

# Multivariate Models

*Zach Proux*

*2/8/2018*

For this assignment will be analyzing data on the Vegetation and Environment in Dutch Dune Meadows.

To import the data and read the metadata run the following:

```
library(vegan)
```

```
## Loading required package: permute
```

```
## Loading required package: lattice
```

```
## This is vegan 2.4-6
```

```
data(dune)
```

```
data(dune.env)
```

```
?dune
```

1. Conduct an indirect ordination on the dune plant community. Specifically, visually examine a NMDS plot using the bray-curtis distance metric. Below is some code to help you develop a potential plot that emphasizes the role of the environmental variable “Moisture”.

```
mds_dune = metaMDS(dune, k=2, trymax=100)
```

```
## Run 0 stress 0.1192678
```

```
## Run 1 stress 0.1980528
```

```
## Run 2 stress 0.1808913
```

```
## Run 3 stress 0.1213932
```

```
## Run 4 stress 0.1183186
```

```
## ... New best solution
```

```
## ... Procrustes: rmse 0.02026923 max resid 0.06495118
```

```
## Run 5 stress 0.1183186
```

```
## ... Procrustes: rmse 1.910451e-05 max resid 6.316843e-05
```

```
## ... Similar to previous best
```

```
## Run 6 stress 0.1183186
```

```
## ... Procrustes: rmse 2.658106e-06 max resid 7.729436e-06
```

```
## ... Similar to previous best
```

```
## Run 7 stress 0.2019172
```

```
## Run 8 stress 0.1183186
```

```
## ... Procrustes: rmse 1.465953e-05 max resid 4.586439e-05
```

```
## ... Similar to previous best
```

```
## Run 9 stress 0.1183186
```

```
## ... Procrustes: rmse 1.026399e-05 max resid 3.161212e-05
```

```
## ... Similar to previous best
```

```
## Run 10 stress 0.1192678
```

```
## Run 11 stress 0.1192679
```

```
## Run 12 stress 0.1192679
```

```
## Run 13 stress 0.1192679
```

```
## Run 14 stress 0.1192687
```

```
## Run 15 stress 0.1183186
```

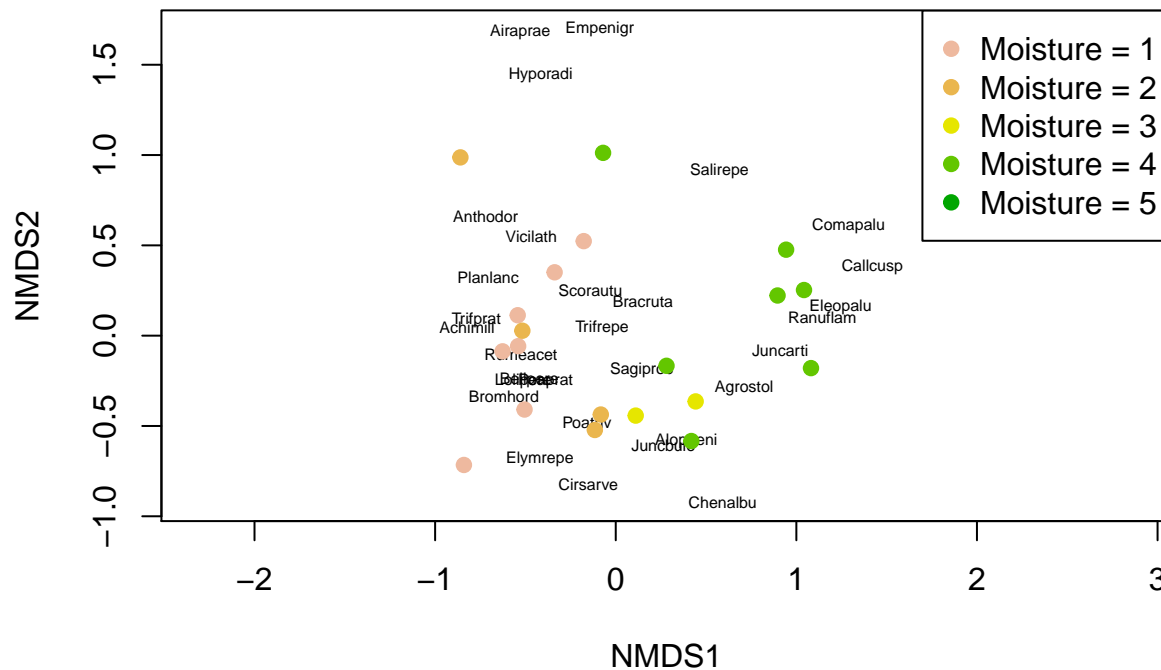
```
## ... Procrustes: rmse 1.816945e-05 max resid 5.663961e-05
```

```
## ... Similar to previous best
```

```
## Run 16 stress 0.1183186
```

```
## ... Procrustes: rmse 1.566725e-05  max resid 5.004719e-05
## ... Similar to previous best
## Run 17 stress 0.1183186
## ... Procrustes: rmse 1.166296e-05  max resid 3.121567e-05
## ... Similar to previous best
## Run 18 stress 0.1192679
## Run 19 stress 0.119268
## Run 20 stress 0.1192682
## *** Solution reached
```

```
plot(mds_dune, type='n')
text(mds_dune, 'sp', cex=.5)
# generate vector of colors
color_vect = rev(terrain.colors(6))[-1]
points(mds_dune, 'sites', pch=19,
       col=color_vect[dune.env$Moisture])
legend('topright', paste("Moisture = ", 1:5, sep=''),
       col=color_vect, pch=19)
```



Describe how you interpret the graphic.

The x-axis is the dominant axis of variation. To interpret the plot, we need to have some knowledge of the data. For example, because we know moisture affects community assemblage we can look at the plot and see that some species were associated with very high moisture and also have higher x values and other species were associated with dry areas and also have lower x values.

What is the goal of creating such a plot?

This approach is used when we are not quite sure what explanatory variables are important in explaining observed variation and we just want to visualize differences in a matrix of explanatory and dependent variables. Once we visualize it, the idea is your background knowledge might allow you to identify a trend in the multidimensional distribution.

Does this analysis suggest any interesting findings with respect to the dune vegetation?

This analysis suggests moisture may be the most important variable in explaining the variation in presence of meadow vegetation.

2. Carry out a direct ordination using CCA in order to test any potential hypotheses that you developed after examining the MDS plot. Specifically, carry out a test of the entire model (i.e., including all constrained axes) and also carry out tests at the scale of individual explanatory variables you included in your model if you included more than one variable. Plot your results.

```
cca_dune = cca(dune ~ ., data = dune)
cca_dune
```

```
## Call: cca(formula = dune ~ Achimill + Agrostol + Airaprae +
## Alop geni + Anthodor + Bellpere + Bromhord + Chenalbu + Cirsarve +
## Comapalu + Eleopal u + Elymrepe + Empenigr + Hyporadi + Juncarti +
## Juncbufo + Lolipere + Planlanc + Poaprat + Poatriv + Ranuflam +
## Rumeacet + Sagiproc + Salirepe + Scora utu + Trifprat + Trifrepe +
## Vicilath + Bracruta + Callcusp, data = dune)
##
##              Inertia Proportion Rank
## Total          2.115      1.000
## Constrained    2.115      1.000   19
## Unconstrained  0.000      0.000    0
## Inertia is mean squared contingency coefficient
## Some constraints were aliased because they were collinear (redundant)
##
## Eigenvalues for constrained axes:
##   CCA1  CCA2  CCA3  CCA4  CCA5  CCA6  CCA7  CCA8  CCA9  CCA10
## 0.5360 0.4001 0.2598 0.1760 0.1448 0.1079 0.0925 0.0809 0.0733 0.0563
##   CCA11 CCA12 CCA13 CCA14 CCA15 CCA16 CCA17 CCA18 CCA19
## 0.0483 0.0412 0.0352 0.0205 0.0149 0.0091 0.0079 0.0070 0.0035
```

```
RsquareAdj(cca_dune, 1000)
```

```
## $r.squared
## [1] 1
##
## $adj.r.squared
## [1] -Inf
```

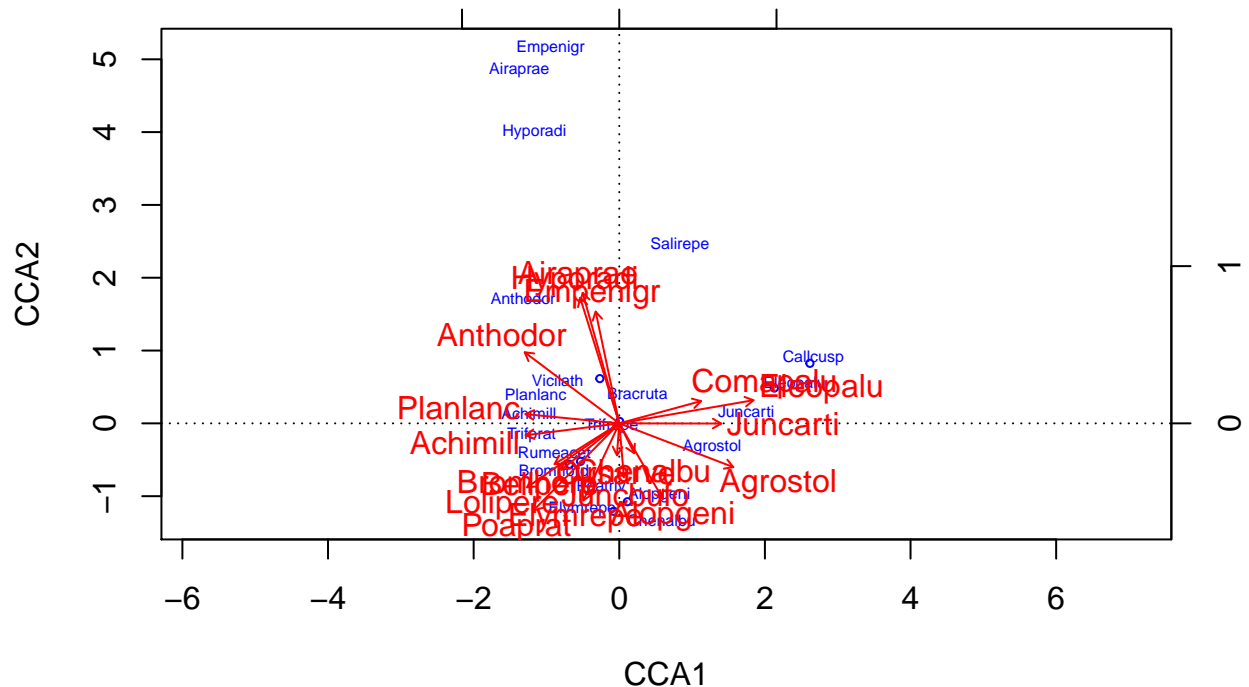
```
anova(cca_dune, permutations = 1000)
```

```
## No residual component
##
## Model: cca(formula = dune ~ Achimill + Agrostol + Airaprae + Alop geni + Anthodor + Bellpere + Bromhord +
##          Df ChiSquare F Pr(>F)
## Model    19    2.1153
## Residual  0    0.0000
```

```
anova(cca_dune, by = 'margin', permutations = 1000)
```

```
## No residual component
##
## Model: cca(formula = dune ~ Achimill + Agrostol + Airaprae + Alop geni + Anthodor + Bellpere + Bromhord +
##          Df ChiSquare F Pr(>F)
## Model    19    2.1153
## Residual  0    0.0000
```

```
plot(cca_dune, type = 'n', scaling = 1)
orditorp(cca_dune, display = 'sp', cex = 0.5, scaling = 1, col = 'blue')
text(cca_dune, display = 'bp', col = 'red')
```



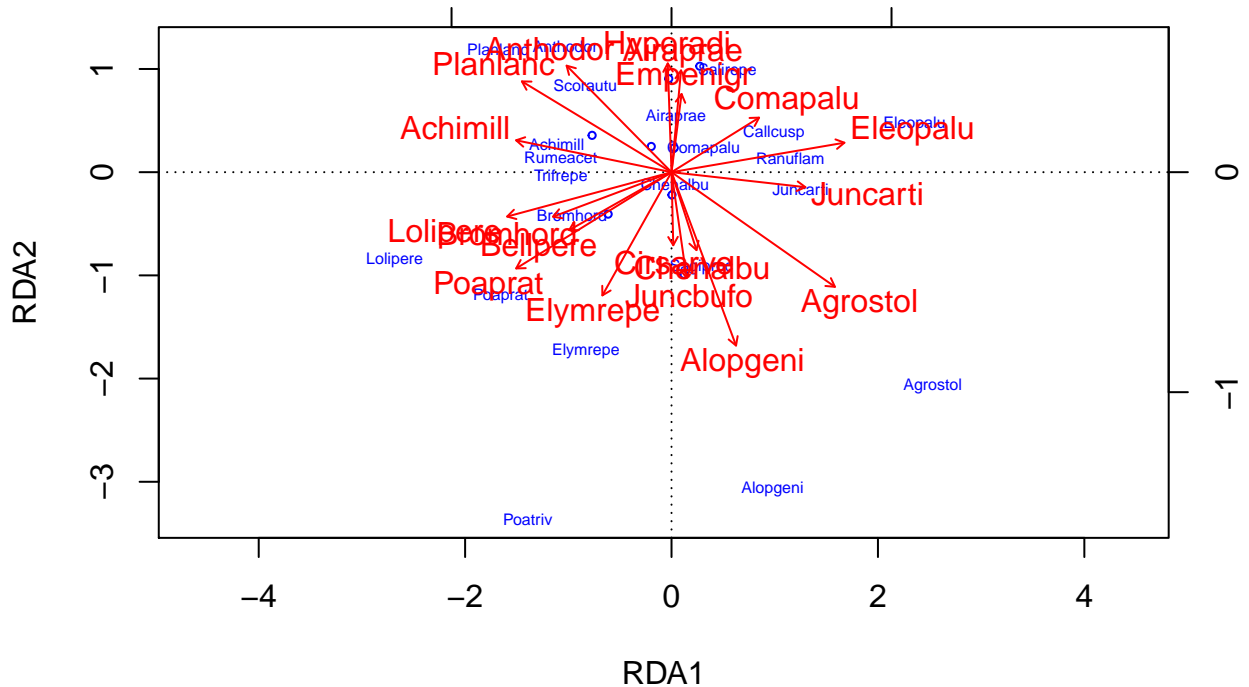
```
rda_dune = rda(dune ~ ., data = dune)
rda_dune
```

```
## Call: rda(formula = dune ~ Achimill + Agrostol + Airaprae +
## Alopogeni + Anthodor + Bellpere + Bromhord + Chenalbu + Cirsarve +
## Comapalu + Eleopal + Elymrepe + Empenigr + Hyporadi + Juncarti +
## Juncbufo + Lolipere + Planlanc + Poaprat + Poatriv + Ranuflam +
## Rumeacet + Sagiproc + Salirepe + Scorautu + Trifprat + Trifrepe +
## Vicilath + Bracruta + Callcusp, data = dune)
##
##              Inertia Proportion Rank
## Total          8.412e+01  1.000e+00
## Constrained    8.412e+01  1.000e+00   19
## Unconstrained  3.050e-30  3.626e-32    1
## Inertia is variance
## Some constraints were aliased because they were collinear (redundant)
##
## Eigenvalues for constrained axes:
##   RDA1   RDA2   RDA3   RDA4   RDA5   RDA6   RDA7   RDA8   RDA9   RDA10
## 24.795 18.147  7.629  7.153  5.695  4.333  3.199  2.782  2.482  1.854
##   RDA11  RDA12  RDA13  RDA14  RDA15  RDA16  RDA17  RDA18  RDA19
##  1.747  1.314  0.991  0.638  0.551  0.351  0.200  0.149  0.116
##
## Eigenvalues for unconstrained axes:
##      PC1
## 3.0504e-30
```

```
RsquareAdj(rda_dune)
```

```
## $r.squared
## [1] 1
##
## $adj.r.squared
## [1] NA
```

```
plot(rda_dune, type = 'n', scaling = 1)
orditorp(rda_dune, display = 'sp', cex = 0.5, scaling = 1, col = 'blue')
text(rda_dune, display = 'cn', col = 'red')
```



3. Do your two analyses agree with one another or complement one another or do these two analyses seem to be suggesting different take home messages? Which analysis do you find to be more useful?

I'm not confident in my understanding of this, but I think the analyses agree with one another. They both plot the species in multidimensional space in similar groupings: agrostol and alop geni opposite anthodor and planlanc. The nmcs was more helpful however, because it actually showed us that moisture could be responsible for the variation as opposed to the cca and rda which just group the species based on the characteristics of the plots they occurred in.