Critique of Laufenberg Paper

Stat 512: Rough Draft of Final Project

Steve Huysman & Parker Levinson

Friday, April 14, 2023

Contents

0.1 Introduction

- White bark pine is very threatened in the GYE.
- it's a keystone and foundational species that helps with the healthy ecosystem function in the park.
- recently, about half of the white bark pine in the GYE has suffered from mortality, decimating this important population
  - higher temps
  - drier conditions
  - insect pathogens which leads to disease
- furthermore, white bark pine is the most vulnerable tree species to climate change

  - this is because they live in high alpine and high topographic areas?

- one way to help is to replant white bark
- INSERT SOMETHING ABOUT ALL THE PHYSICAL REQUIREMENTS THEY NEED
- But those physical requirements haven't been connected with replanting efforts
- RQ: How does climate and competition related to survival of white bark plants planted? They aim to do this by looking at survival and growth rate compared with water balance and competition at the same tree level across a large environmental gradient where seedlings have been planted since 1990.
- Researchers had two different research questions, but we decided to just look at how survival was related to water balance and competition at the same tree level, to simplify the problem a little bit.

0.2 Methods

0.2.1 Field Methods/Study design

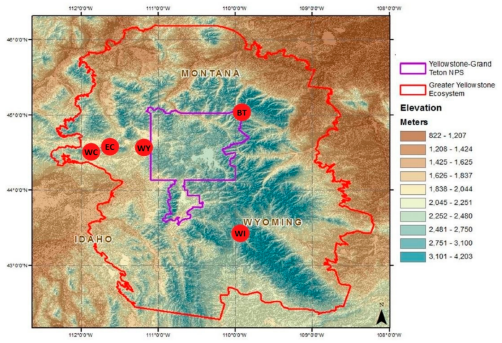

Figure 1: a figure

- Over the past 40 years, US Forest Service and Park service has planted more than 1500 acres of white bark park, including in 5 study areas around the GYE (in FIGURE 1). Each study area has between 5 and 8 planting sites (WHAT DOES THIS MEAN?)

- This study used a hierarchical study design by studying 5 study areas with a combination of 29 planting sites across them and thousands of white bark pine seedlings per planting site.

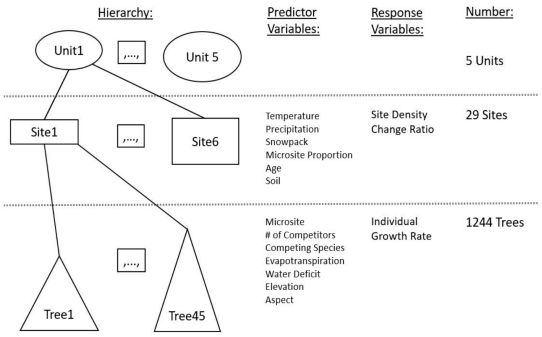

Figure 2. Hierarchical study design, predictor variables, and sample sizes.

Figure 2: another visualization

- Sampling took place from May 2018 to October 2018. A grid cell matrix of 10m x 10m was overlaid on the study site. A random starting point was decided and then every 20th grid cell from that was sampled, equating to sampling between 2-15% of each site. Each seedling within that grid cell was digitally tagged, and Survey123 was used to collect field data (FIGURE 2).

- Researchers were most interested in the annual growth rate. Seedlings were too small to core to measure growth rings, so height was used as a proxy for growth rate, specifically the change in height from the year of study to the relative planted height. This was divided by the years since planting minus 2.5 years to account for the period of time when seedlings sequester carbon instead of focusing on growth.

- The theoretical model is as follows:

$\mu\{growth_{rate}|AET + PET + PPT + T + Micro + CompNnumber + PICO + PIEN + ABLA + random(Unit)\} =$ XXXXX

- Explanatory variables are as follows:
  - microsite - This was a binary variable. 1 if there was a rock or other topographical feature that changed the environmental conditions where the seedling lived.
  - number of competitors
  - competing species
  - evapotranspiration
  - water deficit
  - elevation
  - aspect - The direction that the slope is facing was measured and binned ($0\circ$, $45\circ$, $90\circ$, $135\circ$, $180\circ$, $225\circ$, $270\circ$, $315\circ$, or flat)
  - slope - The steepness of the slope was measured and binned ($0\circ$, $5\circ$, $10\circ$, $15\circ$, $20\circ$, or $25\circ$).
  - WHAT OTHER EXPLANATORY VARIABLES GO HERE?
- We selected the variables of interest and looked for correlation with a pairs plot, which can be seen in Figure **??**. The highest correlation value is -0.681 between annual temperature and aet. The next highest was between aet and group. All other values were below 0.6 and were not of concern. first looked for correlation
- We can visualize the growth rate of each study site (Figure **??**)
- Sample sizes and missing data provided/discussed

0.2.2 Statistical Procedures Used

- Mixed effects models were used to study what covariates were correlated with white bark pine performance. Fixed effects, interactions, linear and polynomial functions,

and different variance structures were explored. A corrected AIC was used to compare models. Colinearity was explored and variables removed if they were more than 0.6 colinear. Authors attempted to find the most parsimonious set of water balance variables while still incorporating the biologically relevant variables and incoporating the variables of questions (including number of competitors, age of planting, and microsite status). Everything was done use R statistical software.

- they did something with step selection, but how do you do that with mixed models?
- Data were provided by Laufenberg, but the process of cleaning and structuring the data were not explicit in the paper. As such, we had to experiment as to cleaning methods in order to recreate numbers and graphs. This includes
  - creating AET which is the total evapotranspiration. To get this, we multiplied the average evapotranspiration by 7 for the 7 months of the growing season.
  - The same was done for PET or potential evapotranspiration.
  - P or the total annual precipitation: "annual precipation" was actually the monthly average precipatation so that needed to be multiplied by 12 to get the total annual precipitation
  - growth rate was normalized via log 10 transformation

0.3 Results/Summary of Statistical Findings

- The final model, decided with drop1, is aetˆ3, petˆ3 + comp_numberˆ3 + PICO

0.4 Scope of Inference

- Random sampling = inferences to large population
- No random treatment (because nature) = correlation/association but not causation.

0.5 Critique

- no clear model selection criteria or process
- difficult data cleaning and processing

0.6 Group Work Statement

0.7 References

0.8 Appendix

- Must include a compiled RMarkdown with all of our results