Supporting Analyses for Lilaeopsis Drought Ecology

Jacob Malcom, Anna Malcom, and William Radke

# Overview

This document captures analyses and observations that are not central to the results of the paper, *Water (or the Lack Thereof), Management, and Conservation of an Endangered Desert Wetland Obligate,* Lilaeopsis schaffneriana subsp. recurva. That is, during data processing we made observations that are informative but which may be distracting if included in the primary code.

All of the data was read from Excel files using readxl and prepped for analysis in prep\_data.R. Code for replicating the results (and tables and figures) of the main paper is found in main\_analyses.R.

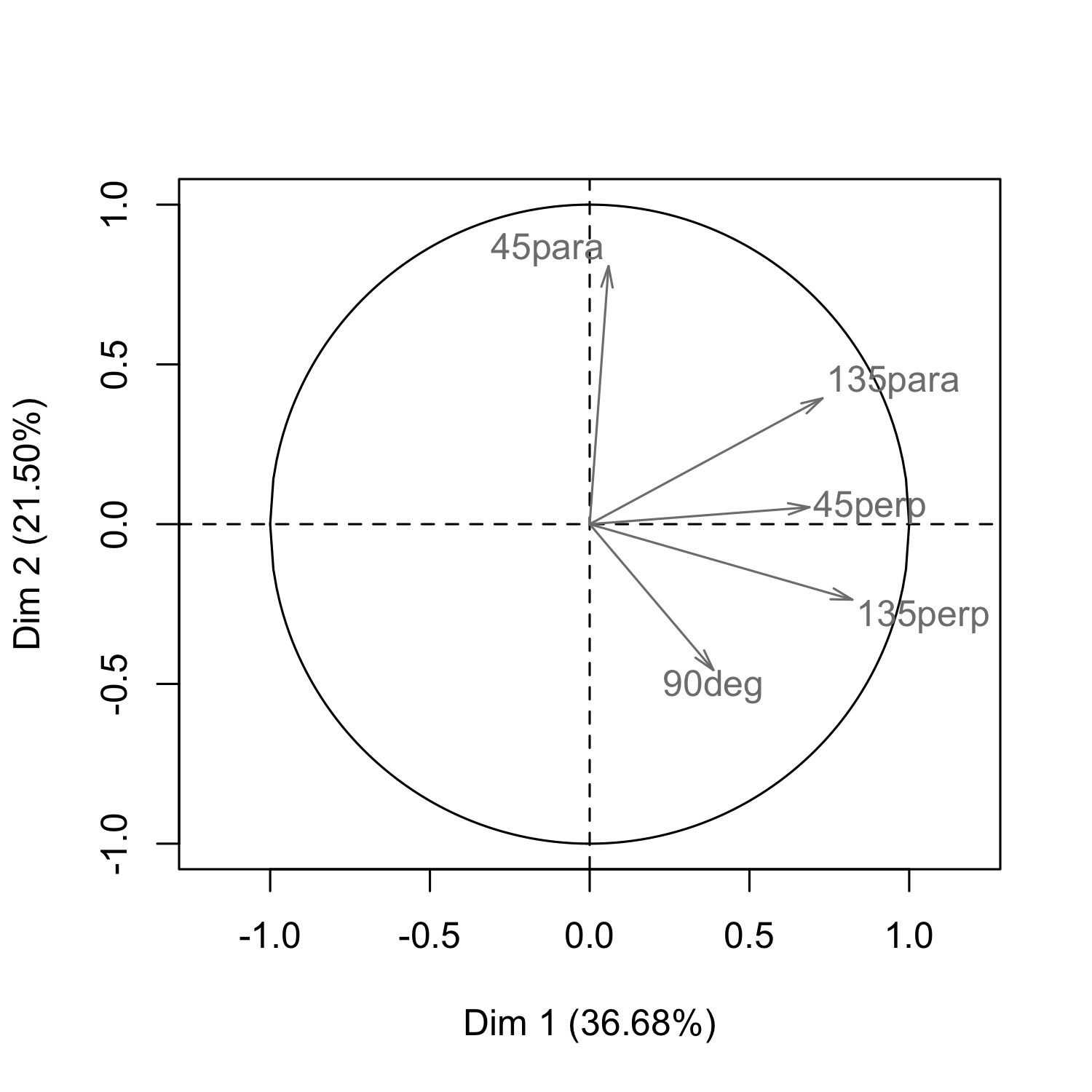
library(dplyr)  
library(FactoMineR)  
library(ggplot2)  
library(ggthemes)  
library(pscl)  
  
load("../data/field\_dat.Rdata")  
  
knitr::opts\_chunk$set(dpi = 300)

# Lilaeopsis presence, absence, and abundance

## Forest canopy PCA

We measured forest canopy openness in five places above or around each sample location along Leslie Creek. We performed a PCA to evaluate the extent to which forest canopy variance was shared.

|  |  |  |  |
| --- | --- | --- | --- |
|  | eigenvalue | percentage of variance | cumulative percentage of variance |
| comp 1 | 1.8339529 | 36.679057 | 36.67906 |
| comp 2 | 1.0750538 | 21.501075 | 58.18013 |
| comp 3 | 1.0095273 | 20.190546 | 78.37068 |
| comp 4 | 0.7089513 | 14.179027 | 92.54971 |
| comp 5 | 0.3725147 | 7.450295 | 100.00000 |

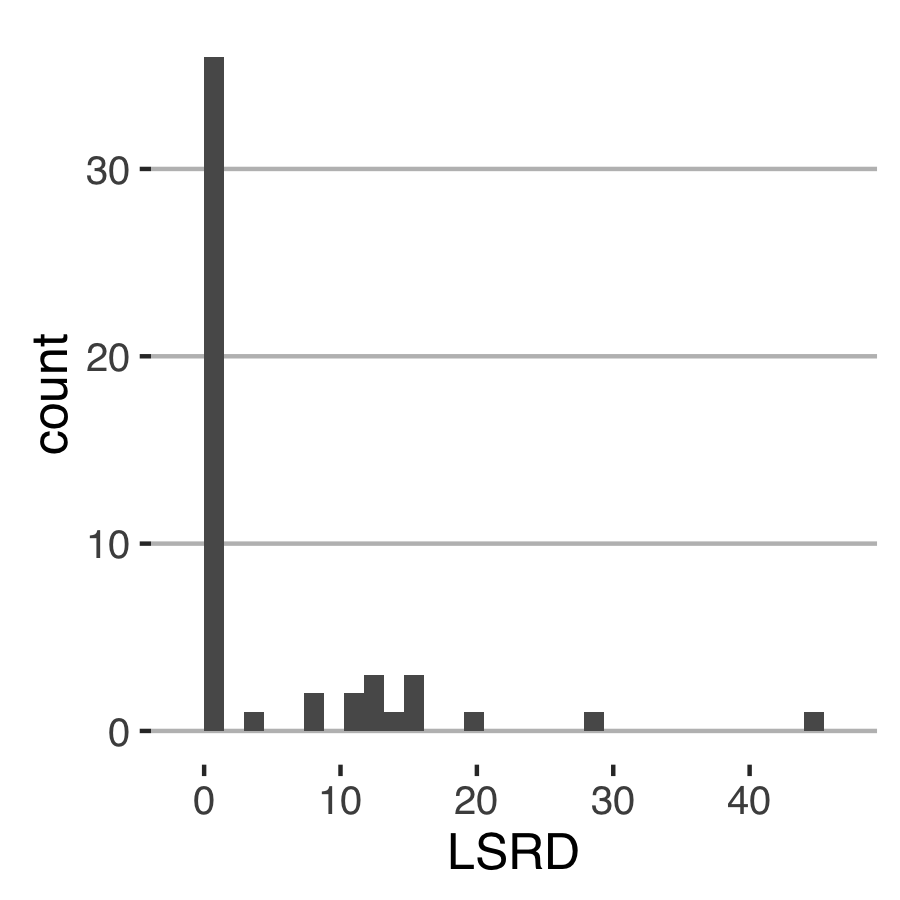


**Biplot of the five riparian canopy measurements.** We used PCs 1 and 2 in some models because they account for 58% of the variance.

The loadings on the first 1-2 PCs are not very strong, but they are sufficient for removing three of five variables.

## Distribution of *Lilaeopsis* densities

We need to see the distribution of densities to choose the best way to model the relationship to the predictor variables.



**Distribution of *Lilaeopsis* densities.** Many sampled points have no *Lilaeopsis*.

Given this distribution we could either (a) model the presence/absence separately, e.g., using a binomial model and a count [neg. binomial or Poisson] model; or (b) model as a zero-inflated count mixture model. We chose the latter option, with a caveat.

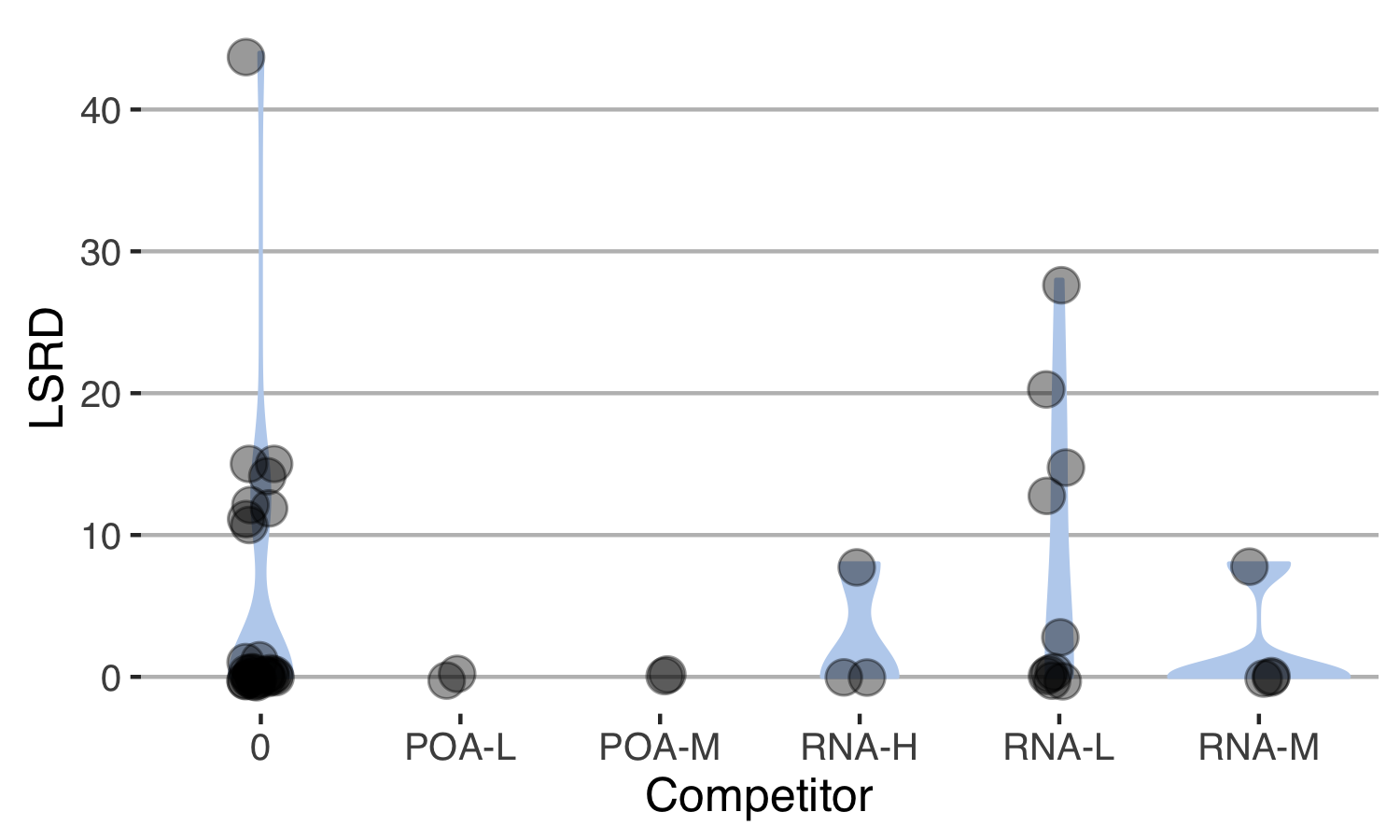
## *Lilaeopsis* density and competitors

The zero-inflated model caveat is that keeping the plant competitors in the zero-inflated model did not work:

model <- zeroinfl(LSRD ~ Competitor, data = field\_dat, dist = "negbin", EM = TRUE)

## Error in solve.default(as.matrix(fit$hessian)): system is computationally singular: reciprocal condition number = 1.50489e-26

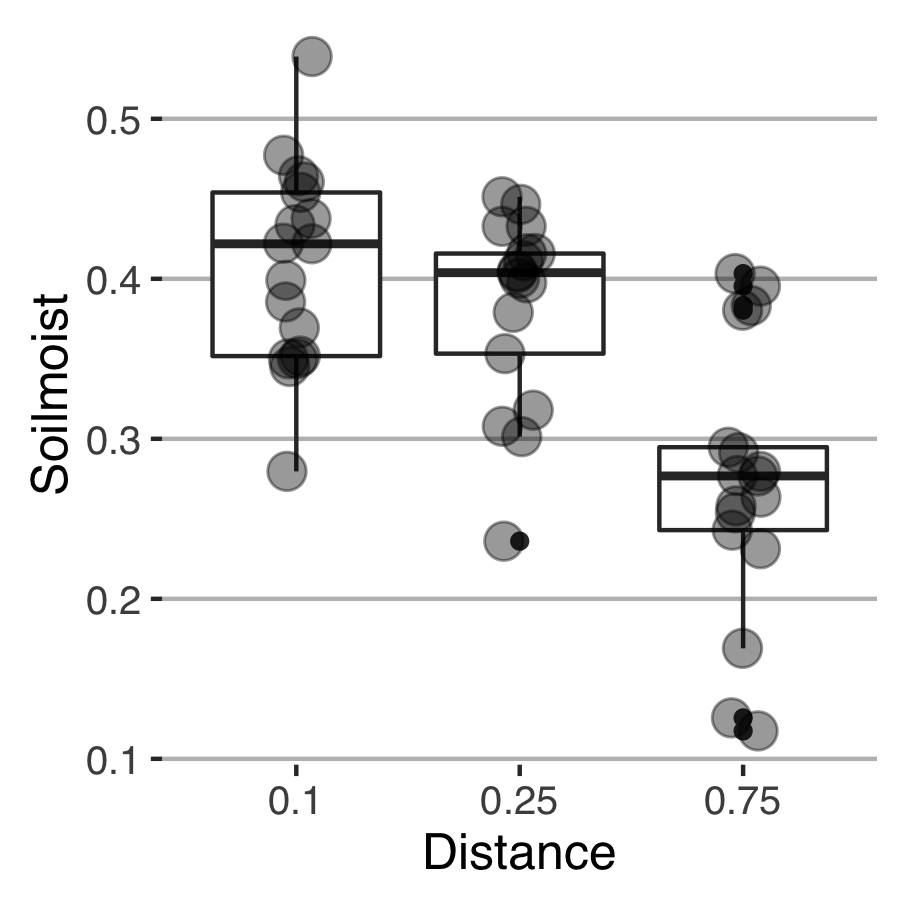
We next checked if *Lilaeopsis* density covaried with competitors; if not, then we can safely drop competitor from the models.



**The distribution of *Lilaeopsis* densities did not covary strongly with any of the competitor species.** 'POA' = grasses (Poaceae), 'RNA' = *Rorippa nasturtium-aquaticum*; H = high, M = medium, L = low competitor density.

# Soil moisture as a function of distance from the creek

We expect that soil moisture will be higher closer to the edge of the creek, so the distance category and soil moisture may be redundant.



**As expected, soil moisture drops as the distance from the edge of the creek increases.** However, we considered models with both soil moisture and distance class in the predictor variable set because the moisture of the near- and mid-distance classes overlaps extensively.