Critique of Laufenberg Paper
Stat 512: Rough Draft of Final Project

Steve Huysman & Parker Levinson

Friday, April 14, 2023

Contents

0.1 Introduction

Whitebark pine (*Pinus albicaulis* Engelm.) is a conifer tree native to the mountains of the western United States and Canada. It inhabits subalpine areas where it can be found growing up to the tree line, at a higher elevation than other tree species found with it.

Whitebark pine (WBP) is an early successional species that is often the first to establish after disturbance such as wildfire. WPB is a keystone species of subalpine environments where it plays important ecological roles such as providing food for wildlife such as Clark's Nutcracker and the threatened Grizzly Bear.

Due to threats from climate change, mountain pine beetle, and the invasive white pine blister rust, Whitebark pine has undergone a rapid and widespread decline. It was recently estimated that over half of all standing WBP in the United States are dead. This decline has

lead to its recent listing as Threatened under the Endangered Species Act. Future climate projections indicate further deterioration of WBP's habitat.

Strategies to conserve this species involve planting WBP seedlings for restoration of high-elevation forests. Successful plantings in the face of climate change require an understanding of the relationship between climate and seedling establishment and growth in this species. Competition from other tree species also plays a role in seedling establishment and was investigated here.

This study investigated USFS plantings of WBP seedlings in the Greater Yellowstone Ecosystem (GYE) to answer two research questions: 1) What is the relationship between climate/competition and WBP seedling establishment, measured by individual growth rate? 2) What is the relationship between climate/competition and WBP seedling survival, measured by density change?

We are interested here in the first question relating to seedling growth rate.

0.2 Methods

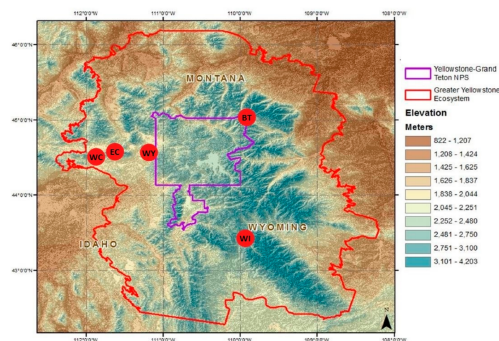### 0.2.1 Field Methods/Study design

Figure 1: Map of planting units included in study in the Greater Yellowstone Ecosystem

Over the past 40 years, the US Forest Service and National Park Service has planted more than 1500 acres of whitebark pine in the GYE. This study investigated five planting units (figure 1) that each contained between two and eight planting sites.

This study used a hierarchical sampling design including 5 planting units, with a total of 29 planting sites across units, and thousands of white bark pine seedlings per planting site. (See figure 2 for model design.)
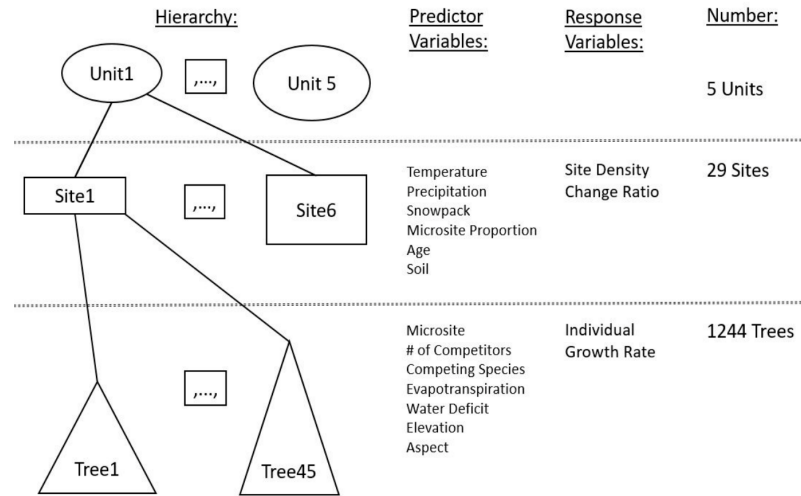

Figure 2. Hierarchical study design, predictor variables, and sample sizes.

Figure 2: Visualization of study design

Sampling took place from May 2018 to October 2018. A grid cell matrix of 10m x 10m was overlaid on the study site or unit. 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.

Researchers were most interested in the annual growth rate. Seedlings were too small to measure growth rings via coring, so height was used as a proxy for growth rate. Specifically, growth rate was calculated as the change in height between the study year (2018) to the relative planted height when the seedling was first planted. This was divided by the number of years since planting minus 2.5 years to account for the period of time when seedlings sequester carbon instead of focusing on their own growth.

Initially, we can look at a box plot of the raw data to see if growth rates may vary between planting units. Figure **??** shows that there might be slight differences between the planting

units, especially for the West Yellowstone (WY) unit. However, any apparent differences may be a result of a difference in the age of the seedlings.

Researchers collected a variety of environmental variables, but for simplification purposes, we chose the most ecologically interesting variables, which are documented below.

- Explanatory variables are as follows:

    - AET - Mean actual evapotranspiration (mm) during growing season (April–October)

    - PET - Mean potential evapotranspiration (mm) during growing season (April–October)

    - PPT - Mean annual precipitation (mm)

    - Micro - This was a binary variable indicating presence of favorable microsite conditions. 1 if there was a rock or other topographical feature that changed the environmental conditions where the seedling lived.

    - Comp_number - number of competitors within a 3.59m radius

    - PICO - Presence of *Pinus contorta* within 3.59m radius of WBP. 1 if PICO is present, 0 otherwise.

    - PIEN - Presence of *Pinus engelmannii* within 3.59m radius of WBP. 1 if PIEN is present, 0 otherwise.

    - ABLA - Presence of *Abies lasiocarpa* within 3.59m radius of WBP. 1 if ABLA is present, 0 otherwise.

Using the variables of interest, we 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.

There are only 5 units and only 29 sites within each of those units, resulting in a somewhat small sample size. However, with over 1000 seedlings within each site/unit combination, there

is a of replication within each unit.

The theoretical model is as follows:

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

0.2.2 Statistical Procedures Used

Mixed effects models were used to study what covariates were correlated with white bark pine performance. Random effects were always used for the unit. Interactions, linear and polynomial functions 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

5

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