Relationship Between Common Stream Health Indicators

Carter Coleman

2021-10-08

# 1 Summary/Abstract

# 2 Introduction

*When monitoring freshwater streams, stream health is typically measured by its physical characteristics (chemistry, surrounding land use area, channelization, etc.); Microbial (fecal indicator), and ecology (diversity of indicator macroinvertebrates). However, these categories are usually analyzed separately to assess stream health. As such, my question is how well are these parameters correlated with each other. Specifically, can you use one of the indicators mentioned (physical, microbial, and ecological) to describe both of the others to make a statement about overall stream health?*

*I would expect that because the all three are correlated with urbanization/anthropocentric impact, a negative impact in one indicator category will indicate a correlated negative impact in the other categories. The only caveat is that because level of noise in stream health data, significant correlations may be difficult to find.*

# 3 Methods and Results

*UOWN has provided a great deal of data dating back to 2003. First,I will need to wrangle missing data values that are a result of inconsistent sampling efforts. This will make the data easier to work with and narrow down a time frame for the data that is consistent and free of sampling bias. Next, I will analyze and compare trends in the change of chemical (conductivity, turbidity, no3.mgL); ecological(Biological Score); and microbial indicator data (e.coli CFU). Generally speaking, increases in chemical properties like conductivity, turbidity, no3.mgL and microbial indicator abundance (i.e. ecoli CFU) indicate decreased stream health. While a decrease in biological score indicates decrease in stream health. Next, since biological scores are used as a measure of broader ecological health, I will run t-tests on if the other factors significantly effect biological score. Finally, I will test the interaction of all three via multivariate analysis.*

## 3.1 Data aquisition

*The data I am using is the publically avaialable data provided by the Upper Oconee Watershed Network (UOWN). This data is part of a citizen science training effort for routine watershed-scale monitoring of the upper oconee river watershed. Trained freshwater scientists assist in training members of the community to collect both chemical and biological stream characteristics that are useful in describing stream health. All data is compliled and posted in raw form on the UOWN website. Of interest to me are the following variables: WSID (ID number), biological\_score, conductivity, turbidity, no3.mgL, e.coli.cfu, month, year, day. The raw data shows 3,075 observations before cleaning. However, while cleaning I had to eliminate “po4.mgl” and “temperature” due to significant amounts of missing data.*

#### 3.1.0.1 Load import packages:

library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5 v purrr 0.3.4  
## v tibble 3.1.3 v dplyr 1.0.7  
## v tidyr 1.1.3 v stringr 1.4.0  
## v readr 2.0.1 v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(readxl)

#### 3.1.0.2 Lodaing data and providing a summary for the raw data

data\_location <- here::here("data/raw\_data/UOWN\_data\_master\_spring2021 (2).xlsx")  
  
rawdata <- readxl::read\_excel(data\_location)  
  
glimpse(rawdata)

## Rows: 3,075  
## Columns: 17  
## $ WS <chr> "BICO", "BICO", "MIDO", "MIDO", "MIDO", "MIDO", "MI~  
## $ ID <chr> "102", "103", "101", "102", "103", "104", "116", "1~  
## $ WSID <chr> "BICO102", "BICO103", "MIDO101", "MIDO102", "MIDO10~  
## $ visual\_score <chr> "44", "NA", "45", "52", "39", "NA", "46", "NA", "44~  
## $ biological\_score <chr> "22", "27", "NA", "NA", "NA", "NA", "NA", "NA", "13~  
## $ conductivity.uscm <chr> "83", "NA", "47", "85", "98", "85", "85", "85", "53~  
## $ turbidity.ntu <chr> "4", "NA", "3", "17", "5", "79", "17", "21", "5", "~  
## $ po4.mgL <chr> "NA", "NA", "NA", "NA", "NA", "NA", "NA", "NA", "NA~  
## $ no3.mgL <chr> "0.8", "NA", "NA", "NA", "NA", "NA", "NA", "NA", "0~  
## $ pH <chr> "7.2", "NA", "NA", "6.8", "NA", "7", "7", "6.9", "6~  
## $ temperature.c <chr> "NA", "NA", "NA", "NA", "NA", "NA", "NA", "NA", "NA~  
## $ e.coli.cfu <chr> "275", "NA", "145", "NA", "NA", "NA", "NA", "NA", "~  
## $ month <dbl> 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, ~  
## $ year <dbl> 2001, 2001, 2001, 2001, 2001, 2001, 2001, 2001, 200~  
## $ day <dbl> 28, 28, 28, 28, 28, 28, 28, 28, 28, 28, 28, 28, 28,~  
## $ quarter <dbl> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, ~  
## $ ecoli.method.known <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~

## 3.2 Data import and cleaning

## 3.3 Exploratory analysis

## 3.4 Full analysis

# 4 Discussion

## 4.1 Summary and Interpretation

## 4.2 Strengths and Limitations

## 4.3 Conclusions

# 5 References