Spatial Analysis of Zoobenthos Data

Trang Nguyen; El Park

05 mars, 2025

Beta Diversity analysis

We begin by loading the zoobenthos data and calculating the Bray-Curtis distance to measure the dissimilarity between sites.

```
# Load packages
library(vegan)
```

```
## Warning: le package 'vegan' a été compilé avec la version R 4.4.2

## Le chargement a nécessité le package : permute

## Warning: le package 'permute' a été compilé avec la version R 4.4.2

## Le chargement a nécessité le package : lattice

## This is vegan 2.6-8
```

```
package.list = c("ade4", "vegan", "viridis", "gplots", "BiodiversityR", "indicspecies")
for (package in package.list) {
  if (!require(package, character.only = TRUE, quietly = TRUE)) {
    install.packages(package, dependencies = TRUE)
    library(package, character.only = TRUE)
  }
}
```

```
## Warning: le package 'ade4' a été compilé avec la version R 4.4.2

## Warning: le package 'viridis' a été compilé avec la version R 4.4.2

## Warning: le package 'gplots' a été compilé avec la version R 4.4.2

##
## Attachement du package : 'gplots'

## L'objet suivant est masqué depuis 'package:stats':
##
##     lowess
```

```
## Warning: le package 'BiodiversityR' a été compilé avec la version R 4.4.2
```

```
## BiodiversityR 2.17-1.1: Use command BiodiversityRGUI() to launch the Graphical User Interface;
## to see changes use BiodiversityRGUI(changeLog=TRUE, backward.compatibility.messages=TRUE)
```

```
## Warning: le package 'indicspecies' a été compilé avec la version R 4.4.2
```

```
# Load SbS dataset
data = read.csv("data/sbys_s013.csv", row.names= 1)
# data
# Calculate Bray-Curtis distance
data.db = vegdist(data, method = "bray")

site_labels = rownames(data)

################################################3

## SAve the plot as a PNG file
png("plots/heatmap_s013.png", width = 1000, height = 1000)
# Heatmap of dissimilarities
levelplot(as.matrix(data.db),
  col.regions = viridis,
  aspect="iso",
  xlab = "Sites", ylab = "Sites",
  scales = list(cex=1,
  x=list(labels=site_labels, rot=90),
  y=list(labels=site_labels)),
  main = "Bray-Curtis Distance")

#  Close the png device
dev.off()
```

```
## pdf
##   2
```

```
#############################################################3
# Perform cluster analysis
data.ward = hclust(data.db, method = "ward.D2")
cut_avg = cutree(data.ward, k = 3)

png("plots/dendrogram_s013.png", width = 1000, height = 1000)
# Plot cluster
par(mar=c(1,5,2,2) + 0.1)
plot(data.ward, main = "Ward's Clustering of SbS data",
  ylab = "squared Bray_Curtis distance")

dev.off()
```

```
## pdf
##   2
```

```
#########################################################
# PCOA analysis for all sites
data.pcoa = cmdscale(data.db, eig = TRUE, k = 3)
data.pcoa
```

```
## $points
##                [,1]         [,2]        [,3]
## S0_1983 -0.15560893 -0.193585510  0.20331257
## S0_1984 -0.24616559 -0.271698535 -0.01294091
## S0_1986 -0.23529482 -0.386167485 -0.07756391
## S0_1988 -0.29483776 -0.106645057 -0.14991417
## S0_1989 -0.19062732 -0.201434831 -0.18490467
## S0_1990 -0.25673701 -0.163010006 -0.14750926
## S0_1991 -0.22731709 -0.259705665 -0.17372074
## S0_1992 -0.12806920 -0.278832762 -0.04064684
## S0_1993 -0.13649587 -0.314097988 -0.02766789
## S0_1994 -0.19289002 -0.065297877 -0.32890401
## S0_1995 -0.24987578 -0.313929190 -0.01768426
## S0_1997 -0.19084212  0.160824672 -0.40569071
## S0_1998 -0.17241732  0.099698848 -0.43572595
## S0_1999 -0.23708659 -0.132447985 -0.24764423
## S0_2000 -0.10077208  0.028390460 -0.32638727
## S0_2001 -0.19029072  0.089003296 -0.42450776
## S0_2002 -0.19461566 -0.348058559  0.07597657
## S0_2003 -0.22576775 -0.074372933 -0.16882568
## S0_2004 -0.16939242  0.041666018 -0.38124742
## S1_1980  0.28095147  0.074480780  0.09128488
## S1_1987  0.31181638  0.077082077  0.07992974
## S1_1991  0.25177956  0.063849927  0.11977704
## S1_1992  0.37702347  0.015518837  0.08308229
## S1_1993  0.54687503 -0.063738507 -0.04380513
## S1_1994  0.49615803 -0.051908444 -0.06413993
## S1_1995  0.47232239 -0.042289993 -0.10521555
## S1_1997  0.56697252 -0.029030762 -0.05028244
## S1_1998  0.54971674 -0.053575724 -0.05753376
## S1_1999  0.50474992 -0.041404191 -0.10392713
## S1_2000  0.56822895 -0.033036459 -0.02997936
## S1_2001  0.49557718 -0.009141980  0.04196198
## S1_2002  0.43501976  0.035170006  0.03785900
## S1_2003  0.56505201 -0.042884607 -0.03360490
## S1_2004  0.52480425 -0.008812771 -0.02202984
## S3_1980 -0.09254835  0.308068457 -0.22729212
## S3_1981 -0.08837966  0.380417130  0.13405771
## S3_1982 -0.13432468  0.110130306  0.34959448
## S3_1983 -0.09733450  0.032133919  0.31981050
## S3_1984 -0.13564434 -0.283346702  0.21262610
## S3_1986 -0.18768115 -0.291526392  0.26611049
## S3_1987 -0.16594921 -0.042353200  0.38587453
## S3_1989 -0.09070486 -0.005192868  0.31864485
## S3_1990 -0.19425708  0.321939607  0.18404027
## S3_1991 -0.11435463  0.360552323  0.04431404
## S3_1993 -0.13089518  0.390981517  0.10535137
## S3_1994 -0.17129172 -0.200488316  0.30912563
```

```
## S3_1995 -0.11739164 -0.111948883  0.29385295
## S3_1997 -0.23077892  0.375048530  0.04391799
## S3_1998 -0.20118621  0.358800305  0.05503038
## S3_2000 -0.22125347 -0.143160967  0.37666094
## S3_2001 -0.23141953  0.264020085  0.09041054
## S3_2002 -0.18085858  0.381987960  0.12127939
## S3_2003 -0.18104904  0.454584819  0.01226069
## S3_2004 -0.18464084  0.138775271 -0.06685108
##
## $eig
##  [1]  4.723028e+00  2.500742e+00  2.288341e+00  1.490361e+00  1.299183e+00
##  [6]  9.748746e-01  8.718222e-01  7.789455e-01  6.869117e-01  6.180092e-01
## [11]  5.256849e-01  4.766561e-01  4.566266e-01  4.178012e-01  4.070245e-01
## [16]  3.447875e-01  3.252391e-01  2.660860e-01  2.495823e-01  2.206050e-01
## [21]  2.068451e-01  1.695843e-01  1.563930e-01  1.087973e-01  1.030388e-01
## [26]  9.862338e-02  9.022524e-02  8.360130e-02  7.687993e-02  6.480135e-02
## [31]  5.784519e-02  4.735596e-02  4.567347e-02  3.264303e-02  2.725644e-02
## [36]  1.989911e-02  1.174053e-02  8.063674e-03  3.262438e-03  1.001407e-03
## [41] -3.885781e-16 -1.003417e-03 -5.456959e-03 -7.670527e-03 -1.164341e-02
## [46] -1.643364e-02 -2.138521e-02 -2.390413e-02 -3.002633e-02 -4.090361e-02
## [51] -4.391739e-02 -6.170906e-02 -7.747057e-02 -8.296922e-02
##
## $x
## NULL
##
## $ac
## [1] 0
##
## $GOF
## [1] 0.4371307 0.4458278
```

```
# Explained variance for first 3 axes
explained_variance = round(sum(data.pcoa$eig[1:3]/sum(data.pcoa$eig)) * 100, 2)
print("Explained variance of the 3 first axes:")
```

```
## [1] "Explained variance of the 3 first axes:"
```

```
print(explained_variance)
```

```
## [1] 45.49
```

```
png("plots/pcoa_s013.png", width = 1000, height = 1000)
# Define Plot Parameters
par(mar=c(1,5,2,2) + 0.1)

# Initiate Plot
plot(data.pcoa$points[ ,1],
  data.pcoa$points[ ,2],
     xlab = "PCoA 1",
     ylab = "PCoA 2",
     pch = 16, cex = 2,  type = "n", cex.lab = 1.5,
     cex.axis = 0.5, axes = FALSE)
```

```r
# # Add 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)

# # Add Points & Labels
points(data.pcoa$points[ ,1],
  data.pcoa$points[ ,2],
  pch = 19, cex = 3, bg = "gray", col = "gray")
text(data.pcoa$points[ ,1], data.pcoa$points[ ,2],
     labels = row.names(data.pcoa$points))

# Add species on top
dataREL <- data
  for(i in 1:nrow(data)){
    dataREL[i, ] = data[i, ] / sum(data[i, ])
    }

# Now, we use this information to calculate and add species scores
data.pcoa <- add.spec.scores(
  data.pcoa,
  dataREL,
  method = "pcoa.scores")

text(
  data.pcoa$cproj[ ,1],
  data.pcoa$cproj[ ,2],
  labels = row.names(data.pcoa$cproj), col = "red",
  cex=1)

dev.off()
```

```
## pdf
##   2
```

Cluster Analysis

The heatmap shows distinct blocks of similarity between sites sampled in different years. Sites that are sampled over multiple years form blocks, which reflects their relatively stable zoobenthos community composition.

I think species nearer to the center are more present in most sites. While species far from the origin are more specialized or strongly associated with particular site conditions. There are several patterns : Macoma balthica is strongly associated with site 1. This suggest that this species may adapt well in this site' specific condition.

Diastylis rathkei is more abundant in these site 3.

Arctica islandica and Hetermastus filiformis are more abundant in site 0.

After our previous analysis of the zoobenthos data, we discovered that over the past 20 years there were only three distinct sites where we had environmental data over time (at least 2+ years). To work around this, we treated each site's data for each year as a separate sample. When we

examined the heatmap of zoobenthos abundance, we saw that the samples clustered into three main groups, which makes sense given that there are three sites. However, we also noticed that sites 0 and 2 tend to be similar to each other, while site 1 maintains a more consistent zoobenthos composition over the years. Furthermore, Ward's Clustering showed quite an interesting pattern: sites 0 and 3 are sometimes more similar to each other than to themselves in certain years

Now, we will now use PERMANOVA to determine if the zoobenthos composition differs significantly between sites and across years.

```
library(vegan)
library(indicspecies)
groups = cut_avg
data$groups = cut_avg
# Run PERMANOVA
adonis2(data ~ groups, method="bray", permutations=999)
```

```
## Permutation test for adonis under reduced model
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = data ~ groups, permutations = 999, method = "bray")
##          Df SumOfSqs      R2      F Pr(>F)
## Model     1   3.7047 0.17797 11.258  0.001 ***
## Residual 52  17.1110 0.82203
## Total    53  20.8156 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## IndVal matrix
indval = multipatt(data, cluster=groups, func = "IndVal.g", control = how(nperm=100))
summary(indval)
```

```
##
##  Multilevel pattern analysis
##  ---------------------------
##
##  Association function: IndVal.g
##  Significance level (alpha): 0.05
##
##  Total number of species: 111
##  Selected number of species: 52
##  Number of species associated to 1 group: 46
##  Number of species associated to 2 groups: 6
##
##  List of species associated to each combination:
##
##  Group 1  #sps.  1
##                    stat p.value
## Astarte.elliptica 0.548  0.0099 **
##
##  Group 2  #sps.  22
##                                stat p.value
```

```
## Arctica.islandica             0.970  0.0099 **
## Corbula.gibba                 0.943  0.0099 **
## Heteromastus.filiformis       0.924  0.0099 **
## Nephtys.hombergii             0.779  0.0099 **
## Abra.alba                     0.760  0.0099 **
## Lagis.koreni                  0.758  0.0099 **
## Gastrosaccus.spinifer         0.607  0.0099 **
## Euchone.papillosa             0.562  0.0099 **
## Terebellides.stroemii         0.555  0.0396 *
## Polydora.ciliata              0.550  0.0495 *
## Nephtys.ciliata               0.530  0.0099 **
## Phyllodoce..Anaitides..mucosa 0.517  0.0198 *
## Pherusa.plumosa               0.513  0.0099 **
## Pholoe.assimilis              0.513  0.0198 *
## Mysella.bidentata             0.510  0.0198 *
## Paraonis.fulgens              0.487  0.0198 *
## Ampharete.baltica             0.486  0.0396 *
## Scalibregma.inflatum          0.459  0.0297 *
## Nemertina                     0.436  0.0198 *
## Metridium.senile              0.397  0.0396 *
## Nephtys.caeca                 0.397  0.0396 *
## Ophiura.albida                0.397  0.0396 *
##
##  Group 3  #sps.  23
##                           stat p.value
## Mytilus.edulis          0.998  0.0099 **
## Macoma.balthica         0.985  0.0099 **
## Pygospio.elegans        0.966  0.0099 **
## Mya.arenaria            0.953  0.0099 **
## Gammarus.salinus        0.856  0.0099 **
## Saduria.entomon         0.809  0.0099 **
## Gammarus.zaddachi       0.775  0.0099 **
## Jaera.albifrons         0.775  0.0099 **
## Hydrobia.ulvae          0.716  0.0099 **
## Gammarus.oceanicus      0.683  0.0099 **
## Alcyonidium.polyoum     0.649  0.0198 *
## Tubifex.costatus        0.632  0.0099 **
## Oligochaeta             0.631  0.0099 **
## Crangon.crangon         0.615  0.0099 **
## Hartlaubella.gelatinosa 0.595  0.0099 **
## Travisia.forbesii       0.577  0.0099 **
## Electra.crustulenta     0.548  0.0495 *
## Bathyporeia.pilosa      0.447  0.0495 *
## Marenzelleria.neglecta  0.447  0.0198 *
## Monoporeia.affinis      0.447  0.0198 *
## Nais.elinguis           0.447  0.0297 *
## Tubificoides.benedeni   0.447  0.0297 *
## Hyperia.galba           0.365  0.0495 *
##
##  Group 1+2  #sps.  3
##                                  stat p.value
## Scoloplos.armiger               0.744  0.0099 **
## Trochochaeta.multisetosa        0.679  0.0198 *
## Polydora..Polydora..quadrilobata 0.641  0.0396 *
```

```
##
##  Group 1+3  #sps.  2
##                   stat p.value
## Neomysis.integer 0.586  0.0297 *
## Mysis.mixta      0.561  0.0396 *
##
##  Group 2+3  #sps.  1
##                          stat p.value
## Halicryptus.spinulosus 0.871  0.0198 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Phi coefficient analysis
data.rel = decostand(data, method='total')
phi = multipatt(data.rel, cluster=groups, func="r.g", control=how(nperm=100))
summary(phi)
```

```
##
##  Multilevel pattern analysis
##  ---------------------------
##
##  Association function: r.g
##  Significance level (alpha): 0.05
##
##  Total number of species: 111
##  Selected number of species: 39
##  Number of species associated to 1 group: 37
##  Number of species associated to 2 groups: 2
##
##  List of species associated to each combination:
##
##  Group 1  #sps.  6
##                                   stat p.value
## Scoloplos.armiger                0.579  0.0099 **
## Pontoporeia.femorata             0.445  0.0099 **
## groups                           0.432  0.0099 **
## Bylgides.sarsi                   0.427  0.0099 **
## Polydora..Polydora..quadrilobata 0.338  0.0495 *
## Capitella.capitata               0.331  0.0099 **
##
##  Group 2  #sps.  12
##                          stat p.value
## Abra.alba              0.572  0.0099 **
## Corbula.gibba          0.511  0.0099 **
## Nephtys.hombergii      0.450  0.0099 **
## Arctica.islandica      0.444  0.0099 **
## Lagis.koreni           0.417  0.0099 **
## Gastrosaccus.spinifer 0.364  0.0099 **
## Pholoe.assimilis       0.361  0.0198 *
## Nemertina              0.335  0.0297 *
## Parvicardium.ovale     0.323  0.0396 *
## Mysella.bidentata      0.319  0.0396 *
## Pherusa.plumosa        0.306  0.0297 *
## Euchone.papillosa      0.293  0.0396 *
```

```
##
##  Group 3  #sps.  19
##                            stat p.value
## Pygospio.elegans          0.920  0.0099 **
## Mytilus.edulis            0.765  0.0099 **
## Macoma.balthica           0.670  0.0099 **
## Mya.arenaria              0.489  0.0099 **
## Saduria.entomon           0.421  0.0099 **
## Tubifex.costatus          0.412  0.0099 **
## Gammarus.oceanicus        0.400  0.0099 **
## Jaera.albifrons           0.398  0.0099 **
## Hartlaubella.gelatinosa 0.384  0.0099 **
## Bathyporeia.pilosa        0.369  0.0396 *
## Gammarus.salinus          0.350  0.0099 **
## Nais.elinguis             0.348  0.0099 **
## Marenzelleria.neglecta    0.334  0.0297 *
## Tubificoides.benedeni     0.334  0.0099 **
## Oligochaeta               0.327  0.0396 *
## Travisia.forbesii         0.327  0.0099 **
## Gammarus.zaddachi         0.320  0.0099 **
## Hyperia.galba             0.287  0.0495 *
## Monoporeia.affinis        0.238  0.0198 *
##
##  Group 1+2  #sps.  2
##                            stat p.value
## Diastylis.rathkei         0.412  0.0198 *
## Heteromastus.filiformis 0.407  0.0099 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Comments about PERMANOVA, IndVal and phi coefficient analyses: Here we see that the PERMANOVA test supports that there are significant differences in zoobenthos composition between the 3 groups that differentiate the sites and some sites in specific years. For IndVal analysis, we saw that there are 34 species that are significantly associated with the groups. For example, species such as Pygospio elegans, Mytilus edulis, and Macoma balthica are strongly associated with GR1. There are 16 species that are significantly linked to GR2, and 3 species show an association with the combination of groups GR2 and GR3. This is interesting as we saw that sites 0 and 2 are sometimes more similar to each other than to themselves in certain years. The phi coefficient analysis showed that the species Pygospio elegans, Mytilus edulis, and Macoma balthica are strongly associated with GR1, which is consistent with the IndVal analysis. Here we see a similar grouping as IndVal.

```r
# data.env = read.csv(paste0(getwd(),"/Week4-Beta/data/env_data.csv"), row.names=1)
data.env = read.csv("data/env_data.csv", header = TRUE)
row.names(data.env) = data.env$site_year
# Delete the site_year column
data.env = data.env[,-1]

# Retain the overlapping row names between the two distance matrices
row_data.env = row.names(data.env)
row_data = row.names(data)

# Calculate the distance matrix for the environmental data
# # Distance matrices of the data data
```

```r
data = data[row_data.env,]

data.dist = vegdist(data, method="bray")
data.env.dist = vegdist(scale(data.env), method="euclid")


# row.names(data.env.dist)
# # Mantel test
mantel(data.dist, data.env.dist, method="pearson", permutations = 1000)
```

```
##
## Mantel statistic based on Pearson's product-moment correlation
##
## Call:
## mantel(xdis = data.dist, ydis = data.env.dist, method = "pearson",      permutations = 1000)
##
## Mantel statistic r: 0.3288
##       Significance: 0.000999
##
## Upper quantiles of permutations (null model):
##    90%    95%  97.5%    99%
## 0.0518 0.0639 0.0813 0.0946
## Permutation: free
## Number of permutations: 1000
```

> **Comments about Mantel test**: In the Mantel test, we compare the dissimilarities in zoobenthos community composition with the differences in environmental conditions (using Euclidean distances on scaled data). First, we see that the statistic r is 0.3288, which is a moderate positive correlation. This means that sites that are more different in their environmental conditions tend to also be more different in their zoobenthos communities. Here, the p-value is 0.001, which is very small. This indicates that the correlation is statistically significant.

```r
# Bray Curtis distance for data
data.db = vegdist(data[rownames(data.env), ], method="bray")

# Creat env matrix
env.vars = as.matrix(data.env)



# dbRDA
png("plots/dbrda_plot_all.png", width = 4000, height = 3000, res = 500)
data.dbrda.modfull = dbrda(data.db ~ ., as.data.frame(env.vars))
```

```
##
## Some constraints or conditions were aliased because they were redundant. This
## can happen if terms are linearly dependent (collinear): 'max_depth'
```

```r
ordiplot(data.dbrda.modfull)
dev.off()
```

```
## pdf
##   2
```

```r
# Check the pairwise correlation
psych::corr.test(env.vars)
```

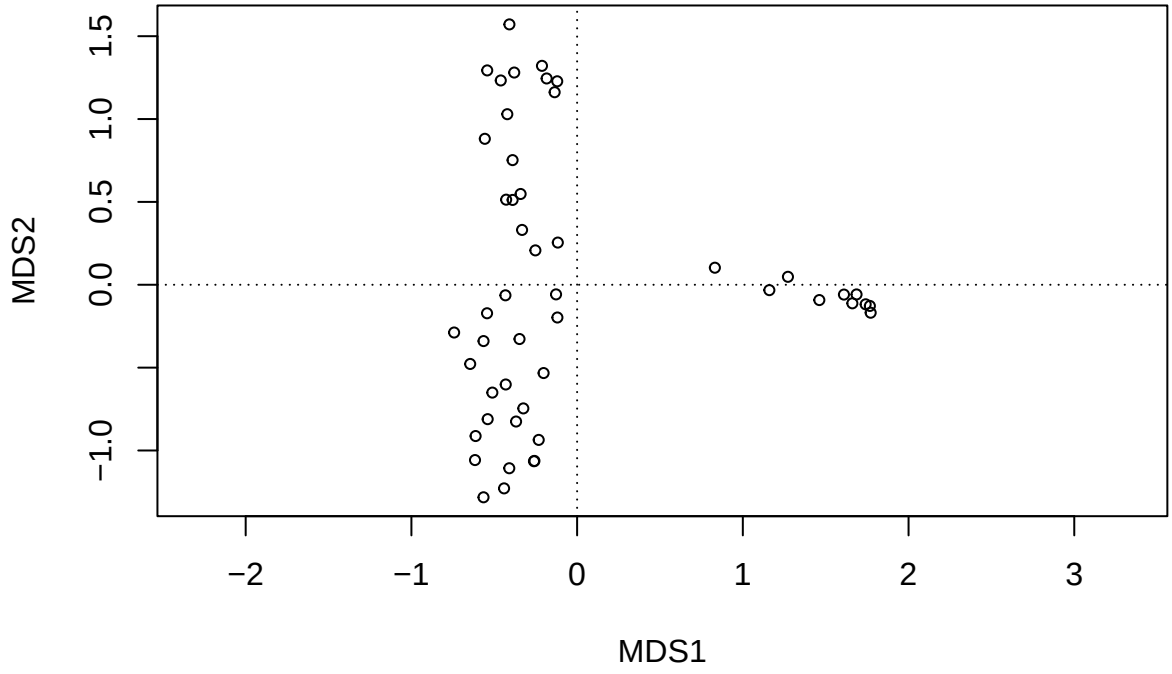
```
## Call:psych::corr.test(x = env.vars)
## Correlation matrix
##                   min_depth max_depth bot_depth secchi_depth depth temperature
## min_depth              1.00      1.00      0.93        -0.01  0.77       -0.42
## max_depth              1.00      1.00      0.93        -0.01  0.77       -0.42
## bot_depth              0.93      0.93      1.00         0.04  0.86       -0.35
## secchi_depth          -0.01     -0.01      0.04         1.00  0.09        0.05
## depth                  0.77      0.77      0.86         0.09  1.00       -0.09
## temperature           -0.42     -0.42     -0.35         0.05 -0.09        1.00
## salinity              -0.56     -0.56     -0.66         0.26 -0.69       -0.01
## oxygen                 0.40      0.40      0.38        -0.15  0.05       -0.71
## phosphate              0.05      0.05      0.05        -0.06 -0.10       -0.12
## total_phosphorus      -0.24     -0.24     -0.25        -0.17 -0.30       -0.04
## silicate              -0.24     -0.24     -0.23        -0.08 -0.26        0.11
## nitrate_nitrite        0.29      0.29      0.31        -0.07  0.30       -0.23
## nitrate                0.10      0.10      0.03         0.16 -0.15       -0.38
## nitrite               -0.15     -0.15     -0.19         0.34 -0.36       -0.16
## ammonium              -0.21     -0.21     -0.25        -0.01 -0.29        0.12
## total_nitrogen         0.15      0.15      0.13        -0.29 -0.03       -0.28
## hydrogen_sulphide      0.09      0.09      0.16        -0.11  0.18        0.06
## ph                     0.12      0.12      0.13        -0.12  0.29        0.23
## total_alkalinity      -0.12     -0.12     -0.12        -0.53 -0.17       -0.05
## chlorophyll_a         -0.20     -0.20     -0.20        -0.23 -0.38       -0.03
##                   salinity oxygen phosphate total_phosphorus silicate
## min_depth            -0.56   0.40      0.05            -0.24    -0.24
## max_depth            -0.56   0.40      0.05            -0.24    -0.24
## bot_depth            -0.66   0.38      0.05            -0.25    -0.23
## secchi_depth          0.26  -0.15     -0.06            -0.17    -0.08
## depth                -0.69   0.05     -0.10            -0.30    -0.26
## temperature          -0.01  -0.71     -0.12            -0.04     0.11
## salinity              1.00  -0.33      0.44             0.43     0.55
## oxygen               -0.33   1.00     -0.21            -0.13    -0.40
## phosphate             0.44  -0.21      1.00             0.43     0.77
## total_phosphorus      0.43  -0.13      0.43             1.00     0.46
## silicate              0.55  -0.40      0.77             0.46     1.00
## nitrate_nitrite      -0.33   0.26     -0.34            -0.04    -0.33
## nitrate               0.58   0.00      0.56             0.47     0.43
## nitrite               0.60  -0.03      0.24             0.23     0.27
## ammonium              0.53  -0.23      0.58             0.42     0.58
## total_nitrogen        0.10   0.13      0.43             0.48     0.31
## hydrogen_sulphide    -0.35   0.07     -0.34            -0.18    -0.20
## ph                   -0.66   0.03     -0.62            -0.59    -0.51
## total_alkalinity     -0.19   0.20     -0.10             0.10     0.10
## chlorophyll_a        -0.02   0.14     -0.12            -0.10    -0.06
##                   nitrate_nitrite nitrate nitrite ammonium total_nitrogen
## min_depth                    0.29    0.10   -0.15    -0.21           0.15
## max_depth                    0.29    0.10   -0.15    -0.21           0.15
## bot_depth                    0.31    0.03   -0.19    -0.25           0.13
## secchi_depth                -0.07    0.16    0.34    -0.01          -0.29
## depth                        0.30   -0.15   -0.36    -0.29          -0.03
```

```
## temperature                -0.23   -0.38   -0.16     0.12          -0.28
## salinity                   -0.33    0.58    0.60     0.53           0.10
## oxygen                      0.26    0.00   -0.03    -0.23           0.13
## phosphate                  -0.34    0.56    0.24     0.58           0.43
## total_phosphorus           -0.04    0.47    0.23     0.42           0.48
## silicate                   -0.33    0.43    0.27     0.58           0.31
## nitrate_nitrite             1.00    0.06    0.04    -0.09           0.06
## nitrate                     0.06    1.00    0.66     0.39           0.31
## nitrite                     0.04    0.66    1.00     0.36           0.01
## ammonium                   -0.09    0.39    0.36     1.00           0.31
## total_nitrogen              0.06    0.31    0.01     0.31           1.00
## hydrogen_sulphide           0.11   -0.29   -0.01    -0.26          -0.23
## ph                          0.26   -0.61   -0.40    -0.53          -0.54
## total_alkalinity            0.12   -0.25   -0.23    -0.16           0.33
## chlorophyll_a              -0.11   -0.17   -0.02    -0.16           0.11
##                    hydrogen_sulphide    ph total_alkalinity chlorophyll_a
## min_depth                       0.09  0.12            -0.12         -0.20
## max_depth                       0.09  0.12            -0.12         -0.20
## bot_depth                       0.16  0.13            -0.12         -0.20
## secchi_depth                   -0.11 -0.12            -0.53         -0.23
## depth                           0.18  0.29            -0.17         -0.38
## temperature                     0.06  0.23            -0.05         -0.03
## salinity                       -0.35 -0.66            -0.19         -0.02
## oxygen                          0.07  0.03             0.20          0.14
## phosphate                      -0.34 -0.62            -0.10         -0.12
## total_phosphorus               -0.18 -0.59             0.10         -0.10
## silicate                       -0.20 -0.51             0.10         -0.06
## nitrate_nitrite                 0.11  0.26             0.12         -0.11
## nitrate                        -0.29 -0.61            -0.25         -0.17
## nitrite                        -0.01 -0.40            -0.23         -0.02
## ammonium                       -0.26 -0.53            -0.16         -0.16
## total_nitrogen                 -0.23 -0.54             0.33          0.11
## hydrogen_sulphide               1.00  0.27             0.20          0.10
## ph                              0.27  1.00             0.15          0.10
## total_alkalinity                0.20  0.15             1.00          0.15
## chlorophyll_a                   0.10  0.10             0.15          1.00
## Sample Size
## [1] 49
## Probability values (Entries above the diagonal are adjusted for multiple tests.)
##                    min_depth max_depth bot_depth secchi_depth depth temperature
## min_depth               0.00      0.00      0.00         1.00  0.00        0.39
## max_depth               0.00      0.00      0.00         1.00  0.00        0.39
## bot_depth               0.00      0.00      0.00         1.00  0.00        1.00
## secchi_depth            0.96      0.96      0.78         0.00  1.00        1.00
## depth                   0.00      0.00      0.00         0.52  0.00        1.00
## temperature             0.00      0.00      0.01         0.75  0.55        0.00
## salinity                0.00      0.00      0.00         0.07  0.00        0.97
## oxygen                  0.00      0.00      0.01         0.30  0.71        0.00
## phosphate               0.71      0.71      0.74         0.69  0.49        0.42
## total_phosphorus        0.10      0.10      0.09         0.23  0.04        0.79
## silicate                0.10      0.10      0.11         0.60  0.07        0.47
## nitrate_nitrite         0.04      0.04      0.03         0.63  0.04        0.11
## nitrate                 0.48      0.48      0.86         0.29  0.31        0.01
## nitrite                 0.29      0.29      0.19         0.02  0.01        0.26
```

```
## ammonium              0.16      0.16      0.08         0.93  0.04        0.43
## total_nitrogen        0.29      0.29      0.38         0.04  0.83        0.05
## hydrogen_sulphide     0.56      0.56      0.26         0.46  0.22        0.71
## ph                    0.42      0.42      0.38         0.43  0.04        0.10
## total_alkalinity      0.43      0.43      0.41         0.00  0.25        0.75
## chlorophyll_a         0.18      0.18      0.16         0.11  0.01        0.84
##                   salinity oxygen phosphate total_phosphorus silicate
## min_depth             0.00   0.73      1.00             1.00     1.00
## max_depth             0.00   0.73      1.00             1.00     1.00
## bot_depth             0.00   1.00      1.00             1.00     1.00
## secchi_depth          1.00   1.00      1.00             1.00     1.00
## depth                 0.00   1.00      1.00             1.00     1.00
## temperature           1.00   0.00      1.00             1.00     1.00
## salinity              0.00   1.00      0.27             0.29     0.01
## oxygen                0.02   0.00      1.00             1.00     0.73
## phosphate             0.00   0.16      0.00             0.35     0.00
## total_phosphorus      0.00   0.36      0.00             0.00     0.15
## silicate              0.00   0.00      0.00             0.00     0.00
## nitrate_nitrite       0.02   0.07      0.02             0.80     0.02
## nitrate               0.00   1.00      0.00             0.00     0.00
## nitrite               0.00   0.82      0.10             0.11     0.06
## ammonium              0.00   0.11      0.00             0.00     0.00
## total_nitrogen        0.49   0.37      0.00             0.00     0.03
## hydrogen_sulphide     0.01   0.62      0.02             0.23     0.18
## ph                    0.00   0.86      0.00             0.00     0.00
## total_alkalinity      0.19   0.16      0.47             0.51     0.51
## chlorophyll_a         0.88   0.32      0.40             0.50     0.70
##                   nitrate_nitrite nitrate nitrite ammonium total_nitrogen
## min_depth                    1.00    1.00    1.00     1.00           1.00
## max_depth                    1.00    1.00    1.00     1.00           1.00
## bot_depth                    1.00    1.00    1.00     1.00           1.00
## secchi_depth                 1.00    1.00    1.00     1.00           1.00
## depth                        1.00    1.00    1.00     1.00           1.00
## temperature                  1.00    0.97    1.00     1.00           1.00
## salinity                     1.00    0.00    0.00     0.01           1.00
## oxygen                       1.00    1.00    1.00     1.00           1.00
## phosphate                    1.00    0.00    1.00     0.00           0.30
## total_phosphorus             1.00    0.11    1.00     0.38           0.09
## silicate                     1.00    0.33    1.00     0.00           1.00
## nitrate_nitrite              0.00    1.00    1.00     1.00           1.00
## nitrate                      0.68    0.00    0.00     0.78           1.00
## nitrite                      0.79    0.00    0.00     1.00           1.00
## ammonium                     0.53    0.01    0.01     0.00           1.00
## total_nitrogen               0.67    0.03    0.94     0.03           0.00
## hydrogen_sulphide            0.46    0.05    0.94     0.07           0.11
## ph                           0.07    0.00    0.00     0.00           0.00
## total_alkalinity             0.40    0.08    0.12     0.27           0.02
## chlorophyll_a                0.44    0.25    0.87     0.26           0.44
##                   hydrogen_sulphide   ph total_alkalinity chlorophyll_a
## min_depth                      1.00 1.00             1.00             1
## max_depth                      1.00 1.00             1.00             1
## bot_depth                      1.00 1.00             1.00             1
## secchi_depth                   1.00 1.00             0.02             1
## depth                          1.00 1.00             1.00             1
```

```
## temperature                1.00 1.00            1.00            1
## salinity                   1.00 0.00            1.00            1
## oxygen                     1.00 1.00            1.00            1
## phosphate                  1.00 0.00            1.00            1
## total_phosphorus           1.00 0.00            1.00            1
## silicate                   1.00 0.03            1.00            1
## nitrate_nitrite            1.00 1.00            1.00            1
## nitrate                    1.00 0.00            1.00            1
## nitrite                    1.00 0.61            1.00            1
## ammonium                   1.00 0.01            1.00            1
## total_nitrogen             1.00 0.01            1.00            1
## hydrogen_sulphide          0.00 1.00            1.00            1
## ph                         0.06 0.00            1.00            1
## total_alkalinity           0.16 0.29            0.00            1
## chlorophyll_a              0.49 0.50            0.32            0
##
##  To see confidence intervals of the correlations, print with the short=FALSE option
```

```r
# ============================================
# Try model with constant explained variables
data.dbrda.mod0 = dbrda(data.db  ~ 1, as.data.frame(env.vars))
ordiplot(data.dbrda.mod0) # no vector as we don't use any variable in env.chem
```




```r
# Next, we will test all combination from 0 explanatory variable to full
# This functions returns the one that has lowest AIC
data.dbrda = ordiR2step(data.dbrda.mod0, data.dbrda.modfull, perm.max=200)
```

```
##
## Some constraints or conditions were aliased because they were redundant. This
## can happen if terms are linearly dependent (collinear): 'max_depth'

## Step: R2.adj= 0
## Call: data.db ~ 1
##
##                       R2.adjusted
## <All variables>      0.3096564523
## + salinity           0.1175763150
## + ph                 0.1096655064
## + nitrate            0.1060525682
## + bot_depth          0.0895187193
## + min_depth          0.0729115257
## + max_depth          0.0729115257
## + phosphate          0.0696859696
## + depth              0.0673383875
## + nitrite            0.0615105459
## + temperature        0.0473378549
## + silicate           0.0459518812
## + ammonium           0.0360156526
## + total_phosphorus   0.0276789974
## + total_alkalinity   0.0207878991
## + secchi_depth       0.0197185479
## + total_nitrogen     0.0184746978
## + hydrogen_sulphide  0.0120994947
## + oxygen             0.0101284237
## + nitrate_nitrite    0.0004714983
## <none>               0.0000000000
## + chlorophyll_a     -0.0051231287
##
##            Df    AIC      F Pr(>F)
## + salinity  1 139.61 7.3956  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.1175763
## Call: data.db ~ salinity
##
##                     R2.adjusted
## <All variables>       0.3096565
## + bot_depth           0.2446849
## + min_depth           0.2146239
## + max_depth           0.2146239
## + nitrate             0.1786475
## + phosphate           0.1723506
## + ph                  0.1722468
## + temperature         0.1681841
## + depth               0.1579125
## + silicate            0.1466048
## + oxygen              0.1373721
## + total_nitrogen      0.1354879
## + total_alkalinity    0.1350943
## + secchi_depth        0.1333245
```

```
## + ammonium            0.1281631
## + nitrite             0.1256039
## + total_phosphorus    0.1244388
## <none>                0.1175763
## + hydrogen_sulphide   0.1158080
## + chlorophyll_a       0.1150144
## + nitrate_nitrite     0.1111906
##
##             Df    AIC      F Pr(>F)
## + bot_depth  1 132.94 8.9094  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.2446849
## Call: data.db ~ salinity + bot_depth
##
##                     R2.adjusted
## <All variables>       0.3096565
## + silicate            0.2750752
## + phosphate           0.2707120
## + min_depth           0.2633371
## + max_depth           0.2633371
## + total_alkalinity    0.2586591
## + secchi_depth        0.2568789
## + total_nitrogen      0.2556170
## + ammonium            0.2551904
## + temperature         0.2541995
## + total_phosphorus    0.2535345
## + ph                  0.2502865
## + oxygen              0.2488866
## + nitrate             0.2461184
## + depth               0.2460308
## <none>                0.2446849
## + hydrogen_sulphide   0.2442424
## + chlorophyll_a       0.2395448
## + nitrate_nitrite     0.2390678
## + nitrite             0.2357921
##
##            Df    AIC      F Pr(>F)
## + silicate  1 131.85 2.9284  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.2750752
## Call: data.db ~ salinity + bot_depth + silicate
##
##                     R2.adjusted
## <All variables>       0.3096565
## + temperature         0.2892493
## + min_depth           0.2884492
## + max_depth           0.2884492
## + secchi_depth        0.2844775
## + total_alkalinity    0.2818712
## + oxygen              0.2793177
```

```
## + ammonium            0.2788269
## + phosphate           0.2774928
## + ph                  0.2771726
## + depth               0.2766497
## + nitrate             0.2757624
## + total_nitrogen      0.2754767
## + hydrogen_sulphide   0.2753668
## <none>                0.2750752
## + total_phosphorus    0.2744675
## + nitrate_nitrite     0.2727531
## + chlorophyll_a       0.2705453
## + nitrite             0.2671563
##
##               Df    AIC      F Pr(>F)
## + temperature  1 131.78 1.8974  0.018 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.2892493
## Call: data.db ~ salinity + bot_depth + silicate + temperature
##
##                     R2.adjusted
## <All variables>       0.3096565
## + min_depth           0.3075098
## + max_depth           0.3075098
## + ammonium            0.2943384
## + total_alkalinity    0.2922455
## + secchi_depth        0.2913564
## + phosphate           0.2908738
## + hydrogen_sulphide   0.2900429
## + ph                  0.2899026
## <none>                0.2892493
## + nitrate             0.2883663
## + total_nitrogen      0.2881577
## + total_phosphorus    0.2881183
## + nitrate_nitrite     0.2868498
## + depth               0.2843742
## + oxygen              0.2832105
## + chlorophyll_a       0.2830584
## + nitrite             0.2813626
##
##             Df    AIC      F Pr(>F)
## + min_depth  1 131.38 2.1603  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.3075098
## Call: data.db ~ salinity + bot_depth + silicate + temperature + min_depth
##


##
## Some constraints or conditions were aliased because they were redundant. This
## can happen if terms are linearly dependent (collinear): 'max_depth'
```

```
##                       R2.adjusted
## + secchi_depth          0.3135735
## + ammonium              0.3134075
## + total_alkalinity      0.3111176
## + phosphate             0.3098350
## <All variables>         0.3096565
## + nitrate               0.3095463
## + hydrogen_sulphide     0.3094044
## <none>                  0.3075098
## + max_depth             0.3075098
## + total_nitrogen        0.3073054
## + nitrate_nitrite       0.3053710
## + total_phosphorus      0.3051166
## + ph                    0.3048054
## + depth                 0.3029465
## + oxygen                0.3019001
## + chlorophyll_a         0.3015660
## + nitrite               0.2999863
```

```
# Summary of selected model
data.dbrda$call   # data.db ~ ph + bot_depth + nitrate + min_depth
```

```
## dbrda(formula = data.db ~ salinity + bot_depth + silicate + temperature +
##     min_depth, data = as.data.frame(env.vars))
```

```
data.dbrda$anova
```

```
##                   R2.adj Df    AIC      F Pr(>F)
## + salinity       0.11758  1 139.61 7.3956  0.002 **
## + bot_depth      0.24469  1 132.94 8.9094  0.002 **
## + silicate       0.27508  1 131.85 2.9284  0.002 **
## + temperature    0.28925  1 131.78 1.8974  0.018 *
## + min_depth      0.30751  1 131.38 2.1603  0.002 **
## <All variables> 0.30966
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
png("plots/dbrda_plot_best.png", width = 6000, height = 6000, res = 500)
ordiplot(data.dbrda)
text(data.dbrda, display = "sites", cex = 0.4, pos = 3)
dev.off()
```

```
## pdf
##   2
```

Driver environmental variables of zoobenthos community composition

Comments about the constrained ordination: To keep this short, we see that 5 environmental variables temperature, silicate, salinity, min_depth, bot_depth are significant predictors of zoobenthos community composition. Together, they explain about 30% of the variation. The stepwise model selection confirms that each of these variables contributes significantly, and the

ordination plot provides a visual representation of these environmental gradients and their effect on community structure. We also note that these variables are not highly correlated with each other, which is good for the model.

```
# Create a matrix model for the envr data
env.mod = model.matrix( ~ salinity + temperature + silicate + bot_depth + min_depth + nitrate, as.data.f



#cCreate spatial model

# first, weight each site by its relative abundance
rs = rowSums(data) / sum(data)
rs = as.data.frame(rs)
rs
```

```
##                   rs
## S0_1983 0.0018659140
## S0_1984 0.0030350626
## S0_1986 0.0022674398
## S0_1988 0.0057158378
## S0_1989 0.0080423255
## S0_1990 0.0044049742
## S0_1991 0.0038263047
## S0_1992 0.0058811720
## S0_1993 0.0040388771
## S0_1994 0.0200999091
## S0_1995 0.0028342998
## S0_1997 0.0267959422
## S0_1998 0.0267841326
## S0_1999 0.0075345135
## S0_2000 0.0617404962
## S0_2001 0.0265125123
## S0_2002 0.0009801953
## S0_2003 0.0075227039
## S0_2004 0.0214580110
## S1_1991 0.0061055540
## S1_1992 0.0119631069
## S1_1993 0.0912172136
## S1_1994 0.1242958537
## S1_1995 0.1458247222
## S1_1998 0.0855013758
## S1_1999 0.1191114470
## S1_2001 0.0224854447
## S1_2002 0.0171947518
## S1_2003 0.0532021682
## S3_1980 0.0184701867
## S3_1981 0.0028579189
## S3_1982 0.0014880074
## S3_1983 0.0006377174
## S3_1984 0.0006140983
## S3_1986 0.0009211474
## S3_1987 0.0006495270
## S3_1989 0.0002243821
```

```
## S3_1990 0.0020666769
## S3_1991 0.0126953010
## S3_1993 0.0069558440
## S3_1994 0.0005314312
## S3_1995 0.0004487641
## S3_1997 0.0044522125
## S3_1998 0.0066487948
## S3_2000 0.0008148612
## S3_2001 0.0031649681
## S3_2002 0.0070621302
## S3_2003 0.0070739398
## S3_2004 0.0039798292
```

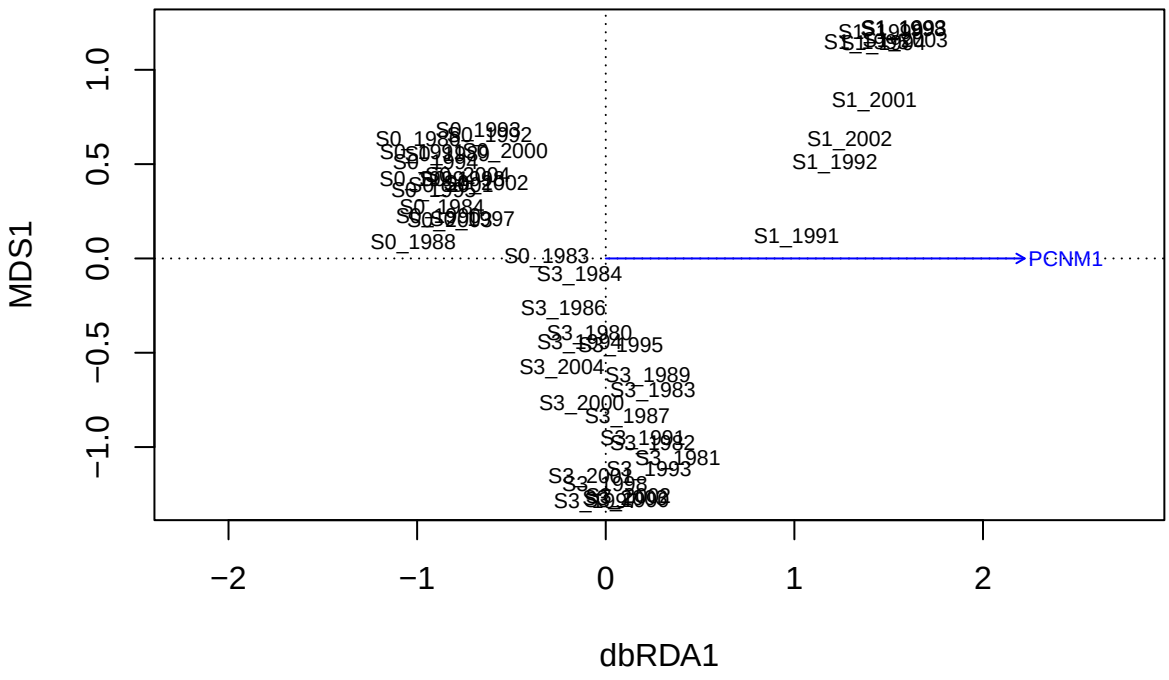
```
# Load spatial data
# data.coords = read.csv(paste0(getwd(),"/Week4-Beta/data/site_coords.csv"), header = TRUE, row.names=1
data.coords = read.csv("data/site_coords.csv", header = TRUE, row.names=1)
data.coords = data.coords[rownames(rs),]
rs = rs[rownames(data.coords),]


data.pcnmw = pcnm(dist(data.coords), w=rs, dist.re=T) # Perform PCNM on the coordinates
data.pcnmw$values > 0 # Extract only eigenvectors associated with positive eigenvalues
```

```
## [1]  TRUE FALSE
```

```
# Perform model selection of spatial data
data.space = as.data.frame(scores(data.pcnmw))
data.pcnmw.mod0 = dbrda(data.db ~ 1, data.space) # no var
data.pcnmw.mod1 = dbrda(data.db ~ ., data.space) # all var
# Stepwise model selection
step.pcnm = ordiR2step(data.pcnmw.mod0, data.pcnmw.mod1, perm.max=200)
```

```
## Step: R2.adj= 0
## Call: data.db ~ 1
##
##                 R2.adjusted
## <All variables>   0.1392301
## + PCNM1           0.1392301
## <none>            0.0000000
##
##         Df    AIC     F Pr(>F)
## + PCNM1  1 138.39 8.764  0.002 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: R2.adj= 0.1392301
## Call: data.db ~ PCNM1
##
```

```
# png("Outputs/step_pcnm.png", width = 4000, height = 3000, res = 500)
plot(step.pcnm)
```




```
# dev.off()
step.pcnm$call
```

```
## dbrda(formula = data.db ~ PCNM1, data = data.space)
```

```
# Check the portion of explained variation of the fish composition using the spatial model
step.pcnm$anova
```

```
##                  R2.adj Df    AIC     F Pr(>F)
## + PCNM1         0.13923  1 138.39 8.764  0.002 **
## <All variables> 0.13923
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Final spatial model
space.mod = model.matrix(~ PCNM1, data.space)[,-1]

# ============================================
# Compare env and spatial model
data.total.env = dbrda(data.db ~ env.mod)
data.total.space = dbrda(data.db ~ space.mod)

#  PArtial constrained ordination
# ?Condition --> control the second explanatory matrix
```

```r
data.env.cond.space = dbrda(data.db ~ env.mod + Condition(space.mod))
data.space.cond.env = dbrda(data.db ~ space.mod + Condition(env.mod))

# Test for significance of the dbRDA fractions
permutest(data.env.cond.space,permutations=999)
```

```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = data.db ~ env.mod + Condition(space.mod))
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     6  4.8724 3.0323  0.001 ***
## Residual 41 10.9801
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = data.db ~ space.mod + Condition(env.mod))
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     1   0.383 1.4302   0.11
## Residual 41  10.980
```

```r
permutest(data.total.env,permutations=999)
```

```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = data.db ~ env.mod)
## Permutation test for all constrained eigenvalues
##          Df Inertia      F Pr(>F)
## Model     6  7.4454 4.5866  0.001 ***
## Residual 42 11.3631
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
permutest(data.total.space,permutations=999)
```

```
##
## Permutation test for dbrda under reduced model
##
## Permutation: free
## Number of permutations: 999
##
## Model: dbrda(formula = data.db ~ space.mod)
## Permutation test for all constrained eigenvalues
##          Df Inertia     F Pr(>F)
## Model     1   2.956 8.764  0.001 ***
## Residual 47  15.852
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Calculate the fraction of variation explained by space alone, by env alone and by both, and by neithe
data.varpart = varpart(data.db, env.mod, space.mod)
data.varpart
```

```
##
## Partition of squared Bray distance in dbRDA
##
## Call: varpart(Y = data.db, X = env.mod, space.mod)
##
## Explanatory tables:
## X1:  env.mod
## X2:  space.mod
##
## No. of explanatory tables: 2
## Total variation (SS): 18.809
## No. of observations: 49
##
## Partition table:
##                       Df R.squared Adj.R.squared Testable
## [a+c] = X1             6   0.39585       0.30955     TRUE
## [b+c] = X2             1   0.15716       0.13923     TRUE
## [a+b+c] = X1+X2        7   0.41622       0.31655     TRUE
## Individual fractions
## [a] = X1|X2            6                 0.17732     TRUE
## [b] = X2|X1            1                 0.00700     TRUE
## [c]                    0                 0.13223    FALSE
## [d] = Residuals                          0.68345    FALSE
## ---
## Use function 'dbrda' to test significance of fractions of interest
```

```r
png("plots/zoobenthos_varpart.png", width = 4000, height = 3000, res = 500)
par(mar = c(2,2,2,2))
plot(data.varpart)
text(1, 0.25, "Space")
text(0, 0.25, "Env")
mtext("Variation Partitioning of\nZoobenthos Diversity", side = 3, line =-3)
dev.off()
```

```
## pdf
##   2
```

Spatial and environmental drivers of zoobenthos community composition

Comments about the variation partitioning including spatial data:

For the spatial model PCNM We identified that there is only a single eigenvector, PCNM1, best captured the spatial structure. This spatial predictor (PCNM1) explains about 1% of the variation in zoobenthos community composition (adjusted R^2= 0.202).

For the model using spatial variables controlling for environment variables, We see that the model is significant (adjusted R-squared 20%). This means that that the environmental variables (pH, bottom depth, nitrate and minimal depth) significantly explain community variation even after accounting for spatial structure.

In contrast, for the model using environemental variables controlling for spatial factors, we see that the model has less explainatory power (R^2 adjusted = 20%). This suggests that the spatial variation is not a significant driver of the zoobenthos biodiversities between sites.

If we take each model excluding the other set of explainatory variables: For the model using only environemental variables, interestingly, the adjusted R squared value decreased by half (21%) For the model using only spatial variable, expectedly, the adjusted R squared value is very low (1%). We also see that when using both environment and space together, the model explains 40.7% of the variation (R&2 adjusted). This shows that we need to take in account additional factors that influence the variation of biodiversities between sites of the zoobenthos.

TEMPORAL Analysis

```
# Load packages
package.list = c("vegan", "tidyr", "dplyr", "ggplot2", "codyn", "cowplot", "MullerPlot",
  "RColorBrewer", "reshape2", "lubridate", "TTR", "xtable", "multcomp", "pander", "png", "grid",
  "tseries", "nlme", "forecast", "emmeans")
for (package in package.list) {
  if (!require(package, character.only = TRUE, quietly = TRUE)) {
    install.packages(package, dependencies = TRUE)
    library(package, character.only = TRUE)
  }
}
```

```
##
## Attachement du package : 'dplyr'

## Les objets suivants sont masqués depuis 'package:stats':
##
##     filter, lag

## Les objets suivants sont masqués depuis 'package:base':
##
##     intersect, setdiff, setequal, union

## Warning: le package 'ggplot2' a été compilé avec la version R 4.4.2
```

```
## Warning: le package 'codyn' a été compilé avec la version R 4.4.2

## Warning: le package 'cowplot' a été compilé avec la version R 4.4.2

## Warning: le package 'MullerPlot' a été compilé avec la version R 4.4.2

## Warning: le package 'reshape2' a été compilé avec la version R 4.4.2

##
## Attachement du package : 'reshape2'

## L'objet suivant est masqué depuis 'package:tidyr':
##
##     smiths

## Warning: le package 'lubridate' a été compilé avec la version R 4.4.2

##
## Attachement du package : 'lubridate'

## L'objet suivant est masqué depuis 'package:cowplot':
##
##     stamp

## Les objets suivants sont masqués depuis 'package:base':
##
##     date, intersect, setdiff, union

## Warning: le package 'TTR' a été compilé avec la version R 4.4.2

## Warning: le package 'xtable' a été compilé avec la version R 4.4.2

## Warning: le package 'multcomp' a été compilé avec la version R 4.4.2

## Warning: le package 'mvtnorm' a été compilé avec la version R 4.4.2

## Warning: le package 'TH.data' a été compilé avec la version R 4.4.2

##
## Attachement du package : 'MASS'

## L'objet suivant est masqué depuis 'package:dplyr':
##
##     select

##
## Attachement du package : 'TH.data'
```

```
## L'objet suivant est masqué depuis 'package:MASS':
##
##     geyser

## Warning: le package 'pander' a été compilé avec la version R 4.4.2

## Warning: le package 'tseries' a été compilé avec la version R 4.4.2

## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo

##
## Attachement du package : 'nlme'

## L'objet suivant est masqué depuis 'package:dplyr':
##
##     collapse

## Warning: le package 'forecast' a été compilé avec la version R 4.4.2

##
## Attachement du package : 'forecast'

## L'objet suivant est masqué depuis 'package:nlme':
##
##     getResponse

## Warning: le package 'emmeans' a été compilé avec la version R 4.4.2

## Welcome to emmeans.
## Caution: You lose important information if you filter this package's results.
## See '? untidy'
```

Next, we will try to make a time by species matrix

```
# Load temporal data
library(dplyr)
library(tidyr)

## ================================================
# S0
## ================================================
S0 = read.csv("data/S0_tbs.csv", row.names = 1)
S0.ts = ts(S0, start = 1984, end = 2004, frequency = 1)

# Calculate variance of each species over time
var_S0 = apply(S0, 2, var)
var_S0 = as.data.frame(var_S0)

# Sort the species by variance
library(tidyverse)
```

```
## Warning: le package 'tidyverse' a été compilé avec la version R 4.4.2

## Warning: le package 'forcats' a été compilé avec la version R 4.4.2

## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v forcats 1.0.0     v stringr 1.5.1
## v purrr   1.0.2     v tibble  3.2.1
## v readr   2.1.5
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x nlme::collapse()   masks dplyr::collapse()
## x dplyr::filter()    masks stats::filter()
## x dplyr::lag()       masks stats::lag()
## x MASS::select()     masks dplyr::select()
## x lubridate::stamp() masks cowplot::stamp()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
most_varspS0 = var_S0 %>%
  arrange(desc(var_S0)) %>%  # arrange in descending order
  slice(1:6)   %>% # return rows 1 through 10
  rownames() # Extract the species names

least_varspS0 = var_S0 %>%
  arrange(var_S0) %>%  # arrange in descending order
  slice(1:6)   %>% # return rows 1 through 10
  rownames() # Extract the species names


# Plot the abundance of the most variable species
png("plots/most_var_S0.png", width = 1000, height = 1000)
plot(S0.ts[, most_varspS0], main = "Most variable species in site 0", xlab = "Year", ylab = "Abundance")
dev.off()
```

```
## pdf
##   2
```

```
# Plot the abundance of the least variable species
png("plots/least_var_S0.png", width = 1000, height = 1000)
plot(S0.ts[, least_varspS0], main = "Least variable species in site 0", xlab = "Year", ylab = "Abundanc
dev.off()
```

```
## pdf
##   2
```

```
## ================================================
# S2
## ================================================
S2 = read.csv("data/S2_tbs.csv", row.names = 1)
S2.ts = ts(S2, start = 1984, end = 2004, frequency = 1)

# Calculate variance of each species over time
var_S2 = apply(S2, 2, var)
var_S2 = as.data.frame(var_S2)
```

```r
# Sort the species by variance
most_varspS2 = var_S2 %>%
  arrange(desc(var_S2)) %>%  # arrange in descending order
  slice(1:6)   %>% # return rows 1 through 10
  rownames() # Extract the species names

least_varspS2 = var_S2 %>%
  arrange(var_S2) %>%  # arrange in descending order
  slice(1:6)   %>% # return rows 1 through 10
  rownames() # Extract the species names


# Plot the abundance of the most variable species
png("plots/most_var_S2.png", width = 1000, height = 1000)
plot(S2.ts[, most_varspS2], main = "Most variable species in site 0", xlab = "Year", ylab = "Abundance")
dev.off()
```

```
## pdf
##   2
```

```r
# Plot the abundance of the least variable species
png("plots/least_var_S2.png", width = 1000, height = 1000)
plot(S2.ts[, least_varspS2], main = "Least variable species in site 0", xlab = "Year", ylab = "Abundanc
dev.off()
```

```
## pdf
##   2
```

```r
## ================================================
# S3
## ================================================
S3 = read.csv("data/S3_tbs.csv", row.names = 1)
S3.ts = ts(S3, start = 1984, end = 2004, frequency = 1)

# Calculate variance of each species over time
var_S3 = apply(S3, 2, var)
var_S3 = as.data.frame(var_S3)

# Sort the species by variance
most_varspS3 = var_S3 %>%
  arrange(desc(var_S3)) %>%  # arrange in descending order
  slice(1:6)   %>% # return rows 1 through 10
  rownames() # Extract the species names

least_varspS3 = var_S3 %>%
  arrange(var_S3) %>%  # arrange in descending order
  slice(1:6)   %>% # return rows 1 through 10
  rownames() # Extract the species names

# Plot the abundance of the most variable species
png("plots/most_var_S3.png", width = 1000, height = 1000)
plot(S3.ts[, most_varspS3], main = "Most variable species in site 0", xlab = "Year", ylab = "Abundance")
dev.off()
```

```
## pdf
##   2
```

```
# Plot the abundance of the least variable species
png("plots/least_var_S3.png", width = 1000, height = 1000)
plot(S3.ts[, least_varspS3], main = "Least variable species in site 0", xlab = "Year", ylab = "Abundanc
dev.off()
```

```
## pdf
##   2
```

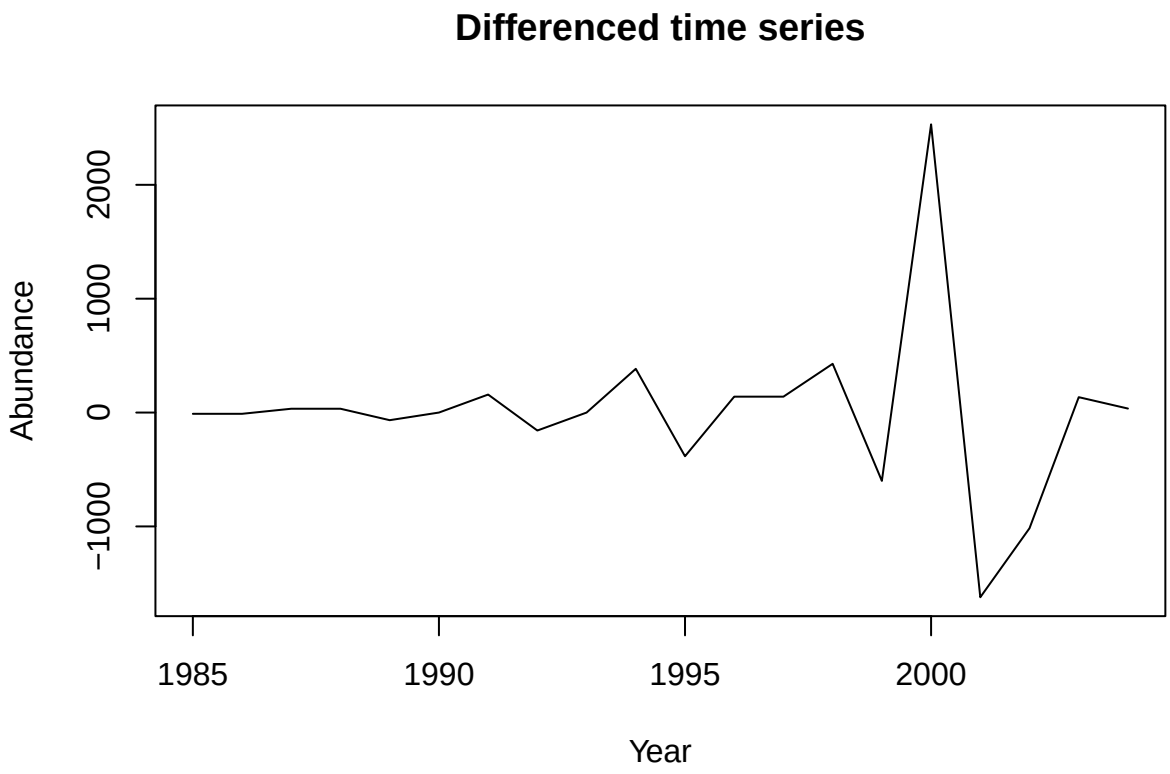
```
## ================================================
# Decompose the time series
## ================================================
# S0
# S0.decomp <- stl(S0.ts[, most_varsp[1]], s.window = "periodic")
library(tseries)
adf.raw = adf.test(S0.ts[, most_varspS0[1]])
adf.raw
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  S0.ts[, most_varspS0[1]]
## Dickey-Fuller = -2.4124, Lag order = 2, p-value = 0.4153
## alternative hypothesis: stationary
```

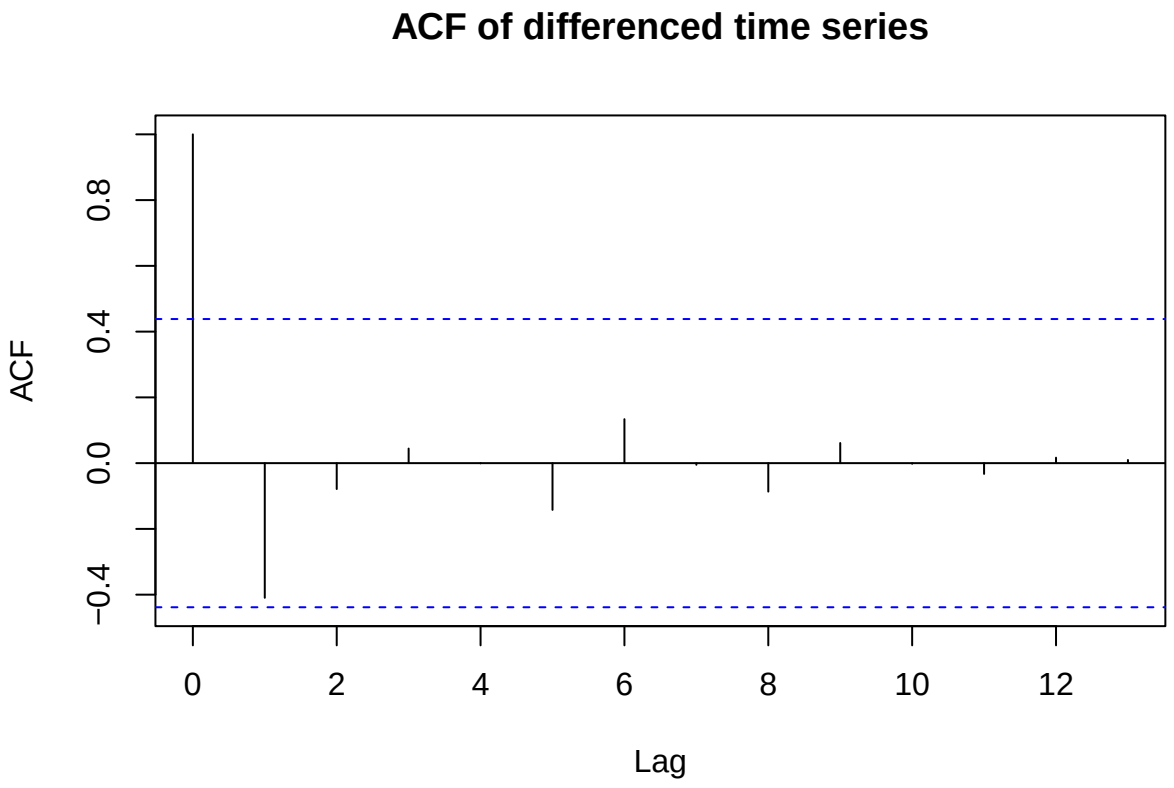
```
S0.ts.diff = diff(S0.ts[, most_varspS0[1]])
adf.diff = adf.test(S0.ts.diff)
adf.diff
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  S0.ts.diff
## Dickey-Fuller = -3.0594, Lag order = 2, p-value = 0.1688
## alternative hypothesis: stationary
```

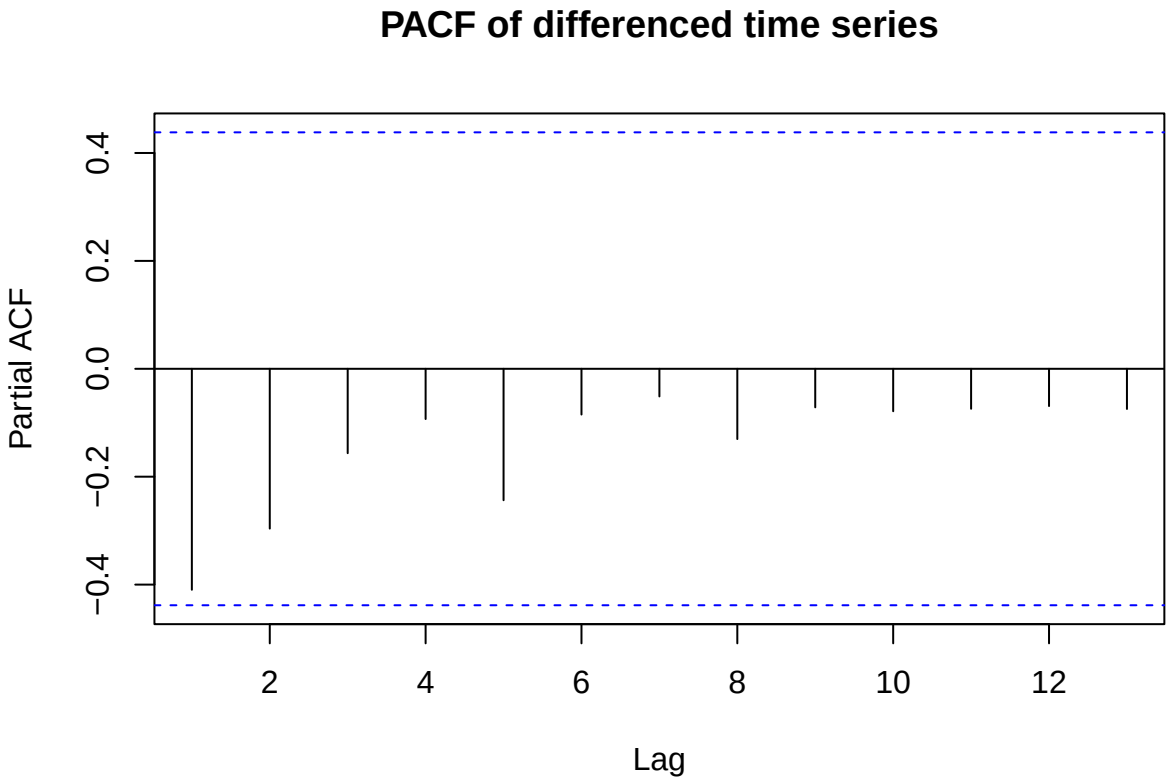
```
plot.ts(S0.ts.diff, main = "Differenced time series",
        xlab = "Year", ylab = "Abundance")
```




```
acf(S0.ts.diff, main = "ACF of differenced time series")
```




```
pacf(S0.ts.diff, main = "PACF of differenced time series")
```



The test suggests that the time series in S0.ts[, most_varsp[1]] is non-stationary because the p-value (0.6225) is greater than common thresholds for significance (typically 0.05). This means the series likely contains a unit root, and you may need to difference or transform the data to achieve stationarity before using it in further analysis or modeling.