Quantitative Community Ecology in R

ESA 2025

Matt Aiello-Lammens, Sarah Supp, Rob Smith, many others

2025-08-10

Good afternoon!

Overview

Overview

1. Introductions

2. Basic syntax and structures

3. Usage and examples

Introductions

Introductions

1. Name

2. Institution/Organization

3. Fave organism or ecosystem

4. What are you working on right now (or want to work on) that you'd like to apply methods from this workshop to? Could be a project a study system, or just an idea.

Software installation

1. Download and install R from: https://cran.r-project.org/

2. Optional: download and install Rstudio from: https://posit.co/download/rstudio-desktop/

3. Install packages (only once!)

```r
pkg <- c('vegan', 'betapart', 'labdsv', 'ade4', 'ecodist', 'ape', 'picante')
install.packages(pkg)
```

4. Load packages

```r
sapply(pkg, require, character.only = TRUE)
```

5. Load some utility functions

```r
source("https://github.com/mlammens/esa2025/raw/main/R/utils_color.R")
```

Get the data

```r
u <- 'https://github.com/mlammens/esa2025/raw/main/data/veg.rda'
u <- url(u)
load(u)
```

Everything in R is one of:

1. object
2. function
3. environment

Kinds of objects in R

1. NULL
2. scalar
3. vector
4. matrix or array
5. data.frame
6. list

Basic R syntax

```
### assignment
x <- 10                                # objects: a scalar
m <- matrix(c(1:8,NA), nrow=3, ncol=3) # objects: a 3x3 matrix
df <- data.frame(                      # objects: a data frame
                 species = c("cewa", "amro", "blja"),
                 count = c(48, NA, 3),
                 month = c(2, 2, 2)
  )

f <- function(a, b) { a + b }          # functions
e <- new.env()                         # environments

### object structure
str(x)                                 # describe structure
head(x)                                # peek at top few rows

### indexing values in an object
m[1,1]                                 # matrix row 1, column 1
lst$the_name                           # list item by name (data.frame column)
```

Types and classes

```r
### examples
x1 <- 1
x2 <- 1.0
x3 <- '1'
x4 <- as.factor(x3)
x5 <- TRUE
x6 <- x5 * 99

### examine classes
list(x1,x2,x3,x4,x5,x6)
lapply(list(x1,x2,x3,x4,x5,x6), class)
```

Handling missing elements

```
### create matrix object
m <- matrix(c(1:8,NA), nrow=3, ncol=3) # a 3x3 matrix

### identify missings
anyNA(m)
is.na(m[1,1])
is.na(m[2,2])
is.na(m[3,3])
is.na(m)

na_matrix <- is.na(df)
na_count <- sum(is.na(df))

### replacing missings
m[is.na(m)] <- 777        # assign one value to any missing values
```

Getting data into R

Ideally, everything is "text"

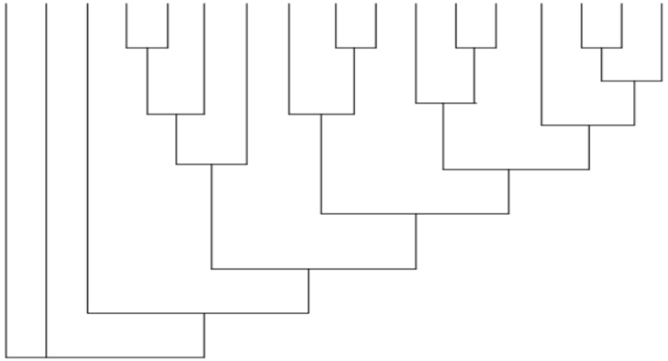
```
### read csv
x <- read.csv('/path/to/your/file.csv')

### load binary RDA from file
load('/path/to/your/file.rda')

### load binary data from package
data('thedata')
```

The Matrix

		SPECIES						ENVIRONMENT			SPATIAL	
		sp1	sp2	sp3	sp4	sp5	sp6	envA	envB	envC	lat	lon
SUs	plot1	82	0	0	0	8	9	6	0.1	1023	45.842	120.44
	plot2	13	9	8	0	0	0	4	0.4	56	46.992	121.01
	plot3	0	1	0	0	0	2	18	0.7	127	45.532	118.84
	plot4	63	0	3	0	0	0	7	0.7	85	45.092	117.82
	plot5	0	0	7	4	0	1	6	0.7	54	45.582	120.72
	plot6	0	0	9	4	9	0	12	0.5	877	45.842	119.04
	plot7	0	5	2	0	0	0	3	0.2	34	45.732	118.45
	plot8	51	2	10	8	1	0	6	0.9	22	45.802	119.03
	plot9	0	0	2	0	0	0	4	0.9	27	46.982	119.38
	plot10	0	10	1	0	10	0	10	0.7	29	46.792	119.16

TRAITS	traitA	long	short	short	long	long	long
	traitB	10.7	8.9	0.1	1.4	9	1.1
	traitC	y	y	n	n	n	y

PHYLOGENY



The data

Mafragh, Algeria vegetation data

We will use one core dataset throughout this tutorial. The `veg` dataset gives information about spatial coordinates, species, environment, traits, and phylogeny for plants on the Mafragh coastal plain in North Africa.

Specifically, `veg` is a list containing five items:

- `xy` 97 observations of 2 spatial coordinates
- `spe` 97 observations of 56 plant species
- `env` 97 observations of 11 soil environmental variables
- `tra` 12 traits for 56 plant species
- `phy` phylogeny for 56 plant species

Load data

```r
load('./data/veg.rda')  # load the object (contains multiple objects)
xy  <- veg$xy           # spatial
spe <- veg$spe          # species
env <- veg$env          # environment
tra <- veg$tra          # traits
phy <- veg$phy          # phylogeny
rm(veg)                 # cleanup
ls()                    # list objects now in this local environment
```

```
## [1] "colvec"      "env"         "get_palette" "phy"         "pkg"
## [6] "snk"         "spe"         "tra"         "xy"
```

How would you examine what's inside one of these objects?

Transformations

Transformations

```r
### Species: express abundances as 0/1 presence/absence
# (spe > 0) * 1

### Species: express abundances on log10 scale
spe <- data.frame(log10(spe + 1))

### Environment: express variables in [0,1] range
env <- data.frame(vegan::decostand(env, 'range'))

### Traits: express traits in [0,1] range
tra <- data.frame(vegan::decostand(tra, 'range'))
```

Why do we transform community data?

Outliers

Outliers

Define multivariate outlier function

```
`outliers` <- function (x, mult=2, method='bray') {
    d       <- as.matrix(vegan::vegdist(x, method=method, binary=F, diag=T, upper=T))
    diag(d) <- as.numeric(1)     # avoid zero-multiplication
    m       <- apply(d, 2, mean) # site means
    z       <- scale(m)          # z-scores
    data.frame(mean_dist = m, z = z, is_outlier = abs(z) >= mult)
}
```

```
o <- outliers(spe, mult=2)
head(o, 3)
```

```
##   mean_dist          z is_outlier
## 1 0.8487643  0.3107496      FALSE
## 2 0.8283680 -0.2028542      FALSE
## 3 0.8214233 -0.3777314      FALSE
```

```
which(o$is_outlier)
```

```
## [1] 32 45 95 97
```

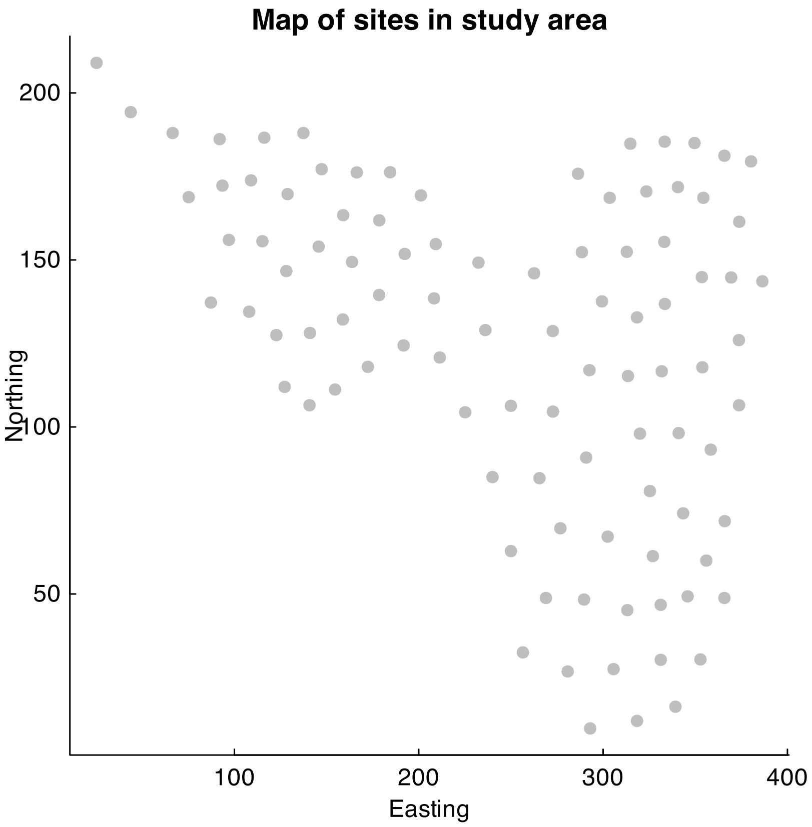
Test validity of species matrix

Species matrix must have no missings, no empty SUs, no empty species.

```
!anyNA(spe)                    # expect TRUE, no missing values

all(rowSums(spe, na.rm=T) > 0) # expect TRUE, no empty sites

all(colSums(spe, na.rm=T) > 0) # expect TRUE, no empty species
```

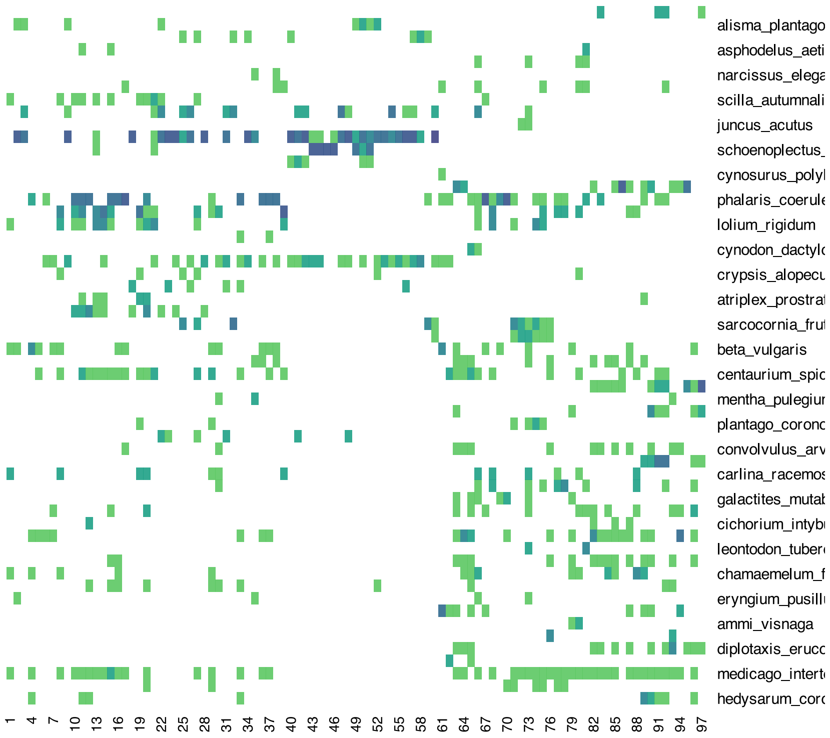
Visualize data

```
### spatial
plot(xy, pch=19, col='grey')

### species (abundances shown as colors)
vegan::tabasco(spe, col=get_palette())

### environment (normalized vals shown as
colors)
vegan::tabasco(env, col=get_palette())

### traits
vegan::tabasco(tra, col=get_palette())

### phylogeny
plot(phy, cex=0.6, no.margin=TRUE)
```



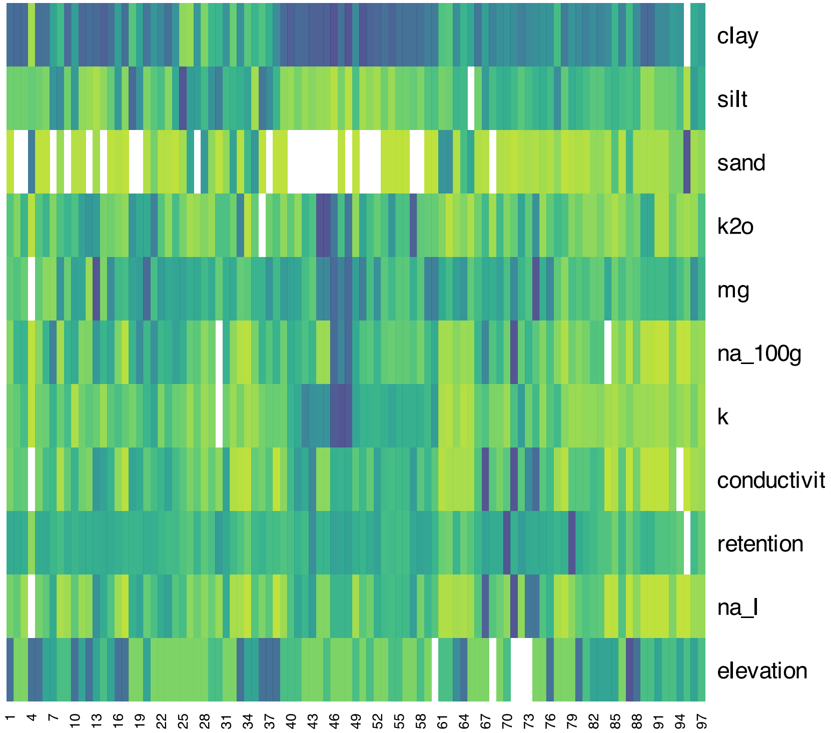
Visualize data

```
### spatial
plot(xy, pch=19, col='grey')

### species
vegan::tabasco(spe, col=get_palette())

### environment
vegan::tabasco(env, col=get_palette())

### traits
vegan::tabasco(tra, col=get_palette())

### phylogeny
plot(phy, cex=0.6, no.margin=TRUE)
```



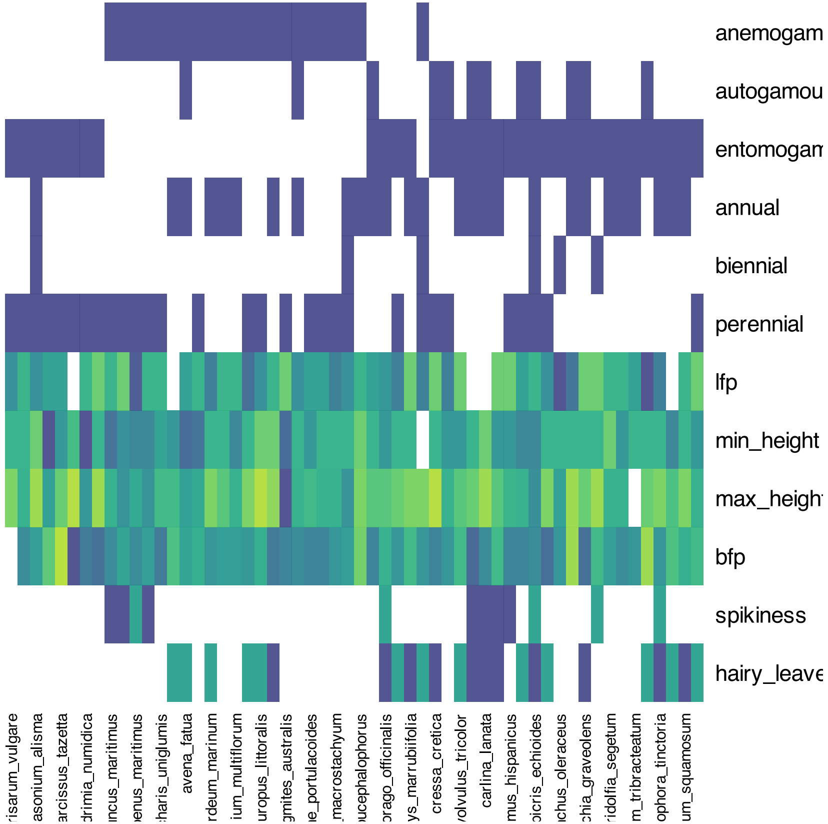
Visualize data

```r
### spatial
plot(xy, pch=19, col='grey')

### species
vegan::tabasco(spe, col=get_palette())

### environment
vegan::tabasco(env, col=get_palette())

### traits
vegan::tabasco(tra, col=get_palette())

### phylogeny
plot(phy, cex=0.6, no.margin=TRUE)
```



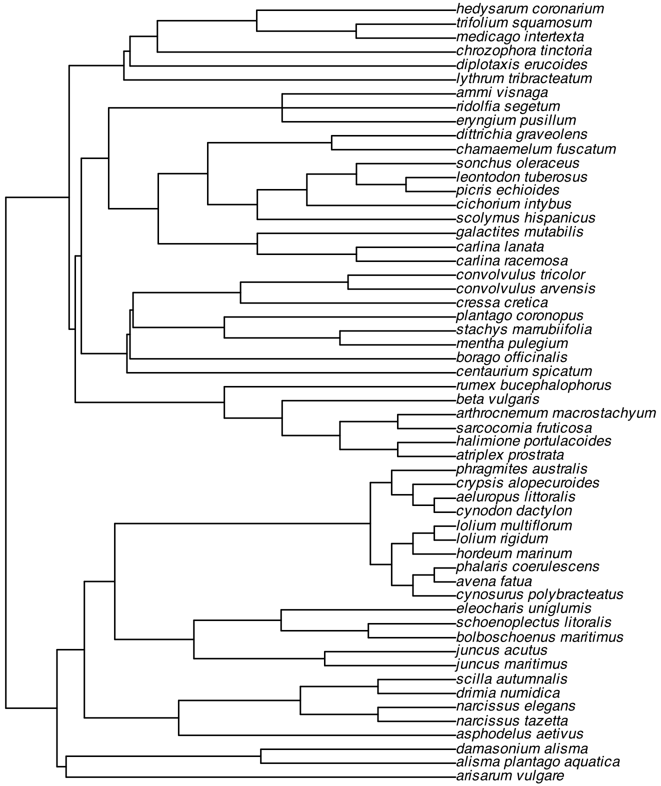
Visualize data

```r
### spatial
plot(xy, pch=19, col='grey')

### species
vegan::tabasco(spe, col=get_palette())

### environment
vegan::tabasco(env, col=get_palette())

### traits
vegan::tabasco(tra, col=get_palette())

### phylogeny
plot(phy, cex=0.6, no.margin=TRUE)
```



Visualize data

```
### spatial
plot(xy, pch=19, col='grey')

### species
vegan::tabasco(spe, col=get_palette())

### environment
vegan::tabasco(env, col=get_palette())

### traits
vegan::tabasco(tra, col=get_palette())

### phylogeny
plot(phy, cex=0.6, no.margin=TRUE)
```



Diversity

Diversity measures

How do you define "diversity"?

```
### Gamma (regional) diversity
gamma <- sum(colSums(spe) > 0)
gamma
```

```
## [1] 56
```

```
### Alpha (per-site) diversity
alpha <- rowSums(spe > 0)  # within-site
avgalpha <- mean(alpha) # average within-site

avgalpha
```

```
## [1] 6.268041
```

```
### Beta (among-site) diversity: Whittaker's
beta <- gamma / avgalpha - 1
beta
```

```
## [1] 7.934211
```

Diversity measures (β-diversity)

```r
### 1 -- proportion of zeros in the matrix (independent of abundance)
propzero <- sum(spe < .Machine$double.eps) / prod(dim(spe))
cat('Proportion of zeros in matrix:', propzero, '\n')
```

```
## Proportion of zeros in matrix: 0.8880707
```

```r
### 2 -- "dust bunny index" of McCune and Root (2015) (uses abundances)
dbi <- 1 - mean(as.matrix(vegan::decostand(spe, method='max')))
cat('Dust bunny index:', dbi, '\n')
```

```
## Dust bunny index: 0.9202329
```

```r
### 3 -- pairs of SUs that don't share species
z <- vegan::no.shared(spe)
propnoshare <- sum(z) / length(z)
cat('', propnoshare, 'proportion of site-pairs share no species in common\n')
```

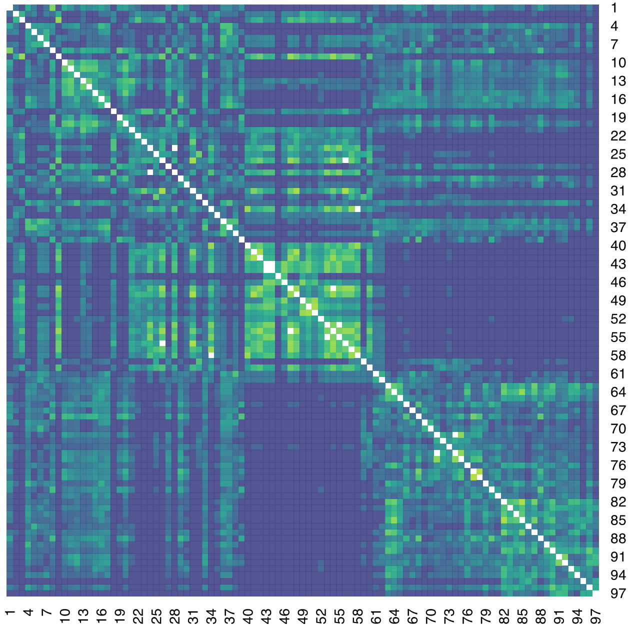
```
##  0.441366 proportion of site-pairs share no species in common
```

Dissimilarities

Dissimilarities

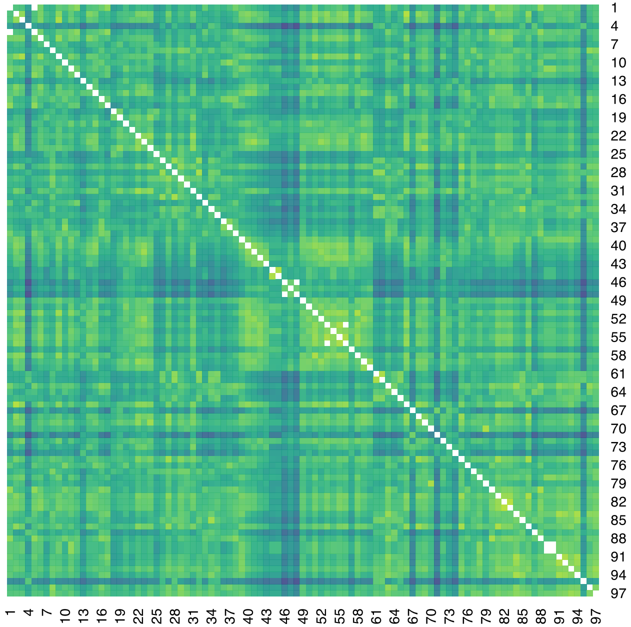
Species

```
d <- vegdist(spe, method='bray', binary=T)
tabasco(as.matrix(d), col=get_palette())
```



Environment

```
E <- vegdist(env, method='euc', binary=F)
tabasco(as.matrix(E), col=get_palette())
```



Dissimilarity Metrics for β-diversity

The specific measure used might depend on your particular ecological question, your data collection methods, or on the variable types you are working with.

High values == more dissimilar (different)

Low values == more similar (same)

Whittaker - variation in species composition across multiple sites.

Jaccard - focuses on shared species, using presence-absence data only. Range 0-1.

Bray-Curtis - considers species composition and abundance. Range 0-1.

Gower - measures general dissimilarity and can be used with continuous, ordinal, or categorical data. Values may appear smaller compared to Jaccard or Bray-Curtis results as it is more conservative (normalizes differences and downweights minor differences in pres-abs)

Dissimilarity Metrics Example Code

The specific measure used might depend on your particular ecological question, your data collection methods, or on the variable types you are working with.

Jaccard - focuses on shared species, using presence-absence data only.

```r
jaccard <- vegdist(spe, method = "jaccard")
```

Bray-Curtis - considers species composition and abundance.

```r
bray_curtis <- vegdist(spe, method = "bray")
```

Gower - measures general dissimilarity and can be used with continuous, ordinal, or categorical data.

```r
gower <- vegdist(spe, method = "gower")
```

Partitioning turnover and nestedness

Turnover is related to species replacement across sites (or time)

Nestedness is the proportion of species that are subset of the species that occur at another site (or time period). In other words, are the species at one site found in another site?

```
# Calculate beta diversity (Uniqueness and Nestedness)
beta_diversity <- betapart::beta.pair(spe)
# spe_pa <- vegan::decostand(spe, method = "pa")
# beta_diversity <- betapart::beta.pair(spe_pa)
```

There are three items in the list output.

beta.sim: turnover component; represents dissimilarity due to species turnover/replacement. High SIM suggests species replacement is a main driver of site differences (not richness)

beta.sne: nestedness component; represents dissimilarity due to shared species or richness differences. High SNE suggests species loss and nestedness (richness differences) are main drivers of site differences.

beta.sor: total dissimilarity, using Sorenson or Jaccard dissimilarity

Loss of sensitivity problem

How dissimilar are SUs A and C? They share no species in common...

```
##     sp1 sp2 sp3 sp4 sp5
## suA   1   1   0   0   0
## suB   0   1   1   1   0
## suC   0   0   0   1   1
```

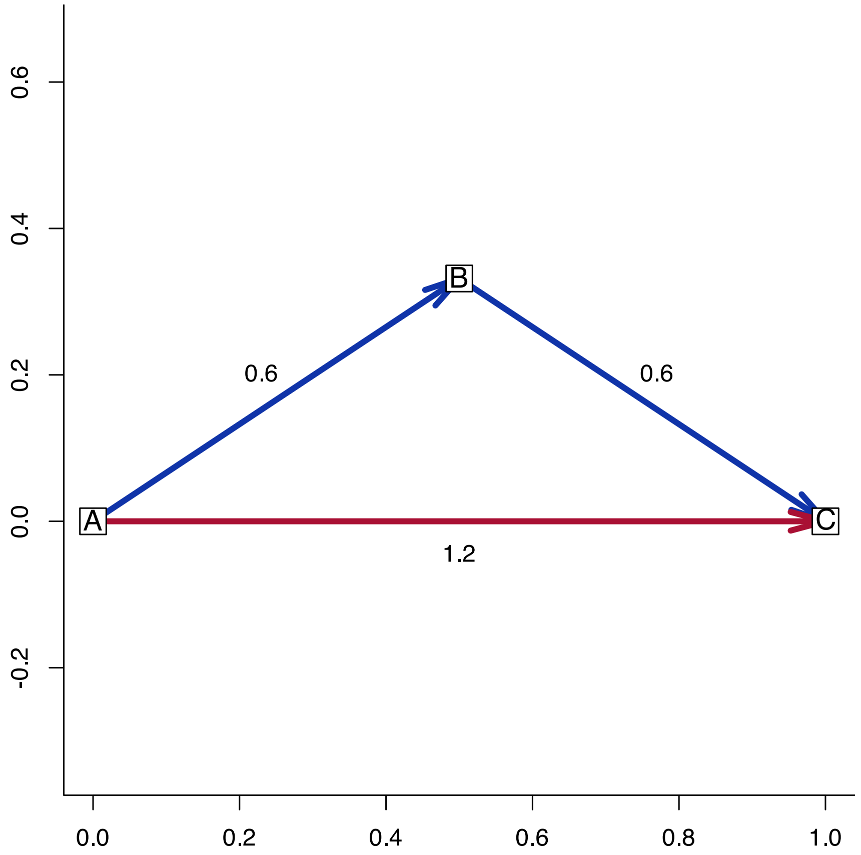
...Bray-Curtis will max out at 1.0...

```
##     suA suB
## suB 0.6
## suC 1.0 0.6
```

...but `stepacross()` replaces "too long" distances with shortest path.

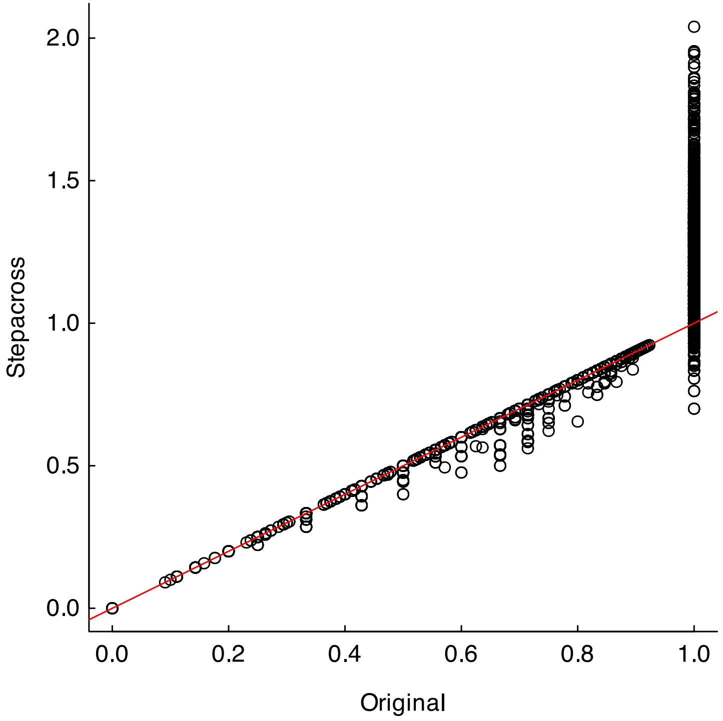
```
##     suA suB
## suB 0.6
## suC 1.2 0.6
```

Stepacross adjustment



Stepacross adjustment

```
D <- stepacross(d, 'shortest', toolong = 1)
plot(d, D, xlab = 'Original', ylab = 'Stepacross')
```

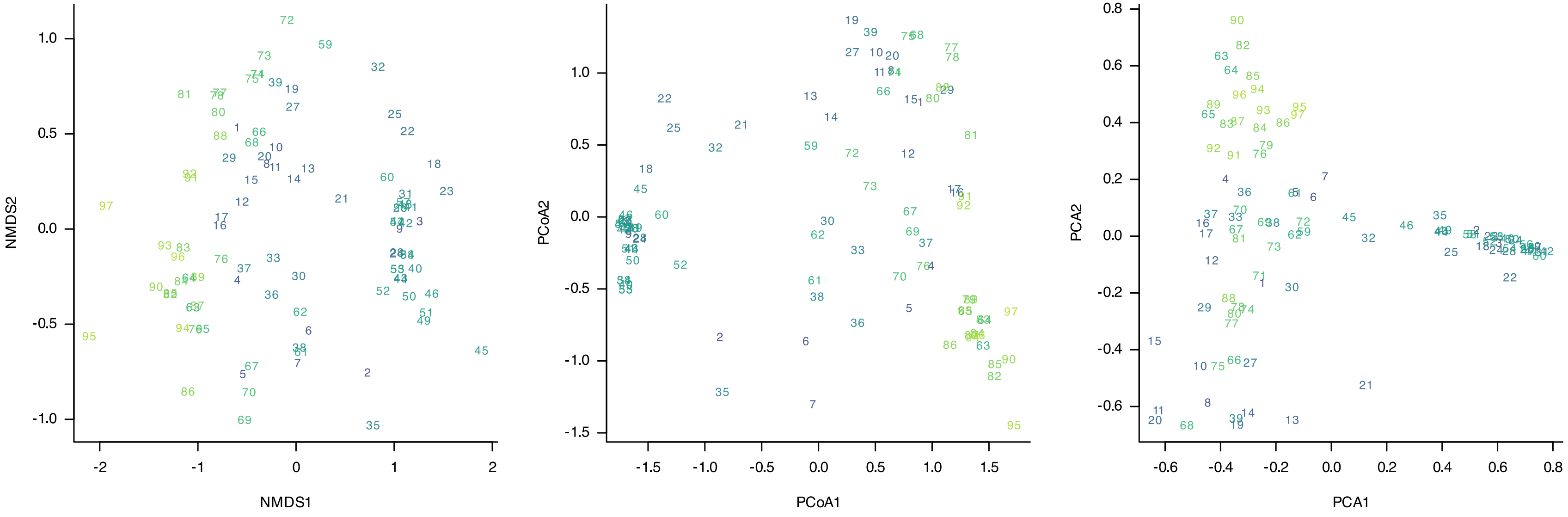


Ordination

Ordination: unconstrained

Three different algorithms

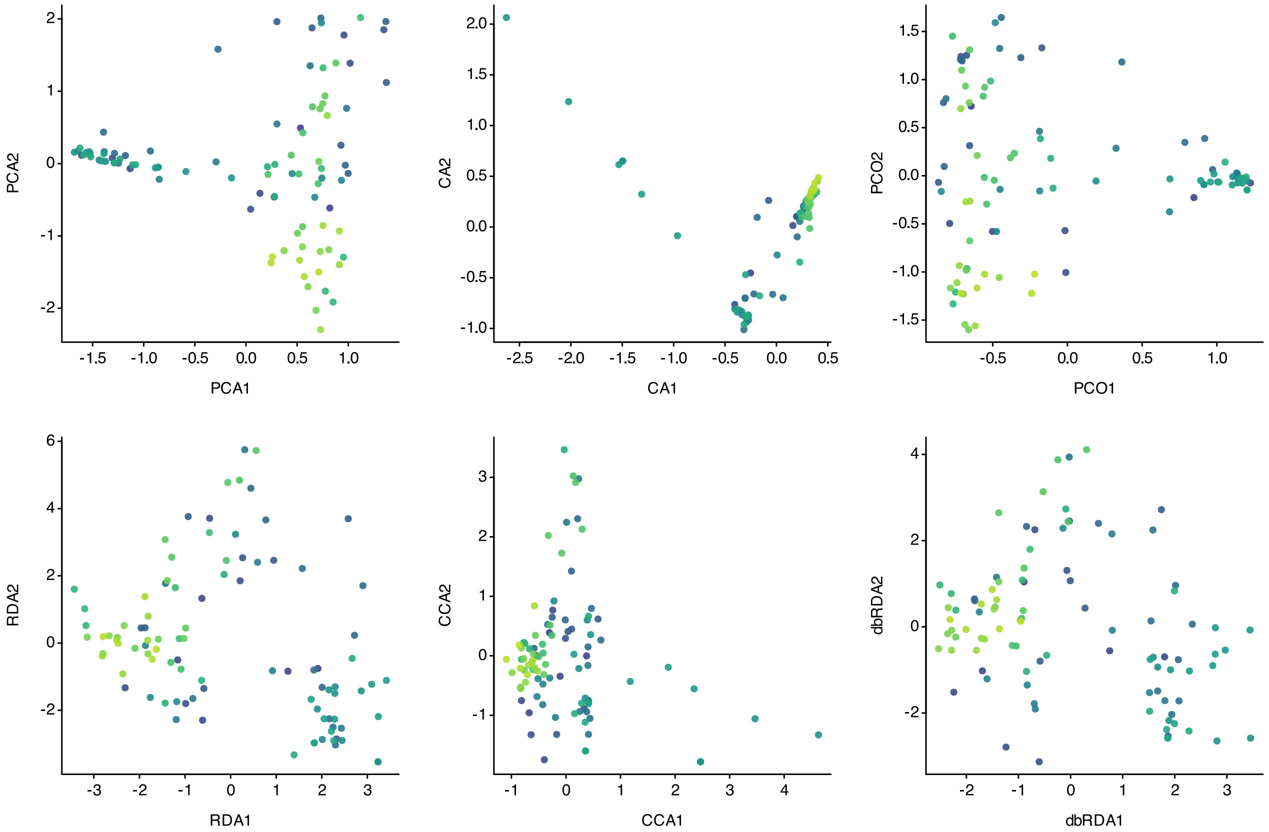
```
m1 <- metaMDS(D, k=2, maxit=250, try=100, trymax=101, trace=0) # NMS
m2 <- cmdscale(D, k=2, add=T)                                  # PCoA
m3 <- prcomp(spe)                                              # PCA
```



dbRDA: the Swiss army knife

```r
### dissimilarities
D_euc  <- vegdist(spe, 'euc')  # Euclidean
D_chi  <- vegdist(spe, 'chi')  # Chi-sq
D_bc   <- vegdist(spe, 'bray') # Bray-Curtis (Sorensen)

### PCA: PCO based on Euclidean distances (see `stats::prcomp`)
pca <- dbrda(D_euc ~ 1)

### CA/RA: PCO based on Chi-sq distances (see `vegan::cca`)
ca  <- dbrda(D_chi ~ 1)

### PCO: generalizes to any dissimilarity (see `labdsv::pco`)
pco <- dbrda(D_bc  ~ 1)

### RDA: constrained form of PCA (see `vegan::rda`)
rda <- dbrda(D_euc ~ k2o + mg, data = env)

### CCA: constrained form of CA (see `vegan::cca`)
cca <- dbrda(D_chi ~ k2o + mg, data = env)

### dbRDA: constrained form of PCO; can handle neg eigenvalues
dbr <- dbrda(D_bc  ~ k2o + mg, data = env, add = 'lingoes')
```

dbRDA: the Swiss army knife



What are the points?

What is the ordination space?

What does distance between points mean?

Scaling scores

Using `scores(x, scaling = ...)`:

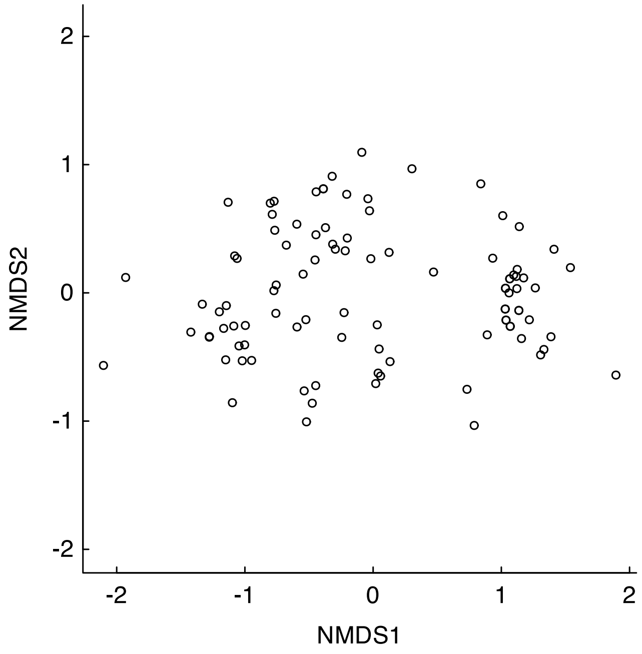
- `scaling = 1` — focus on sites, scale site scores by λ_i
- `scaling = 2` — focus on species, scale species scores by λ_i
- `scaling = 3` — **symmetric scaling**, scale both scores by $\sqrt{\lambda_i}$
- `scaling = -1` — as above, but for `rda()` get correlation scores
- `scaling = -2` — for `cca()` multiply results by $\sqrt{(1/(1 - \lambda_i))}$
- `scaling = -3` — this is Hill's scaling
- `scaling < 0` — for `rda()` divide species scores by species' σ
- `scaling = 0` — raw scores

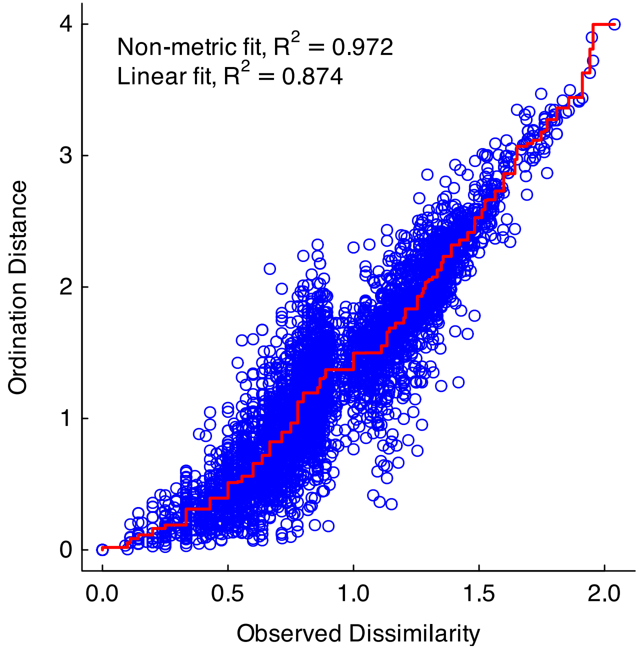
...where λ_i is the *i*th eigenvalue.

Credit: Gavin Simpson

NMS: a sensible default

```r
nms <- vegan::metaMDS(D, k = 2, maxit = 500, try = 500, trymax = 501)
plot(nms) # barebones -- what would you do to interpret this?
stressplot(nms)
```



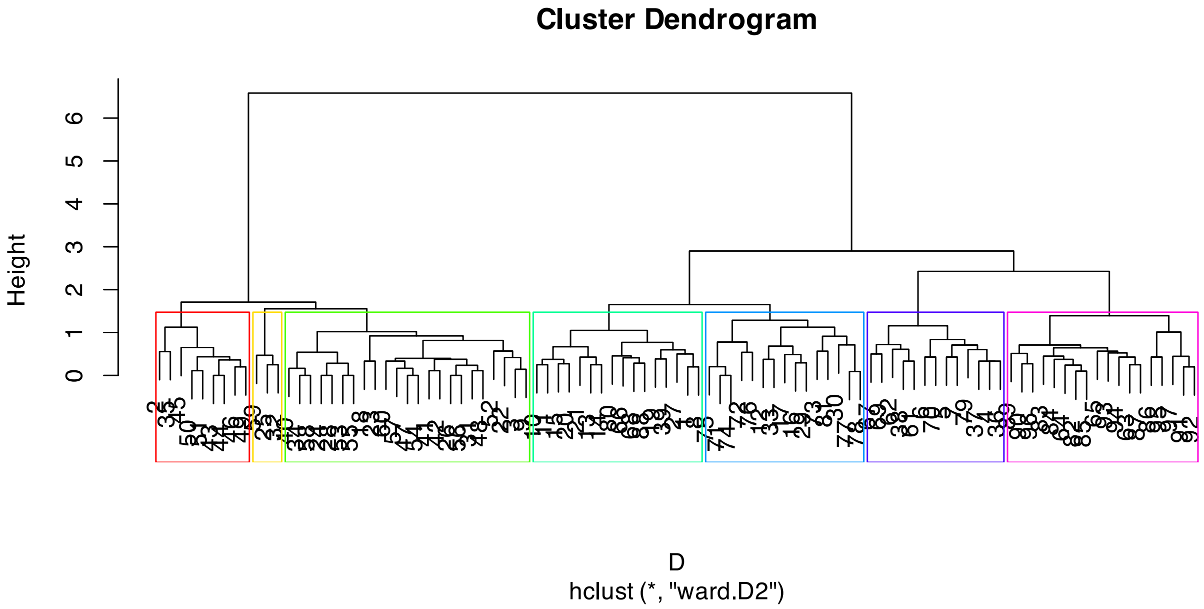


Group clustering

Group clustering

Hierarchical: Ward's clustering

```
k   <- 7                                             # specify number of groups
cl  <- hclust(D, method='ward.D2')                   # clustering solution
grp <- cutree(cl, k)                                 # group memberships
plot(cl) ; rect.hclust(cl, k, border=rainbow(k))     # plot the dendrogram
```



Group clustering

Non-hierarchical: fuzzy clustering

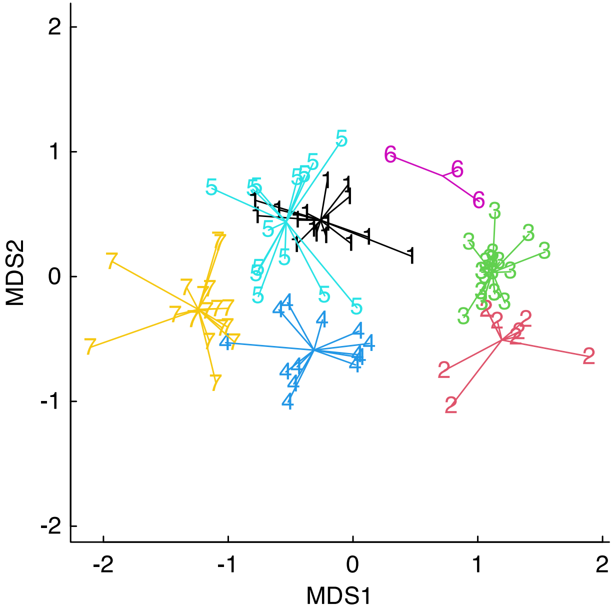
```
install.packages(vegclust)
require(vegclust)
k     <- 7                                    # specify number of groups
cl    <- vegclust::vegclustdist(D, mobileMemb=k)   # clustering solution
grp_f <- cl$memb                              # fuzzy membership
grp_c <- vegclust::defuzzify(cl, 'cut', alpha=0.8) # crisp membership
```

Group differences

Group differences

First, define and visualize groups (arbitrary example)

```
plot(m1$points, pch=NA, asp=1)                      # visualize on NMS
text(m1$points, labels=grp, col=as.numeric(grp))    # group memberships
ordispider(m1, groups=grp, col=1:k)                 # group centroids
```



Group differences

PERMANOVA: test for differences in multivariate *centroid*

```
adonis2(D ~ grp, permu=99)
```

```
## Permutation test for adonis under reduced model
## Permutation: free
## Number of permutations: 99
##
## adonis2(formula = D ~ grp, permutations = 99)
##          Df SumOfSqs      R2      F Pr(>F)
## Model     1    9.510 0.20354 24.278   0.01 **
## Residual 95   37.214 0.79646
## Total    96   46.724 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Recall how we defined the groups in `grp`; are you comfortable with this hypothesis test?

Group differences

PERMDISP: test for differences in multivariate *dispersion*

```
permutest(betadisper(D, grp), pairwise=TRUE, permu=99)
```

```
##             Df    Sum Sq    Mean Sq        F N.Perm Pr(>F)
## Groups       6 0.4708883 0.07848139 4.364745     99   0.01
## Residuals   90 1.6182676 0.01798075       NA     NA     NA
```

```
## Permuted p-values:
```

```
##  1-2  1-3  1-4  1-5  1-6  1-7  2-3  2-4  2-5  2-6  2-7  3-4  3-5  3-6  3-7  4-5
## 0.60 0.10 0.03 0.02 0.03 0.36 0.54 0.11 0.02 0.37 0.29 0.02 0.01 0.39 0.01 0.37
##  4-6  4-7  5-6  5-7  6-7
## 0.02 0.43 0.02 0.11 0.08
```

PERMANOVA extensions

Can use categorical and/or continuous predictors

```
adonis2(D ~ grp + elevation + clay + silt + conductivity, permu=99)
```

Blocked design: permutations must occur *within* strata

```
blk <- factor(letters[sample(rep(1:12,len=nrow(env)))]) # arbitrary 'blocks'
pm  <- how(nperm=999)             # setup permutation object
setBlocks(pm) <- blk              # permute only *within* blocks
adonis2(D ~ grp, permu=pm)        # correct test
```

Group indicator species

Group indicator species

Real groups

```
iv  <- labdsv::indval(spe, grp) # indicator species analysis for *real* groups
summary(iv)                     # IndVal observed
```

```
##                            cluster indicator_value probability
## halimione_portulacoides          1          0.1854       0.047
## atriplex_prostrata               1          0.2731       0.009
## carlina_racemosa                 1          0.4144       0.002
## centaurium_spicatum              1          0.4172       0.002
## scilla_autumnalis                1          0.5288       0.001
## hordeum_marinum                  1          0.6991       0.001
## lolium_rigidum                   1          0.7014       0.001
## alisma_plantago_aquatica         2          0.3194       0.020
## schoenoplectus_litoralis         2          0.7283       0.001
## aeluropus_littoralis             3          0.3929       0.008
## bolboschoenus_maritimus          3          0.5295       0.001
## phalaris_coerulescens            4          0.2832       0.037
## beta_vulgaris                    4          0.4280       0.005
## carlina_lanata                   5          0.2305       0.035
## plantago_coronopus               5          0.2854       0.009
## trifolium_squamosum              5          0.2966       0.023
```

Group indicator species

Random groups

```
rnd <- sample(grp, length(grp), replace=T) # define random groups by bootstrapping
ivr <- labdsv::indval(spe, rnd) # indicator species analysis for *random* groups
summary(ivr)                    # IndVal expected at random
```

```
##                   cluster indicator_value probability
## drimia_numidica         6          0.2903       0.008
```

Null expectation, setting alpha = 0.05

```
ceiling(ncol(spe) * 0.05)
```

```
## [1] 3
```

Community traits

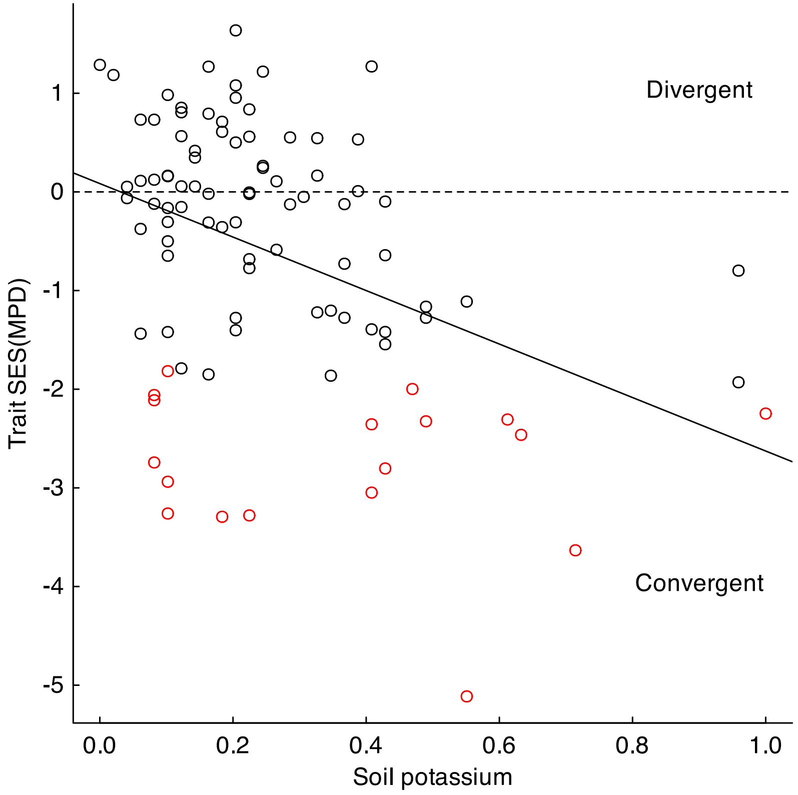
Community traits

```
head(tra, 4)
```

```
##                          anemogamous autogamous entomogamous annual biennial
## arisarum_vulgare                   0          0            1      0        0
## alisma_plantago_aquatica           0          0            1      0        0
## damasonium_alisma                  0          0            1      1        1
## asphodelus_aetivus                 0          0            1      0        0
##                          perennial       lfp min_height max_height       bfp
## arisarum_vulgare                 1 0.6444426  0.4114081  0.1879018 0.0000000
## alisma_plantago_aquatica         1 0.4065980  0.4114081  0.4362945 0.6666667
## damasonium_alisma                1 0.6444426  0.2342243  0.1099155 0.5555556
## asphodelus_aetivus               1 0.5374933  1.0000000  0.5462100 0.2777778
##                          spikiness hairy_leaves
## arisarum_vulgare                 0            0
## alisma_plantago_aquatica         0            0
## damasonium_alisma                0            0
## asphodelus_aetivus               0            0
```

Community traits

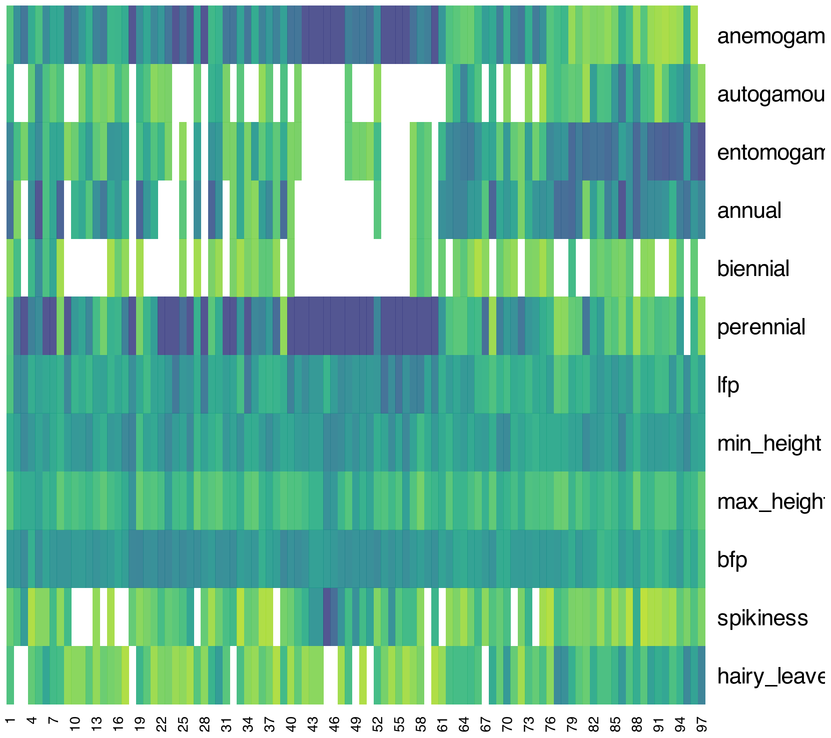
```r
### Euclidean trait dissimilarity
###   (traits already scaled 0-1)
Dt  <- dist(tra, method='euc')

### calculate trait SES of mean
###   pairwise distances in sites
ses <- picante::ses.mpd(spe, Dt,
         null.model='richness')

### plot SES across a nutrient gradient
u <- ifelse(ses$mpd.obs.p < 0.05,2,1)
plot(ses$mpd.obs.z ~ env$k,
     ylab='Trait SES(MPD)',
     xlab='Soil potassium', col=u)
abline(lm(ses$mpd.obs.z ~ env$k)) # regression
abline(h=0, lty=2) # random-traits line
text(0.9, 1, 'Divergent')
text(0.9, -4, 'Convergent')
```



Community weighted means

(Weighted) mean trait value per SU

```
### function to make CWM matrix
`makecwm` <- function (spe, tra) {
  spe <- as.matrix(spe)
  tra <- as.matrix(tra)
  `stdz` <- function(x) {
    (x - min(x, na.rm=TRUE)) /
        diff(range(x, na.rm=TRUE))
  }
  tra <- apply(tra, MARGIN = 2, FUN = stdz)
  awt <- spe %*% tra # abund-weighted totals
  awt / rowSums(spe, na.rm=TRUE) # CWM matrix
}

### make the CWM traits matrix
cwm <- data.frame(makecwm(spe, tra))

### visualize
tabasco(cwm, col=get_palette())
```



Ordination of CWM traits

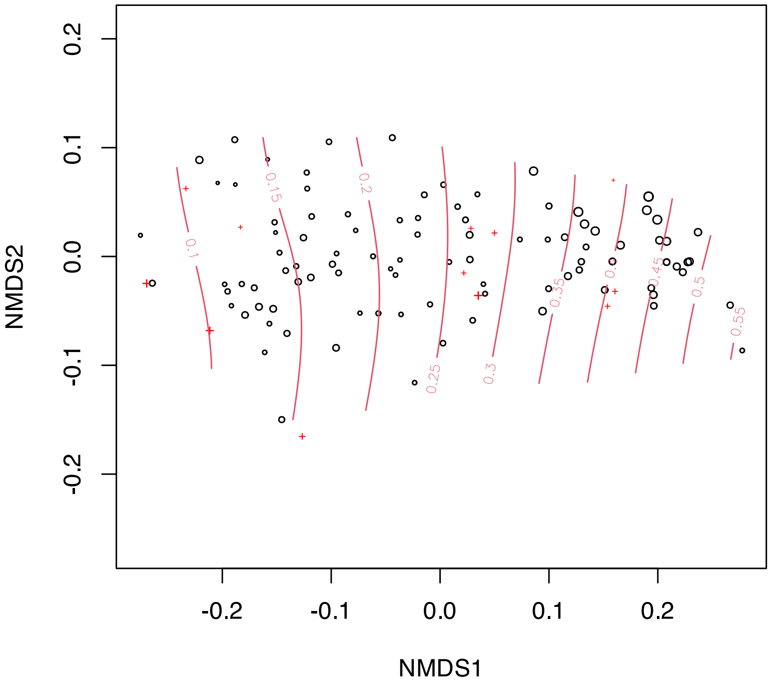
Admits nonlinear trait-enviro relationship

```
### NMS based on abundance-weighted traits
m <- metaMDS(cwm, 'altGower', k=2, trace=0)

### SUs sized relative to a leaf trait
plot(m, cex=cwm$lfp)

### overlay enviro variable
o <- ordisurf(m, env$k, col=2, add=T)
```

What are the points?

What is the ordination space?

What does distance between points indicate?

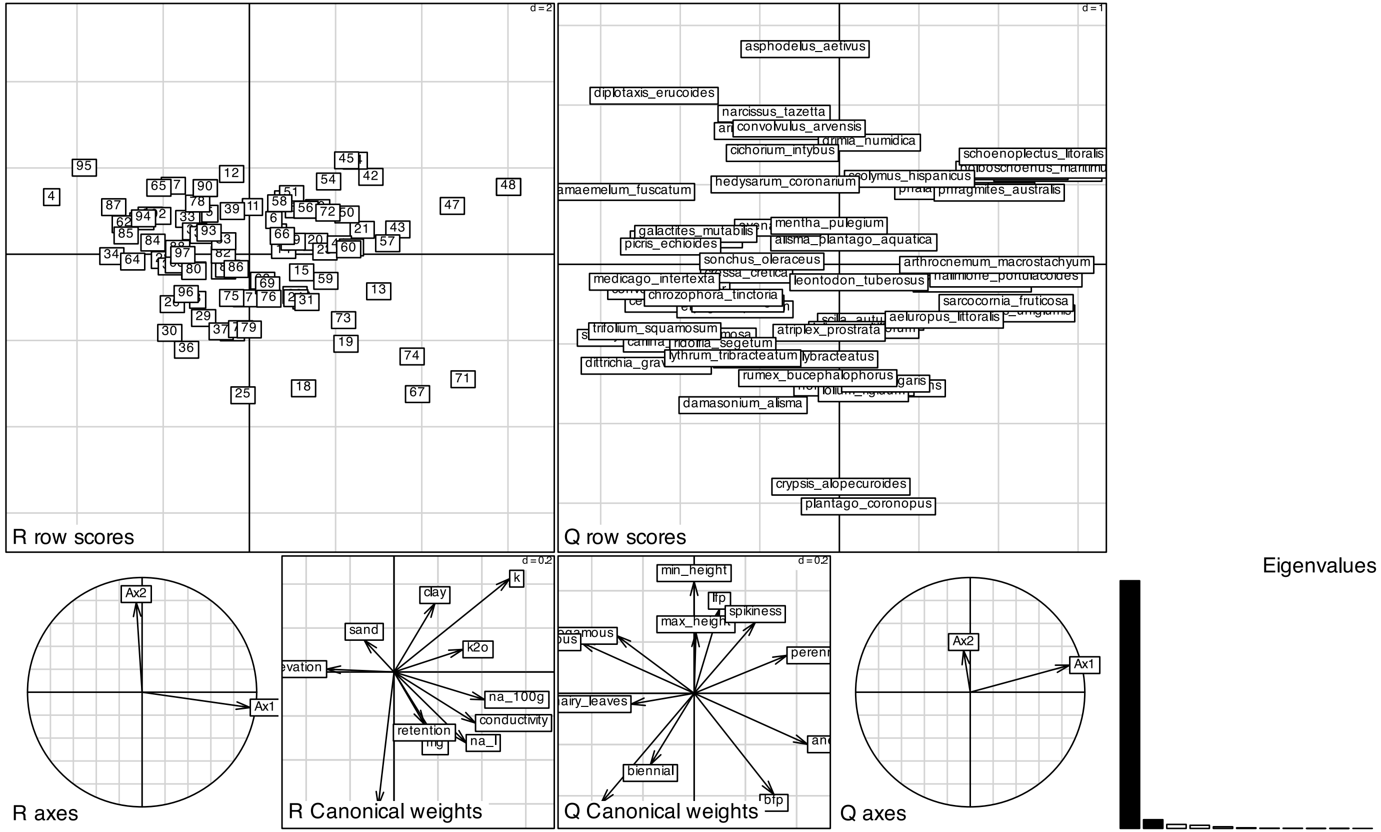


Fourth-corner analysis and RLQ

the RLQ method

```
require(ade4)
o_spe <- dudi.coa(spe, F) # Correspondence Analysis
o_env <- dudi.hillsmith(env, F, row.w = o_spe$lw)
o_tra <- dudi.hillsmith(tra, F, row.w = o_spe$cw)
r     <- rlq(o_env, o_spe, o_tra, F)
plot(r)
summary(r)
randtest(r)
```

Fourth-corner analysis and RLQ



Fourth-corner analysis and RLQ

Trait correlations

```r
fourthcorner.rlq(r, type='Q.axes')
```

```
## Fourth-corner Statistics
## ------------------------
## Permutation method  Comb. 2 and 4  ( 999  permutations)
##
## Adjustment method for multiple comparisons:   holm
## call:  fourthcorner.rlq(xtest = r, typetest = "Q.axes")
##
## ---
##
##                    Test Stat           Obs     Std.Obs     Alter Pvalue
## 1   AxcR1 / anemogamous    r  0.3305888147  5.69292462 two-sided  0.001
## 2   AxcR2 / anemogamous    r -0.0458582587 -0.77960338 two-sided  0.444
## 3    AxcR1 / autogamous    r -0.2245440975 -2.08793024 two-sided   0.03
## 4    AxcR2 / autogamous    r  0.0516421890  1.11307304 two-sided  0.259
## 5  AxcR1 / entomogamous    r -0.3305888147 -5.69292462 two-sided  0.001
## 6  AxcR2 / entomogamous    r  0.0458582587  0.77960338 two-sided  0.444
## 7        AxcR1 / annual    r -0.2710550381 -2.52925916 two-sided  0.007
## 8        AxcR2 / annual    r -0.0984989149 -1.67108085 two-sided  0.089
```

Fourth-corner analysis and RLQ

Environment correlations

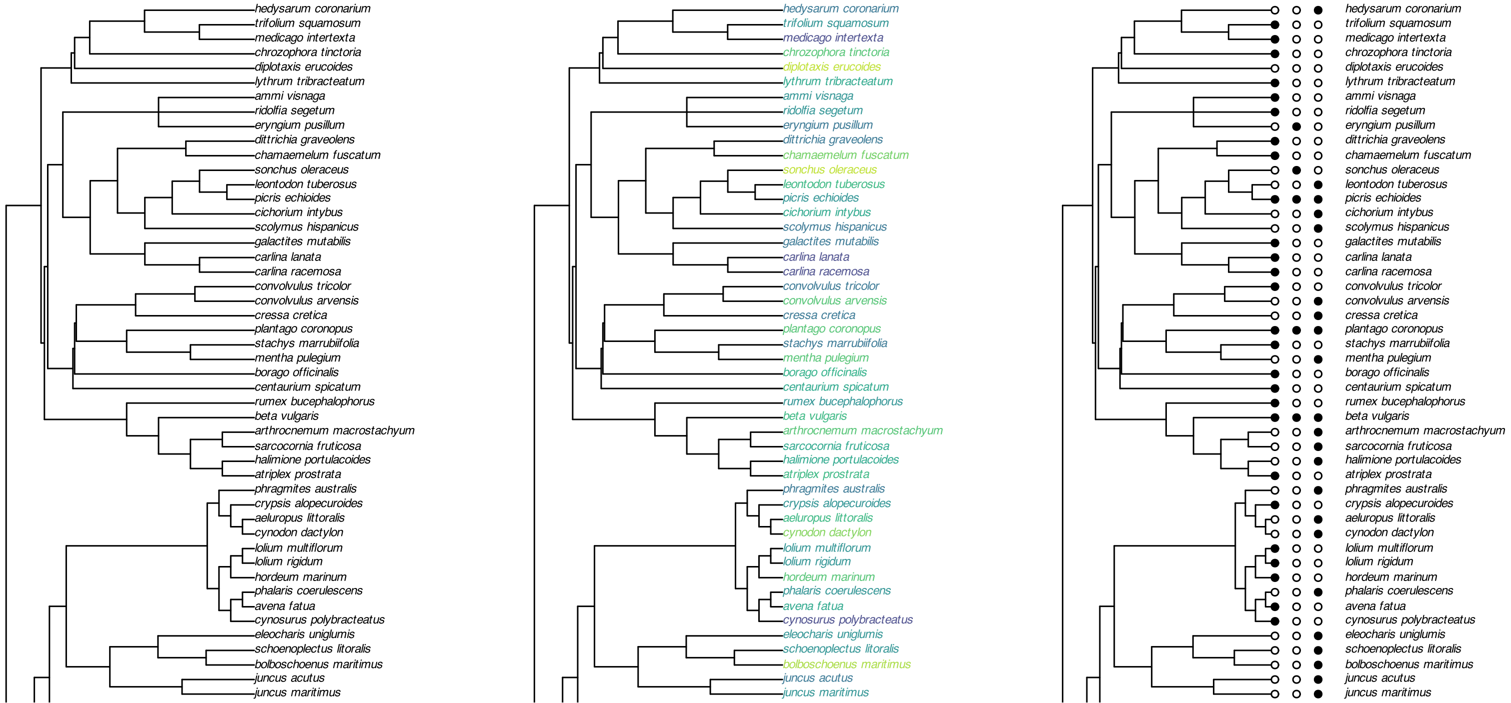
```r
fourthcorner.rlq(r, type='R.axes')
```

```
## Fourth-corner Statistics
## ------------------------
## Permutation method  Comb. 2 and 4  ( 999  permutations)
##
## Adjustment method for multiple comparisons:   holm
## call:  fourthcorner.rlq(xtest = r, typetest = "R.axes")
##
## ---
##
##                     Test Stat          Obs     Std.Obs     Alter Pvalue
## 1          clay / AxcQ1    r  0.158920982  2.44189264 two-sided  0.014
## 2          silt / AxcQ1    r -0.061153914 -0.96433329 two-sided  0.326
## 3          sand / AxcQ1    r -0.115820713 -1.79748542 two-sided  0.077
## 4           k2o / AxcQ1    r  0.268313148  4.11998443 two-sided  0.001
## 5            mg / AxcQ1    r  0.160253315  2.39412055 two-sided  0.014
## 6       na_100g / AxcQ1    r  0.353355511  5.34613443 two-sided  0.001
## 7             k / AxcQ1    r  0.449746890  6.69092568 two-sided  0.001
## 8  conductivity / AxcQ1    r  0.316892896  4.83476275 two-sided  0.001
```

Community phylogenetics

Community phylogenetics: basic plotting

```
par(mfrow=c(1,3), mar=c(0,0,0,0), oma=c(0,0,0,0))          # plotting parameters
plot(phy, cex=0.75, no.margin=T)                           # basic plotting
plot(phy, cex=0.75, tip.color=colvec(tra$lfp), no.margin=T) # color labels by traits
plot(phy, cex=0.75, label.offset=48, no.margin=T)          # symbolize trait values at tips
tiplabels(pch=21, bg = c(tra$annual),    adj=0)
tiplabels(pch=21, bg = c(tra$biennial),  adj=15)
tiplabels(pch=21, bg = c(tra$perennial), adj=30)
```



Phylogenetic diversity - alpha

```r
### phylogenetic distance matrix
Dp  <- cophenetic(phy)

### Faith's phylogenetic diversity:
###  total branch length connecting
###   all species per site
fpd <- pd(spe, phy)

### mean pairwise distance
mpd <- ses.mpd(
  spe, Dp, null.model='independentswap')

### mean nearest taxon distance
mnd <- ses.mntd(
  spe, Dp, null.model='independentswap')

### bring all together
phy_div <-
  cbind(fpd,
        mpd  = mpd$mpd.obs.z,
        mntd = mnd$mntd.obs.z)
```

```
##          PD SR         mpd        mntd
## 1  767.32  6  0.92152467  1.08223033
## 2  551.60  4  1.00127789  0.89956948
## 3  364.80  3 -0.74186332 -0.52641802
## 4  694.38  6 -0.23269719  0.08822563
## 5  395.80  3 -0.26034867  0.33585163
## 6  323.60  3 -1.03913225 -1.66716963
## 7  485.76  4 -0.17221377 -0.18773460
## 8  932.32  9  0.04080037  0.08182749
## 9  476.80  4 -1.01997248 -0.09066436
## 10 579.60  6 -1.65459527 -1.13524065
## 11 879.94 10 -0.48729928 -1.97117056
## 12 817.38  7 -0.09923170  0.53220249
```

Phylogenetic diversity - beta

```
### correlation between phylogenetic and taxonomic beta–diversity
Dp <- picante::phylosor(spe, phy) # phylogenetic distances
protest(Dp, D)                    # procrustes correlation
```

```
##
## Call:
## protest(X = Dp, Y = D)
##
## Procrustes Sum of Squares (m12 squared):        0.2736
## Correlation in a symmetric Procrustes rotation: 0.8523
## Significance:  0.001
##
## Permutation: free
## Number of permutations: 999
```

Phylogenetic signal

```r
sapply(tra, FUN=function(j){
  names(j) <- rownames(tra)
  round(picante::phylosignal(j, ape::multi2di(phy)), 4)})
```

```
##                         anemogamous autogamous entomogamous annual  biennial
## K                       3.2136      0.3401     3.2136       0.3466  0.4564
## PIC.variance.obs        6e-04       0.0039     6e-04        0.0057  0.0016
## PIC.variance.rnd.mean 0.0066      0.0047     0.0066       0.0069  0.0026
## PIC.variance.P          0.001       0.251      0.001        0.148   0.158
## PIC.variance.Z          -5.3551     -0.6838    -5.316       -1.1289 -1.0962
##                         perennial lfp    min_height max_height bfp    spikiness
## K                       0.3951    0.3512 0.3451     0.3043     0.4064 0.9436
## PIC.variance.obs        0.005     0.0014 9e-04      9e-04      9e-04  0.001
## PIC.variance.rnd.mean 0.0069    0.0017 0.0011     0.001      0.0012 0.0033
## PIC.variance.P          0.04      0.17   0.266      0.464      0.124  0.001
## PIC.variance.Z          -1.7276   -0.8997 -0.6953   -0.2408    -1.083 -2.6172
##                         hairy_leaves
## K                       0.5273
## PIC.variance.obs        0.0024
## PIC.variance.rnd.mean 0.0044
## PIC.variance.P          0.007
## PIC.variance.Z          -2.3522
```

Community spatial analysis

Mantel test

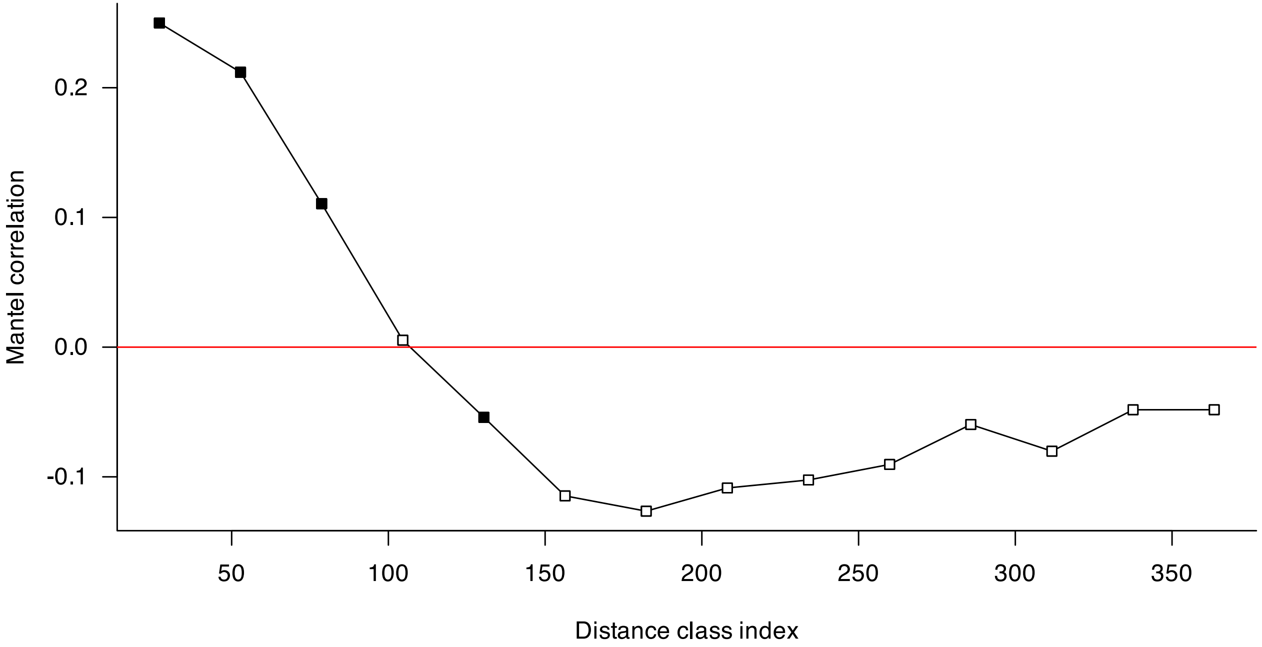
```r
E <- dist(xy)                                 # spatial distance matrix
vegan::mantel(D, E, method = 'spearman')   # spearman *rank* correlation
```

```
##
## Mantel statistic based on Spearman's rank correlation rho
##
## Call:
## vegan::mantel(xdis = D, ydis = E, method = "spearman")
##
## Mantel statistic r: 0.4093
##       Significance: 0.001
##
## Upper quantiles of permutations (null model):
##    90%    95%  97.5%    99%
## 0.0257 0.0350 0.0426 0.0495
## Permutation: free
## Number of permutations: 999
```

```r
ecodist::mantel(D ~ E, mrank = T) # alternative with useful bootstrap CIs
```

Mantel correlogram

```
plot(vegan::mantel.correlog(D, E, cutoff=F, r.type='spearman', nperm=99, mult='holm'))
```



Multiple regression on distance matrices

```r
# species dissimilarities are related
#  to space (extremely weakly) and
#   potassium (moderately)
ecodist::MRM(
  D
   ~ dist(xy) + dist(env$k))
```

```
## $coef
##                        D  pval
## Int         0.627953130 1.000
## dist(xy)    0.001755839 0.001
## dist(env$k) 0.301671934 0.001
##
## $r.squared
##        R2      pval
## 0.1994291 0.0010000
##
## $F.test
##        F   F.pval
## 579.5514   0.0010
```

```r
# abundance of cosmopolitan bulrush
#  is NOT related to space, but is
#   related to potassium (moderately)
ecodist::MRM(
  dist(spe$bolboschoenus)
   ~ dist(xy) + dist(env$k))
```

```
## $coef
##             dist(spe$bolboschoenus)  pval
## Int                     0.201552720 1.000
## dist(xy)                0.000223733 0.091
## dist(env$k)             0.411223852 0.001
##
## $r.squared
##         R2       pval
## 0.06277826 0.00100000
##
## $F.test
##        F   F.pval
## 155.8368   0.0010
```

Wrap up

Acknowledgments

ESA Vegetation Section

Martin Dovciak

Carissa Brown

Morgan D. Frost

Anita Thompson

Further details

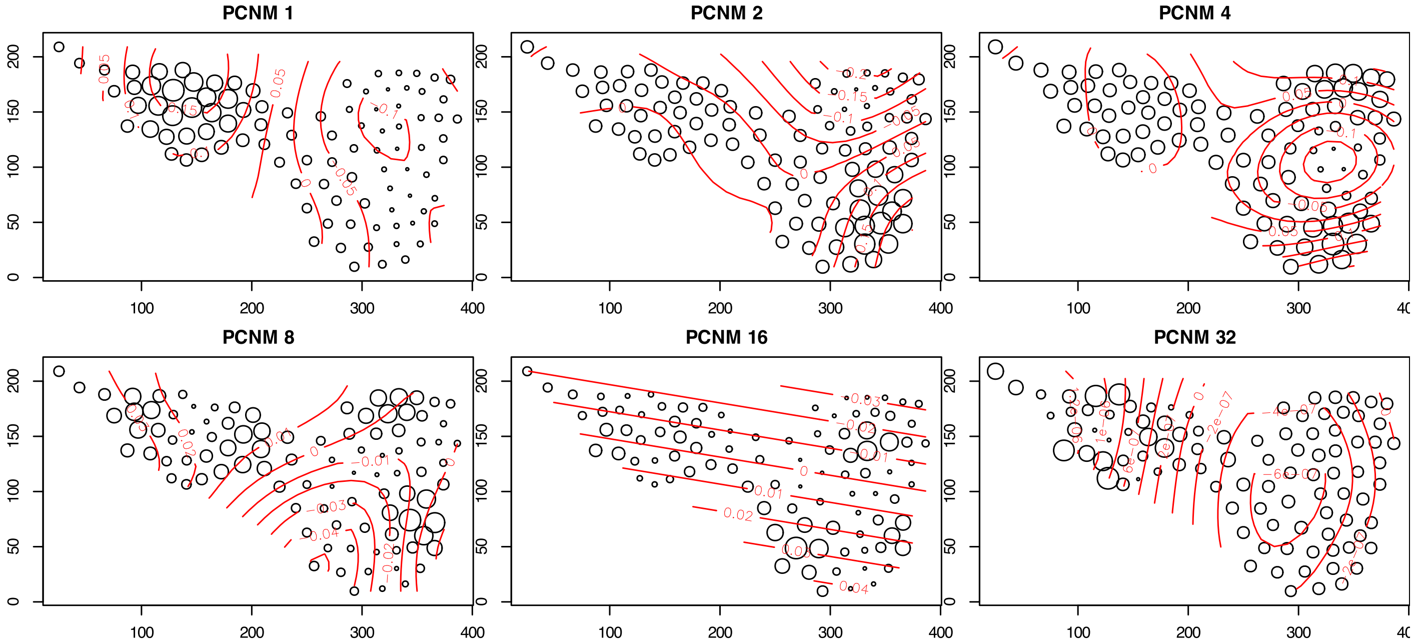
https://ecol.shinyapps.io/esa_tutorial/

More??

PCNM

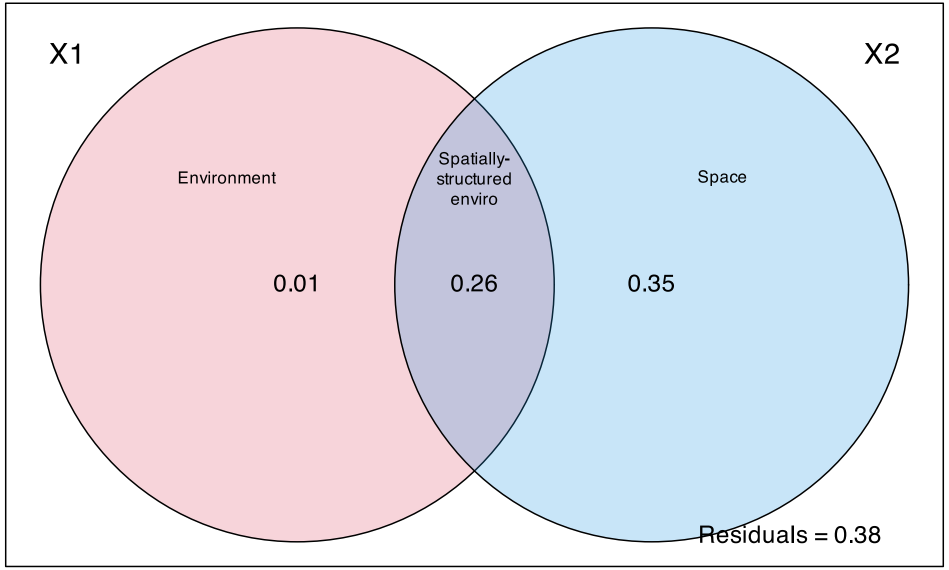
PCNM: principle coordinates of neighbor matrices

```
E   <- dist(xy)                 # euclidean distances between sites
pc  <- pcnm(E)                  # principal coordinates of neighbor matrices
```



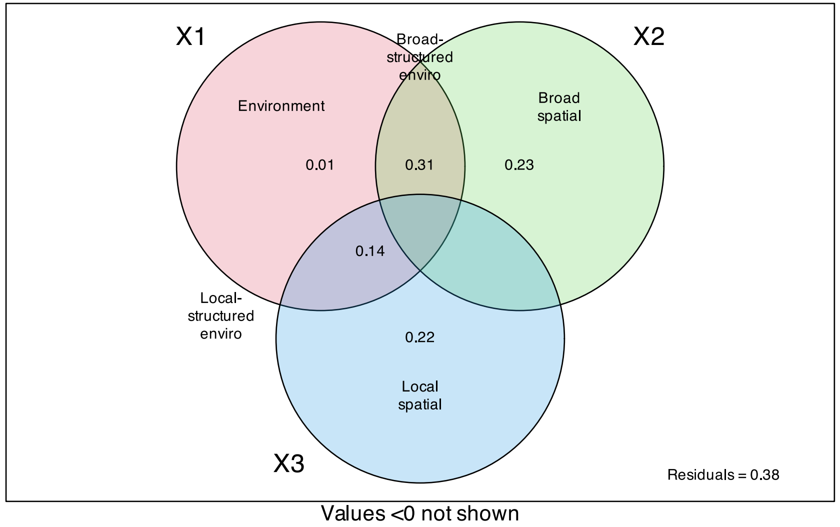
PCNM

```
rs <- rowSums(spe) / sum(spe)      # sites weighted by abundances
pc <- pcnm(E, w=rs)                # *weighted* PCNMs
vp <- varpart(D, env, vegan::scores(pc))  # variation partitioning
```



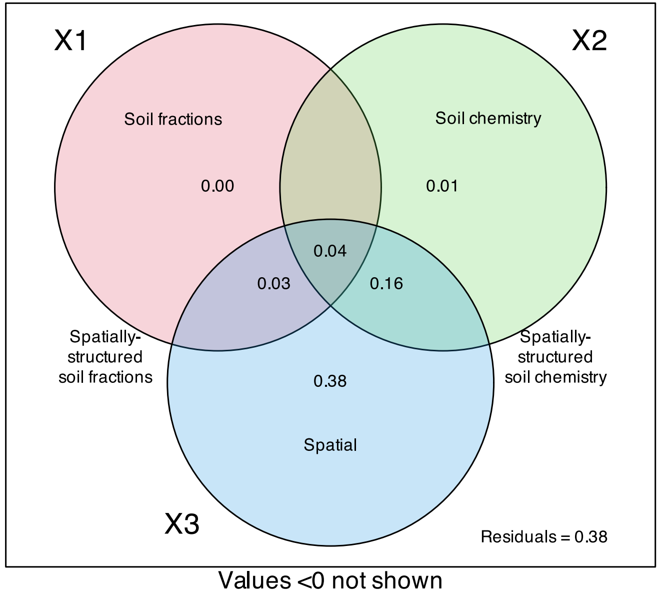
PCNM

```
### broad and local structure
pc_broad <- vegan::scores(pc)[,1:10]
pc_local <- vegan::scores(pc)[,11:59]
vp       <- varpart(D, env, pc_broad,
pc_local)
```



PCNM

```r
### soil physical, soil chemistry, and spatial predictors
vp <- varpart(
  D,                                 # dissimilarities
  ~ clay + sand + silt,              # soil fractions
  ~ mg + k + conductivity + na_l,  # soil chemistry
  vegan::scores(pc),                 # spatial predictors
  data = env)                        # environmental dataset
```



Conclusion