8. Worksheet: Among Site (Beta) Diversity – Part 2

Thomas Zambiasi; Z620: Quantitative Biodiversity, Indiana University

08 February, 2023

OVERVIEW

In this worksheet, we continue to explore concepts, statistics, and visualizations related to β-diversity. Now that you know how to formally quantify β-diversity, we will learn how to test hypotheses about β-diversity using multivariate statistics.

Directions:

1. In the Markdown version of this document in your cloned repo, change "Student Name" on line 3 (above) with your name.
2. Complete as much of the worksheet as possible during class.
3. Use the handout as a guide; it contains a more complete description of data sets along with examples of proper scripting needed to carry out the exercises.
4. Answer questions in the worksheet. Space for your answers is provided in this document and is indicated by the ">" character. If you need a second paragraph be sure to start the first line with ">". You should notice that the answer is highlighted in green by RStudio (color may vary if you changed the editor theme).
5. Before you leave the classroom today, it is *imperative* that you **push** this file to your GitHub repo, at whatever stage you are. This will enable you to pull your work onto your own computer.
6. When you have completed the worksheet, **Knit** the text and code into a single PDF file by pressing the `Knit` button in the RStudio scripting panel. This will save the PDF output in your '8.BetaDiversity' folder.
7. After Knitting, please submit the worksheet by making a **push** to your GitHub repo and then create a **pull request** via GitHub. Your pull request should include this file (**8.BetaDiversity_2_Worksheet.Rmd**) with all code blocks filled out and questions answered) and the PDF output of `Knitr` (**8.BetaDiversity_2_Worksheet.pdf**).

The completed exercise is due on **Wednesday, February 8th, 2023 before 12:00 PM (noon)**.

1) R SETUP

Typically, the first thing you will do in either an R script or an RMarkdown file is setup your environment. This includes things such as setting the working directory and loading any packages that you will need.

In the R code chunk below, provide the code to:

1. clear your R environment,
2. print your current working directory,
3. set your working directory to your "*/6.BetaDiversity*" folder, and

4. load the **vegan** R package (be sure to install if needed).

```
rm(list=ls())
getwd()
```

```
## [1] "C:/Users/tmzam/GitHub/QB2023_Zambiasi/2.Worksheets/6.BetaDiversity"
```

```
setwd("C:/Users/tmzam/GitHub/QB2023_Zambiasi/2.Worksheets/6.BetaDiversity")
library("vegan")
```

```
## Warning: package 'vegan' was built under R version 4.2.2

## Loading required package: permute

## Loading required package: lattice

## This is vegan 2.6-4
```

2) LOADING DATA

Load dataset

In the R code chunk below, load the **doubs** dataset from the **ade4** package

```
# note, please do not print the dataset when submitting
library("ade4")
data(doubs)
```

3) HYPOTHESIS TESTING

A. Multivariate Procedures for Categorical Designs

Earlier work done in the Doubs River suggested that the river has four distinct regions of habitat quality: the first region (sites 1-14) of "high quality"; the second (sites 15 - 19) and fourth (sites 26 - 30) of "moderate quality"; and the third (sites 20 - 25) of "low quality".

In the code chunk below, test the hypothesis that fish community composition varies with river quality.

1. create a factor vector that categorizes habitat quality in the Doubs River,
2. use the multivariate analyses for categorical predictors to describe how fish community structure relates to habitat quality.

```
#making factor vector
qual.factors<-c(rep("HQ",13),rep("MQ",5),rep("LQ",6),rep("MQ",5))
#recreating fish object from Beta Diversity 1
#making fish abundance data its own object
fish<-doubs$fish
#removing sites with no fish observed
fish<-fish[rowSums(fish)!=0,]

#multivariate analyses
#PERMANOVA using adonis function
adonis2(fish~qual.factors,method="bray",permutations=999)
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = fish ~ qual.factors, permutations = 999, method = "bray")
##              Df SumOfSqs      R2     F Pr(>F)
## qual.factors  2   3.0947 0.45765 10.97  0.001 ***
## Residual     26   3.6674 0.54235
## Total        28   6.7621 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#Indicator Value test (ind. spp. relation to site group)
library("indicspecies") #need to load this package
```

```
## Warning: package 'indicspecies' was built under R version 4.2.2
```

```
indval<-multipatt(fish,cluster=qual.factors,func="IndVal.g",
                  control=how(nperm=999))
summary(indval)
```

```
##
##  Multilevel pattern analysis
##  ---------------------------
##
##  Association function: IndVal.g
##  Significance level (alpha): 0.05
##
##  Total number of species: 27
##  Selected number of species: 23
##  Number of species associated to 1 group: 1
##  Number of species associated to 2 groups: 22
##
##  List of species associated to each combination:
##
##  Group MQ  #sps.  1
##       stat p.value
## Teso 0.686   0.024 *
##
##  Group HQ+MQ  #sps.  2
##       stat p.value
## Satr 0.860   0.002 **
## Phph 0.859   0.010 **
##
##  Group LQ+MQ  #sps.  20
##       stat p.value
## Alal 0.935   0.001 ***
## Gogo 0.933   0.001 ***
## Ruru 0.916   0.001 ***
## Legi 0.901   0.001 ***
## Baba 0.895   0.001 ***
## Chna 0.866   0.002 **
```

```
## Spbi 0.866   0.001 ***
## Cyca 0.866   0.001 ***
## Acce 0.866   0.001 ***
## Lele 0.863   0.005 **
## Titi 0.853   0.005 **
## Chto 0.829   0.003 **
## Rham 0.829   0.002 **
## Anan 0.829   0.001 ***
## Eslu 0.827   0.027 *
## Pefl 0.806   0.016 *
## Blbj 0.791   0.002 **
## Scer 0.766   0.004 **
## Abbr 0.750   0.006 **
## Icme 0.661   0.014 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#finding phi coeff. of association (spp. habitat preference)
fish.rel<-decostand(fish,method="total")
phi<-multipatt(fish.rel,cluster=qual.factors,func="r.g",
                control=how(nperm=999))
summary(phi)
```

```
##
##  Multilevel pattern analysis
##  ---------------------------
##
##  Association function: r.g
##  Significance level (alpha): 0.05
##
##  Total number of species: 27
##  Selected number of species: 18
##  Number of species associated to 1 group: 9
##  Number of species associated to 2 groups: 9
##
##  List of species associated to each combination:
##
##  Group HQ  #sps.  3
##       stat p.value
## Phph 0.802   0.002 **
## Neba 0.734   0.001 ***
## Satr 0.650   0.001 ***
##
##  Group LQ  #sps.  2
##       stat p.value
## Alal 0.693   0.001 ***
## Ruru 0.473   0.031 *
##
##  Group MQ  #sps.  4
##       stat p.value
## Anan 0.571   0.008 **
## Spbi 0.557   0.007 **
## Chto 0.542   0.011 *
## Icme 0.475   0.035 *
```

```
##
##  Group LQ+MQ  #sps.  9
##       stat p.value
## Legi 0.658   0.003 **
## Baba 0.645   0.003 **
## Rham 0.600   0.005 **
## Acce 0.594   0.006 **
## Cyca 0.586   0.008 **
## Chna 0.571   0.004 **
## Blbj 0.571   0.011 *
## Gogo 0.523   0.019 *
## Abbr 0.499   0.038 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Question 1: Based on the PERMANOVA, IndVal, and phi coefficient analyses, what did you learn about the relationship between habitat quality and the fish species composition? Are the different analyses consistent with one another and do they agree with the visualizations (heat maps, cluster dendograms, ordinations) that you created?

> ***Answer 1***: The PERMANOVA analysis indicates that the fish community composition does vary between the three levels of river quality ($F_{2,26} = 10.97$, $p = 0.001$). The Indicator Value analysis showed that 23 of the 27 fish species had a significant association with a river quality group (or pairing of groups). Only two of these, brown trout (*Salmo trutta fario*) and minnow (*Phoxinus phoxinus*) were associated with the combination of high-quality and mid-quality habitats. We can also assess habitat preferences via the phi coefficient. Three species had a significant preference for the high-quality habitat, with minnows having the highest (0.802), stone loach (*Nemacheilus barbatulus*) having the second highest (0.734) and brown trout having the least preference of these species (0.650). Overall, it seems that high-quality habitats contain fish species that exhibit a strong preference for them and tend to primarily be associated with these environments. Here, these would be *S. trutta fario*, *P. phoxinus*, and *N. barbatulus*; these generally wouldn't be found in lower-quality environments, so their presence can be predictive of water quality. Fish communities present in lower quality habitats appeared to have less of a preference for their environments (phi mostly between 0.4 and 0.7). These analyses are all consistent with one another and the groupings (particularly for the phi coefficient) match what I saw in the PCoA visualization I created. In both the coefficient groupings and the visualization, minnows, brown trout, and stone loach fell into their own high-quality group, as did bleak (*Alburnus alburnus*) and roach (*Rutilus rutilus*) in the low-quality group, while the rest of the species (the majority) were within the mid- to low-quality groupings.

B. Multivariate Procedures for Continuous Designs

i. Mantel Test

In the R code chunk below, do the following:

1. create distance matrices for both fish communities and environmental factors, and
2. use a Mantel test to test the hypothesis that fish assemblages are correlated with stream environmental variables.

```
#creating matrices
fish.dist<-vegdist(fish,method="bray") #using Bray-Curtis distance
```

```
env.dist<-vegdist(scale(doubs$env[-8,]),
                  method="euclid") #using physical distance
#performing mantel test
mantel(fish.dist,env.dist)
```

```
##
## Mantel statistic based on Pearson's product-moment correlation
##
## Call:
## mantel(xdis = fish.dist, ydis = env.dist)
##
## Mantel statistic r: 0.604
##       Significance: 0.001
##
## Upper quantiles of permutations (null model):
##   90%   95% 97.5%   99%
## 0.113 0.140 0.173 0.207
## Permutation: free
## Number of permutations: 999
```

Question 2: What do the results from our Mantel test suggest about fish diversity and stream environmental conditions? How does this relate to your hypothesis about stream quality influencing fish communities?

> ***Answer 2***: The Mantel test is a further assessment of whether fish community diversity is correlated with habitat quality. The r statistic for the correlation between fish in the Doubs river and the environmental conditions is 0.604. Values of this metric closer to 1 indicate a stronger association between the two matrices, meaning that there is a fairly strong, significant correlation between the fish communities and habitat ($p = 0.001$). Because this result suggests that fish community composition in the Doubs is correlated with environmental conditions throughout the river, this supports my hypothesis that the structure of fish communities is determined by water quality.

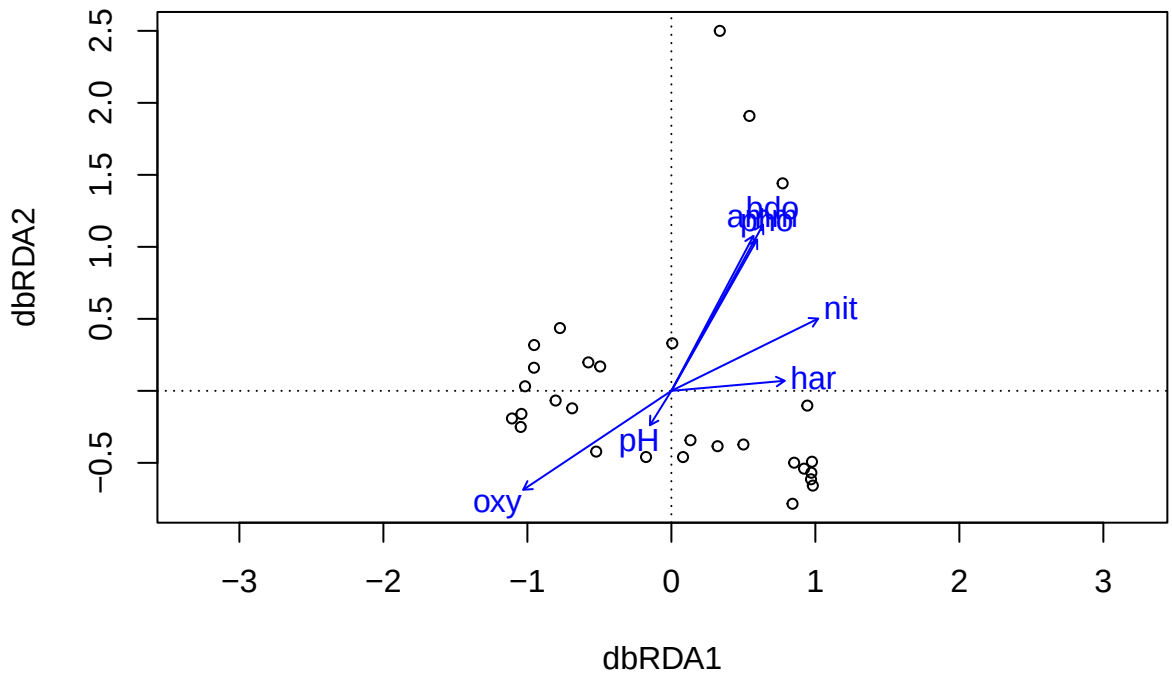
ii. Constrained Ordination

In the R code chunk below, do the following:

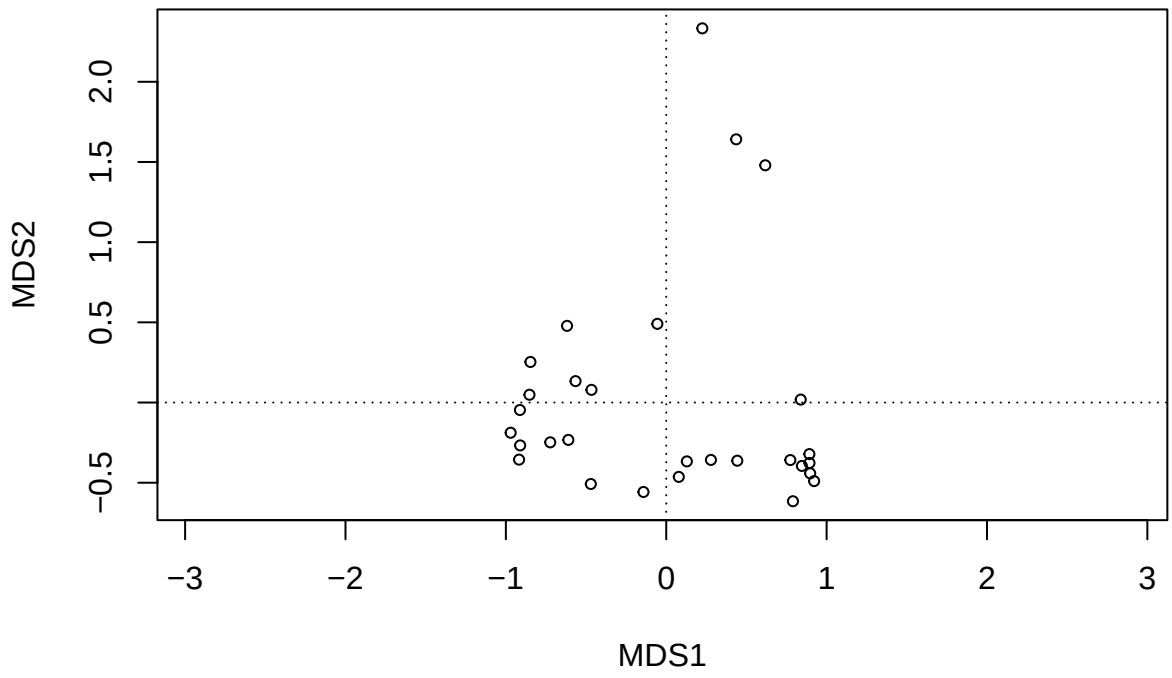
1. create an environmental matrix of the water chemistry data included in the `doubs` dataset using forward and reverse selection of variables,
2. conduct a redundancy analysis on the fish assemblages of the Doubs River,
3. use a permutation test to determine the significance of the constrained analysis,
4. use a permutation test to determine the correlation of each environmental factor on the constrained axes,
5. calculate the explained variation on the first and second constrained axes,
6. plot the constrained ordination results including labeled points for each site, and
7. add vectors that demonstrate the influence of each environmental factor the constrained ordination.

```
#making environmental matrix for water chemistry
water.chem<-as.matrix(doubs$env[-8,5:11])
#doing distance-based redundancy analysis (dbRDA)
#first need to make resemblance matrix using Bray-Curtis
```

```
fish.db<-vegdist(fish,method="bray")
doubs.dbrda<-dbrda(fish.db~.,as.data.frame(water.chem))
ordiplot(doubs.dbrda)
```




```
#model selection to avoid over-fitting
#first modeling only intercept
doubs.dbrda.mod0<-dbrda(fish.db~1,as.data.frame(water.chem))
#no vectors b/c didn't constrain anything, axes mean this is simple MDS (PCoA)
ordiplot(doubs.dbrda.mod0)
```




```
#now modeling full model w/ all explanatory variables
doubs.dbrda.mod1<-dbrda(fish.db~.,as.data.frame(water.chem))
#now going through all combos of explanatory variables in model to return
#the one w/ lowest AIC value
doubs.dbrda<-ordiR2step(doubs.dbrda.mod0,doubs.dbrda.mod1,perm.max=200)
```

```
## Step: R2.adj= 0
## Call: fish.db ~ 1
##
##                 R2.adjusted
## <All variables>  0.53032584
## + oxy            0.27727176
## + nit            0.25755208
## + bdo            0.17477787
## + pho            0.14568614
## + har            0.14174915
## + amm            0.14142804
## <none>           0.00000000
## + pH            -0.01827054
##
##       Df    AIC      F Pr(>F)
## + oxy  1 47.939 11.742  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Step: R2.adj= 0.2772718
## Call: fish.db ~ oxy
##
##                 R2.adjusted
## <All variables>   0.5303258
## + bdo             0.4009000
## + amm             0.3474192
## + pho             0.3452702
## + har             0.3331357
## + nit             0.3316120
## <none>            0.2772718
## + pH              0.2586983
##
##       Df    AIC      F Pr(>F)
## + bdo  1 43.404 6.5716  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.4009
## Call: fish.db ~ oxy + bdo
##
##                 R2.adjusted
## <All variables>   0.5303258
## + nit             0.4980793
## + har             0.4695121
## <none>            0.4009000
## + pho             0.3938042
## + amm             0.3869134
## + pH              0.3865240
##
##       Df    AIC     F Pr(>F)
## + nit  1 39.134 6.034  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.4980793
## Call: fish.db ~ oxy + bdo + nit
##
##                 R2.adjusted
## + amm             0.5415705
## <All variables>   0.5303258
## + pho             0.5277128
## + har             0.5218852
## <none>            0.4980793
## + pH              0.4843267
```

```
#checking out model that has been chosen
doubs.dbrda$call
```

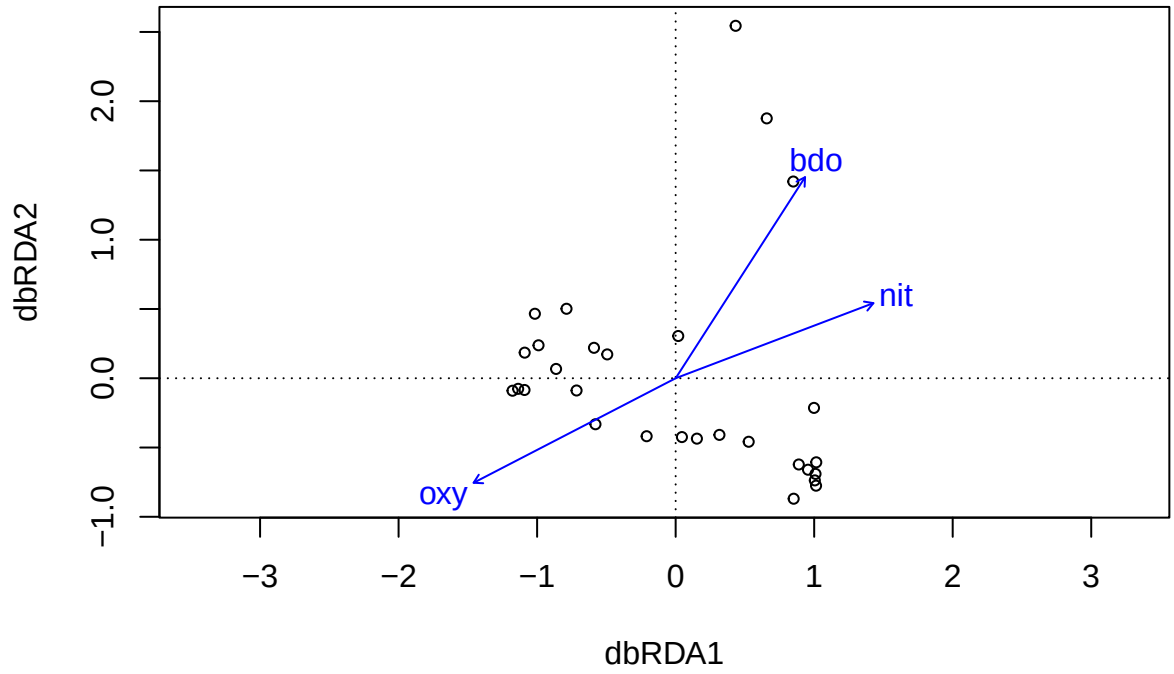
```
## dbrda(formula = fish.db ~ oxy + bdo + nit, data = as.data.frame(water.chem))
```

```
doubs.dbrda$anova
```

```
##                  R2.adj Df    AIC       F Pr(>F)
```

```
## + oxy           0.27727  1 47.939 11.7421  0.002 **
## + bdo           0.40090  1 43.404  6.5716  0.002 **
## + nit           0.49808  1 39.134  6.0340  0.002 **
## <All variables> 0.53033
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
ordiplot(doubs.dbrda)
```




```
#permutation looking for significance of constrained analysis
permutest(doubs.dbrda,permutations=999)
```

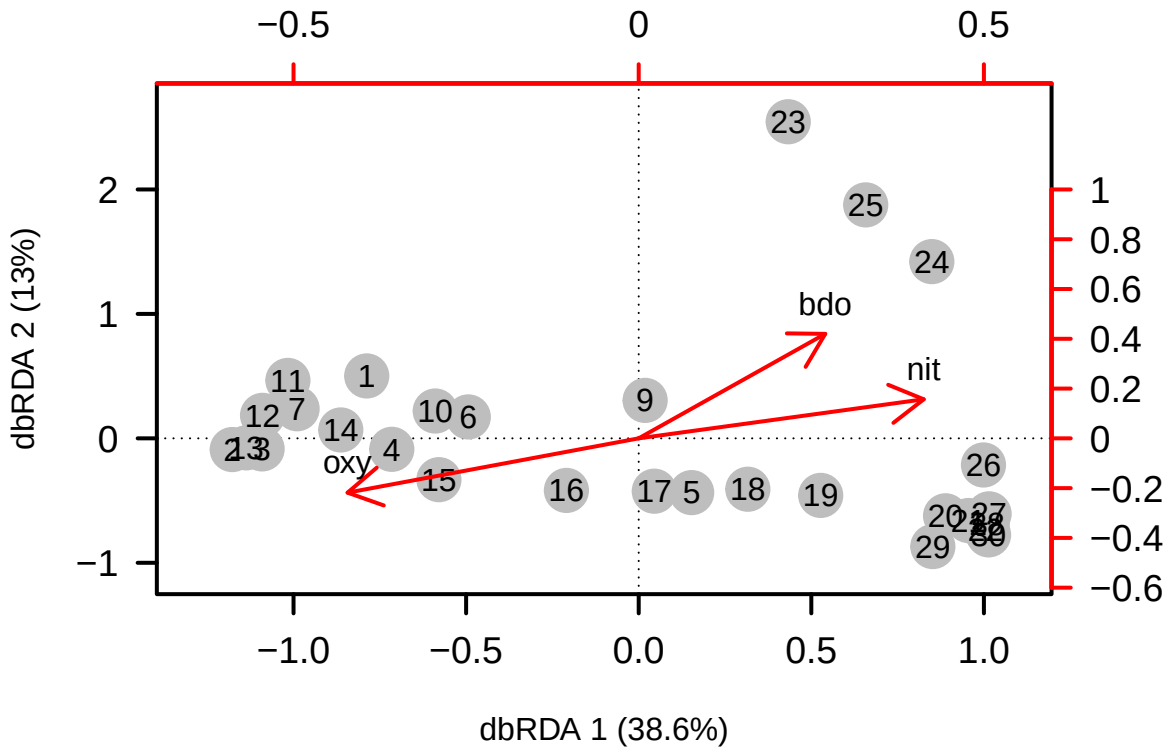
```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.db ~ oxy + bdo + nit, data =
## as.data.frame(water.chem))
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     3  3.7317 10.262  0.001 ***
## Residual 25  3.0304
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
envfit(doubs.dbrda,water.chem[,c(4,6,7)],perm=999)
```

```
##
## ***VECTORS
##
##       dbRDA1   dbRDA2     r2 Pr(>r)
## nit  0.87724  0.48005 0.6431  0.001 ***
## oxy -0.82864 -0.55979 0.7656  0.001 ***
## bdo  0.55603  0.83116 0.8939  0.001 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Permutation: free
## Number of permutations: 999
```

```
#calculating explained variation for constrained axes
dbrda.explvar1<-round(doubs.dbrda$CCA$eig[1]/sum(c(doubs.dbrda$CCA$eig,
                                                   doubs.dbrda$CA$eig)),3)*100
dbrda.explvar2<-round(doubs.dbrda$CCA$eig[2]/sum(c(doubs.dbrda$CCA$eig,
                                                   doubs.dbrda$CA$eig)),3)*100

#plotting ordination for the chosen model
#plot parameters
par(mar=c(5,5,4,4)+0.1)
#starting plot
plot(scores(doubs.dbrda,display="wa"),xlim=c(-1.3,1.1),ylim=c(-1.1,2.7),
     xlab=paste("dbRDA 1 (",dbrda.explvar1,"%)",sep=""),
     ylab=paste("dbRDA 2 (",dbrda.explvar2,"%)",sep=""),cex.axis=1.2,
     axes=FALSE)
#making axes
axis(side=1,labels=T,lwd.ticks=2,cex.axis=1.2,las=1)
axis(side=2,labels=T,lwd.ticks=2,cex.axis=1.2,las=1)
abline(h=0,v=0,lty=3)
box(lwd=2)
#adding points and labels
points(scores(doubs.dbrda,display="wa"),pch=19,cex=3,bg="gray",col="gray")
text(scores(doubs.dbrda,display="wa"),labels=row.names(scores(
  doubs.dbrda,display="wa")))
#adding environmental vectors
env.vectors<-scores(doubs.dbrda,display="bp")
arrows(0,0,env.vectors[,1],env.vectors[,2],lwd=2,lty=1,length=0.2,col="red")
text(env.vectors[,1],env.vectors[,2],pos=3,labels=row.names(env.vectors))
axis(side=3,lwd.ticks=2,cex.axis=1.2,las=1,col="red",lwd=2.2,
     at=pretty(range(env.vectors[,1]))*2,labels=pretty(range(env.vectors[,1])))
axis(side=4,lwd.ticks=2,cex.axis=1.2,las=1,col="red",lwd=2.2,
     at=pretty(range(env.vectors[,2]))*2,labels=pretty(range(env.vectors[,2])))
```



Question 3: Based on the constrained ordination, what are the environmental variables (or groups of correlated variables) that seem to be contributing to variation in fish community structure?

> ***Answer 3***: The results of the constrained ordinatoin seem to suggest that the environmental variables that impact fish community composition the most are dissolved oxygen (oxy), nitrates (nit), and biological demand for oxygen (bdo). The site placement on this ordination plot matches what we saw in the PCoA plot for site/species associations, so we can easily see which communities tend to be associated with certain environmental variables. The left-most site cluster contains sites that have high water quality and communities characterized by brown trout, minnows, and stone loach. The fish community composition in these high-quality habitats seems to be driven by higher levels of dissolved oxygen, and lower levels of both nitrates and biological demand for oxygen. The lowest-quality sites (upper right), identifiable by their populations of bleak and roach, appear to have their fish community composition driven primarily by a higher level of biological demand for oxygen, increased nitrates (less directly), and lower amounts of dissolved oxygen. The remaining mid- and low-quality sites, associated primarily with all of the other species in the dataset, are marked by moderate to low amounts of dissolved oxygen, moderate to high amounts of nitrates (more directly than the lowest-quality sites), and a moderate to high biological demand for oxygen. Overall, it seems that availability of oxygen in the river is most indicative of both water quality and which species are able to live in particular portions of the stream.

iii. Variation Partitioning

In the code chunk below,

1. Create a matrix model of the selected environmental variables,

2. Create a matrix model of the selected PCNM axes,
3. Perform constrained and partial constrained ordinations using the spatial and environmental models you just created,
4. Test the significance of each of your constrained ordinations using permutation tests,
5. Partition the variation among sites into the relative importance of space, environment, spatially structured environment, and residuals,
6. Plot the variation partitioning output to visualize it.

```
#making matrix model for environmental data
env.model<-model.matrix(~oxy+bdo+nit,as.data.frame(water.chem))[,-1]

#making matrix model for PCNM axes
#first weighting sites by relative abundance
rs<-rowSums(fish)/sum(fish)
#now conducting PCNM
doubs.pcnmw<-pcnm(dist(doubs$xy[-8,]),w=rs,dist.ret=T)
#PCNM can give negative eigenvalues, but only eigenvectors associated w/
#positive eigenvalues are useful so will isolate those
doubs.pcnmw$values>0
```

```
##  [1]  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
## [13]  TRUE  TRUE  TRUE  TRUE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [25] FALSE FALSE
```

```
#more model selection to find eigenvalues that make the best model w/ few param.
doubs.space<-as.data.frame(scores(doubs.pcnmw))
doubs.pcnm.mod0<-dbrda(fish.db~1,doubs.space)
doubs.pcnm.mod1<-dbrda(fish.db~.,doubs.space)
step.pcnm<-ordiR2step(doubs.pcnm.mod0,doubs.pcnm.mod1,perm.max=200)
```

```
## Step: R2.adj= 0
## Call: fish.db ~ 1
##
##                   R2.adjusted
## <All variables>  0.626011301
## + PCNM2          0.235370423
## + PCNM3          0.078394885
## + PCNM13         0.065305668
## + PCNM5          0.046185074
## + PCNM6          0.032809156
## + PCNM16         0.030486700
## + PCNM14         0.029680999
## + PCNM9          0.020357410
## + PCNM15         0.013632610
## + PCNM8          0.009411968
## + PCNM1          0.003986221
## + PCNM17         0.002415012
## + PCNM10         0.001326442
## <none>           0.000000000
## + PCNM7         -0.001861430
## + PCNM11        -0.006841522
## + PCNM4         -0.007089863
```

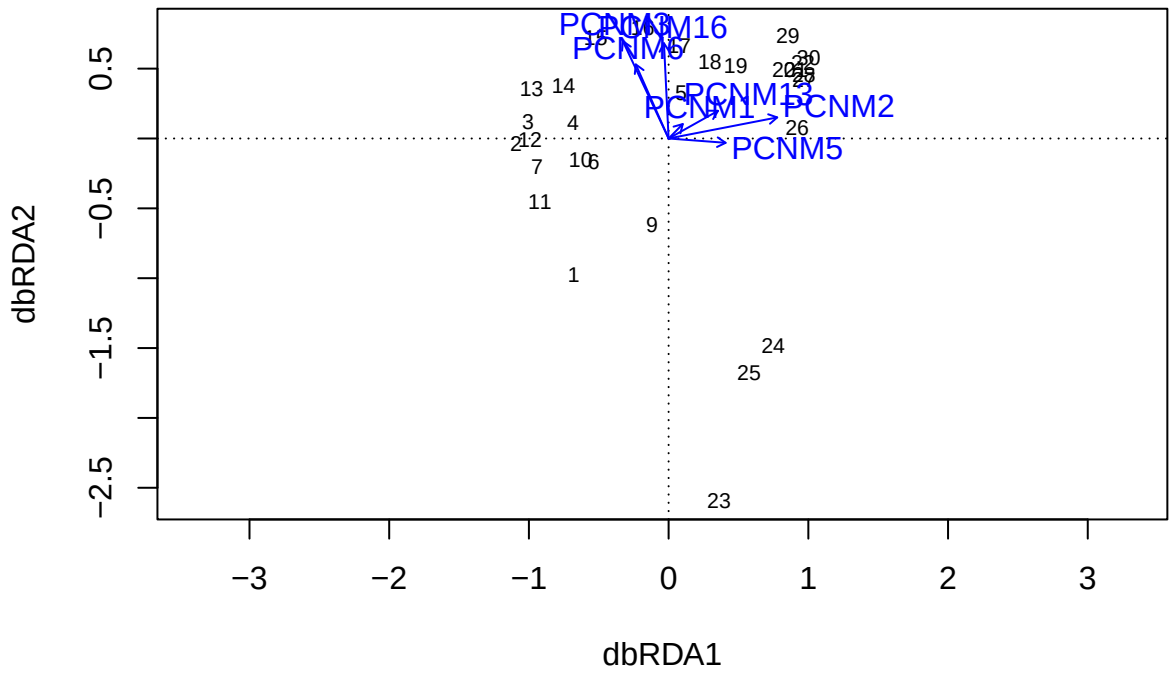
```
## + PCNM12        -0.014396973
##
##         Df    AIC     F Pr(>F)
## + PCNM2  1 49.574 9.619  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.2353704
## Call: fish.db ~ PCNM2
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM3           0.3429270
## + PCNM5           0.3057368
## + PCNM1           0.2885396
## + PCNM16          0.2786746
## + PCNM14          0.2744520
## + PCNM15          0.2692809
## + PCNM6           0.2659866
## + PCNM13          0.2636194
## + PCNM9           0.2517847
## + PCNM8           0.2496240
## + PCNM10          0.2434688
## + PCNM7           0.2431476
## + PCNM17          0.2404343
## + PCNM11          0.2366833
## <none>            0.2353704
## + PCNM12          0.2288789
## + PCNM4           0.2189522
##
##         Df    AIC      F Pr(>F)
## + PCNM3  1 46.083 5.4196  0.004 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.342927
## Call: fish.db ~ PCNM2 + PCNM3
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM5           0.4076020
## + PCNM1           0.3970300
## + PCNM16          0.3853210
## + PCNM15          0.3828748
## + PCNM14          0.3781827
## + PCNM13          0.3770376
## + PCNM6           0.3595644
## + PCNM8           0.3556885
## + PCNM7           0.3541631
## + PCNM10          0.3526775
## + PCNM17          0.3513683
## + PCNM9           0.3433672
## <none>            0.3429270
## + PCNM11          0.3416399
```

```
## + PCNM12          0.3396547
## + PCNM4           0.3311509
##
##         Df    AIC      F Pr(>F)
## + PCNM5  1 43.941 3.8385  0.012 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.407602
## Call: fish.db ~ PCNM2 + PCNM3 + PCNM5
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM1           0.4721469
## + PCNM16          0.4631976
## + PCNM15          0.4589111
## + PCNM14          0.4535248
## + PCNM13          0.4511582
## + PCNM6           0.4305640
## + PCNM7           0.4261965
## + PCNM8           0.4224505
## + PCNM17          0.4181666
## + PCNM10          0.4154485
## + PCNM11          0.4112178
## + PCNM9           0.4111995
## + PCNM12          0.4087602
## <none>            0.4076020
## + PCNM4           0.3976526
##
##         Df    AIC     F Pr(>F)
## + PCNM1  1 41.411 4.057   0.01 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.4721469
## Call: fish.db ~ PCNM2 + PCNM3 + PCNM5 + PCNM1
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM13          0.5212427
## + PCNM16          0.5208668
## + PCNM15          0.5161770
## + PCNM14          0.5147355
## + PCNM6           0.4999020
## + PCNM7           0.4936559
## + PCNM8           0.4904113
## + PCNM17          0.4856884
## + PCNM10          0.4835952
## + PCNM11          0.4760087
## + PCNM9           0.4751424
## + PCNM12          0.4747221
## <none>            0.4721469
## + PCNM4           0.4651218
##
```

```
##          Df    AIC      F Pr(>F)
## + PCNM13  1 39.346 3.4612  0.018 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.5212427
## Call: fish.db ~ PCNM2 + PCNM3 + PCNM5 + PCNM1 + PCNM13
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM16          0.5767968
## + PCNM15          0.5715331
## + PCNM14          0.5698343
## + PCNM6           0.5475140
## + PCNM7           0.5392074
## + PCNM8           0.5379134
## + PCNM11          0.5281106
## + PCNM9           0.5267003
## + PCNM10          0.5265029
## + PCNM12          0.5255581
## <none>            0.5212427
## + PCNM17          0.5171800
## + PCNM4           0.5152311
##
##          Df   AIC      F Pr(>F)
## + PCNM16  1 36.48 4.0192  0.014 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.5767968
## Call: fish.db ~ PCNM2 + PCNM3 + PCNM5 + PCNM1 + PCNM13 + PCNM16
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM6           0.6043089
## + PCNM8           0.5970286
## + PCNM12          0.5946888
## + PCNM7           0.5946475
## + PCNM9           0.5883735
## + PCNM10          0.5851333
## + PCNM15          0.5846468
## <none>            0.5767968
## + PCNM17          0.5748533
## + PCNM4           0.5733749
## + PCNM11          0.5711176
## + PCNM14          0.5652509
##
##         Df    AIC      F Pr(>F)
## + PCNM6  1 35.182 2.5296  0.048 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.6043089
## Call: fish.db ~ PCNM2 + PCNM3 + PCNM5 + PCNM1 + PCNM13 + PCNM16 + PCNM6
```

```
##
##                 R2.adjusted
## <All variables>   0.6260113
## + PCNM8           0.6248697
## + PCNM12          0.6208788
## + PCNM10          0.6170988
## + PCNM7           0.6142419
## + PCNM15          0.6140369
## + PCNM9           0.6107110
## <none>            0.6043089
## + PCNM17          0.6037430
## + PCNM11          0.5978305
## + PCNM4           0.5963667
## + PCNM14          0.5932113
##
##         Df    AIC     F Pr(>F)
## + PCNM8  1 34.219 2.151   0.07 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#this is another dbRDA, so we can use a biplot to show how vectors explain
#variation for the different sites
plot(step.pcnm)
```



```
#object "step.pcnm" includes selected model
step.pcnm$anova
```

```
##                  R2.adj Df    AIC      F Pr(>F)
## + PCNM2         0.23537  1 49.574 9.6190  0.002 **
## + PCNM3         0.34293  1 46.083 5.4196  0.004 **
## + PCNM5         0.40760  1 43.941 3.8385  0.012 *
## + PCNM1         0.47215  1 41.411 4.0570  0.010 **
## + PCNM13        0.52124  1 39.346 3.4612  0.018 *
## + PCNM16        0.57680  1 36.480 4.0192  0.014 *
## + PCNM6         0.60431  1 35.182 2.5296  0.048 *
## <All variables> 0.62601
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#now making spatial model with only selected PCNM axes
space.model<-model.matrix(~PCNM2+PCNM3+PCNM5+PCNM1+PCNM13+PCNM16+PCNM6,
                          doubs.space)[,-1]

#constrained/partial constrained ordinations for spatial and env. data
#making constrained ordinations first
doubs.total.env<-dbrda(fish.db~env.model)
doubs.total.space<-dbrda(fish.db~space.model)
#now making partial constrained ordinations
doubs.env.cond.space<-dbrda(fish.db~env.model+Condition(space.model))
doubs.space.cond.env<-dbrda(fish.db~space.model+Condition(env.model))

#testing significance of constrained ordinations w/ permutation tests
permutest(doubs.env.cond.space,permutations=999)
```

```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.db ~ env.model + Condition(space.model))
## Permutation test for all constrained eigenvalues
##          Df Inertia     F Pr(>F)
## Model     3 0.85158 4.423  0.002 **
## Residual 18 1.15519
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
permutest(doubs.space.cond.env,permutations=999)
```

```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
```

```
## Model: dbrda(formula = fish.db ~ space.model + Condition(env.model))
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     7  1.8752 4.1741  0.001 ***
## Residual 18  1.1552
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
permutest(doubs.total.env,permutations=999)
```

```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.db ~ env.model)
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     3  3.7317 10.262  0.001 ***
## Residual 25  3.0304
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
permutest(doubs.total.space,permutations=999)
```

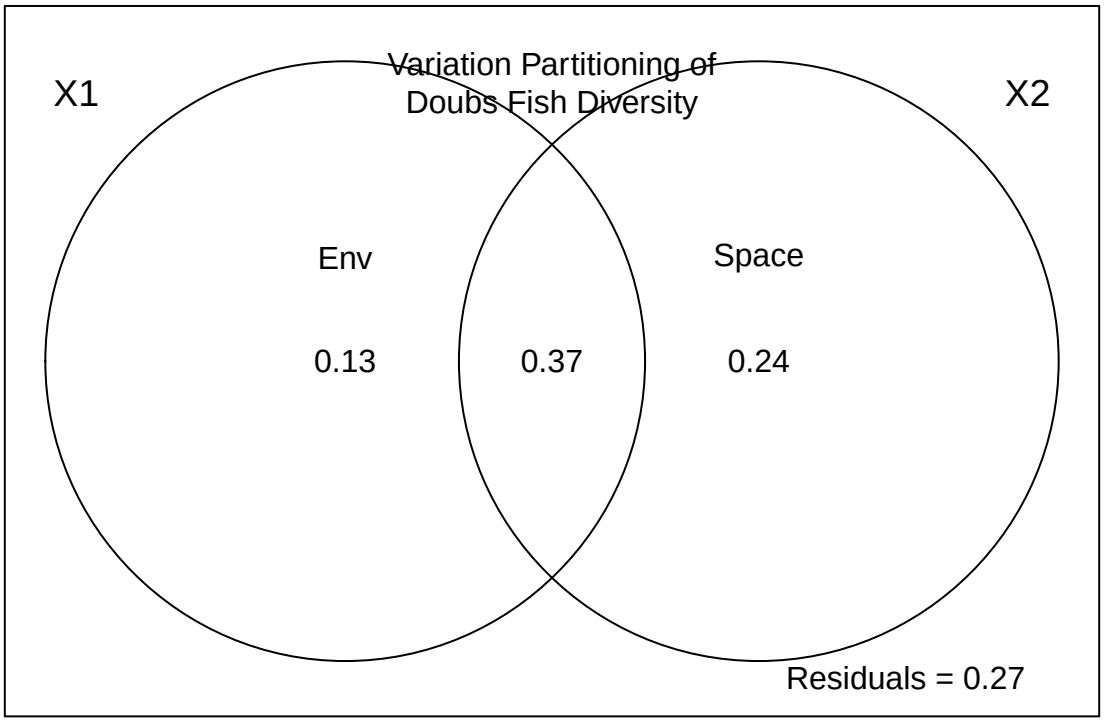
```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.db ~ space.model)
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     7  4.7553 7.1089  0.001 ***
## Residual 21  2.0068
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#partitioning variation and visualizing it
doubs.varpart<-varpart(fish.db,env.model,space.model)
doubs.varpart
```

```
##
## Partition of squared Bray distance in dbRDA
##
## Call: varpart(Y = fish.db, X = env.model, space.model)
##
## Explanatory tables:
## X1:  env.model
## X2:  space.model
```

```
##
## No. of explanatory tables: 2
## Total variation (SS): 6.7621
## No. of observations: 29
##
## Partition table:
##                       Df R.squared Adj.R.squared Testable
## [a+c] = X1             3   0.55186       0.49808     TRUE
## [b+c] = X2             7   0.70323       0.60431     TRUE
## [a+b+c] = X1+X2       10   0.82917       0.73426     TRUE
## Individual fractions
## [a] = X1|X2            3                 0.12995     TRUE
## [b] = X2|X1            7                 0.23618     TRUE
## [c]                    0                 0.36813    FALSE
## [d] = Residuals                          0.26574    FALSE
## ---
## Use function 'dbrda' to test significance of fractions of interest
```

```
#figure code
par(mar=c(2,2,2,2))
plot(doubs.varpart)
text(1,0.25,"Space")
text(0,0.25,"Env")
mtext("Variation Partitioning of\nDoubs Fish Diversity",side=3,line=-3)
```



Question 4: Interpret the variation partitioning results.

Answer 4: The variance partitioning figure above shows the percentage of variation in fish community composition explained by environmental variables, space, or the combination of the two. Environmental variables explained the least amount of variation in community structure, only 13%. Physical space explained 24% of the variation in community composition. However, the interaction of space and environmental variables explains 37% of the variation in community composition, much more than either of the two predictors on their own. Based on these results, I would conclude that fish community composition in the Doubs River is driven by a combination of both the water quality and location along the river. High-quality habitats, for example, appeared in previous analyses to be associated with high amounts of dissolved oxygen, paired with low levels of nitrite and biological demand for oxygen. Highly-oxygenated habitats tend to be near the headwaters of a stream, and the combination of these factors means that indicator species like brown trout, minnow, and stone loach will primarily be found in these areas. Less-oxygenated water would be found further downstream where the river's current has slowed considerably; compounds such as nitrites may have accumulated in the water and a reduced oxygen supply for an area capable of containing more species results in a higher biological demand for oxygen. The other species included in this data might be more capable of surviving in these conditions (although they may not be entirely preferred, as seen in the phi coefficients) than the indicator species, and therefore comprise the communities found in poorer habitats further downstream on the Doubs River.

SYNTHESIS

Load the dataset from that you and your partner are using for the team project. Use one of the hypothesis-testing tools introduced in the beta diversity module. Interpret the findings of your data with respect to principles of biodiversity.

```
#loading KBS dataset (same as used last week)
#sorry the code for making the matrix is still a bit of a mess!

#loading dataset
kbsdata1<-read.csv("kbs_set_1.csv")
#removing first row (has extra unneeded characters)
kbsdata1<-kbsdata1[-1,]
#isolating relevant data columns
kbsdata2<-kbsdata1[,c("replicate","disturbed","fertilized",
                      "species.name","biomass_g_m2")]
#sites will be determined by rep+treatment combo
#need to combine rep+treatment combos into single label column
kbsdata2$site<-paste(kbsdata2$replicate,kbsdata2$disturbed,kbsdata2$fertilized)
#rearranging data again
kbsdata3<-kbsdata2[,c("site","species.name","biomass_g_m2")]
#creating list of sites
kbssitelist<-as.data.frame(unique(kbsdata3$site))
#creating list of species
kbsspplist<-unique(kbsdata3$species.name)
#reworking into dataframe
kbsspplist<-as.data.frame(kbsspplist)
#renaming column to be able to join matrix components easier
colnames(kbsspplist)<-c("species.name")
#need to load dpylr to do joining
library("dplyr")
```

```
##
```

```
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
#creating site-specific data and adding to sXspp. by spp. name, steps
#repeat for each site
#r1, dist. unfert.
R1du<-kbsdata3[kbsdata3$site=="R1 disturbed unfertilized",] #isolating site data
R1du<-R1du[,2:3] #removing site column
sXspp<-left_join(kbsspplist,R1du,by="species.name") #joining to spp. list
colnames(sXspp)<-c("species.name","R1du") #renaming sXspp columns
#r1, dist. fert.
R1df<-kbsdata3[kbsdata3$site=="R1 disturbed fertilized",]
R1df<-R1df[,2:3]
sXspp<-left_join(sXspp,R1df,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df")
#r1, undist. unfert.
R1uu<-kbsdata3[kbsdata3$site=="R1 undisturbed unfertilized",]
R1uu<-R1uu[,2:3]
sXspp<-left_join(sXspp,R1uu,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu")
#r1, undist. fert.
R1uf<-kbsdata3[kbsdata3$site=="R1 undisturbed fertilized",]
R1uf<-R1uf[,2:3]
sXspp<-left_join(sXspp,R1uf,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf")
#r2, dist. unfert.
R2du<-kbsdata3[kbsdata3$site=="R2 disturbed unfertilized",]
R2du<-R2du[,2:3]
sXspp<-left_join(sXspp,R2du,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du")
#r2, dist. fert.
R2df<-kbsdata3[kbsdata3$site=="R2 disturbed fertilized",]
R2df<-R2df[,2:3]
sXspp<-left_join(sXspp,R2df,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df")
#r2, undist. unfert.
R2uu<-kbsdata3[kbsdata3$site=="R2 undisturbed unfertilized",]
R2uu<-R2uu[,2:3]
sXspp<-left_join(sXspp,R2uu,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                   "R2uu")
#r2, undist. fert.
R2uf<-kbsdata3[kbsdata3$site=="R2 undisturbed fertilized",]
R2uf<-R2uf[,2:3]
sXspp<-left_join(sXspp,R2uf,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
```

```
                    "R2uu","R2uf")
#r3, dist. unfert.
R3du<-kbsdata3[kbsdata3$site=="R3 disturbed unfertilized",]
R3du<-R3du[,2:3]
sXspp<-left_join(sXspp,R3du,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                    "R2uu","R2uf","R3du")
#r3, dist. fert.
R3df<-kbsdata3[kbsdata3$site=="R3 disturbed fertilized",]
R3df<-R3df[,2:3]
sXspp<-left_join(sXspp,R3df,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                    "R2uu","R2uf","R3du","R3df")
#r3, undist. unfert.
R3uu<-kbsdata3[kbsdata3$site=="R3 undisturbed unfertilized",]
R3uu<-R3uu[,2:3]
sXspp<-left_join(sXspp,R3uu,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                    "R2uu","R2uf","R3du","R3df","R3uu")
#r3, undist. fert.
R3uf<-kbsdata3[kbsdata3$site=="R3 undisturbed fertilized",]
R3uf<-R3uf[,2:3]
sXspp<-left_join(sXspp,R3uf,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                    "R2uu","R2uf","R3du","R3df","R3uu","R3uf")
#r4, dist. unfert.
R4du<-kbsdata3[kbsdata3$site=="R4 disturbed unfertilized",]
R4du<-R4du[,2:3]
sXspp<-left_join(sXspp,R4du,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                    "R2uu","R2uf","R3du","R3df","R3uu","R3uf","R4du")
#r4, dist. fert.
R4df<-kbsdata3[kbsdata3$site=="R4 disturbed fertilized",]
R4df<-R4df[,2:3]
sXspp<-left_join(sXspp,R4df,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                    "R2uu","R2uf","R3du","R3df","R3uu","R3uf","R4du","R4df")
#r4, undist. unfert.
R4uu<-kbsdata3[kbsdata3$site=="R4 undisturbed unfertilized",]
R4uu<-R4uu[,2:3]
sXspp<-left_join(sXspp,R4uu,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                    "R2uu","R2uf","R3du","R3df","R3uu","R3uf","R4du","R4df",
                    "R4uu")
#r4, undist. fert.
R4uf<-kbsdata3[kbsdata3$site=="R4 undisturbed fertilized",]
R4uf<-R4uf[,2:3]
sXspp<-left_join(sXspp,R4uf,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                    "R2uu","R2uf","R3du","R3df","R3uu","R3uf","R4du","R4df",
                    "R4uu","R4uf")
#r5, dist. unfert.
R5du<-kbsdata3[kbsdata3$site=="R5 disturbed unfertilized",]
```

```
R5du<-R5du[,2:3]
sXspp<-left_join(sXspp,R5du,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                   "R2uu","R2uf","R3du","R3df","R3uu","R3uf","R4du","R4df",
                   "R4uu","R4uf","R5du")
#r5, dist. fert.
R5df<-kbsdata3[kbsdata3$site=="R5 disturbed fertilized",]
R5df<-R5df[,2:3]
sXspp<-left_join(sXspp,R5df,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                   "R2uu","R2uf","R3du","R3df","R3uu","R3uf","R4du","R4df",
                   "R4uu","R4uf","R5du","R5df")
#r5, undist. unfert.
R5uu<-kbsdata3[kbsdata3$site=="R5 undisturbed unfertilized",]
R5uu<-R5uu[,2:3]
sXspp<-left_join(sXspp,R5uu,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                   "R2uu","R2uf","R3du","R3df","R3uu","R3uf","R4du","R4df",
                   "R4uu","R4uf","R5du","R5df","R5uu")
#r5, undist. fert.
R5uf<-kbsdata3[kbsdata3$site=="R5 undisturbed fertilized",]
R5uf<-R5uf[,2:3]
sXspp<-left_join(sXspp,R5uf,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                   "R2uu","R2uf","R3du","R3df","R3uu","R3uf","R4du","R4df",
                   "R4uu","R4uf","R5du","R5df","R5uu","R5uf")
#r6, dist. unfert.
R6du<-kbsdata3[kbsdata3$site=="R6 disturbed unfertilized",]
R6du<-R6du[,2:3]
sXspp<-left_join(sXspp,R6du,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                   "R2uu","R2uf","R3du","R3df","R3uu","R3uf","R4du","R4df",
                   "R4uu","R4uf","R5du","R5df","R5uu","R5uf","R6du")
#r6, dist. fert.
R6df<-kbsdata3[kbsdata3$site=="R6 disturbed fertilized",]
R6df<-R6df[,2:3]
sXspp<-left_join(sXspp,R6df,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                   "R2uu","R2uf","R3du","R3df","R3uu","R3uf","R4du","R4df",
                   "R4uu","R4uf","R5du","R5df","R5uu","R5uf","R6du","R6df")
#r6, undist. unfert.
R6uu<-kbsdata3[kbsdata3$site=="R6 undisturbed unfertilized",]
R6uu<-R6uu[,2:3]
sXspp<-left_join(sXspp,R6uu,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
                   "R2uu","R2uf","R3du","R3df","R3uu","R3uf","R4du","R4df",
                   "R4uu","R4uf","R5du","R5df","R5uu","R5uf","R6du","R6df",
                   "R6uu")
#r6, undist. fert.
R6uf<-kbsdata3[kbsdata3$site=="R6 undisturbed fertilized",]
R6uf<-R6uf[,2:3]
sXspp<-left_join(sXspp,R6uf,by="species.name")
colnames(sXspp)<-c("species.name","R1du","R1df","R1uu","R1uf","R2du","R2df",
```

```
                  "R2uu","R2uf","R3du","R3df","R3uu","R3uf","R4du","R4df",
                  "R4uu","R4uf","R5du","R5df","R5uu","R5uf","R6du","R6df",
                  "R6uu","R6uf")
#converting column 1 to row names (species names)
row.names(sXspp)<-sXspp$species.name
sXspp<-sXspp[,-1]
#converting all NAs to zeros
sXspp[is.na(sXspp)]<-0
#converting to numeric (columns all chr)
sXspp<-mutate_all(sXspp,function(x)as.numeric(x))
#also need to convert to matrix and transform
sXspp.mtx<-t(as.matrix(sXspp))

#reorganizing sXspp matrix so it will line up better with factor vector
#will be using treatments (df, du, uf, uu) as distinct treatment groups
sXspp.treat<-sXspp.mtx[c("R1du","R2du","R3du","R4du","R5du","R6du",
                         "R1df","R2df","R3df","R4df","R5df","R6df",
                         "R1uu","R2uu","R3uu","R4uu","R5uu","R6uu",
                         "R1uf","R2uf","R3uf","R4uf","R5uf","R6uf"),]
#Will make three factor vectors; one for treat combo, one for each individually
#treatment combo factors
combo.factor<-c(rep("DU",6),rep("DF",6),rep("UU",6),rep("UF",6))
#disturbance factors
dist.factor<-c(rep("DIST",12),rep("UNDIST",12))
#fertilizer factors
fert.factor<-c(rep("UNFERT",6),rep("FERT",6),rep("UNFERT",6),rep("FERT",6))

#will use multivariate/categorical hypothesis tests to look for spp. association
#w/ particular treatments

#will start with full combo treatments
#PERMANOVA - does community comp. differ between treatment combos
adonis2(sXspp.treat~combo.factor,method="bray",permutations=999)
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = sXspp.treat ~ combo.factor, permutations = 999, method = "bray")
##              Df SumOfSqs      R2      F Pr(>F)
## combo.factor  3   4.7811 0.57434 8.9955  0.001 ***
## Residual     20   3.5433 0.42566
## Total        23   8.3244 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#Indicator Value - do any spp. presences indicate particular treatment combo
combo.indval<-multipatt(sXspp.treat,cluster=combo.factor,func="IndVal.g",
                        control=how(nperm=999))
summary(combo.indval)
```

```
##
```

```
##  Multilevel pattern analysis
##  ---------------------------
##
##  Association function: IndVal.g
##  Significance level (alpha): 0.05
##
##  Total number of species: 60
##  Selected number of species: 22
##  Number of species associated to 1 group: 14
##  Number of species associated to 2 groups: 8
##  Number of species associated to 3 groups: 0
##
##  List of species associated to each combination:
##
##  Group DU  #sps.  3
##                                                        stat p.value
## Digitaria ischaemum (Schreb. Ex Schweig.) Schreb. ex Muhl. 0.976   0.001 ***
## Setaria pumila (Poir.) Roem. & Schult                      0.778   0.025 *
## Arabidopsis thaliana (L.) Heynh.                           0.707   0.038 *
##
##  Group UF  #sps.  3
##                                  stat p.value
## Silene alba (Mill.) E.H.L.Krause 0.913   0.001 ***
## Rumex crispus L.                 0.913   0.002 **
## Apocynum cannabinum L.           0.677   0.033 *
##
##  Group UU  #sps.  8
##                                                 stat p.value
## Poa pratensis L.                               0.996   0.001 ***
## Trifolium pratense L.                          0.969   0.001 ***
## Trifolium repens L.                            0.783   0.030 *
## Hieracium sp. (*)                              0.707   0.035 *
## Hypericum perforatum L.                        0.707   0.037 *
## Poa compressa L.                               0.707   0.039 *
## Taraxacum officinale Weber in                  0.705   0.047 *
## Centaurea stoebe L. ssp. micranthos (Gugler) Hayek 0.705   0.035 *
##
##  Group DF+DU  #sps.  7
##                                  stat p.value
## Abutilon theophrasti Medikus     1.000   0.001 ***
## Chenopodium album L.             1.000   0.001 ***
## Digitaria sanguinalis (L.) Scop. 1.000   0.001 ***
## Setaria faberi Herrm.            1.000   0.001 ***
## Cyperus esculentus L.            0.913   0.001 ***
## Amaranthus retroflexus L.        0.816   0.021 *
## Panicum dichotomiflorum Michx.   0.707   0.049 *
##
##  Group UF+UU  #sps.  1
##                          stat p.value
## Solidago canadensis L. 0.957   0.001 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#disturbance treatments
#PERMANOVA - does community comp. differ between disturbed and not
adonis2(sXspp.treat~dist.factor,method="bray",permutations=999)
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = sXspp.treat ~ dist.factor, permutations = 999, method = "bray")
##             Df SumOfSqs      R2      F Pr(>F)
## dist.factor  1   3.6726 0.44118 17.369  0.001 ***
## Residual    22   4.6518 0.55882
## Total       23   8.3244 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#Indicator Value - do any spp. presences indicate disturbed v. undisturbed
dist.indval<-multipatt(sXspp.treat,cluster=dist.factor,func="IndVal.g",
                       control=how(nperm=999))
summary(dist.indval)
```

```
##
##  Multilevel pattern analysis
##  ---------------------------
##
##  Association function: IndVal.g
##  Significance level (alpha): 0.05
##
##  Total number of species: 60
##  Selected number of species: 25
##  Number of species associated to 1 group: 25
##
##  List of species associated to each combination:
##
##  Group DIST  #sps.  13
##                                                      stat p.value
## Abutilon theophrasti Medikus                        1.000   0.001 ***
## Chenopodium album L.                                1.000   0.001 ***
## Digitaria sanguinalis (L.) Scop.                    1.000   0.001 ***
## Setaria faberi Herrm.                               1.000   0.001 ***
## Cyperus esculentus L.                               0.913   0.001 ***
## Digitaria ischaemum (Schreb. Ex Schweig.) Schreb. ex Muhl. 0.866   0.001 ***
## Amaranthus retroflexus L.                           0.816   0.003 **
## Ambrosia artemisiifolia L.                          0.764   0.023 *
## Oxalis stricta L.                                   0.720   0.038 *
## Eragrostis cilianensis (All.) E.Mosher              0.707   0.024 *
## Panicum dichotomiflorum Michx.                      0.707   0.024 *
## Setaria pumila (Poir.) Roem. & Schult               0.707   0.014 *
## Polygonum pensylvanicum L.                          0.645   0.032 *
##
##  Group UNDIST  #sps.  12
##                                  stat p.value
```

```
## Solidago canadensis L.           0.957   0.001 ***
## Poa pratensis L.                 0.913   0.001 ***
## Trifolium pratense L.            0.816   0.003 **
## Elymus repens (L.) Gould         0.763   0.010 **
## Asclepias syriaca L.             0.707   0.012 *
## Trifolium repens L.              0.707   0.020 *
## Apocynum cannabinum L.           0.645   0.034 *
## Aster pilosus Willd.             0.645   0.038 *
## Aster sagittifolius              0.645   0.041 *
## Phleum pratense L.               0.645   0.038 *
## Silene alba (Mill.) E.H.L.Krause 0.645   0.032 *
## Rumex crispus L.                 0.645   0.041 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#fertilizer treatments
#PERMANOVA - does community comp. differ between fertilized and not
adonis2(sXspp.treat~fert.factor,method="bray",permutations=999)
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = sXspp.treat ~ fert.factor, permutations = 999, method = "bray")
##             Df SumOfSqs      R2      F Pr(>F)
## fert.factor  1   0.5543 0.06659 1.5694  0.147
## Residual    22   7.7701 0.93341
## Total       23   8.3244 1.00000
```

```
#does not seem to show differences between fertilized and unfertilized
#communities, so will focus on dist/undist and interaction treatments
```

After loading in my dataset, I decided to test whether specific plant communities from the KBS microplot were associated with particular treatments: fertilized/unfertilized, disturbed/undisturbed, or fully-factorial combinations of the these. Since these treatments were categorical, I used the multivariate procedures for categorical designs to test my hypothesis. I initially used a PERMANOVA to assess whether there were notable differences in plant community between the different treatment categories. This testing did not find a significant difference in plant community between fertilized and unfertilized groups ($F_{1,22} = 1.5694$, $p = 0.157$), but there were notable differences between disturbed and undisturbed communities ($F_{1,22} = 17.369$, $p = 0.001$) *and* differences between disturbed/unfertilized, disturbed/fertilized, undisturbed/unfertilized, and undisturbed/fertilized communities ($F_{3,20} = 8.9955$, $p = 0.001$).

I conducted more tests for the disturbance and factorial treatments to gather more information about the community differences. By calculating an indicator value, I was able to assess whether any species may be particularly associated with a specific treatment. 25 of 60 species were significantly associated with one of the disturbance treatments (detailed in the summary output above). The species most associated with disturbance plots were *Abutilon theophrasti Medikus* (1.000, $p = 0.001$), *Chenopodium album L.* (1.000, $p = 0.001$), *Digitaria sanguinalis (L.) Scop.* (1.000, $p = 0.001$), and *Setaria faberi Herrm.* (1.000, $p = 0.001$); species most associated with undisturbed plots included *Solidago canadensis L.* (0.957, $p = 0.001$), *Poa pratensis L.* (0.913,

p = 0.002), *Trifolium pratense L.* (0.816, p = 0.003), and *Elymus repens (L.) Gould* (0.763, p = 0.016). Notably, no species was significantly associated with both disturbed and undisturbed plots.

I also calculated indicator values for the factorial treatments to see if any species were significantly associated with any of the disturbance/fertilization treatment combinations. In this analysis, 22 of the 60 species were significantly associated with a treatment group, with 15 being connected to only 1 group and 7 with 2 groups. Three species were associated with the disturbed/unfertilized group (*D. ischaemum*, *S. pumila*, *A. thaliana*), two were associated with the undisturbed/fertilized group (*S. alba*, *R. crispus*), and ten were associated with the undisturbed/unfertilized group (*P. pratensis*, *T. pratense*, *T. repens*, *A. millefolium*, *H. sp.*, *H. perforatum*, *P. compressa*, *T. hybridum*, *T. officinale*, *C. stoebe*). Some of the strongest associations were with both of the disturbed treatments, including *A. theophrasti Medikus* (1.000, p = 0.001), *C. album L.* (1.000, p = 0.001), *D. sanguinalis (L.) Scop.* (1.000, p = 0.001), and *S. faberi Herrm.* (1.000, p = 0.001). The undisturbed treatments were also strongly associated with the presence of *S. canadensis L.* (0.957, p = 0.001).

Based on the results of these tests, I would say that environmental treatment does drive plant community assembly in the KBS LTER microplots. While fertilization does seem to play a role in undisturbed environments, the strength of associations between species and treatments suggests that disturbance regime may be a primary driver of community assembly in these plots. Several plants seem to prioritize colonization following disturbance, while goldenrod is really only found in undisturbed areas. If I can incorporate data from more years of the LTER dataset, I would be curious to see if this pattern has been consistent over time.