

Progress Report 2.3.3 Primary Paper

Charles Bond

10/25/2020

Current Status: *Partial Success!* Now I'm trying to separate the data by site and treatment to produce 2 labeled NMDS plots.

Sequence of what I did:

I was able to set up labdsv functions for nmDS: `dsvdist` and `nmDS`. The output of the normal `dist` function and vegan `vegandist` functions generate NA values which cause an error in all the various NMDS functions I've tried to use. I was hoping the `labdsv` functions would fit together. I did find the `bray/curtis` similarity index is available in `dsvdist`, which is the index the paper used, so I am getting warmer. Still, I get an error that appears related to NAs: "Error in `dsvdis(Dia.RelAbun.ext, "bray/curtis", weight = rep(1, ncol(x))`, : NA/NaN/Inf in foreign function call (arg 1)" I believe this means that in argument 1, where the dataset is listed, the function stops because there are NA values? But now I'm confused. Okay. So I used the function `complete.cases()` to look for missing values, and am just now realizing that the last two rows are totally blank. Now I've removed those. YAY, and now the NA error is gone, and the `dsvdist` now gives me a Bray-Curtis dissimilarity matrix! Yay! AND THE NMDS FUNCTION RUNS. Test plot below.

I'm a bit stuck

splitting the data by the two site categories, just need to grind through this for a while really. I split the data and ran 2 separate NMDS analyses, but these plots look different from the originals. Meanwhile, the single test plot looks like the author's two plots overlayed, which is asking me to think this might be way they took the data and made their graphs in excel manually: **how can one separate product of the nmDS function by site?** The product of the NMDS function looks like some special list that I'm not sure how one would plot if joined with other data, or am I wrong? Would it be possible to run the abundance data altogether through the NMDS function and then post-facto join that list with?

So, my next task is to try and join the nmDS function product with the factors by which I will group the data, and then to plot that in ggplot, if possible.

```
## -- Attaching packages ----- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.2      v purrr   0.3.4
## v tibble  3.0.4      v dplyr   1.0.2
## v tidyr   1.1.2      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.0
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag() masks stats::lag()

## # A tibble: 48 x 64
##   Sample Date Site Chamber Treatment TrialID 'Achnanthes min-
##   <fct> <chr> <fct> <dbl> <fct> <chr> <dbl>
## 1 T1 Oct-- FW 6 C M1FW6C 0
## 2 T1 Oct-- FW 8 C M1FW8C 0
## 3 T1 Oct-- FW 10 C M1FW10C 0
## 4 T1 Oct-- FW 12 S M1FW12S 0
## 5 T1 Oct-- FW 14 S M1FW14S 0
## 6 T1 Oct-- FW 16 S M1FW16S 0
## 7 T1 Oct-- OH 5 C M1BW5C 0
## 8 T1 Oct-- OH 7 C M1BW7C 0
## 9 T1 Oct-- OH 9 C M1BW9C 0
## 10 T1 Oct-- OH 11 S M1BW11S 0
## # ... with 38 more rows, and 57 more variables: 'Amphora montana' <dbl>,
## # 'Amphora veneta' <dbl>, 'Amphora coffeaeformis aponina' <dbl>, 'Amphora
## # punctata' <dbl>, 'Amphora sulcata' <dbl>, 'Amphora acutiuscula' <dbl>,
## # 'Brachysira microcephala' <dbl>, 'Caloneis bacillum' <dbl>, 'Caponea
## # caribbea' <dbl>, 'Cyclotella iris' <dbl>, 'Diploneis oblongella' <dbl>,
## # 'Diploneis parma' <dbl>, 'Encyonema evergladianum' <dbl>, 'Encyonema
## # mesianum' <dbl>, 'Encyonema silesiacum elegans' <dbl>, 'Encyonopsis
## # microcephala' <dbl>, 'Entomoneis cf. paludosa' <dbl>, 'Envekadea
## # metezeltinii' <dbl>, 'Envekadea vanlandinghamii' <dbl>, 'Eunotia
## # flexuosa' <dbl>, 'Eunotia naegeli' <dbl>, 'Fragilaria ftsp16' <dbl>,
## # 'Fragilaria minuscula' <dbl>, 'Fragilaria nanana' <dbl>, 'Fragilaria
## # synegrotesca' <dbl>, 'Gomphonema coronatum' <dbl>, 'Gomphonema intricatum
## # vibrio' <dbl>, 'Kobayasiella parasubtilissima' <dbl>, 'Kobayasiella
## # subtilissima' <dbl>, 'Mastogloia pseduosmithii' <dbl>, 'Mastogloia
## # braunii' <dbl>, 'Mastogloia calcarea' <dbl>, 'Mastogloia lanceolata' <dbl>,
## # 'Navicula bulnheimii' <dbl>, 'Navicula cryptotenella' <dbl>, 'Navicula
## # radiosa' <dbl>, 'Navicula salinicola' <dbl>, 'Navicula salinarum' <dbl>,
## # 'Navicula vmsp2' <dbl>, 'Nitzschia acicularis' <dbl>, 'Nitzschia amphibia
## # amphibia' <dbl>, 'Nitzschia ftsp16' <dbl>, 'Nitzschia lacunarum' <dbl>,
## # 'Nitzschia nana' <dbl>, 'Nitzschia palea debilis' <dbl>, 'Nitzschia
## # serpentiraphe' <dbl>, 'Nitzschia vmsp4 (curved)' <dbl>, 'Nitzschia
## # microcephala' <dbl>, 'Planothidium rostratum' <dbl>, 'Pleurosigma
## # salinarum' <dbl>, 'Rhopalodia pacifica' <dbl>, 'Seminavis pusilla' <dbl>,
## # 'Stauroneis phoenicentron' <dbl>, 'Stephanocyclus menghiniana' <dbl>,
## # 'Synedra filiformis exilis' <dbl>, 'Ulnaria delicatissima' <dbl>, 'Unknown
## # vmsp2' <dbl>
```

Preliminary NMDS test

```
##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
## select
```

```
## Loading required package: permute

## Loading required package: lattice

## This is vegan 2.5-6

## Loading required package: mgcv

## Loading required package: nlme

##
## Attaching package: 'nlme'

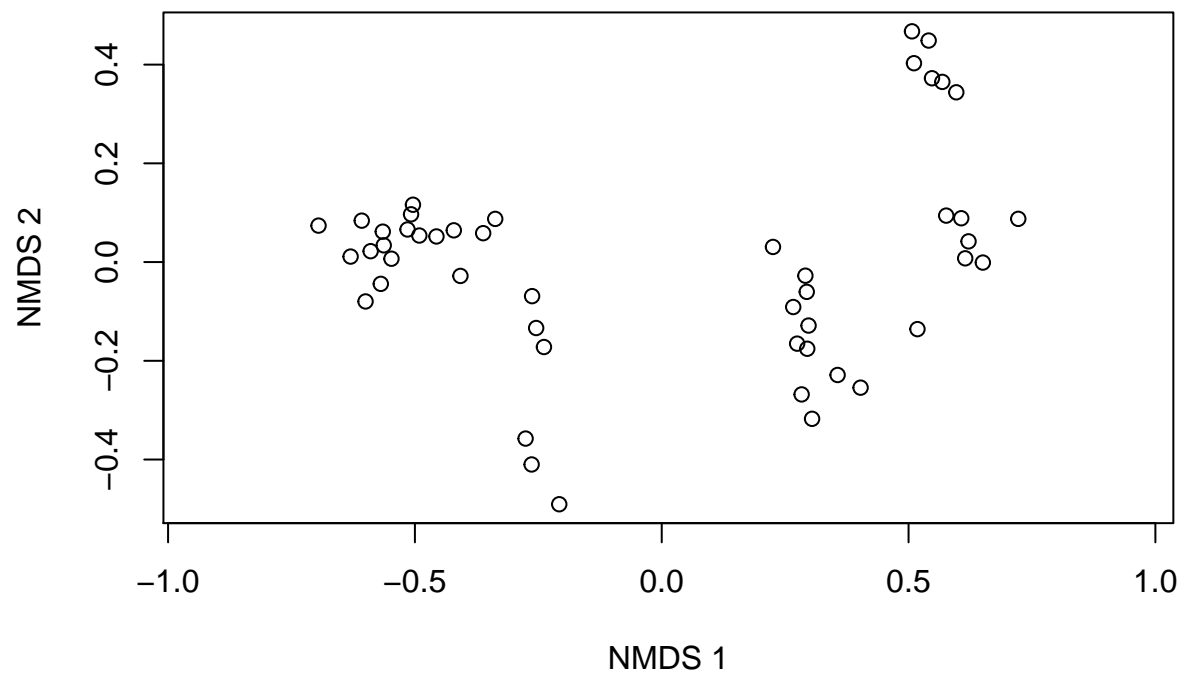
## The following object is masked from 'package:dplyr':
##
##      collapse

## This is mgcv 1.8-33. For overview type 'help("mgcv-package")'.

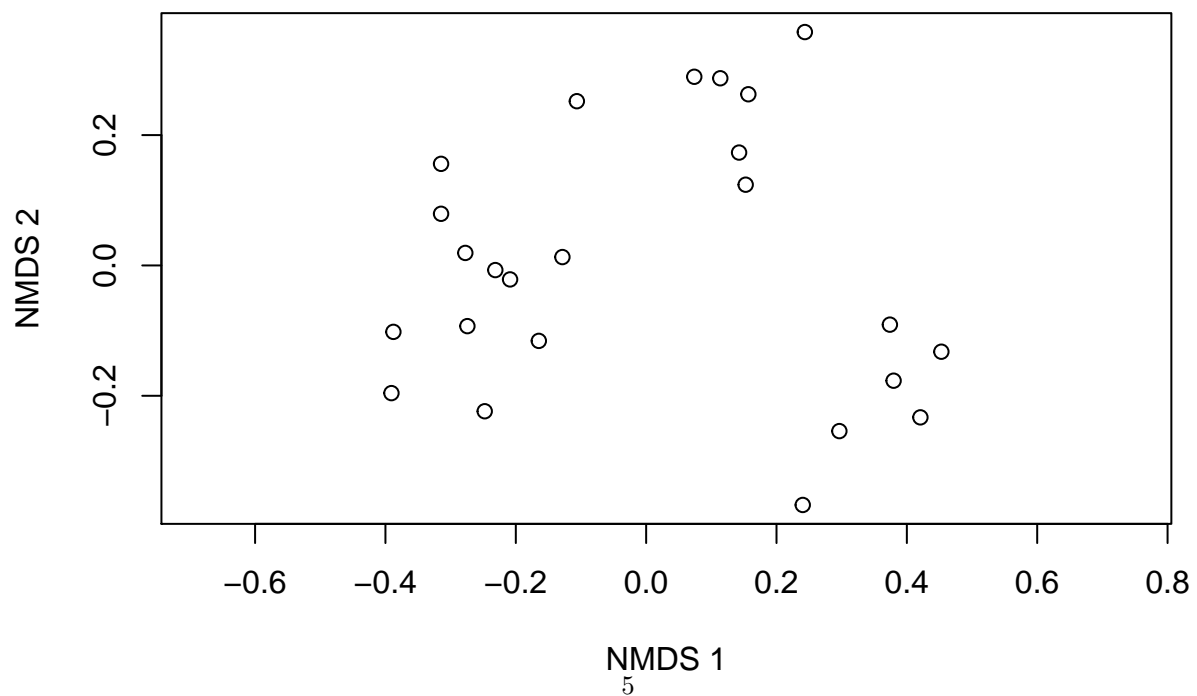
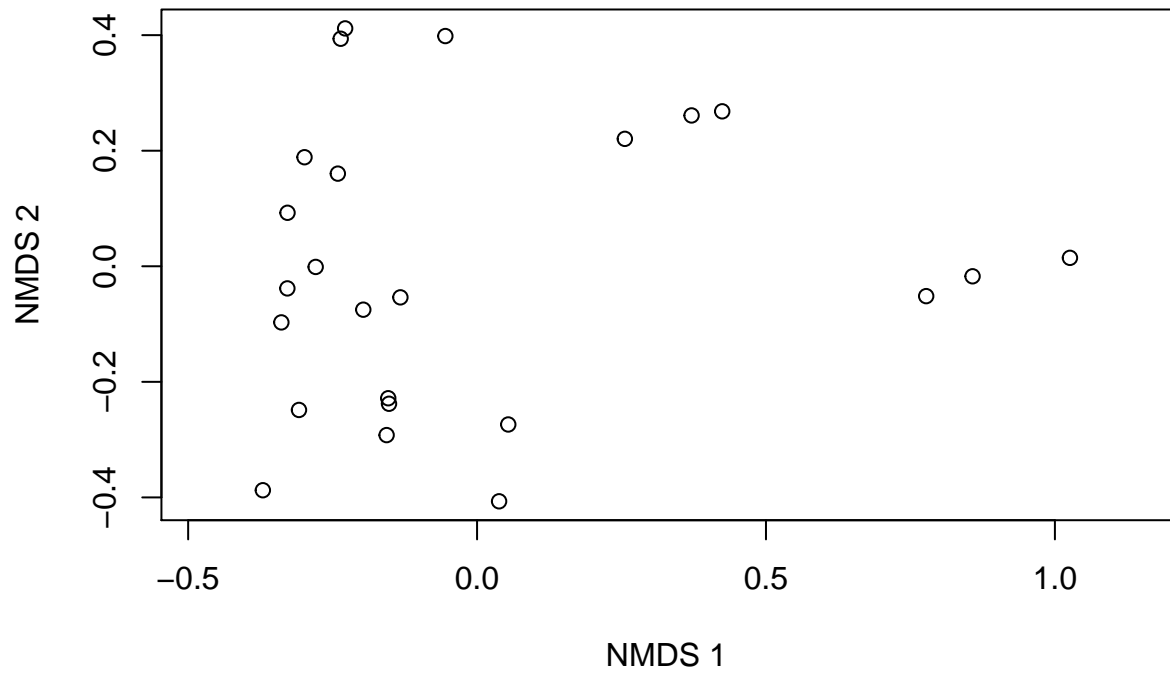
## This is labdsv 2.0-1
## convert existing ordinations with as.dsvord()

##
## Attaching package: 'labdsv'

## The following object is masked from 'package:stats':
##
##      density
```



Seperate NMDS plots for fresh and brackish water sites, evidently not what the paper did.



Background on the Article

For 2 years before I came to Tulane I worked in the Periphyton Lab at FIU, where a Ph.D. student/postdoc in our lab was part of a big mesocosm experiment looking at the responses of diatoms, among other things, to saltwater intrusion in the Florida Everglades (Mazzei et. al. 2018). Some of the wetland/soil microbial folks in our class may also find this interesting, the experiment had soil carbon and bacteria community components published in different papers. (I did not participate in this paper or this kind of data analysis).

Final Report Structure

1. Give the full reference of the paper
2. Name the task that you wanted to accomplish – i.e. which part of the research did you want to reproduce? If you are reproducing a figure, say what kind of figure it is.
3. Were the data available that you could use to complete your task? Did you manage to obtain the data?
 - i. Did you have to do anything “special” to get the data, e.g. email the author, create an account somewhere
4. Did you manage to read the data into R? i.If no, where did you get stuck or give up?
 - ii. If yes, did you have to do anything to the data to read them into R? i.e. use excel and export as csv, modify the data in any way. Describe any code you wrote to wrangle the data
5. Was code available to reproduce the task you set yourself?
 - i. If yes, did you manage to use the code exactly or closely reproduce the task?
 - a. If yes – describe what you did, put the code into github. Did you stop there or did you extend the task in some way? If the latter, go back to step 2 and describe the next task
 - b. If no, did you change the task? Or write new code yourself?
 - ii. If no, did you write (or try to write) the code to reproduce the task
 - a. If yes, did it work out? I should be able to run your code
 - b. If no, explain why not.

References

[1] Mazzei, V., Gaiser, E., Kominoski, E., Wilson, J., Servais, S., Bauman, B., . . . Troxler, F. (2018). Functional and Compositional Responses of Periphyton Mats to Simulated Saltwater Intrusion in the Southern Everglades. *Estuaries and Coasts*, 41(7), 2105-2119.