NMDS to Study Dead Wood Fungi Communities in Parks of New Jersey

Christopher Zambell, Maria Shumskaya. Kean University, NJ, USA

**Instructor notes**

**Learning objectives:**

In this activity, students will:

* Discuss quantitative methods such as ordination used in community ecology with the faculty. Presentation slides are provided to guide this discussion.
* Use basics of statistical software R to perform ordination analysis of datasets.
* Familiarize with the global biodiversity portals such as inaturalist.org or gbif.org and methods of storage and sharing of data.
* Use open biodiversity dataset to perform NMDS ordination analysis.
* Compare and contrast dead wood fungal communities at different types of forests (hardwood vs softwood).

**Activity software**:

Software used was: R / RStudio; packages: dplyr and vegan.

**Activity datasets:**

deadwood.csv

The dataset used in this module is an adapted version from the full occurrence dataset produced by a group of both undergraduate and master’s student researchers and faculty Maria Shumskaya and Christopher Zambell of Kean University, NJ, USA.

The full dataset is available at the Global Biodiversity Information Facility portal:

<https://www.gbif.org/dataset/d1d59f9f-e130-4715-bc9b-8a25e7b62e89>

Please cite the dataset as follows:

Shumskaya M, Zambell C, Mishra S, Bell E, Blue S, Yearwood-Marut J, Marut W, Vindas-Cruz A, Jennings A, Hylton N, Burghardt J (2019). Survey of saproxylic fungi across parks of New Jersey. Kean University. Occurrence dataset https://doi.org/10.15468/ngpb5m accessed via GBIF.org on YYYY-MM-DD.

All observations made by the researchers were recorded using iNaturalist phone app and shared on iNaturalist.org web-site:

<https://www.inaturalist.org/projects/saproxylic-fungi-of-new-jersey>

iNaturalist is a free app that allows everybody to make observations and share it with the world, you can read more about it at inaturalist.org.

**Related documents:**

* **NMDS\_TeachingModule\_KeanU\_saproxylicFungi\_instructorsNotes.docx** – this document. Provides instructor information on the activity.
* **NMDS\_TeachingModule\_KeanU\_saproxylicFungi\_studentHandout.docx** – a handout to distribute to the students to guide them through the activity. Also contains questions that students must answer while working with the R script.
* **NMDS\_TeachingModule\_KeanU\_saproxylicFungi\_Slides.pptx** – presentation that can be used to introduce students into the dead wood fungi ecology and ordination. Contains a simple exercise on calculation of Jaccard similarity index. Slide notes contain necessary information for the instructor.
* **NMDS\_TeachingModule\_KeanU\_saproxylicFungi\_students.R** – incomplete R script to distribute to the students so they can progress through the activity. Requires editing to fully work!
* **NMDS\_TeachingModule\_KeanU\_saproxylicFungi\_instructor.R** – complete and working R script for instructors to serve as control.
* **NMDS\_TeachingModule\_KeanU\_saproxylicFungi\_sampleAnswer.docx** – the document provides sample answers to the questions from the student handout.

**References:**

Bobiec A, Gutowski JM, Zub K, Pawlaczyk P, Laudenslayer WF (2005) The afterlife of a tree. WWF Poland.

Galimanas et al. (2014). Bacterial community composition of chronic periodontitis and novel oral sampling sites for detecting disease indicators. Microbiome, 2:32. (used in slides to provide an example of NMDS ordination in clinical research).

**Image credits**: all images in the slides unless specified are provided by the authors and are shared via iNaturalist.org. License CC BY 4.0.

**Instructor notes**

1. The assignment is developed for one laboratory session of 2.5 hours. If the students are unfamiliar with R, then another laboratory session would be required prior to this activity. During the first lab session we assign an introductory course on R at Datacamp: <https://www.datacamp.com/>

Datacamp is free for educational purposes. <https://www.datacamp.com/groups/education>

Instructors will need to register as an educator and then create a class where so students can enroll. We suggest a course “Introduction to R” and chapters from “Importing Data in R” to be added to this custom course.

1. During the activity, the instructor should introduce the subject using the provided slides. We highly recommend not only to discuss the ordination methods but also calculate Jaccard distance in the class for a simple dataset in the provided slides. This way students will have an idea of what kind of analysis is performed by the software.
2. After the introduction, students proceed with the analysis of the dataset. We suggest to always use rm(list= ls())as a first line in the script to remove all previous work from the R workspace; this might be important when students have tried to import or modify a file previously and made mistakes.
3. Make sure dplyr and vegan packages are installed, and checkmarks are placed against them in the list of packages. We find it a common issue that packages are not installed properly.
4. Have your students import the deadwood.csv file into the R-studio and let them proceed with the script.
5. Emphasize that R is sensitive to spelling, capital letters, and spaces. Students often do not notice the spelling errors.
6. Provide a Word handout to the students so they can type questions answers directly in the handout and submit to you later.

Here is a short recap of the script used in this assignment:

**Major commands of the script:**

rm(list= ls())

deadwood2 <- select(deadwood,-locality,-habitat,-decimalLatitude,-decimalLongitude, -geodeticDatum, -eventDate, -basisOfRecord, -countryCode, -stateProvince, -kingdom, -recordedBy)

View(deadwood2)

wood.dis<-vegdist(deadwood2, method=”jaccard”)

wood.dis

wood.mds<- metaMDS(deadwood2, try=1000, distance="jaccard",k=2)

wood.mds

stressplot(wood.mds)

pl<-ordiplot(wood.mds)

ordiplot(wood.mds, type="points", cex=1)

ordihull(pl,deadwood$habitat,col=c("green","brown"))

text(pl,what="sites", deadwood$locality,cex=1)

identify(pl, "species", col="purple" )

**Explanation of the script:**

1. Using command select **{it belongs to dplyr package}**, remove any columns that do not contain numerical species data from the imported file deadwood, so what is left is saved as a new vector with a new name (deadwood2). This way you will retain the original dataset (deadwood) with its variables and labels for later.

deadwood2 <- select(deadwood,-locality,-habitat,-decimalLatitude,-decimalLongitude, -geodeticDatum, -eventDate, -basisOfRecord, -countryCode, -stateProvince, -kingdom, -recordedBy)

You can view the result using command View(deadwood2).

1. Using command vegdist **{it belongs to vegan package}** create a distance matrix using Jaccard index.

wood.dis<-vegdist(deadwood2, method=”jaccard”)

wood.dis

1. Using command metaMDS **{it belongs to vegan package},** perform NMDS analysis. Utilize 1000 iterations and choose appropriate distance metric (as instructed). Save the result as a vector wood.mds.

wood.mds<- metaMDS(deadwood2, try=1000, distance="jaccard",k=2)

1. Display the object you just created to see the stress value and other variables:

wood.mds

1. To check the goodness of fit using a Shepard plot, use command stressplot. If the fit is good, the dots should follow the curve.

stressplot(wood.mds)

1. Create the ordination plot (ordiplot command from vegan package) object. We call it pl (short for the plot).

pl<-ordiplot(wood.mds)

1. Create the plot to view, again using ordiplot command and adjust size of points if needed (cex).

ordiplot(wood.mds, type="points", cex=1)

1. Add hulls around appropriate predefined categories (such as pine barrens vs. hardwood, localities, or habitats). Use command ordihull to do this.

ordihull(pl,deadwood$habitat,col=c("green","brown"))

**Variations:**
Alternatively, you could name the habitat vector in a separate step:

habitat<-deadwood$habitat

ordihull(wood.mds, habitat, col =c("green","brown"))

Or create a vector from scratch:

habitat<- c("H","H","H","H","H","H","H","H","P","P","P","P","P","P","P","P")

ordihull(pl,habitat,col=c("green","brown"))

1. Now add the text labeling, in a similar way to how hulls were added. Text could be added before or after adding the hulls. Example as follows, labeling each of the black circles by the name of the samples:

text(pl,what="sites", deadwood$locality,cex=1)

1. Finally, you can add species labels using identify **{it belongs to the graphics package – automatically loaded with R}**

identify(pl, "species", col="purple")

**Assessment:**

In order to assess the skills students learned after this class, we check their answers to the questions in the handout and assign a grade based on the amount of correct answers. For the cumulative final exams we suggest to offer a slightly different dataset (e.g. remove a couple of columns and some localities) and ask students to perform the analysis. Beginner students can be asked just to make changes in the existing script to color the species names red rather than purple, or change cex in the plot and provide screenshots to the instructors.

**Sample exam question:**  Download the dataset of ground dwelling mushrooms sampled from four New Jersey locations. Perform NMDS.

1. Tell the stress value.
2. Plot the sites, and label each point using the “Sample” column entries.
3. Hull the sites in different colors by Hardwood vs. Pine.
4. Identify/Label six of the species in the plot.
5. Export the plot.
6. Tell whether the sites seemed to cluster by “Hardwood” vs. “Pine” categories or not.
7. Explain in a short paragraph what can be a reason for that.