

QUANTITATIVE METHODS: Multivariate Homework Assignment

Lloyd Hill

February 11, 2019

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?

```
library(vegan)
```

```
data(dune)
```

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

```
#create NMDS model for the dune data
```

```
dune_mds <- metaMDS(dune)
```

```
#plot this model overlaid with moisture data
```

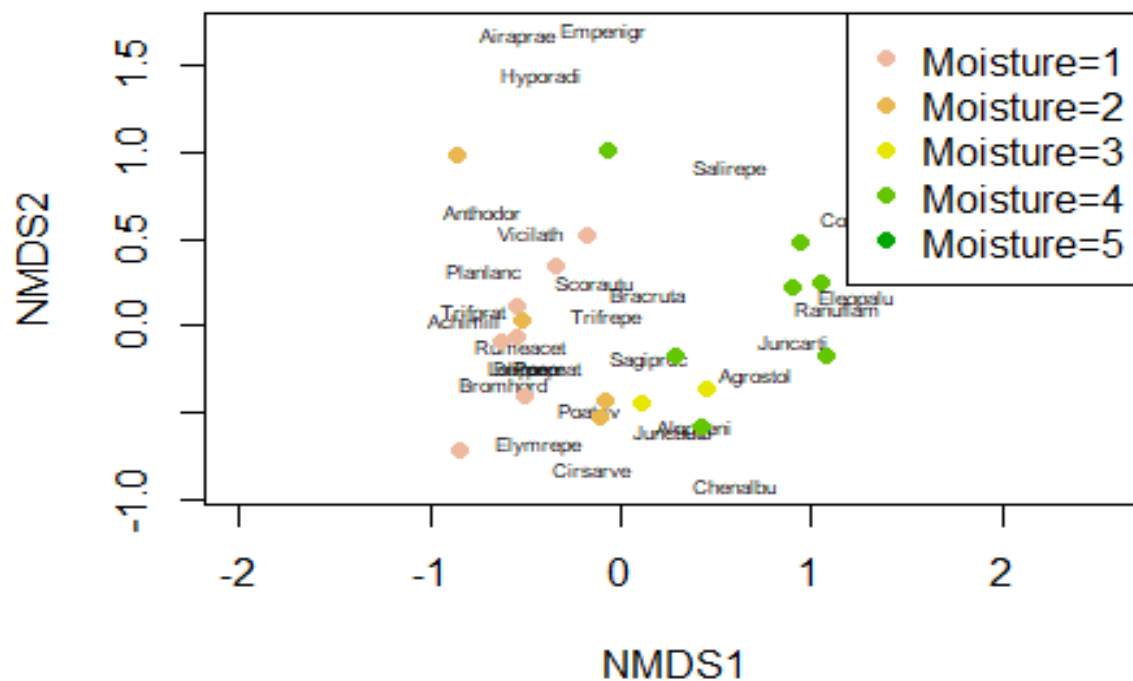
```
plot(dune_mds, type='n')
```

```
text(dune_mds, 'sp', cex=.5)
```

```
color_vect = rev(terrain.colors(6))[-1]
```

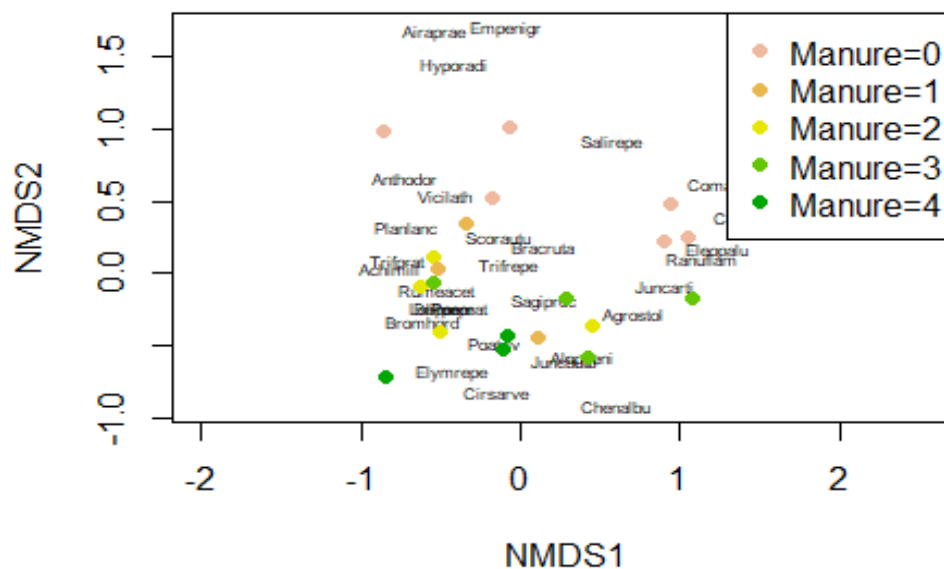
```
points(dune_mds, 'sites', pch=19,  
       col=color_vect[dune.env$Moisture])
```

```
legend('topright', paste("Moisture=", 1:5, sep=""),  
      col=color_vect, pch=19)
```



```
#repeat with other data from dune.env, such as 'Manure'
```

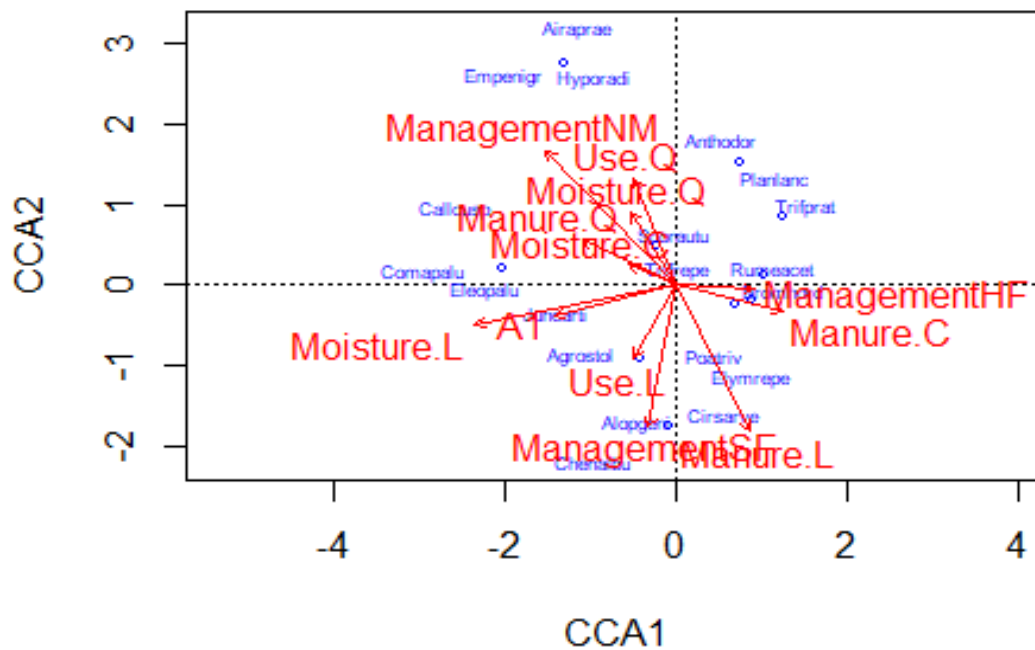
```
plot(dune_mds, type='n')
text(dune_mds, 'sp', cex=.5)
color_vect = rev(terrain.colors(6))[-1]
points(dune_mds, 'sites', pch=19,
       col=color_vect[dune.env$Manure])
legend('topright', paste("Manure=", 0:4, sep=""),
      col=color_vect, pch=19)
```



In my interpretation of these plots, I first looked for any groupings of plant species. For example, there appears to be a trio of species that prefer similar habitat types (airaprae, emperigr, hyporadi). The NDMS plot may be useful in identifying plant communities. To get to the underlying environmental factors responsible for these plant communities, I overlaid this plot with moisture data from 'dune.env.' Moisture data appears to follow the x-axis of this plot (dry to wet) suggesting that moisture has some explanatory value in the distribution of these plants. I repeated this process with the 'manure' variable. I had a harder time detecting a pattern, which suggests that manure may not offer the best explanation for the species found at each site.

- 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.

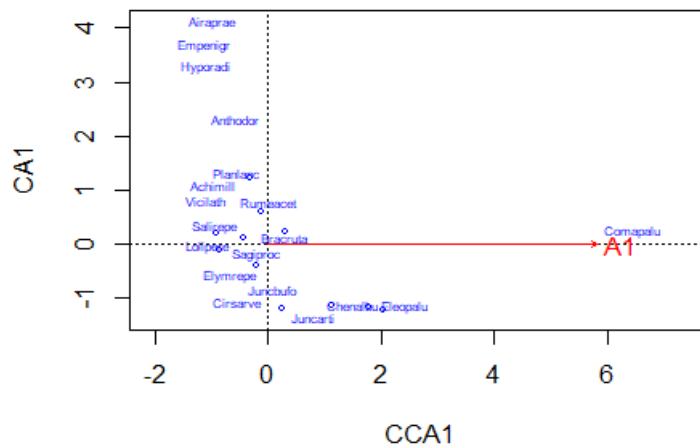
```
#create and plot CCA model for all environmental variables
dune_cca <- cca(dune ~ ., data=dune.env)
plot(dune_cca, type='n', scaling=1)
orditorp(dune_cca, display='sp', cex=0.5, scaling=1, col='blue')
text(dune_cca, display='bp', col='red')
```



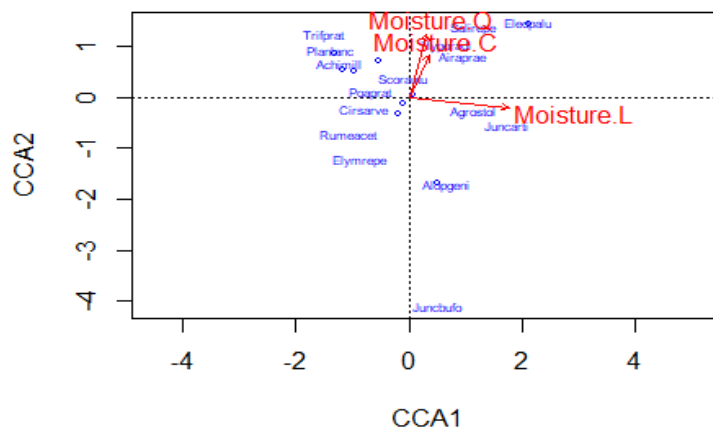
```
#create CCA models for each variable
dune_cca_A1 <- cca(dune ~ A1, data=dune.env)
dune_cca_moist <- cca(dune ~ Moisture, data=dune.env)
dune_cca_manage <- cca(dune ~ Management, data=dune.env)
dune_cca_use <- cca(dune ~ Use, data=dune.env)
dune_cca_manure <- cca(dune ~ Manure, data=dune.env)
```

```
#plot models
```

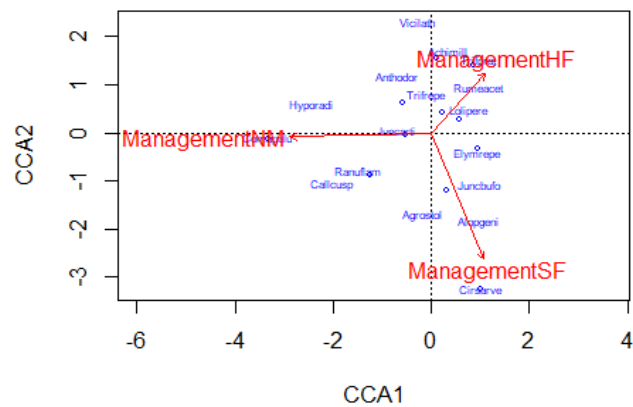
```
plot(dune_cca_A1, type='n', scaling=1)
orditorp(dune_cca_A1, display='sp', cex=0.5, scaling=1, col='blue')
text(dune_cca_A1, display='bp', col='red')
```



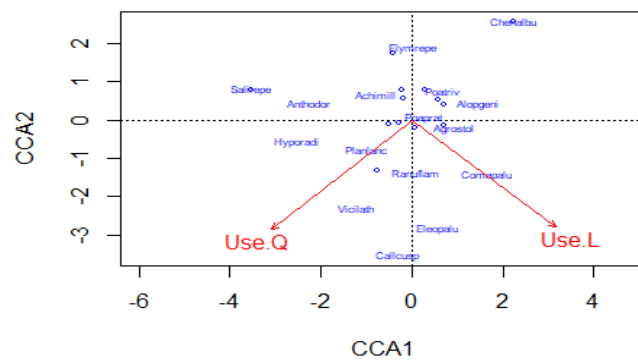
```
plot(dune_cca_moist, type='n', scaling=1)
orditorp(dune_cca_moist, display='sp', cex=0.5, scaling=1, col='blue')
text(dune_cca_moist, display='bp', col='red')
```



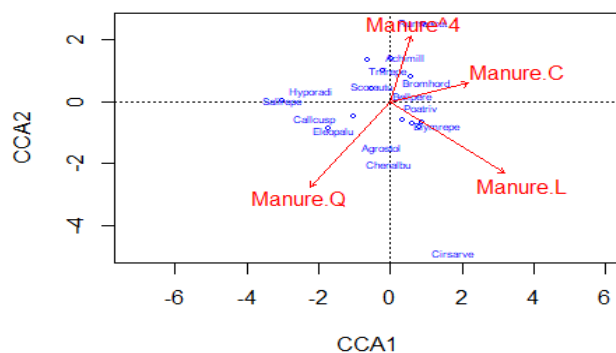
```
plot(dune_cca_manage, type='n', scaling=1)
orditorp(dune_cca_manage, display='sp', cex=0.5, scaling=1, col='blue')
text(dune_cca_manage, display='bp', col='red')
```



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



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



#These plots help visualize the distribution of species with respect to their growing environment. To better explore the influence of environmental variables on species distribution, I used the anova function and vegan's r-squared function.

#First, I ran tests for the CCA model with all environmental variables

```
anova(dune_cca)
```

```
## 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 0.026 *
## Residual  7    0.6121
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(dune_cca, by='margin')
```

```
## 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 0.241
## Moisture  3  0.31587 1.2041 0.223
## Management 2  0.15882 0.9081 0.549
## Use       2  0.13010 0.7439 0.779
## Manure    3  0.25490 0.9717 0.461
## Residual  7  0.61210
```

```
RsquareAdj(dune_cca, 999)
```

```
## $r.squared
## [1] 0.7106267
##
## $adj.r.squared
## [1] 0.2179055
```

#These tests show a significant P-value (0.024) and a good r-squared of 0.71 (however, the adjusted r-squared is much lower). When the anova is broken down into parts (by margin) there is no significant P-value for any individual environmental variable, although 'moisture' had the smallest P-value.

#I ran the same tests on my CCA model with moisture as the only environmental variable.

```
anova(dune_cca_moist)
```

```
## Permutation test for cca under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: cca(formula = dune ~ Moisture, data = dune.env)
##      Df ChiSquare    F Pr(>F)
## Model   3  0.62831 2.2536 0.001 ***
## Residual 16  1.48695
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RsquareAdj(dune_cca_moist, 999)
```

```
## $r.squared
## [1] 0.2970359
##
## $adj.r.squared
## [1] 0.1651024
```

#These tests produced an even smaller p-value than the complete model (more significance), but the adjusted R-squared value shrank as well (less explained variance). I then compared these two models to see if one is better at explaining the species distribution. The resulting p-value of 0.27 suggests that the difference between these two models is not significant. In other words, the inclusion or exclusion of more environmental variables does not significantly alter the outcome of either model.

```
anova(dune_cca_moist, dune_cca)
```

```
## Permutation tests for cca under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model 1: dune ~ Moisture
## Model 2: dune ~ A1 + Moisture + Management + Use + Manure
## ResDf ResChiSquare Df ChiSquare    F Pr(>F)
## 1   16    1.4869
## 2    7    0.6121 9  0.87485 1.1117 0.28
```

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 two approaches I took in this assignment (NMDS vs. CCA) represent two different strategies of data analysis with their own set of pros and cons. The first approach, NMDS, is an example of unconstrained ordination. This technique allowed me to view the data (and any possible structure/trends) all at once. This could be useful for quickly picking out any associations among species. However, it was more difficult to pinpoint individual environmental variables behind the species assemblages. The second technique, CCA, is an example of constrained ordination. In these analyses, I could better visualize the environmental variables' effect on species distribution. One thing I noticed about the CCA plots was when I plotted soil horizon data (A1). Here, one species stood out on one end of the axis. I could see how a fluke in the dataset could offset the rest of the ordination and/or be mistaken for significance.