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

Emmi Mueller; Z620: Quantitative Biodiversity, Indiana University

13 February, 2019

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 13th, 2019 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 "*/8.BetaDiversity*" folder, and
4. load the `vegan` R package (be sure to install if needed).

```
rm(list=ls())
setwd("~/GitHub/QB2019_Mueller/2.Worksheets/8.BetaDiversity")
getwd
```

```
## function ()
## .Internal(getwd())
## <bytecode: 0x00000000193956b8>
## <environment: namespace:base>
```

```r
require(vegan)
```

```
## Loading required package: vegan
## Loading required package: permute
## Loading required package: lattice
## This is vegan 2.5-4
```

2) LOADING DATA

Load dataset

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

```r
# note, pleae do not print the dataset when submitting
require(ade4)
```

```
## Loading required package: ade4
```

```r
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.

```r
fish <- doubs$fish
fish <- fish[-8,]
quality <- c(rep("HQ", 13), rep("MQ", 5), rep("LQ", 6), rep("MQ", 5))

adonis(fish ~ quality, method = "bray", permutations = 999)
```

```
##
## Call:
## adonis(formula = fish ~ quality, permutations = 999, method = "bray")
##
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
##
##           Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
## quality    2    3.0947 1.54733   10.97 0.45765  0.001 ***
## Residuals 26    3.6674 0.14105         0.54235
## Total     28    6.7621                 1.00000
## ---
## 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***: Quality is a significant predictor of fish community composition. This is consistent with the visualizations that showed that sites were grouped together by region of the river. If river quality is grouped by sites then it makes sense that sites grouped together would be dependent on quality.

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.

```
fish.dist <- vegdist(fish, method = "bray")
env.dist <- vegdist(scale(doubs$env[-8,]), method = "euclid")

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%
## 0.103 0.143 0.173 0.193
## 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***: Fish community composition is correlated with environmental variables. This means that river quality is likely correlated with the environmental factors that are measured in the env matrix.

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.

```
env.chem <- as.matrix(doubs$env[-8,5:11])

doubs.dbrda <- dbrda(fish.dist ~ ., as.data.frame(env.chem))

require(psych)
```

```
## Loading required package: psych
```

```
psych::corr.test(env.chem)
```

```
## Call:psych::corr.test(x = env.chem)
## Correlation matrix
##        pH   har   pho   nit   amm   oxy   bdo
## pH   1.00  0.08 -0.08 -0.04 -0.12  0.19 -0.16
## har  0.08  1.00  0.37  0.53  0.30 -0.37  0.34
## pho -0.08  0.37  1.00  0.80  0.97 -0.76  0.91
## nit -0.04  0.53  0.80  1.00  0.80 -0.69  0.68
## amm -0.12  0.30  0.97  0.80  1.00 -0.75  0.90
## oxy  0.19 -0.37 -0.76 -0.69 -0.75  1.00 -0.84
## bdo -0.16  0.34  0.91  0.68  0.90 -0.84  1.00
## Sample Size
## [1] 29
## Probability values (Entries above the diagonal are adjusted for multiple tests.)
##       pH  har  pho  nit  amm  oxy  bdo
## pH  0.00 1.00 1.00 1.00 1.00 1.00 1.00
## har 0.66 0.00 0.46 0.03 0.83 0.46 0.59
## pho 0.68 0.05 0.00 0.00 0.00 0.00 0.00
## nit 0.83 0.00 0.00 0.00 0.00 0.00 0.00
## amm 0.53 0.12 0.00 0.00 0.00 0.00 0.00
## oxy 0.32 0.05 0.00 0.00 0.00 0.00 0.00
## bdo 0.40 0.07 0.00 0.00 0.00 0.00 0.00
##
##  To see confidence intervals of the correlations, print with the short=FALSE option
```

```
doubs.dbrda.mod0 <- dbrda(fish.dist ~ 1, as.data.frame(env.chem))

doubs.dbrda.mod1 <- dbrda(fish.dist ~ ., as.data.frame(env.chem))
doubs.dbrda <- ordiR2step(doubs.dbrda.mod0, doubs.dbrda.mod1, perm.max = 200)
```

```
## Step: R2.adj= 0
## Call: fish.dist ~ 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.dist ~ 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.dist ~ 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.dist ~ oxy + bdo + nit
##
##                 R2.adjusted
## + amm             0.5415705
## <All variables>   0.5303258
## + pho             0.5277128
## + har             0.5218852
## <none>            0.4980793
## + pH              0.4843267
```

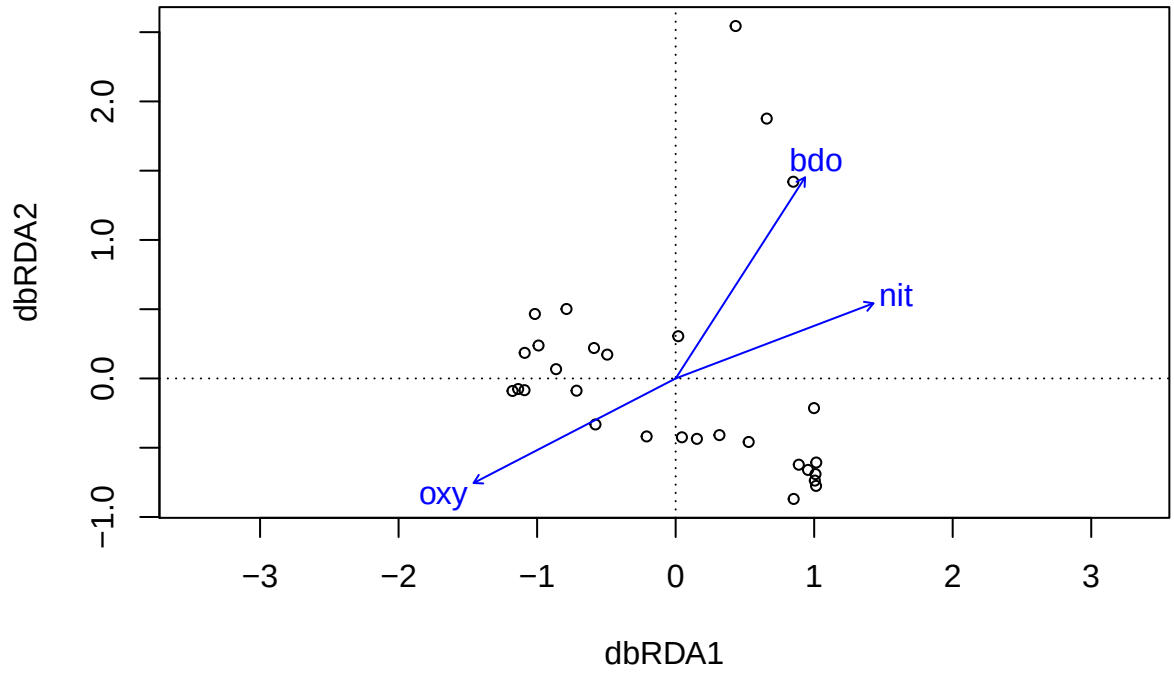
```
doubs.dbrda$call
```

```
## dbrda(formula = fish.dist ~ oxy + bdo + nit, data = as.data.frame(env.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)
```




```
permutest(doubs.dbrda, permutations = 999)
```

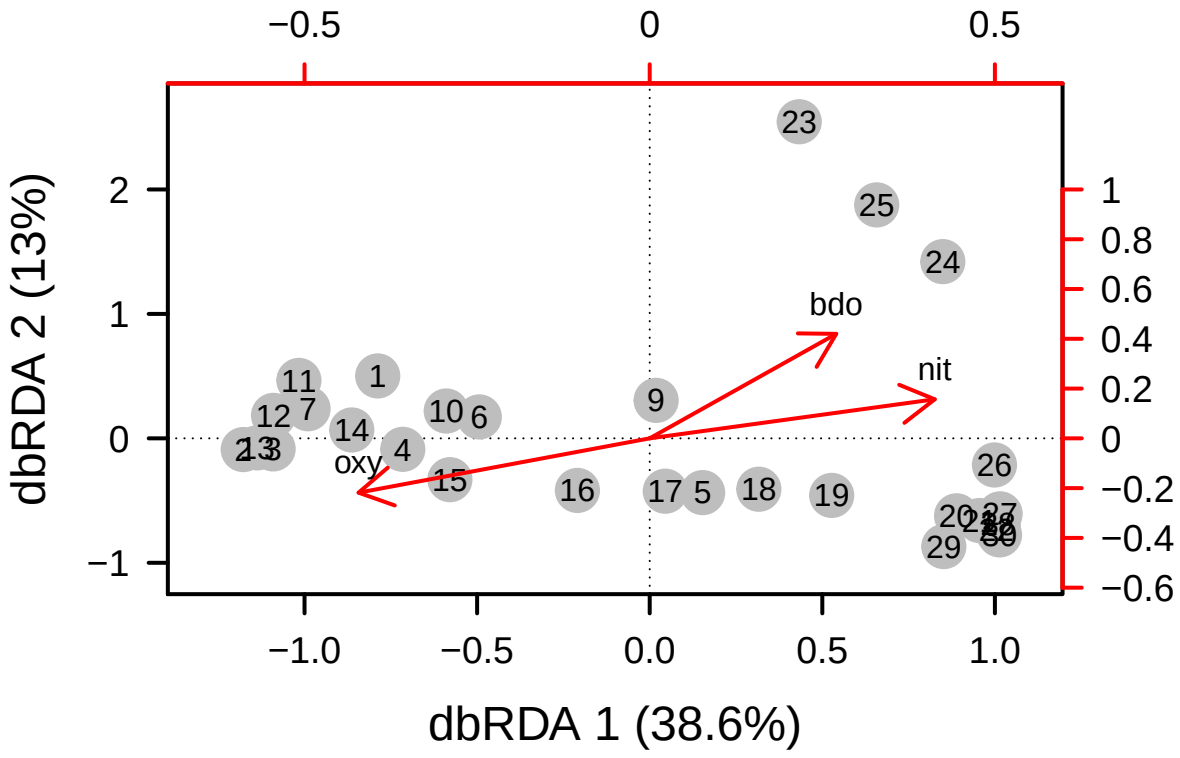
```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.dist ~ oxy + bdo + nit, data =
## as.data.frame(env.chem))
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     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, env.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
```

```
dbrda.explainvar1 <- round(doubs.dbrda$CCA$eig[1]/ sum(c(doubs.dbrda$CCA$eig, doubs.dbrda$CA$eig)), 3)
dbrda.explainvar2 <- round(doubs.dbrda$CCA$eig[2]/ sum(c(doubs.dbrda$CCA$eig, doubs.dbrda$CA$eig)), 3)

par(mar = c(5,5,4,4) + 0.1)

plot(scores(doubs.dbrda, display = "wa"), xlim = c(-1.3, 1.1), ylim = c(-1.1, 2.7), xlab = paste("dbRDA

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)

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")))

vectors <- scores(doubs.dbrda, display = "bp")

arrows(0, 0, vectors[,1], vectors[,2], lwd = 2, lty = 1, length = 0.2, col = "red")
text(vectors[,1], vectors[,2], pos = 3, labels = row.names(vectors))
axis(side=3, lwd.ticks = 2, cex.axis=1.2, las = 1, col = "red", lwd = 2.2, at = pretty(range(vectors[,1]
axis(side=4, lwd.ticks = 2, cex.axis=1.2, las = 1, col = "red", lwd = 2.2, at = pretty(range(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 groups of correlated variables that are likely contributing to fish community structure are dissolved oxygen levels, nitrate levels and biological demand for oxygen.

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.

```
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
```

```
env.mod <- model.matrix(~ oxy + bdo+ nit, as.data.frame(env.chem))[,-1]

rs <- rowSums(fish)/sum(fish)
doubs.pcnmw <- pcnm(dist(doubs$xy[-8,]), w = rs, dist.ret = T)
doubs.pcnmw$values > 0
```

```
##  [1]  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
## [12]  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE FALSE FALSE FALSE FALSE FALSE
## [23] FALSE FALSE FALSE FALSE
```

```
doubs.space <-as.data.frame(scores(doubs.pcnmw))
doubs.pcnm.mod0 <- dbrda(fish.dist ~ 1, doubs.space)
doubs.pcnm.mod1 <- dbrda(fish.dist ~ ., doubs.space)
step.pcnm <- ordiR2step(doubs.pcnm.mod0, doubs.pcnm.mod1, perm.max = 200)
```

```
## Step: R2.adj= 0
## Call: fish.dist ~ 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.dist ~ PCNM2
##
##                  R2.adjusted
## <All variables>   0.6260113
## + PCNM3           0.3429270
## + PCNM5           0.3057368
## + PCNM1           0.2885396
## + PCNM16          0.2786746
## + PCNM14          0.2744520
```

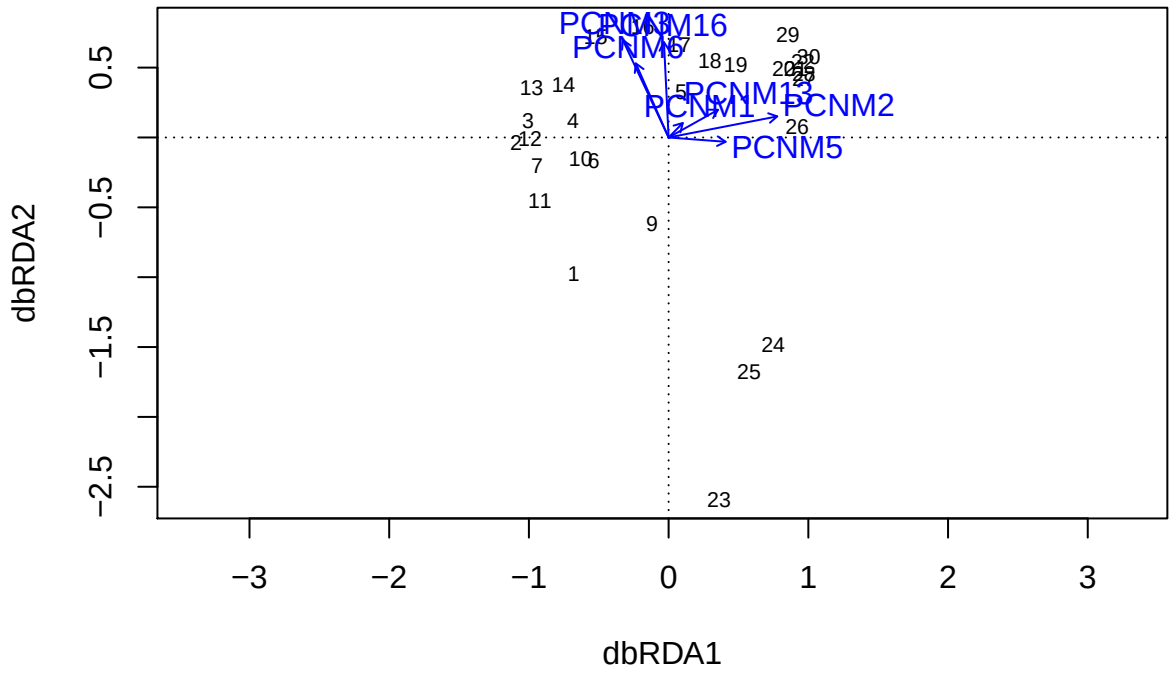
```
## + PCNM15          0.2692809
## + PCNM6           0.2659866
## + PCNM13          0.2636194
## + PCNM9           0.2517847
## + PCNM8           0.2496240
## + PCNM10          0.2434688
## + PCNM7           0.2431476
## + PCNM17          0.2404343
## + PCNM11          0.2366833
## <none>            0.2353704
## + PCNM12          0.2288789
## + PCNM4           0.2189522
##
##         Df    AIC      F Pr(>F)
## + PCNM3  1 46.083 5.4196  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.342927
## Call: fish.dist ~ 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   0.01 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.407602
## Call: fish.dist ~ 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  0.004 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.4721469
## Call: fish.dist ~ 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  0.004 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.5212427
## Call: fish.dist ~ 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  0.022 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.5767968
## Call: fish.dist ~ 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   0.04 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.6043089
## Call: fish.dist ~ 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  0.056 .
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
plot(step.pcnm)
```




```
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  0.002 **
## + PCNM5         0.40760  1 43.941 3.8385  0.010 **
## + PCNM1         0.47215  1 41.411 4.0570  0.004 **
## + PCNM13        0.52124  1 39.346 3.4612  0.004 **
## + PCNM16        0.57680  1 36.480 4.0192  0.022 *
## + PCNM6         0.60431  1 35.182 2.5296  0.040 *
## <All variables> 0.62601
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
space.mod <- model.matrix(~ PCNM2 + PCNM3 + PCNM5 + PCNM1 + PCNM13 +PCNM16 + PCNM6, doubs.space)[,-1]

doubs.total.env <- dbrda(fish.dist ~ env.mod)
doubs.total.space <- dbrda(fish.dist ~ space.mod)

doubs.env.cond.space <- dbrda(fish.dist ~ env.mod + Condition(space.mod))
doubs.space.cond.env <- dbrda(fish.dist ~ space.mod + Condition(env.mod))

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

```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.dist ~ env.mod + Condition(space.mod))
## Permutation test for all constrained eigenvalues
##          Df Inertia     F Pr(>F)
## Model     3 0.85158 4.423  0.001 ***
## 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.dist ~ space.mod + Condition(env.mod))
## 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.dist ~ env.mod)
## 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)
```

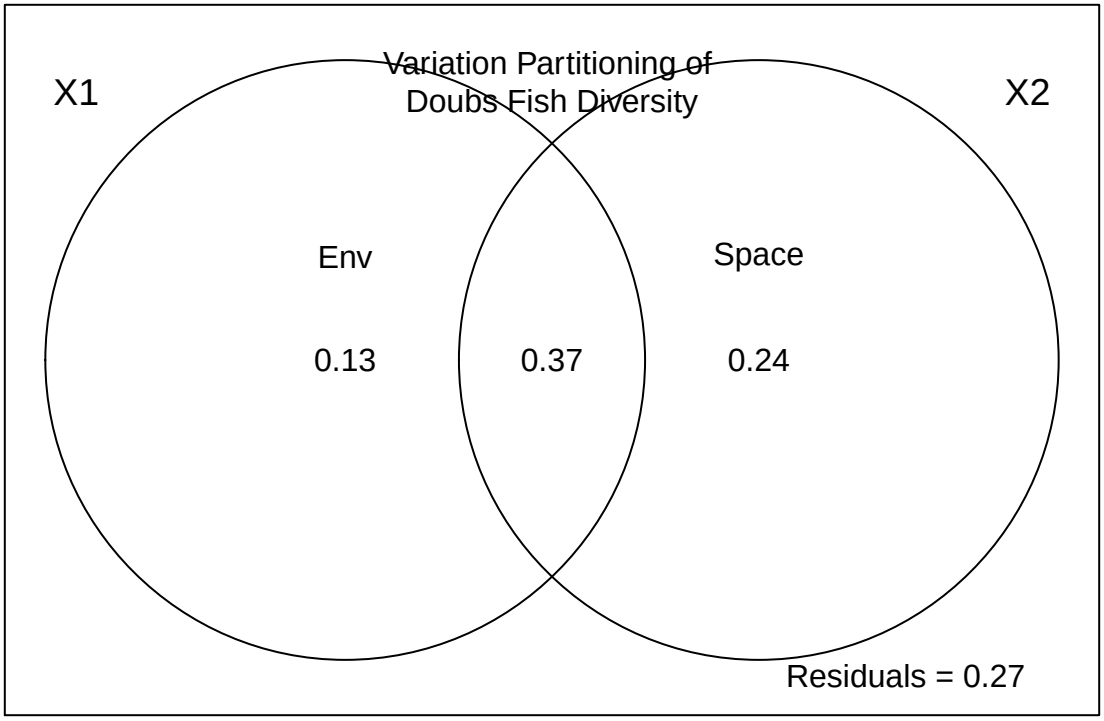
```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = fish.dist ~ space.mod)
```

```
## 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
```

```
doubs.varpart <- varpart(fish.dist, env.mod, space.mod)
doubs.varpart
```

```
##
## Partition of squared Bray distance in dbRDA
##
## Call: varpart(Y = fish.dist, X = env.mod, space.mod)
##
## Explanatory tables:
## X1:  env.mod
## X2:  space.mod
##
## No. of explanatory tables: 2
## Total variation (SS): 6.7621
## No. of observations: 29
##
## Partition table:
##                       Df R.squared Adj.R.squared Testable
## [a+b] = 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]                    0                 0.36813    FALSE
## [c] = X2|X1            7                 0.23618     TRUE
## [d] = Residuals                          0.26574    FALSE
## ---
## Use function 'dbrda' to test significance of fractions of interest
```

```
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***: Spatially structured environments have the largest effect on fish diversity. Space and environment individually play less of a role in

SYNTHESIS

1) Using the jelly bean data from class (i.e., JellyBeans.txt), perform a PERMANOVA to test whether or not the vicariance event (random splitting of source community) had an affect on jelly bean composition. Based on your previous analyses with this data set, what are your thoughts about the importance of stochastic vs. deterministic factors on estimates of biodiversity?

```
jellyAB <- read.table("./JellyBeans_site.txt", header = TRUE, sep = "\t", row.names = 1)
groups <- c("A", "A", "A","B","B","A","B","A", "B")
jellyAB[,"GreenTrans"] <- jellyAB[,"GreenTrans"]+jellyAB[,"GreenTrans2"]
jellyAB[,"Rainbow"] <- jellyAB[,"WhiteSolid"] + jellyAB[,"Rainbow"]
jellyAB[,c("Group", "GreenTrans2", "WhiteSolid")] <- list(NULL)

adonis(jellyAB ~ groups, method = "bray", permutations = 999)
```

```
##
## Call:
## adonis(formula = jellyAB ~ groups, permutations = 999, method = "bray")
##
## Permutation: free
## Number of permutations: 999
##
## Terms added sequentially (first to last)
```

```
##
##           Df SumsOfSqs  MeanSqs F.Model     R2 Pr(>F)
## groups     1   0.08741 0.087413  1.9732 0.2199   0.04 *
## Residuals  7   0.31011 0.044301         0.7801
## Total      8   0.39752                  1.0000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Synthesis 1: The PERMANOVA shows that the group is a significant predictor of jelly bean community composition. This means that the deterministic factor of the splitting of the source community plays an important role in determining final community compositions, something that has been supported by clustering yet stochastic factors still play a role in shaping the community as can be seen by the large R^2 value of the residuals.

2) Load the dataset you are using for your Team Project. Perform an ordination to visualize your dataset. Using this ordination, develop some hypotheses relevant to β-diversity. Use a statistic to test one of these hypotheses. Succinctly explain the finding and its relevance to your system.

dbRDA

require("OTUtable") data(otu_table) otu_table <- as.data.frame(t(otu_table)) JUL07 <- c("CBE02JUL07.R1", "FBE02JUL07.R1", "HKE04JUL07.R1", "MAE04JUL07.R1", "NSE03JUL07.R1", "SSE03JUL07.R1", "TBE02JUL07.R1", "WSE03JUL07") otu_JUL07 <- otu_table[grep(paste(JUL07, collapse = "|"), row.names(otu_table)),] row.names(otu_JUL07) <- c("Crystal Bog", "Forestry Bog", "Hell's Kitchen", "Mary Lake", "North Sparkling Bog", "South Sparkling Bog", "Trout Bog", "West Sparkling Bog")

Set up metadata for JUL07 sites data(metadata) rows <- c(4962, 4971,4977,4982, 4988, 4998, 5004,5025) metadata_JUL07 <- metadata[rows, -1] row.names(metadata_JUL07) <- c("Trout Bog", "Crystal Bog", "Forestry Bog","North Sparkling Bog", "Sourth Sparkling Bog", "West Sparkling Bog", "Hell's Kitchen", "Mary Lake") num = 1 rows <- c(4970,4976,4987,4981,4997, 5003, 5024, 5048) while(num < nrow(metadata_JUL07)){ depth_current <- metadata$Depth[rows[num]] metadata_JUL07Depth[num] <- depth_current num <- num + 1 } metadata_JUL07 <- metadata_JUL07[,c("Depth", "DO", "Temperature")] bogcca <- cca(otu_JUL07, metadata_JUL07)

Mantel Test jul07.dist <- vegdist(otu_JUL07, method = "bray") jul07.env.dist <- vegdist(scale(metadata_JUL07), method = "euclid") mantel(jul07.dist, jul07.env.dist) dbRDA from metadata metadata_JUL07 <- as.matrix(metadata_JUL07) print(as.data.frame(metadata_JUL07))

bog.dbrda <- dbrda(jul07.dist ~ ., as.data.frame(metadata_JUL07))

require(psych) psych::corr.test(metadata_JUL07)

bog.dbrda.modO <- dbrda(jul07.dist ~ 1, as.data.frame(metadata_JUL07)) bog.dbrda.mod1 <- dbrda(jul07.dist ~ ., as.data.frame(metadata_JUL07)) bog.dbrda <- ordiR2step(bog.dbrda.modO, bog.dbrda.mod1, perm.max = 200)

bog.dbrda$call bog.dbrda anova ordiplot(bog.dbrda)

permutest(bog.dbrda, permutations = 999) envfit(bog.dbrda, metadata_JUL07, perm = 999)

dbrda.explainvar1 <- round(bog.dbrda$CCA eig[1]/ sum(c(bog.dbrda$CCA eig, bog.dbrda$CA eig)), 3) * 100 dbrda.explainvar2 <- round(bog.dbrda$CCA eig[2]/ sum(c(bog.dbrda$CCA eig, bog.dbrda$CA eig)), 3) * 100

par(mar = c(5,5,4,4) + 0.1)

plot(scores(bog.dbrda, display = "wa"), xlim = c(-1.3, 1.1), ylim = c(-1.1, 2.7), xlab = paste("dbRDA 1 (", dbrda.explainvar1, "%)", sep = ""), ylab = paste("dbRDA 2 (", dbrda.explainvar2, "%)", sep = ""), pch = 16, cex = 2.0, type = "n", cex.lab = 1.5, cex.axis = 1.2, axes = FALSE)

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)

points(scores(bog.dbrda, display = “wa”), pch = 19, cex = 3, bg = “gray”, col = “gray”) text(scores(bog.dbrda, display = “wa”), labels = row.names(scores(bog.dbrda, display = “wa”)))

vectors <- scores(bog.dbrda, display = “bp”)

arrows(0, 0, vectors[,1], vectors[,2], lwd = 2, lty = 1, length = 0.2, col = “red”) text(vectors[,1], vectors[,2], pos = 3, labels = row.names(vectors)) axis(side=3, lwd.ticks = 2, cex.axis=1.2, las = 1, col = “red”, lwd = 2.2, at = pretty(range(vectors[,1]))*2, labels = pretty(range(vectors[,1]))) axis(side=4, lwd.ticks = 2, cex.axis=1.2, las = 1, col = “red”, lwd = 2.2, at = pretty(range(vectors[,2]))*2, labels = pretty(range(vectors[,2])))

Synthesis 2: We predict that beta diversity will decrease with increasing distance between sites. If this is not the case, we predict that beta diversity will be correlated with an environmental variable such as DO or pH.