# Assignment 4: Water Quality in Rivers

*Gabi Richichi*

## OVERVIEW

This exercise accompanies the lessons in Hydrologic Data Analysis on water quality in rivers.

### Directions

1. Change "Student Name" on line 3 (above) with your name.
2. Work through the steps, **creating code and output** that fulfill each instruction.
3. Be sure to **answer the questions** in this assignment document.
4. When you have completed the assignment, **Knit** the text and code into a single HTML file.
5. After Knitting, submit the completed exercise (HTML file) to the dropbox in Sakai. Add your last name into the file name (e.g., "A04_Chamberlin.html") prior to submission.

The completed exercise is due on 25 September 2019 at 9:00 am.

### Setup

1. Verify your working directory is set to the R project file,
2. Load the tidyverse, dataRetrieval, cowplot, xts and dygraphs packages.
3. Set your ggplot theme (can be theme_classic or something else)

```
getwd()
```

```
## [1] "/Users/gabriellerichichi/Documents/5th year @ Duke/Stats/Hydrologic_Data_Analysis/Assignments"
```

```
library(tidyverse)
```

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

## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   0.8.3     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0

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

```
library(dataRetrieval)
library(cowplot)
```

```
##
## ********************************************************

## Note: As of version 1.0.0, cowplot does not change the

##   default ggplot2 theme anymore. To recover the previous

##   behavior, execute:
##   theme_set(theme_cowplot())

## ********************************************************
```

```r
#install.packages("xts")
#install.packages("dygraphs")
library(xts)
```

```
## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

## Registered S3 method overwritten by 'xts':
##   method     from
##   as.zoo.xts zoo

##
## Attaching package: 'xts'

## The following objects are masked from 'package:dplyr':
##
##     first, last
```

```r
library(dygraphs)
```

## Hypoxia

This assignment will look at another measure of water quality - oxygen concentration. Though not directly important for human health, oxygen in the water column is very important for aquatic life, and so is considered a measure of water quality. Hypoxia (low oxygen) has many different definitions. For this assignment, we will use 2 mg/L $O_2$ as our cut-off.

4. Import the oxygen water quality data from New Hope Creek at Blands (using `readNWISqw()`, site code `02097314`, parameter code `00300`). Make a data frame called `O2.dat` that includes only the Date and $O_2$ concentration values. Give your data frame understandable column names.

```r
NewHopeData <- readNWISqw(siteNumbers = "02097314",
                          parameterCd = "00300",
                          startDate = "",
                          endDate = "")
names(NewHopeData)[3] <- c("Date")
names(NewHopeData)[21] <- c("O2")
O2.dat <- NewHopeData %>%
  select(3, 21)
```

5. Create a ggplot of oxygen concentrations over time. Include a horizonal line at 2 mg/l to show the hypoxia cutoff.

```r
OxygenPlot <-
  ggplot(O2.dat, aes(x = Date, y = O2)) +
  geom_point() +
  geom_line(y = 2) +
  ggtitle("Dissolved Oxygen Levels over Time") +
  xlab("Time") +
  ylab(expression("Dissolved O" [2]* " (mg/L)"))
print(OxygenPlot)
```

A04_RiverWQ_files/figure-latex/unnamed-chunk-1-1.pdf

6. What do you notice about the frequency of hypoxia overtime?

   It appears that hypoxia occurred a few times in the 1980s but has not occurred since. In the '80s, there was a wide range of dissolved oxygen levels measured, from below 2.0 mg/L to almost 12.5 mg/L. Over time, the dissolved oxygen levels seem to stabilize and centralize around 6 mg/L.

## Nutrients

7. Often times hypoxia is associated with high nutrient concentrations, because abundant nutrients promote biomass growth which increases respiration and depletes oxygen concentrations in the water (remember how oxygen concentrations were very low in the hypolimnion from the Physical Properties of Lakes week). Create a new data frame, called `nutrients.dat` with total nitrogen (parameter code `00600`) and total phosphorus (parameter code `00665`) data from the USGS. Your data frame should have 3 columns, `Date`, `TotalNitrogen_mgl-N`, and `TotalPhosphorus_mgl-P`.

```
NutrientData <- readNWISqw(siteNumbers = "02097314",
                           parameterCd = c("00600", "00665"),
                           startDate = "",
                           endDate = "")

names(NutrientData)[3] = "Date"
names(NutrientData)[21] = "Measurements"
Nitrogen <- filter(NutrientData, parm_cd == "00600")
Phosphorus <- filter(NutrientData, parm_cd == "00665")
NitrogenCut <- Nitrogen %>%
  select(3, 21)
names(NitrogenCut)[2] = "TotalNitrogen_mglN"
PhosphorusCut <- Phosphorus %>%
  select(3,21)
names(PhosphorusCut)[2] = "TotalPhosphorus_mglP"
nutrients.dat <- left_join(NitrogenCut, PhosphorusCut, by = "Date")
```

8. Create two ggplots stacked with cowplot that show nutrient concentrations over time.

```
NPlot <-
  ggplot(nutrients.dat, aes(x = Date, y = TotalNitrogen_mglN)) +
  geom_point() +
  xlab("Date") +
  ylab("Total Nitrogen (mg/L)")
print(NPlot)
```

-1.pdf)

```
PPlot <-
  ggplot(nutrients.dat, aes(x = Date, y = TotalPhosphorus_mglP)) +
  geom_point() +
  xlab("Date") +
  ylab("Total Phosphorus (mg/L)")
print(PPlot)
```

-2.pdf)

```
CombinedNP <-
  plot_grid(NPlot, PPlot, ncol = 1)
print(CombinedNP)
```

-3.pdf)

9. What do these plots tell you about nutrient concentrations over time? How might this relate to your previous plot of hypoxia events?

   Nutrient concentrations were very high in the 1980s. One data point for Total Nitrogen reached about 30 mg/L, and one data point for Total Phosphorus reached about 5 mg/L. After the mid-1980s, total nutrient levels stabilized at below 5 mg/L for nitrogen and below 1 mg/L for phosphorus. Nutrient levels are correlated with dissolved oxygen levels. As nutrient levels increase, phytoplankton are able to utilize those additional nutrients to synthesize food, and they subsequently expel oxygen. As a result, oxygen levels increase, the phytoplankton respire the oxygen at night, some phytoplankton cells die and are decomposed, and DO levels decrease. The river is then at risk of hypoxia. Therefore, when nutrients levels increased in the 1980s, it makes sense that oxygen levels both increased and decreased, even to the point of hypoxia.

## Discharge and Temperature

10. Turbulent flow in rivers mixes oxygen into the water column. As discharge decreases, water moves slower, and oxygen diffuses slower into the water from the atmosphere. Download and reformat the daily discharge data for New Hope Creek (function `readNWISdv()`, site `02097314`, parameter `00060`).

```
NewHopeDischarge <- readNWISdv(siteNumbers = "02097314",
                               parameterCd = "00060")
names(NewHopeDischarge)[4:5] = c("Discharge", "Approval.Code")
```

11. Create a dygraph of discharge and oxygen concentrations. You should have discharge on the y axis and oxygen concentration on the y2 axis. Set your y-axes limits so that they don't change as you zoom in and out. (hint: use `dyAxis(..., valueRange = c(0, 6500))` for discharge & `dyAxis(..., valueRange = c(0, 13))` for oxygen).

```
#Discharge <- NewHopeDischarge$Discharge
#OxygenConc <- NewHopeData$O2
#NewPlot <-
#  dygraph(NewHopeDischarge, NewHopeData, aes(x = Date, y1 = Discharge, y2 = O2)) +
#  dyAxis(valueRange = c(0, 6500)) +
#  dyAxis(valueRange = c(0, 13))
#print(NewPlot)
```

12. Do hypoxic events happen in the wet or the dry season? How many days after storm events (of at least 100 cfs) do they occur?

    Hypoxic events occur in the wet season, when runoff transports nutrients from agricultural fields into rivers and other bodies of water.

13. Temperature also impacts the solubility of oxygen in water. More oxygen can dissolve when temperatures are *lower*. How would you analyze the effects of temperature on dissolved oxygen concentrations? What data would you need? How would you separate the effects of temperature from the effects of flow?

    In order to analyze the effects of temepraturе on dissolved oxygen concentrations, you would need to measure both the temeperature and the dissolved oxygen concentrations of a river over time. You would then plot the data and see if there is a correlation. You would separate the effects of temperature from the effects of flow by measuring the flow at various points in the river and plotting it against the temperature and dissolved oxygen data.

14. What do you think is the biggest driver of hypoxia in New Hope Creek? What data do you use to support your conclusion?

    The biggest driver of hypoxia in New Hope Creek appears to be nutrient concentration, since nutrient concentration and dissolved oxygen levels are highly correlated over a long time frame, from the 1980s until present day.

## Reflection

15. What are 2-3 conclusions or summary points about river water quality you learned through your analysis?

    I have learned: 1) That nutrient levels correlate with dissolved oxygen levels. 2) That hypoxia is the absence of oxygen, and that hypoxia is caused by a boom in production followed by algal respiration and death. 3) That temperature can also affect dissolved oxygen levels, which could be compounded by flow.

16. What data, visualizations, and/or models supported your conclusions from 15?

    The ggplots of Dissolved Oxygen, Nitrogen, and Phosphorous over time were very convincing. There is a clear correlation.

17. Did hands-on data analysis impact your learning about water quality relative to a theory-based lesson? If so, how?

    Yes, I believe that actually creating the graphs was more effortful than just viewing graphs that were already created in a textbook, for example. I will remember the data trends more since I put in more effort.

18. How did the real-world data compare with your expectations from theory?

    The real-world data aligned with my expectations in terms of the fact that increased nutrient concentrations are correlated with hypoxic events. However, I would have assumed that hypoxia is worse now, in present day, than it was in the 1980s, just because some environmental problems such as climate change have been exacerbated recently.