Are the Thermoclines on Wisconsin Lakes Moving Deeper?

https://github.com/keithbollt/ENV872BadgerThermoclines.git

Keith Bollt

Abstract

For this project, I was interested in determining if thermoclines in five Wisconsin lakes are trending deeper over time, possibly due to climate change. I looked at physical and chemical data from the NTLR dataset, specifically temperature and dissolved oxygen levels at 7 meters depth between June 20 and September 21 from every year data was available between 1984 and 2016. Initial data visualization did not promise significant trends that would point to climate change. I ran four total tests on the dataset: a repeated measures ANOVA on the combined dataset for both variables, and a seasonal Mann Kendall on each of the five lakes for both variables. I found that there was not a significant trend in temperature or oxygen on the five lakes as a whole, and that no individual lake exhibited both a significant increase in temperature and a significant decrease in dissolved oxygen at 7 meters depth. In my conclusion, I discuss that for incremental climate change, the thermocline may not move because it is a result of a lake's relative physical conditions. Further tests would likely reveal that the thermocline is steeper.

*

*Table of Contents**

List of Tables

List of Figures

Setting Up RStudio for Analysis

```
#Setting up session and loading all packages I think I might need
getwd()
```

```
## [1] "V:/ENV872BadgerThermoclines"
```

```
library(tidyverse)
```

```
## -- Attaching packages -----------------------------------------------------------------

## v ggplot2 3.1.0     v purrr   0.2.5
## v tibble  1.4.2     v dplyr   0.7.8
## v tidyr   0.8.2     v stringr 1.3.1
## v readr   1.3.1     v forcats 0.3.0

## -- Conflicts --------------------------------------------------------------------------
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':
##
##     date
```

```
library(ggplot2)
library(multcompView)
library(nlme)
```

```
##
## Attaching package: 'nlme'

## The following object is masked from 'package:dplyr':
##
##     collapse
```

```
library(lsmeans)
```

```
## Loading required package: emmeans

## The 'lsmeans' package is now basically a front end for 'emmeans'.
## Users are encouraged to switch the rest of the way.
## See help('transition') for more information, including how to
## convert old 'lsmeans' objects and scripts to work with 'emmeans'.
```

```
#install.packages("trend")
library(trend)
```

```r
#setting a theme for plots
mytheme <- theme_classic(base_size = 12.2696)+
  theme(axis.text = element_text(color = "Blue"),
  legend.position = "top")
  plot.title = element_text(hjust = 0.5)
theme_set(mytheme)
```

```r
#Reading in the raw lake dataset
NTLR_raw <-
  read.csv("V:/ENV_872_Project_Directory/Data/Raw/NTL-LTER_Lake_ChemistryPhysics_Raw.csv
View(NTLR_raw)

#Changing the date column to read as a date
class(NTLR_raw$sampledate)
NTLR_raw$sampledate <- as.Date(NTLR_raw$sampledate, format="%m/%d/%y")
class(NTLR_raw$sampledate)
```

*

*Research Question and Rationale**

The crux of this project is to determine if the thermoclines on each of the 9 Wisconsin lakes in the NTL-LTER chemistry and physics dataset have moved over the course of the 33 years of data between 1984 and 2016. I am interested in this question because as a flyfisherman, I have an interest in coldwater fisheries that rely on thermoclines to survive the summer weather. Perhaps climate change is affecting where the thermocline sets up, and therefore shrinking the available summer habitat of trout.

In order to determine whether thermoclines on these lakes are moving deeper in the water column, I need to set a benchmark definition for a thermocline. It looks like there is not enough temporal resolution to measure close-to-continuous change in thermocline depth over the course of a given season. Likewise, there is not enough close-to-continous depth measurements taken at each lake, nor is there consistant data taken below 10 meters of depth. As such, I will compare what is occurring at a constant depth near the expected thermocline location at each lake over time. The depth that I chose for the purpose of this study is 7 meters. 7 meters, I know from my experience as a fisherman, is on the shallow end of where a thermocline sets up in a northern US lake. In addition, there is enough temporal resolution at seven meters to perform statistical analysis on five of the lakes in the raw dataset. Therefore, evaluating what is happening at 7m depth in each lake will give a good idea what sorts of conditions trout are dealing with in these lakes in the summmer. I know from data visualization that did not make my final report that there is much more variation by year at 7 meters depth than at, say, 10 meters depth. This indicates that 7 meters is indeed a pretty good estimation of thermocline location for these lakes. Looking for change at 7 meters should give me a good picture of whether the thermocline in these lakes is changing over time.

I will look at two indicators of thermocline establishment at 7 meters: temperature and dissolved oxygen content. I would expect water at or below the thermocline to have low temperatures and hypoxic conditions (low oxygen levels). If the thermoclines on these lakes are trending deeper over the course of my dataset, I would expect most of the lakes to show increasing temperatures and increasing dissolved oxygen levels at a 7 meter depth.

My research question, then, is as follows:

How have temperature and oxygen conditions changed at 7 meters depth in a series of Wisconsin Lakes? Is climate change affecting where the thermocline sets up in these lakes?

*

*Dataset Information**

The dataset used for this project was prepared for Environmental Data Analytics (ENV 872L) at Duke University's Nicholas School of the Environment for the spring 2019 semester by Professor Kateri Salk.The dataset contains physical and chemical data from nine lakes in Wisconsin. The data was collected between 1984 and 2016 as part of the NSF-funded North Temperate Lakes Long Term Ecological Research Station. The data is measured at a station in the middle of each of the nine lakes. Generally, the data was collected in the morning. The data is taken at increments, and most data is taken at or below ten meters depth. The temporal resolution varies across lakes. Some lakes only have a few years of data, while others have ample data from most or all years.Data was collected at irregular 1 to 7 day increments from May through August of each year. The data was taken during periods of no ice on the lake. This means that there is little data during the winter months. While there are several different measurements taken at these lakes, the two variables that the thermocline study focuses on are dissolved oxygen and temperature. A table summarizing the data structure of this project is provided at the end of this section. Data collection techniques used during the period 1984-1990 were described by Carpenter and Kitchell (1993) and data collection techniques for 1991-1997 were described by Carpenter et al. (2001).

Carpenter, S.R. and J.F. Kitchell (eds.). 1993. The Trophic Cascade in Lakes. Cambridge University Press, Cambridge, England.

Carpenter, S.R., J.J. Cole, J.R. Hodgson, J.F. Kitchell, M.L. Pace,D. Bade, K.L. Cottingham, T.E. Essington, J.N. Houser and D.E. Schindler. 2001. Trophic cascades, nutrients and lake productivity: whole-lake experiments. Ecological Monographs 71: 163-186."

Table 1: Table Summarizing Project Data Structure:

Element	Details
Data Source	North Temperate Lakes Long Term Ecological Research Station
Data Scraper	Professor Kateri Salk, PhD, Duke University; kateri.salk@duke.edu
More Information about Data	https://lter.limnology.wisc.edu/about/overview
Link to Access Data	https://lter.limnology.wisc.edu/data
Dimensions of Raw Dataset	38,614 Observations of 11 variables over nine lakes
Dimensions of Processed Dataset	1,116 Observations of 8 variables over five lakes
Variables in Processed Dataset	lakename (lake name), year4 (year), daynum (Julian day number), sampledate (date of observation), depth of observation (in meters), temperature_C (lake temperature in degrees Celsius), dissolved Oxygen (percent dissolved Oxygen), Week (numbered week of the year)

Element	Details
Temporal Resolution of Processed Dataset	Irregular collection intervals from June 20 to September 21, 1984-2016

Exploratory Data Analysis Wrangling

In order to answer my research question, I need to wrangle my data to only include the data relevant to my study and to organize it by lake. I only want data during the summer months (June 20 through September 21) because this is the time of year when the thermocline is the steepest and coldwater fish are the most thermally stressed. I also only want data collected at a 7 meter depth. My first step is to filter my raw data to generate a dataset for each of the nine lakes.This data summary gives me the names of the nine lakes.

```
unique(NTLR_raw$lakename)
Paullake_raw <- NTLR_raw %>%
  filter(lakename == "Paul Lake")
Peterlake_raw <- NTLR_raw %>%
  filter(lakename == "Peter Lake")
Tuesdaylake_raw <- NTLR_raw %>%
  filter(lakename == "Tuesday Lake")
Eastlonglake_raw <- NTLR_raw %>%
  filter(lakename == "East Long Lake")
Westlonglake_raw <- NTLR_raw %>%
  filter(lakename == "West Long Lake")
Centrallonglake_raw <- NTLR_raw %>%
  filter(lakename == "Central Long Lake")
Hummingbirdlake_raw <- NTLR_raw %>%
  filter(lakename == "Hummingbird Lake")
Cramptonlake_raw <- NTLR_raw %>%
  filter(lakename == "Crampton Lake")
Wardlake_raw <- NTLR_raw %>%
  filter(lakename == "Ward Lake")
```

My next step is to filter my dataset to only include data collected at 7 meters depths between June 20 and September 21. I also checked the dimensions of my wrangled data to make sure each lake still had enough datapoints to perform statistical analysis on.

```
Paullake_processed <- Paullake_raw %>%
  filter(depth == 7, daynum %in% 172:264) %>%
  select(lakename:dissolvedOxygen) %>%
  mutate(Week = week(sampledate)) %>%
  na.exclude()
dim(Paullake_processed)

Peterlake_processed <- Peterlake_raw %>%
  filter(depth == 7, daynum %in% 172:264)%>%
  select(lakename:dissolvedOxygen) %>%
  mutate(Week = week(sampledate)) %>%
  na.exclude()
dim(Peterlake_processed)
```

```r
#Enough datapoints to analyse this lake.

Tuesdaylake_processed <- Tuesdaylake_raw %>%
  filter(depth == 7, daynum %in% 172:264)%>%
  select(lakename:dissolvedOxygen) %>%
  mutate(Week = week(sampledate)) %>%
  na.exclude()
dim(Tuesdaylake_processed) #Enough datapoints to analyse this lake.

Eastlonglake_processed <- Eastlonglake_raw %>%
  filter(depth == 7, daynum %in% 172:264)%>%
  select(lakename:dissolvedOxygen) %>%
  mutate(Week = week(sampledate)) %>%
  na.exclude()
dim(Eastlonglake_processed)
#Enough datapoints to analyse this lake.

Westlonglake_processed <- Westlonglake_raw %>%
  filter(depth == 7, daynum %in% 172:264)%>%
  select(lakename:dissolvedOxygen) %>%
  mutate(Week = week(sampledate)) %>%
  na.exclude()
dim(Westlonglake_processed)
#Enough datapoints to analyse this lake.

Centrallonglake_processed <- Centrallonglake_raw %>%
  filter(depth == 7, daynum %in% 172:264)%>%
  select(lakename:dissolvedOxygen) %>%
  mutate(Week = week(sampledate)) %>%
  na.exclude()
dim(Centrallonglake_processed)
#There are 0 data points taken on Central Long Lake below 4 meters.

Hummingbirdlake_processed <- Hummingbirdlake_raw %>%
  filter(depth == 7, daynum %in% 172:264)%>%
  select(lakename:dissolvedOxygen) %>%
  mutate(Week = week(sampledate)) %>%
  na.exclude()
dim(Hummingbirdlake_processed)
#Hummingbird lake only has one datapoint at 7 meters.

Cramptonlake_processed <- Cramptonlake_raw %>%
  filter(depth == 7, daynum %in% 172:264)%>%
  select(lakename:dissolvedOxygen) %>%
```

```
  mutate(Week = week(sampledate)) %>%
  na.exclude()
dim(Cramptonlake_processed)
#There are only 34 datapoints, not enough datapoints to analyse this lake.

Wardlake_processed <- Wardlake_raw %>%
  filter(depth == 7, daynum %in% 172:264)%>%
  select(lakename:dissolvedOxygen) %>%
  mutate(Week = week(sampledate)) %>%
  na.exclude()
dim(Wardlake_processed)
#Ward lake only has 11 datapoints at 7 meters.
```

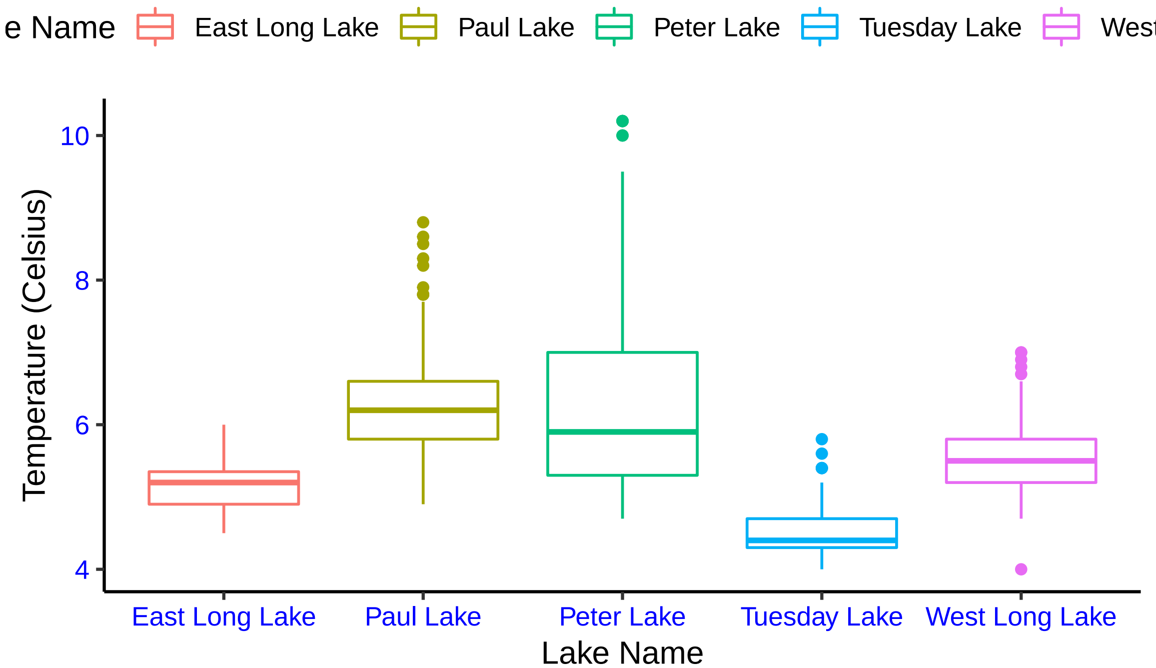
After deciding on what question I wanted to answer and what data I wanted to use to answer my question, I unfortunately had to eliminate four lakes from my analysis. I still have five lakes to perform my analysis on. I also combined my five processed datasets into one dataframe for part of my statistical analysis. My first series of statistical tests will look at the combined dataset and my second series of statistical tests will look at each lake individually.

```
NTL_processed <- rbind(Paullake_processed, Peterlake_processed,
                        Tuesdaylake_processed,
                       Eastlonglake_processed, Westlonglake_processed)
View(NTL_processed)
unique(NTL_processed$lakename)
dim(NTL_processed)
head(NTL_processed)
```

Now, I will perform a little data visualization before my data analysis. I hope to get a feel for the data I am working with, and maybe anticipate the types of results I will find when I perform statistical analyses in the next section of this report.

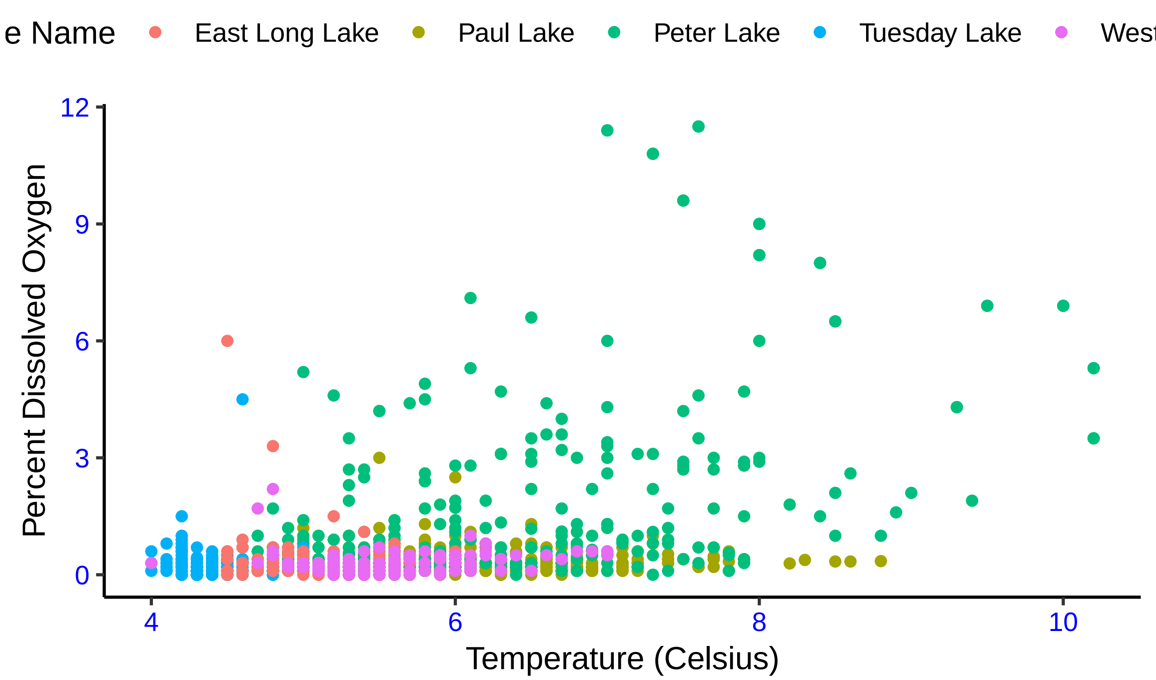
See Figure 1 below, which graphs the temperature data distribution for each of the five lakes at 7m depth. This figure tells us which lakes are coldest in the summer.

Figure 1: Temperature Distribution at 7 Meters

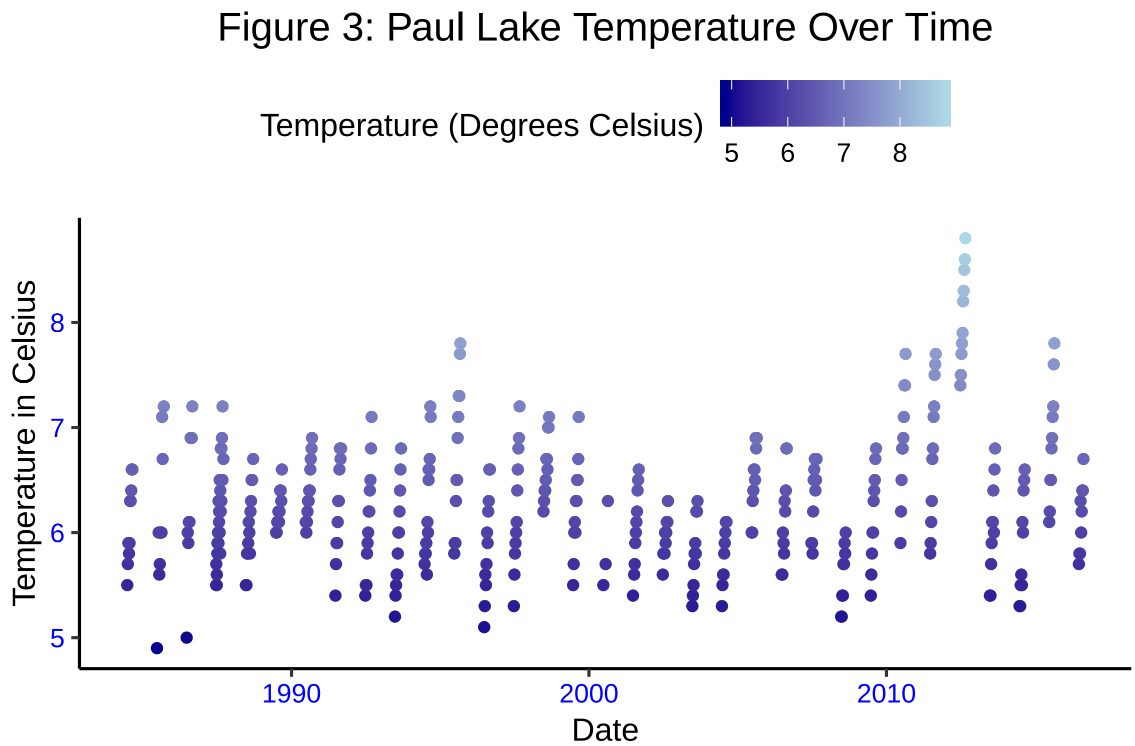


See in Figure 2: "Temperature versus Dissolved Oxygen" that there is no real correlation between temperature and dissolved oxygen at 7 meters depth across the five lakes.

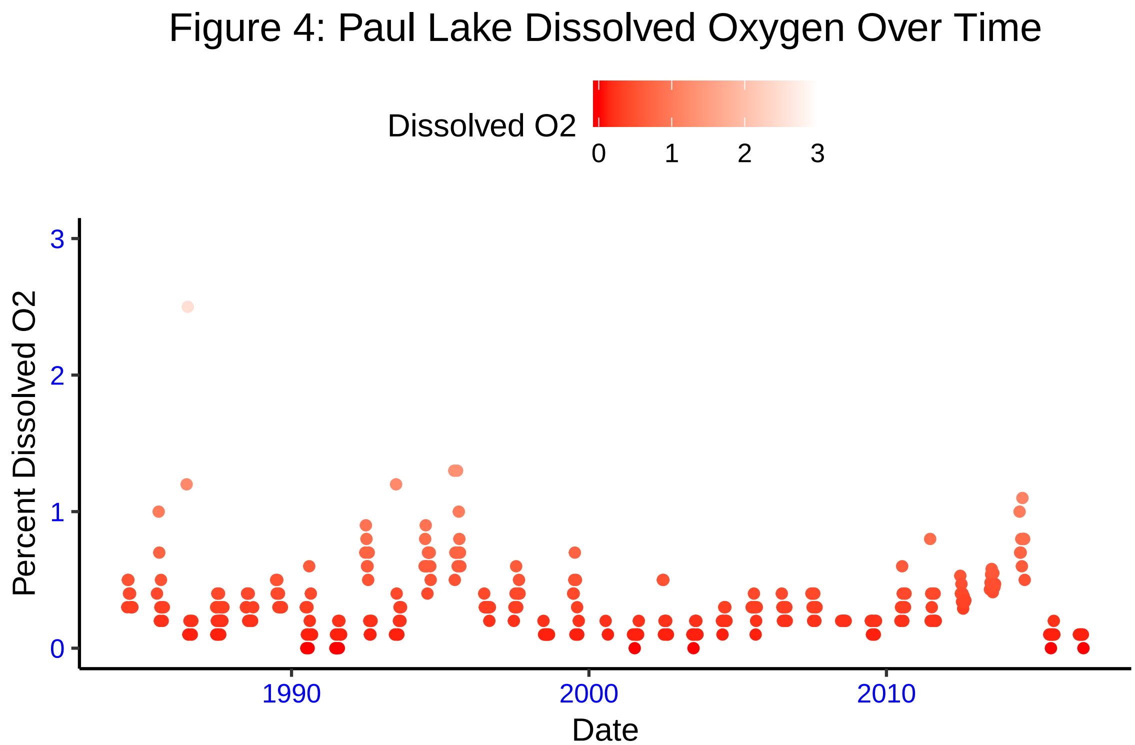
Figure 2: Temperature vs Dissolved Oxygen



See in Figure 3: "Paul Lake Temperature Over Time" there is not really much of a trend in temperature in Paul Lake at 7 m depth over time.



See in Figure 4: "Paul Lake Dissolved Oxygen Over Time" there is not really much of a trend in dissolved oxygen in Paul Lake at 7 m depth over time.



Overall, my data visualization is not too promising yet. Paul Lake does not seem to have a trend in either temperature or dissolved oxygen over time. Likewise, the combined dataset does not appear to have a trend in temperature over time. However, just because there is not a trend visible to the naked eye does not mean that a trend does not exist. I will run a series of statistical analyses on the combined dataset and on each of the five lakes individually to try and tease out a relationship that would indicate that thermocline location is moving in response to climate change.

*

*Analysis**

Now that I have visualized my data, it is time to start running statistical tests on it in order to answer my research question. I am interested in two parameters at 7 meters depth, temperature and dissolved oxygen content. First, I will perform two repeated measures ANOVAs on the combined processed dataset. I will first run the test on temperature and then on dissolved oxygen. These two tests takes into account autocorrelation within a given year and within a given lake.

```
#Accounting for autocorrelation
Alllakestemptest.auto <- lme(data = NTL_processed,
                    temperature_C ~ sampledate * lakename, #fixed effects portion
                    random = ~1|Week)  # this is the random effect portion
Alllakestemptest.auto
# we care about the Stddeviation between each week
ACF(Alllakestemptest.auto)
# we care about the lag of 1's value. This tells us how much temperature
#is autocorrelated within a given year (it's about 23%)

#running the ANOVA
Alllakestemptest.mixed <- lme(data = NTL_processed,
                 temperature_C ~ sampledate * lakename,
                 random = ~1|Week,
            correlation = corAR1(form = ~ sampledate/lakename|Week, value = 0.2323),
                 method = "REML")
summary(Alllakestemptest.mixed)
#There is not a significant trend among all of the lakes at 7m.
```

The results from our mixed effects test demonstrates an important finding. Observe that the p-value for sampledate is 0.94. This means that we can reject the null hypothesis that there is a significant linear correlation between date and temperature of the combined dataset at 7 meters. This does not supports the idea that the temperature at 7 meters in these lakes as a group is changing.

```
#oxygen may be autocorrelated with temperature
#copy from above, and plug in the oxygen data
#Accounting for autocorrelation
Alllakesoxygentest.auto <- lme(data = NTL_processed,
                    dissolvedOxygen ~ sampledate * lakename,
                    random = ~1|Week)
Alllakesoxygentest.auto
# we care about the Stddeviation between each week
ACF(Alllakesoxygentest.auto)
# autocorrelation with temperature is about 6%
```

```r
#running the ANOVA
Alllakesoxygentest.mixed <- lme(data = NTL_processed,
                     dissolvedOxygen ~ sampledate * lakename,
                     random = ~1|Week,
                     correlation =
                corAR1(form = ~ sampledate/lakename|Week, value = 0.0645),
                     method = "REML")
summary(Alllakesoxygentest.mixed)
#There is not a significant trend among all of the lakes at 7m.
```

The results from our mixed effects test demonstrates an important finding. Observe that the p-value for sampledate is 0.43. This means that we can reject the null hypothesis that there is a significant linear correlation between date and dissolved oxygen content of combined dataset at 7 meters. This does not supports the idea that the dissolved oxygen at 7 meters in these lakes as a group is changing. The results of the two repeated measures ANOVA tests are summarized in Table 2 below.

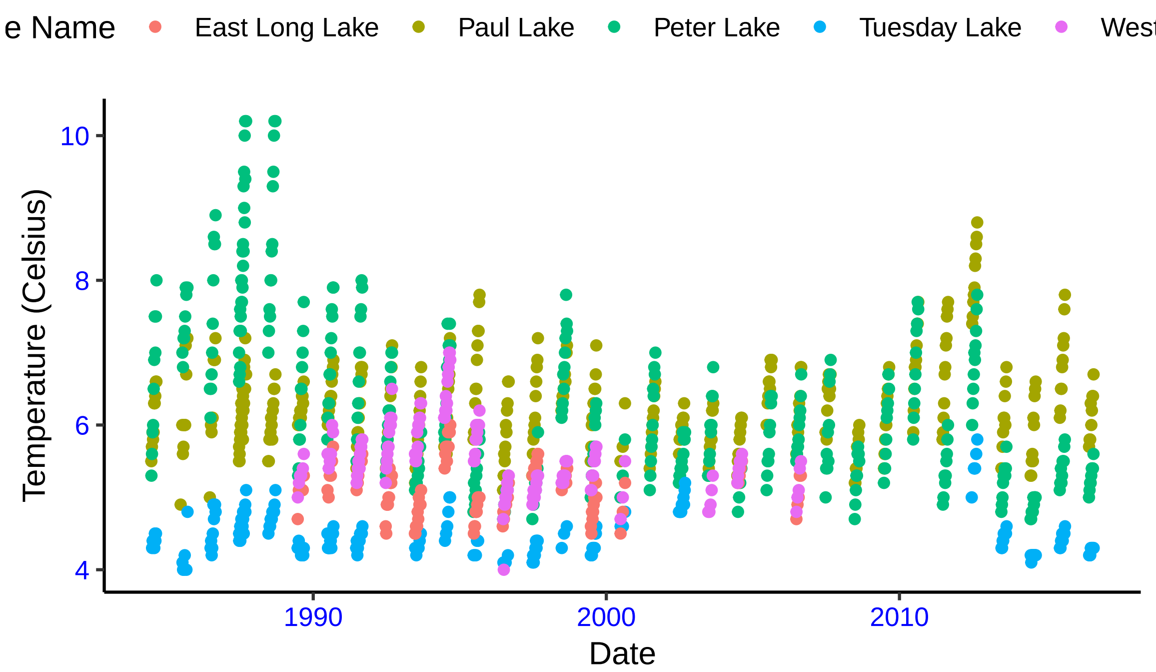
Table 2: Summary of ANOVA Results

Test Run	Variable Measured	Trend	Significance of Trend
repeated measures ANOVA on combined five-lake datset	dissolved oxygen	negative	not significant
repeated measures ANOVA on combined five-lake datset	temperature	negative	not significant

Let's visualize our data for both temperature and oxygen over time in all five lakes.

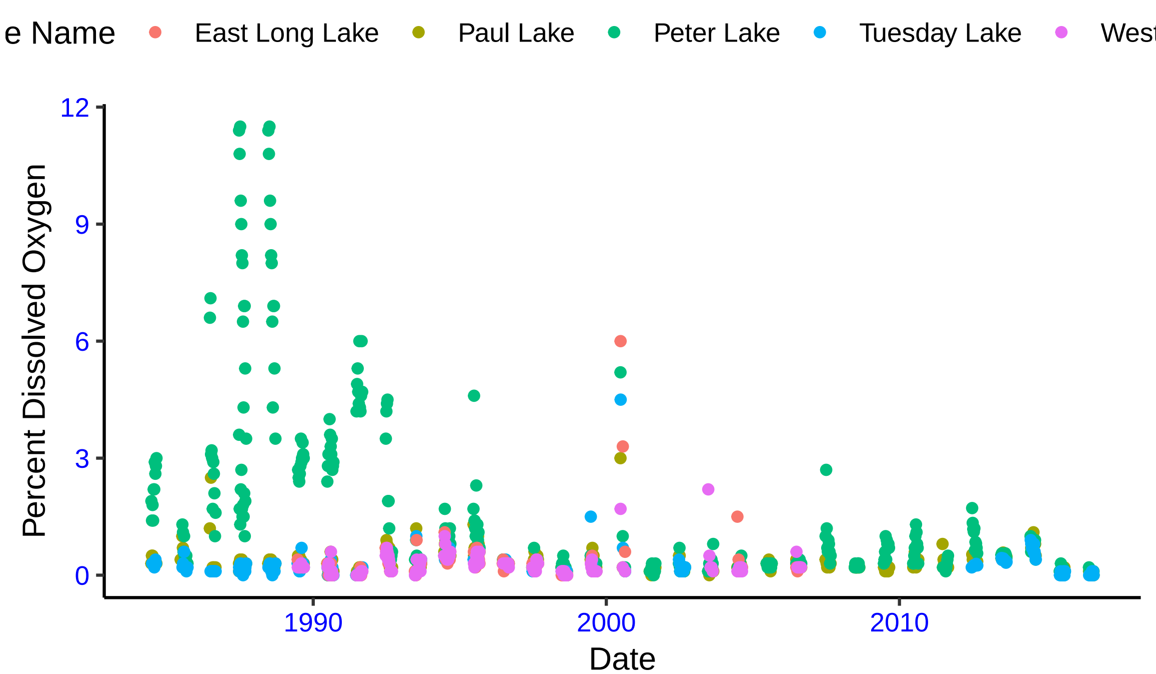
See in Figure 5: "Temperature at 7m of All Lakes Over Time" that it is difficult to make out a temperature trend in the combined dataset.

Figure 5: Temperature at 7m of All Lakes Over Time



Likewise, observe in Figure 6: "Dissolved Oxygen at 7m of All Lakes Over Time" that it is difficult to make out a dissolved oxygen trend in the combined dataset.

Figure 6: Dissolved Oxygen at 7m of All Lakes Over Time



While we have rejected the null hypothesis that there is a significant linear trend between the five lakes for either temperature or dissolved oxygen, we can still look at trends in the individual lakes. In order to do this, a seasonal Mann-Kendall test is appropriate. I will run 10 seasonal Mann Kendall tests (five lakes and two parameters per lake). I will set each year's summer as its own season.

```
#Paul Lake
length(unique(Paullake_processed$year4))
#Tells me how many summers are in the dataset.
Paullaketemp_ts <- ts(Paullake_processed$temperature_C,
                      start = c(1984), frequency = 33)
Paullaketemp_ts
Paul_temp_smk <- smk.test(Paullaketemp_ts)
Paul_temp_smk
summary(Paul_temp_smk)
#The seasonal Mann Kendall test for temperature at Paul Lake had an
#overall z-score of 2.2304, an overall p-value of 0.02572 and an
#overall S value of 159. This test shows a significant positive
#temperature trend over time at Paul Lake.

Paullakeo2_ts <- ts(Paullake_processed$dissolvedOxygen,
                      start = c(1984), frequency = 33)
Paullakeo2_ts
Paul_o2_smk <- smk.test(Paullakeo2_ts)
Paul_o2_smk
summary(Paul_o2_smk) #The seasonal Mann Kendall test for dissolved oxygen
#at Paul Lake had an overall z-score of -0.434, an overall p-value of 0.66
#and an overall S value of -31. This test shows a nonsignificant negative
#dissolved oxygen trend over time at Paul Lake.

#Peter Lake
length(unique(Peterlake_processed$year4))
#Tells me how many summers are in the dataset.
Peterlaketemp_ts <- ts(Peterlake_processed$temperature_C,
                      start = c(1984), frequency = 33)
Peterlaketemp_ts
Peter_temp_smk <- smk.test(Peterlaketemp_ts)
Peter_temp_smk
summary(Peter_temp_smk) #The seasonal Mann Kendall test for temperature
#at Peter Lake had an overall z-score of -9.637, an overall p-value of
#<2.2x10^-16 and an overall S value of -676. This test shows a significant
#negative temperature trend over time at Peter Lake.

Peterlakeo2_ts <- ts(Peterlake_processed$dissolvedOxygen,
                      start = c(1984), frequency = 33)
Peterlakeo2_ts
```

```
Peter_o2_smk <- smk.test(Peterlakeo2_ts)
Peter_o2_smk
summary(Peter_o2_smk) #The seasonal Mann Kendall test for dissolved oxygen
#at Peter Lake had an overall z-score of -10.557, an overall p-value of
#<2.2x10^-16 and an overall S value of -738. This test shows a significant
#negative dissolved oxygen trend over time at Peter Lake.

#Tuesday Lake
unique(Tuesdaylake_processed$year4) # There are too many gaps in the data,
#including a 10 year gap from 2002 to 2012 (which is 30% of the length of
#the timeseries), for me to interpolate the whole dataset for purposes of
#analysis. I will look at 1984 to 1991, the longest set of continous
#data collection.
Tuesday_smk <- Tuesdaylake_processed %>%
  filter(year4 >= 1984, year4 <= 1991)
Tuesdaylaketemp_ts <- ts(Tuesday_smk$temperature_C,
                      start = c(1984), frequency = 8)
Tuesdaylaketemp_ts
Tuesday_temp_smk <- smk.test(Tuesdaylaketemp_ts)
Tuesday_temp_smk
summary(Tuesday_temp_smk) #The seasonal Mann Kendall test for temperature at
#Tuesday Lake had an overall z-score of -0.56, an overall p-value of 0.57 and
#an overall S value of -23. This test shows a nonsignificant negative
#temperature trend over time at Tuesday Lake.

Tuesdaylakeo2_ts <- ts(Tuesday_smk$dissolvedOxygen,
                      start = c(1984), frequency = 8)
Tuesdaylakeo2_ts
Tuesday_o2_smk <- smk.test(Tuesdaylakeo2_ts)
Tuesday_o2_smk
summary(Tuesday_o2_smk) #The seasonal Mann Kendall test for dissolved
#oxygen at Tuesday Lake had an overall z-score of -3.39, an overall
#p-value of 0.0007 and an overall S value of -128. This test shows a
#significant negative dissolved oxygen trend over time at Tuesday Lake.

#East Long Lake
unique(Eastlonglake_processed$year4) #There are too many gaps in the data,
#including no data after 2006 (which is 30% of the length of the timeseries),
#for me to interpolate the whole dataset for purposes of analysis. I will
#look at 1989 to 2000, the longest set of continous data collection.
Eastlong_smk <- Eastlonglake_processed %>%
  filter(year4 >= 1989, year4 <= 2000)
Eastlonglaketemp_ts <- ts(Eastlong_smk$temperature_C,
                      start = c(1989), frequency = 12)
```

```r
Eastlonglaketemp_ts
Eastlong_temp_smk <- smk.test(Eastlonglaketemp_ts)
Eastlong_temp_smk
summary(Eastlong_temp_smk) #The seasonal Mann Kendall test for temperature
#at East Long Lake had an overall z-score of -0.84, an overall p-value
#of 0.40 and an overall S value of -32. This test shows a nonsignificant
#negative temperature trend over time at East Long Lake.

Eastlonglakeo2_ts <- ts(Eastlong_smk$dissolvedOxygen,
                      start = c(1989), frequency = 12)
Eastlonglakeo2_ts
Eastlong_o2_smk <- smk.test(Eastlonglakeo2_ts)
Eastlong_o2_smk
summary(Eastlong_o2_smk) #The seasonal Mann Kendall test for dissolved
#oxygen at East Long Lake had an overall z-score of 0.83, an overall
#p-value of 0.41 and an overall S value of 31. This test shows a
#nonsignificant positive dissolved oxygen trend over time at East Long Lake.

#West Long Lake
unique(Westlonglake_processed$year4) #There are too many gaps in the data,
#including no data after 2006 (which is 30% of the length of the timeseries),
#for me to interpolate the whole dataset for purposes of analysis. I will
#look at 1989 to 2000, the longest set of continous data collection.
Westlong_smk <- Westlonglake_processed %>%
  filter(year4 >= 1989, year4 <= 2000)
Westlonglaketemp_ts <- ts(Westlong_smk$temperature_C,
                      start = c(1989), frequency = 12)
Westlonglaketemp_ts
Westlong_temp_smk <- smk.test(Westlonglaketemp_ts)
Westlong_temp_smk
summary(Westlong_temp_smk) #The seasonal Mann Kendall test for temperature
#at West Long Lake had an overall z-score of -1.65, an overall p-value of
#0.10 and an overall S value of -62. This test shows a nonsignificant
#negative temperature trend over time at West Long Lake.

Westlonglakeo2_ts <- ts(Westlong_smk$dissolvedOxygen,
                      start = c(1989), frequency = 12)
Westlonglakeo2_ts
Westlong_o2_smk <- smk.test(Westlonglakeo2_ts)
Westlong_o2_smk
summary(Westlong_o2_smk) #The seasonal Mann Kendall test for dissolved
#oxygen at West Long Lake had an overall z-score of -0.47, an overall
#p-value of 0.64 and an overall S value of -18. This test shows a
#nonsignificant negative dissolved oxygen trend over time at West Long Lake.
```

My data does not appear to be showing a significant change in the thermocline of any of the five individual lakes, as defined by both higher temperatures and lower oxygen levels at 7 meters depth. The results of these tests are summarized in the table below.

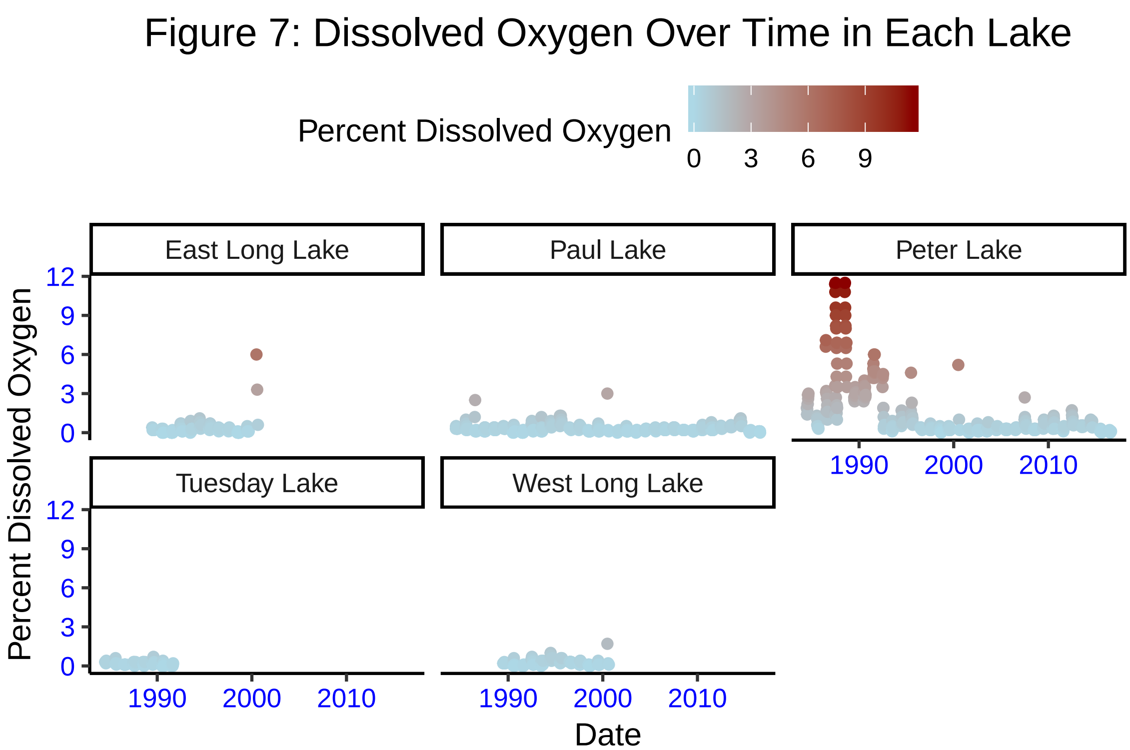
Table 3: Summary of Seasonal Mann Kendall results

test run	variable measured	trend	significance of trend
seasonal Mann Kendall on Paul Lake	temperature	positive	significant
seasonal Mann Kendall on Paul Lake	dissolved oxygen	negative	nonsignificant
seasonal Mann Kendall on Peter Lake	temperature	negative	significant
seasonal Mann Kendall on Peter Lake	dissolved oxygen	negative	significant
seasonal Mann Kendall on Tuesday Lake	temperature	negative	nonsignificant
seasonal Mann Kendall on Tuesday Lake	dissolved oxygen	negative	significant
seasonal Mann Kendall on East Long Lake	temperature	negative	nonsignificant
seasonal Mann Kendall on East Long Lake	dissolved oxygen	positive	significant

test run	variable measured	trend	significance of trend
seasonal Mann Kendall on West Long Lake	temperature	negative	nonsignificant
seasonal Mann Kendall on West Long Lake	dissolved oxygen	negative	nonsignificant

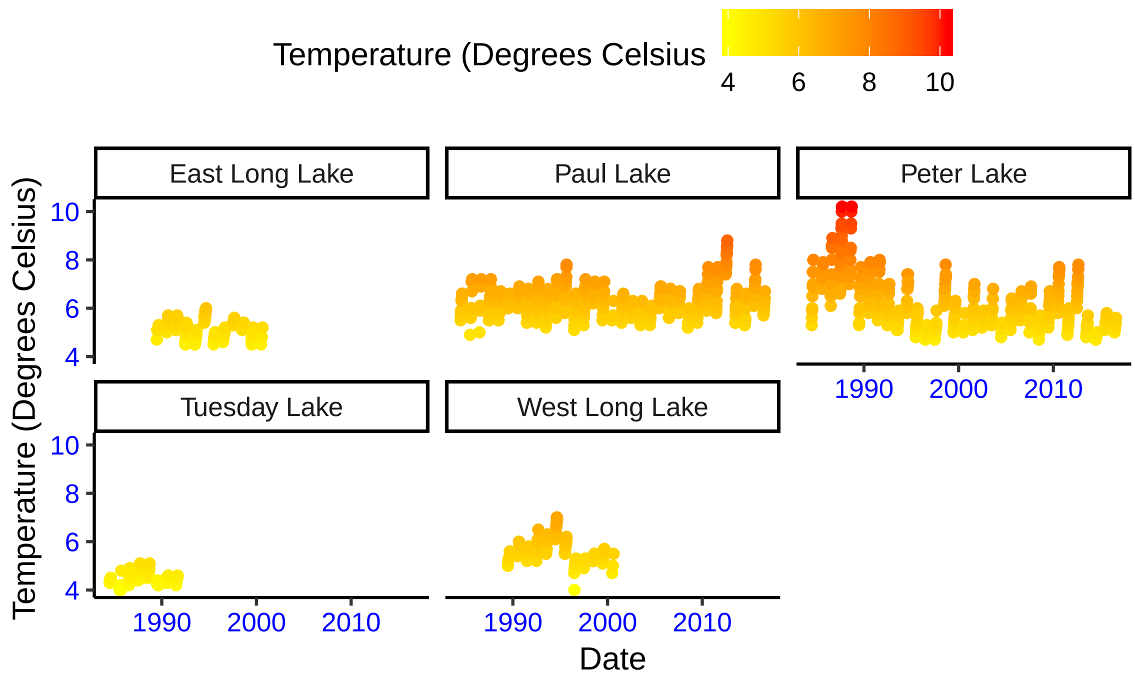
As the above table demonstrates, only two of the ten results (20%) are both significant and in the predicted direction (positive temperature trend or negative dissolved oxygen trend); no one lake has both conditions met. This suggests that the thermoclines on these lakes are not moving. Let's visualize the temperature and oxygen over the period of the ten tests I just ran to see if the naked eye agrees with my statistical analysis.

See in Figure 7: "Dissolved Oxygen Over Time in Each Lake" that the seasonal Mann Kendall results for the five lakes do not show an overall dissolved oxygen trend at 7 meters in a particular direction.



See in Figure 8: "Temperature Over Time in Each Lake" that the seasonal Mann Kendall results for the five lakes do not show an overall temperature trend at 7 meters in a particular direction.

Figure 8: Temperature Over Time in Each Lake



*

*Discussion** I used repeated measures ANOVA and seasonal Mann Kendall statistical evaluations of trends in dissolved oxygen and temperature in five Wisconsin lakes to determine if the thermoclines in these lakes are moving deeper over time. I found that neither the five lakes in combination nor any of the five lakes individually are observing both a significant positive temperature trend and a significant negative dissolved oxygen trend. The results of my tests seem to run contrary to the idea that climate change is affecting the water temperatures of the lakes in this study. However, it may be the case that my tests are inconclusive by design. I am testing whether the thermocline on these lakes is moving over time. The location of a lake's thermocline is a product of water chemistry and physics, especially water density. It may be the case that the basic physics of water that determine where a thermocline sets up are not affected by 1 °C of global warming. Perhaps, given relatively modest warming, it is the relative water density that determines where the thermocline location sets up. It may be the case that the steepness and not the location of the thermocline is what is changing under relatively small amounts of warming. If I were to test surface water temperature or water temperature at, say, 15 meters, I would likely see evidence of climate change. It might require larger levels of warming for the thermocline itself to also move. Unfortunately, testing this hypothesis is beyond the scope of the dataset I am analyzing.

Appendix: Project README File

1 badger-thermoclines repository information

This respository was created as part of a final project for ENVIRON 872 at Duke University. I am looking at NTLR data to determine thermocline movement on 9 Wisconsin lakes between the period 1984 and 2016. There are potentially biological ramifications from a shrinking coldwater summer refuge for organisms such as trout if in fact thermoclines are setting up at lower levels.

1.1 Summary

The dataset used for this project was prepared for Environmental Data Analytics (ENV 872L) at Duke University's Nicholas School of the Environment for the spring 2019 semester by Kateri Salk.

The dataset contains physical and chemical data from nine lakes in Wisconsin. The data was collected between 1984 and 2016 as part of the NSF-funded North Temperate Lakes Long Term Ecological Research Station.

1.2 Database Information

To find out more information on the North Temperate Lakes Long Term Ecological Research Station, visit https://lter.limnology.wisc.edu/about/overview

The following tool was used to acquire the NTLER data for this project: (https://lter.limnology.wisc.edu/data)

At the above link, the following data selections were made: * Cascade Project at North Temperate Lakes LTER Core Data Physical and Chemical Limnology 1984 - 2016 * Download All Data (csv)

Data was accessed on December 6, 2018.

1.2.1 Physical and chemical limnology

The data is measured at a station in the middle of each of the nine lakes. Generally, the data was collected in the morning. The data is taken at increments, and most data is taken at or below ten meters depth. The temporal resolution varies across lakes. Some lakes only have a few years of data, while others have ample data from most or all years. The data was taken during periods of no ice on the lake. This means that there is little data during the winter months. While there are several different measurements taken at these lakes, the two variables that the thermocline study focuses on are dissolved oxygen and temperature.

Data collection techniques used during the period 1984-1990 were described by Carpenter and Kitchell (1993) and data collection techniques for 1991-1997 were described by Carpenter et al. (2001).

Carpenter, S.R. and J.F. Kitchell (eds.). 1993. The Trophic Cascade in Lakes. Cambridge University Press, Cambridge, England.

Carpenter, S.R., J.J. Cole, J.R. Hodgson, J.F. Kitchell, M.L. Pace,D. Bade, K.L. Cottingham, T.E. Essington, J.N. Houser and D.E. Schindler. 2001. Trophic cascades, nutrients and lake productivity: whole-lake experiments. Ecological Monographs 71: 163-186."

1.3 Additional Information and Support

For more information, please contact the data analyser, **Keith Bollt** at keith.bollt@duke.edu