homework4

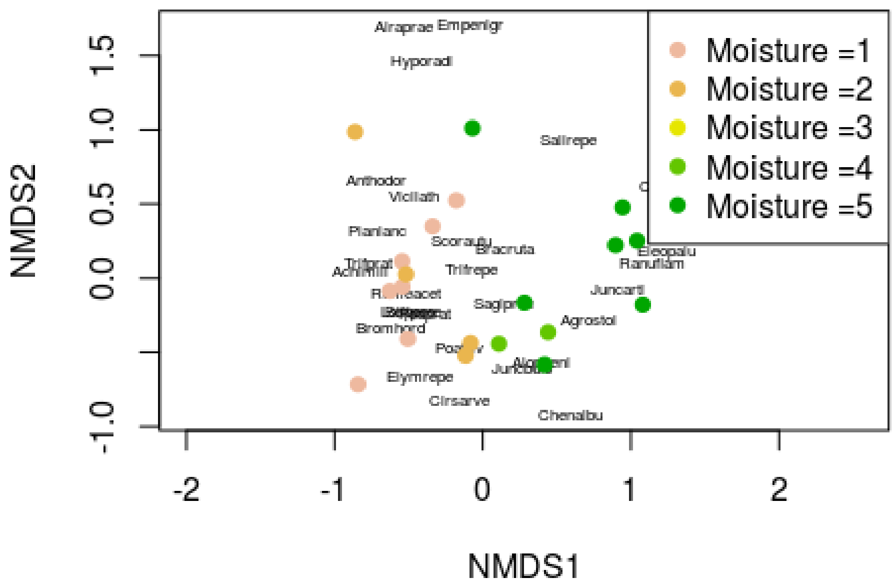
Taylor Miller

2/6/2020

```r
library(vegan)

data(dune)
data(dune.env)
?dune

dune.env$Moisture

##  [1] 1 1 2 2 1 1 1 5 4 2 1 4 5 5 5 5 2 1 5 5
## Levels: 1 < 2 < 4 < 5

dune.env$Moisture = factor(dune.env$Moisture, levels = 1:
                           5, ordered = FALSE)
```

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”. Describe how you interpret the graphic. What is the goal of creating such a plot? Does this analysis suggest any interesting findings with respect to the dune vegetation? To interpret a NMDS, objects that are closer together on the plot are more alike than those further apart. For this plot, species from the same habitat are more similiar than species from other habitats, Dan pointed out that many od the green dotted variables are wetland plants. This means that the plant species moisture category differs across their dune.

```r
dune_mds =  metaMDS(dune)

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



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. I ran a step function and the results indiacted that Use and Manure are not necesasry for the best function, the new cc_dune2: cca_dune2 <- update(cca_dune, . ~ . - Use - Manure). The adjusted r-squared only improved 1%, but the cca anova test suggests that the less complex model excluding the varibales Use and Manure are strongly supported.

```
cca_dune = cca(dune ~ ., data=dune.env)
RsquareAdj(cca_dune, 100)

## $r.squared
## [1] 0.7106267
##
## $adj.r.squared
## [1] 0.2167025

anova(cca_dune, permutations = 999)

## Permutation test for cca under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: cca(formula = dune ~ A1 + Moisture + Management + Use + Manure,
data = dune.env)
```

```
##          Df ChiSquare      F Pr(>F)
## Model    12    1.5032 1.4325  0.029 *
## Residual  7    0.6121
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(cca_dune, by='margin', permutations = 999)
```

```
## Permutation test for cca under reduced model
## Marginal effects of terms
## Permutation: free
## Number of permutations: 999
##
## Model: cca(formula = dune ~ A1 + Moisture + Management + Use + Manure,
data = dune.env)
##            Df ChiSquare      F Pr(>F)
## A1          1   0.11070 1.2660  0.224
## Moisture    3   0.31587 1.2041  0.220
## Management  2   0.15882 0.9081  0.575
## Use         2   0.13010 0.7439  0.779
## Manure      3   0.25490 0.9717  0.518
## Residual    7   0.61210
```

```r
step(cca_dune)
```

```
## Start:  AIC=86.86
## dune ~ A1 + Moisture + Management + Use + Manure
##
##              Df    AIC
## - Use         2 86.711
## <none>          86.857
## - Management  2 87.470
## - Manure      3 87.819
## - A1          1 88.181
## - Moisture    3 89.179
##
## Step:  AIC=86.71
## dune ~ A1 + Moisture + Management + Manure
##
##              Df    AIC
## - Manure      3 86.190
## - Management  2 86.446
## <none>          86.711
## - Moisture    3 87.873
## - A1          1 88.430
##
## Step:  AIC=86.19
## dune ~ A1 + Moisture + Management
##
##              Df    AIC
## <none>          86.190
```

```
## - Moisture    3 86.460
## - A1          1 86.813
## - Management  3 86.992

## Call: cca(formula = dune ~ A1 + Moisture + Management, data = dune.env)
## 
##               Inertia Proportion Rank
## Total          2.1153     1.0000     
## Constrained    1.1392     0.5385    7
## Unconstrained  0.9761     0.4615   12
## Inertia is scaled Chi-square 
## 
## Eigenvalues for constrained axes:
##   CCA1   CCA2   CCA3   CCA4   CCA5   CCA6   CCA7 
## 0.4483 0.3001 0.1499 0.1073 0.0567 0.0434 0.0335 
## 
## Eigenvalues for unconstrained axes:
##     CA1     CA2     CA3     CA4     CA5     CA6     CA7     CA8     CA9 
CA10 
## 0.30637 0.13191 0.11516 0.10947 0.07724 0.07575 0.04871 0.03758 0.03106 
0.02102 
##    CA11    CA12 
## 0.01254 0.00928
```

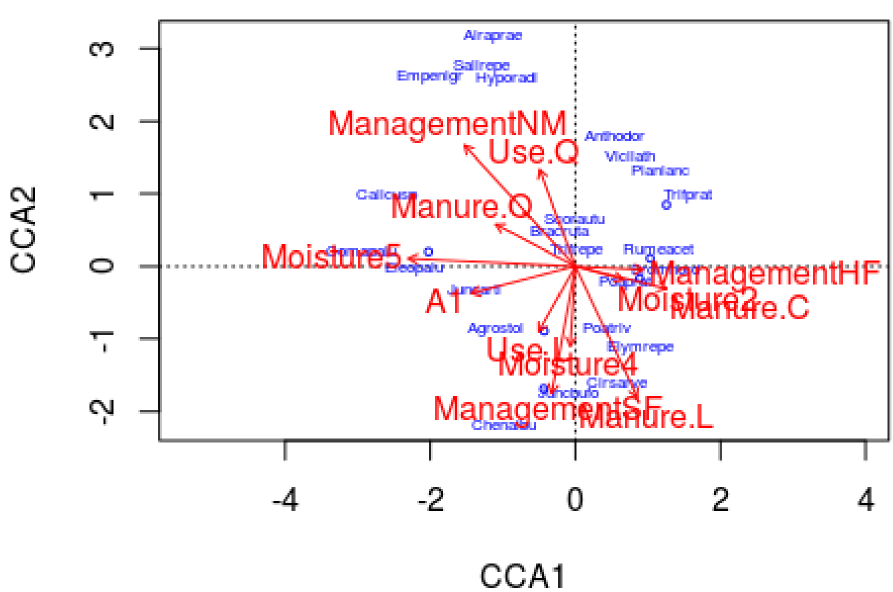
```
anova.cca(cca_dune)
```

```
## Permutation test for cca under reduced model
## Permutation: free
## Number of permutations: 999
## 
## Model: cca(formula = dune ~ A1 + Moisture + Management + Use + Manure, 
data = dune.env)
##          Df ChiSquare      F Pr(>F)  
## Model    12    1.5032 1.4325  0.024 *
## Residual  7    0.6121                
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cca_dune2 <- update(cca_dune, . ~ . -Use - Manure)
anova(cca_dune, cca_dune2)
```

```
## Permutation tests for cca under reduced model
## Permutation: free
## Number of permutations: 999
## 
## Model 1: dune ~ A1 + Moisture + Management + Use + Manure
## Model 2: dune ~ A1 + Moisture + Management
##   ResDf ResChiSquare Df ChiSquare      F Pr(>F)
## 1     7       0.6121                           
## 2    12       0.9761 -5    -0.364 0.8325  0.767
```

```
RsquareAdj(cca_dune2, 100)

## $r.squared
## [1] 0.5385459
##
## $adj.r.squared
## [1] 0.2756167

anova.cca(cca_dune2)

## Permutation test for cca under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: cca(formula = dune ~ A1 + Moisture + Management, data = dune.env)
##          Df ChiSquare      F Pr(>F)
## Model     7    1.1392 2.0007  0.001 ***
## Residual 12    0.9761
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

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')
```



Hypothesis Testing

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

```
## Call: rda(formula = dune ~ A1 + Moisture + Management + Use + Manure,
## data = dune.env)
##
##                 Inertia Proportion Rank
## Total          84.1237     1.0000
## Constrained    63.2062     0.7513   12
## Unconstrained 20.9175     0.2487    7
## 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
RDA11
## 22.396 16.208  7.039  4.038  3.760  2.609  2.167  1.803  1.404  0.917
0.582
##  RDA12
##  0.284
##
## Eigenvalues for unconstrained axes:
##   PC1   PC2   PC3   PC4   PC5   PC6   PC7
## 6.627 4.309 3.549 2.546 2.340 0.934 0.612
```

```r
anova(rda_dune, permutations=10)
```

```
## Permutation test for rda under reduced model
## Permutation: free
## Number of permutations: 10
##
## Model: rda(formula = dune ~ A1 + Moisture + Management + Use + Manure,
data = dune.env)
##          Df Variance      F  Pr(>F)
## Model    12   63.206 1.7627 0.09091 .
## Residual  7   20.917
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

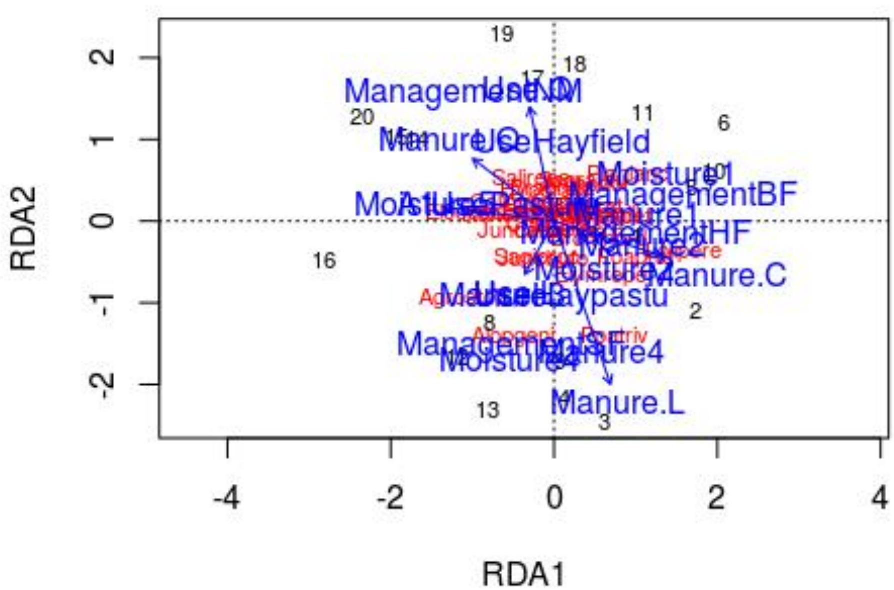
```r
RsquareAdj(rda_dune)
```

```
## $r.squared
## [1] 0.7513483
##
## $adj.r.squared
## [1] 0.3250882
```

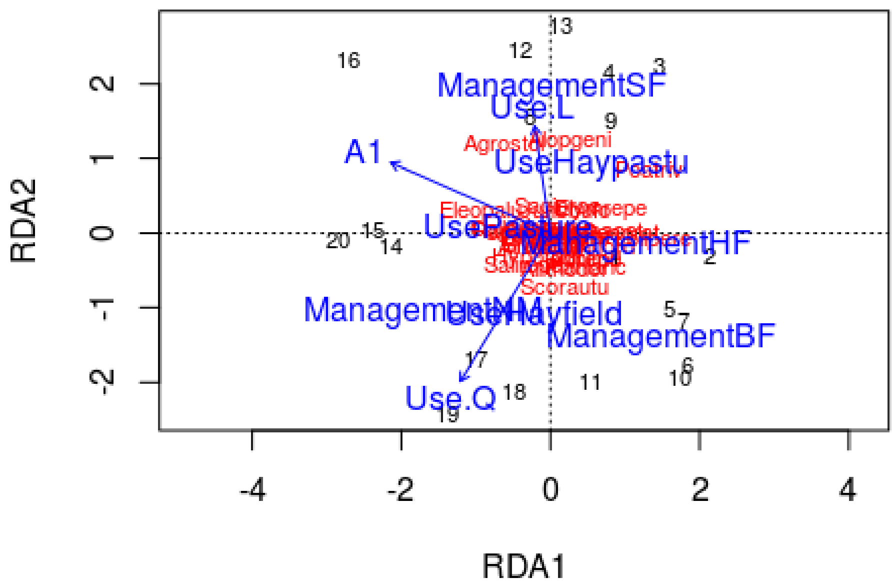
```
anova(rda_dune, by='margin', permutations=10)

## Permutation test for rda under reduced model
## Marginal effects of terms
## Permutation: free
## Number of permutations: 10
##
## Model: rda(formula = dune ~ A1 + Moisture + Management + Use + Manure,
data = dune.env)
##            Df Variance      F Pr(>F)
## A1          1   2.3704 0.7933 0.6364
## Moisture    3  11.9409 1.3320 0.2727
## Management  2   7.1574 1.1976 0.3636
## Use         2   4.9785 0.8330 0.3636
## Manure      3   9.6257 1.0737 0.4545
## Residual    7  20.9175



rda_dune_simple <- update(rda_dune, . ~ . - Moisture - Manure)
anova(rda_dune_simple, rda_dune)

## Permutation tests for rda under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model 1: dune ~ A1 + Management + Use
## Model 2: dune ~ A1 + Moisture + Management + Use + Manure
##   ResDf ResChiSquare Df ChiSquare      F Pr(>F)
## 1    13       43.582
## 2     7       20.917  6    22.665 1.2641  0.173

RsquareAdj(rda_dune_simple)

## $r.squared
## [1] 0.481925
##
## $adj.r.squared
## [1] 0.2428135

step(rda_dune)

## Start:  AIC=85.79
## dune ~ A1 + Moisture + Management + Use + Manure
##
##              Df    AIC
## <none>          85.786
## - A1          1 85.933
## - Use         2 86.056
## - Manure      3 87.357
## - Management  2 87.672
## - Moisture    3 88.818
```

```
## Call: rda(formula = dune ~ A1 + Moisture + Management + Use + Manure,
## data = dune.env)
##
##               Inertia Proportion Rank
## Total         84.1237     1.0000
## Constrained   63.2062     0.7513   12
## Unconstrained 20.9175     0.2487    7
## 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
RDA11
## 22.396 16.208  7.039  4.038  3.760  2.609  2.167  1.803  1.404  0.917
0.582
##  RDA12
##  0.284
##
## Eigenvalues for unconstrained axes:
##   PC1   PC2   PC3   PC4   PC5   PC6   PC7
## 6.627 4.309 3.549 2.546 2.340 0.934 0.612
```

```
plot(rda_dune)
```



```
plot(rda_dune_simple)
```



The anova test inidicated that Moisture and Manure had the highest Df, so excluded them from the rda_dune_simple formula.

The above tests suggests that the more complex model including the variables moisture and manure are strongly supported. Using the r-squared results rda_dune: 33% and rda_dune_simple: 24%

I will run the same test with the varaibales used in the cca model and see if I can generate better r-squared and significance in my anova test.

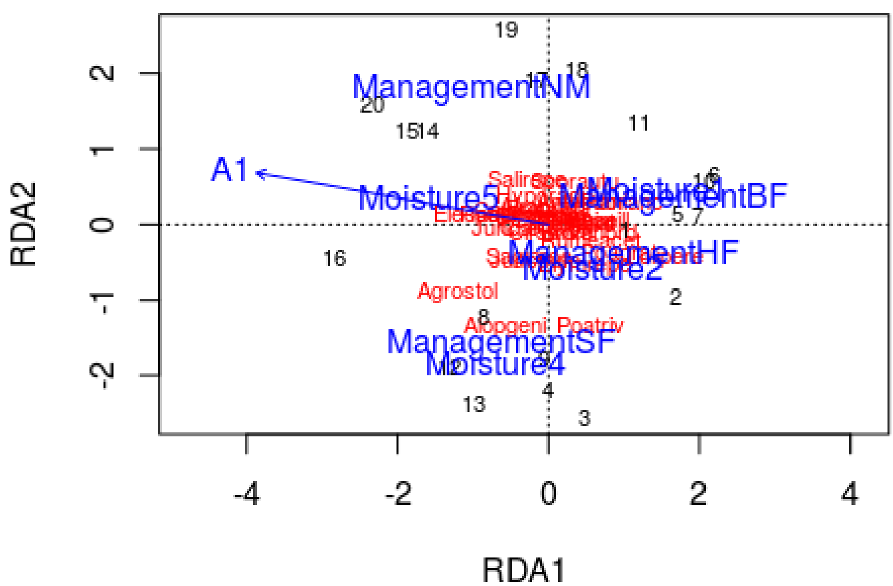
```
rda_dune_2 <- update(rda_dune, . ~ . - Use - Manure)
anova(rda_dune_2)
```

```
## Permutation test for rda under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: rda(formula = dune ~ A1 + Moisture + Management, data = dune.env)
##          Df Variance      F Pr(>F)
## Model     7   48.880 2.3775  0.001 ***
```

```
## Residual 12   35.244
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The anova test suggests that the less complex model excluding the varibales Use and Manure are strongly supported

```r
plot(rda_dune_2)
```



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?

The plots of the two analysis differ in clarity for me, I find the Non-metric Mulidimensional Scaling much easier to understand in developing a conclusion about the data. The Canonical Correspondence Analysis was not my favorite and really intimadted my ability to make an assumoption with the eratic plot. As far as data analyisis I did prefer running cca and rda models. Being able to idnetify the best indicators for the model is helpful, which allows you to make more complex assumptions tham with the NMDS model. Overall I think both models guide you to make the assumotions about the best variables to run in a model, but i prefer the CCA modeling to better understand the data and establish indiactors. I'm sure there is an easier way to plot the data generated, may need to be transformed, but i like the qualitative rersults.