb{R}^{D \times n}$  is a full rank matrix containing the set of points on the tangent space of  $P_0^\mu$  and  $\prod_{T^\mu}(P_0^\mu)$  denotes the projection of  $P_0^\mu$  onto the subspace  $T^\mu$ . This projection is derived from the solution of the normal equation  $A^T A x = A^T P_0^\mu$  which is equivalent to the associated least squares solution of  $A x = P_0^\mu$ . Due to the difficulty associated with inverting a general matrix that may be singular or non-square depending on the number of neighboring points selected in the manifold method, the problem (1) can be posed as a minimization problem in which the Moore-Penrose pseudoinverse  $A^+$  [Golub and Van Loan, 2013] of the original matrix  $A$  is used. The pseudoinverse  $A^+$  generalizes the concept of matrix inverse and arises in the minimum norm (that is, approximate as opposed to exact) or best-fit (in a least squares sense) solution to a system of linear equations. The problem  $\min_x \|Ax - P_0^\mu\|_2$  has the solution:  $x = A^+ P_0^\mu$ . The pseudoinverse can be computed using SVD as follows: if  $A = U \Sigma V^T$ , where  $U$ ,  $V$  denote unitary matrices and  $\Sigma$  is a diagonal matrix containing the singular values of  $A$ , then  $A^+ = V \Sigma^+ U^T$ . The function `pinv` was used to compute the pseudoinverse in MATLAB.

3. Find a linear representation coefficient vector  $\alpha$  of that projection onto the tangent space:

This coefficient is calculated by solving the following equation:

$$\prod_{T^\mu} (P_0^\mu) = A \cdot \alpha \quad (9)$$

4. Estimate the missing components of the point  $P_0$  ( $P_0^\nu$ ):

The last step is finding a point on the subspace  $T^\nu$  that is closest (in norm) to the point  $P_0$ .

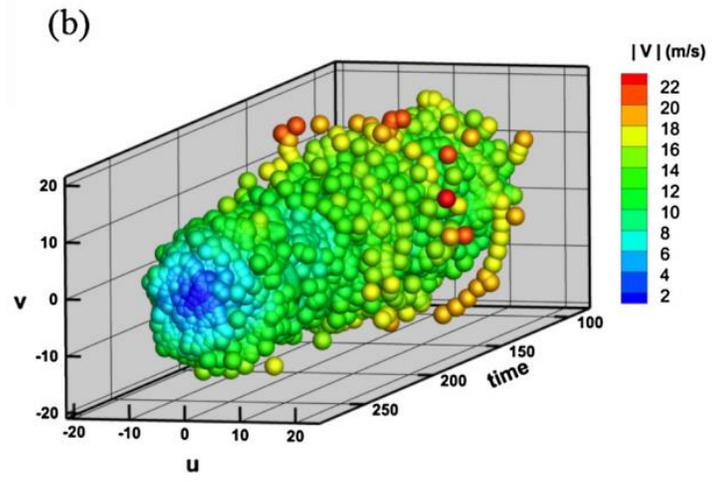
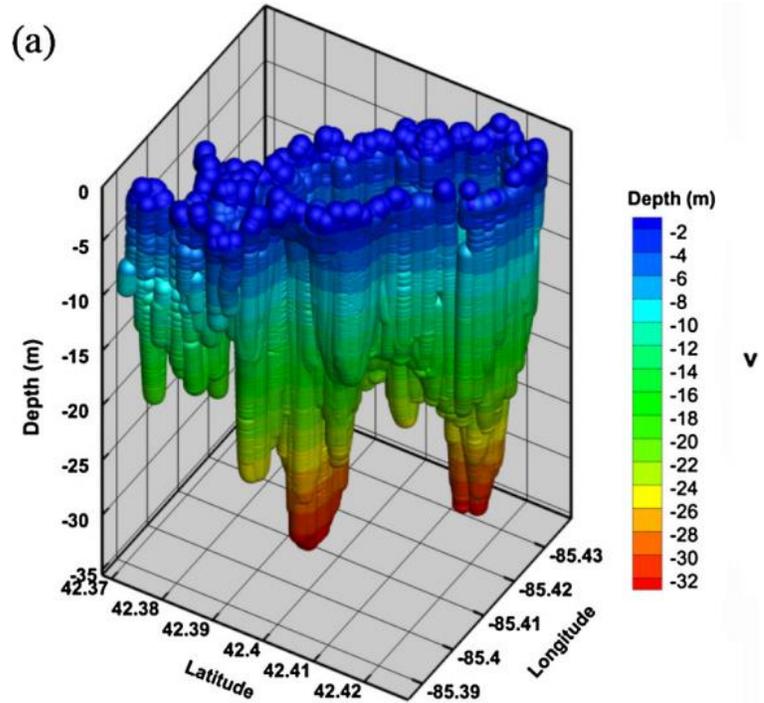
In order to do that,  $T^\nu$  is projected using the projection coefficient  $\alpha$  computed in step 3:

$$P_0^\nu \cong T^\nu \cdot \alpha \quad (10)$$

The result of this projection is the closest point to  $P_0^\nu$  that belongs to its subspace. In this algorithm, high-dimensional coordinates of selected neighborhoods on the manifold are projected to a low-dimensional subspace. An alternative to this approach is to use kernel regression to assign a weight to each neighbor based on the distance from  $P_0^\mu$  (Figure 2-3b). A weight for each selected neighborhood can be computed using the following Gaussian kernel function:

$$W_i = e^{-\frac{(P_i^\mu - P_o^\mu)^2}{2\sigma^2}}, \sigma = \sqrt{\text{var}(d(P_i^\mu, P_o^\mu))} \quad (11)$$

Examples of manifolds representing geophysical (bathymetry) and meteorological (wind) data are shown in Figures 2-4(a) and (b). These figures support the assumption that the manifold can be considered as being linear locally, but with complicated topology overall.



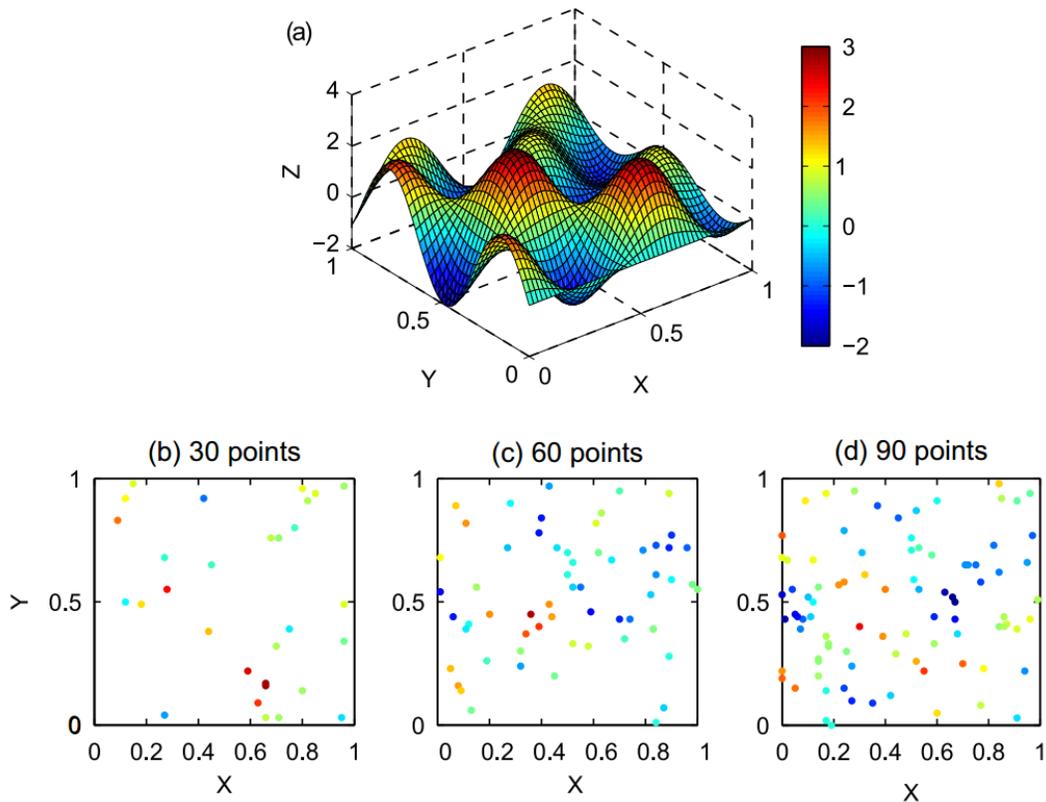
**Figure 2-4** Manifolds representing (a) bathymetry of Gull Lake and (b) wind components over Lake Michigan in three dimensional space.

### 2.2.1.2 Test case: Analytical function

Before applying the manifold method to reconstruct complex geophysical and meteorological data, we first evaluate the effectiveness of the method in reproducing an analytical function, since errors can be computed relative to the known function values; therefore, the F7 function suggested by *Lazzaro and Montefusco* [2002] and *Renka and Brown* [1999] is used:

$$F7(x, y) = 2 \cos(10x) \cdot \sin(10y) + \sin(10x \cdot y) \quad (12)$$

where the domain of F7 is restricted to  $0 \leq x \leq 1$  and  $0 \leq y \leq 1$  (Figure 2-5a). Three sets of sparse random points from a normal distribution were generated in the domain with numbers of sampling points of 30, 60, and 90. The F7 function was sampled randomly as shown in Figure 2-5b.



**Figure 2-5** (a) Analytical function used to test the manifold method for interpolation of scattered data. Random sampling was used to generate scatter points as shown in figures (b, 30 points), (c, 60 points) and (d, 90 points) to reconstruct the function.

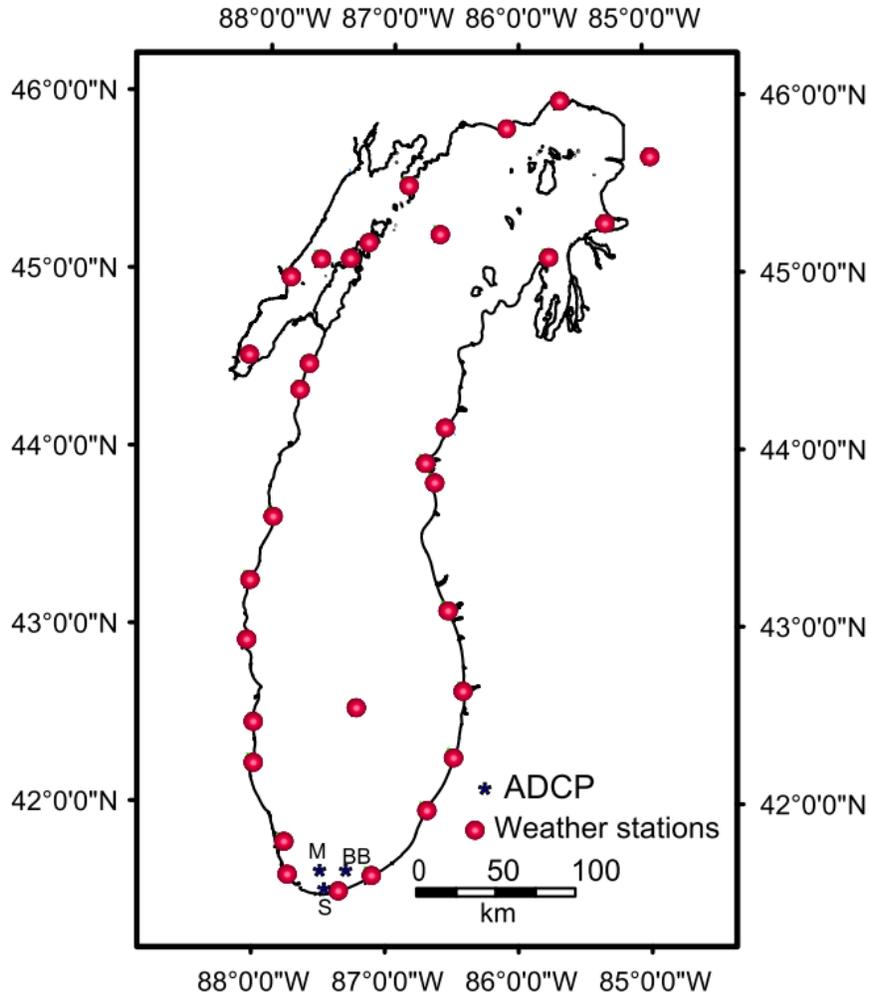
The manifold method was tested by withholding one point at a time and estimating its associated value from the remaining points using the manifold method, in addition to other methods such as the natural neighbor, nearest neighbor, and IDW interpolations. Since known components of the scatter points are located in the two-dimensional X-Y plane, at least two neighboring points are needed to form a tangent space for the manifold method. Therefore, for simplicity, only two nearest neighbors are used in both manifold and IDW interpolation methods.

### **2.2.1.3 Assimilating meteorological data for improved lake circulation modeling: Lake Michigan**

The proposed method was first applied for the reconstruction of wind fields (time-varying vector fields) over Lake Michigan. Hourly wind speed and direction data during April-September 2008 were obtained from the National Data Buoy Center (NDBC) weather stations surrounding the lake (Figure 2-6). The wind measurements were adjusted to a 10 m anemometer height using the profile methods described in Schwab (1987). Since the aerodynamic roughness over the lake is much lower compared to its counterpart over the land, an empirical overland-overlake adjustment was applied to the wind speeds recorded by overland stations [*Schwab and Beletsky, 1998*]. The datasets of wind speed and direction were converted to two coordinates in the Cartesian coordinate system ( $x$  and  $y$  directions).

Instead of using the cross-validation method to evaluate the interpolated wind data, results from the hydrodynamic model of the lake were compared with current measurements to test the applied method. To this end, a well-tested three-dimensional hydrodynamic model of the lake [*Safaie et al., 2016*] was used. The model was based on the unstructured grid Finite Volume Community Ocean Model (FVCOM; *Chen et al. [2006]*) which was successfully used in the past in ocean [*Li et al., 2014b*], lake [*Nguyen et al., 2014*] and river [*Anderson and Phanikumar, 2011*]

modeling. Details of the unstructured mesh used in the hydrodynamic model are presented in Table 2-1.



**Figure 2-6** Locations of the ADCPs deployed during summer 2008 and weather stations surrounding Lake Michigan.

Wind fields from April to September 2008 were reconstructed at the locations of nodes in the numerical mesh. Other hourly meteorological observations related to heat flux fields, including air temperature, cloud cover, dew point, long-wave solar radiation, short-wave solar radiation, and relative humidity, obtained from the National Climatic Data Center (NCDC) and NDBC stations, were interpolated over the computational grid using a smoothed natural neighbor method with a

smoothing radius of 30 km. Air pressure was assumed to be constant ( $10^5$  Pa) through the course of the study and a constant startup water temperature with a value of 2.5 °C was used in the model.

**Table 2-1** Properties of the numerical grids used for the hydrodynamic models.

Site	Grid Classification	Element shape	Grid Resolution	# Nodes	# Elements	#Vertical layers
Lake Michigan	Unstructured	Triangle	4m -5 km	12,684	23,602	20
Gull Lake	Unstructured	Triangle	8-100m	5,132	9,361	20

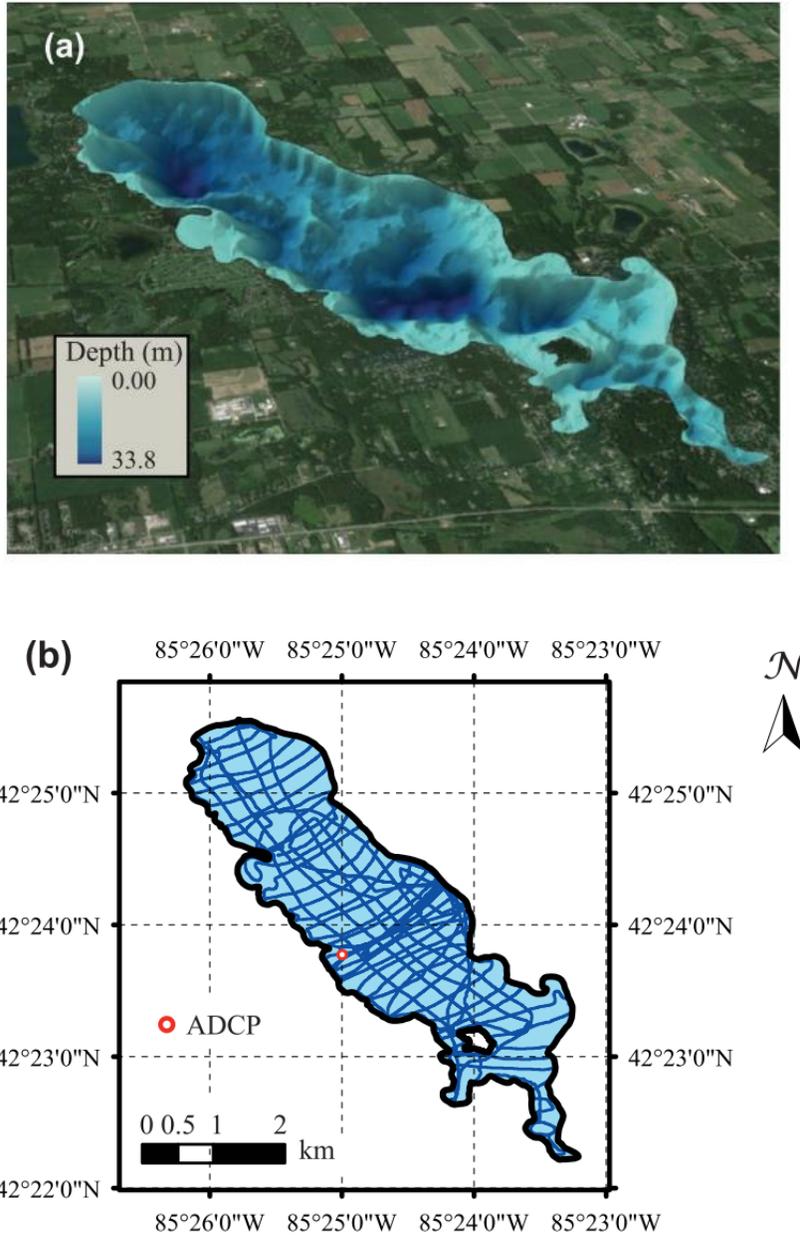
The overlake dew points were estimated from overland observations using an empirical formula described in [Schwab and Beletsky, 1998]. Air temperature and cloud cover were used to estimate long-wave solar radiation [Parkinson and Washington, 1979] and short-wave solar radiation was modeled using the clear-sky value and cloud cover [Nguyen et al., 2014]. Six arc-second bathymetric data obtained from the NOAA National Geophysical Data Center (NGDC) combined with two-meter resolution LIDAR data along the Indiana coast from the National Oceanic and Atmospheric Administration (NOAA) were interpolated to the numerical mesh using the natural neighbor method [Safaie et al., 2016].

Three bottom-mounted, upward-looking Acoustic Doppler Current Profilers (ADCPs) were deployed at stations M, BB and S (Figure 2-6) in southern Lake Michigan from early June to late August 2008 to measure nearshore currents for model testing [Thupaki et al., 2013; Safaie et al., 2016]. The hydrodynamic model was run from April to August 2008 to have a two-month spin-up period. Evaluation of the manifold method was carried out by comparing the simulated currents with data collected by the ADCPs. Comparisons between simulated and observed currents can be improved by identifying an optimal set of parameters in the manifold method. These parameters

include: an optimum number of the nearest neighbors to create a tangent space, the time scale  $c$ , and parameters of the Gaussian kernel function. In addition, the method used for creating a tangent space from a set of data points can be changed to improve the agreement between simulated and observed currents. The manifold method for the reconstruction of wind fields was directly applied to reconstruct the other six scalar observations, including air temperatures, cloud cover, dew point, relative humidity, shortwave and longwave solar radiation, to calculate the heat flux fields. This time, however,  $P^v$  is a scalar, rather than a vector.

#### **2.2.1.4 Assimilating geophysical data for improved lake circulation modeling: Gull Lake**

In the second example, the bathymetry of Gull Lake was reconstructed using a manifold method. The lake bathymetry data were collected using a SonTek RiverSurveyor M9 system. The M9 system has an Acoustic Doppler Profiler (ADP) with two sets of four profiling beams and one vertical acoustic beam (0.5-MHz echo-sounder) for river discharge measurements and bathymetric surveys. The system was equipped with differential GPS with sub-meter precision and mounted on a SonTek hydroboard to avoid high pitch and roll angles. The vertical acoustic beam has a range of 0.2 m to 80 m with an accuracy of 1% and a resolution of 0.001 m. The bathymetry survey was performed in four days (June 9 – June 12, 2015) by collecting data along longitudinal and transverse transects of the lake with an approximate interval of 200 m between each transect pair and sampling interval of 0.2 m- 2 m along the transects depending on the boat speed (Figure 2-7).



**Figure 2-7** (a) Bathymetry of Gull Lake. (b) Boat tracks generated during the sampling survey in Gull Lake.

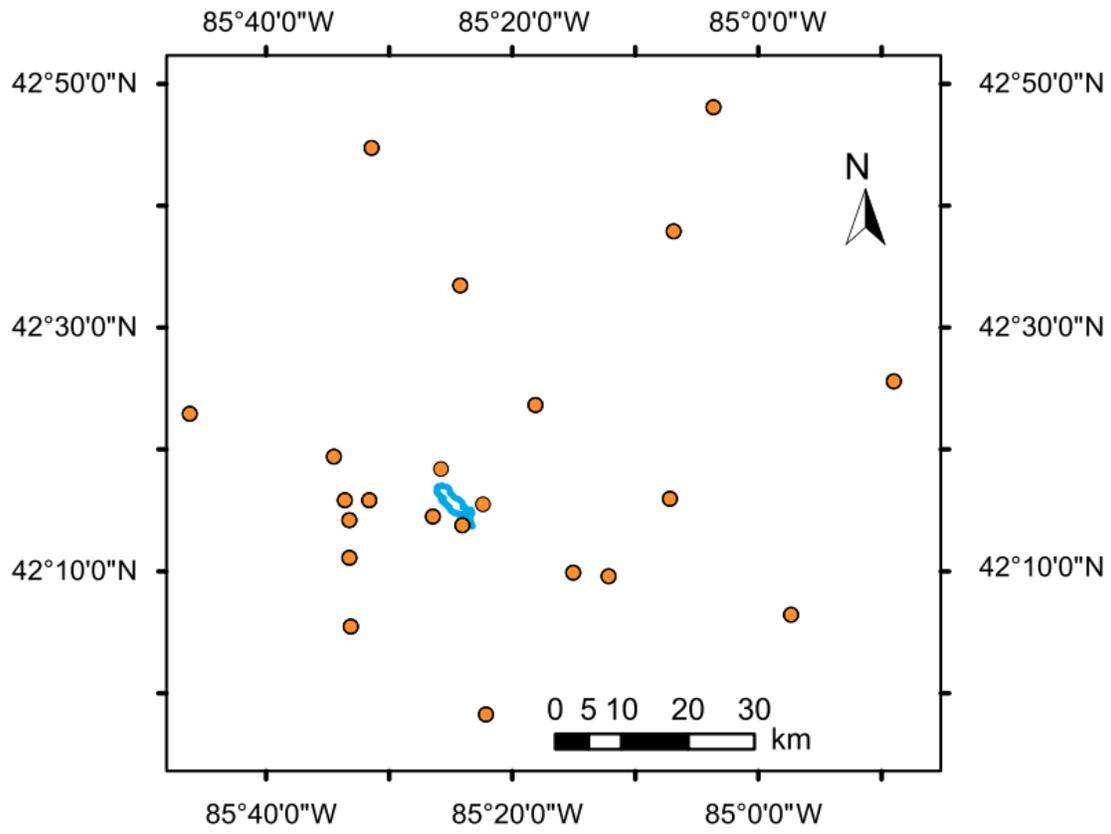
In order to assimilate the bathymetry of the lake, a three-dimensional hydrodynamic model based on FVCOM has been developed for the lake during the period of thermal stratification (June-August of 2014). The hydrodynamic equations were solved by the numerical model on an unstructured grid and details are given in Table 2-1.

The surface water temperature was collected using an Onset HOBO Pro v2 sensor with an accuracy of 0.2 °C. A linearly varying startup water temperature was used with a value of 12 °C at the water surface and 4 °C at the depth of 10 m. The hydrodynamic model was tested using observed current data measured using a Teledyne - RDI Sentinel-V ADCP (1000 kHz frequency with a bin size of 0.3 m) deployed in the nearshore waters of the lake in approximately 10 m of water (Figure 2-7b). Finally, the bathymetry of the lake interpolated to grid nodes using the manifold method was assimilated into the model.

## **2.2.2 Reconstruction of meteorological data in Gull Lake**

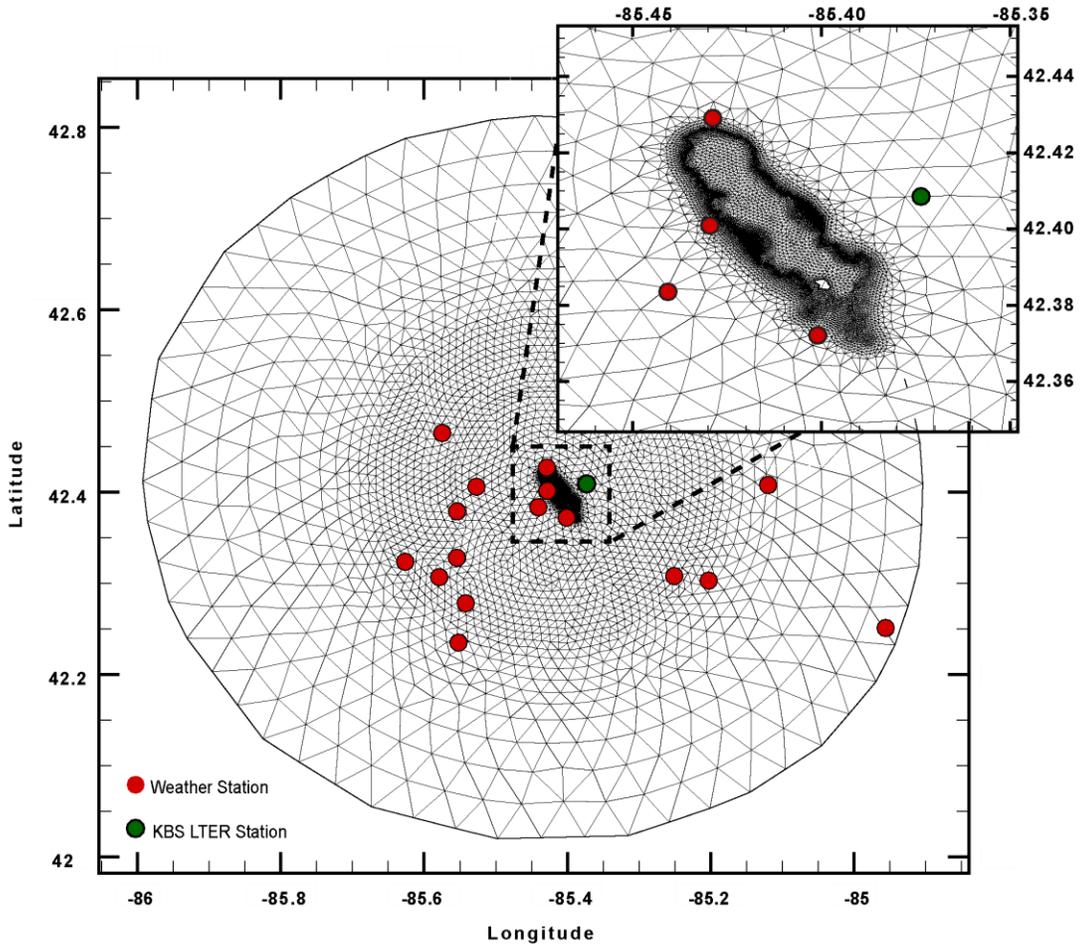
### **2.2.2.1 Land-based weather station network**

Hourly meteorological observations, including wind speed and direction, air temperature, air pressure, solar radiation, and relative humidity are needed for calculation of wind and heat flux fields. The meteorological observations were obtained from NCDC, Weather Underground (<https://www.wunderground.com>), and the Kellogg Biological Station Long-Term Ecological Research (KBS LTER, <http://lter.kbs.msu.edu>) stations, a total of 22 locations surrounding Gull Lake from May to August of 2014 and 2015 (Figure 2-8). Instead of a constant air pressure, hourly air pressure data recorded by the KBS LTER stations were used to improve the performance of the model. This also helped in the calculation of water density in FVCOM based on a polynomial expression [*Jackett and McDougall, 1995*] that takes pressure into account. Long wave solar radiation was estimated using air temperature and cloud cover [*Parkinson and Washington, 1979*].



**Figure 2-8** Selected weather stations surrounding Gull Lake.

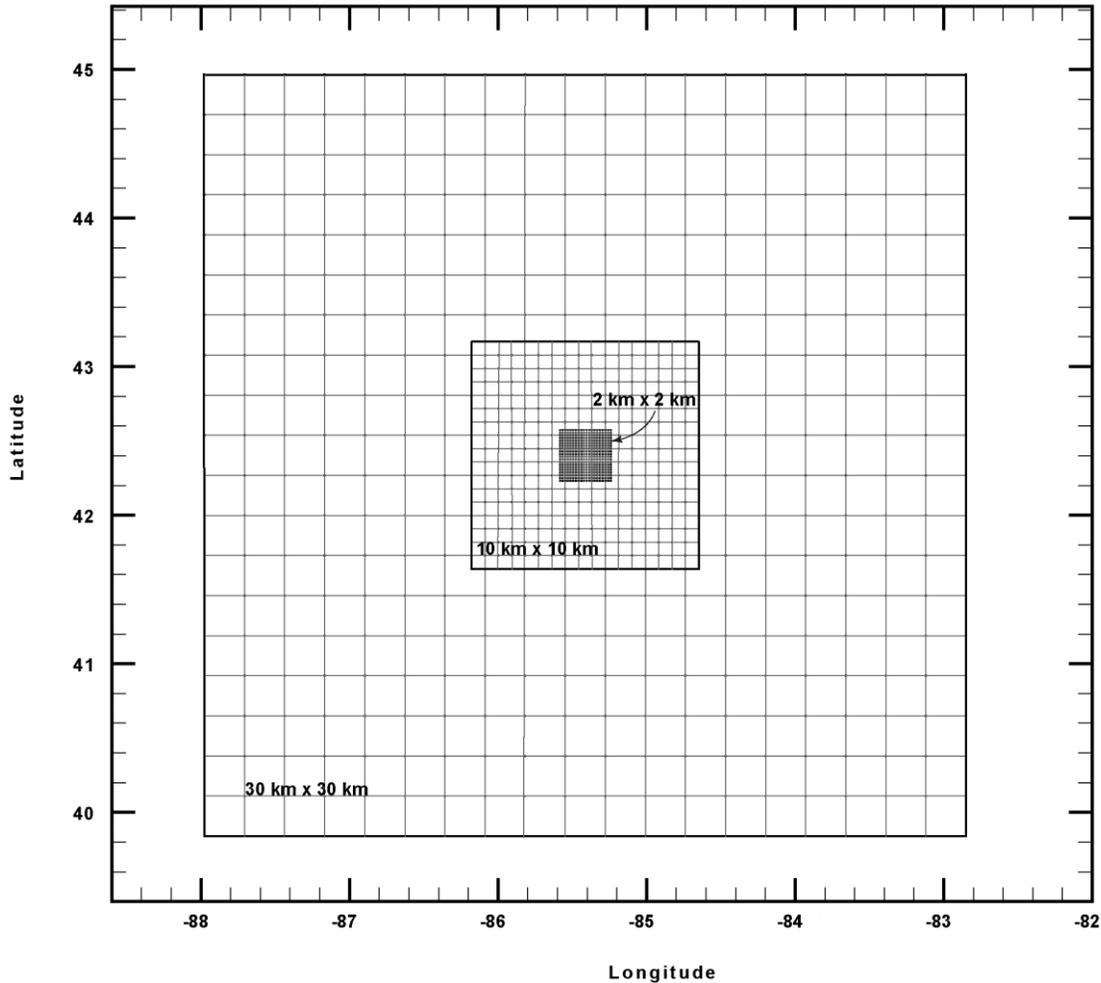
Since all weather stations are located within a 42 km radius of Gull Lake, a circular grid with a radius of 42 km was created for the reconstruction of meteorological fields (Figure 2-9). After applying overland-overlake adjustments, all observations were interpolated over the circular grid using a smoothed natural neighbor method. In order to have smooth and consistent wind and heat flux fields, a spatial moving averaged filter with a radius of 6 km was applied to the fields. This radius provided the best simulated results between the ranges of 0 to 30 km. Magnitudes of the smoothed fields were then adjusted based on the observations at the KBS LTER station to compensate for the reduction in magnitudes caused by the smoothing filter. Finally, the results were interpolated to the FVCOM’s computational grid using the natural neighbor interpolation method.



**Figure 2-9** Circular grid with a radius of 42 km used for the reconstruction of meteorological fields from the land-based weather station network.

### 2.2.2.2 Mesoscale weather prediction model

The Weather Research and Forecasting model (WRF3.7.1, <http://www.wrf-model.org>) with a lake physics component was used in this study to simulate meteorological forcing over Gull Lake. Three nested simulation domains with resolutions of 30 km, 10 km, and 2 km were defined and centered at (42.402778 °N, 85.41295 °W) on Gull Lake (Figure 2-10).



**Figure 2-10** Three nested simulation domains with resolutions of 30 km, 10 km, and 2 km used for the WRF model.

The 20-category MODIS-based land use data with a 30 s resolution of static data were interpolated to the model grids. USGS 24-category data were used as well, as an alternative set of land use, if a category from the MODIS-based data was not available. There are a number of meteorological reanalysis datasets which can be used as input to the WRF model. National Centers for Environmental Prediction / National Center for Atmospheric Research (NCEP/NCAR) reanalysis data provides four-times daily data with a spatial resolution of 206 km. The 32 km, three-hourly NCEP North American Regional Reanalysis (NARR) data can also be

used to generate initial and lateral boundary conditions. We selected Community Land Model Version 4 (CLM4, *Oleson et al.* [2010]) for the WRF's land surface scheme in order to model the land - atmosphere interaction processes. The hourly wind field, air temperature, air pressure, shortwave and longwave solar radiations, and relative humidity were simulated from early April to September in years 2014 and 2015.

## 2.3 Results and Discussion

### 2.3.1 Manifold method

True values of the analytical function at each of the randomly selected sampling locations were compared with the estimated values obtained by the manifold method as well as other standard interpolation methods. The performance statistics for this example are provided in Table 2-2. Definitions of metrics used in this study are provided in the Appendix. For all methods, the approximation of the F7 function improved by increasing the number of sampling points. In this particular example, the results show that the manifold method produced better overall performance compared to the other three methods considered. However, the best method in this example might perform differently on another test function or for a different sampling point distribution. Therefore, we examine the performance of the method for other datasets in the Lake Michigan and Gull Lake.

Due to the sparse distribution of weather stations around Lake Michigan, it was not clear *a priori* how many neighboring stations would provide an adequate representation of the data. Since choosing a relatively few (e.g., three) neighboring stations in this situation would involve using information from stations that are far apart as neighbors, we used kernel regression to assign weights to each station depending on the distance from the point of interest. For each node of the

numerical grid of Lake Michigan,  $k$  number of nearest neighbors were selected and their assigned weights were projected to a low-dimensional subspace.

**Table 2-2** Cross-validation results for the analytical function based on different sampling points selected randomly.

Sample size	Method	R <sup>2</sup>	RMSE	Fn	PBIAS	NASH	APB (%)
30	Manifold	0.667	0.778	0.710	14.420	0.416	0.652
	Natural neighbor	0.582	0.876	0.799	-8.866	0.259	0.713
	Nearest neighbor	0.619	0.882	0.804	-14.689	0.250	0.669
	IDW	0.577	0.870	0.793	35.847	0.270	0.727
60	Manifold	0.846	0.579	0.531	-33.924	0.703	0.472
	Natural neighbor	0.816	0.615	0.564	16.912	0.664	0.466
	Nearest neighbor	0.779	0.720	0.660	-34.524	0.540	0.512
	IDW	0.832	0.603	0.553	-44.292	0.677	0.469
90	Manifold	0.891	0.502	0.432	-14.103	0.791	0.400
	Natural neighbor	0.874	0.539	0.464	-10.666	0.759	0.344
	Nearest neighbor	0.867	0.571	0.491	-28.777	0.730	0.446
	IDW	0.859	0.567	0.487	-5.974	0.735	0.416

The free parameters in the method are  $c$  (time scale),  $\sigma$  (the parameter used in kernel regression), and  $k$ . The standard deviation of weather station distances from the point of interest was used for the parameter  $\sigma$  in kernel regression. Performance of the manifold method as measured by a comparison of simulated and observed currents in Lake Michigan is summarized in Table 2-3 relative to the other standard methods considered. We note that the manifold method based on kernel weighting considering all stations produced the best overall performance as measured by the root mean squared error (RMSE) between the observed and simulated currents. The performance of the method without kernel regression and with only three neighboring stations

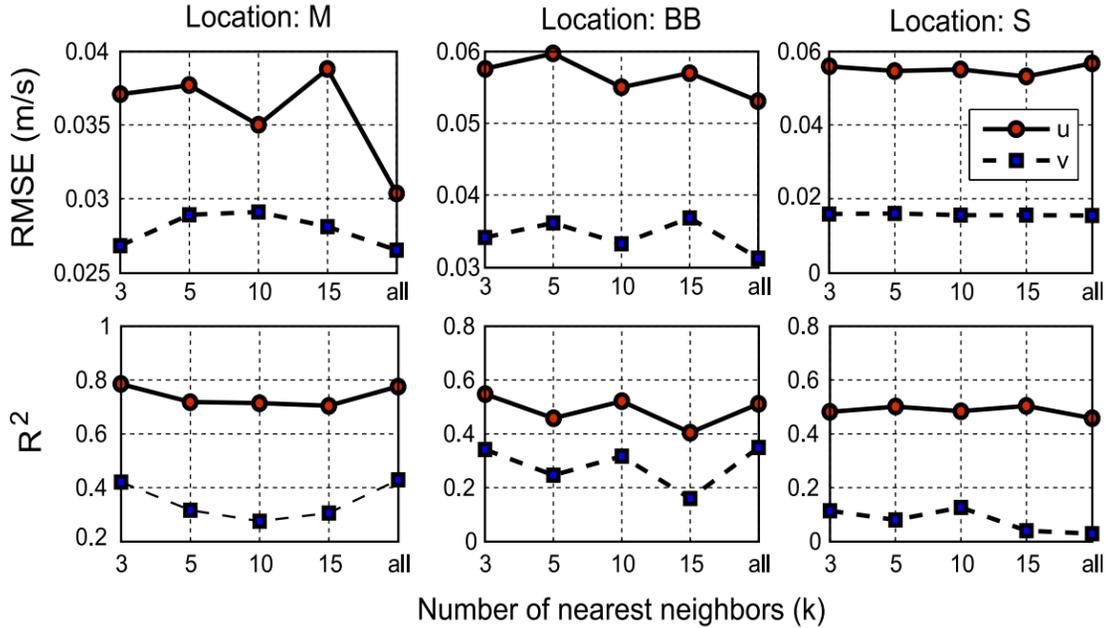
was comparable to the other methods, but slightly inferior to the natural and nearest neighbor methods.

**Table 2-3** RMSE values (m/s) of alongshore and cross-shore velocities in Lake Michigan for comparison of the manifold method with other standard methods used in limnology and oceanography.

Method	Location: M		Location: BB		Location: S	
	RMSEu	RMSv	RMSEu	RMSEv	RMSEu	RMSEv
O-kriging	0.0385	0.0290	0.0590	0.0349	0.0540	0.0152
Nearest Neighbor	0.0363	0.0286	0.0580	0.0348	0.0545	0.0152
Natural Neighbor	0.0366	0.0275	0.0553	0.0334	0.0515	0.0158
Manifold (3 NBR)	0.0383	0.0276	0.0594	0.0346	0.0568	0.0158
Manifold+Kernel (3 NBR)	0.0371	0.0268	0.0576	0.0341	0.0559	0.0158
Manifold+Kernel (all NBR)	0.0304	0.0265	0.0531	0.0312	0.0568	0.0154
IDW (all NBR)	0.0328	0.0267	0.0535	0.0316	0.0498	0.0155

Figure 2-11 shows the RMSE and  $R^2$  values for different numbers of nearest neighbors at different ADCP locations. Having all stations to create the tangent space for the manifold method resulted in a better representation of wind fields, and improved the results of the hydrodynamic model (Figure 2-12). Cross-validation was also used to compare the performance of the manifold method with other standard methods for the same Lake Michigan datasets. The performance metrics are summarized in Table 2-4. In this cross-validation method, one weather station was withheld to be used later for validating the manifold method, and all other stations surrounding the lake were used for the manifold training set. This process was repeated so that each weather station was given a chance to be part of this validation process. Based on these results the proposed manifold method with three nearest neighbors gave better results compared to other standard methods. However, the performance of the hydrodynamic model based on these methods was

relatively inferior compared to the performance of the model when the manifold method used all neighboring points.



**Figure 2-11** Performance of the manifold method evaluated using observed and simulated currents at different stations in Lake Michigan. Different number of nearest neighbors were used to reconstruct the wind field using the manifold method with kernel regression.

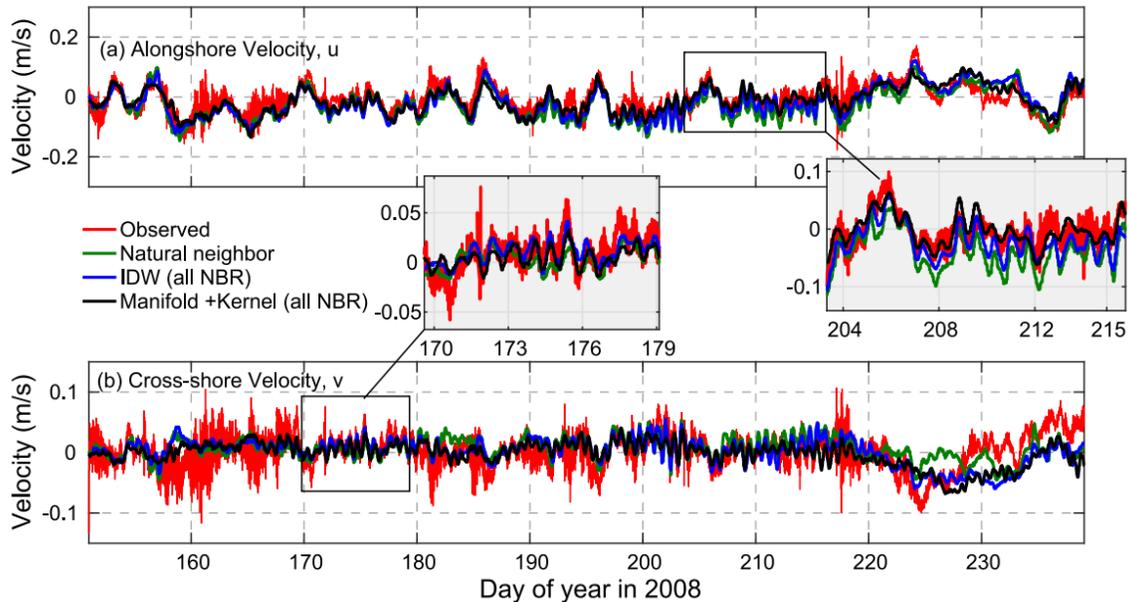
**Table 2-4** Cross-validation results for wind field over Lake Michigan.

Method	$R^2_u$	$R^2_v$	RMSE <sub>u</sub>	RMSE <sub>v</sub>	Computational time (s)
O-kriging	0.441	0.572	3.497	3.853	92463.8
Nearest Neighbor	0.666	0.743	2.792	3.044	18.6
Natural Neighbor	0.693	0.794	2.558	2.750	183.6
Manifold (3 NBR)	0.690	0.801	2.433	2.595	28.1
Manifold+Kernel (3 NBR)	0.710	0.806	2.392	2.566	55.1
Manifold+Kernel (all NBR)	0.547	0.681	2.884	3.129	77.3
IDW (3 NBR)	0.724	0.822	2.278	2.458	69.3

All different versions of the manifold methods had reasonable computational efficiency. The computational time for the O-kriging was high due to the time needed for finding the best

variogram at each time step. Computations were performed using MATLAB, on an Intel Core i7-4771 3.5 GHz platform.

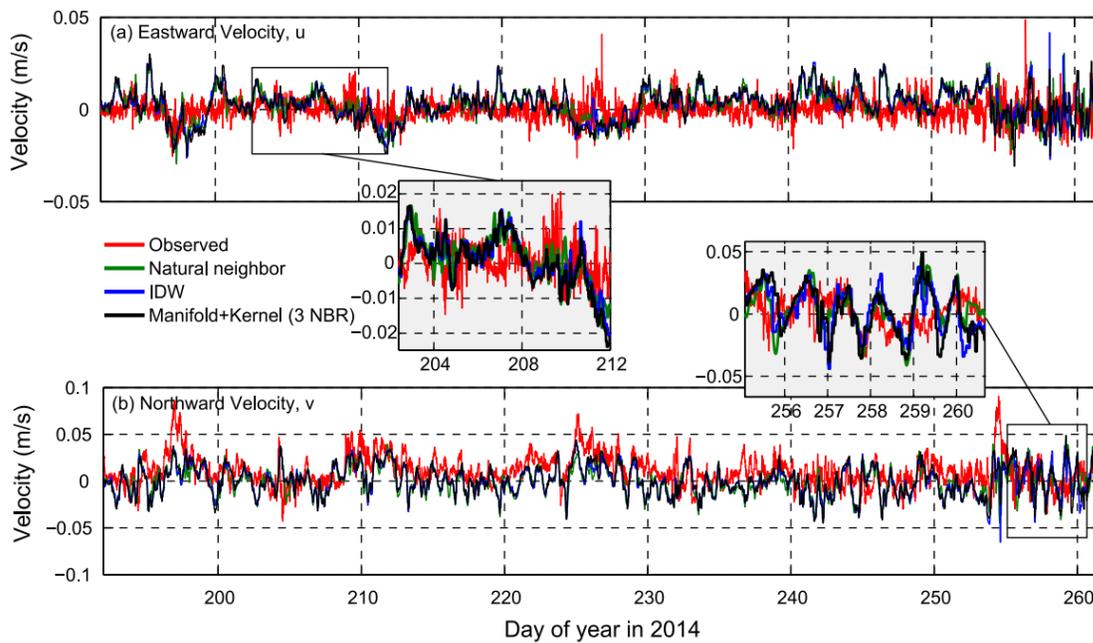
As the last case study, wind and heat flux fields of the Gull Lake were interpolated over the numerical mesh of the lake using the natural neighbor method. Then, the bathymetry of the lake was interpolated over the mesh using the same natural neighbor method to develop the initial version of the lake hydrodynamic model. The raw bathymetry data, which has some regions of steep bathymetry change, created artificial currents in the model due to an error in the pressure gradient force introduced by the sigma-coordinate system of FVCOM [Mellor *et al.*, 1998]. Therefore, the interpolated bathymetry was smoothed with a radius of 100 m in order to reduce the errors.



**Figure 2-12** Comparison of simulated and observed vertically averaged currents at the location M in Lake Michigan. (a) Alongshore velocity (b) Cross-shore velocity

The developed model was used to assimilate the bathymetry of the lake based on the manifold method. First, the bathymetry data were reconstructed from the tangent space of the manifold with

three nearest neighbors and smoothed with the same method described above. Then, the hydrodynamic model was run with the reconstructed bathymetry. The comparisons of the vertically-averaged velocity profiles at the ADCP location using natural neighbor method, IDW with three nearest neighbors, and manifold method are presented in Figure 2-13. The best value of  $\sigma$  used in kernel regression was equal to the standard deviation of distances of observational points where water depth values are available within a search radius of 50 m from the point of interest.



**Figure 2-13** Comparison of simulated (black lines) and observed (red lines) vertically averaged currents at the ADCP location in Gull Lake. (a) Eastward velocity and (b) Northward velocity

When the number of samples within this radius was smaller than 100,  $\sigma$  value was calculated based on locations of 100 nearest samples. This method is more accurate when enough samples are available around an estimated point, unless selecting 100 samples itself does a reasonable job. RMSE values (m/s) of eastward and northward velocities in Gull Lake for comparison of the manifold method with other standard methods used in limnology and oceanography are presented in Table 2-5.

**Table 2-5** RMSE values (m/s) of eastward and northward velocities in Gull Lake for comparison of the manifold method with other standard methods used in limnology and oceanography.

Method	RMSE u	RMSE v
Natural Neighbor	0.0090	0.0205
Manifold+Kernel (3 NBR)	0.0098	0.0200
IDW (3 NBR)	0.0095	0.0204

The statistics of cross-validation for all (=71) measured longitudinal and transverse transects are shown in Table 2-6. The cross-validation was performed by omitting one transect at each step and calculating the bathymetry for that transect from the rest of the observation data and repeating the process for all other transects.

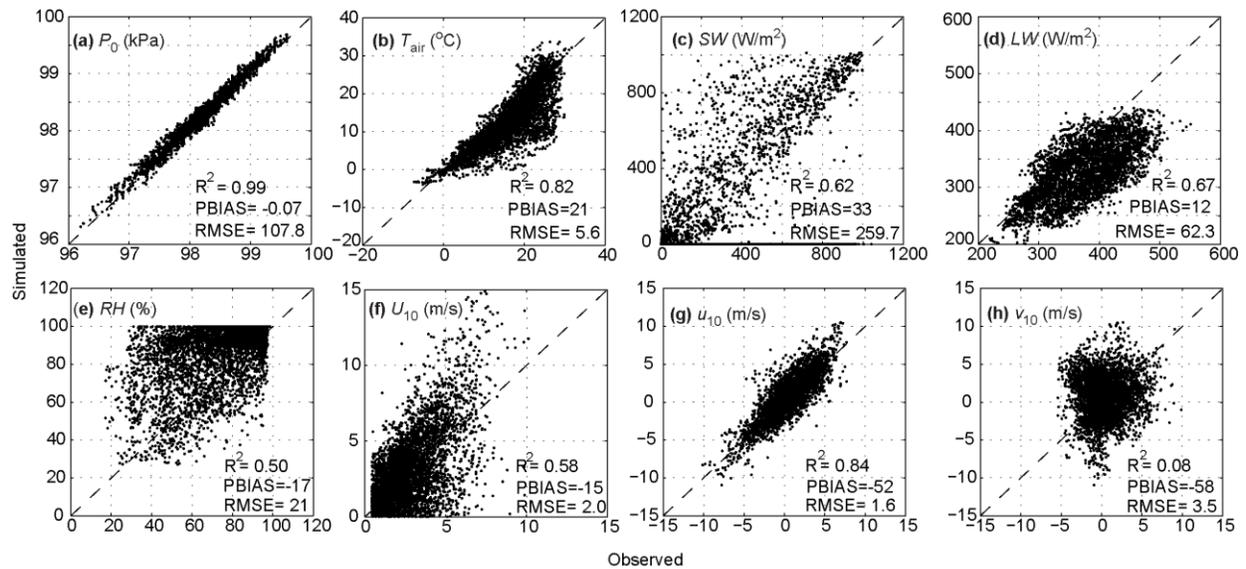
**Table 2-6** Cross-validation results for Gull Lake bathymetry.

Method	R <sup>2</sup>	RMSE (m)	Fn	NASH	PBIAS
Manifold	0.890	2.011	0.222	0.678	-14.016
Natural Neighbor	0.925	1.288	0.170	0.468	-13.301
Nearest Neighbor	0.888	2.039	0.230	0.670	-17.132
IDW (3 NBR)	0.839	3.282	0.6065	0.540	-15.918

### 2.3.2 Observed and predicted meteorological data in Gull Lake

Two types of reanalysis data (NCEP/NCAR and NARR) were tested for use as input data of the WRF model. The WRF simulated weather data were compared with the observed data at the KBS LTER station to evaluate the performance of the weather forecast model. Comparisons of observed meteorological data at KBS LTER station with simulated results based on NCEP/NCAR are presented in Figure 2-14. Figure 2-15 also shows scatter plots of WRF-simulated results based on NARR data versus observed meteorological forcing at KBS LTER. The simulated results, particularly wind speed and air temperature, indicated that NARR data would be a better choice

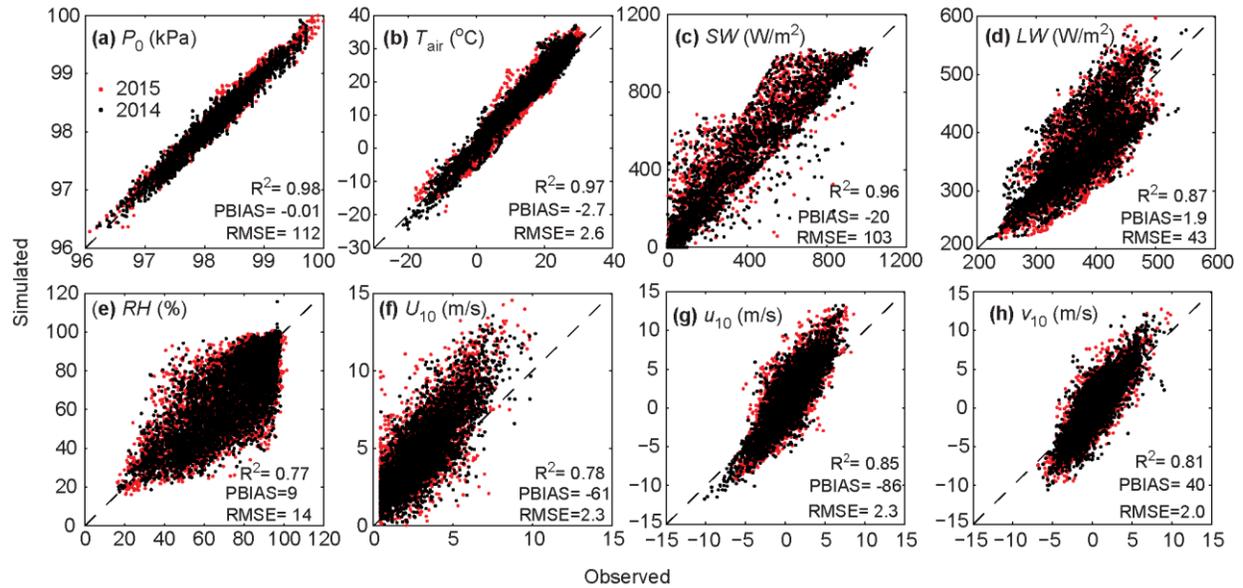
for the input of the hydrodynamic model. Thus, NARR data will be utilized in Chapter 3 to simulate the meteorological forcing for the coupled WRF-lake model. The results showed promising performance in simulation of the weather data at the KBS LTER station. Comparisons of averaged simulated forcing and meteorological forcing over Gull Lake during April to mid-September of 2014 and 2015 are presented in Figure 2-16 and 2-17, respectively.



**Figure 2-14** Comparisons of observed meteorological data at KBS LTER station with simulated results based on NCEP/NCAR data during April to mid-September of 2014 (a) air pressure (kPa), (b) air temperature ( $^{\circ}C$ ), (c) downward shortwave radiation ( $W/m^2$ ), (d) downward longwave radiation ( $W/m^2$ ), (e) relative humidity (%), (f) wind speed at 10-meter height (m/s), (g) eastward wind velocity (m/s), (h) northward wind velocity (m/s).

Since the raw wind fields reconstructed using weather station data alone were in good agreement with the overlake WRF model results, no overlake-overland adjustment [Schwab and Morton 1984] was applied to the meteorological forcing. An empirical overland-overlake adjustment has usually been applied to wind speeds recorded by land-based weather stations to determine overlake wind speed in the Great Lakes [Schwab and Beletsky, 1998; Thupaki et al., 2013; Nguyen et al., 2014; Safaie et al., 2016]. However, this study shows that the overland-overlake adjustment can generate high values of overlake wind speeds for a small inland lake,

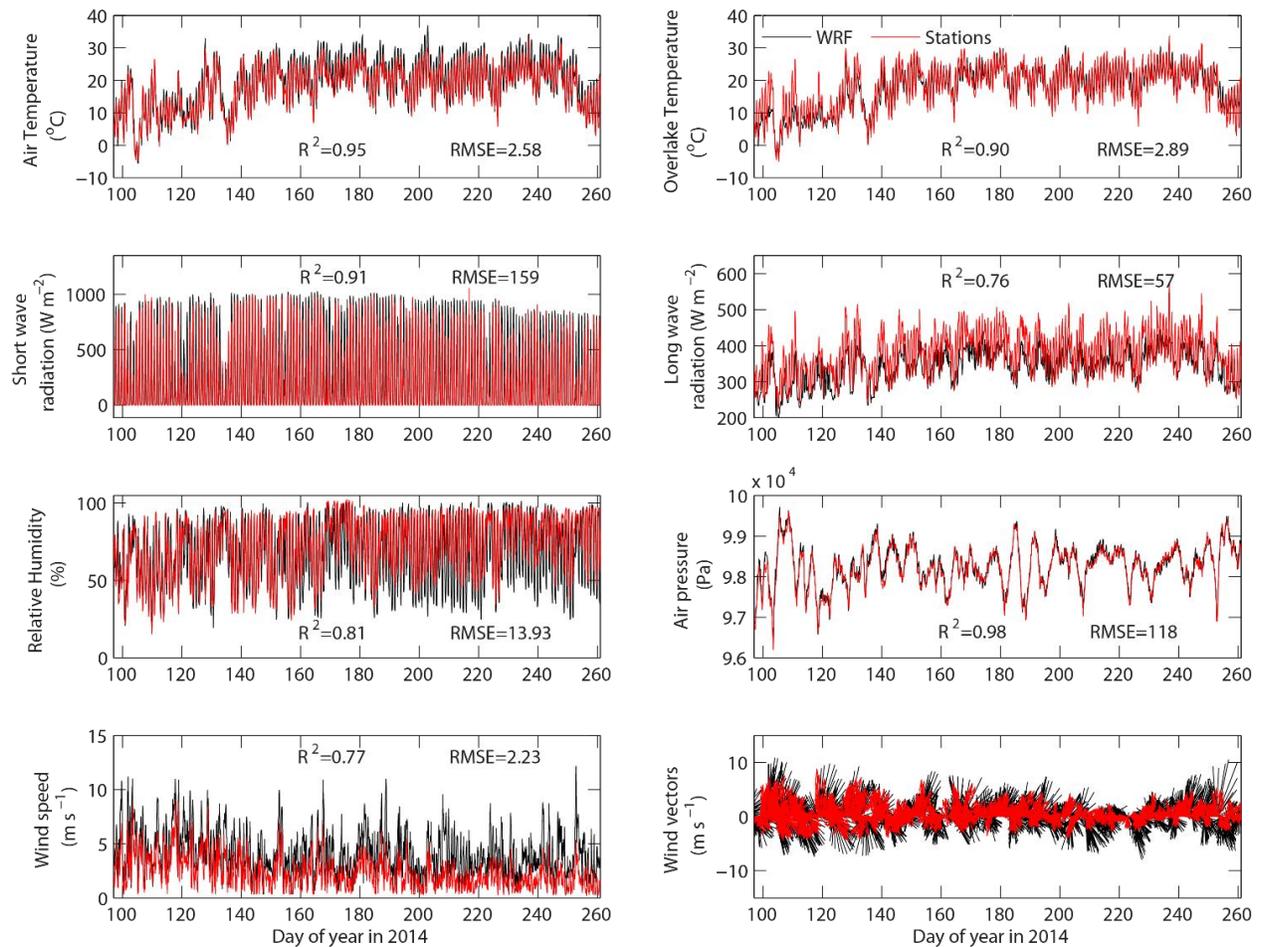
which has a short fetch length compared to large lakes. On the other hand, the primary results of the coupled WRF-lake model show that current speeds were overestimated due to the relatively high simulated wind speed.



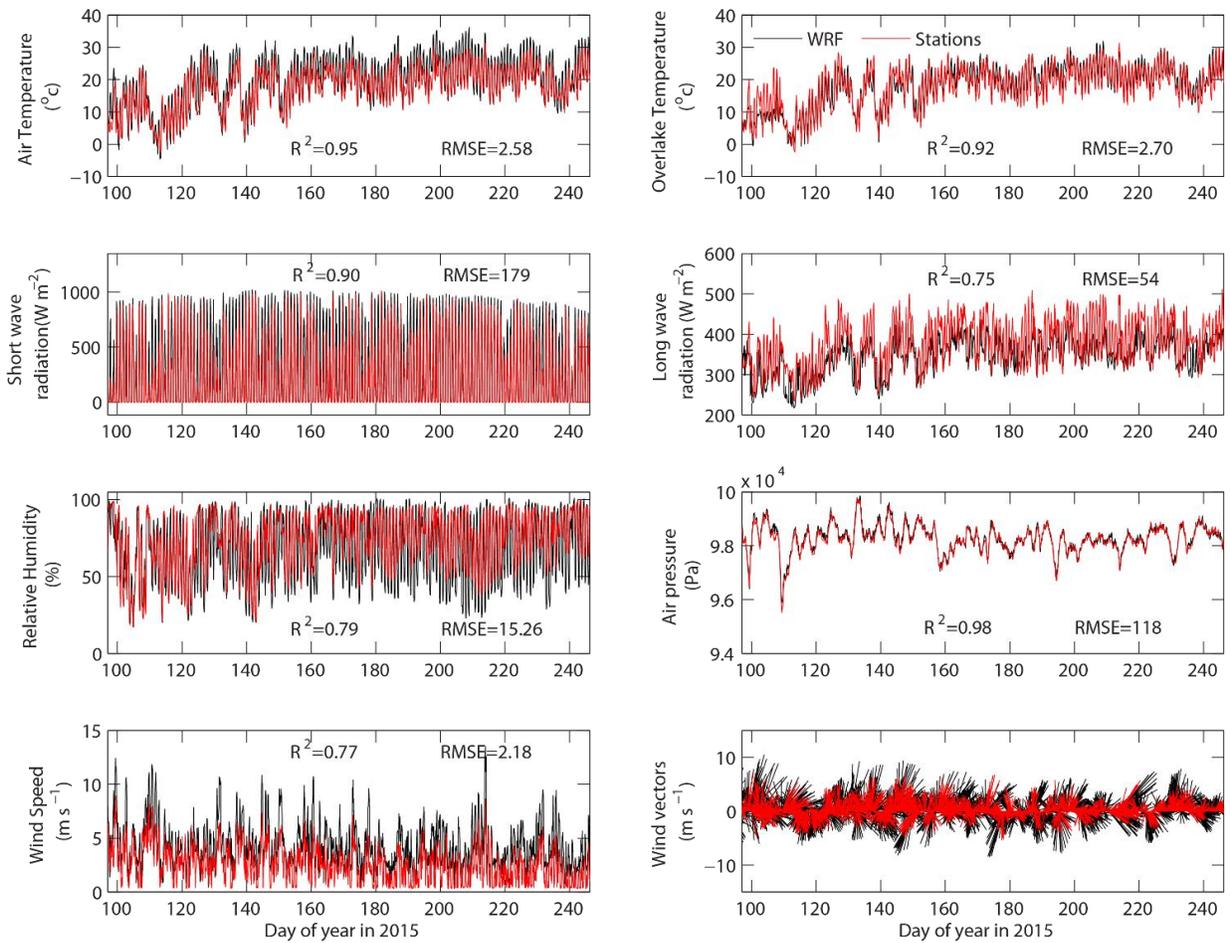
**Figure 2-15** Scatter plots of WRF-simulated results versus observed meteorological forcing at KBS LTER during April to mid-September of 2014 (black dot) and 2015 (red dot): (a) air pressure (kPa), (b) air temperature ( $^{\circ}\text{C}$ ), (c) downward shortwave radiation ( $\text{W}/\text{m}^2$ ), (d) downward longwave radiation ( $\text{W}/\text{m}^2$ ), (e) relative humidity(%), (f) wind speed at 10-meter height (m/s), (g) eastward wind velocity (m/s), (h) northward wind velocity (m/s).

A comparison of the simulated and observed wind speeds for years 2014 and 2015 at the KBS LTER station is shown in Figure 2-18. These comparisons indicate that for both 2014 and 2015, the simulated wind speeds are about 30 percent higher than observations. WRF is known to over predict the wind speed depends on the topographic complexity, drag parameterization, and its vertical and horizontal resolutions [Jiménez and Dudhia, 2012; Brunner et al., 2015; Staffell and Pfenninger, 2016]. Therefore, an adjustment factor of 0.7 should be applied into the WRF-simulated wind speeds to be used as inputs of the Gull lake model. It is also worth mentioning that initially the shortwave solar radiation was calculated using theoretical estimates of clear-sky solar

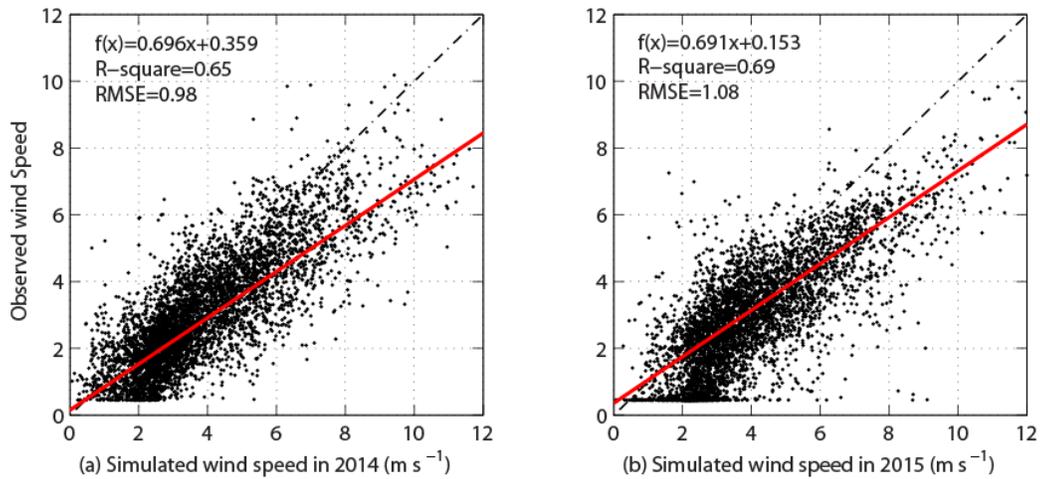
radiation [Annear and Wells, 2007], and adjusted by an empirical relation between the clear-sky value and the measured cloud cover [Bunker, 1976] to assess the WRF results. Since the nearest weather station (a NCDC station) with the cloud cover data was 23 km far from the lake, the calculated shortwave solar radiation becomes 30% smaller compared with WRF-simulated results. However, using the observed solar radiation data that were obtained from the KBS LTER station resulted in a better agreement with prediction of WRF.



**Figure 2-16** Comparisons of averaged simulated forcing (WRF) and observed meteorological forcing (Stations) over Gull Lake during April to mid-September of 2014.



**Figure 2-17** Comparisons of averaged simulated forcing and meteorological forcing over Gull Lake during April to mid-September of 2015.



**Figure 2-18** Comparisons of the simulated and observed wind speed in 2014 and 2015 at the KBS LTER station.

## 2.4 Conclusions

We presented a novel manifold method of reconstructing spatio-temporal data, which could be used for assimilating geophysical and meteorological data integrated land surface subsurface, and lake models. All case studies illustrate the superior performance of the presented manifold algorithm over standard methods in terms of accuracy and computational efficiency. The hydrodynamic model of Lake Michigan based on the manifold method of reconstructing wind fields produced better performance relative to the other methods. The best results were obtained using kernel regression applied to all weather stations (neighbors). However, the cross-validation results show that the results of the three nearest neighbors were better than the other methods. We can see that the manifold method performs better than the IDW method at two of the stations (M and BB) but not at the nearshore point S. We believe that the reason for this has to do with the fact that in the nearshore region there are a number of additional processes (waves, wave-current interactions etc) which are not simulated in our model. In other words, in that region the flow is not directly driven by the wind fields but rather indirectly through exchange between offshore and nearshore water and wave propagation. Therefore model performance in that region cannot be directly related to the wind field. At the other two offshore stations M and BB, the flow is predominantly wind-driven and an improvement in the simulated hydrodynamic fields can be seen.

This also brings us to two relevant points: (1) Details of the manifold method such as the tangent space estimation, the distance metric that defines spatiotemporal proximity and other details can all be further improved to improve the performance of the manifold method, but these topics are beyond the scope of the present study. (2) We do not claim that the manifold method provides superior performance on all datasets and for all performance metrics, but from the examples considered here it appears that the manifold method may offer an attractive method that

is comparable or superior to other standard methods. Clearly more research is needed to understand the relative strengths and weaknesses of different manifold-based approaches compared to standard methods. The Gull Lake model results indicated that the proposed method has the ability to reconstruct geophysical data at unsampled locations. The results highlight that evaluating the performance of interpolation methods using the cross-validation method without considering the data structure and the purpose of interpolating data can lead to misleading conclusions about the relative performance of the methods considered.

The WRF model was found to be predictive of meteorological data. Given the uncertainty involved due to the sparse distribution of weather stations, lack of quality assurance of raw weather data, and the overland-overlake effect, the weather prediction model could be utilized to assess the reconstructed meteorological forcing. The fact that meteorological forcing based on the outputs of a mesoscale weather prediction model (WRF) could provide results comparable to the forcing based on a network of weather stations is encouraging. However, further examination is needed to assess the accuracy of the WRF in representing the spatio-temporal fields of interest. Indeed, we can use the same approach presented here for evaluating the performance of the manifold method. Therefore, in the next chapter, a coupled WRF-lake model will be used to further assess the accuracy of meteorological forcing reconstructed from land-based weather stations and vice versa based on their ability to describe circulation and thermal structure of Gull Lake.

# CHAPTER 3

## 3 Evaluating the Role of Groundwater in Circulation and Thermal Structure

### 3.1 Introduction

Lake ecosystems readily respond to the influences of climate through a number of coupled surface and subsurface processes within their watersheds and therefore serve as sentinels of climate change [Adrian *et al.*, 2009; Schindler, 2009]. A recent global synthesis of *in situ* and remotely-sensed lake data indicated that lake surface temperatures increased quickly between 1985 and 2009 at a mean global rate of  $0.34^{\circ}\text{C}$  per decade. Similar warming trends were noted in streams, rivers and shallow groundwater as well [Kaushal *et al.*, 2010; Menberg *et al.*, 2014], although the rates of warming can be expected to be different for surface and groundwater systems. Due to the higher heat capacity of soils and the buffering effects of vegetation, groundwater temperatures remain relatively constant throughout the year partly offsetting the effects of surface warming in groundwater-fed lake ecosystems. While the effects of climate change on the Laurentian Great Lakes are well-documented (e.g., growing annual average temperatures, shorter winters, decreasing lake ice covers and ice-albedo feedback) and continue to receive significant attention (e.g., Austin and Colman [2007]; Nguyen *et al.* [2014]), there is growing evidence to indicate that

smaller inland lakes may respond differently to climate change than the Great Lakes of the world [Winslow *et al.*, 2015].

Temperature is a key driver that affects the structure and function of ecosystems including biological productivity. Therefore, important changes related to nutrient cycling, eutrophication and hypoxia can be linked to changes in lake circulation and thermal structure. Most lake models, however, do not explicitly consider the effects of groundwater and quantifying these processes for small lakes presents unique challenges. For example, choices associated with the representation of meteorological forcing, lake morphometry, and turbulent mixing within the water column can give rise to model uncertainties in temperatures comparable to the global mean warming rate noted above. Therefore, there is a need to understand and quantify the role of groundwater in circulation and thermal structure within lake ecosystems beyond simple water budgets. In particular, there is a need to systematically evaluate the impact of several modeling choices on model outcomes in order to identify the best choices relative to the spatial and temporal scales of interest. In inland shallow lakes, heat exchange between water and lake sediment was taken into consideration for vertical thermal diffusivity analyses [Hondzo *et al.*, 1991; Hondzo and Stefan, 1993]. Although, the role of groundwater in circulation and thermal structure of lakes are often assumed to be negligible, a shallow groundwater-fed lake may have significant bottom heating during winter and bottom cooling in summer [Kettle *et al.*, 2012]

Relatively small lakes such as Gull Lake, the focus of this study, have shorter fetch lengths compared to large lakes and wind speed can be significantly decreased due to sheltering effects [Hondzo and Stefan, 1993; Markfort *et al.*, 2010]. Therefore, wind- and wave-driven turbulent mixing are weaker in small lakes, and consequently, influences of water clarity on mixed-layer depth of small lakes are more pronounced than in large lakes [Heiskanen *et al.*, 2015]. Physical,

chemical, and biological processes at the air-water interface, mixed layer, and thermocline are directly influenced by the distribution of solar radiation in the water column [*Simpson and Dickey, 1981b*]. For example, phytoplankton competition for nutrients and light in a stratified water column is controlled by the ratio of light attenuation to their nutrient consumption [*Yoshiyama et al., 2009*]. A minor change in parameter values such as the mixed layer depth or light attenuation length may drastically change the vertical distribution of phytoplankton [*Mellard et al., 2011*]. Accurate mechanistic models capturing physical aspects of these processes hold the key to the understanding and predicting biophysical processes. Despite the availability of high-quality data based on decades of water quality monitoring and biological research in Gull Lake, a numerical hydrodynamic model of the lake has not been developed to date. One of the objectives of this chapter is to fill this gap as a first step towards building a coupled physical-chemical-biological modeling system.

The aim of this study was to evaluate the role of groundwater in circulation and thermal structure within Gull Lake, which is typical of many other groundwater-fed inland lakes in Michigan. Groundwater effects on lakes can be studied at different spatial and temporal scales. The focus of the present work is on examining lake hydrodynamics and thermal structure during the summer stratified period using a “lake modeling perspective” in which the effects of groundwater are introduced via boundary conditions without explicitly coupling with groundwater models. We seek to address the following questions as part of this research: (a) How will summer stratification and circulation change if groundwater contribution is ignored? This question is important as groundwater levels are declining in many parts of the world reducing or eliminating groundwater contributions to lakes. (b) Can mesoscale weather prediction model outputs provide sufficiently accurate forcing fields to run hydrodynamic models for relatively small lakes? If so,

how do results compare with those obtained using observations from a network of weather stations? This question is important since many small lakes throughout the world do not have a network of meteorological stations around them. (c) To what extent do alternative turbulence models such as the two-equation  $k-\varepsilon$  model [Rodi, 1987] improve descriptions of vertical mixing and thermal structure in small lakes such as the Gull Lake? The Mellor-Yamada 2.5 level turbulence model (MY2.5, Mellor and Yamada [1982]) is often used to describe vertical mixing in lake and ocean circulation models. The performance of the  $k-\varepsilon$  turbulence model for small, groundwater-fed lakes will be evaluated to identify the best model.

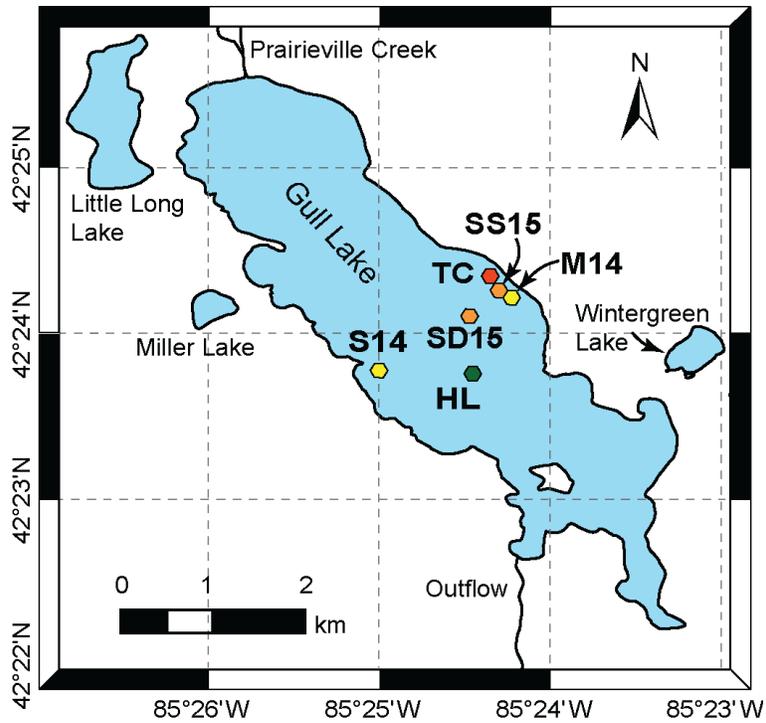
The chapter is structured as follows. Following a description of the study site, we describe data collected in the lake to evaluate the performance of the numerical models. After a description of the hydrodynamic model, we describe how the numerical mesh was refined using a bathymetric map of the lake for a more accurate simulation of thermal structure. Surface heat fluxes are usually the primary sources driving heat transfer in the system; therefore, we describe an approach to assess the accuracy of our station-based meteorological forcing by coupling a mesoscale weather prediction model with our lake model. Model descriptions of vertical turbulent mixing, a key process for transferring heat downward and for the onset of stratification, may not be accurate enough under conditions of strong stratification [Li *et al.*, 2005]. Therefore, we compare the performance of two turbulence models. Changes in internal heating of the water column associated with fluctuations in water clarity and photic zone depth are known to influence stratification and vertical mixing [Chen, 2003a]. Therefore, we describe the use of *in situ* observations of photosynthetically available radiation (PAR) to enhance the performance of the shortwave penetration model in the hydrodynamic model. After these improvements, we describe how

groundwater effects are simulated in the lake model. Finally, an assessment of simulation results is presented along with a discussion and conclusions.

## **3.2 Materials and methods**

### **3.2.1 Observational data**

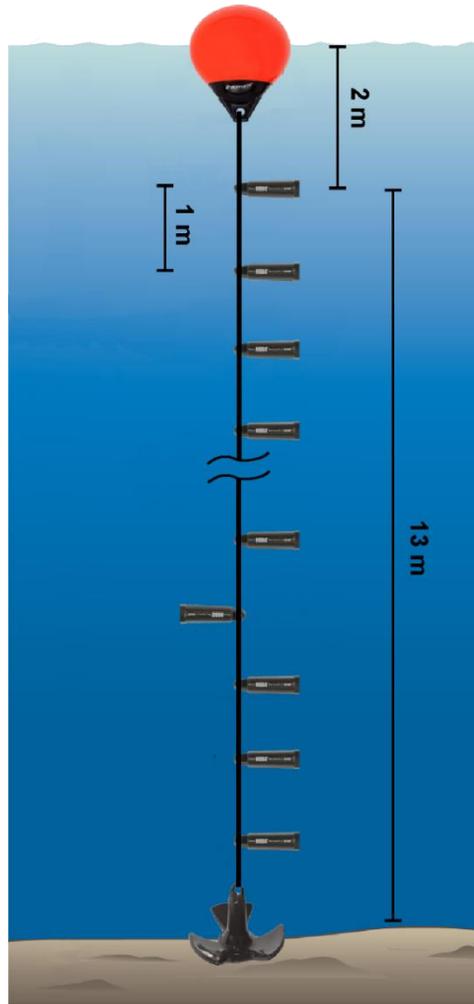
Field observations were made during the summer months of 2014 and 2015 to test the hydrodynamic model. Bathymetry data were collected along longitudinal and transverse transects of the lake using a SonTek RiverSurveyor M9 system, as described in Chapter 2. Figure 3-1 is a map of Gull Lake with locations of these *in situ* measurements marked. In the summer of 2014, we deployed, for the first time in the lake, two upward-looking ADCPs manufactured by Teledyne RD Instruments (at locations marked M14 and S14 in Figure 3-1 and obtained high-resolution current and temperature data. A thermistor chain using Onset HOBO Pro v2 temperature sensors (Onset Computers Inc, Cape Cod, Massachusetts, USA) with an accuracy of 0.2 °C was deployed near the ADCPs (at location marked TC in Figure 3-1) in 15 m of water. This thermistor chain had sensors between 2-13 m depth of waters with a one-meter interval (Figure 3-2). We also had a surface buoy near TC location to record surface water temperature. A Hydrolab multi-parameter sonde (OTT Hydromet, Kempten, Germany) was used to measure the vertical temperature profile at another location (marked HL in Figure 3-1) in 32 m of water. In 2015, two other ADCPs were deployed at locations SS15 and SM15 for 91 days from early June to early September. A SCAMP (Self-Contained Autonomous Micro-Profiler, <http://pme.com>) was used to collect salinity and photosynthetic active radiation (PAR) data at the center of Gull Lake (HL location) during the summers of 2014 and 2015. Further details of instruments deployed in Gull Lake are provided in Table 3-1.



**Figure 3-1** Map of Gull Lake showing locations of the *in situ* measurements.

**Table 3-1** Details of instruments deployed in Gull Lake.

ID	Instrument	Year	Depth	Deployment Location	
				Latitude	Longitude
M14	600 kHz Monitor ADCP	2014	11.6	42.4047	-85.403
S14	1000 kHz Sentinel V20 ADCP	2014	9.45	42.3963	-85.416
SD15	1000 kHz Sentinel V20 ADCP	2015	19	42.402	-85.407
SS15	1000 kHz Sentinel V20 ADCP	2015	10.4	42.4051	-85.404
TC	Thermistor Chain	2014-2015	12.96	42.4054	-85.404
HL	Hydrolab and SCAMP	2014-2015	32	42.3959	-85.407



**Figure 3-2** Schematic of the thermistor chain.

### **3.2.2 Numerical hydrodynamic model**

The mechanistic model of Gull Lake has been developed based on a three-dimensional, unstructured grid, Finite-Volume Community Ocean Model (FVCOM; *Chen et al.* [2003b, 2006]). FVCOM has been successfully applied to all Great Lakes of North America [*Bai et al.*, 2013], including Lake Michigan [*Luo et al.*, 2012; *Rowe et al.*, 2015; *Safaie et al.*, 2016, 2017a], Lake Huron [*Nguyen et al.*, 2014], Lake Superior [*Xue et al.*, 2015], Lake Erie [*Jiang et al.*, 2015; *Niu et al.*, 2015], Lake Ontario [*Shore*, 2009; *Wilson et al.*, 2013] and large rivers such as the St. Clair

River [Anderson and Phanikumar, 2011]. The Reynolds-averaged Navier–Stokes equations of momentum (Eq. 1-3), continuity (Eq. 4), temperature (Eq. 5), and salinity (Eq. 6) are as follows:

$$\frac{\partial \bar{u}}{\partial t} + \bar{u} \frac{\partial \bar{u}}{\partial x} + \bar{v} \frac{\partial \bar{u}}{\partial y} + \bar{w} \frac{\partial \bar{u}}{\partial z} - f\bar{v} = -\frac{1}{\rho_0} \frac{\partial(\bar{P})}{\partial x} + \frac{\partial}{\partial z} \left( K_m \frac{\partial \bar{u}}{\partial z} \right) + F_u \quad (1)$$

$$\frac{\partial \bar{v}}{\partial t} + \bar{u} \frac{\partial \bar{v}}{\partial x} + \bar{v} \frac{\partial \bar{v}}{\partial y} + \bar{w} \frac{\partial \bar{v}}{\partial z} - f\bar{u} = -\frac{1}{\rho_0} \frac{\partial(\bar{P})}{\partial y} + \frac{\partial}{\partial z} \left( K_m \frac{\partial \bar{v}}{\partial z} \right) + F_v \quad (2)$$

$$\frac{\partial \bar{w}}{\partial t} + \bar{u} \frac{\partial \bar{w}}{\partial x} + \bar{v} \frac{\partial \bar{w}}{\partial y} + \bar{w} \frac{\partial \bar{w}}{\partial z} = -\frac{1}{\bar{\rho}} \frac{\partial \bar{P}}{\partial z} - g + \frac{\partial}{\partial z} \left( K_m \frac{\partial \bar{w}}{\partial z} \right) + F_w \quad (3)$$

$$\frac{\partial \bar{u}}{\partial x} + \frac{\partial \bar{v}}{\partial y} + \frac{\partial \bar{w}}{\partial z} = 0 \quad (4)$$

$$\frac{\partial \bar{T}}{\partial t} + \bar{u} \frac{\partial \bar{T}}{\partial x} + \bar{v} \frac{\partial \bar{T}}{\partial y} + \bar{w} \frac{\partial \bar{T}}{\partial z} = \frac{\partial}{\partial z} \left( K_h \frac{\partial \bar{T}}{\partial z} \right) + F_T \quad (5)$$

$$\frac{\partial \bar{S}}{\partial t} + \bar{u} \frac{\partial \bar{S}}{\partial x} + \bar{v} \frac{\partial \bar{S}}{\partial y} + \bar{w} \frac{\partial \bar{S}}{\partial z} = \frac{\partial}{\partial z} \left( K_h \frac{\partial \bar{S}}{\partial z} \right) + F_S \quad (6)$$

where  $\bar{u}$ ,  $\bar{v}$ , and  $\bar{w}$  are mean velocity components in  $x$ ,  $y$ , and  $z$  directions, respectively.  $\bar{T}$  and  $\bar{S}$  are time-averaged temperature and salinity.  $\bar{\rho}$  denotes the mean density and  $\rho_0$  is the reference density.  $\bar{P}$  is the mean pressure.  $f$  is the Coriolis parameter and  $g$  is the gravitational acceleration.  $K_m$  and  $K_h$  are vertical eddy viscosity and thermal vertical eddy diffusion coefficients, respectively.  $F_u$  and  $F_v$  are the horizontal momentum;  $F_w$ ,  $F_T$  and  $F_S$  are vertical momentum, thermal, and salt diffusion terms. By a scaling argument for vertical velocity, the vertical momentum equation (Eq. 3) is reduced to the following hydrostatic equation:

$$0 = -g - \frac{1}{\bar{\rho}} \frac{\partial \bar{P}}{\partial z} \quad (7)$$

$F_u, F_v, F_T$  and  $F_S$  are modeled using the following forms:

$$F_u = \frac{\partial}{\partial x} (2A_m \frac{\partial \bar{u}}{\partial x}) + \frac{\partial}{\partial y} (A_m (\frac{\partial \bar{u}}{\partial y} + \frac{\partial \bar{v}}{\partial x})) \quad (8)$$

$$F_v = \frac{\partial}{\partial y} (2A_m \frac{\partial \bar{v}}{\partial y}) + \frac{\partial}{\partial x} (A_m (\frac{\partial \bar{v}}{\partial x} + \frac{\partial \bar{u}}{\partial y})) \quad (9)$$

$$F_T = \frac{\partial}{\partial x} (A_h \frac{\partial \bar{T}}{\partial x}) + \frac{\partial}{\partial y} (A_h \frac{\partial \bar{T}}{\partial y}) \quad (10)$$

$$F_S = \frac{\partial}{\partial x} (A_h \frac{\partial \bar{S}}{\partial x}) + \frac{\partial}{\partial y} (A_h \frac{\partial \bar{S}}{\partial y}) \quad (11)$$

where  $A_m$  and  $A_h$  are horizontal diffusion coefficients for momentum and scalars, respectively. The governing equations of FVCOM as the default are closed with the modified Mellor - Yamada 2.5 level [Mellor and Yamada, 1982] and Smagorinsky turbulent closure schemes for vertical and horizontal mixing, respectively. A two-equation turbulence model [Rodi, 1987] was utilized as an alternative turbulence closure scheme for vertical mixing in this work. For this purpose, the General Ocean Turbulence Model (GOTM4.0, <http://www.gotm.net>), originally developed by Burchard *et al.* [1998] was coupled to FVCOM. Details of the turbulence closure models and parameter values are described in Section 2.3.

The three-dimensional hydrodynamic equations of the lake were solved using a mode splitting method [Simons, 1974; Madala and Piacseki, 1977]. In this method, the vertically integrated equations (external mode) and the vertical structure equations (internal mode) are solved separately with different time steps. Since the internal waves or the mean currents travel much

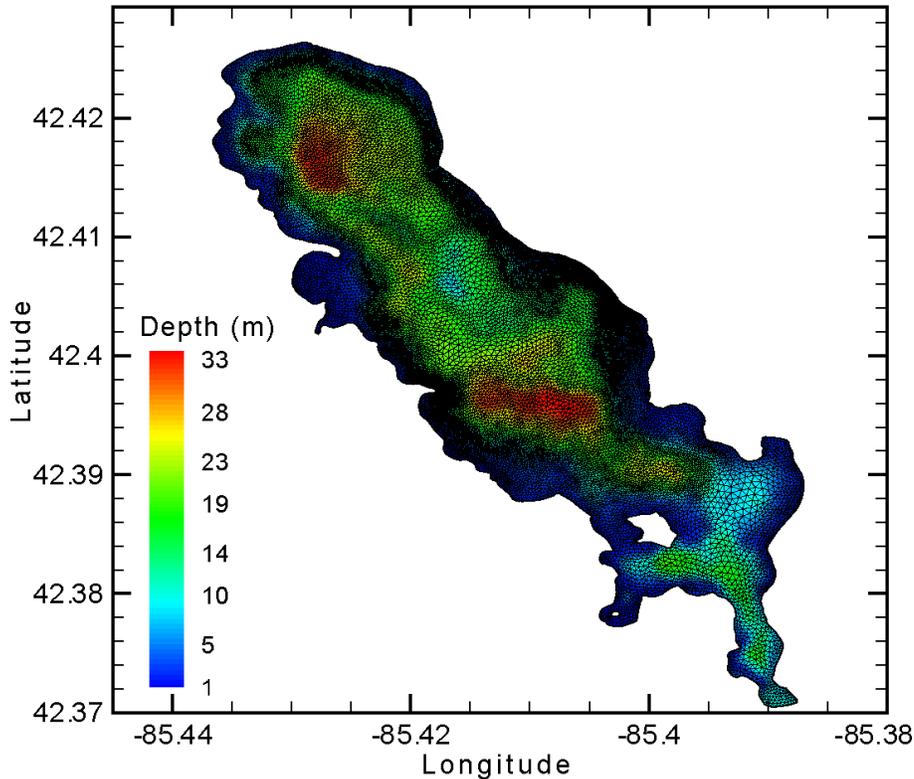
more slowly than the external mode, a larger time step can be used for the internal mode to reduce the computational time. Initial conditions for water temperature in the lake used the measured temperature profiles obtained using the Hydrolab sonde. For salinity we used the SCAMP data. The mean density, which is a function of water temperature, salinity, and pressure [Jackett and McDougall, 1995], was recalculated in pressure coordinates every 30 minutes. This recalculation of the mean density helps by producing more stable stratification.

Average depth of the lake is approximately 12.5 m with a maximum water depth of 33 m. While typical lake-bed slopes in Gull Lake are less than 0.2, the bottom slopes at some points, especially near the deployment sites, are as steep as 0.55. To follow the bottom topography, FVCOM, which is a terrain-following model, uses a vertical  $\sigma$ -coordinate transformation. Although the advantages of the  $\sigma$ -coordinate system are well-known, this coordinate system introduces pressure gradient force errors over steep topography [Mellor *et al.*, 1998]. Previous research showed that bathymetry smoothing, to some extent, would be required to achieve stability and accuracy in simulations with an adequate representation of topography [Barnier *et al.*, 1998; Haidvogel *et al.*, 2000; Martinho and Batteen, 2006; Sikirić *et al.*, 2009]. In Chapter 2, we used a moving average filter with a radius of 100 m to smooth the bathymetry and to reduce artificial noise in simulated currents. However, detailed bathymetry data with fine horizontal and vertical grid resolutions are needed to simulate hydrodynamics and thermal structure accurately. Similar to bottom steepness, vertical and horizontal resolutions, and the stratification strength have impacts on the pressure gradient errors [Haidvogel *et al.*, 2000], and because Gull Lake exhibits strong stratification during summer months, adequate mesh resolution is a key to the success of numerical modeling. One way to avoid pressure gradient errors in  $\sigma$ -coordinates is to generate a high-quality unstructured computational mesh [Gorman *et al.*, 2006]. We used the Surface Water Model

System software (SMS 11.1, <http://www.aquaveo.com>) developed by *Zundel and Jones* [1996], to generate a triangular horizontal mesh. We used the following bathymetry-based algorithm to resolve topographic features and to reduce the pressure gradient errors. First, we created contour lines with one meter elevation interval from the collected bathymetry. Then, we selected contour lines that follow the key topographic features of the lake, and treated them as constraint lines, along which grid nodes must be placed. The unstructured mesh was generated to meet the quality criteria for maximum slope. Elements with slopes greater than 0.2 were refined. A maximum slope of 0.1 is recommended for FVCOM [*Chen et al.*, 2006], although it is preferable to use larger values (such as 0.3) to retain the accuracy of the raw bathymetry data [*Foreman et al.*, 2009]. The computational mesh was also refined where the slope parameter, recommended by *Mellor et al.* [1994], exceeded the limit of 0.2. The slope parameter for each element is defined as:

$$r = \frac{|\Delta h|}{2\bar{h}} \quad (12)$$

where  $\Delta h$  is the maximum depth difference between adjacent grid points, and  $\bar{h}$  is the average of the depths for the two adjacent grid cells. We limited the minimum grid resolution to 5 m to avoid high computational cost and numerical instability issues. Therefore, in regions where either maximum slope or slope parameter conditions needed a finer resolution, the bathymetry was smoothed using the method described in *Mellor et al.* [1998]. A numerical mesh with 20,604 nodes and 40,260 triangular elements in the horizontal with a resolution range of 5-75 m was created following the bathymetry-based refinement algorithm (Figure 3-3). In the vertical, a sigma coordinate system with 30 sigma layers was used.



**Figure 3-3.** The unstructured computational mesh created based on the bathymetry-based refinement algorithm

The hydrodynamic model of Gull Lake based on the meteorological forcing for summer 2014 was run for three different cases. To systematically quantify the improvements in model performance due to the choices in reconstructing meteorological forcing fields, turbulence models and inclusion of groundwater, precipitation and evaporation processes, several cases (described in Table 3-2) were considered. First the model was run with the  $k-\varepsilon$  turbulence model and without the groundwater module (Model 1). Then for the second case, the groundwater module was taken into account in order to examine the effect of groundwater on the physical behavior of the lake (Model 2). This case also was run with the MY2.5 turbulence model (Model 3). Since our primary interest is in evaluating the role of groundwater in circulation and thermal structure of Gull Lake, the coupled WRF-lake model was run by including the contribution from groundwater and with

**Table 3-2.** Description of different models evaluated

Year	Model name	Model description			Remark
		Meteorological Forcing	Turbulence Model	Additional Processes*	
2014	Model1	Weather Stations	k-ε	---	Simulation shows the effect of ignoring groundwater.
	Model2	Weather Stations	k-ε	GW	Together with model 3, simulation provides a comparison of two popular turbulence models for their ability to describe vertical mixing
	Model3	Weather Stations	MY2.5	GW	Together with model 2, simulation provides a comparison of two popular turbulence models for their ability to describe vertical mixing
	Model4	WRF Model	k-ε	GW	Together with model 2 which is based on weather station data, simulation allows a comparison of results based on two different meteorological forcing fields (weather station data versus WRF model outputs)
2015	Model1	Weather Stations	k-ε	---	Same as model 1 but for year 2015
	Model2	Weather Stations	k-ε	GW	Same as model 2 but for year 2015
	Model4	WRF Model	k-ε	GW	Same as model 4 but for year 2015

\*GW = Groundwater

the  $k-\varepsilon$  turbulence model (Model 4). At the end, the developed models were also tested using data obtained during the summer of 2015. Details of the reconstruction of meteorological forcing fields, model of shortwave radiation penetration and the groundwater-precipitation-evaporation module are explained in the following sections.

### 3.2.3 Turbulent closure models

#### 3.2.3.1 Mellor-Yamada 2.5 level turbulence model

The governing equations of FVCOM, including momentum, continuity, temperature, salinity, and density equations, as the default are closed with the modified Mellor and Yamada level 2.5 (MY-2.5; *Mellor and Yamada* [1982]) and Smagorinsky turbulent closure schemes for vertical and horizontal mixing, respectively. The Smagorinsky horizontal diffusion for momentum is given as:

$$A_m = 0.5C\Omega \sqrt{\left(\frac{\partial \bar{u}}{\partial x}\right)^2 + 0.5\left(\frac{\partial \bar{v}}{\partial x} + \frac{\partial \bar{u}}{\partial y}\right)^2 + \left(\frac{\partial \bar{v}}{\partial y}\right)^2} \quad (13)$$

where  $C$  is a constant parameter and  $\Omega$  is the area of the individual momentum control element. Then, the horizontal diffusion coefficients for scalars can be calculated as  $A_m / A_n = \text{Pr}$ , where  $\text{Pr}$  is the turbulent Prandtl number.

MY-2.5 is a two equation model that solves transport equations for variables  $q^2$  (twice of turbulent kinetic energy) and  $q^2l$ , where  $l$  is the turbulence length scale. The set of  $q-ql$  equations in FVCOM is defined as [*Chen et al.*, 2003b]:

$$\frac{\partial q^2}{\partial t} + \bar{u} \frac{\partial q^2}{\partial x} + \bar{v} \frac{\partial q^2}{\partial y} + \bar{w} \frac{\partial q^2}{\partial z} = 2(P_s + P_b - \varepsilon) + \frac{\partial}{\partial z} \left( K_q \frac{\partial q^2}{\partial z} \right) + F_q \quad (14)$$

$$\frac{\partial q^2 l}{\partial t} + \bar{u} \frac{\partial q^2 l}{\partial x} + \bar{v} \frac{\partial q^2 l}{\partial y} + \bar{w} \frac{\partial q^2 l}{\partial z} = l E_1 (P_s + P_b - \frac{\tilde{W}}{E_1} \varepsilon) + \frac{\partial}{\partial z} (K_q \frac{\partial q^2 l}{\partial z}) + F_l \quad (15)$$

where  $\bar{u}$ ,  $\bar{v}$ , and  $\bar{w}$  are components of the mean velocity in  $x$ ,  $y$ , and  $z$  directions.  $F_q$  and  $F_l$  denote the horizontal diffusion of turbulent kinetic energy and turbulence macroscale, respectively.  $K_q$  is the vertical eddy diffusion coefficient and  $\varepsilon = q^3 / B_1 l$  is the dissipation rate of turbulent kinetic energy, where  $B_1$  is a model constant.  $\tilde{W} = 1 + E_2 l^2 / (\kappa L)^2$  is the wall proximity function where  $L^{-1} = (\zeta - z)^{-1} + (H + z)^{-1}$ ,  $\zeta$  is free surface elevation,  $H$  is mean water depth, and  $\kappa$  is the von Kármán constant, equal to 0.4. The constant parameters of the  $q-ql$  model are  $(E_1, E_2) = (1.8, 1.33) \cdot P_s$  and  $P_b$  are shear and buoyancy production terms, respectively:

$$P_s = -\overline{u'w'} \frac{\partial \bar{u}}{\partial z} - \overline{v'w'} \frac{\partial \bar{v}}{\partial z} = K_m [(\frac{\partial \bar{u}}{\partial z})^2 + (\frac{\partial \bar{v}}{\partial z})^2] = K_m M^2 \quad (16)$$

$$P_b = -\frac{g}{\rho_0} \overline{\rho'w'} = -K_h (-\frac{g}{\rho_0} \frac{\partial \bar{\rho}}{\partial z}) = -K_h N^2 \quad (17)$$

where  $u'$ ,  $v'$ , and  $w'$  are turbulent velocity components and  $N$  denotes the Brunt-Väisälä frequency,  $M$  is the shear frequency,  $\bar{\rho}$  is the mean water density,  $\rho_0$  is the reference density, and  $g$  is the acceleration due to gravity. To close the system Eqs. (14) and (15), the vertical eddy viscosity coefficient ( $K_m$ ), the thermal vertical eddy diffusion ( $K_h$ ) and  $K_q$  are modeled as:

$$K_m = lqS_m, K_h = lqS_h, K_q = lqS_q \quad (18)$$

where  $S_m$ ,  $S_h$ , and  $S_q$  are the stability functions originally proposed by *Mellor and Yamada* [1982] and simplified by *Galperin et al.* [1988] as the following [*Allen et al.*, 1995]:

$$S_h = A_2 \frac{1 - 6A_1 B_1^{-1}}{1 - 3A_2(6A_1 + B_2)G_H} \quad (19)$$

$$S_m = A_1 \frac{(1 - 3C_1 - 6A_1 B_1^{-1}) + 9(2A_1 + A_2)S_H G_H}{1 - 9A_1 A_2 G_H} \quad (20)$$

$$S_q = \kappa S_m \quad (21)$$

where  $G_H = -l^2 N^2 / q^2$  restricted to the maximum value of 0.023 for unstable stratification conditions, and the minimum value of -0.28 for conditions of stable stratification [*Galperin et al.*, 1988]. The values of the constant parameters  $A_1$ ,  $B_1$ ,  $A_2$ ,  $B_2$ , and  $C_1$  are 0.92, 16.6, 0.74, 10.1, and 0.08 [*Mellor and Yamada*, 1982], respectively. The surface and bottom boundary conditions for the turbulent kinetic flux are given as:

$$q^2 l = 0, \quad q^2 = B_1^3 u_{\tau s}^2 \quad \text{at } z = \zeta(x, y, t) \quad (22)$$

$$q^2 l = 0, \quad q^2 = B_1^3 u_{\tau b}^2 \quad \text{at } z = -H(x, y, t) \quad (23)$$

where  $u_{\tau s}$  and  $u_{\tau b}$  are surface and bottom friction velocities, respectively.

### 3.2.3.2 The $k - \varepsilon$ turbulence model

The simplified form of the transport equations for the turbulent kinetic energy ( $k$ ) and the rate of dissipation ( $\varepsilon$ ) are given as following:

$$\frac{\partial k}{\partial t} + \bar{u} \frac{\partial k}{\partial x} + \bar{v} \frac{\partial k}{\partial y} + \bar{w} \frac{\partial k}{\partial z} = \frac{\partial}{\partial z} \left( \frac{\nu_t}{\sigma_k} \frac{\partial k}{\partial z} \right) + P + B - \varepsilon \quad (24)$$

$$\frac{\partial \varepsilon}{\partial t} + \bar{u} \frac{\partial \varepsilon}{\partial x} + \bar{v} \frac{\partial \varepsilon}{\partial y} + \bar{w} \frac{\partial \varepsilon}{\partial z} = \frac{\partial}{\partial z} \left( \frac{\nu_t}{\sigma_\varepsilon} \frac{\partial \varepsilon}{\partial z} \right) + \frac{\varepsilon}{k} (c_{\varepsilon 1} P + c_{\varepsilon 3} B - c_{\varepsilon 2} \varepsilon) \quad (25)$$

where  $\sigma_k$  and  $\sigma_\varepsilon$  are constant turbulent Schmidt numbers of  $k$  and  $\varepsilon$ , respectively.  $\nu_t$  denotes the eddy viscosity;  $P$  and  $B$  are the shear and buoyancy production terms, respectively, with the aforementioned definitions in the MY2.5 model. The turbulent diffusivities of momentum and heat can be expressed as:

$$\nu_t = c_\mu k^{1/2} l, \quad \nu_t' = c_\mu' k^{1/2} l \quad (26)$$

where  $c_\mu$  and  $c_\mu'$  are stability functions computed according to the model of *Schumann and Gerz* [1995]:

$$c_\mu = c_\mu^0, \quad c_\mu' = \frac{c_\mu^0}{\text{Pr}_t} \quad (27)$$

where  $c_\mu^0$  is a model constant, and  $\text{Pr}_t$  is the turbulent Prandtl number with the following empirical relation:

$$\text{Pr}_t = \text{Pr}_t^0 \exp\left(-\frac{Ri}{\text{Pr}_t^0 Ri^\infty}\right) - \frac{Ri}{Ri^\infty} \quad (28)$$

where  $Ri$  denotes the gradient Richardson number,  $\text{Pr}_t^0$  is the turbulent Prandtl number for neutral flows ( $Ri \rightarrow 0$ ), and  $Ri^\infty$  is the desired steady-state Richardson number. The values of  $\text{Pr}_t^0$  and  $Ri^\infty$  were set as 4 and 0.25, respectively. The constant parameters of the  $k-\varepsilon$  model are

$c_\mu^0 = 0.5577$ ,  $\sigma_k = 1.0$ ,  $\sigma_\varepsilon = 1.3$ ,  $c_{\varepsilon_1} = 1.44$ , and  $c_{\varepsilon_2} = 1.92$  [Rodi, 1987]. The parameter  $c_{\varepsilon_3}$  can be obtained as a function of a prescribed steady-state Richardson number,  $Ri_{st}$  [Umlauf and Burchard, 2003]:

$$c_{\varepsilon_3} = c_{\varepsilon_2} + (c_{\varepsilon_1} - c_{\varepsilon_2}) \frac{c_\mu}{c_\mu'} \frac{1}{Ri_{st}} \quad (29)$$

### 3.2.4 Light attenuation lengths

To accurately represent the thermal structure, surface temperatures and the depth of thermocline, it is important to correctly model how sunlight penetrates into water. In particular, absorption of downward irradiance in the stratified region has a strong influence on water temperature to the extent that different shortwave penetration models produce different stratification regimes and residual currents [Chen, 2003a]. In FVCOM, the depth distribution of the downward shortwave solar radiation flux,  $SW(x, y, z, t)$ , is calculated as:

$$SW(x, y, z, t) = SW(x, y, \zeta, t) \cdot \left[ R \cdot e^{-z/a} + (1 - R) \cdot e^{-z/b} \right] \quad (30)$$

where  $SW(x, y, z, t)$  is the downward shortwave radiation that penetrates into water as a function of  $x$ ,  $y$ , water depth ( $z$ ), and time ( $t$ ).  $SW(x, y, \zeta, t)$  represents the downward shortwave radiation at the water surface ( $z = \zeta$ ).  $R$  is an empirical constant,  $a$  is a short wavelength due to the near-surface absorption of red spectral components of solar radiation in the upper few meters, and  $b$  is a long wavelength related to the blue-green spectral components in a deeper water depth ( $a < b$ ) [Kraus and Businger, 1995]. Using two distinct sets of attenuation lengths produces a better model performance than a single attenuation wavelength model [Paulson and Simpson, 1977; Simpson and Dickey, 1981b, 1981a; Chen, 2003a]. Paulson and Simpson [1977] determined the values of  $a$ ,  $b$ , and  $R$  corresponding to different water types by fitting Eq. (30) to downward irradiance data.

Chen [2003a] used *in situ* profiles of PAR to re-estimate  $b$ . In this work, we used PAR data collected using the SCAMP at the center of the lake where the depth was 32 m (at HL location). Ensembles of the vertical profiles (in groups of six) were collected weekly during the summers of 2014 and 2015. A nonlinear least squares fitting method was used for determining parameters of the shortwave penetration model. The bi-exponential model (Eq. (30)) was fitted to the observed vertical profiles of PAR to find the parameters  $a$ ,  $b$ , and  $R$ .

### 3.2.5 Precipitation/Evaporation

Hourly precipitation data at the Kellogg Biological Station Longterm Ecological Research Station (KBS-LTER) close to Gull Lake were obtained from Enviro-weather Automated Weather Station Network (<https://mawn.geo.msu.edu>) and used as input of the model. Evaporation of water from the lake was calculated based on the mass-transfer approach using the following equation [Dingman, 2002]:

$$E = K_E \cdot W_{2m} \cdot (e_s - e_a) \quad (31)$$

where  $E$  denotes the evaporation rate ( $\text{m s}^{-1}$ ),  $e_s$  and  $e_a$  are the vapor pressures of the open-water surface and the overlying air, respectively (kPa), and  $W_{2m}$  is the wind speed in  $\text{m s}^{-1}$  at 2 m height. Since the anemometer at the KBS LTER station has a 10-meter height, the power-law correction with an exponent of  $1/7$  [Schwab and Morton, 1984] was applied into the wind speed in order to reduce the wind speed by a factor of  $(2/10)^{1/7}$ .  $K_E$  is a mass-transfer coefficient (kPa) which has the following empirical relation with the lake surface area:

$$K_E = 1.69 \times 10^{-8} \cdot A_L^{-0.05} \quad (32)$$

where  $A_L$  is the lake surface area in  $\text{km}^2$ . The vapor pressure  $e_s$  of an evaporating water surface is equal to the saturation vapor pressure of water,  $e_s^*$ . The vapor pressure in the air was calculate as:

$$e_a = RH \cdot e_a^* \quad (33)$$

where  $RH$  is the relative humidity, and  $e_a^*$  is the saturation air vapor pressure at the air temperature.

The saturation vapor pressures at both air and water surface temperatures can be estimated as:

$$e^* = 0.611 \cdot \exp\left(\frac{17.3 \cdot T}{T + 237.3}\right) \quad (34)$$

where  $e^*$  is in  $\text{kPa}$  and  $T$ , temperature, is in  $^\circ\text{C}$ . When the air temperature is lower than the water temperature, *Rasmussen et al.* [1995] suggested to use the following equation to account for the effect of density instabilities in the air above the water surface:

$$E = \frac{\left[2.33 \cdot T_s - T_a^{1/3} + 2.68 \cdot W_{2m}\right] \cdot e_s - e_a}{\lambda_v}, \quad T_s > T_a \quad T_s > T_a \quad (35)$$

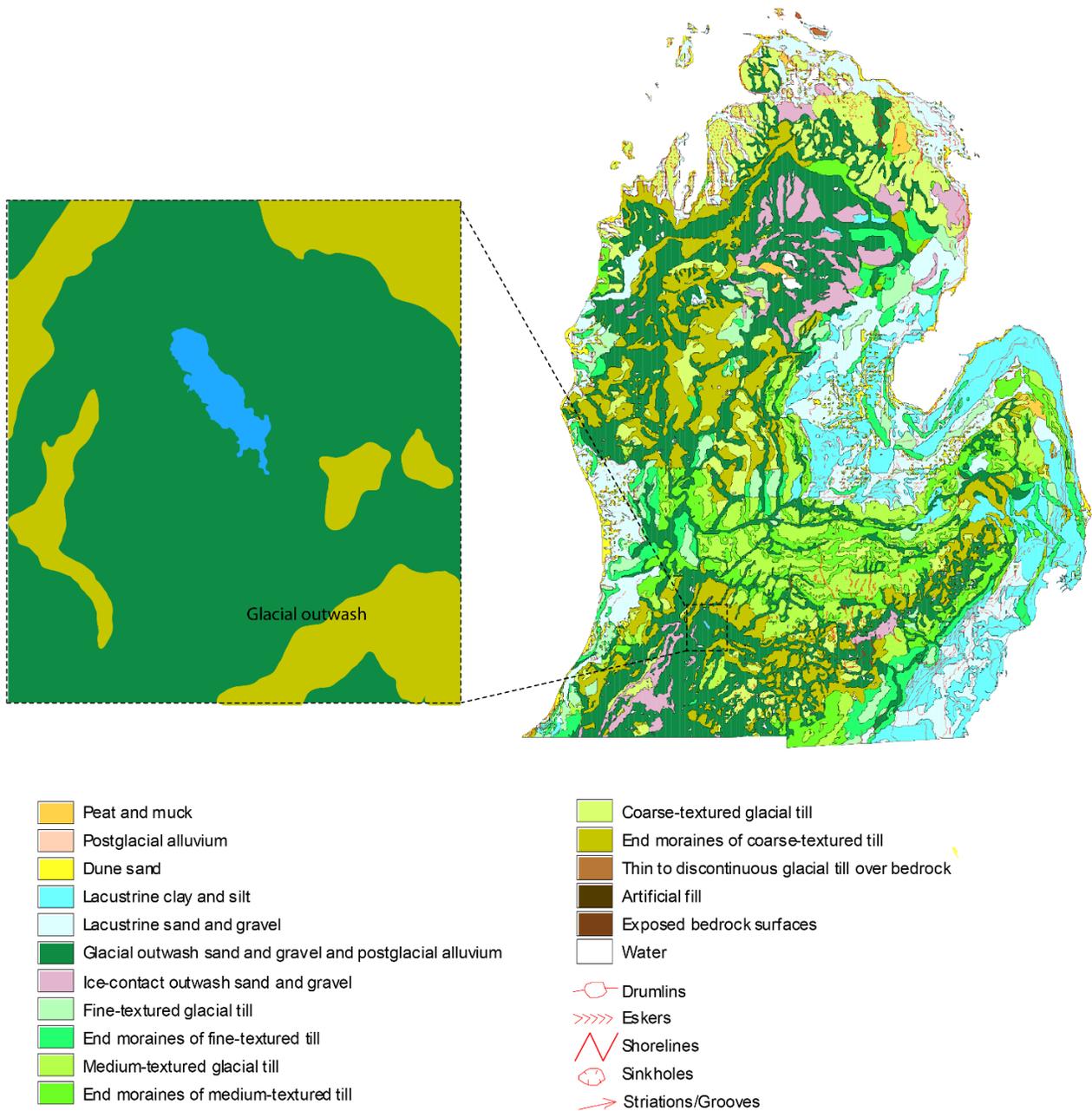
In this equation,  $E$  is evaporation rate in  $\text{mm day}^{-1}$ ,  $W_{2m}$  is wind speed at two meters above ground level ( $\text{m s}^{-1}$ ),  $T_s$  and  $T_a$  are the surface and air temperatures in  $^\circ\text{C}$ , respectively,  $e_s$  and  $e_a$  are the surface and air vapor pressures in  $\text{kPa}$ , and  $\lambda_v$  is the latent heat of evaporation ( $\text{MJ kg}^{-1}$ ) which is given by:

$$\lambda_v = 2.50 - 2.36 \times 10^{-3} \cdot T_s \quad (36)$$

### 3.2.6 Groundwater

Quaternary geology of the Lower Peninsula of Michigan and the Gull Lake area is shown in Figure 3-4. Statewide topographic features are attributed to the glacial processes. The glacial drift

in the lake area is primarily glacial outwash. Predominant outwash materials are medium- to very-coarse sand and gravel [Monaghan *et al.*, 1983]. Some scattered lenses of lacustrine clay are also located in the shallow aquifer [Brewer, 1991].



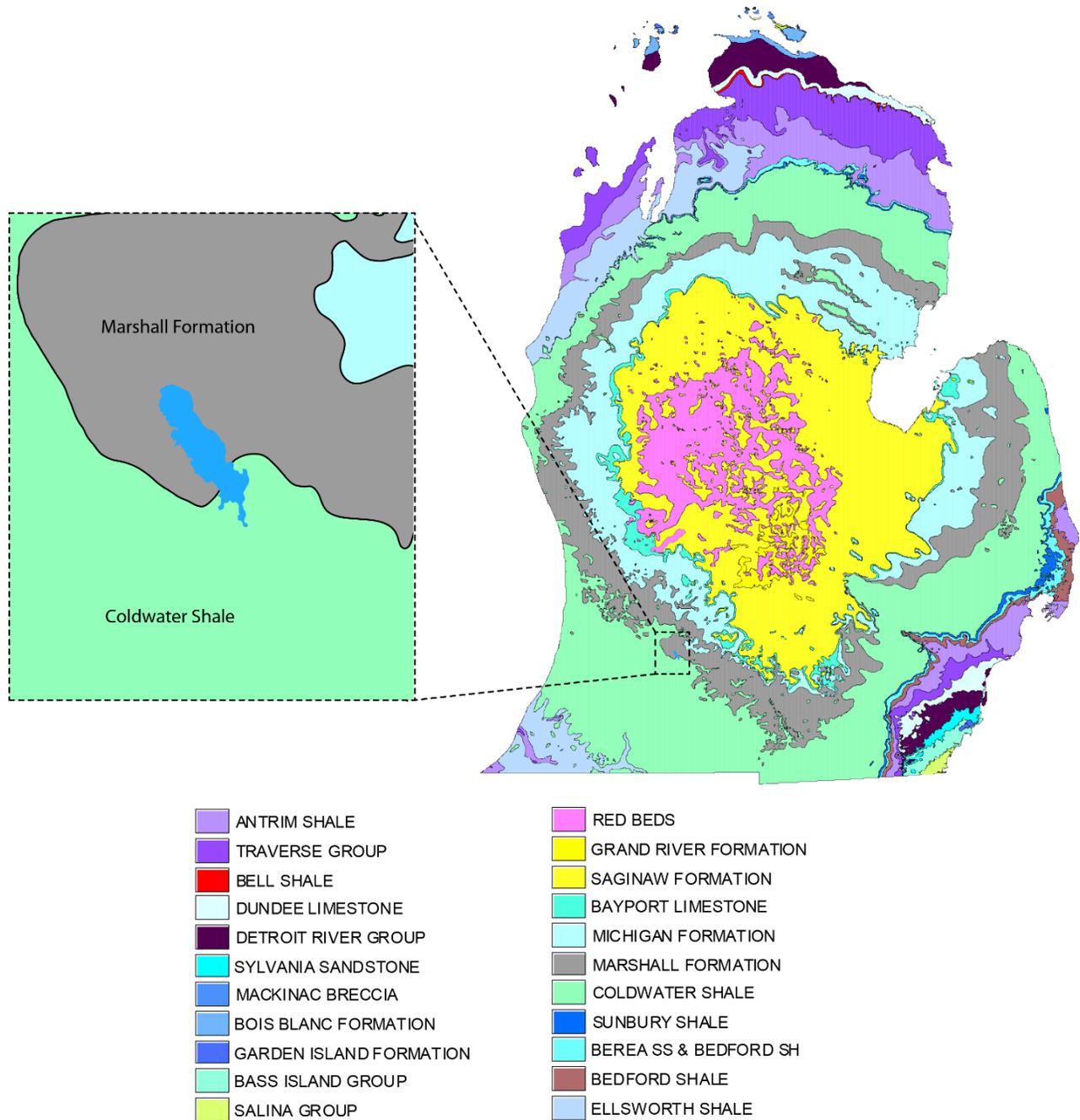
**Figure 3-4.** Quaternary geology of Lower Peninsula and Gull Lake. Figure adapted by author from *Farrand [1982]*

Underlying the shallow aquifer, there are two bedrock formations, including Coldwater shale and Marshall Formation (Figure 3-5). The estimated transmissivity for glacial wells in this region has a range of 116 to 5314 m<sup>2</sup>/day [Apple and Reeves, 2007]. As a result of high transmissivity, groundwater can move relatively quickly through the outwash aquifer, following the hydraulic gradient.

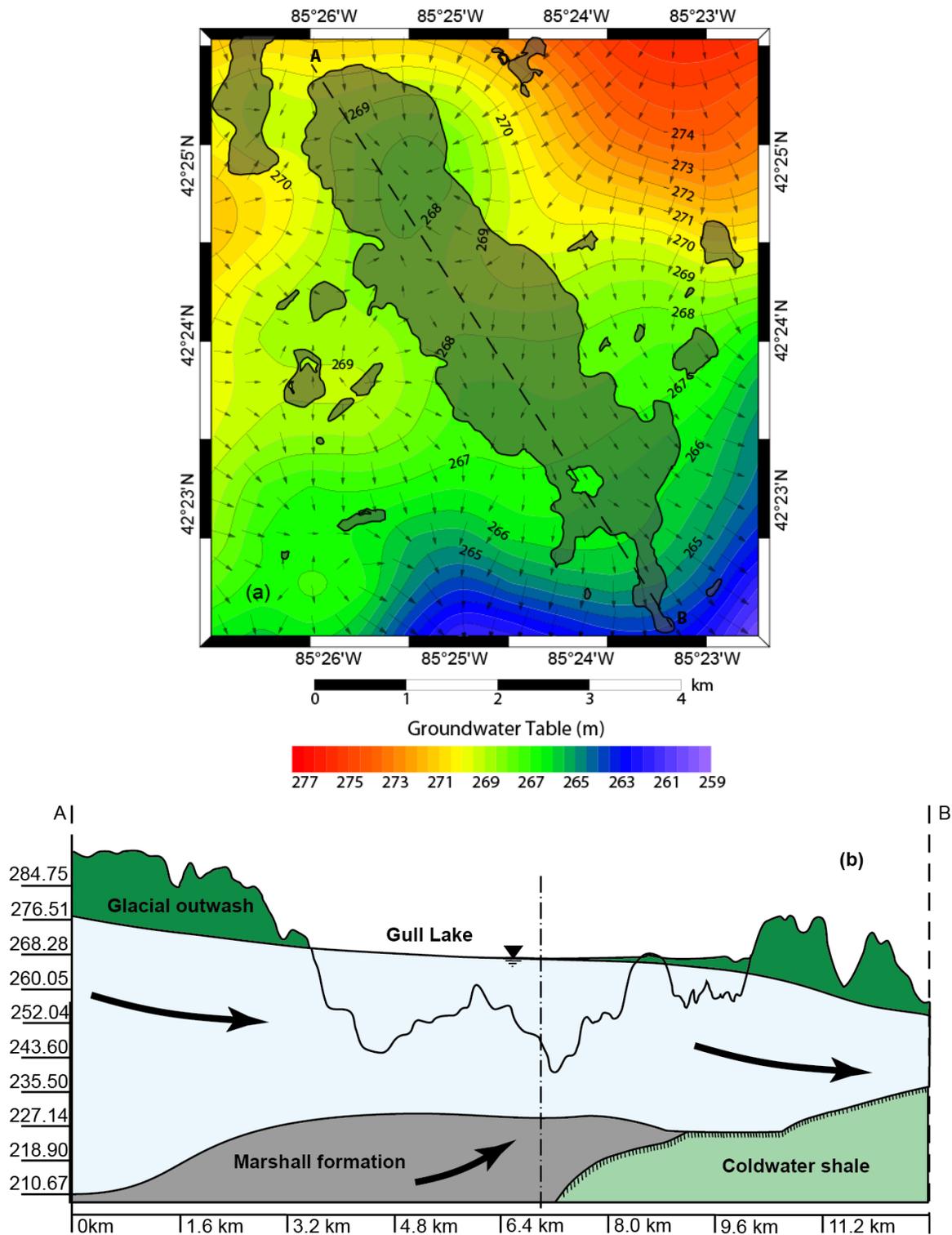
Information about groundwater flow directions and the relative extent of gaining and losing portions of the lake are important inputs to the lake model. The State of Michigan maintains an extensive record of groundwater wells (the Wellogic database) and the data were successfully used to calibrate watershed and groundwater flow models in the past [Niu *et al.*, 2014]. Static water table elevations from the Wellogic database associated with wells from the glacial unconfined aquifer were used to create a groundwater flow map which was used to visualize the groundwater flow patterns through Gull Lake. As expected the water table and the groundwater flow follows the topography in the north-southeast direction (Figure 3-6a). The surface elevation of Gull Lake is about 268 m above sea level. Figure 3-6a shows that where groundwater levels are higher than the lake surface elevation, groundwater discharges into the lake. The groundwater levels and the flow patterns in Figure 3-6a also suggest that the upper half of the lake is groundwater-fed, but the lower part loses water through the lake bed.

A cross-section of the aquifer along the longitudinal transect of the lake is presented in Figure 3-6b (cross section A-B) and overlapped with the lake's bathymetry. Figure 3-5 and 3-6a indicate that the Marshall aquifer pinches out beneath the northern part of the lake. While the Marshall aquifer is composed of permeable sandstones, the Coldwater shale is a confining unit [Apple and Reeves, 2007]. Therefore, Coldwater shale creates an impermeable barrier to horizontal flow, and forces regional groundwater to move upward toward the lake. Furthermore, decreasing

thickness of the shallow aquifer along the groundwater flow directions (Figure 3-6b), combined with high transmissivity and the present of the Coldwater shale, enhances the convergence of groundwater flow into the lake.



**Figure 3-5.** Bedrock Geology of the Lower Peninsula and the Gull Lake area. Figure adapted by author from *Wilson [1987]*



**Figure 3-6.** (a) Groundwater flow map in the study area. (b) A cross-section of the aquifer along the longitudinal transect of the lake.

Gaining and losing portions of the lake can be defined based on Figure 3-6a and 3-6b. Thus, the net groundwater flow is given by:

$$G_{net} = G_{in} - G_{out} = (q_{in} \cdot A_{in} - q_{out} \cdot A_{out}) \quad (37)$$

where  $G_{net}$ ,  $G_{in}$ , and  $G_{out}$  denote the net groundwater flow, groundwater inflow and outflow ( $\text{m}^3/\text{s}$ ), respectively.  $q_{in}$  is the groundwater inflow rate ( $\text{m}/\text{s}$ ) received by the upper part of the lake with a bottom area of  $A_{in}$  ( $\text{m}^2$ ),  $q_{out}$  is groundwater outflow rate ( $\text{m}/\text{s}$ ) associated with the loss of water from the remaining portion of the lake with a bottom area of  $A_{out}$  ( $\text{m}^2$ ). These constant rates were taken into account in the governing equations of FVCOM to simulate groundwater outflow/inflow through the bottom of the lake.

Groundwater fluxes to lakes and wetlands can be measured using seepage meters (e.g., *Mendoza-Sanchez et al.* [2013]). An alternative approach, especially if the interest is in quantifying the contribution of groundwater and other components of the water budget to the entire lake, is to use a water balance approach. However, the water balance approach can only provide an estimation of the net groundwater flow. To quantify contributions of inflow and outflow of groundwater, one approach is to use an additional budget of a chemical tracer [*von Rohden et al.*, 2009; *Langston et al.*, 2013], a stable isotope [*Turner and Townley*, 2006; *Wollschläger et al.*, 2007] or a heat-budget analysis [*Chikita et al.*, 2000; *Kettle et al.*, 2012; *Langston et al.*, 2013]. In this study, we used heat as a groundwater tracer [*Anderson*, 2005; *Constantz*, 2008] to determine the groundwater flow components of Gull Lake. Heat and water budgets of Gull Lake were calculated simultaneously to quantify the lake-groundwater interactions. In this study we followed the second approach to estimate the net contribution of groundwater to the lake. Assuming negligible direct overland runoff during the summer [*Tague*, 1977], the water budget of Gull Lake is given by:

$$\frac{dV}{dt} = P \cdot A_s - E \cdot A_s + G_{net} + Q_{PC} + Q_{LL} + Q_{WL} - Q_{out} \quad (38)$$

where  $dV/dt$  is the rate of change in storage ( $m^3/s$ ), and  $A_s$  is the lake surface area ( $m^2$ ).  $P$  is the precipitation ( $m/s$ ), and  $E$  denotes the evaporation ( $m/s$ ).  $Q_{PC}$ ,  $Q_{LL}$ , and  $Q_{WL}$  are stream flows ( $m^3/s$ ) from Prairieville Creek, Little Long Lake, and Wintergreen Lake into Gull Lake, respectively.  $Q_{out}$  is the discharge ( $m^3/s$ ) from the outlet of Gull Lake.

The energy balance of Gull Lake can be written as:

$$\begin{aligned} V \rho C_p \frac{dT}{dt} = & A_s(SW) + A_s(LW) - A_s H_s - A_s H_L - A_s H_{br} + A_s H_{rain} \\ & + \rho C_p (Q_{PC} T_{PC} + Q_{LW} T_{LW} + Q_{LL} T_{LL} - Q_{out} T_{out}) \\ & + \rho C_p G_{in} T_G - \rho C_p G_{out} T - \rho C_p A_{in} J_G \end{aligned} \quad (39)$$

where  $T$  = water temperature ( $^{\circ}C$ ),  $\rho$  = water density ( $kg/m^3$ ),  $C_p$  = specific heat of water ( $J/kg$   $^{\circ}C$ ),  $V$  = lake water volume ( $m^3$ ),  $A_s$  = lake's surface area ( $m^2$ ),  $SW$  = downward shortwave radiation ( $W/m^2$ ),  $LW$  = downward longwave radiation ( $W/m^2$ ),  $H_s$  = sensible heat flux ( $W/m^2$ ),  $H_L$  = latent heat flux ( $W/m^2$ ),  $H_{br}$  = long wave back radiation from the water ( $W/m^2$ ),  $H_{rain}$  = heat input by rainfall ( $W/m^2$ ).  $T_{PC}$ ,  $T_{LW}$ , and  $T_{LL}$  are stream water temperatures ( $^{\circ}C$ ).  $Q_{out}$  is outflow of Gull Lake ( $m^3/s$ ).  $G_{in}$  and  $G_{out}$  are inflow and outflow of groundwater ( $m^3/s$ ).  $T_G$  denotes the groundwater temperature ( $^{\circ}C$ ).  $J_G$  is the heat flux exchange between lake and groundwater, which was calculated using the Fick's first law:

$$J_G = -D \left. \frac{\partial T}{\partial z} \right|_{z=-H} \quad (40)$$

where  $D$  is the vertical thermal diffusion coefficient ( $\text{m}^2\text{s}^{-1}$ ). To quantify groundwater inflow and outflow to the lake, Eqs. (37), (38), and (39) were solved simultaneously. Observed lake stages were used to calculate  $dV/dt$  in Eq. (38) during the simulation period. Inflows and outflow were measured in 2014, and mean values based on the measurements were used to estimate stream flows in 2015. Observed stream temperatures were used to calculate heat input/output by stream flows. Precipitation was obtained from KBS LTER station.

Evaporation rate in Eq. (38) and the remaining terms in Eq. (39) depend on the spatial and temporal distribution of water temperature in the lake. One approach to determining evaporation and surface heat fluxes is to assume that the lake has homogeneous water surface temperature equal to the observed surface temperature at one location at each time step. Although this method works well for estimation of surface heat fluxes in a one-dimensional model [Stepanenko *et al.*, 2014], it may introduce errors in simulating a three-dimensional temperature field [Abbasi *et al.*, 2016]. Therefore, to simulate lake hydrodynamic and temperature fields by including the groundwater contribution, the groundwater flow and temperature components ( $q_{in}$ ,  $q_{out}$ ,  $T_G$ ) are needed, which in turn, are coupled to lake temperatures via Eqs. (38) and (39). To resolve this and to start the computations we used an iterative approach as described below. As a first step, we calculated time series of evaporation based on observed values of water surface temperature using a mass-transfer method for the entire simulation period replaced. Therefore,  $G_{net}$  was the only unknown in Eq. (38). By substituting the value of  $G_{net}$  into Eq. (37), and assuming that  $q_{in} = q_{out}$  (a reasonable assumption based on the observed water table elevations and flow directions computed using groundwater heads, see Fig. 3-5a), the groundwater flow components were obtained as an initial guess to run the hydrodynamic model. Observed water surface temperatures and measured vertical temperature profiles were used for assessing simulated water temperature

dynamics and heat transfer of the lake. Then, the lake evaporation, sensible heat, latent heat, longwave back radiation, and rain heat fluxes were calculated based on the simulated surface water temperatures using the COARE algorithm [Fairall *et al.*, 1996; Edson *et al.*, 2013] implemented in FVCOM. The rate of change in heat storage ( $V\rho C_p \frac{\partial T}{\partial t}$ ) in Eq. (39) was also determined based on the simulated thermal structure of the lake. Finally, Eqs. (37), (38), and (39) were solved to find the values of  $G_{net}$ ,  $G_{in}$  and  $G_{out}$ . Values of  $q_{in}$ ,  $q_{out}$ ,  $A_m$  and  $A_{out}$  were assumed to be constant over the simulation period. In order to quantify groundwater contribution to the entire lake, water level fluctuations recorded by Acoustic Doppler Current Profilers (ADCPs) were used as a basis for evaluating the water budget of the lake.

Approximate average temperature of shallow groundwater obtained from shallow wells at depths of 9.14 to 18.28 m ranges from 8 °C to 11 °C in Western Michigan [Collins, 1925]. Near the surface, however, the groundwater temperature follows the changes in air temperature. Bottom water temperature at 32 m depth measured using a Hydrolab from 2011 to 2015 and is presented in Figure 3-7. On the basis of this figure when the air temperature goes above the freezing point in early spring, the lake bottom temperature at 32 m was relatively constant compared with the average surface water temperature. In the summer, water temperatures at 32 m depth fall within the range of measured groundwater temperatures.

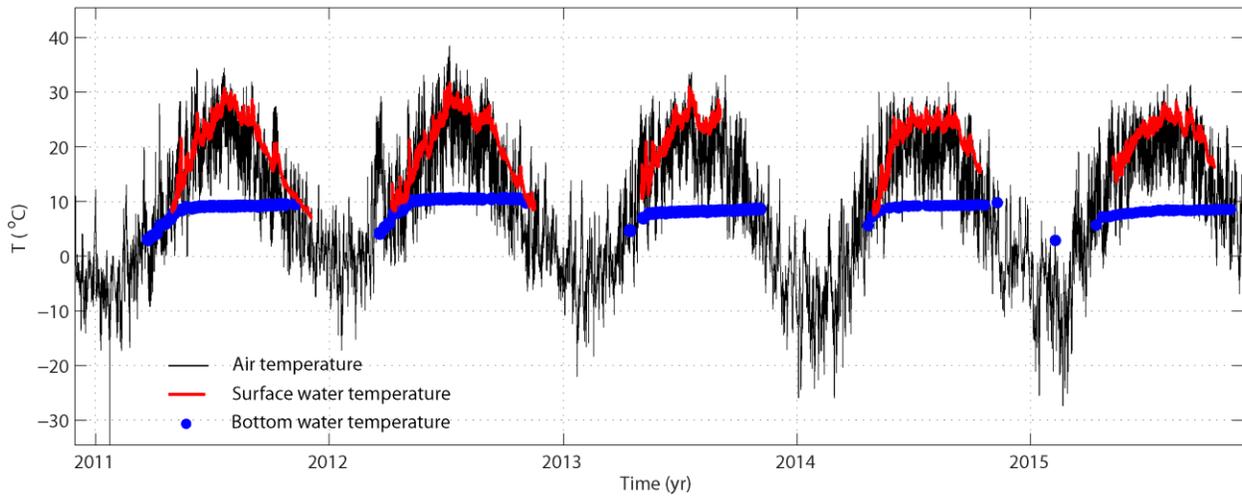
This observation indicates that there may be bottom cooling in the summer due to the groundwater inflow into the lake. To test this hypothesis in our FVCOM model, a benthic layer with a temperature equal to the temperature of groundwater was added to the lake bed. The resulting temperature profile involves a balance between bottom advection and diffusion of water temperature near the lake bed. Benthic groundwater fluxes were prescribed at the bottom by adding

a benthic layer on the bed of the lake. The following equation temperature boundary condition was solved in the benthic layer (at  $z = -H$ ) that received groundwater inflow ( $T = T_G$ ):

$$\frac{\partial T}{\partial t} + q_{in} \frac{\partial T}{\partial z} = D_z \frac{\partial^2 T}{\partial z^2} \quad (41)$$

where  $T$  is the water temperature in °C as a function of depth ( $z$ ) and time ( $t$ ),  $q_{in}$  is the rate of groundwater inflow ( $\text{ms}^{-1}$ ),  $D_z$  is the vertical thermal diffusion coefficient ( $\text{m}^2\text{s}^{-1}$ ), and  $H(x,y)$  is the bottom depth (m). Values of  $D_z$  the parameters in Eq. (40) was estimated to provide the best description of thermal structure. For the losing portion of the lake, thermal discharge from the lake bed was defined as:

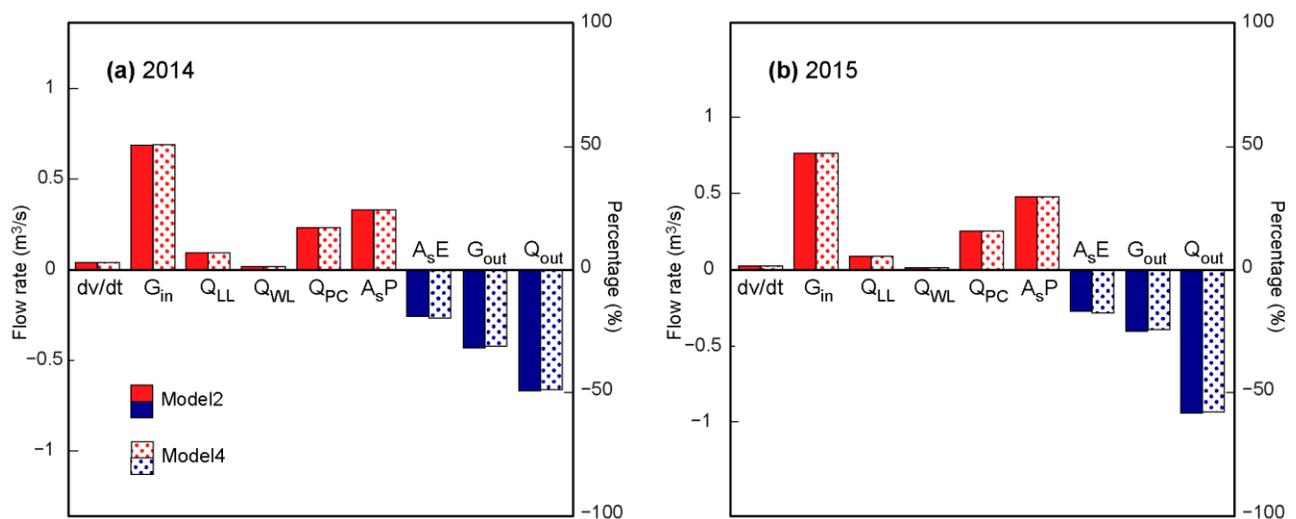
$$\frac{\partial T}{\partial t} - q_{out} \frac{\partial T}{\partial z} = 0 \quad (42)$$



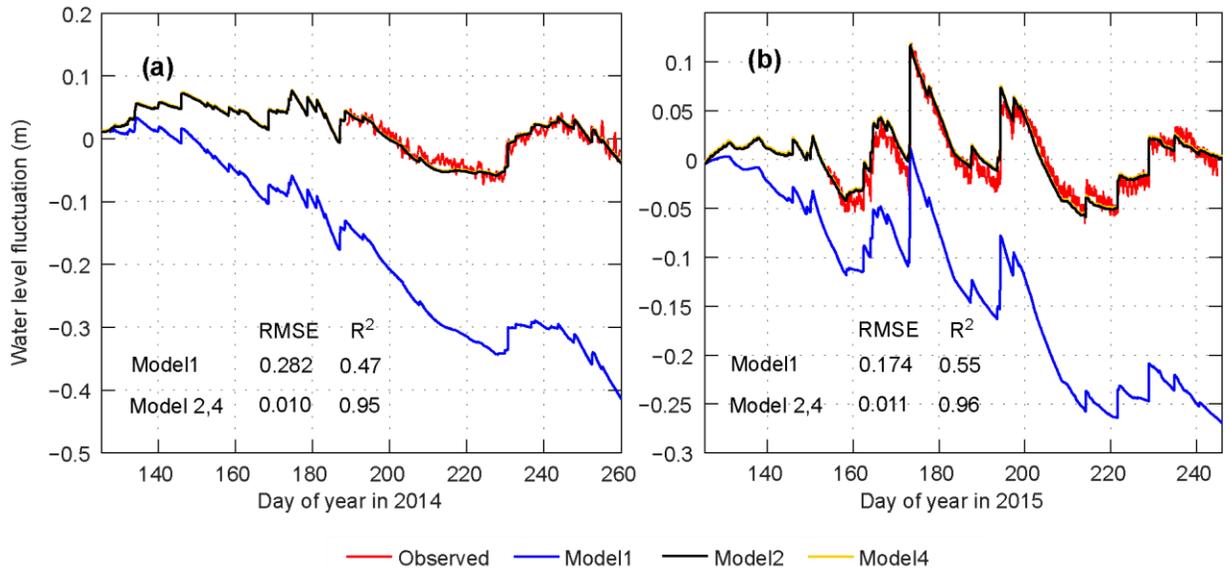
**Figure 3-7.** Time series of relatively constant bottom water temperatures (blue symbols) measured at a depth of 32 m between 2011 and 2015 compared to the surface water temperature and air temperature for the period 2011-2015 (black line).

### 3.3 Results and discussion

Water budget components of Gull Lake are shown in Figure 3-8. Water budget analyses indicated that groundwater contribution to the water budget of the lake was higher than the other components (about 50% percent of total water sources). Sum of inflows into the lake was  $1.17 \times 10^5 \text{ m}^3/\text{day}$  in 2014. The residence time of a lake is defined as the ratio of the reservoir volume and total inflows. Total water volume of Gull Lake is  $1.08 \times 10^8 \text{ m}^3$ . Therefore, the residence time of the lake, which is defined as the ratio of the reservoir volume to the total inflows, is approximately 2.25 yr. However, if there would be no groundwater flow into the lake, then the resident time would be 5.12 yr. To quantify the groundwater contribution to the entire lake, simulated and observed water level fluctuations recorded by the ADCPs are presented in Figure 3-9. As expected, the water budget was simulated much more accurately by considering the contributions of precipitation/evaporation, inflows/outflow and groundwater. The simulated water fluctuations in both 2014 and 2015 followed the trend of the observations. However, the simulated lake levels dropped by 0.3 m in the absence of groundwater.



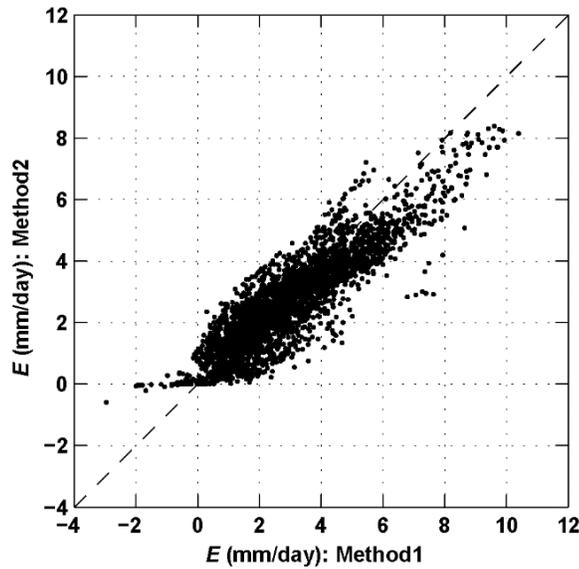
**Figure 3-8.** Water budget components in Gull Lake in (a) 2014 and (b) 2015



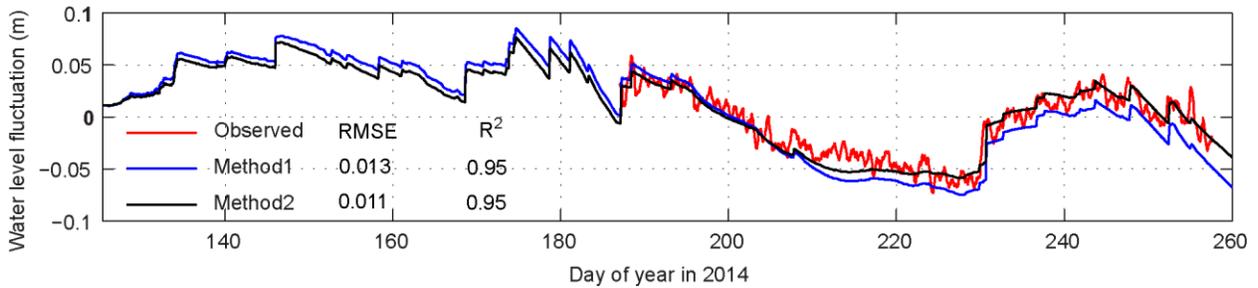
**Figure 3-9.** Comparisons of water level fluctuations recorded by ADCPs (red line), simulated without groundwater (blue line), and simulated with groundwater (black and yellow lines) in (a) 2014 and (b) 2015.

Evaporation rates computed using the mass transfer approach based on the observed water surface temperature were compared with the results of COARE algorithm (Figure 3-10). To compare the mass transfer approach, which was used to as a starting point for water budget analysis, with the COARE algorithm, we ran the model with the same set of parameter values based on the evaporation rates calculated by the two algorithms. The total amount of evaporation computed by the mass transfer approach (373 mm) over the entire simulation period was 8.4 % larger than that calculated using the COARE algorithm (344 mm). Therefore, there was a difference of about 30 mm in the simulated water levels at the end of the simulation (Figure 3-11).

Heat budget components of Gull Lake in the summers of 2015 and 2015 simulated by Model 2 and Model 4 are presented in Table 3-3. Both models produced comparable results. Average longwave solar radiation rates simulated by Model 2 were about 53 (W/m<sup>2</sup>) higher than those simulated by Model 4. On the other hand, shortwave solar radiation rates were lower by 45 (W/m<sup>2</sup>).



**Figure 3-10.** Comparison of Gull Lake evaporation in 2014 computed by the mass transfer approach (method1) and the COARE algorithm (method2)



**Figure 3-11.** Comparisons of simulated water level fluctuations in Gull Lake using evaporation computed by the mass transfer approach (method1) and the COARE algorithm (method2).

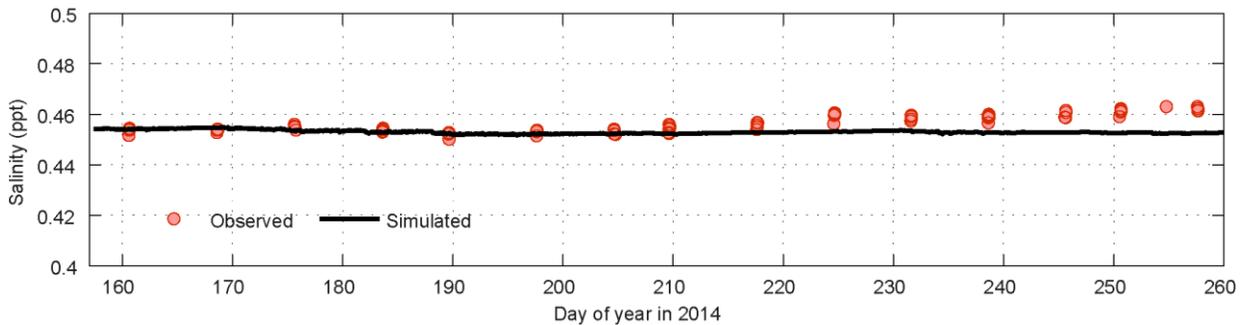
Values of PBIAS for WRF-simulated shortwave (Figure 2-15c) and longwave (Figure 2-15d) radiations also indicate that the simulated shortwave and longwave radiations have overestimation bias (PBIAS=-20%), and underestimation bias (PBIAS=2%), respectively. However, differences between net surface heat rates computed by Model 2 and 4 are as low as 3.5 (W/m<sup>2</sup>). The groundwater heat flux components were found to be the most important heat loss terms after the longwave back radiation from the lake [Kettle *et al.*, 2012]. Due to the summer cooling effect induced by lake - groundwater interaction, the total heat storage of Gull Lake over

the entire simulation period was only one-third of the total surface heat fluxes. This result highlights the ability of the lake to buffer seasonal water temperature variations in the hypolimnion.

**Table 3-3.** Averaged heat budget components of Gull Lake in the summers of 2014 and 2015.

Heat rate ( $\text{W}/\text{m}^2$ )	2014		2015	
	Model 2	Model 4	Model 2	Model 4
Net surface heat	117.52	120.82	115.68	119.16
$LW$	391.18	347.28	393.84	340.74
$SW$	234.37	277.68	225.20	275.15
$H_S$	-16.72	-9.55	-14.23	-9.98
$H_L$	-71.79	-79.40	-74.75	-82.55
$H_{br}$	-419.52	-415.20	-414.38	-404.21
$H_{PC}$	2.00	2.00	2.82	2.82
$H_{WL}$	0.15	0.15	0.18	0.18
$H_{LL}$	1.01	1.01	1.01	1.01
$H_{out}$	-7.05	-7.05	-9.30	-9.30
$H_{Gin}$	2.13	1.97	3.07	2.99
$H_{Gout}$	-3.51	-3.47	-3.33	-3.29
Heat storage	-33.57	-34.96	-39.34	-38.04
$J_G$	-79.17	-80.97	-71.36	-76.11
$H_{rain}$	0.50	0.50	0.58	0.58

Changes in density due to changes in salinity and temperature were simulated by the model (Figure 3-12). However, the small changes in density did not produce any appreciable changes in simulated circulation and thermal structure due to small salinity values.



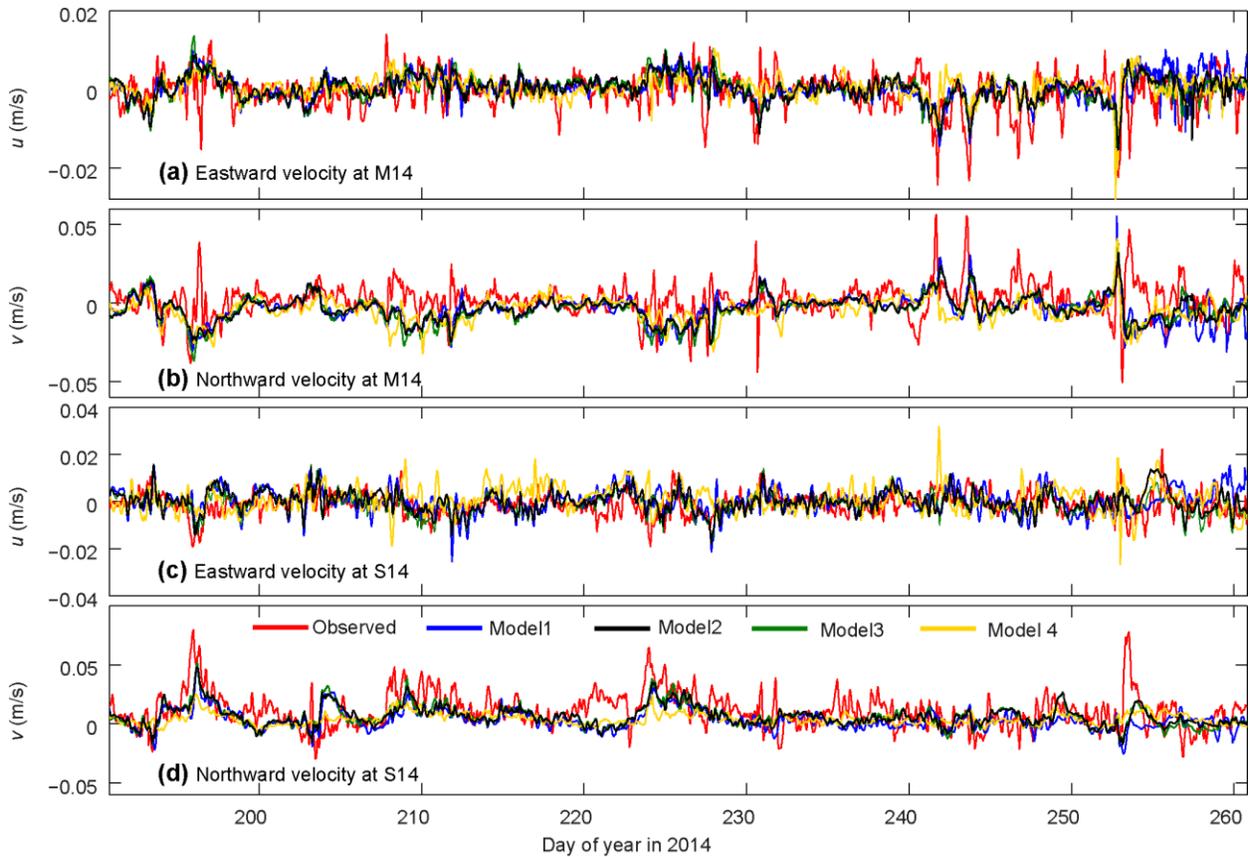
**Figure 3-12.** Comparison of simulated and observed vertically averaged salinity at HL location for year 2014.

Comparisons of observed and simulated vertically-averaged velocity profiles at M14 and S14 locations for Models 1-4 are presented in Figures 3-13. RMSE values (m/s) of eastward ( $u$ ) and northward ( $v$ ) velocities at these locations are presented in Table 3-4. The results show that simulated current velocities at S14 and M14 locations are in good agreement with observations. Overall, all four models produced almost similar performance based on the vertically-averaged velocities. However, the results of the hydrodynamic models based on the k- $\epsilon$  turbulence model are slightly better than those based on the MY2.5 turbulence model.

**Table 3-4.** RMSE values (m/s) of eastward and northward velocities at ADCP locations.

Year	Model	M14		S14	
		RMSE <sub>u</sub>	RMSE <sub>v</sub>	RMSE <sub>u</sub>	RMSE <sub>v</sub>
2014	Model 1	0.0052	0.0137	0.0063	0.0164
	Model 2	0.0052	0.0138	0.0055	0.0148
	Model 3	0.0052	0.0146	0.0052	0.0147
	Model 4	0.0052	0.0139	0.0069	0.0146
2015		SD15		SS15	
	Model1	0.0052	0.0070	0.0060	0.0116
	Model 2	0.0052	0.0068	0.0059	0.0115
	Model 4	0.0053	0.0076	0.0063	0.0118

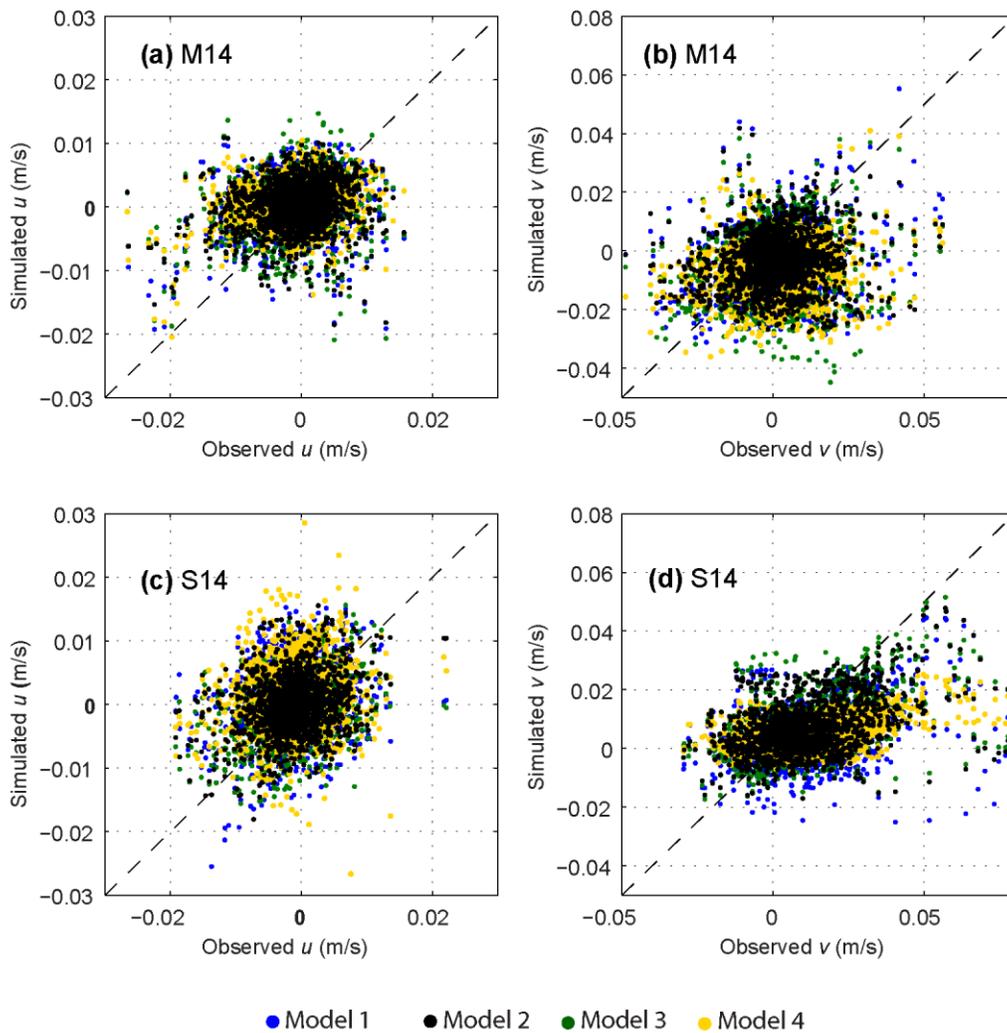
Comparisons of the simulated vertically-averaged velocities with observations (Figures 3-13) show that the reconstructed forcing fields obtained from the WRF model did just as in reproducing the hydrodynamics as the forcing fields based on weather stations data alone. Scatter plots of the simulated vertically-averaged velocities versus observed data at M14 and S14 are shown in Figure 3-14. RMSE values of these comparisons, presented in Table 3-4, show that Models 2 and 4 produced similar performance at the M14 location.



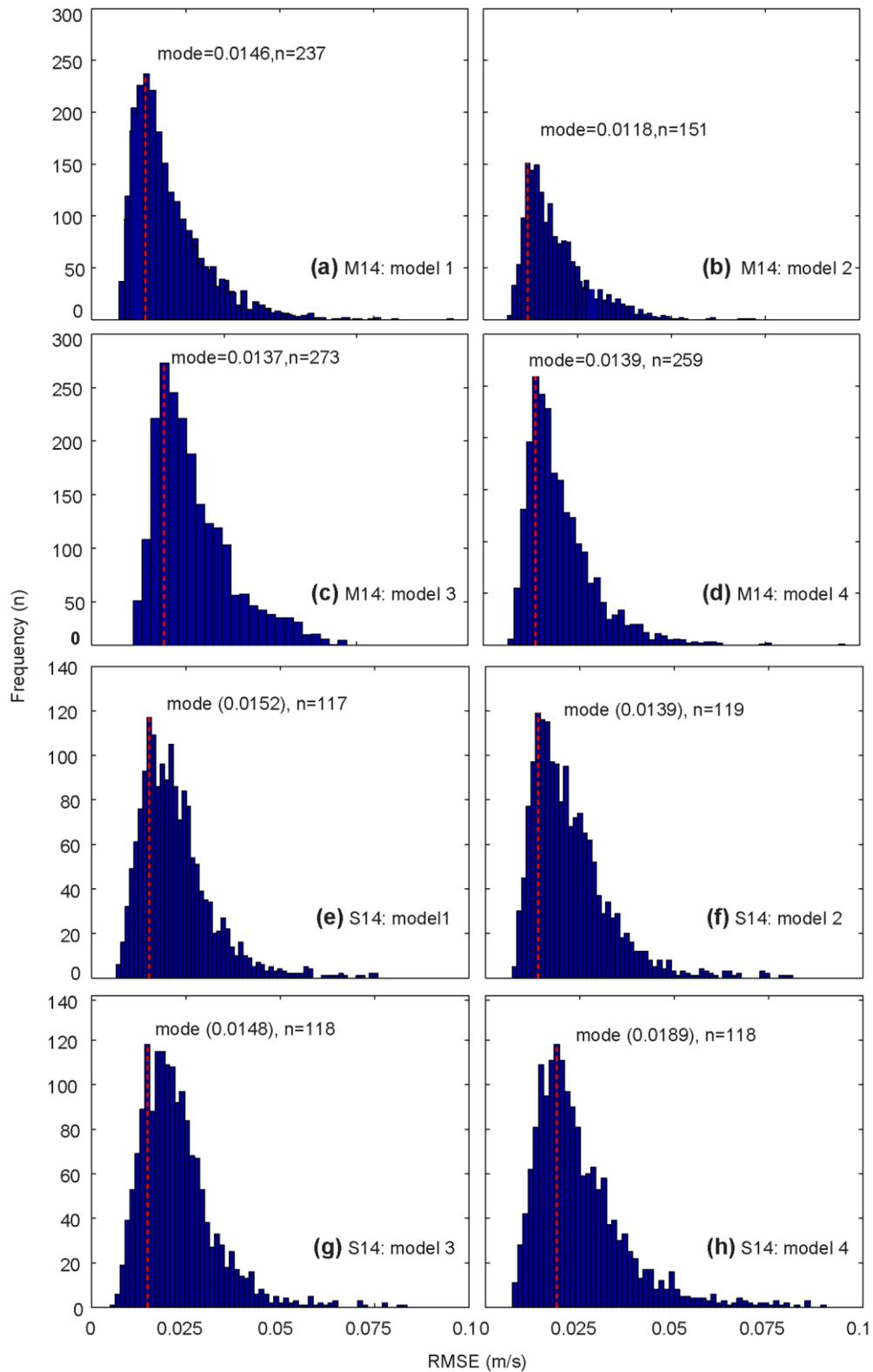
**Figure 3-13.** Comparisons of observed and simulated vertically-averaged velocity profiles at ADCP locations in the summer of 2014.

Since comparisons of vertically-integrated horizontal currents mask important processes within the water column, we assessed the effectiveness of the models for their ability to accurately simulate the vertical velocity profiles at the ADCP locations using RMSE values of observed and simulated vertical velocity profiles for all time steps during the simulation. The probability distributions of RMSE values at the M14 location (Figure 3-15a-d) indicated that Model 2 (mode=0.0118 m s<sup>-1</sup>, n=151) produced the best performance (n=number of time steps with the most frequent RMSE), followed by Model 3 (mode=0.0137 m s<sup>-1</sup>, n=273), Model 4 (mode=0.0139 m s<sup>-1</sup>, n=259), and, lastly, Model 1 (mode=0.0146 m s<sup>-1</sup>, n=237). At the S14 location also Model 2 (mode=0.0139 m s<sup>-1</sup>, n=119) followed by the Model 3 (mode=0.0148 m s<sup>-1</sup>, n=118) performed better than Model 1 (mode=0.0152 m s<sup>-1</sup>, n=117). However, the relative performance of Model 4

(mode=0.0189 m s<sup>-1</sup>, n=118) in simulating the vertical profiles was somewhat inferior at this location compared to the other models (Figure 3-15e-h).

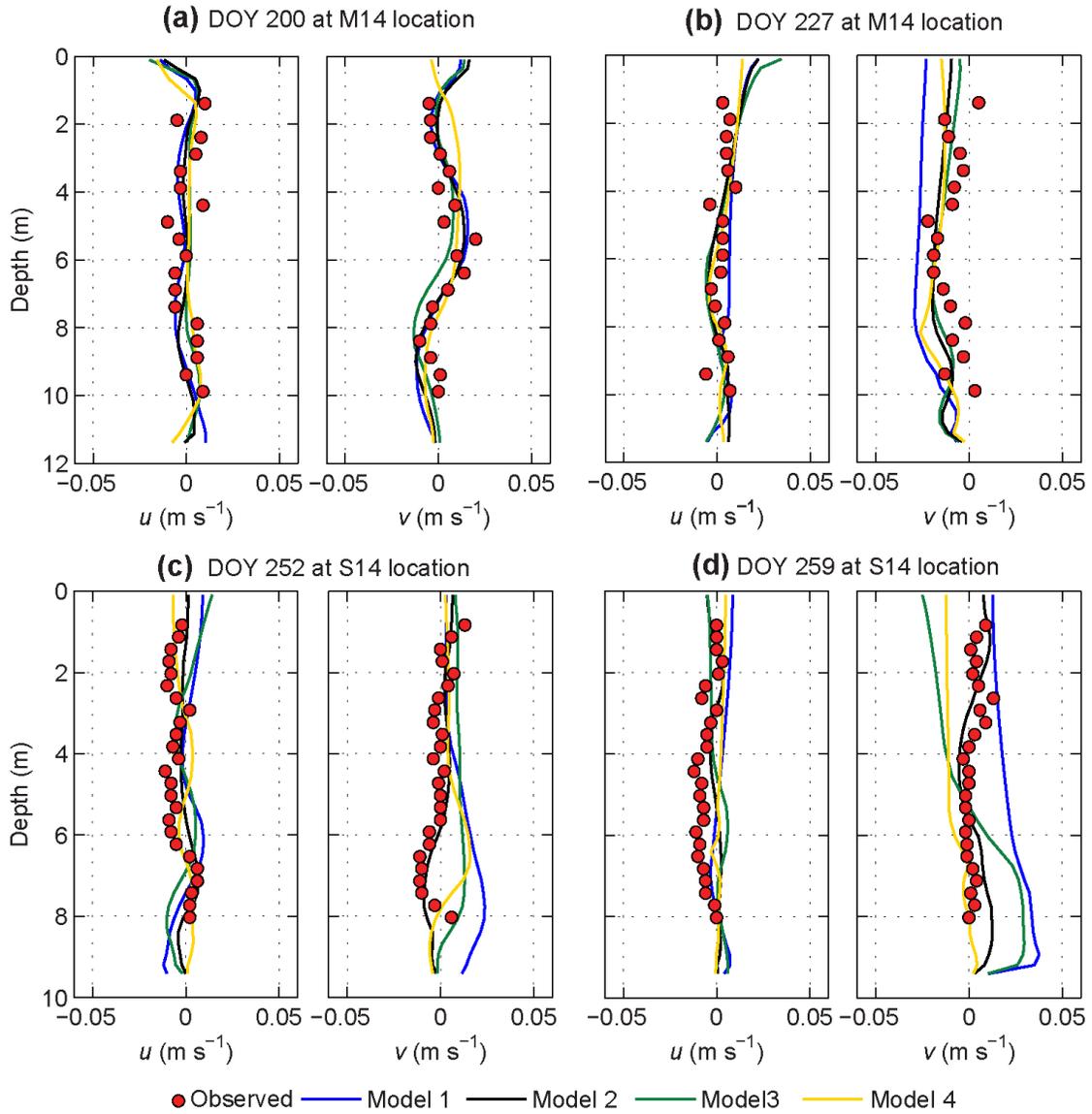


**Figure 3-14.** Scatter plots of velocities (observed versus simulated for different models) for year 2014.



**Figure 3-15.** Probability distributions of RMSE values based on observed and simulated vertical velocity profiles

In figure 3-16, comparisons of simulated and observed vertical velocity profiles at the ADCP locations are shown for a few selected time steps. RMSE values of these comparisons are presented in Table 3-5.



**Figure 3-16.** Examples of comparisons of the observed and simulated vertical velocity profiles at ADCP locations in the summer of 2014.

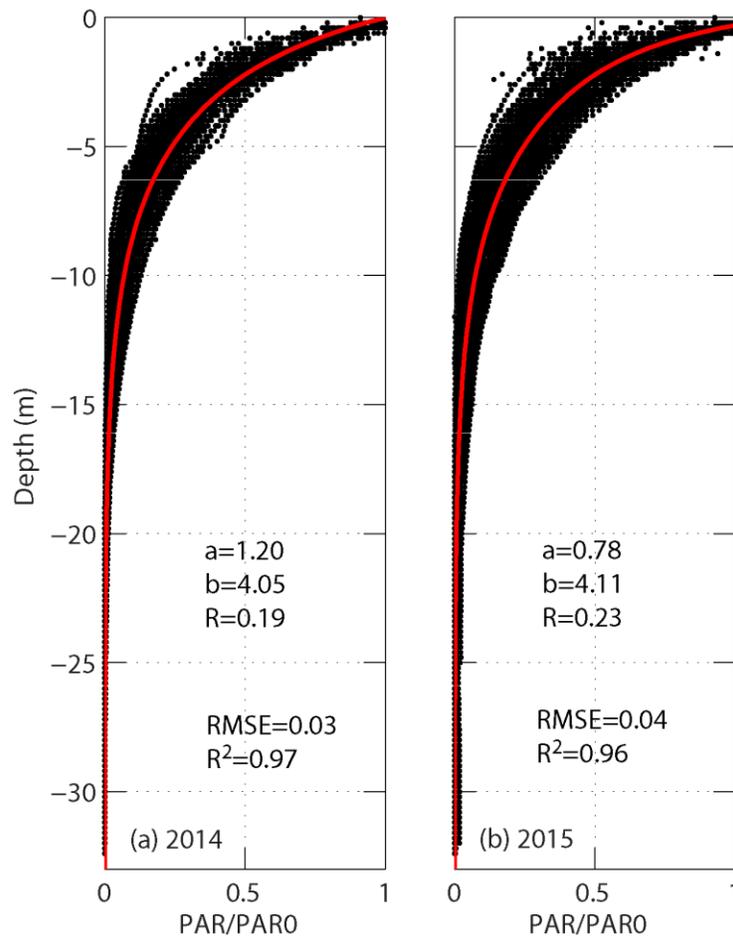
The simulated vertical velocity profiles for Model 2 at these time steps provided the best agreement with observations. Figures 3-16b, 3-16c, and 3-16d indicate that the model without the groundwater contribution (Model1) sometimes completely failed to reproduce the observed vertical velocity profiles.

**Table 3-5.** RMSE values (m/s) of the observed and simulated vertical velocity profiles at ADCP locations for some selected time steps.

	S14				M14			
	(a) Day 250		(b) Day 259		(c) Day 200		(d) Day 227	
	RMSEu	RMSEv	RMSEu	RMSEv	RMSEu	RMSEv	RMSEu	RMSEv
Model1	0.0099	0.0169	0.0083	0.0202	0.0067	0.0063	0.0060	0.0173
Model2	0.0054	0.0051	0.0069	0.0066	0.0066	0.0058	0.0055	0.0086
Model3	0.0095	0.0135	0.0084	0.0199	0.0060	0.0070	0.0058	0.0072
Model4	0.0058	0.0124	0.0079	0.0126	0.0060	0.0077	0.0050	0.0103
	SD15				SS15			
	(a)Day 242		(b) Day 244		(c) Day 156		(d) Day 234	
	RMSEu	RMSEv	RMSEu	RMSEv	RMSEu	RMSEv	RMSEu	RMSEv
Model2	0.0080	0.0067	0.0048	0.0082	0.0094	0.0083	0.0054	0.0067
Model4	0.0086	0.0085	0.0096	0.0080	0.0133	0.0183	0.0052	0.0068

For the shortwave radiation model (Eq. 30), the parameters  $R=0.81$ ,  $a=1.2$  m, and  $b= 4.05$  m provided the best fit to the measured PAR data in 2014 with an averaged regression factor of  $R^2=0.87$  (Figure 3-17a). For 2015 simulations, we used the parameters of  $R=0.77$ ,  $a=0.78$  m, and  $b=4.11$  m based on PAR data ( $R^2=0.96$ ) obtained in the summer of 2015 (Figure 3-17b). Use this shortwave radiation model improved the simulation of surface temperatures and depth of thermocline. However, further improvement would be expected by utilizing time-dependent

parameters. An optimal set of parameters for the groundwater module was identified by comparing the simulated temperature with data collected by ADCPs, thermistor chains, and Hydrolab (Table 3-6). The groundwater flow components of Gull Lake estimated using the method described in an earlier section (3.2.6) are shown in this table. The magnitude of groundwater flow rates computed in this study are within the same range of values (0.14–90 mm/day) determined by other works (Table 4 in *Kettle et al.* [2012]). For instance, groundwater exchange rates of 8.7-14.7 mm/day were reported for Williams Lake in Wisconsin, USA, which is also located in the glacial outwash region in northern USA [*Schuster et al.*, 2003].

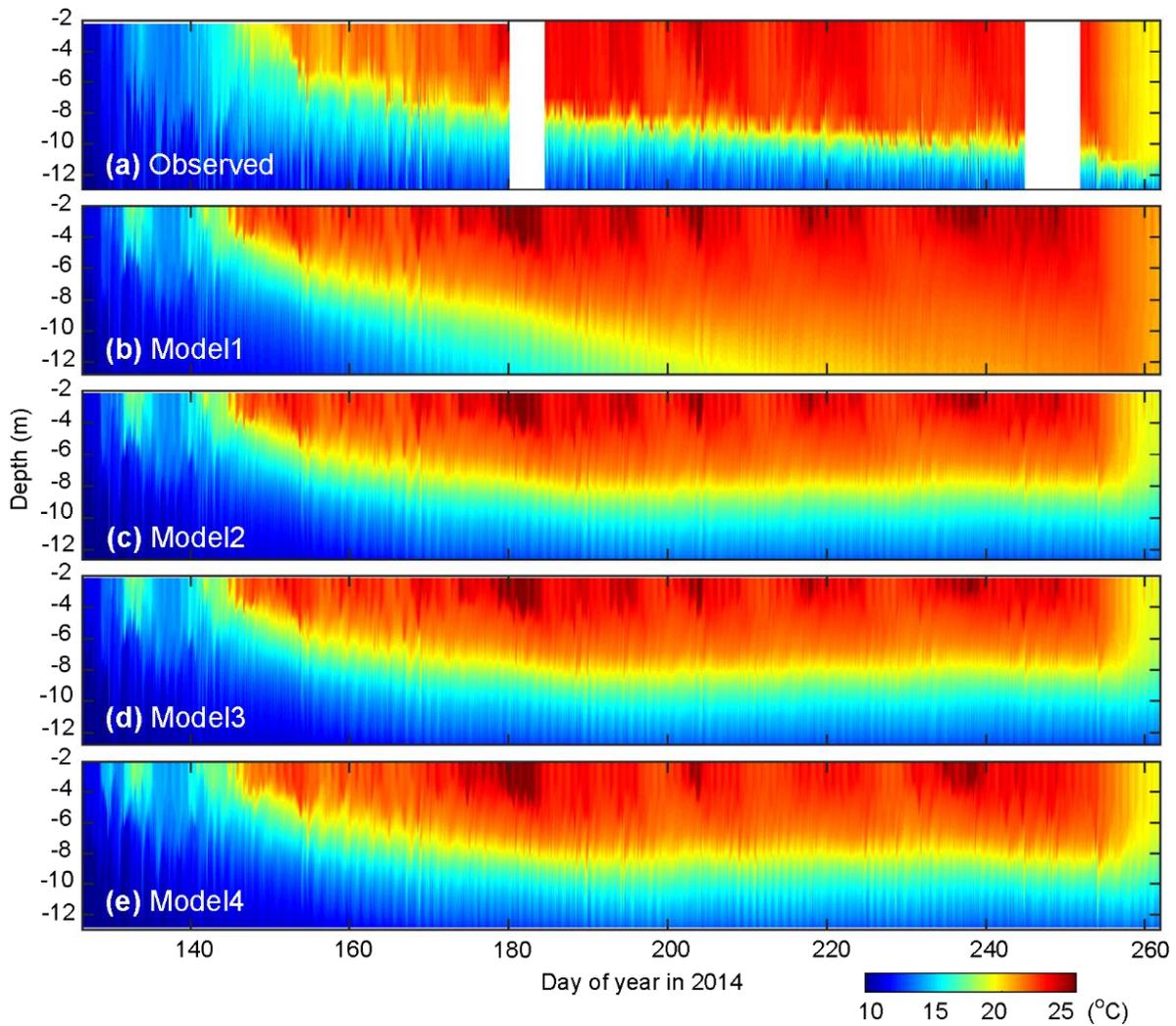


**Figure 3-17.** Observed PAR data used to determine the parameters of the shortwave penetration model.

**Table 3-6.** Set of parameters used for the groundwater module.

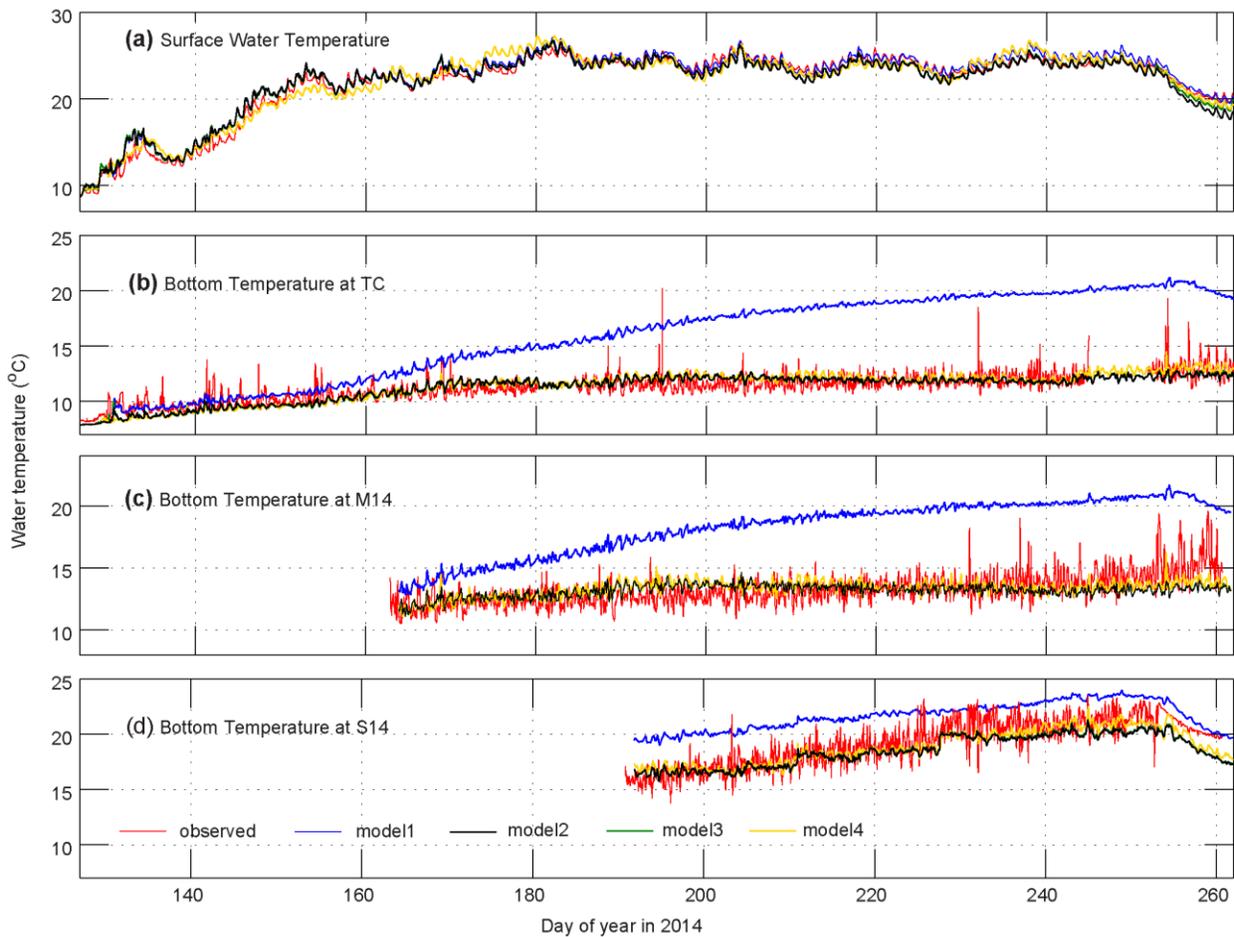
Year	Model	$T_G$ (°C)	$D_z$ (m <sup>2</sup> /day)	$A_{in}$ (km <sup>2</sup> )	$A_{out}$ (km <sup>2</sup> )	$G_{in}$ (m <sup>3</sup> /s)	$G_{out}$ (m <sup>3</sup> /s)	$q_{in}$ (mm/day)	$q_{out}$ (mm/day)
2014	Model2	9	1	5.09	3.45	0.687	0.418	11.66	10.47
	Model4								
2015	Model2	8	1	5.09	3.45	0.769	0.408	13.06	10.23
	Model4								

Contour plots of observed and simulated water temperatures in year 2014 at location TC are presented in Figure 3-18 for different models.



**Figure 3-18.** Contour plots of the observed and simulated water temperature at the TC location in the summer of 2014.

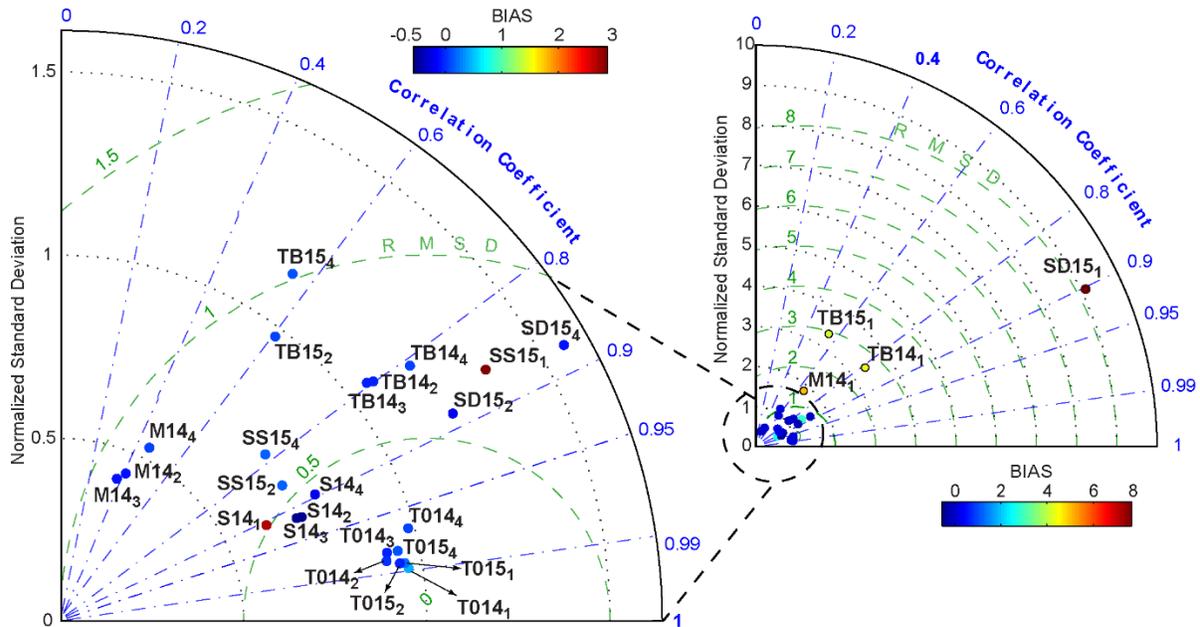
Without the groundwater contribution, both the location and thickness of thermocline were not captured accurately by the model. After about 15 days of simulation, the bottom temperature of the lake became too high relative to observations. This can be clearly seen by comparing simulated and observed time series of the water temperature at TC, M14, and S14 locations (Figure 3-19).



**Figure 3-19.** Comparisons of simulated and observed time series of (a) surface water temperature ( $T_0$ ), and bottom temperature at (b) TC, (c) M14, and (d) S14 locations in 2014.

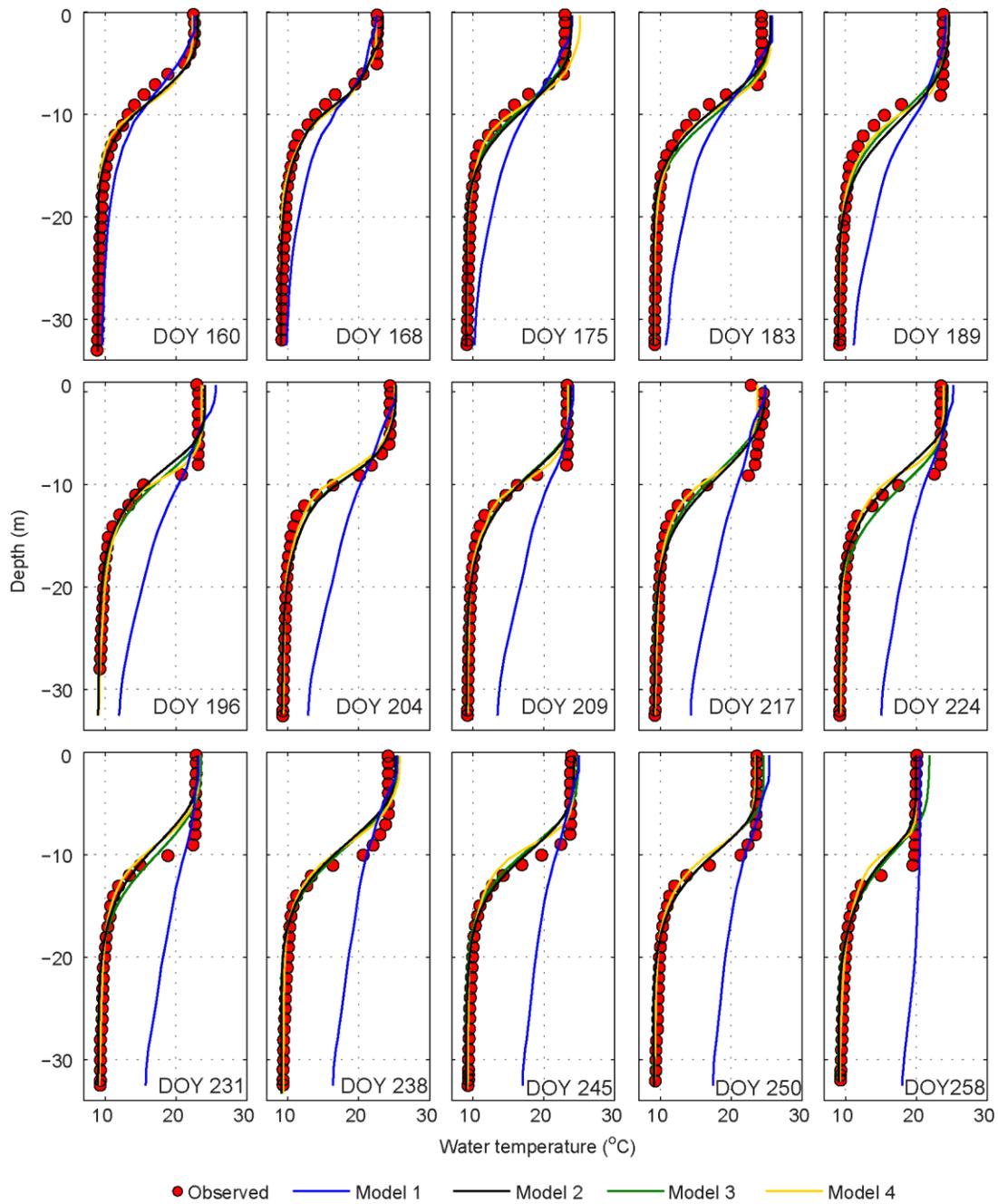
Statistical comparisons of observed water temperature with simulated results of models 1 to 4 were presented in a Taylor diagram in Figure 3-20. The statistics are normalized by the standard deviation of the observations. Definitions of metrics used in the Taylor diagram are provided in the Appendix. The cold layer below the thermocline, however, was captured by all other models

after including the contribution of groundwater. The coupled WRF-lake model and the FVCOM model with a forcing based on the weather stations alone produced almost comparable performance in simulating the thermal structure.



**Figure 3-20.** Taylor diagram which shows statistical comparisons of observed water temperature with simulated results of model 1 to 4. The statistics are normalized by the standard deviation of the observations

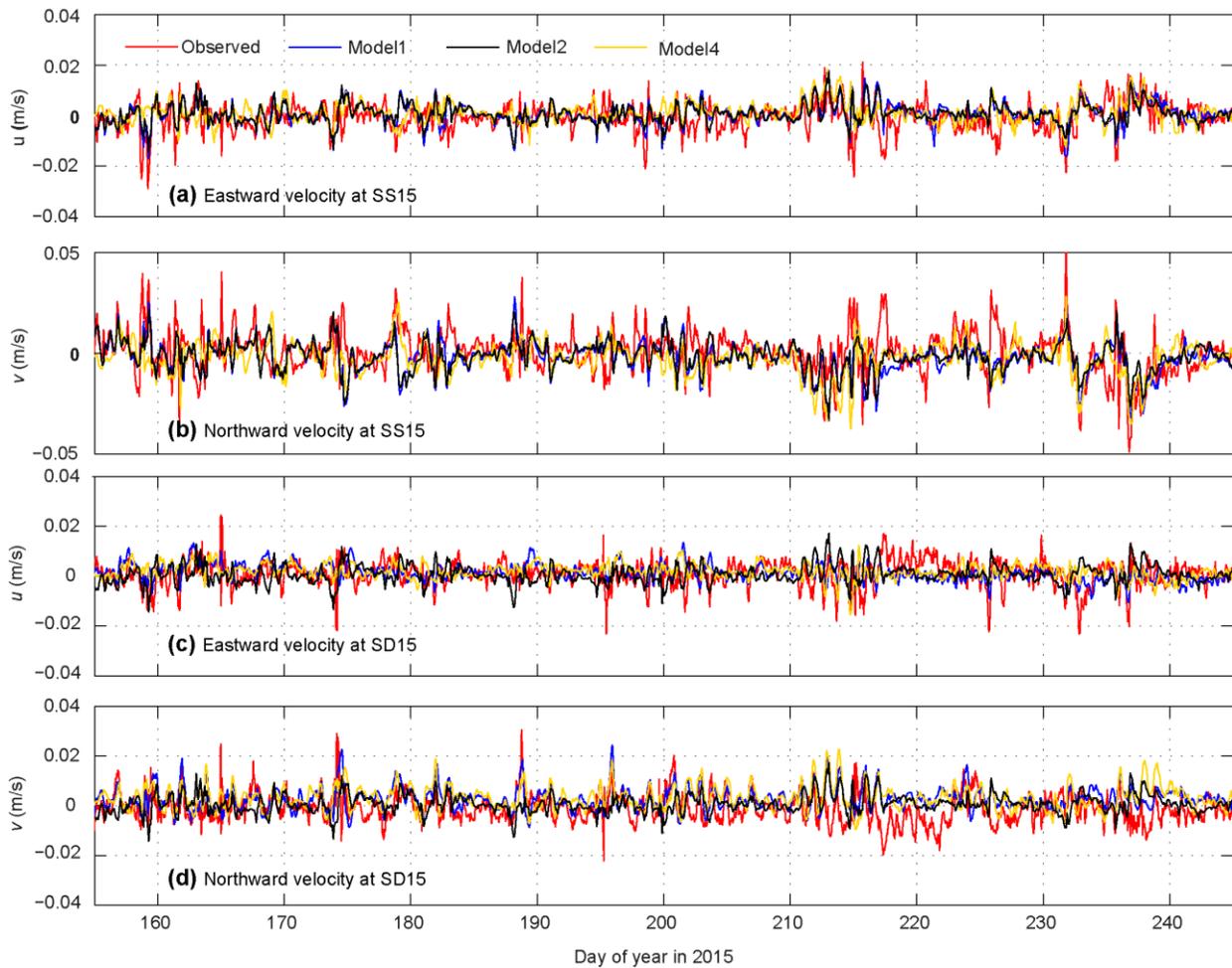
Comparisons of simulated vertical profiles of temperature with weekly Hydrolab data are presented in Figure 3-21. These comparisons illustrate the significant improvement that results from including the bottom cooling in the summer due to groundwater inflow into the lake. The vertical temperature structure was well captured by the models with the groundwater module. However, the model without the groundwater contribution failed to simulate the temperature profiles adequately as the thermal stratification was getting stronger. It should be noted that when the water temperature gradient was sharp, the simulated temperature profiles appeared diffuse as a result of the vertical resolution.



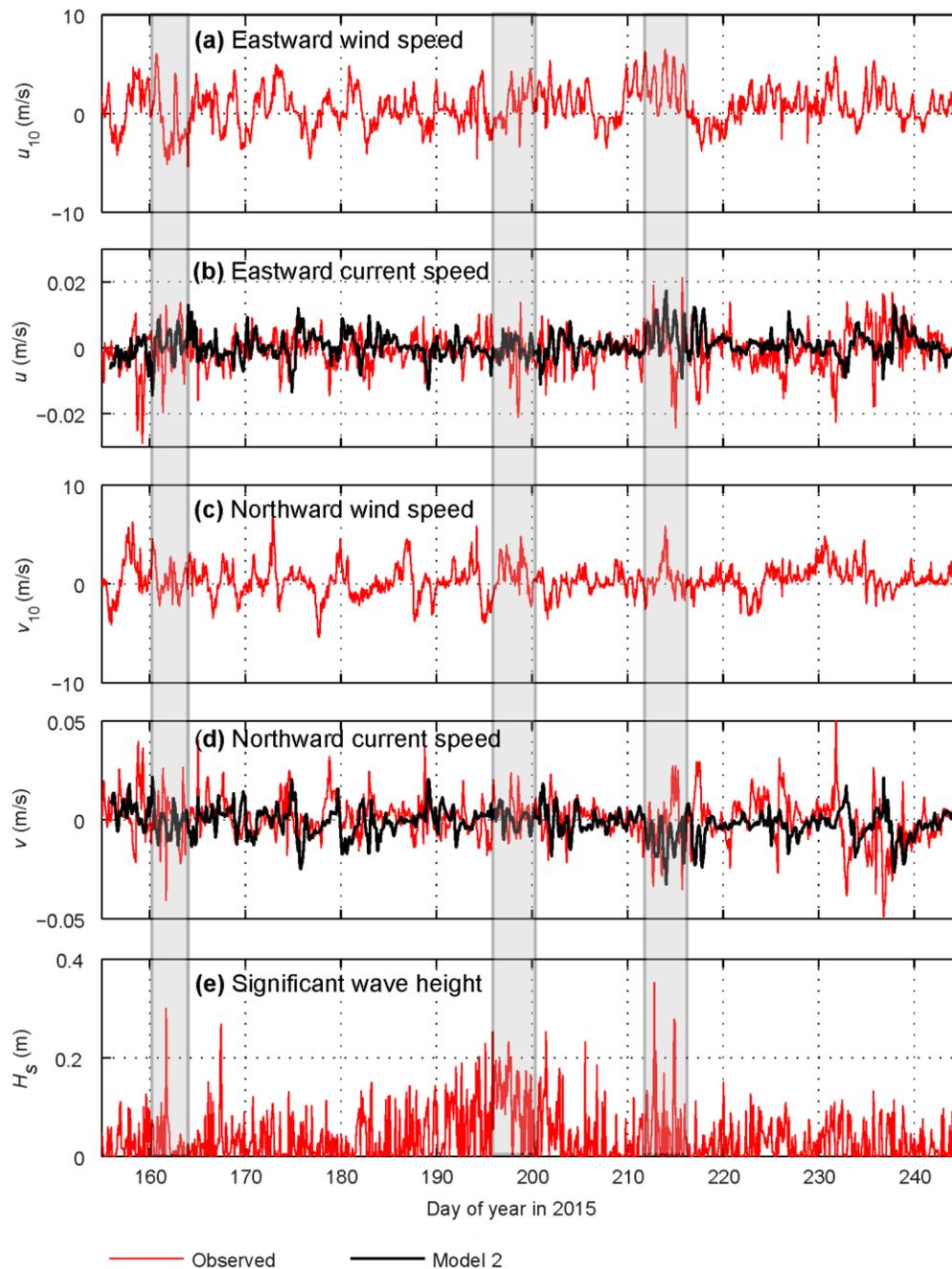
**Figure 3-21.** Comparisons of simulated vertical profiles of temperature with weekly Hydrolab data in 2014.

Finally, the developed models were also tested using data obtained during the summer of 2015. Only the groundwater temperature and flow rates were changed according to the observed 2015 data (Table 3-6). Evaluation of meteorological forcing based on weather stations alone and

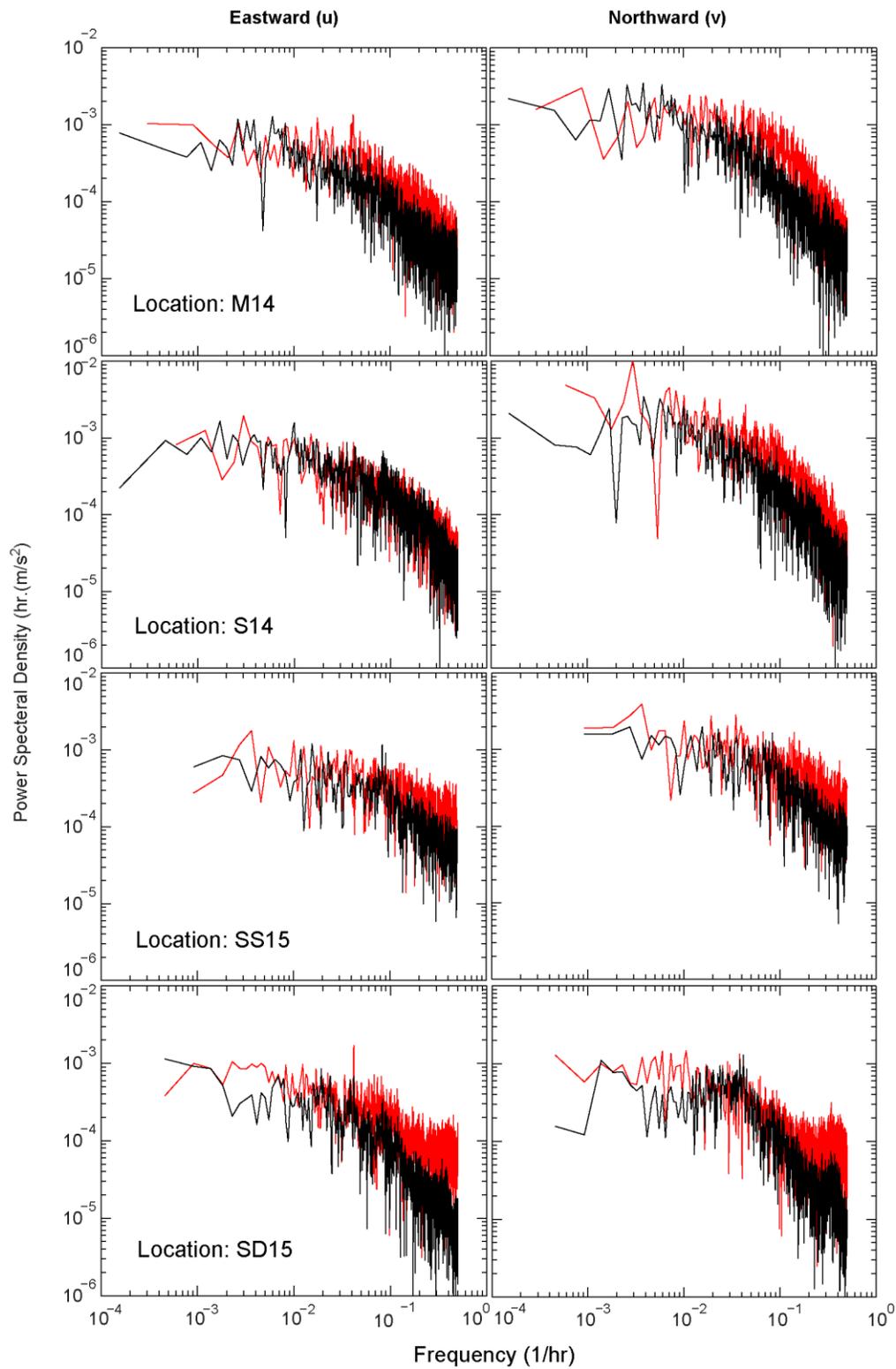
the WRF model outputs was conducted by comparing the performance of the lake model in simulating currents at the ADCP locations (Figures 3-22 to 3-25) as well as thermal structure in the lake (Figure 3-26 to 3-28). The circulation results based on both types of meteorological forcing are comparable with observations at the ADCP locations. The mismatch between observed and simulated currents is found on days with significant wave activity (see DOY 195 in Figure 3-22).



**Figure 3-22.** Comparisons of the simulated vertically averaged velocities and observations at ADCP locations in the summer of 2015.



**Figure 3-23.** Time series of wind, current and significant wave height at SS15 location. (a) Eastward wind speed (m/s), (b) comparison of observed and simulated eastward current speeds (m/s), (c) northward wind speed (m/s), (d) comparison of observed and simulated northward current speeds (m/s), (e) observed significant wave height (m). Gray areas represent examples of periods with significant wave activity during which the model could not adequately capture the observed currents.

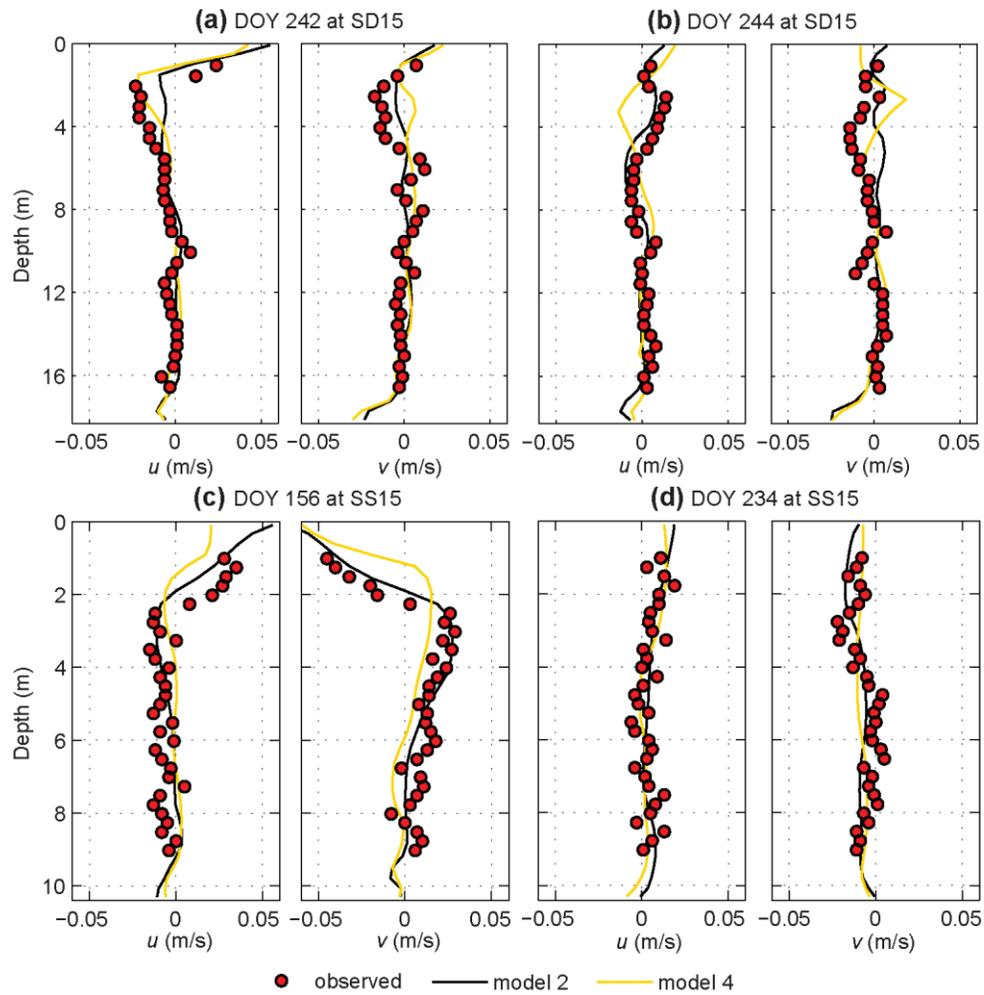


**Figure 3-24.** Comparisons between observed (red lines) and simulated (black lines) power spectral densities at ADCP locations in 2014 and 2015.

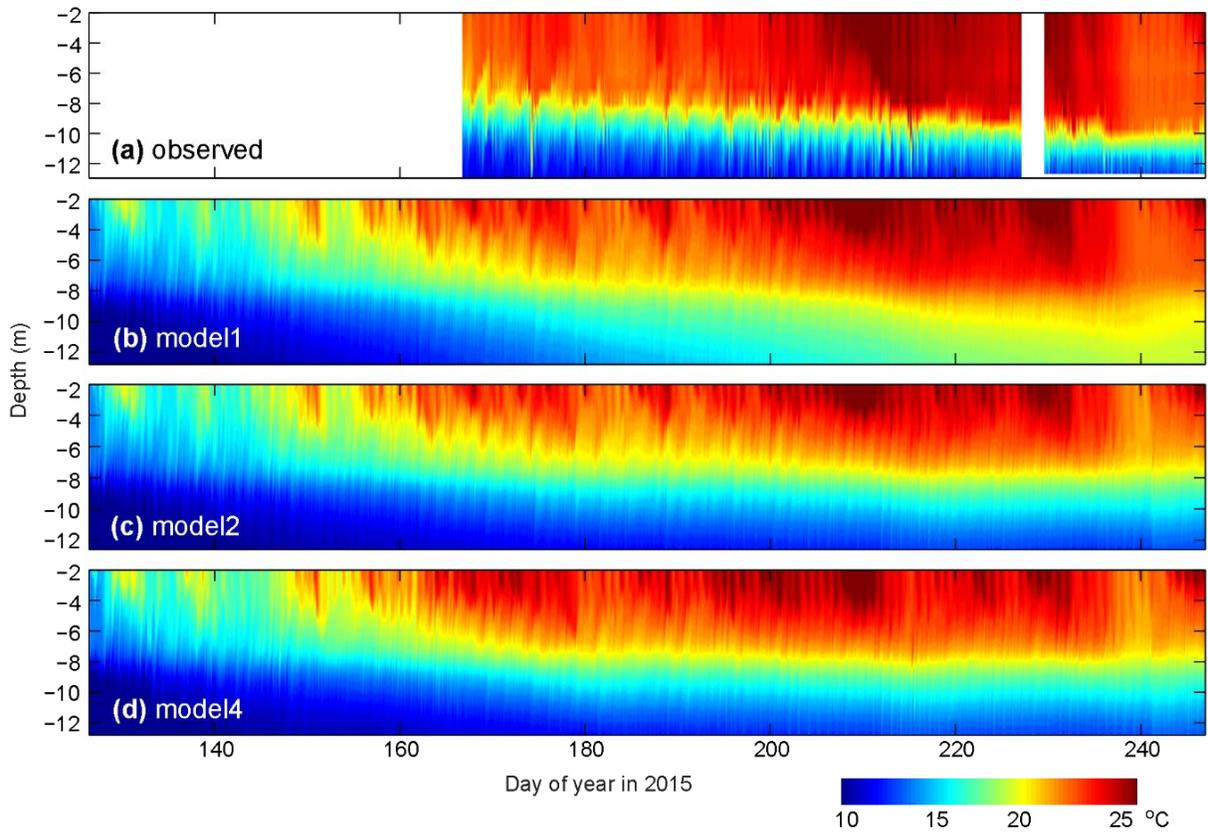
Since wave - current interactions are not simulated, the hydrodynamic model was not able to capture the wave-induced currents on these days (see Figure 3-23). Observed and simulated power spectral densities of vertically averaged velocities were also compared at the ADCP locations for further model evaluation (Figure 3-24). Although wave - current interactions were not taken into account, simulated water currents in time and spectral domains were in good agreement with observed data.

In Figure 3-25, comparisons of the simulated and observed vertical velocity profiles for Model 2 and Model 4 in 2015 at the ADCP locations are plotted for few selected time steps. RMSE values of these comparisons presented in Table 3-5 show that although both Model 2 and Model 4 did a reasonable job of simulating vertical velocity profiles in 2015, Model 2 performed slightly better. The probability distributions of RMSE values at the SD15 location indicated that Model 2 (mode=0.0163 m s<sup>-1</sup>, n=245) had slightly better performance than Model 4 (mode=0.0183 m s<sup>-1</sup>, n=238). At the SS15 location, Model 2 (mode=0.0143 m s<sup>-1</sup>, n=95) produced better results than model 4 (mode=0.0151 m s<sup>-1</sup>, n=121).

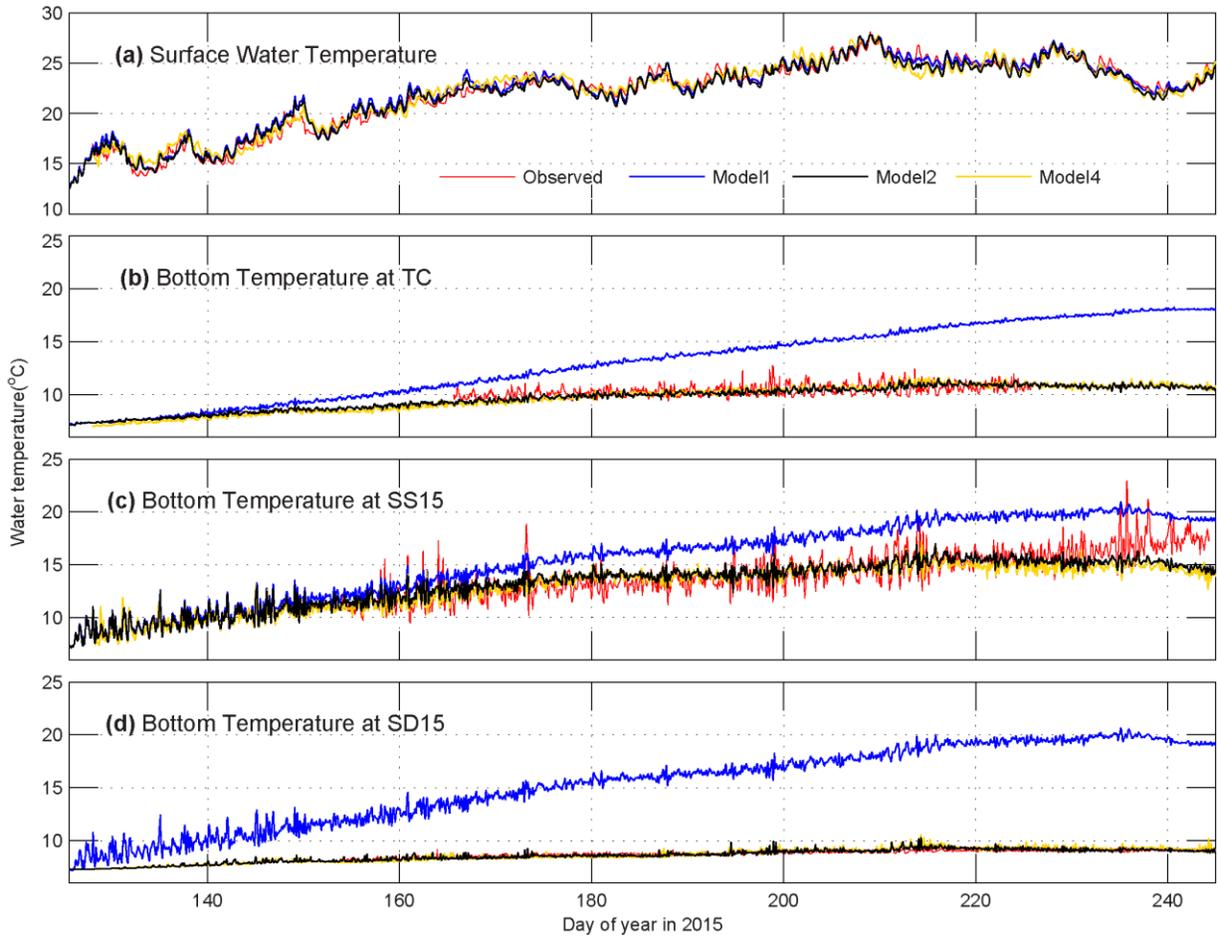
Contour plots of the observed and simulated water temperatures (Figure 3-26) indicate the ability of the models to capture thermal stratification by considering the groundwater contribution. The meteorological forcing used with the FVCOM model resulted in a better simulation of water temperature based on the Taylor diagram presented previously in Figure 3-20. Comparisons of simulated vertical profiles of temperature with weekly Hydrolab data (Figure 3-28) show that both models can accurately simulate vertical temperature profiles, especially in the hypolimnion.



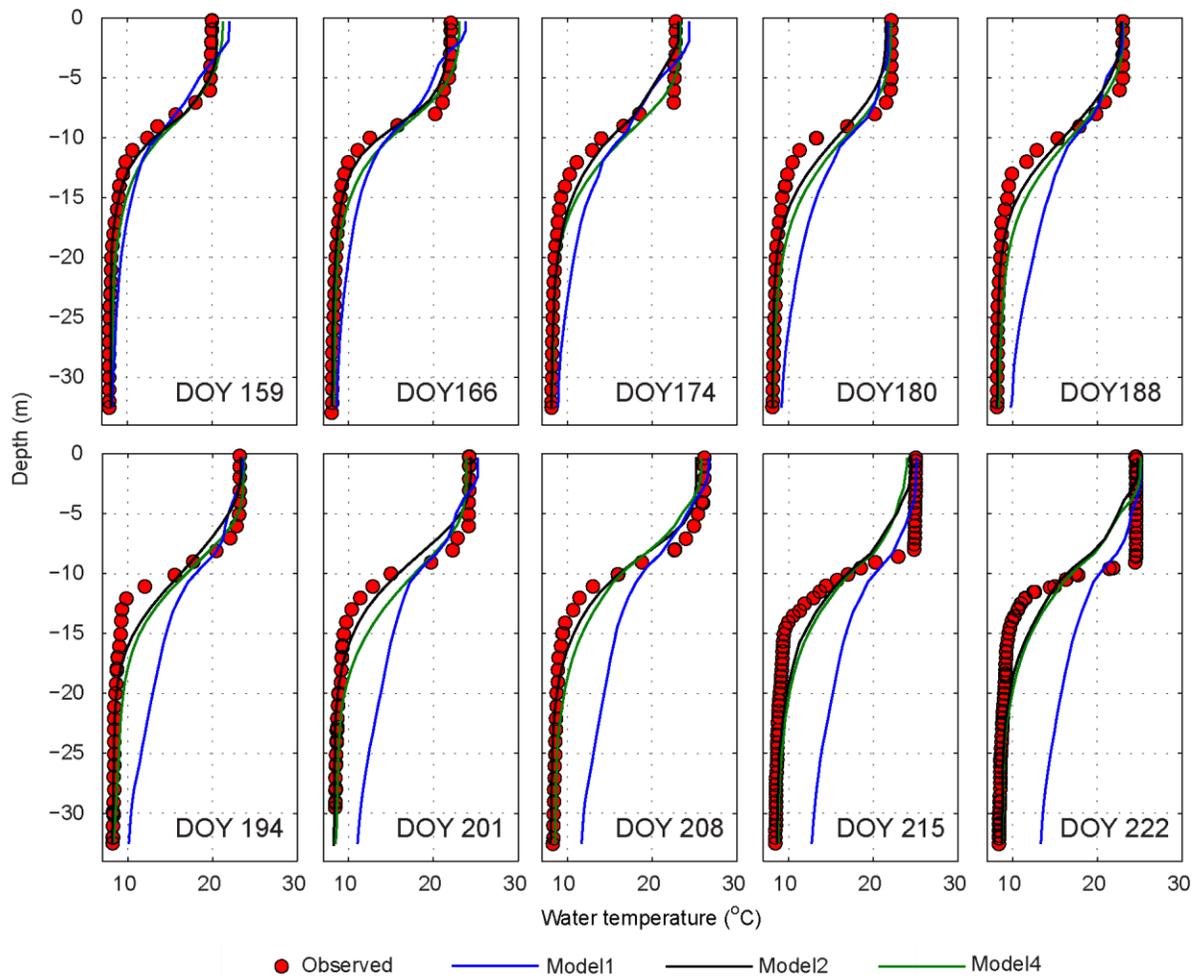
**Figure 3-25.** Examples of comparisons of the observed and simulated vertical velocity profiles at ADCP locations in the summer of 2015.



**Figure 3-26.** Contour plots of the observed and simulated water temperature at the thermistor chain location (TC) in the summer of 2015.



**Figure 3-27.** Comparisons of simulated and observed time series of (a) surface water temperature at TC, and bottom temperature at (b) TC, (c) SS15, and (d) SD15 locations in 2015.



**Figure 3-28.** Comparisons of simulated vertical profiles of temperature with weekly Hydrolab data in 2015.

### 3.4 Conclusions

Due to the dual pressures of climate change and anthropogenic activities such as high-capacity pumping rates exceeding the natural recharge rates due to increasing demand for water, it is important to understand how lake ecosystems respond to warming surface temperatures and declining groundwater levels. Our results clearly bring out the important buffering role played by groundwater in modulating hypolimnetic temperatures and vertical mixing within a deep inland lake in Michigan. Results based on a model simulation without the groundwater contribution

indicated that hypolimnetic temperatures close to the lake bottom can increase by as much as 8 °C or more if lake levels were to drop to a level where the connection with the underlying aquifer is cutoff. This degree of water column heating (in the absence of groundwater) is known to produce important lake-wide changes including decreased dissolved inorganic nitrogen due to higher productivity [Sommaruga-Wögrath *et al.*, 1997] and changes in underwater light climate [Sommaruga *et al.*, 1999].

One feature that was not captured by any of the models considered here is the observed deepening of the thermocline as season progresses (Figure 3-18a and 3-26a). The main reason is due to the assumption that groundwater exchange rate is constant during the simulation and this aspect can be easily improved using either observed groundwater inflow rates or using a fully coupled modeling approach or both. Future modeling using fully coupled atmosphere-lake-groundwater/watershed models is expected to remedy this situation. The fact that meteorological forcing based on the outputs of a mesoscale weather prediction model (WRF) could provide results comparable to the forcing based on a network of weather stations is encouraging, because small lakes in some parts of the world may not have access to any weather networks. As expected, the comparisons between observed and simulated vertically-averaged currents in the lake do not show significant differences between the simulations with and without groundwater as differences within the water column tend to get averaged. However, the vertical velocity and temperature profiles clearly indicate significant differences between the simulations. In some cases, the simulations without the groundwater contribution were in complete disagreement with observations.

In a summary, The hydrodynamic model of Gull lake was developed and further improved using several techniques including (a) the use of a bathymetry-based refinement for the horizontal

mesh (b) the use of *in situ* observations for estimation of light attenuation lengths (c) the use of a coupled WRF-lake model to assess the accuracy of meteorological forcing reconstructed from land-based weather stations and vice versa, and d) the use of the  $k-\varepsilon$  turbulence closure model to improve vertical mixing. All of these model improvements, however, did not resolve the issue of the overheated water column suggesting that hypolimnetic temperatures were primarily controlled by groundwater inflow into the lake and this contribution must be acknowledged to accurately simulate lake thermal structure. The results illustrate the significant improvement in describing thermal structure of Gull Lake corresponding to the bottom cooling in the summer that is controlled by the groundwater exchange.

# CHAPTER 4

## 4 Evaluating the Role of Groundwater in Nutrient - Algal Dynamics

### 4.1 Introduction

Lake surface temperatures are on the rise in response to a warming climate [Magnuson, 2000; Brooks and Zastrow, 2002; Hampton *et al.*, 2008; Schneider and Hook, 2010; Austin, 2013]. Temperature is a critical factor in aquatic systems, and holds the key to controlling vertical stratification, mixing of nutrients, and regulating rate of aquatic ecosystem metabolism [Williamson *et al.*, 1996]. Increased water temperature and the presence of subsequent intensified stratification have substantial effects on water quality of inland lakes [Chang *et al.*, 2015]. The effect of temperature changes in rates of chemical reactions, biological growth rates, photosynthesis and respiration can disturb the equilibrium of the ecosystem [Woolway *et al.*, 2016]. In addition, the solubility of oxygen in water decreases as water temperature increases, thus a lack of dissolved oxygen required for respiration by aquatic organisms can have serious consequences [Carpenter *et al.*, 2011; Zhang *et al.*, 2015]. The deficiency of dissolved oxygen not only has a strong influence on the solubility of phosphorus and other inorganic nutrient, but also results a higher risk of algal blooms and eutrophication [Zhang *et al.*, 2015]. Recent studies indicate that many lakes are facing oxygen depletion and more frequent periods of hypoxia to

anoxia in the hypolimnion have been observed due to lake eutrophication and climate change [Goto *et al.*, 2012; Umaña, 2014; Zhang *et al.*, 2015, 2016b]. Long-term planning to protect freshwater ecosystems depends upon our ability to identify key processes driving the changes [Combes, 2003].

Resilience of an ecosystem depends on interactions across rapidly-changing and slowly-changing variables [Carpenter and Turner, 2000]. Research, monitoring and management efforts have tended to focus more on fast variables, which exhibit tangible variability at a short timescale, rather than the slow variables. Slow variables, however, have an important role in determining the resilience of a system and its regime shifts [Ludwig *et al.*, 2003]. For example, surface fluxes are usually assumed to be the only important driving forces responsible for climate-induced changes in fresh water systems. From a system dynamics point of view, if a system is composed of only a reinforcing loop, the closed loop of cause and effect would lead to instability via accelerating growth [Sterman, 2000]. However, lake responses to climate induce changes, such as drought, vary from lake to lake, which is likely due to the nature and strength of their interaction with groundwater [Baines *et al.*, 2000; Ala-aho *et al.*, 2013]. In fact, groundwater serves as a balancing loop to resist attempted changes in order to maintain a balance and keep the system at a desired state. It is known that rivers that are primarily groundwater-fed are buffered against increasing seasonal temperature variation as the incoming groundwater flow at relatively lower temperatures holds more oxygen to support aquatic organisms [Combes, 2003]. A shallow groundwater-fed lakes can have a significant bottom cooling in summer [Kettle *et al.*, 2012], which may act as a balancing loop within a lake ecosystem, and enhance the ability of the system to resist disturbances from surface-induced changes.

To maintain the resilience of groundwater-dependent ecosystems, it is crucial to monitor and control slow variable associated with groundwater. However, numerous studies indicate long-term groundwater depletion in different regions of the world [Rodell *et al.*, 2009; Wada *et al.*, 2010; Konikow, 2013; Joodaki *et al.*, 2014; Pokhrel *et al.*, 2015; Dalin *et al.*, 2017]. In addition to socio-economic consequences, groundwater depletion can also cause severe environmental impacts on groundwater-dependent aquatic ecosystems by altering surface water-groundwater interactions. Climatic and human-induced groundwater table drawdown could cause a shift in the physical, chemical, and biological structure of groundwater-fed lakes [Webster *et al.*, 1996; Williamson *et al.*, 1996; Gurrieri and Furniss, 2004; Turner and Townley, 2006]. Lack of monitoring slow variables and feedbacks often results in the loss of ecosystem resilience [Biggs *et al.*, 2012]. The history of Lake Urumia, in north-western Iran, is one example of a dramatic regime shift in response to drought and heat stress, combined with increased human-induced stress. The lake water level has experienced significant depletion, and its surface area shrank by 88% in the past decades [AghaKouchak *et al.*, 2015]. Additionally, rising water temperatures and salinity levels accelerate algae growth, and threaten the fragile food web and biodiversity of the ecosystem [AghaKouchak *et al.*, 2015]. Moreover, a color shift from blue to red has often been observed in Urmia Lake in response to environmental stress, which is attributed to high growths of Archaeobacteria and microalgae of *Dunaliella* [Asem *et al.*, 2014].

The present study was motivated by the need to understand the role of groundwater in the resilience of a groundwater-fed inland lake, and determine what regime shifts may occur in the ecosystem in the absence of groundwater. To identify the role of groundwater, we combined extensive field datasets with a coupled biophysical model of Gull Lake, a relatively small but deep inland lake in Michigan, USA. Gull Lake is a groundwater-fed lake with bottom cooling and strong

stratification during summer, and supports diverse warm and cold water fisheries. The performance of the biophysical model was first evaluated against *in situ* observations, including high-resolution current, lake levels, and temperature, nutrient, dissolved oxygen, and chlorophyll data. Then, the developed model was used to predict the ability of the lake ecosystem to resist the disturbance caused by disconnection of the lake from groundwater due to groundwater depletion.

## **4.2 Materials and methods**

### **4.2.1 Field sampling and laboratory analyses**

*In situ* measurements of water quality parameters were carried out weekly by Dr. Elena Litchman's lab at the Kellogg Biological Station during the summer of 2015. A semidiurnal intensive sampling also was conducted around 10:00 AM and 2:00 PM for two weeks in August, whenever the weather cooperated. Water samples were regularly collected weekly for nutrient analysis, including total phosphorus, dissolved inorganic phosphorus, total nitrogen, dissolved nitrate, and chlorophyll concentrations. A four-meter integrated sampler was used for collecting depth-integrated samples of the upper water column. In addition, lake water samples were collected at the surface, deep chlorophyll maxima (DCM), and 32 m depth close to the bottom of the lake. *In situ* fluorescence profiles were measured using a SCAMP (Self-Contained Autonomous Micro-Profiler, <http://pme.com>) to find a depth of DCM within the water column where the fluorescence reaches the maximum value. Moreover, a Hydrolab multi-parameter sonde was used to measure vertical profiles of temperature, Dissolved Oxygen (DO), and chlorophyll concentration in 32 m. Data were recorded at 1 m and 0.5 m depth intervals for weekly sampling and intensive sampling, respectively. DO concentrations were measured using a luminescent-based optical sensor [ASTM D888, 2012]

Nitrogen analyses were conducted with second-derivative spectroscopy [Crumpton *et al.*, 1992]. Dissolved nitrate was determined based on second-derivative UV spectroscopy for filtered samples. Total N was determined using second derivative analyses of nitrate, following persulfate digestion for unfiltered samples. Filtered samples were analyzed using a Lachat Quickchem Autoanalyzer to measure dissolved inorganic phosphorus (orthophosphate or soluble reactive phosphorus, SRP). For total P, unfiltered samples were digested by the persulfate digestion method. Initial conditions for water quality parameters were defined based on observed nutrient compounds in a nearby stream (Prairieville Creek) and a monitoring well near Gull Lake. Nutrient concentration ranges in Prairieville Creek and in the water supply well at the Pond Laboratory at Kellogg Biological Station are presented in Table 4-1. The data were obtained from KBS-LTER database (<https://lter.kbs.msu.edu/datatables>). Dissolved nitrate (NO<sub>3</sub>), ammonium (NH<sub>4</sub>), total dissolved phosphorus (TDP), and soluble reactive phosphorus (PO<sub>4</sub>) were measured on filtered water (0.45 micron).

**Table 4-1.** Nutrient concentration ranges in Prairieville Creek and a pond lab reservoir well near Gull Lake from 2009-2014.

Nutrient forms	Prairieville Creek			Pond Lab Reservoir well		
	average	min	max	average	min	max
NO <sub>3</sub> (mg/L)	5.70	4.90	6.40	0.026	0.001	0.5
NH <sub>4</sub> (ug/L)	28.0	8.30	68.0	14.0	0.73	33.0
TDP (ug/L)	5.97	2.30	9.14	5.13	0.66	14.0
PO <sub>4</sub> (ug/L)	1.80	0.72	3.87	4.12	0.36	14.0

#### 4.2.2 Biophysical model

A fully coupled biophysical model of Gull Lake was developed based on the three-dimensional, unstructured grid, Finite-Volume Community Ocean Model (FVCOM; *Chen et al.*

[2003b; 2006]) to simulate biological and chemical processes in the lake. Thermal structure and circulation in Gull Lake were well-simulated by FVCOM [Safaie *et al.*, 2017b]. A water quality model implemented into FVCOM is based on the EPA Water quality Analysis Simulation Program (WASP5; Ambrose *et al.*, 1993). The water quality model is coupled with FVCOM (FVCOM-WQM), and includes dissolved oxygen budget, nutrient cycling, and phytoplankton biomass. Average values of observed nutrient compounds (Table 4-1) were used to set initial conditions for water quality parameters, and to estimate the nutrient concentrations in stream flows. The performance of the biophysical model was first evaluated against *in situ* observations using the water quality model coupled with Model 2 (Table 3-2 in Chapter 3). Concentrations of the water quality parameters in inflowing groundwater were assumed to be equal to measured bottom value of their concentrations in the lake. The model performance was assessed by comparisons of simulated results and *in situ* observations collected in the summer of 2015. Then, Model 1 (Table 3-2 in Chapter 3) was used to predict the ability of the lake ecosystem in the absence of groundwater effects. Parameter values used in the water quality model of Gull Lake are summarized in Table 4-2.

The model simulates the temporal and spatial dynamics of Dissolved Oxygen (DO), Carbonaceous Biochemical Oxygen Demand (CBOD), Ammonia Nitrogen (NH<sub>3</sub>), Nitrate Nitrogen (NO<sub>3</sub>), Organic Nitrogen (ON), Orthophosphorus (OPO<sub>4</sub>), Organic Phosphorus (OP), and phytoplankton. The mass balance equation for the water quality components can be written as:

$$\frac{\partial C_i}{\partial t} + \frac{\partial(uC_i)}{\partial x} + \frac{\partial(vC_i)}{\partial y} + \frac{\partial(wC_i)}{\partial z} = \frac{\partial}{\partial x}(A_h \frac{\partial C_i}{\partial x}) + \frac{\partial}{\partial y}(A_h \frac{\partial C_i}{\partial y}) + \frac{\partial}{\partial z}(K_h \frac{\partial C_i}{\partial z}) + S_i \quad (1)$$

where  $C_i (i=1, \dots, 8)$  are the concentrations of water quality state variables, which are (1) DO (mg O<sub>2</sub>/L), (2) CBOD (mg C/L), (3) phytoplankton (mg C/L), (4) NH<sub>3</sub> (mg N/L), (5) NO<sub>3</sub> (mg N/L),

(6) ON (mg N/L), (7) OPO<sub>4</sub> (mg P/L), and (8) OP (mg P/L), respectively;  $u$ ,  $v$ , and  $w$  are velocity components in  $x$ ,  $y$ , and  $z$  directions;  $A_h$  and  $K_h$  denote the horizontal mixing coefficient and vertical eddy diffusion coefficient, respectively;  $S_i$  ( $i=1, \dots, 8$ ) is the internal source/sink terms. Since advection and dispersion terms are the same for all transport equations, we describe the source/sink terms in more detail below.

### 1. Dissolved oxygen ( $C_1$ )

The dynamics of dissolved oxygen are controlled by the following chemical and biological processes [Ambrose *et al.*, 1993]: (1) reaeration, (2) oxidation, (3) nitrification, (4) phytoplankton loss, (5) phytoplankton growth, (6) sediment oxygen demand, and (7) bacterial respiration. The term that includes all these processes can be written as follows:

$$S_1 = k_{reae} \theta_{reae}^{(T-20)} (C_s - C_1) - k_{deox} \theta_{deox}^{(T-20)} \frac{C_1 C_2}{K_{BOD} + C_1} - \frac{64}{14} k_{nitr} \theta_{nitr}^{(T-20)} \frac{C_1 C_4}{K_{nitr} + C_1} - \frac{32}{12} D_P C_3 + G_P \left[ \frac{32}{12} + \frac{48}{14} a_{nc} (1 - P_{NH_3}) \right] C_3 - k_{bresp} \quad (2)$$

The first term in Eq. (2) is the reaeration term, where  $k_{reae}$  is the reaeration rate coefficient at 20 °C (day<sup>-1</sup>) and  $\theta_{reae}$  is the temperature coefficient of reaeration.  $C_s$  and  $C_1$  denote the dissolved oxygen saturation and dissolved oxygen concentration (mg O<sub>2</sub>/L), respectively.  $C_s$  is calculated as a function of temperature and salinity [APHA, 1998]:

$$\ln C_s = -139.34411 + \frac{1.575701 \times 10^5}{T_K} - \frac{6.642308 \times 10^7}{T_K^2} + \frac{1.243800 \times 10^{10}}{T_K^3} - \frac{8.621949 \times 10^{11}}{T_K^4} - 0.5535 \cdot S \left( 0.031929 - \frac{19.428}{T_K} + \frac{3867.3}{T_K^2} \right) \quad (3)$$

where  $T_k$  is the water temperature in Kelvin (K).  $k_{ree}$  in the first term of Eq. (2) is calculated as the maximum of wind-induced and flow-induced reaeration. Oxygen reaeration induced by wind is determined using O'Conner's method (O'Connor, 1983). Flow-induced reaeration is calculated as a power function of average hydraulic depth and velocity using the Covar formulation [Covar, 1976].

The second term of Eq. (2) is the CBOD oxidation, where  $k_{deox}$  is the CBOD deoxygenation rate at 20 °C (day<sup>-1</sup>) and  $\theta_{ree}$  is the temperature coefficient of deoxygenation.  $C_2$  and  $K_{BOD}$  denote CBOD concentration and the half-saturation constant for oxygen limitation of CBOD oxidation (mg O<sub>2</sub>/L), respectively. The third term of Eq. (2) represents the nitrification, where  $k_{nitr}$  is the nitrification rate at 20 °C (day<sup>-1</sup>) and  $\theta_{nitr}$  is the temperature coefficient of nitrification.  $C_4$  and  $K_{nitr}$  are concentration of NH<sub>3</sub> and the half-saturation constant for oxygen limitation of nitrification (mg O<sub>2</sub>/L), respectively. In this term, 64/14 mg of oxygen per mg of NH<sub>3</sub> is consumed in the nitrification process. The fourth term in Eq. (2) shows phytoplankton oxygen consumption due to its respiration and death, where  $D_p$  is the phytoplankton loss (day<sup>-1</sup>), and  $C_3$  is the concentration of phytoplankton (mg C/L). The respiration process in this term would consume 32/12 mg of oxygen per mg of phytoplankton carbon.

The fifth term in Eq. (2) represents oxygen production by phytoplankton photosynthesis and nitrogen reduction, where  $G_p$  is the growth rate of phytoplankton (day<sup>-1</sup>),  $a_{nc}$  is the stoichiometric ratio of nitrogen to carbon in phytoplankton. In this term, the stoichiometric oxygen to carbon ratio of 32/12 mg of O<sub>2</sub> are generated for 1 mg of phytoplankton carbon produced in the growth process. In addition, phytoplankton nitrogen reduction process, which reduces nitrate nitrogen to ammonia nitrogen, produce  $48/14 a_{nc}$  mg of oxygen.  $P_{NH_3}$  is the constant of ammonia preference. Dissolved

inorganic nitrogen is taken up by for phytoplankton growth. Both ammonia and nitrate can be consumed by phytoplankton. However, ammonia nitrogen is the preferred form, so the ammonia preference term was used to model their ammonia preference [Ambrose *et al.*, 1993]:

$$P_{\text{NH}_3} = \frac{C_4 C_5}{(C_4 + 0.0001 K_{\text{mN}})(C_5 + 0.0001 K_{\text{mN}})} + \frac{0.0001 C_4 K_{\text{mN}}}{(C_4 + C_5)(C_5 + 0.0001 K_{\text{mN}})} \quad (4)$$

The last term of Eq. (2) is the oxygen consumption by bacterial respiration, where  $k_{\text{bresp}}$  is the bacterial respiration rate (mg O<sub>2</sub>/day).

At the bottom of the lake, the following equation was solved to include the DO flux exchange from the benthic layer to the water column:

$$\left. \frac{\partial C_1}{\partial t} \right|_{z=-H} = -H = D_z \frac{\partial^2 C_1}{\partial z^2} \quad (5)$$

where  $D_z$  denotes diffusive exchange coefficient (m<sup>2</sup>/day).

## 2. Carbonaceous biochemical oxygen demand (C<sub>2</sub>)

The amount of Carbonaceous Biochemical Oxygen Demand (CBOD) in a water body depends on (1) phytoplankton loss, (2) oxidation, (3) denitrification, and (4) settling. These processes can be described by the following equation:

$$S_2 = \frac{32}{12} D_P C_3 - k_{\text{deox}} \theta_{\text{deox}}^{(T-20)} \frac{C_1 C_2}{K_{\text{BOD}} + C_1} - \frac{5}{4} \frac{32}{14} k_{\text{deni}} \theta_{\text{deni}}^{(T-20)} \frac{K_{\text{NO}_3} C_5}{K_{\text{NO}_3} + C_1} - (1 - f_{\text{D2}}) v_{\text{S2}} \frac{\partial C_2}{\partial z} \quad (6)$$

where  $K_{\text{NO}_3}$  = half-saturation concentration for oxygen limitation of denitrification (mg O<sub>2</sub>/L),  $f_{\text{D2}}$  = fraction of dissolved CBOD,  $C_s$  is the dissolved oxygen saturation (mg O<sub>2</sub>/L) (Eq. (3)), and  $v_{\text{S2}}$  = organic matter settling velocity (m.day<sup>-1</sup>).

### 3. Phytoplankton ( $C_3$ )

Sources and sinks of phytoplankton are mainly described by the following processes: (1) phytoplankton growth, (2) phytoplankton loss, and (3) phytoplankton settling:

$$S_3 = G_P C_3 - D_P C_3 - v_{s3} \frac{\partial C_3}{\partial z} \quad (7)$$

where  $v_{s2}$  = phytoplankton settling velocity (m.day<sup>-1</sup>), and the remaining variables have the same definition as those in Eq. (2). The growth rate of phytoplankton ( $G_P$ ) is modeled as:

$$G_P = k_{grow} R_N R_I \theta^{(T-20)} \quad (8)$$

where  $k_{grow}$  is the optimum phytoplankton growth rate at 20 °C (day<sup>-1</sup>),  $R_N$  is the growth rate reduction due to nutrient limitation,  $R_I$  denotes growth rate reduction due to light limitation, and  $\theta_{grow}$  is the temperature coefficient of optimum growth.  $R_N$  is determined using the Michaelis-Menten model for inorganic nutrients:

$$R_N = \min \left( \frac{C_4 + C_5}{K_{MN} + C_4 + C_5}, \frac{C_7}{K_{MP} + C_7} \right) \quad (9)$$

where  $K_{MN}$  and  $K_{MP}$  are half-saturation constants for uptake of inorganic nitrogen (mg N/L) and phosphorus (mg P/L), respectively.  $R_I$  is determined using the model proposed by (Steele, 1962), which has the following form:

$$R_I = \frac{I_z}{I_s} \exp \left( 1 - \frac{I_z}{I_s} \right) \quad (10)$$

where  $I_s$  is the optimum light intensity ( $\text{W}/\text{m}^2$ ), and  $I_z$  is the light intensity in a water column.  $I_z$  is an exponential function of water depth that can be calculated using the Beer's law equation:

$$I_z = I_0 \exp(-k_e z) \quad (11)$$

where  $I_0$  is the light intensity at the surface ( $\text{W}/\text{m}^2$ ),  $z$  is the water depth (m), and  $k_e$  is the light extinction coefficient ( $\text{m}^{-1}$ ).  $k_e$  is calculated as the sum of the vertical light attenuation coefficient for pure water ( $k_w$ ) and the phytoplankton self-shading attenuation ( $k_{shd}$ ). Light intensity attenuates by the presence of phytoplankton biomass in a column of water [Shigesada and Okubo, 1981]. Self-shading of light by algae growing in a column of water can be expressed as [Ambrose *et al.*, 1993]:

$$k_{shd} = 0.0088Chl + 0.054Chl^{0.667} \quad (12)$$

where  $Chl$  is the chlorophyll concentration ( $\mu\text{g}/\text{L}$ ) at water depth of  $z$ . Chlorophyll concentrations can be estimated by  $Chl = C_2 / a_{chl}$ , where  $a_{chl}$  is the ratio of carbon to chlorophyll. Substitution of Eq. (11) into Eq. (10) in a layer integrated form is used to calculate  $R_l$  in FVCOM:

$$R_l = \frac{2.718}{k_e \Delta z} \left( \exp \left( -\frac{I_0}{I_s} \exp -k_e z_{i+1} \right) - \exp \left( -\frac{I_0}{I_s} \exp -k_e z_i \right) \right) \quad (13)$$

where  $z_i$  and  $z_{i+1}$  are depths from the free surface to the bottom and top of each sigma layer (m), respectively, and  $\Delta z$  is the layer thickness (m).

The phytoplankton loss rate ( $D_p$ ) used in Eq. (2) and Eq. (7) is determined by considering phytoplankton respiration and death:

$$D_P = k_{resp} \theta_{resp}^{(T-20)} + k_{mort} \theta_{mort}^{(T-20)} \quad (14)$$

where  $k_{resp}$  = phytoplankton respiration rate at 20 °C (day<sup>-1</sup>);  $\theta_{resp}$  = temperature coefficient of phytoplankton respiration;  $k_{mort}$  = rate of phytoplankton mortality at 20 °C (day<sup>-1</sup>); and  $\theta_{mort}$  = temperature coefficient of phytoplankton mortality.

#### 4. Ammonia (C4):

Ammonia, NH<sub>3</sub>, is determined by the balance between the following processes: (1) phytoplankton loss, (2) mineralization from organic nitrogen, (3) phytoplankton uptake, and (4) nitrification:

$$S_4 = a_{nc} D_P (1 - f_{ON}) C_3 + k_{mine1} \theta_{mine1}^{(T-20)} \frac{C_3 C_6}{K_{mPC} + C_3} - a_{nc} G_P P_{NH_3} C_3 - k_{nitr} \theta_{nitr}^{(T-20)} \frac{C_1 C_4}{K_{nitr} + C_1} \quad (15)$$

where  $a_{nc}$  = the stoichiometric ratio of nitrogen to carbon in phytoplankton;  $D_P$  = phytoplankton loss rate (day<sup>-1</sup>);  $f_{ON}$  = fraction of dead and respired phytoplankton recycled to the organic nitrogen pool;  $k_{mine1}$  = organic nitrogen mineralization at 20 °C (day<sup>-1</sup>);  $\theta_{mine1}$  = temperature coefficient of organic nitrogen mineralization;  $K_{mPC}$  = half-saturation constant of phytoplankton limitation of phosphorus recycle (mg C/L);  $G_P$  = phytoplankton growth rate (day<sup>-1</sup>);  $P_{NH_3}$  = constant of ammonia preference;  $k_{nitr}$  = nitrification rate at 20 °C (day<sup>-1</sup>);  $\theta_{nitr}$  = temperature coefficient of nitrification; and  $K_{nitr}$  = half-saturation constant for oxygen limitation of nitrification (mg O<sub>2</sub>/L).

## 5. Nitrate and nitrite nitrogen ( $C_5$ ):

Nitrogen oxides, including  $\text{NO}_2$  and  $\text{NO}_3$ , are inorganic compounds of nitrogen, which can be described via the following processes: (1) nitrification, (2) phytoplankton uptake, (3) denitrification:

$$S_5 = k_{\text{nit}} \theta_{\text{nit}}^{(T-20)} \frac{C_1 C_4}{K_{\text{nit}} + C_1} - a_{\text{nc}} G_P (1 - P_{\text{NH}_3}) C_3 - k_{\text{denit}} \theta_{\text{denit}}^{(T-20)} \frac{K_{\text{NO}_3} C_5}{K_{\text{NO}_3} + C_1} \quad (16)$$

where  $k_{\text{denit}}$  = denitrification rate at  $20^\circ\text{C}$  ( $\text{day}^{-1}$ );  $\theta_{\text{denit}}$  = temperature coefficient of denitrification;  $K_{\text{NO}_3}$  = half-saturation constant for oxygen limitation of denitrification ( $\text{mg O}_2/\text{L}$ ); and the remaining variables have the same definition as those in Eq.(15). The first term of Eq. (16) represents the nitrification that ammonia ( $\text{NH}_3$ ) is oxidized to nitrite ( $\text{NO}_2^-$ ) and then to nitrate ( $\text{NO}_3^-$ ). The second term shows the preferred amount of  $\text{NO}_3$  that is consumed by phytoplankton via the photosynthetic process. The last term in Eq. (16) is the denitrification that corresponds to the reduction of nitrate to nitrogen gas ( $\text{N}_2$ ).

## 6. Organic nitrogen ( $C_6$ )

Sources and sinks for organic nitrogen (ON) are determined by (1) phytoplankton loss, (2) organic nitrogen mineralization, and (3) ON settling. These processes can be describe as:

$$S_6 = a_{\text{nc}} D_P f_{\text{ON}} C_3 - k_{\text{minel}} \theta_{\text{minel}}^{(T-20)} \frac{C_3 C_6}{K_{\text{mPC}} + C_3} - (1 - f_{D6}) v_{S6} \frac{\partial C_6}{\partial z} \quad (17)$$

where  $f_{D6}$  = fraction of dissolved organic nitrogen, and  $v_{S6}$  = settling velocity of organic nitrogen ( $\text{m.day}^{-1}$ ).

## 7. Inorganic phosphorus ( $C_7$ )

The amount of inorganic phosphorus (Orthophosphorus),  $OPO_4$ , in a water body depends on (1) phytoplankton loss, (2) mineralization from organic phosphorus, and (3) phytoplankton uptake.

These processes are determined in the following form:

$$S_7 = a_{pc} D_P (1 - f_{OP}) C_3 + k_{mine2} \theta_{mine2}^{(T-20)} \frac{C_3 C_8}{K_{mPC} + C_3} - a_{pc} G_P C_3 \quad (18)$$

where  $a_{pc}$  = the stoichiometric ratio of phosphorus to carbon in phytoplankton;  $f_{ON}$  = fraction of dead and respired phytoplankton recycled to the organic phosphorus pool;  $k_{mine2}$  = organic phosphorus mineralization at 20 °C (day<sup>-1</sup>);  $\theta_{mine2}$  = temperature coefficient of organic phosphorus mineralization;  $K_{mPC}$  = half-saturation constant of phytoplankton limitation of phosphorus recycle (mg C/L).

## 8. Organic phosphorus ( $C_8$ )

Sources and sinks for organic phosphorus (OP), in a water body include (1) phytoplankton loss, (2) OP mineralization, and (3) OP settling velocity. These processes are described by the following equation:

$$S_8 = a_{pc} D_P f_{OP} C_3 - k_{mine2} \theta_{mine2}^{(T-20)} \frac{C_3 C_8}{K_{mPC} + C_3} - (1 - f_{D8}) v_{s8} \frac{\partial C_8}{\partial z} \quad (19)$$

where  $f_{D8}$  denotes fraction of organic phosphorus, and  $v_{s8}$  is the settling velocity of organic phosphorus (m day<sup>-1</sup>).

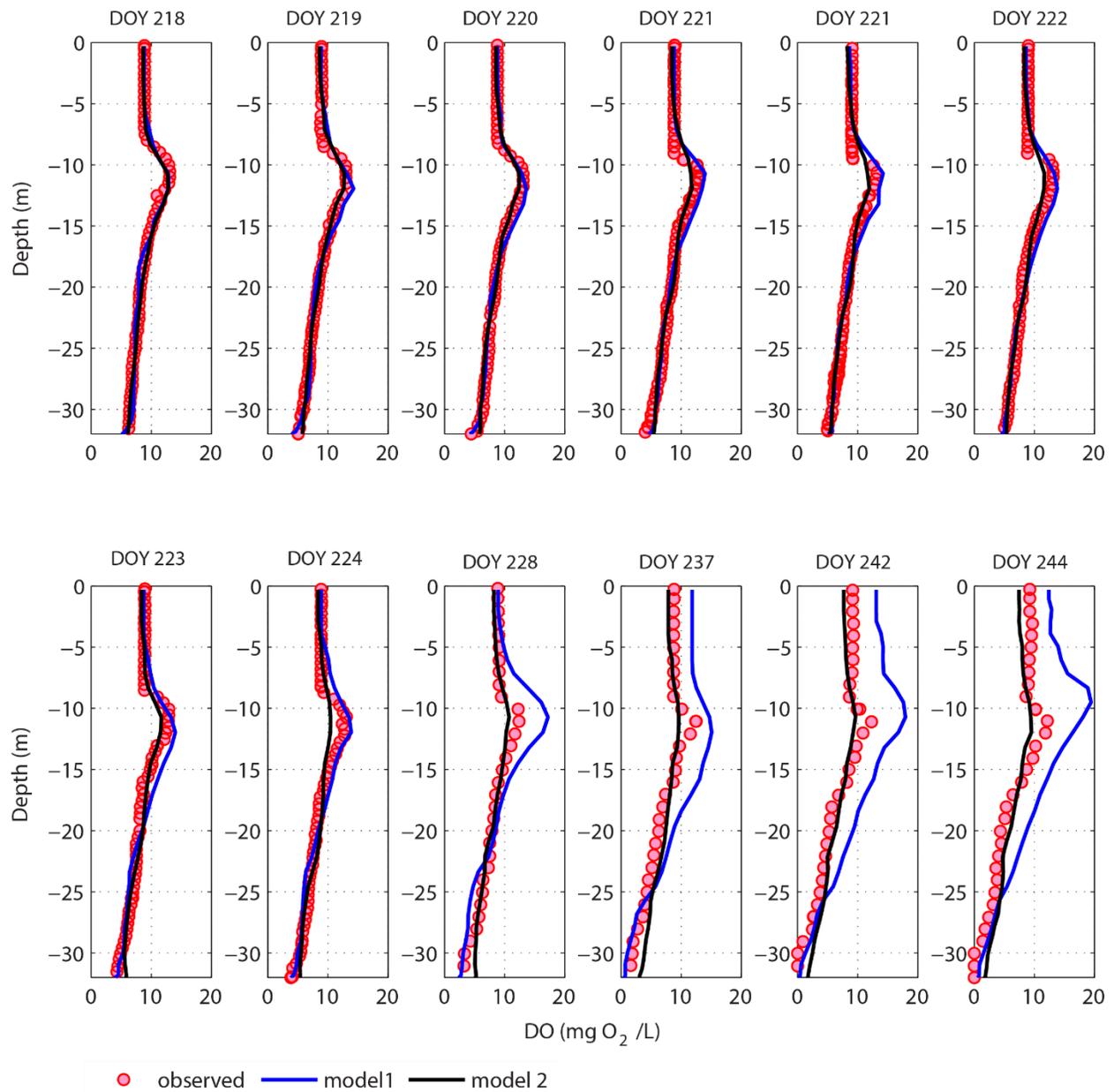
**Table 4-2** Parameter values used in the water quality model

Symbol	Value	Definition
$k_{deox}$	0.05 <sup>a</sup>	Deoxygenation rate at 20 degree (day <sup>-1</sup> )
$k_{nitr}$	0.09 <sup>a</sup>	Nitrification rate at 20 degree (day <sup>-1</sup> )
$k_{resp}$	0.071 <sup>b</sup>	Phytoplankton respiration rate at 20 degree (day <sup>-1</sup> )
$k_{bresp}$	0.20 <sup>c</sup>	Bacterial respiration rate (uM/h)
$k_{deni}$	0.09 <sup>a</sup>	Denitrification rate at 20 degree (day <sup>-1</sup> )
$k_{grow}$	1.21 <sup>b</sup>	Optimum phytoplankton growth rate at 20 degree, (day <sup>-1</sup> )
$k_{mort}$	0.04 <sup>b</sup>	The Mortality rate of phytoplankton at 20 degree, (day <sup>-1</sup> )
$k_{mine1}$	0.075 <sup>b</sup>	Organic nitrogen mineralization at 20 degree (day <sup>-1</sup> )
$k_{mine2}$	0.22 <sup>b</sup>	Organic phosphorus mineralization at 20 degree (day <sup>-1</sup> )
$\theta_{reae}$	1.028 <sup>a</sup>	Temperature coefficient of reaeration
$\theta_{deox}$	1.047 <sup>a</sup>	Temperature coefficient of deoxygenation
$\theta_{nitr}$	1.08 <sup>a</sup>	Temperature coefficient of nitrification
$\theta_{resp}$	1.08 <sup>a</sup>	Temperature coefficient of phytoplankton respiration
$\theta_{deni}$	1.08 <sup>a</sup>	Temperature coefficient of denitrification
$\theta_{grow}$	1.07 <sup>a</sup>	Temperature coefficient of optimum growth
$\theta_{mort}$	1.00 <sup>a</sup>	Temperature coefficient of phytoplankton mortality
$\theta_{mine1}$	1.08 <sup>a</sup>	Temperature coefficient of nitrogen mineralization
$\theta_{mine2}$	1.08 <sup>a</sup>	Temperature coefficient of phosphorus mineralization
$K_{BOD}$	0.5 <sup>a</sup>	Half-saturation constant for oxygen limitation of CBOD oxidation (mg O <sub>2</sub> /l)
$K_{nitr}$	0.5 <sup>a</sup>	Half-saturation constant for oxygen limitation of nitrification (mg O <sub>2</sub> /l)
$K_{mN}$	25 <sup>a</sup>	Half-saturation constant for uptake of inorganic nitrogen (µg N/l)
$K_{mP}$	1 <sup>a</sup>	Half-saturation constant for uptake of inorganic phosphorus (µg P/l)
$K_{NO_3}$	0.1 <sup>a</sup>	Half-saturation constant for oxygen limitation of denitrification (mg O <sub>2</sub> /l)
$K_{mPC}$	1 <sup>a</sup>	Half-saturation constant of phytoplankton limitation of phosphorus recycle (mg C/l)
$D_z$	0.76	Diffusive exchange coefficient (m <sup>2</sup> /day)
$v_{s2}$	0.5 <sup>a</sup>	Organic matter sinking velocity (m/day)
$v_{s3}$	0.14 <sup>d</sup>	Phytoplankton settling velocity (m/day)
$v_{s6}$	0.5 <sup>a</sup>	Settling velocity of organic nitrogen (m.day <sup>-1</sup> )
$v_{s8}$	0.5 <sup>a</sup>	Settling velocity of organic phosphorus (m.day <sup>-1</sup> )
$f_{D2}$	0.5 <sup>a</sup>	Fraction of dissolved CBOD
$f_{D6}$	1 <sup>a</sup>	Fraction of dissolved organic nitrogen
$f_{D8}$	1 <sup>a</sup>	Fraction of dissolved organic phosphorus
$f_{ON}$	0.65 <sup>e</sup>	Fraction of dead and respired phytoplankton recycled to the organic nitrogen pool
$f_{OP}$	0.65 <sup>e</sup>	Fraction of dead and respired phytoplankton recycled to the organic phosphorus pool
$a_{chl}$	60 <sup>f</sup>	Ratio of carbon to chlorophyll
$a_{nc}$	1/12.5 <sup>g</sup>	Ratio of nitrogen to carbon in phytoplankton (mg N/mg C)
$a_{pc}$	1/412 <sup>g</sup>	Ratio of phosphorus to carbon in phytoplankton (mg P/mg C)

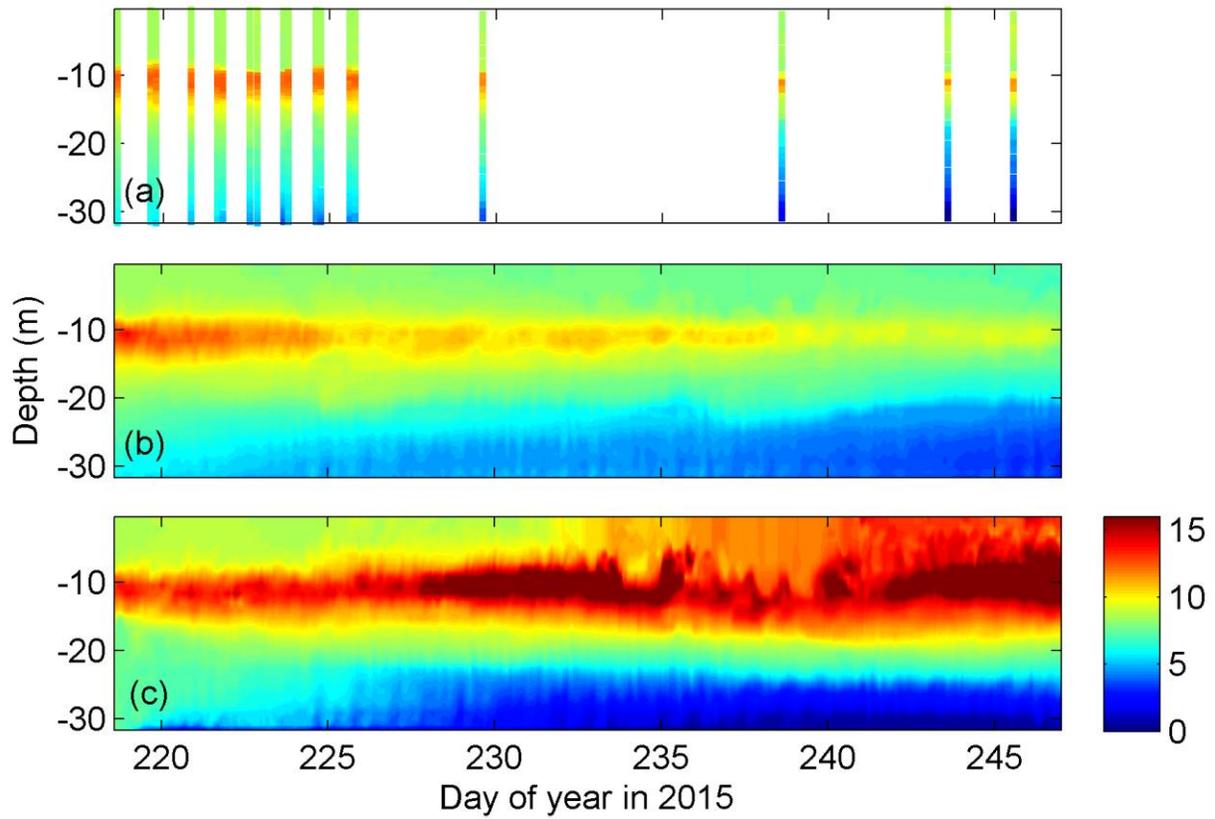
<sup>a</sup>[Ambrose et al., 1993]<sup>b</sup>[Schladow and Hamilton, 1997]<sup>c</sup>[Zheng et al., 2004]<sup>d</sup>This study<sup>e</sup>[Yassuda et al., 2000]<sup>f</sup>[Yacobi and Zohary, 2010]<sup>g</sup>[Hecky et al., 1993]

### 4.3 Results and discussion

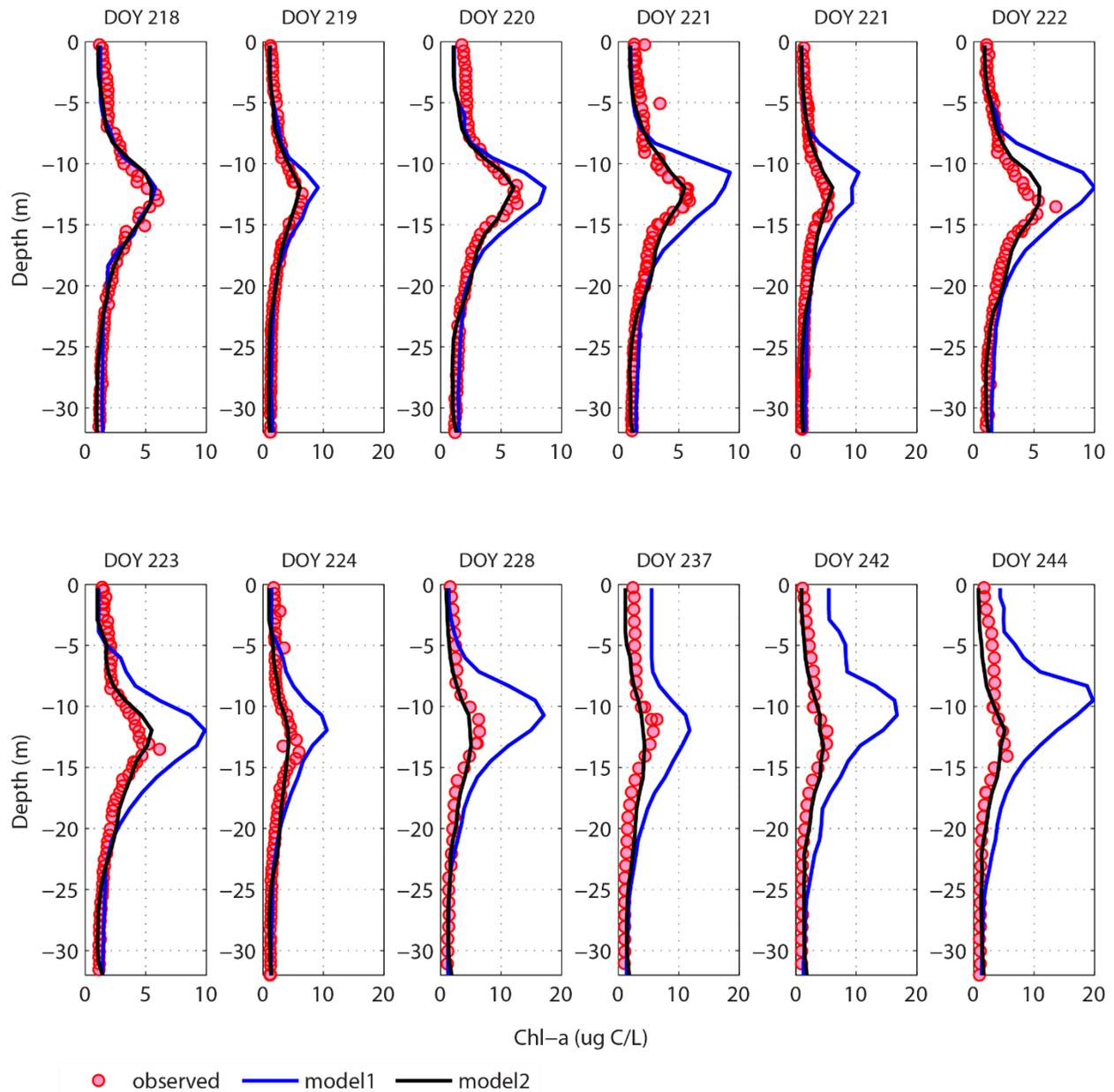
The lake model with groundwater contribution (Model 2 in Chapter 3) closely approximated the observed temporal and spatial dynamics of water quality components. Simulated results show that the developed water quality model has the ability to accurately predict vertical distribution of dissolved oxygen (Figure 4-1 and 4-2b) and phytoplankton biomass (Figure 4-3 and 4-4b). In Figure 4-1, comparisons of observed vertical velocity profiles of DO were compared with the results of Model 2. This figure shows that the vertical DO structure is well captured by the models with the groundwater module (Model 2). Comparisons of simulated vertical profiles of chlorophyll with the Hydrolab data are presented in Figure 4-3. Results indicate that algal dynamics in Gull Lake was mainly controlled by temperature and light limitation (Eq. (10)). During the simulation period, deep chlorophyll maxima (DCM) occurred usually at the center of thermocline located at 13 m depth of water (Model 2 in Figure 4-3 and Figure 4-4b). The intensity of the DCM fluctuation with diel cycle depends on light availability and stability of the water column, and reached its maximum value in the late afternoon [Lucas *et al.*, 2016]. DO stratification was linked to vertical structure of phytoplankton and thermal stratification pattern of the lake [Zhang *et al.*, 2015]. Phytoplankton photosynthesis had a strong effect on oxygen production, which often increased DO saturations up to 135% (Figure 4-5b). In Figures 4-6 and 4-7, nutrient data from water samples collected at DCM, close to the bottom of the lake, and from depth-integrated samples of the upper water column were compared with simulated vertical profiles. In these comparisons, vertical error bars were used to show the standard deviation of observed data. Overall, the comparisons show that Model 2 was able to predict the nutrient-algal dynamics.



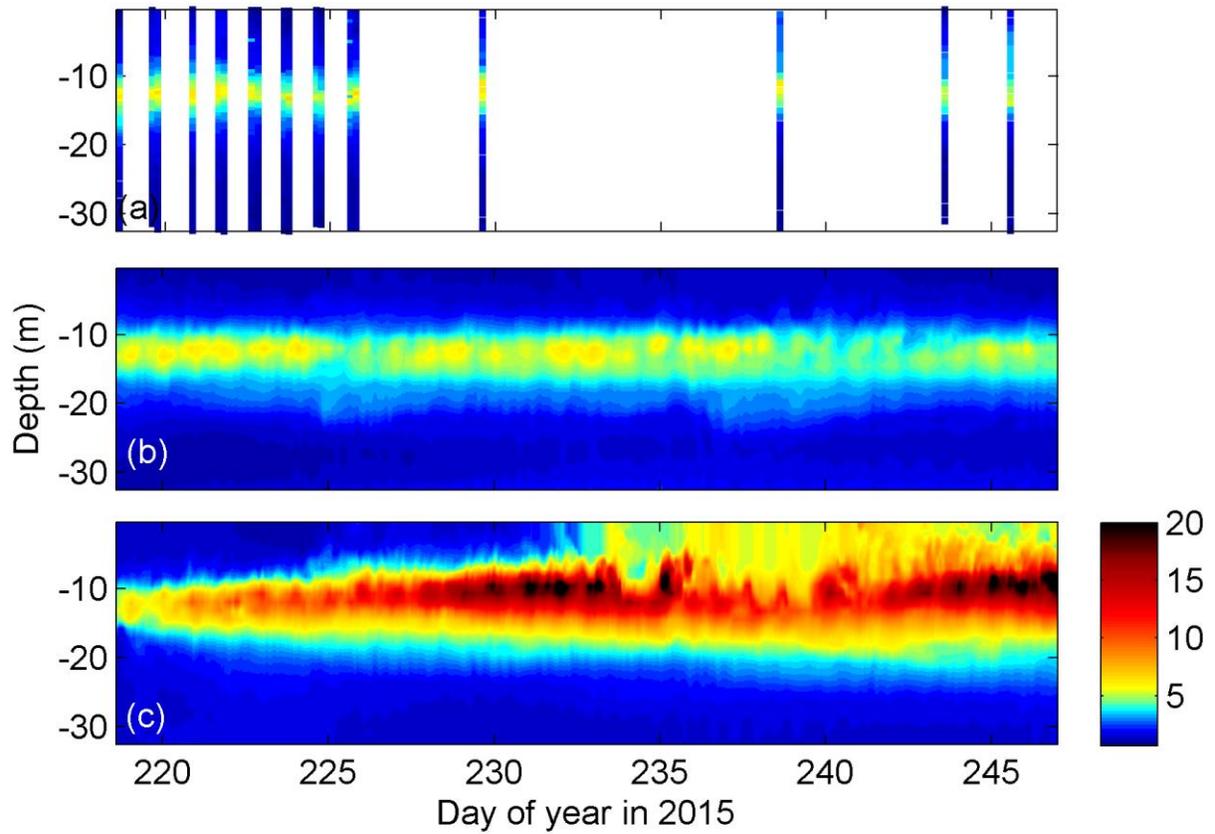
**Figure 4-1** Comparisons of simulated vertical profiles of dissolved oxygen (mg O<sub>2</sub>/L) with Hydrolab data in 2015. Simulated results of Model 1 (blue line) are in the absence of groundwater effects, and results of Model 2 (black line) are with groundwater effects.



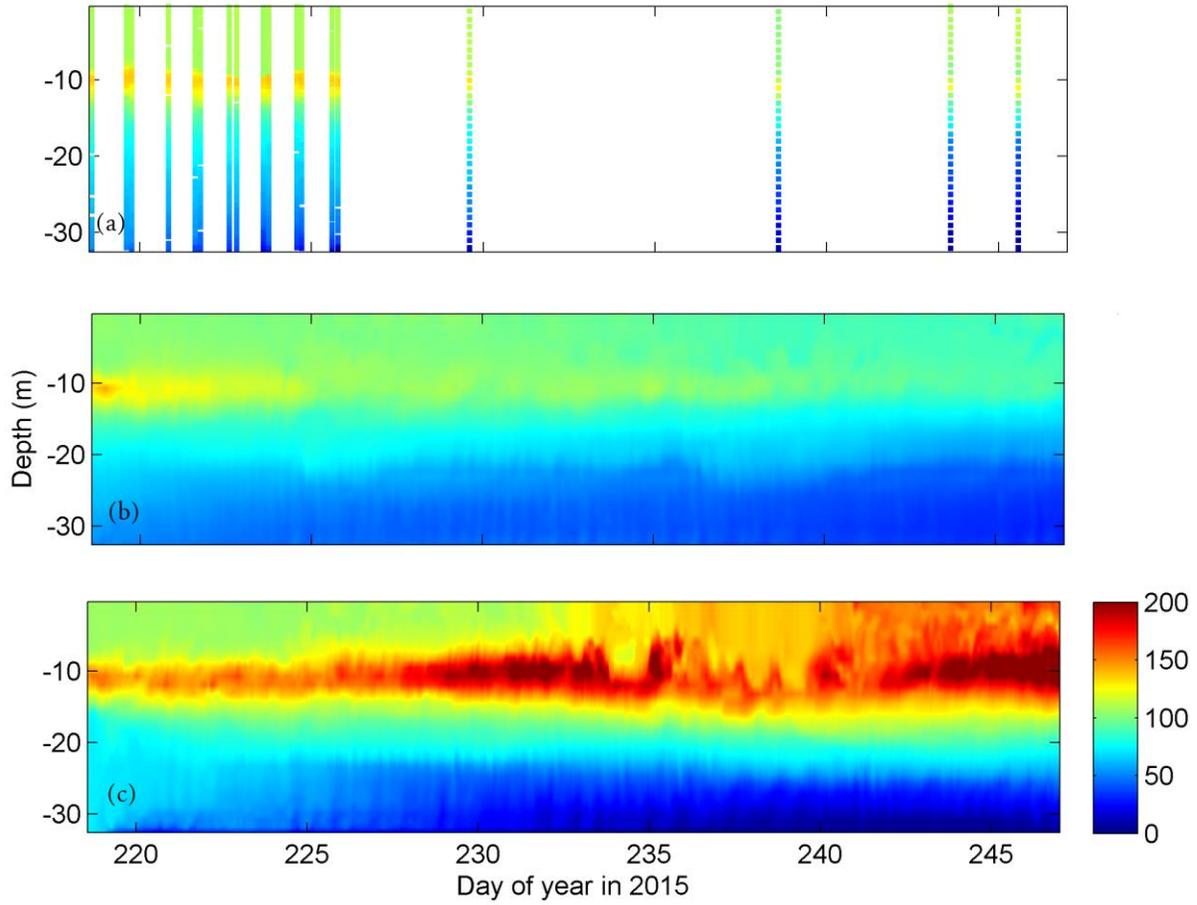
**Figure 4-2.** Contour plots of (a) observed DO ( $\text{mg O}_2/\text{L}$ ), (b) simulated DO with groundwater effects (Model 2) and (c) simulated DO in the absence of groundwater effects (Model 1) in Gull Lake in 2015.



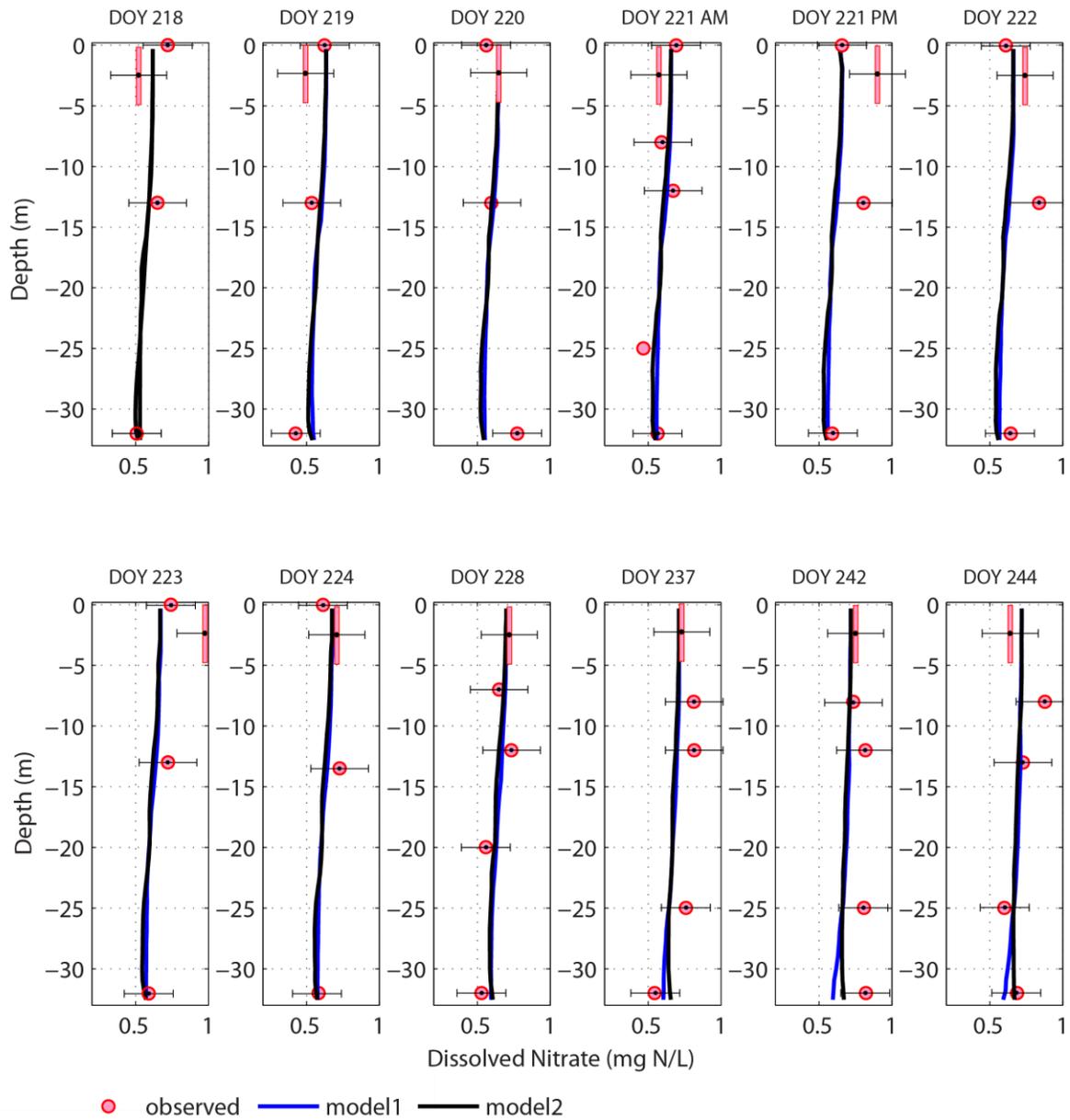
**Figure 4-3.** Comparisons of simulated vertical profiles of Chl-a ( $\mu\text{g C/L}$ ) with Hydrolab data in 2015. Simulated results of Model 1 (blue line) are in the absence of groundwater effects, and results of Model 2 (black line) are with groundwater effects.



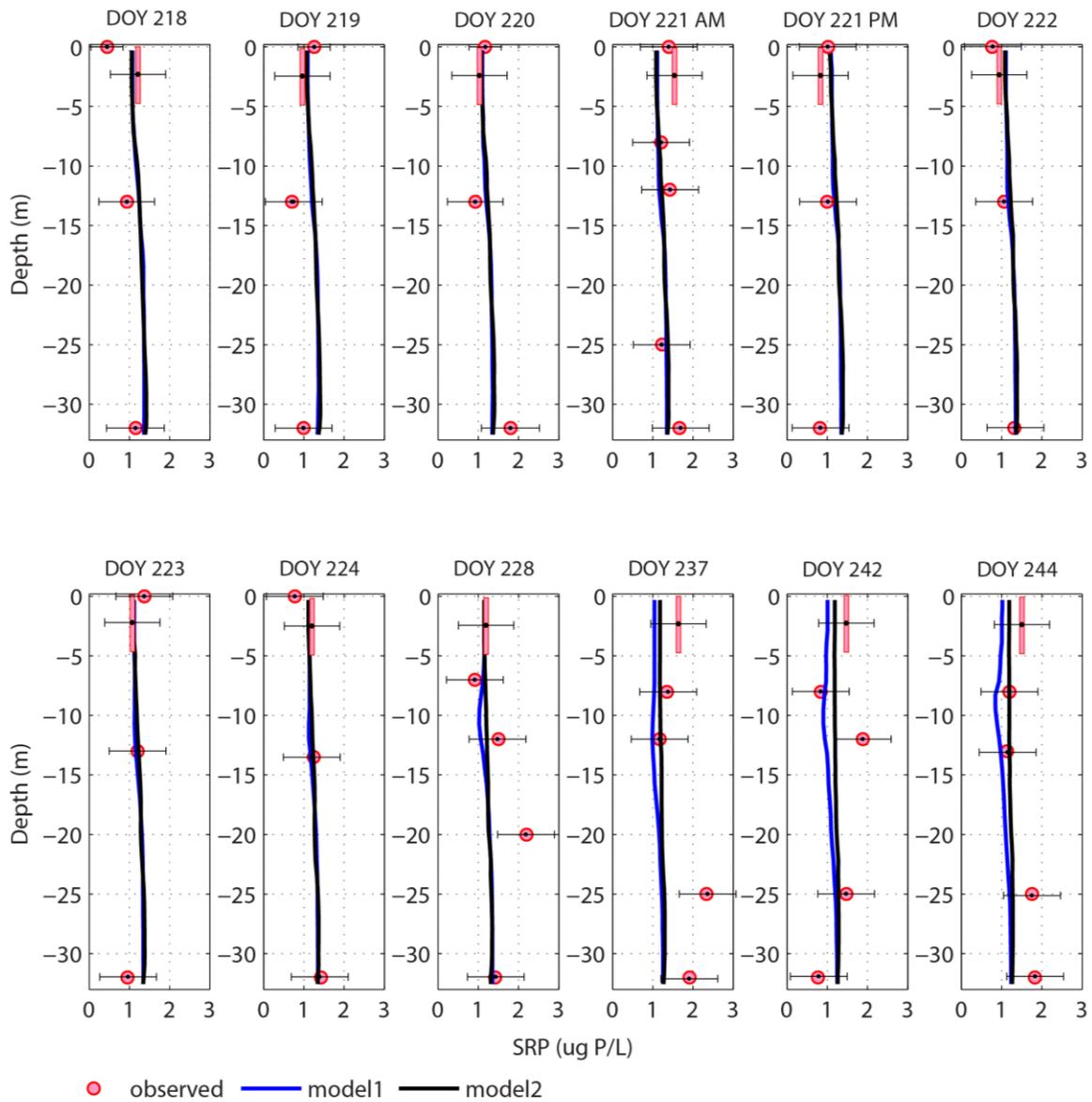
**Figure 4-4.** Contour plots of (a) observed Chl-a ( $\mu\text{g C/L}$ ), (b) simulated Chl-a with groundwater effects (Model 2) and (c) simulated Chl-a in the absence of groundwater effects (Model 1) in Gull Lake in 2015.



**Figure 4-5.** Contour plots of (a) observed DO saturation (%), (b) simulated DO saturation with groundwater effects and (c) simulated DO saturations in the absence of groundwater effects in Gull Lake in 2015.



**Figure 4-6.** Comparisons of observed and simulated vertical profiles of dissolved nitrate (mg N/L) in 2015. Simulated results of Model 1 (blue line) are in the absence of groundwater effects, and results of Model 2 (black line) are with groundwater effects.



**Figure 4-7.** Comparisons of observed and simulated vertical profiles of SRP ( $\mu\text{g P/L}$ ) in 2015. Simulated results of Model 1 (blue line) are in the absence of groundwater effects, and results of Model 2 (black line) are with groundwater effects.

In order to explore the role of groundwater in potential regime shifts in Gull Lake, the developed model was utilized to predict a scenario in which there is no feedback from the groundwater to the lake due to groundwater depletion. Water budget analyses (Chapter 3) showed that groundwater had a high contribution to the water budget of the lake (40-56 percent of total

water sources). In the sense that if no groundwater flow occurs into the lake, the simulated lake stage will be dropped down by 30 cm within the four-month stratification period (from early May to mid-September). Since a fully disconnected system will start to lose more water due to higher infiltration rates [Brunner *et al.*, 2009], the actual decline in the lake stage is expected to be more than 30 cm in transition from fully connected to fully disconnected flow regime. The change can also be reflected in lake water residence times, to such an extent that the 2.16 year residence time of Gull lake at the current condition would be extended by 120% (see Chapter 3). Moreover, results suggest that cold temperatures observed in the hypolimnion within a narrow range of 8 to 10 °C, will be increased by 8 °C by the absence of groundwater. Therefore, the thickness of epilimnion will be expanded from 8 m to 11 m, which could increase warm-water fish distributions and their suitable thermal habitat [Comte *et al.*, 2013]. However, it will also increase growth of algal blooms (Figure 4-4c) and invasive species, such as zebra mussels. Furthermore, heat stress would influence the ability of aquatic organisms, especially those in early life stages, to be immune from toxic substances and diseases.

Excessive algal photosynthesis (Figure 4-4c), influenced by high temperature, produced strengthened DO stratification, but decreased flow of DO to the benthic waters due to the DO stratification (Figure 4-2c). Predicated dissolved oxygen above the thermocline reached supersaturation with DO saturations ranging between 120 and 200% (Figure 4-5c). This supersaturation condition would increase risk of gas bubble disease, bacterial infection, and mortality in fish and invertebrates [Weitkamp and Katz, 1980; Elston and Wood, 1983; Harris *et al.*, 2005]. Additionally, it resulted in quick depletion of nutrients (e.g. DOY 242 in Figure 4-7).

The development of algal biomass also attenuated light intensity through the self-shading effect [Shigesada and Okubo, 1981]. Thus, the DCM locations moved upward to the surface

(Model 1 in Figure 4-3), and the light limitation changed the vertical distributions of DO (Figures 4-1 and 4-2), nutrients and phytoplankton, accordingly. For instance, the decrease in light penetration reduced depths of DCM and oxycline by 3 m in day of the year (DOY) 235 (Figure 4-4c). Warmer bottom temperatures lead to hypoxia and anoxia in the deep waters of the lake (Figure 4-2c and Figure 4-5c). Early and strong hypolimnion anoxic conditions were projected due to the increased sediment oxygen demand, and bacterial respiration. The observed DO concentrations at the 32 m depth of water had a depletion rate of  $0.22 \text{ mg O}_2 \text{ L}^{-1} \text{ day}^{-1}$  during the simulation period, so the initial bottom DO level of  $6.18 \text{ mg O}_2 \text{ L}^{-1}$  reached anoxia within four weeks. However, in the absence of groundwater effects, anoxia occurred 12 days sooner with a depletion rate of  $0.36 \text{ mg O}_2 \text{ L}^{-1} \text{ day}^{-1}$ . The high bottom temperature along with the low DO levels could negatively impact cold-water habitat in the lake. Results suggest that cold-water species with thermal preferences of  $8\text{-}12 \text{ }^\circ\text{C}$  and a benchmark oxygen concentration of  $3 \text{ mg L}^{-1}$  will experience high oxythermal stress [Jiang and Fang, 2016]. It was predicted that a habitat volume for  $\text{DO} > 3 \text{ mg O}_2 \text{ L}^{-1}$  and  $8 \text{ }^\circ\text{C} < T < 12 \text{ }^\circ\text{C}$ , which was 27% of the lake volume in DOY 247, would completely disappear in the absence of groundwater inputs.

## 4.4 Conclusions

Groundwater contribution to Gull Lake was found to have a significant effect on coupled processes within the lake. The results presented in this chapter (with and without groundwater contribution) indicate that feedback from the groundwater-lake interactions, acts as a balancing loop that can help resist environmental disturbances. Cold groundwater flow through the lake causes significant bottom cooling in summer, and this has an important role in water quality, biodiversity, and habitat structure of the lake. For instance, the ability to buffer seasonal water temperature variations in the hypolimnion allows cold-water fish to thrive. This finding

emphasizes the important role of groundwater in controlling the lake temperature whose changes can affect the behavior of the lake ecosystem. This research highlights that changes on groundwater level and its temperature due to natural climate variability and human activities have the potential to lead to an ecological regime shift in groundwater-fed lakes. Groundwater depletion was shown to have potential negative impacts on sustainability of lake ecosystems, e.g. lake level changes, water temperature rising, accelerating algal growth and dissolved oxygen depletion, early anoxia, reduction of light availability, and eutrophication. These changes could cause ecological consequences on food-web functioning and biodiversity of the ecosystem. These results are expected to provide improved understanding of the role of groundwater in physical, chemical and biological processes within groundwater-fed lakes.

# CHAPTER 5

## 5 Conclusions

The specific aim of the present study was to evaluate the role of groundwater in biophysical processes of groundwater-fed lakes. To gain insights into physical, chemical, and biological processes in inland lakes, detailed three-dimensional hydrodynamic and temperature models of Gull Lake coupled to nutrient and algal dynamics were developed using field observations and numerical modeling during the summer stratification period. As a first step towards building a coupled physical-chemical-biological modeling system, we evaluated the role played by groundwater in thermal structure and circulation of the lake. A three-dimensional hydrodynamic model (FVCOM, Finite-Volume Community Ocean Model) and a mesoscale weather prediction model (WRF, Weather Research and Forecasting model) were utilized to simulate hydrodynamic and thermal behaviors of the lake. Wind fields and heat fluxes, which are major driving forces of lake hydrodynamics, were reconstructed from both weather station data and WRF-simulated results. This study shows that WRF-simulated meteorological forcing fields could provide comparable results to those reconstructed from the land-based weather station network. This is encouraging since many lakes in remote parts of the world may not have a network of weather stations around them.

In addition, a novel manifold method of reconstructing spatiotemporal data was proposed for assimilating geophysical and meteorological data into numerical lake circulation and transport models. The effectiveness of the presented manifold algorithm was evaluated through assimilation of geophysical and meteorological data in lake models. Results indicate the superior performance of the manifold method over standard methods in terms of accuracy and computational efficiency for reconstructing meteorological data. Combining manifold methods with assimilation methods such as the ensemble Kalman filter [Moradkhani *et al.*, 2005; Evensen, 2007; Pathiraja *et al.*, 2016] could be an important direction of future work to further improve process-based modeling of land surface, subsurface and lake/ocean models.

In this study, high-resolution currents, lake levels, and temperature data were collected to evaluate the performance of the hydrodynamic model. Two well-known turbulence models, the modified Mellor and Yamada level 2.5 level and the two-equation  $k-\varepsilon$  turbulence models, were compared in order to better describe the observed vertical mixing and thermal structure in the lake. Although both turbulence models produced similar trends in thermal structure, the  $k-\varepsilon$  turbulence model produced better model performance in describing surface temperatures and circulations in the surface mixed layer. To quantify groundwater contribution to the entire lake, water level fluctuations recorded by Acoustic Doppler Current Profilers (ADCPs) were used as a basis for evaluating the water budget of the lake. The developed model achieved promising results when the groundwater contribution was taken into account. However, the simulated lake stage dropped down by 30 cm during the summer in the absence of groundwater contribution.

The water quality model was fully coupled with FVCOM for integrated analysis of dissolved oxygen budget, nutrient cycling, and phytoplankton biomass. The model performance was assessed by comparisons of simulated results and *in situ* water quality observations, including

vertical profiles of nutrients, dissolved oxygen, and chlorophyll concentration in 32 m of water. The developed model was used to evaluate the role of groundwater in dissolved oxygen, nutrients and algal dynamics of Gull Lake. Simulated results show that the developed water quality model has the ability to accurately predict vertical distribution of dissolved oxygen and phytoplankton biomass. Then, the developed model was used to predict the ability of the lake ecosystem to resist disturbance caused by disconnection of the lake from groundwater due to groundwater depletion.

Results suggest that groundwater-fed lakes have the ability to buffer seasonal water temperature variations in hypolimnion. However, the low temperatures observed in the hypolimnion will be increased by 8 °C by neglecting the groundwater effects. This finding emphasizes the important role of groundwater in controlling the lake temperatures. Rising water temperature in the absence of groundwater contribution to the lake could alter water quality, biodiversity, and habitat structure of the lake. Our predicted results indicated that lake level changes, rising water temperatures, increased growth rates of algae, oxygen depletion, early anoxia, reduction of light availability, and eutrophication are some possible consequential effects of groundwater depletion on lake ecosystems. This finding is expected to aid in understanding the role of groundwater in several key biophysical processes in groundwater-fed lakes, and could strengthen management efforts to maintain the resilience of ecosystem functions in the face of external stressors.

# **APPENDIX**

## APPENDIX

The following metrics were used in this study to evaluate model performance:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2} \quad (A1)$$

$$R^2 = \frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^n (P_i - \bar{P})^2}} \quad (A2)$$

$$PBIAS = \frac{\sum_{i=1}^n (O_i - P_i) \times 100}{\sum_{i=1}^n O_i} \quad (A3)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (A4)$$

$$F_n = \frac{\|O_i, P_i\|}{\|O_i, 0\|}, \text{ where } \|O_i, P_i\| = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2} \quad (A5)$$

where  $O_i$  and  $P_i$  are observed and predicted values of a variable, respectively.  $\bar{O}$  and  $\bar{P}$  are the mean values of  $O_i$  and  $P_i$ . The  $R^2$  and RMSE are well-known metrics while PBIAS is a measure of the tendency of the simulated data to be higher or lower relative to the observations. NSE in Eq. (A.4) represents the Nash – Sutcliffe model efficiency, and  $F_n$  in Eq. (A.5) is the Fourier norm provides an indication of the variance in the observed data that is not captured by the model.

The Taylor diagram (Figure 2-30) is based on the normalized standard deviation ( $\sigma$ ), centered root mean square difference ( $E'$ ) between observed and simulated data, correlation coefficient (R) and bias with the following definitions [Taylor, 2001]:

$$R = \frac{\frac{1}{N} \sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sigma_o \sigma_p} \quad (\text{A6})$$

$$BAIS = \bar{P} - \bar{O} \quad (\text{A7})$$

$$E' = \sqrt{\frac{1}{n} \sum_{i=1}^n [(P_i - \bar{P}) - (O_i - \bar{O})]^2} \quad (\text{A8})$$

where  $\sigma_o$  and  $\sigma_p$  are standard deviation of  $O$  and  $P$ . Note that Taylor [2001] used the  $R^2$ , as we defined here in (Eq. A2), for the correlation coefficient (R). Eq. A2 is identical to the definition of R in (Eq. A6) used by Taylor (2001). To be consistent with Taylor's definition, we used the definition in Eq. (A6) for the Taylor diagram. The statistics used in the Taylor diagram have the following relationship:

$$E'^2 = \sigma_o^2 + \sigma_p^2 - 2\sigma_o\sigma_p R \quad (\text{A9})$$

## **BIBLIOGRAPHY**

# BIBLIOGRAPHY

- Abbasi, A., F. Annor, and N. van de Giesen (2016), Investigation of Temperature Dynamics in Small and Shallow Reservoirs, Case Study: Lake Binaba, Upper East Region of Ghana, *Water*, 8(3), 84, doi:10.3390/w8030084.
- Adrian, R. et al. (2009), Lakes as sentinels of climate change, *Limnol. Oceanogr.*, 54(6), 2283.
- AghaKouchak, A., H. Norouzi, K. Madani, A. Mirchi, M. Azarderakhsh, A. Nazemi, N. Nasrollahi, A. Farahmand, A. Mehran, and E. Hasanzadeh (2015), Aral Sea syndrome desiccates Lake Urmia: Call for action, *J. Gt. Lakes Res.*, 41(1), 307–311, doi:10.1016/j.jglr.2014.12.007.
- Ala-aho, P., P. M. Rossi, and B. Kløve (2013), Interaction of esker groundwater with headwater lakes and streams, *J. Hydrol.*, 500, 144–156, doi:10.1016/j.jhydrol.2013.07.014.
- Allard, W. K., G. Chen, and M. Maggioni (2012), Multi-scale geometric methods for data sets II: Geometric Multi-Resolution Analysis, *Appl. Comput. Harmon. Anal.*, 32(3), 435–462, doi:10.1016/j.acha.2011.08.001.
- Allen, J. S., P. A. Newberger, and J. Federiuk (1995), Upwelling Circulation on the Oregon Continental Shelf. Part I: Response to Idealized Forcing, *J. Phys. Oceanogr.*, 25(8), 1843–1866, doi:10.1175/1520-0485(1995)025<1843:UCOTOC>2.0.CO;2.
- Ambrose, R. B., T. A. Wool, and J. L. Martin (1993), *The Water Quality Analysis Simulation Program, WASP5, Part A: Model Documentation*, Environmental Research Laboratory, U.S. Environmental Protection Agency, Athens, Georgia.
- Anderson, M. P. (2005), Heat as a Ground Water Tracer, *Ground Water*, 43(6), 951–968, doi:10.1111/j.1745-6584.2005.00052.x.
- Anderson, E. J., and M. S. Phanikumar (2011), Surface storage dynamics in large rivers: Comparing three-dimensional particle transport, one-dimensional fractional derivative, and multirate transient storage models, *Water Resour. Res.*, 47(9), 1–15, doi:10.1029/2010WR010228.
- Annear, R. L., and S. A. Wells (2007), A comparison of five models for estimating clear-sky solar radiation, *Water Resour. Res.*, 43(10), 1–15, doi:10.1029/2006WR005055.
- APHA (1998), *Standard methods for the examination of water and wastewater*, 20th ed., American Public Health Association, Washington, DC.
- Apple, B. A., and H. W. Reeves (2007), *Summary of Hydrogeologic Conditions by County for the State of Michigan*, Open-File Report, U.S. Geological Survey, Reston, Virginia, United States.

- Asem, A., A. Eimanifar, M. Djamali, P. De los Rios, and M. Wink (2014), Biodiversity of the Hypersaline Urmia Lake National Park (NW Iran), *Diversity*, 6(1), 102–132, doi:10.3390/d6010102.
- ASTM D888 (2012), *Standard Test Methods for Dissolved Oxygen in Water*, ASTM International, West Conshohocken, PA.
- Austin, J. (2013), Observations of near-inertial energy in Lake Superior, *Limnol. Oceanogr.*, 58(2), 715–728, doi:10.4319/lo.2013.58.2.0715.
- Austin, J. A., and S. M. Colman (2007), Lake Superior summer water temperatures are increasing more rapidly than regional air temperatures: A positive ice-albedo feedback, *Geophys. Res. Lett.*, 34(6), doi:10.1029/2006GL029021.
- Bai, X., J. Wang, D. J. Schwab, Y. Yang, L. Luo, G. A. Leshkevich, and S. Liu (2013), Modeling 1993–2008 climatology of seasonal general circulation and thermal structure in the Great Lakes using FVCOM, *Ocean Model.*, 65, 40–63, doi:10.1016/j.ocemod.2013.02.003.
- Baines, S. B., K. E. Webster, T. K. Kratz, S. R. Carpenter, and J. J. Magnuson (2000), Synchronous behavior of temperature, calcium, and chlorophyll in lakes of northern Wisconsin, *Ecology*, 81(3), 815–825.
- Barnier, B., P. Marchesiello, A. P. De Miranda, J.-M. Molines, and M. Coulibaly (1998), A sigma-coordinate primitive equation model for studying the circulation in the South Atlantic. Part I: Model configuration with error estimates, *Deep Sea Res. Part Oceanogr. Res. Pap.*, 45(4), 543–572.
- Barth, A., A. A. Azcárate, P. Joassin, B. Jean-Marie, and C. Troupin (2008), Introduction to Optimal Interpolation and Variational Analysis, SESAME Summer School, Varna, Bulgaria.
- Belkin, M., and P. Niyogi (2003), Laplacian eigenmaps for dimensionality, *Speech Commun.*, 1(2–3), 349–367.
- Biggs, R. et al. (2012), Toward Principles for Enhancing the Resilience of Ecosystem Services, *Annu. Rev. Environ. Resour.*, 37(1), 421–448, doi:10.1146/annurev-environ-051211-123836.
- Blumberg, A. F., and G. L. Mellor (1987), A description of a three-dimensional coastal ocean circulation model, *Am. Geophys. Union*, (edited by N. Heaps), 1–16.
- Boegman, L., M. R. Loewen, P. F. Hamblin, and D. A. Culver (2008), Vertical mixing and weak stratification over zebra mussel colonies in western Lake Erie, *Limnol. Oceanogr.*, 53(3), 1093–1110.
- Brewer, M. K. (1991), A regional study of ground-water quality in Barry County, Michigan, M.S., Western Michigan University, United States -- Michigan.

- Brooks, A. S., and J. C. Zastrow (2002), The Potential Influence of Climate Change on Offshore Primary Production in Lake Michigan, *J. Gt. Lakes Res.*, 28(4), 597–607, doi:10.1016/S0380-1330(02)70608-4.
- Bruesewitz, D. A., J. L. Tank, and S. K. Hamilton (2009), Seasonal effects of zebra mussels on littoral nitrogen transformation rates in Gull Lake, Michigan, U.S.A., *Freshw. Biol.*, 54(7), 1427–1443, doi:10.1111/j.1365-2427.2009.02195.x.
- Brunner, P., C. T. Simmons, and P. G. Cook (2009), Spatial and temporal aspects of the transition from connection to disconnection between rivers, lakes and groundwater, *Journal of Hydrology*, 376(1–2), 159–169, doi:10.1016/j.jhydrol.2009.07.023.
- Brunner, D. et al. (2015), Comparative analysis of meteorological performance of coupled chemistry-meteorology models in the context of AQMEII phase 2, *Atmos. Environ.*, 115, 470–498, doi:10.1016/j.atmosenv.2014.12.032.
- Bunker, A. F. (1976), Computations of Surface Energy Flux and Annual Air–Sea Interaction Cycles of the North Atlantic Ocean, *Mon. Weather Rev.*, 104(9), 1122–1140, doi:10.1175/1520-0493(1976)104<1122:COSEFA>2.0.CO;2.
- Burchard, H., O. Petersen, and T. P. Rippeth (1998), Comparing the performance of the Mellor–Yamada and the  $\kappa$ - $\epsilon$  two-equation turbulence models, *J. Geophys. Res. Oceans*, 103(C5), 10543–10554, doi:10.1029/98JC00261.
- Burns, N. M., D. C. Rockwell, P. E. Bertram, D. M. Dolan, and J. J. H. Ciborowski (2005), Trends in Temperature, Secchi Depth, and Dissolved Oxygen Depletion Rates in the Central Basin of Lake Erie, 1983–2002, *J. Gt. Lakes Res.*, 31, 35–49, doi:10.1016/S0380-1330(05)70303-8.
- Carin, L., R. G. Baraniuk, V. Cevher, D. Dunson, M. I. Jordan, G. Sapiro, and M. B. Wakin (2011), Learning Low-Dimensional Signal Models, *IEEE Signal Process. Mag.*, 28(2), doi:10.1109/MSP.2010.939733.
- Carpenter, S. R., and M. G. Turner (2000), Hares and Tortoises: Interactions of Fast and Slow Variables in Ecosystems, *Ecosystems*, 3(6), 495–497, doi:10.1007/s100210000043.
- Carpenter, S. R., E. H. Stanley, and M. J. Vander Zanden (2011), State of the World’s Freshwater Ecosystems: Physical, Chemical, and Biological Changes, *Annu. Rev. Environ. Resour.*, 36(1), 75–99, doi:10.1146/annurev-environ-021810-094524.
- Chang, C.-H., L.-Y. Cai, T.-F. Lin, C.-L. Chung, L. van der Linden, and M. Burch (2015), Assessment of the Impacts of Climate Change on the Water Quality of a Small Deep Reservoir in a Humid-Subtropical Climatic Region, *Water*, 7(4), 1687–1711, doi:10.3390/w7041687.
- Chen, C., Beardsley, R.C., Franks, P.J.S., Van Keuren, J., (2003a), Influence of diurnal heating on stratification and residual circulation of Georges Bank, *J. Geophys. Res.*, 108(C11), doi:10.1029/2001JC001245.

- Chen, C., H. Liu, and R. C. Beardsley (2003b), An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: application to coastal ocean and estuaries, *J. Atmospheric Ocean. Technol.*, 20(1), 159–186.
- Chen, C., R. Beardsley, and G. Cowles (2006), An Unstructured Grid, Finite-Volume Coastal Ocean Model (FVCOM) System, *Oceanography*, 19(1), 78–89, doi:10.5670/oceanog.2006.92.
- Chikita, K., S. P. Joshi, J. Jha, and H. Hasegawa (2000), Hydrological and thermal regimes in a supra-glacial lake: Imja, Khumbu, Nepal Himalaya, *Hydrol. Sci. J.*, 45(4), 507–521.
- Combes, S. (2003), Protecting freshwater ecosystems in the face of global climate change, *Buy. Time User's Man. Build. Resist. Resil. Clim. Change Nat. Syst. WWF Clim. Change Program*.
- Comte, L., L. Buisson, M. Daufresne, and G. Grenouillet (2013), Climate-induced changes in the distribution of freshwater fish: observed and predicted trends: *Climate change and freshwater fish*, *Freshw. Biol.*, 58(4), 625–639, doi:10.1111/fwb.12081.
- Constantz, J. (2008), Heat as a tracer to determine streambed water exchanges, *Water Resour. Res.*, 44(4), doi:10.1029/2008WR006996.
- Covar, A. P. (1976), Selecting the Proper Reaeration Coefficient for Use in Water Quality Models, in *Proceedings of the EPA Conference on Environmental Modeling and Simulation*, Cincinnati, Ohio.
- Crumpton, W. G., T. M. Isenhardt, and P. D. Mitchell (1992), Nitrate and organic N analyses with second-derivative spectroscopy, *Limnol. Oceanogr.*, 37(4), 907.
- Dalin, C., Y. Wada, T. Kastner, and M. J. Puma (2017), Groundwater depletion embedded in international food trade, *Nature*, 543(7647), 700–704, doi:10.1038/nature21403.
- Dang, C., and H. Radha (2014), Single Image Super Resolution via Manifold Approximation, *ArXiv Prepr. ArXiv14103349*.
- Dang, C., and H. Radha (2015), Fast Image Super Resolution via Selective Manifold Learning of High Resolution Patches, Québec City, Canada.
- Dang, C., M. Aghagolzadeh, and H. Radha (2014), Image Super-Resolution via Local Self-Learning Manifold Approximation, *IEEE Signal Process. Lett.*, 21(10), 1245–1249, doi:10.1109/LSP.2014.2332118.
- Dang, C. T., M. Aghagolzadeh, A. A. Moghadam, and H. Radha (2013), Single image super resolution via manifold linear approximation using sparse subspace clustering, in *Global Conference on Signal and Information Processing (GlobalSIP), 2013 IEEE*, pp. 949–952, IEEE, Austin, Texas, U.S.A.

- Denman, K. L., and A. E. Gargett (1983), Time and space scales of vertical mixing and advection of phytoplankton in the upper ocean, *Oceanography*, 28(5).
- Dingman, S. L. (2002), *Physical hydrology*, 2nd ed., Prentice Hall, Upper Saddle River, N.J.
- Edson, J. B., V. Jampana, R. A. Weller, S. P. Bigorre, A. J. Plueddemann, C. W. Fairall, S. D. Miller, L. Mahrt, D. Vickers, and H. Hersbach (2013), On the Exchange of Momentum over the Open Ocean, *J. Phys. Oceanogr.*, 43(8), 1589–1610, doi:10.1175/JPO-D-12-0173.1.
- Edwards, W. J., C. R. Rehmann, E. McDonald, and D. A. Culver (2005), The impact of a benthic filter feeder: limitations imposed by physical transport of algae to the benthos, *Can. J. Fish. Aquat. Sci.*, 62(1), 205–214, doi:10.1139/f04-188.
- Elston, R., and G. S. L. Wood (1983), Pathogenesis of vibriosis in cultured juvenile red abalone, *Haliotis rufescens* Swainson, *J. Fish Dis.*, 6(2), 111–128, doi:10.1111/j.1365-2761.1983.tb00059.x.
- Evensen, G. (2007), *Data assimilation: the ensemble Kalman filter*, Springer, Berlin; New York.
- Fairall, C. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S. Young (1996), Bulk parameterization of air-sea fluxes for Tropical Ocean-Global Atmosphere Coupled-Ocean Atmosphere Response Experiment, *J. Geophys. Res. Oceans*, 101(C2), 3747–3764, doi:10.1029/95JC03205.
- Farrand, W. R. (1982), *Quaternary Geology of Michigan*, Michigan Department of Environmental Quality, Geological Survey Division, Lansing, MI.
- Foreman, M. G. G., P. Czajko, D. J. Stucchi, and M. Guo (2009), A finite volume model simulation for the Broughton Archipelago, Canada, *Ocean Model.*, 30(1), 29–47, doi:10.1016/j.ocemod.2009.05.009.
- Galperin, B., L. H. Kantha, S. Hassid, and A. Rosati (1988), A Quasi-equilibrium Turbulent Energy Model for Geophysical Flows, *J. Atmospheric Sci.*, 45(1), 55–62, doi:10.1175/1520-0469(1988)045<0055:AQETEM>2.0.CO;2.
- Golub, G., and C. F. Van Loan (2013), *Matrix Computations*, 4th ed., Baltimore: Johns Hopkins.
- Gorman, G. J., M. D. Piggott, C. C. Pain, C. R. E. de Oliveira, A. P. Umpleby, and A. J. H. Goddard (2006), Optimisation based bathymetry approximation through constrained unstructured mesh adaptivity, *Ocean Model.*, 12(3–4), 436–452, doi:10.1016/j.ocemod.2005.09.004.
- Goto, D., K. Lindelof, D. Fanslow, S. Ludsin, S. Pothoven, J. Roberts, H. Vanderploeg, A. Wilson, and T. Höök (2012), Indirect consequences of hypolimnetic hypoxia on zooplankton growth in a large eutrophic lake, *Aquat. Biol.*, 16(3), 217–227, doi:10.3354/ab00442.
- Guillemin, V., and A. Pollack (2010), *Differential topology*, AMS Chelsea Pub, American Mathematical Society, Providence, R.I.

- Gurrieri, J. T., and G. Furniss (2004), Estimation of groundwater exchange in alpine lakes using non-steady mass-balance methods, *J. Hydrol.*, 297(1–4), 187–208, doi:10.1016/j.jhydrol.2004.04.021.
- Haidvogel, D. B., and A. Beckmann (1999), Numerical Ocean Circulation Modeling, *Imp. Coll. Press*.
- Haidvogel, D. B., H. G. Arango, K. Hedstrom, A. Beckmann, P. Malanotte-Rizzoli, and A. F. Shchepetkin (2000), Model evaluation experiments in the North Atlantic Basin: simulations in nonlinear terrain-following coordinates, *Dyn. Atmospheres Oceans*, 32(3), 239–281.
- Hampton, S. E., L. R. Izmet'eva, M. V. Moore, S. L. Katz, B. Dennis, and E. A. Silow (2008), Sixty years of environmental change in the world's largest freshwater lake - Lake Baikal, Siberia, *Glob. Change Biol.*, 14(8), 1947–1958, doi:10.1111/j.1365-2486.2008.01616.x.
- Harris, J. O., C. M. Burke, S. J. Edwards, and D. R. Johns (2005), Effects of oxygen supersaturation and temperature on juvenile greenlip, *Haliotis laevigata* Donovan, and blacklip, *Haliotis rubra* Leach, abalone, *Aquac. Res.*, 36(14), 1400–1407, doi:10.1111/j.1365-2109.2005.01360.x.
- Hecky, R. E., P. Campbell, and L. L. Hendzel (1993), The stoichiometry of carbon, nitrogen, and phosphorus in particulate matter of lakes and oceans, *Limnol. Oceanogr.*, 38(4), 709–724, doi:10.4319/lo.1993.38.4.0709.
- Heiskanen, J. J. et al. (2015), Effects of water clarity on lake stratification and lake-atmosphere heat exchange, *J. Geophys. Res. Atmospheres*, 120(15), 2014JD022938, doi:10.1002/2014JD022938.
- Hondzo, M., and H. G. Stefan (1993), Lake Water Temperature Simulation Model, *J. Hydraul. Eng.*, 119(11), 1251–1273, doi:10.1061/(ASCE)0733-9429(1993)119:11(1251).
- Hondzo, M., C. R. Ellis, and H. G. Stefan (1991), Vertical diffusion in small stratified lake: Data and error analysis, *J. Hydraul. Eng.*, 117(10), 1352–1369.
- Horst, G. P., O. Sarnelle, J. D. White, S. K. Hamilton, R. B. Kaul, and J. D. Bressie (2014), Nitrogen availability increases the toxin quota of a harmful cyanobacterium, *Microcystis aeruginosa*, *Water Res.*, 54, 188–198, doi:10.1016/j.watres.2014.01.063.
- Jackett, D. R., and T. J. McDougall (1995), Minimal Adjustment of Hydrographic Profiles to Achieve Static Stability, *J. Atmospheric Ocean. Technol.*, 12(2), 381–389, doi:10.1175/1520-0426(1995)012<0381:MAOHPT>2.0.CO;2.
- Jiang, L., and X. Fang (2016), Simulation and Validation of Cisco Lethal Conditions in Minnesota Lakes under Past and Future Climate Scenarios Using Constant Survival Limits, *Water*, 8(7), 279, doi:10.3390/w8070279.
- Jiang, L., M. Xia, S. A. Ludsin, E. S. Rutherford, D. M. Mason, J. Marin Jarrin, and K. L. Pangle (2015), Biophysical modeling assessment of the drivers for plankton dynamics in

- dreissenid-colonized western Lake Erie, *Ecol. Model.*, 308, 18–33, doi:10.1016/j.ecolmodel.2015.04.004.
- Jiménez, P. A., and J. Dudhia (2012), Improving the Representation of Resolved and Unresolved Topographic Effects on Surface Wind in the WRF Model, *J. Appl. Meteorol. Climatol.*, 51(2), 300–316, doi:10.1175/JAMC-D-11-084.1.
- Joodaki, G., J. Wahr, and S. Swenson (2014), Estimating the human contribution to groundwater depletion in the Middle East, from GRACE data, land surface models, and well observations, *Water Resour. Res.*, 50(3), 2679–2692, doi:10.1002/2013WR014633.
- Kaushal, S. S., G. E. Likens, N. A. Jaworski, M. L. Pace, A. M. Sides, D. Seekell, K. T. Belt, D. H. Secor, and R. L. Wingate (2010), Rising stream and river temperatures in the United States, *Front. Ecol. Environ.*, 8(9), 461–466, doi:10.1890/090037.
- Kettle, A. J., C. Hughes, G. A. Unazi, L. Birch, H. Mohie-El-Din, and M. R. Jones (2012), Role of groundwater exchange on the energy budget and seasonal stratification of a shallow temperate lake, *J. Hydrol.*, 470–471, 12–27, doi:10.1016/j.jhydrol.2012.07.004.
- Kinsman-Costello, L. E., J. M. O'Brien, and S. K. Hamilton (2015), Natural stressors in uncontaminated sediments of shallow freshwaters: The prevalence of sulfide, ammonia, and reduced iron: Sulfide, ammonia, and iron in shallow freshwater sediments, *Environ. Toxicol. Chem.*, 34(3), 467–479, doi:10.1002/etc.2801.
- Konikow, L. F. (2013), *Groundwater Depletion in the United States (1900–2008)*, U.S. Geological Survey Scientific Investigations Report 2013–5079, U.S. Geological Survey, Reston, Virginia, United States.
- Kraus, E. B., and J. A. Businger (1995), *Atmosphere-Ocean Interaction*, Oxford University Press (US), Cary, US.
- Lane, P. A. (1979), Vertebrate and invertebrate predation intensity on freshwater zooplankton communities, *Nature*, 280(5721), 391–393, doi:10.1038/280391a0.
- Lane, P. V. Z., L. Llinás, S. L. Smith, and D. Pilz (2008), Zooplankton distribution in the western Arctic during summer 2002: Hydrographic habitats and implications for food chain dynamics, *J. Mar. Syst.*, 70(1–2), 97–133, doi:10.1016/j.jmarsys.2007.04.001.
- Langston, G., M. Hayashi, and J. W. Roy (2013), Quantifying groundwater-surface water interactions in a proglacial moraine using heat and solute tracers: GW-SW Interaction in Proglacial Moraine, *Water Resour. Res.*, 49(9), 5411–5426, doi:10.1002/wrcr.20372.
- Lazzaro, D., and L. B. Montefusco (2002), Radial basis functions for the multivariate interpolation of large scattered data sets, *J. Comput. Appl. Math.*, 140(1–2), 521–536, doi:10.1016/S0377-0427(01)00485-X.
- Li, J., and A. D. Heap (2008), *A review of spatial interpolation methods for environmental scientists*, Record 2008/23, Geoscience Australia, Canberra.

- Li, L., and P. Revesz (2004), Interpolation methods for spatio-temporal geographic data, *Comput. Environ. Urban Syst.*, 28(3), 201–227, doi:10.1016/S0198-9715(03)00018-8.
- Li, L., T. Losser, C. Yorke, and R. Piltner (2014a), Fast Inverse Distance Weighting-Based Spatiotemporal Interpolation: A Web-Based Application of Interpolating Daily Fine Particulate Matter PM<sub>2.5</sub> in the Contiguous U.S. Using Parallel Programming and k-d Tree, *Int. J. Environ. Res. Public Health*, 11(9), 9101–9141, doi:10.3390/ijerph110909101.
- Li, M., L. Zhong, and W. C. Boicourt (2005), Simulations of Chesapeake Bay estuary: Sensitivity to turbulence mixing parameterizations and comparison with observations, *J. Geophys. Res.*, 110(C12), doi:10.1029/2004JC002585.
- Li, R. et al. (2014b), Observed wintertime tidal and subtidal currents over the continental shelf in the northern South China Sea, *J. Geophys. Res. Oceans*, 119(8), 5289–5310, doi:10.1002/2014JC009931.
- Lucas, A. et al. (2016), Adrift Upon a Salinity-Stratified Sea: A View of Upper-Ocean Processes in the Bay of Bengal During the Southwest Monsoon, *Oceanography*, 29(2), 134–145, doi:10.5670/oceanog.2016.46.
- Ludwig, D., S. Carpenter, and W. Brock (2003), Optimal phosphorus loading for a potentially eutrophic lake, *Ecol. Appl.*, 13(4), 1135–1152.
- Luo, L., J. Wang, D. J. Schwab, H. Vanderploeg, G. Leshkevich, X. Bai, H. Hu, and D. Wang (2012), Simulating the 1998 spring bloom in Lake Michigan using a coupled physical-biological model, *J. Geophys. Res. Oceans*, 117(C10), n/a-n/a, doi:10.1029/2012JC008216.
- Luo, W., M. C. Taylor, and S. R. Parker (2008), A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales, *Int. J. Climatol.*, 28(7), 947–959, doi:10.1002/joc.1583.
- Ma, Y., P. Niyogi, G. Sapiro, and R. Vidal (2011), Dimensionality reduction via subspace and submanifold learning, *IEEE Signal Process. Mag.*, 28(2), 14–126, doi:10.1109/MSP.2010.940005.
- van der Maaten, L. J., E. O. Postma, and H. J. van den Herik (2009), Dimensionality reduction: A comparative review, *J. Mach. Learn. Res.*, 10(1–41), 66–71.
- MacEachren, A. M., and J. V. Davidson (1987), Sampling and isometric mapping of continuous geographic surfaces, *Am. Cartogr.*, 14(4), 299–320.
- MacKay, M. D. et al. (2009), Modeling lakes and reservoirs in the climate system, *Limnol. Oceanogr.*, 54(6), 2315.
- Madala, R. V., and S. A. Piacsek (1977), A semi-implicit numerical model for baroclinic oceans, *J. Comput. Phys.*, 23(2), 167–178, doi:10.1016/0021-9991(77)90119-X.

- Magnuson, J. J. (2000), Historical Trends in Lake and River Ice Cover in the Northern Hemisphere, *Science*, 289(5485), 1743–1746, doi:10.1126/science.289.5485.1743.
- Markfort, C. D., A. L. S. Perez, J. W. Thill, D. A. Jaster, F. Porté-Agel, and H. G. Stefan (2010), Wind sheltering of a lake by a tree canopy or bluff topography, *Water Resour. Res.*, 46(3), W03530, doi:10.1029/2009WR007759.
- Martinho, A. S., and M. L. Batteen (2006), On reducing the slope parameter in terrain-following numerical ocean models, *Ocean Model.*, 13(2), 166–175, doi:10.1016/j.ocemod.2006.01.003.
- Mellard, J. P., K. Yoshiyama, E. Litchman, and C. A. Klausmeier (2011), The vertical distribution of phytoplankton in stratified water columns, *J. Theor. Biol.*, 269(1), 16–30, doi:10.1016/j.jtbi.2010.09.041.
- Mellor, G. L., and T. Yamada (1982), Development of a turbulence closure model for geophysical fluid problems, *Rev. Geophys.*, 20(4), 851–875, doi:10.1029/RG020i004p00851.
- Mellor, G. L., T. Ezer, and L.-Y. Oey (1994), The Pressure Gradient Conundrum of Sigma Coordinate Ocean Models, *J. Atmospheric Ocean. Technol.*, 11(4), 1126–1134, doi:10.1175/1520-0426(1994)011<1126:TPGCOS>2.0.CO;2.
- Mellor, G. L., L. Y. Oey, and T. Ezer (1998), Sigma coordinate pressure gradient errors and the seamount problem, *J. Atmospheric Ocean. Technol.*, 15(5), 1122–1131.
- Menberg, K., P. Blum, B. L. Kurylyk, and P. Bayer (2014), Observed groundwater temperature response to recent climate change, *Hydrol. Earth Syst. Sci.*, 18(11), 4453–4466, doi:10.5194/hess-18-4453-2014.
- Mendoza-Sanchez, I., M. S. Phanikumar, J. Niu, J. R. Masoner, I. M. Cozzarelli, and J. T. McGuire (2013), Quantifying wetland-aquifer interactions in a humid subtropical climate region: An integrated approach, *J. Hydrol.*, 498, 237–253.
- Merwade, V. (2009), Effect of spatial trends on interpolation of river bathymetry, *J. Hydrol.*, 371(1–4), 169–181, doi:10.1016/j.jhydrol.2009.03.026.
- Monaghan, G. W., D. W. Forstater, and H. Q. Sorenson (1983), *Selected geologic maps of Kalamazoo County, Michigan: County geologic map series*, Michigan Geologic Survey Division, Map.
- Moradkhani, H., S. Sorooshian, H. V. Gupta, and P. R. Houser (2005), Dual state–parameter estimation of hydrological models using ensemble Kalman filter, *Adv. Water Resour.*, 28(2), 135–147, doi:10.1016/j.advwatres.2004.09.002.
- Mordohai, P., and G. Medioni (2010), Dimensionality estimation, manifold learning and function approximation using tensor voting, *J. Mach. Learn. Res.*, 11, 411–450.

- Moss, B. (1972), Studies on Gull Lake, Michigan: I. Seasonal and depth distribution of phytoplankton, *Freshw. Biol.*, 2(4), 289–307.
- Nguyen, T. D., P. Thupaki, E. J. Anderson, and M. S. Phanikumar (2014), Summer circulation and exchange in the Saginaw Bay-Lake Huron system, *J. Geophys. Res. Oceans*, 119(4), 2713–2734, doi:10.1002/2014JC009828.
- Niu, J., C. Shen, S.-G. Li, and M. S. Phanikumar (2014), Quantifying storage changes in regional Great Lakes watersheds using a coupled subsurface-land surface process model and GRACE, MODIS products, *Water Resour. Res.*, 50(9), 7359–7377, doi:10.1002/2014WR015589.
- Niu, Q., M. Xia, E. S. Rutherford, D. M. Mason, E. J. Anderson, and D. J. Schwab (2015), Investigation of interbasin exchange and interannual variability in Lake Erie using an unstructured-grid hydrodynamic model, *J. Geophys. Res. Oceans*, 120(3), 2212–2232, doi:10.1002/2014JC010457.
- Oleson, K. W. et al. (2010), *Technical description of version 4.0 of the Community Land Model (CLM)*, NCAR Technical Note, National Center for Atmospheric Research, Boulder, Colorado.
- O'Reilly, C. M., S. R. Alin, P.-D. Plisnier, A. S. Cohen, and B. A. McKee (2003), Climate change decreases aquatic ecosystem productivity of Lake Tanganyika, Africa, *Nature*, 424(6950), 766–768, doi:10.1038/nature01833.
- Parkinson, C. L., and W. M. Washington (1979), A large-scale numerical model of sea ice, *J. Geophys. Res. Oceans*, 84(C1), 311–337, doi:10.1029/JC084iC01p00311.
- Pathiraja, S., L. Marshall, A. Sharma, and H. Moradkhani (2016), Hydrologic modeling in dynamic catchments: A data assimilation approach, *Water Resour. Res.*, n/a-n/a, doi:10.1002/2015WR017192.
- Paulson, C. A., and J. J. Simpson (1977), Irradiance Measurements in the Upper Ocean, *J. Phys. Oceanogr.*, 7(6), 952–956, doi:10.1175/1520-0485(1977)007<0952:IMITUO>2.0.CO;2.
- Perry, L., and C. J. D. Brown (1942), *A fisheries survey of Gull Lake, Kalamazoo and Barry counties*, Fisheries Research Report No. 725, Michigan Department of Natural Resources, Fisheries Division, Ann Arbor, Michigan, USA.
- Pokhrel, Y. N., S. Koirala, P. J.-F. Yeh, N. Hanasaki, L. Longuevergne, S. Kanae, and T. Oki (2015), Incorporation of groundwater pumping in a global Land Surface Model with the representation of human impacts, *Water Resour. Res.*, 51(1), 78–96, doi:10.1002/2014WR015602.
- Press, W. H. (Ed.) (2007), *Numerical recipes: the art of scientific computing*, 3rd ed., Cambridge University Press, Cambridge, UK ; New York.

- Rasmussen, A. H., M. Hondzo, and H. G. Stefan (1995), A TEST OF SEVERAL EVAPORATION EQUATIONS FOR WATER TEMPERATURE SIMULATIONS IN LAKES, *J. Am. Water Resour. Assoc.*, 31(6), 1023–1028, doi:10.1111/j.1752-1688.1995.tb03418.x.
- Renka, R. J., and R. Brown (1999), Algorithm 792: Accuracy Test of ACM Algorithms for Interpolation of Scattered Data in the Plane, *ACM Trans Math Softw*, 25(1), 78–94, doi:10.1145/305658.305745.
- Rodell, M., I. Velicogna, and J. S. Famiglietti (2009), Satellite-based estimates of groundwater depletion in India, *Nature*, 460(7258), 999–1002, doi:10.1038/nature08238.
- Rodi, W. (1987), Examples of calculation methods for flow and mixing in stratified fluids, *J. Geophys. Res. Oceans*, 92(C5), 5305–5328.
- Rowe, M. D., E. J. Anderson, J. Wang, and H. A. Vanderploeg (2015), Modeling the effect of invasive quagga mussels on the spring phytoplankton bloom in Lake Michigan, *J. Gt. Lakes Res.*, doi:10.1016/j.jglr.2014.12.018.
- Roweis, S. T., and L. K. Saul (2000), Nonlinear dimensionality reduction by locally linear embedding, *Science*, 290(5500), 2323–2326, doi:10.1126/science.290.5500.2323.
- Rueda, F. J., S. G. Schladow, S. G. Monismith, and M. T. Stacey (2005), On the effects of topography on wind and the generation of currents in a large multi-basin lake, *Hydrobiologia*, 532(1–3), 139–151.
- Rühland, K. M., A. M. Paterson, and J. P. Smol (2015), Lake diatom responses to warming: reviewing the evidence, *J. Paleolimnol.*, 54(1), 1–35, doi:10.1007/s10933-015-9837-3.
- Safaie, A., A. Wendzel, Z. Ge, M. B. Nevers, R. L. Whitman, S. R. Corsi, and M. S. Phanikumar (2016), Comparative Evaluation of Statistical and Mechanistic Models of *Escherichia coli* at Beaches in Southern Lake Michigan, *Environ. Sci. Technol.*, 50(5), 2442–2449, doi:10.1021/acs.est.5b05378.
- Safaie, A., C. Dang, H. Qiu, H. Radha, and M. S. Phanikumar (2017a), Manifold methods for assimilating geophysical and meteorological data in Earth system models and their components, *J. Hydrol.*, 544, 383–396, doi:10.1016/j.jhydrol.2016.11.009.
- Safaie, A., E. Litchman, and M. S. Phanikumar (2017b), Evaluating the Role of Groundwater in Circulation and Thermal Structure within a Deep Inland Lake, *Advances in Water Resources*, doi:10.1016/j.advwatres.2017.08.002.
- Sarnelle, O., A. E. Wilson, S. K. Hamilton, L. B. Knoll, and D. F. Raikow (2005), Complex interactions between the zebra mussel, *Dreissena polymorpha*, and the harmful phytoplankton, *Microcystis aeruginosa*, *Limnol. Oceanogr.*, 50(3), 896–904, doi:10.4319/lo.2005.50.3.0896.
- Schindler, D. W. (2009), Lakes as sentinels and integrators for the effects of climate change on watersheds, airsheds, and landscapes, *Limnol. Oceanogr.*, 54(6), 2349.

- Schladow, S. G., and D. P. Hamilton (1997), Prediction of water quality in lakes and reservoirs: Part II - Model calibration, sensitivity analysis and application, *Ecol. Model.*, 96(1–3), 111–123, doi:10.1016/S0304-3800(96)00063-4.
- Schneider, P., and S. J. Hook (2010), Space observations of inland water bodies show rapid surface warming since 1985, *Geophys. Res. Lett.*, 37(22), n/a-n/a, doi:10.1029/2010GL045059.
- Schumann, U., and T. Gerz (1995), Turbulent Mixing in Stably Stratified Shear Flows, *J. Appl. Meteorol.*, 34(1), 33–48, doi:10.1175/1520-0450-34.1.33.
- Schuster, P. F., M. M. Reddy, J. W. LaBaugh, R. S. Parkhurst, D. O. Rosenberry, T. C. Winter, R. C. Antweiler, and W. E. Dean (2003), Characterization of lake water and ground water movement in the littoral zone of Williams Lake, a closed-basin lake in north central Minnesota, *Hydrol. Process.*, 17(4), 823–838, doi:10.1002/hyp.1211.
- Schwab, D. J., and D. Beletsky (1998), *Lake Michigan Mass Balance Study: Hydrodynamic modeling project*, Great Lakes Environmental Research Laboratory, Ann Arbor, MI.
- Schwab, D. J., and J. A. Morton (1984), Estimation of Overlake Wind Speed from Overland Wind Speed: A Comparison of Three Methods, *J. Gt. Lakes Res.*, 10(1), 68–72, doi:10.1016/S0380-1330(84)71808-9.
- Shigesada, N., and A. Okubo (1981), Analysis of the self-shading effect on algal vertical distribution in natural waters, *J. Math. Biol.*, 12(3), 311–326.
- Shore, J. A. (2009), Modelling the circulation and exchange of Kingston Basin and Lake Ontario with FVCOM, *Ocean Model.*, 30(2–3), 106–114, doi:10.1016/j.ocemod.2009.06.007.
- Sikirić, M. D., I. Janeković, and M. Kuzmić (2009), A new approach to bathymetry smoothing in sigma-coordinate ocean models, *Ocean Model.*, 29(2), 128–136, doi:10.1016/j.ocemod.2009.03.009.
- Šiljeg, A., S. Lozić, and S. Šiljeg (2015), A comparison of interpolation methods on the basis of data obtained from a bathymetric survey of Lake Vrana, Croatia, *Hydrol. Earth Syst. Sci.*, 19(8), 3653–3666, doi:10.5194/hess-19-3653-2015.
- Simons, T. J. (1974), Verification of Numerical Models of Lake Ontario: Part I. Circulation in Spring and Early Summer, *J. Phys. Oceanogr.*, 4(4), 507–523, doi:10.1175/1520-0485(1974)004<0507:VONMOL>2.0.CO;2.
- Simpson, J. J., and T. D. Dickey (1981a), Alternative Parameterizations of Downward Irradiance and Their Dynamical Significance, *J. Phys. Oceanogr.*, 11(6), 876–882, doi:10.1175/1520-0485(1981)011<0876:APODIA>2.0.CO;2.
- Simpson, J. J., and T. D. Dickey (1981b), The Relationship between Downward Irradiance and Upper Ocean Structure, *J. Phys. Oceanogr.*, 11(3), 309–323, doi:10.1175/1520-0485(1981)011<0309:TRBDIA>2.0.CO;2.

- Sommaruga, R., R. Psenner, E. Schaffner, K. A. Koinig, and S. Sommaruga-Wögrath (1999), Dissolved Organic Carbon Concentration and Phytoplankton Biomass in High-Mountain Lakes of the Austrian Alps: Potential Effect of Climatic Warming on UV Underwater Attenuation, *Arct. Antarct. Alp. Res.*, *31*(3), 247, doi:10.2307/1552253.
- Sommaruga-Wögrath, S., K. A. Koinig, R. Schmidt, R. Sommaruga, R. Tessadri, and R. Psenner (1997), Temperature effects on the acidity of remote alpine lakes, *Nature*, *387*(6628), 64–67, doi:10.1038/387064a0.
- Staffell, I., and S. Pfenninger (2016), Using bias-corrected reanalysis to simulate current and future wind power output, *Energy*, *114*, 1224–1239, doi:10.1016/j.energy.2016.08.068.
- Stepanenko, V., K. D. Jöhnk, E. Machulskaya, M. Perroud, Z. Subin, A. Nordbo, I. Mammarella, and D. Mironov (2014), Simulation of surface energy fluxes and stratification of a small boreal lake by a set of one-dimensional models, *Tellus Dyn. Meteorol. Oceanogr.*, *66*(1), 21389, doi:10.3402/tellusa.v66.21389.
- Sterman, J. (2000), *Business dynamics: systems thinking and modeling for a complex world*, Irwin/McGraw-Hill, Boston.
- Suparta, W., and R. Rahman (2016), Spatial interpolation of GPS PWV and meteorological variables over the west coast of Peninsular Malaysia during 2013 Klang Valley Flash Flood, *Atmospheric Res.*, *168*, 205–219, doi:10.1016/j.atmosres.2015.09.023.
- Tague, D. F. (1977), The hydrologic and total phosphorus budgets of Gull Lake, Michigan, Michigan State University, East Lansing, Michigan, United States.
- Tenenbaum, J. B., V. De Silva, and J. C. Langford (2000), A global geometric framework for nonlinear dimensionality reduction, *Science*, *290*(5500), 2319–2323.
- Thupaki, P., M. S. Phanikumar, and R. L. Whitman (2013), Solute dispersion in the coastal boundary layer of southern Lake Michigan, *J. Geophys. Res. Oceans*, *118*(3), 1606–1617, doi:10.1002/jgrc.20136.
- Turner, J. V., and L. R. Townley (2006), Determination of groundwater flow-through regimes of shallow lakes and wetlands from numerical analysis of stable isotope and chloride tracer distribution patterns, *J. Hydrol.*, *320*(3–4), 451–483, doi:10.1016/j.jhydrol.2005.07.050.
- Umaña, G. (2014), Ten years of limnological monitoring of a modified natural lake in the tropics: Cote Lake, Costa Rica, *Rev. Biol. Trop.*, *62*(2), 576–578.
- Umlauf, L., and H. Burchard (2003), A generic length-scale equation for geophysical turbulence models, *J. Mar. Res.*, *61*(2), 235–265, doi:10.1357/002224003322005087.
- von Rohden, C., J. Ilmberger, and B. Boehrer (2009), Assessing groundwater coupling and vertical exchange in a meromictic mining lake with an SF<sub>6</sub>-tracer experiment, *J. Hydrol.*, *372*(1–4), 102–108, doi:10.1016/j.jhydrol.2009.04.004.

- Wada, Y., L. P. H. van Beek, C. M. van Kempen, J. W. T. M. Reckman, S. Vasak, and M. F. P. Bierkens (2010), Global depletion of groundwater resources: GLOBAL GROUNDWATER DEPLETION, *Geophys. Res. Lett.*, 37(20), n/a-n/a, doi:10.1029/2010GL044571.
- Webster, K. E., T. K. Kratz, C. J. Bowser, J. J. Magnuson, and W. J. Rose (1996), The influence of landscape position on lake chemical responses to drought in northern Wisconsin, *Limnol. Oceanogr.*, 41(5), 977–984.
- Weitkamp, D. E., and M. Katz (1980), A Review of Dissolved Gas Supersaturation Literature, *Trans. Am. Fish. Soc.*, 109(6), 659–702, doi:10.1577/1548-8659(1980)109<659:ARODGS>2.0.CO;2.
- White, J., and S. K. Hamilton (2014), *Gull Lake watershed monitoring Program 2014 annual report*, The Gull Lake Quality Organization, Hickory Corners, MI.
- White, J. D., S. K. Hamilton, O. Sarnelle, and K. Tierney (2015), Heat-induced mass mortality of invasive zebra mussels (*Dreissena polymorpha*) at sublethal water temperatures, *Can. J. Fish. Aquat. Sci.*, 72(8), 1221–1229, doi:10.1139/cjfas-2015-0064.
- Williamson, C. E., R. S. Stemberger, D. P. Morris, T. M. Frost, and S. G. Paulsen (1996), Ultraviolet radiation in North American lakes: attenuation estimates from DOC measurements and implications for plankton communities, *Limnol. Oceanogr.*, 41(5), 1024–1034.
- Wilson, A. E., and O. Sarnelle (2002), Relationship between zebra mussel biomass and total phosphorus in European and North American lakes, *Arch. Hydrobiol.*, 153(2).
- Wilson, M. C., J. A. Shore, and Y. R. Rao (2013), Sensitivity of the Simulated Kingston Basin—Lake Ontario Summer Temperature Profile using FVCOM, *Atmosphere-Ocean*, 51(3), 319–331, doi:10.1080/07055900.2013.800017.
- Wilson, S. E. (1987), *Bedrock Geology Map of Michigan*, Michigan Department of Environmental Quality, Geological Survey Division, Lansing, MI.
- Winslow, L. A., J. S. Read, G. J. A. Hansen, and P. C. Hanson (2015), Small lakes show muted climate change signal in deepwater temperatures, *Geophys. Res. Lett.*, 42(2), 2014GL062325, doi:10.1002/2014GL062325.
- Wollschläger, U., J. Ilmberger, M. Isenbeck-Schröter, A. M. Kreuzer, C. von Rohden, K. Roth, and W. Schäfer (2007), Coupling of groundwater and surface water at Lake Willersinnweiher: Groundwater modeling and tracer studies, *Aquat. Sci.*, 69(1), 138–152, doi:10.1007/s00027-006-0825-6.
- Woolway, R. I. et al. (2016), Diel surface temperature range scales with lake size, *PloS One*, 11(3), e0152466.

- Xue, P., D. J. Schwab, and S. Hu (2015), An investigation of the thermal response to meteorological forcing in a hydrodynamic model of Lake Superior, *J. Geophys. Res. Oceans*, 120(7), 5233–5253, doi:10.1002/2015JC010740.
- Yacobi, Y. Z., and T. Zohary (2010), Carbon:chlorophyll a ratio, assimilation numbers and turnover times of Lake Kinneret phytoplankton, *Hydrobiologia*, 639(1), 185–196, doi:10.1007/s10750-009-0023-3.
- Yan, Y., F. Xiao, and Y. Du (2014), Construction of lake bathymetry from MODIS satellite data and GIS from 2003 to 2011, *Chin. J. Oceanol. Limnol.*, 32(3), 720–731, doi:10.1007/s00343-014-3185-4.
- Yassuda, E. A., S. R. Davie, D. L. Mendelsohn, T. Isaji, and S. J. Peene (2000), Development of a waste load allocation model for the Charleston Harbor estuary, phase II: water quality, *Estuar. Coast. Shelf Sci.*, 50(1), 99–107.
- Yoshiyama, K., J. P. Mellard, E. Litchman, and C. A. Klausmeier (2009), Phytoplankton Competition for Nutrients and Light in a Stratified Water Column, *Am. Nat.*, 174(2), 190–203, doi:10.1086/600113.
- Zhang, H., I. Mendoza-Sanchez, E. L. Miller, and L. M. Abriola (2016a), Manifold Regression Framework for Characterizing Source Zone Architecture, *IEEE Trans. Geosci. Remote Sens.*, 54(1), 3–17, doi:10.1109/TGRS.2015.2448086.
- Zhang, Y., Z. Wu, M. Liu, J. He, K. Shi, Y. Zhou, M. Wang, and X. Liu (2015), Dissolved oxygen stratification and response to thermal structure and long-term climate change in a large and deep subtropical reservoir (Lake Qiandaohu, China), *Water Res.*, 75, 249–258, doi:10.1016/j.watres.2015.02.052.
- Zhang, Y., Y. Su, Z. Liu, E. Jeppesen, J. Yu, and M. Jin (2016b), Geochemical records of anoxic water mass expansion in an oligotrophic alpine lake (Yunnan Province, SW China) in response to climate warming since the 1980s, *The Holocene*, 0959683616645948.
- Zheng, L., C. Chen, and F. Y. Zhang (2004), Development of water quality model in the Satilla River Estuary, Georgia, *Ecol. Model.*, 178(3–4), 457–482, doi:10.1016/j.ecolmodel.2004.01.016.
- Zundel, A. K., and J. L. Jones (1996), *An Integrated Surface Water Modeling System*, International Association for Hydraulic Research HYDROINFORMATICS, Zurich, Switzerland.