**Instructions for Model Selection:**

We have two overarching study aims:

**How does seasonal variability influence soil and litter invertebrate communities?**

1. To measure this, we are using the following model structure and mixed models, to do model selection with AIC to determine the best model (with most explanatory variables):
2. Write out the following model for your response variable – note that one of these models is for macrofauna and one is for nematodes.
3. Calculate the AIC score for this model using the *AIC(model)* function. Make a note of this value.
4. Write out the model again but excluding the response variable with the highest p-value. Be sure to consider the interactions when doing this (ex. If the highest p-value is part of an interaction with the lowest p-value, leave it in for now, and go to the next variable) and make sure to exclude p-values of NA prior to numeric values.
5. Run this model and calculate the AIC score.
6. Repeat until all only significant variables remain in the model.
7. Compare all the AIC scores and choose the model with the lowest AIC score. This is the best model.
8. Keep the full (initial) model and the best model in the script. We will use the best model for our interpretation. (Note: If there are multiple models with the same AIC score, then choose the model with fewer variables as the best model)
9. Run normality tests in the model.
10. We are also looking at the community structure (community composition) using NMDS (non-metric multidimensional scaling) and *adonis* (permanova) methods:
11. Conduct an NMDS test using a matrix using relative abundance measures for either the nematodes or macrofauna.
12. Conduct a permanova test with the *adonis2()* function in R using the following equation:

In R, this would look like:

1. Choose significant variables to visualize using ellipses in your NMDS plot.

**What influences the community composition within seasons?**

Repeat the instructions above for nematode and macrofauna total abundance, nematode trophic groups, and community matrices by each season. Use the following model structures:

For mixed models:

For community ordinations:

To measure communities, we have measured the following **response variables**:

Nematode Community:

* Total Abundance- a sum of all nematodes, calculated from trophic groups below, as well as *fungal.feeders.aph*, *plant.parasites*, and *unknown* columns in dataset(*total.nematodes*)
* Simpson Diversity Index- calculated using the trophic groups listed below (*simpson*)
* Trophic groups:
  + Bacterial Feeders (*bacterial.feeders*)
  + Fungal Feeders (*fungal.feeders*)
  + Plant Parasites (*plant.parasites.aph*)
  + Omnivorous (*omnivorous*)
  + Predators (*predators*)
* Relative abundance – this is calculated for the matrix of trophic groups using the formula:
  + This is measured for the entire matrix, with total abundance variables being measured by sample, so each cell has a relative abundance value from 0-100.

Macrofauna:

* Total Abundance- a sum of all macrofauna groups identified, including unknowns (*totalAbundance*)
* Shannon Diversity Index – calculated using all macrofauna groups, excluding unknown *(shannon*)
* Relative abundance – this is calculated for the matrix of trophic groups using the formula:
  + This is measured for the entire matrix, with total abundance variables being measured by sample, so each cell has a relative abundance value from 0-100.
  + The group *unknown* is excluded from this matrix but is part of the total abundance calculations.

We are retaining (testing) the following **explanatory variables**:

* Season (*season*)
* Topographic Position (*landscape.position*)
* Mycorrhizal Fungi Type (*mycorrhizal.fungi.type*)
* Sample Type for macrofauna only (*sample.type*)
* Average weekly precipitation (*PRECIP*)
* Average weekly ambient temperature (*AIRTEMP*)
* Soil Moisture, calculated from soil samples taken at the same time as invertebrate sampling (*soil.moisture*)

Soil temperature was calculated from data from the weather station, but was excluded due to correlation with ambient temperature, and the variability of soil temperature at smaller scales compared to ambient temperature. Measurements were taken from the weather station on the golf course near sampling locations – while the air temperature may not vary drastically at this scale, soil temperature likely varies between the forest and golf course.

*Document History*

*Created by GL on 9/13/2024*