**An investigation of fish community composition in the Chicago Area Waterway System**

Github repository with data and R code: https://github.com/mcgilll/ENTM642-Final.git

**Intro**

The Chicago Area Waterway System, initialized as CAWS, is a system comprised of over 140km of waterways (United States Geological Survey, 2015). Most of the system has been dredged, straightened, channelized, and the flow has even been reversed. Several waterways make up the CAWS, including the North Shore Channel (NSC), the Chicago River proper, the North and South Branches (NBCR and SBCR, respectively), and the Chicago Sanitary and Ship canal (CSSC) (figure 1).

Like many urban rivers, it faces many ecological challenges brought on mainly through anthropogenic effects, such as boat traffic, lack of habitat, and pollution from both raw and treated sewage (Happel & Gallagher, 2020). The CAWS uses an outdated combined sewer system (United States Geological Survey, 2015). During heavy rains, this sewer system can overflow and back-up into the river, filling the water with thousands of gallons of untreated sewage leading to hypoxic conditions (Gaulke et al., 2015).

The CAWS also features 6 water reclamation plants that have outflows into the river. The Stickney Water Reclamation Plant, located on the CSSC, is the largest reclamation plant in the world, processing an average of 700 million gallons of wastewater per day (MWRD, 2019a). The O’Brien Water Reclamation Plant is located on the NSC and treats an average of 230 million gallons per day (MWRD, 2019b). Up to 70% of the flow of the CAWS comes from sewage treatment plant effluent (United States Geological Survey, 2015).

Fish communities can serve as biological indicators of water quality (Fausch et al., 1990). Wastewater treatment plant outflows can severely impact fish communities. Sites closest to wastewater treatment plant outfalls have a higher abundance of species and have more stress tolerant and non-native species (McCallum et al., 2019). Because these branches of the Chicago River receive water from different sources, I hypothesize that the fish communities in the different waterways will be significantly different. I hypothesize that the Chicago River communities will be an outlier because it receives water from Lake Michigan, which is on average colder and less productive than the river. It also does not receive any treated wastewater, unlike the NSC and CSSC.

**Methods**

*Fish Collection*

To test this hypothesis, I used fish abundance data from the Metropolitan Water Reclamation District of Greater Chicago. This dataset is part of a much larger collection of fish species abundances dating back to 1974. The fish abundance data is from annual electrofishing surveys throughout the CAWS. Every year at each site, 2 electrofishing passes were made, 1 pass on each side of the river. I used the past 20 years of data and used 5 sites from each waterway. The only exception is the SBCR, which only had 4 sites.

*Statistical analysis*

All statistical analyses were performed in R version 4.2.2 (R core Team, 2022). Before any analysis, I removed any species that had a mean of less than 1 from each site. I also scaled down the data by taking the square root of each abundance measure. This was to reduce the weight of some more abundant species.

To test my hypothesis, I first created a Bray-Curtis and Jaccard dissimilarity matrix. I chose Bray-Curtis because it measures dissimilarity between sites based on abundance (Bray & Curtis, 1957). Bray-Curtis is a great index for species data because it considers double zeros, meaning that a species could be absent for different reasons at different sites (Legendre & Gallagher, 2001). Jaccard index measures the dissimilarity between sites based on presence and absence of species (Ricotta & Podani, 2017).

I visualized these matrices with dendrograms using the vegan and dendextend packages (Galili, 2015) and a PCoA using the ape package (Paradis & Schliep, 2019). To test for differences in species composition between sites, I used an adonis test and calculated the PERMANOVA (Anderson, 2017). I used the betadisper function of the Vegan package to check the assumption of similar dispersions (Oksanen et al., 2022). PCoA ordinates were then visualized by waterway by plotting the adonis.

**Results**

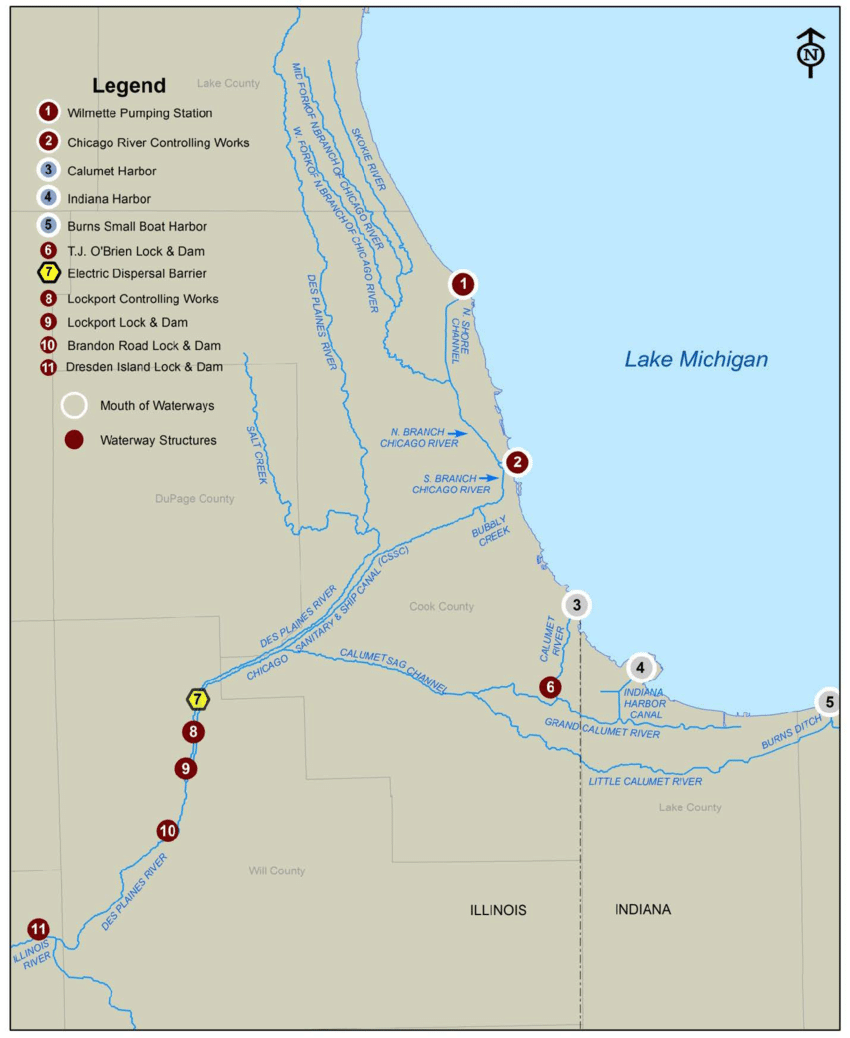
The dendrograms did not show any clear grouping of sites by waterways (Figure 2). The PCoA ordination was also similar, showing no clear separations (Figure 3). The ANOVA of the beta-dispersions of the Bray-Curtis matrix showed that the 5 waterways met the assumption of homogenous dispersions (P=0.028, DF=4, P=0.280). The PERMANOVA test showed that the waterways were significantly different (p=0.004, DF=4, R2=0.307). Further analysis using adonis shows that some of the groups were significantly different by waterway, with NSC significantly different from NBCR and SBCR, and CSSC significantly different from NBCR (figure 4).

**Discussion**

Though the fish communities of each waterway were not as separate as I hypothesized, the waterways were still significantly different. The two sites with wastewater treatment outflows (CSSC and NSC) were different compared to the other three waterways. This confirms previous studies that show that wastewater treatment outflows affect the fish community composition (McCallum et al., 2019). The Chicago River, SBCR, and NBCR had considerable overlap on the PCoA ordination, which shows that the Bray-Curtis matrices of fish abundances were similar.

Further studies should be completed to better understand how and why these fish communities differ. Though these waterways are connected, the fish communities differ, which could mean that each waterway has different environmental conditions. Further studies investigating the environmental variables at each waterway could lead to better management planning for species richness.

**Figures**



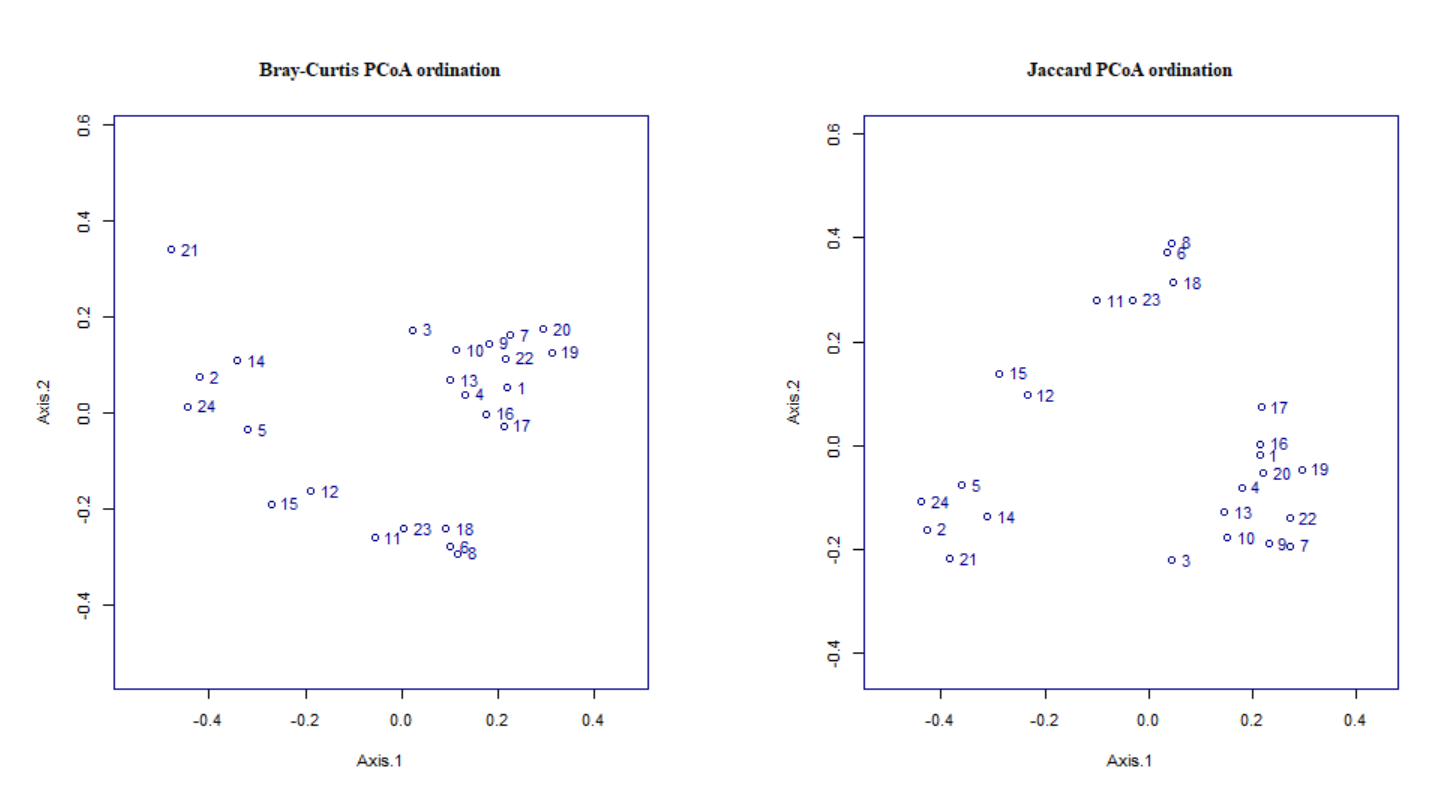
Chicago River

Figure 1: Map of the Chicago Area Waterway System (From Fisheries and Oceans Canada, 2017).

Study areas are highlighted with boxes. Chart

Description automatically generated

Figure 2. Dendrogram of both Bray-Curtis and Jaccard dissimilarity matrix.

Figure 3: PCoA ordinations of Bray-Curtis and Jaccard dissimilarity indices.

Diagram

Description automatically generated

Figure 4: PCoA ordination of Bray-Curtis dissimilarity matrix. Sites grouped by waterway.

**References**

Anderson, M. J. (2017). Permutational multivariate analysis of variance (PERMANOVA). *Wiley*

*StatsRef: Statistics Reference Online*, 1–15.

https://doi.org/10.1002/9781118445112.stat07841

Bray, J. R., & Curtis, J. T. (1957). An ordination of the upland forest communities of Southern Wisconsin.

*Ecological Monographs*, *27*(4), 325–349. <https://doi.org/10.2307/1942268>

Fausch, K.D., Lyons, J., Karr, J.R., Angermeier, P.L. (1990). Fish communities as indicators of

environmental degradation. *American Fisheries Society Symposium, 8*, 123-144.

Fisheries and Oceans Canada. (2017). *Ecological risk assessment of grass carp (Ctenopharyngodon*

*Idella) for the Great Lakes basin.*

Gaulke, G. L., Wolfe, J. R., Bradley,, D. L., Moskus, P. E., Wahl, D. H., & Suski, C. D. (2015). Behavioral and

physiological responses of largemouth bass to rain-induced reductions in dissolved oxygen in an

urban system. *Transactions of the American Fisheries Society*, *144*(5), 927–941.

<https://doi.org/10.1080/00028487.2015.1054517>

Galili, Tal (2015). dendextend: an R package for visualizing, adjusting, and comparing trees of

hierarchical clustering. Bioinformatics. DOI: 10.1093/bioinformatics/btv428

Happel, A., & Gallagher, D. (2020). Chicago’s fish assemblage over ~30 years – more fish and

more native species. *Urban Ecosystems*, *24*(2), 311–325.

<https://doi.org/10.1007/s11252-020-01020-3>

Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, *129*(2), 271–280. <https://doi.org/10.1007/s004420100716>

McCallum, E. S., Nikel, K. E., Mehdi, H., Du, S. N. N., Bowman, J. E., Midwood, J. D., Kidd, K. A., Scott, G.

R., & Balshine, S. (2019). Municipal wastewater effluent affects fish communities: A multi-year

study involving two wastewater treatment plants. *Environmental Pollution*, *252*, 1730–1741.

<https://doi.org/10.1016/j.envpol.2019.06.075>

MWRD. (2019a). *Stickney Water Reclamation Plant Fact Sheet, Metropolitan Water Reclamation District*

*of Greater Chicago Fact Sheet.*

<https://mwrd.org/sites/default/files/documents/Fact_Sheet_Stickney.pdf>

MWRD. (2019a). *O’Brien Water Reclamation Plant Fact Sheet, Metropolitan Water Reclamation District*

*of Greater Chicago Fact Sheet.*

<https://mwrd.org/sites/default/files/documents/Fact_Sheet_O%27Brien.pdf>

Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M,

Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De

Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M,

Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J (2022).

\_vegan: Community Ecology Package\_. R package version 2.6-4, <https://CRAN.R-

project.org/package=vegan>.

Paradis E, Schliep K (2019). “ape 5.0: an environment for modern phylogenetics and evolutionary

analyses in R.” \_Bioinformatics\_, \*35\*, 526-528. doi:10.1093/bioinformatics/bty633

<https://doi.org/10.1093/bioinformatics/bty633>.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for

Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Ricotta, C., & Podani, J. (2017). On some properties of the Bray-Curtis dissimilarity and their ecological

meaning. *Ecological Complexity*, *31*, 201–205. <https://doi.org/10.1016/j.ecocom.2017.07.003>

United States Geological Survey. (2015). *Hydrology of and Current Monitoring Issues for the Chicago*

*Area Waterway System, Northeastern Illinois. https://pubs.usgs.gov/sir/2015/5115/sir20155115.pdf*

**R code**

Setwd(“Yourdirectoryhere”)

install.packages("vegan")

install.packages("ape")

install.packages("dendextend")

library(vegan)

library(ape)

library(dendextend)

###### Read in and clean up data ####

CCDS<-read.csv("CAWS fish abundance data.csv")

str(CCDS)

CCDS[is.na(CCDS)] <- 0 ### Replaced NA values with zeros

str(CCDS) ### Checked to make sure

dim(CCDS)

my.means <- colMeans(CCDS[,3:79])

my.means ### Wide range of species abundance

str(my.means)

common.species <- subset(CCDS[3:79], select=which(my.means>1))

colMeans(common.species) ### Subset df to only include common species

scaled.ccd<-sqrt(common.species) ### I scaled down the data because some

###species are much more abundant than others.

scaled.ccd ### Square root of original data

###### PCOA #######

waterways <-as.factor(CCDS[,1]) ### Made waterways a factor

caws.full <- cbind(scaled.ccd, waterways) ### Combined waterways and scaled fish data

dim(caws.full)

library(ape) ### Package for pcoa

fish.d.bc <- vegdist(caws.full[,1:37], "bray")### Created dissimilarity matrix using bray-Curtis dissimilarity

fish.bc.pcoa <- pcoa(fish.d.bc) ### Created PCoA from dissimilarity matrix

biplot(fish.bc.pcoa) ### Plot PCoA

fish.d.j <- vegdist(scaled.ccd, "jaccard") ### Created dissimilarity matrix using jaccard ###dissimilarity based on presence/absence

fish.pcoa.j <- pcoa(fish.d.j)

biplot(fish.pcoa.j)

par(mfrow=c(1,2), cex=0.7) ### Plot both PCoA ordinations

biplot(fish.bc.pcoa, main = "Bray-Curtis PCoA ordination")

biplot(fish.pcoa.j, main= "Jaccard PCoA ordination")

########## Dendrogram ########

clust.j.w <- hclust(fish.d.j, method="ward.D2") ## Use 2 different cluster methods

clust.bc.w <- hclust(fish.d.bc, method="ward.D2")

clust.bc.w <- as.dendrogram(clust.bc.w)

clust.j.w <-as.dendrogram(clust.j.w)

par(mfrow=c(2,1),cex=0.6, col="darkblue", mar=c(6,2,2,10)) ### Plotting dendrogram of both jaccard and B-C to compare

plot(clust.bc.w, horiz=TRUE, main="Bray Curtis", xlab="Distance in Species Space") ### Very similar

plot(clust.j.w, horiz=TRUE, main="Jaccard", xlab="Distance in Species Space")

par(mfrow=c(1,1), cex=1)

##### Testing for differences between waterways ####

waterway <- as.factor(CCDS[,1])

fish.full <- cbind(scaled.ccd, waterway)

attach(fish.full)

str(fish.full)

species <- fish.full[,1:37] ## Tell R to use the 4 measures as input

adon.disp <- betadisper(vegdist(fish.full[,1:37], method="bray"), waterway) ## Check assumption of ### similar dispersions

anova(adon.disp) ## Group dispersions are homogenous

waterways.adon <- adonis(species ~ waterway, method="bray", data=fish.full, ## Do adonis to check ## for differences

control=permControl(strata=waterway), permuations=9999) ## Test against 9999 permutations as null

waterways.adon ## Check results.

plot(adon.disp) ## Plot the dispersions.

########################