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

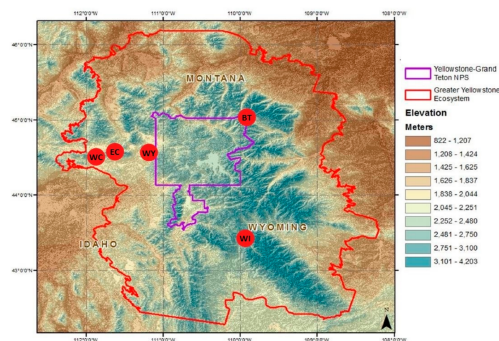### 0.2.1 Field Methods/Study design

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.)
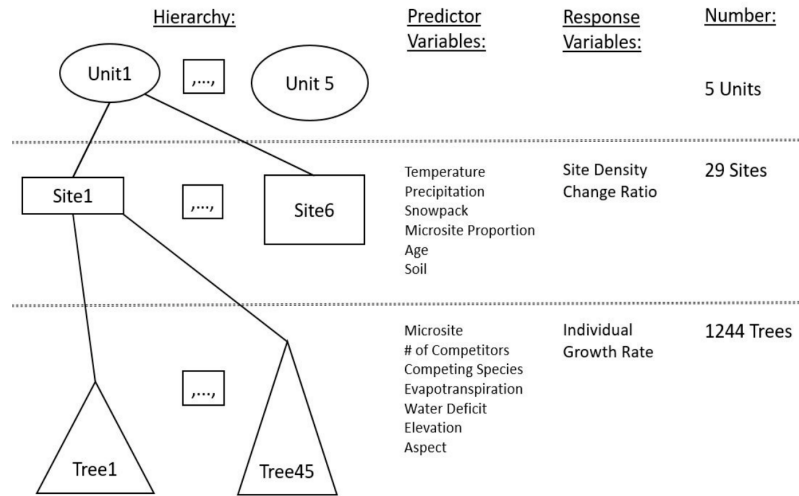

Figure 2. Hierarchical study design, predictor variables, and sample sizes.

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.

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.

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 3 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. This plot is equivalent to Figure 4 in the original paper.
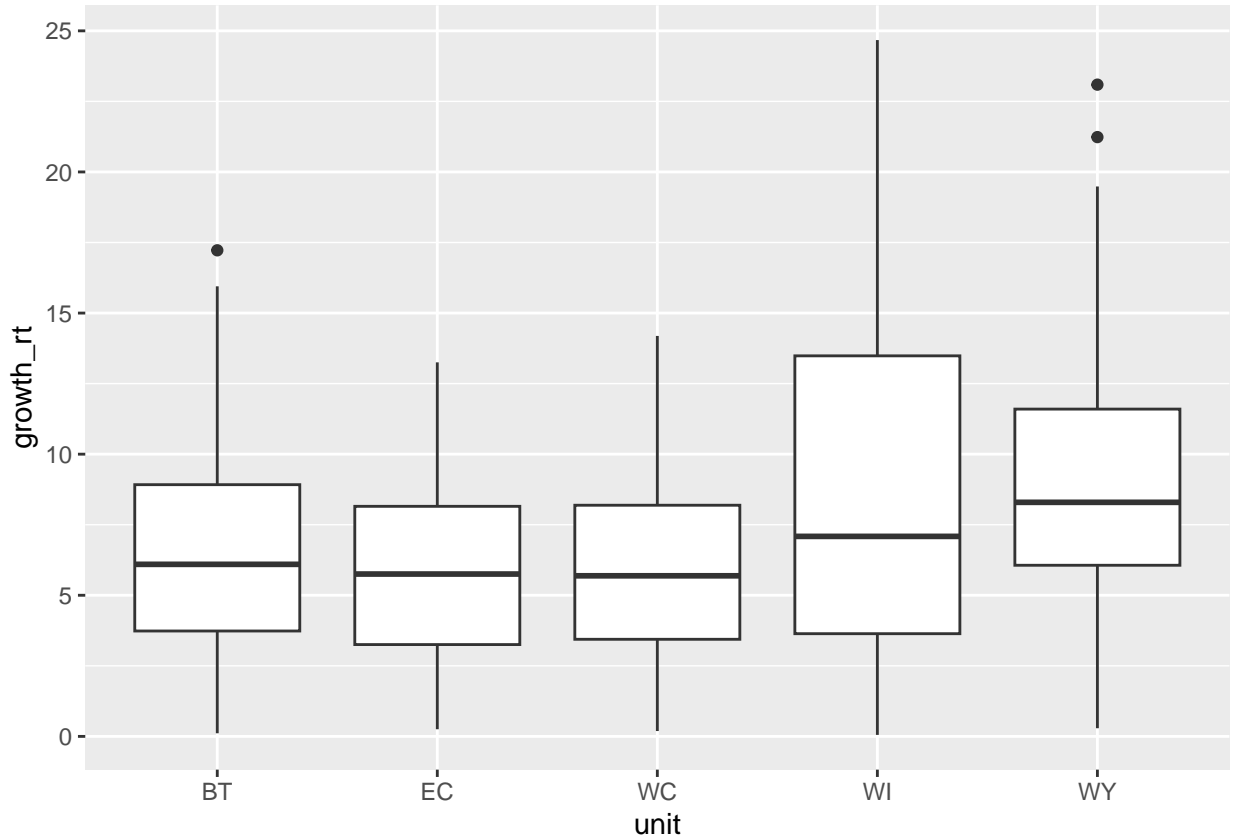

Figure 3: Growth rate by unit

At each site, data on seedling heights (used to estimate growth rate) and competing plant species was recorded. Water balance variables were estimated using Daymet as inputs for temperature and precipitation. The following predictor variables were evaluated:

- Predictor variables are as follows:
    - Age - Years since planting
    - T_{mean} - Mean annual temperature (°C)
    - T_{max} - Max monthly temperature (°C)

4

- PPT (mm) - Mean annual precipitation.
- Snowpack (mm) - Mean spring (March-May) snowpack
- WD_{annual_mean} (mm)
- WD_{annual_max} (mm)
- WD_{month_max} (mm)
- AET - Mean actual evapotranspiration (mm) during growing season (April–October)
- PET - Mean potential evapotranspiration (mm) during growing season (April–October)
- GDD - Mean annual growing degree days (April-October)
- 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.
- $Micro_{prop}$ - Proportion of WBP with a microsite at the site-level
- 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.
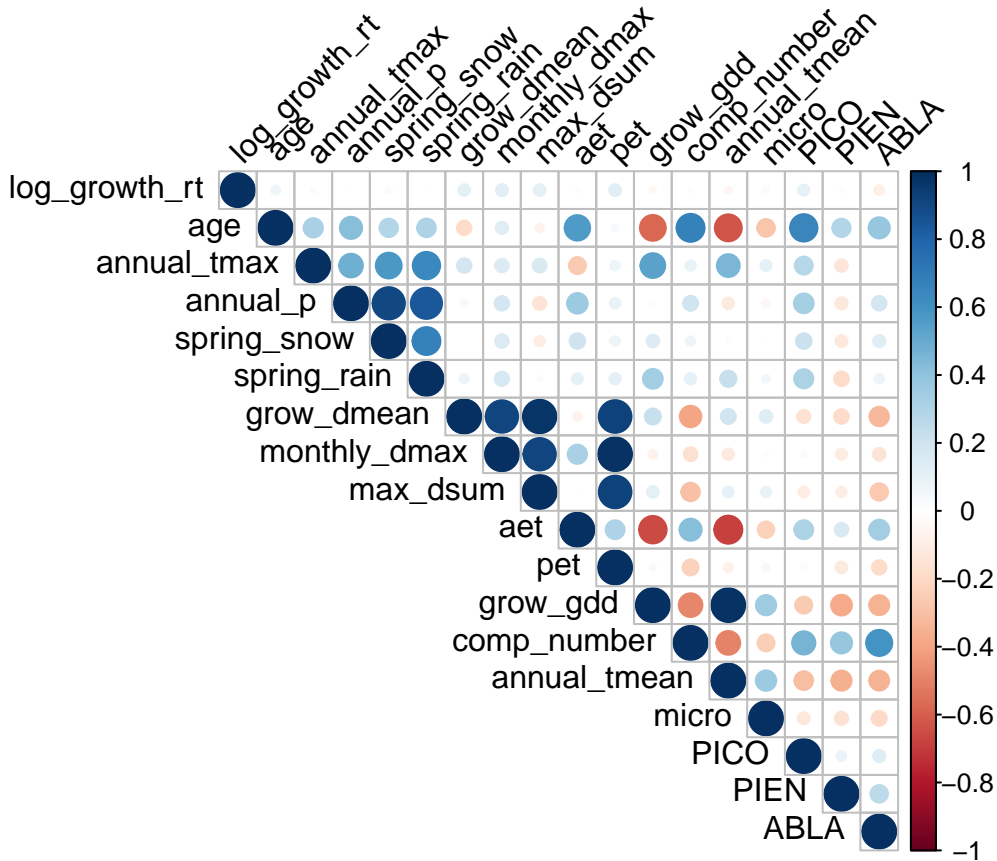
Data were provided by Laufenberg, but the process of cleaning and structuring the data were not explicit in the paper. As such, we had to experiment to figure out how data were cleaned. We did this by modifying data in an ecologically reasonable way while comparing raw data plots to the paper. A few examples of data cleaning are: - Creating AET which is the total annual evapotranspiration. To get this, we multiplied the average monthly evapotranspiration by 7 for the 7 months of the growing season. The same was done for PET or potential evapotranspiration. - multiplying p by 12 to get the total "annual precipitation". - using log_{10} to transform growth rate.

Climate and water balance predictor variables were tested for colinearity with a cutoff of r $>= +/- $ 0.6 (@ref(fig:corr_matrix). A parsimonious list of water balance variables was selected by choosing the more biologically relevant variable from pairs that exceeded this threshold. This list of variables was then combined with the variables pertinent to the research question, which were for the number of competitors, age of the planting site, and microsite presence/absence to create a list of candidate variables to evaluate in model selection. Following this process, our list of candidate variables to evaluate in model selection is T_{max}, PPT, PET, AET, comp_number, micro, PICO, PIEN, ABLA. This differs from the "full model" evaluated in Table 6 of Laufenberg et al. 2020 by the inclusion of T_{max} and the exclusion of T_{mean}.

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. A variety of linear and polynomial functions, and fixed effects were explore. A corrected AIC was used to compare models. We attempted to find the most parsimonious set of water balance variables while still incorporating the biologically relevant variables and incorporating environmental variables related to the research questions (including number of competitors, age of planting, and microsite status).

model	AICc
log_growth_rt ~ aet^3 + aet^2 + aet + pet^3+ pet^2 + pet + comp_number^3 + comp_number^2 + comp_number + annual_p^3 +annual_p^2 + annual_p + annual_tmax^3 + annual_tmax^2 + annual_tmax +PICO + PIEN +ABLA + (1 | unit)	1126.504
log_growth_rt ~ aet^3 + aet^2 + aet + pet^3+ pet^2 + pet + comp_number^3 + comp_number^2 + comp_number + (1 | unit)	1130.444

Table 1: Model Suite Tested

We ran the saturated model without interactions (0.2.2), and then refined it based on AIC score (less than 2). We used the $drop1()$ function to conduct backwards stepwise analysis using single term deletions following a χ^2 distribution. All analysis was done use R statistical software.

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

0.3 Results/Summary of Statistical Findings

The initial full model, with all variables of interest, had an AICc of 1126.504 , and our final model has an AICc of 1122.34 (Table (**??**)

ONLY GOT ONE DIAGNOSTIC PLOT FOR THE MODEL.

The final model suggests the WBP seedling growth is most correlated with the annual evapotranspiration, the potential evapotranspirtaion, and the number of competitors around the seedling, while accounting for the random effect of planting unit. The cubic forms of these variables suggest WHAT DOES THE CUBIC FORM MEAN.

- ALSO NEED TO INTERPRET COEFFICIENTS AND 95% CI

0.4   Scope of Inference

- Random sampling of WPB = inferences to large population

- No random treatment (because nature) = correlation/association but not causation.

0.5   Critique

- paper had no clear model selection criteria or process

- difficult data cleaning and processing without instruction

- unfamiliar with drop1() technique for model selection, so hard to critique that with what we are used to, but struggled using step selection with random effects

0.6   Group Work Statement

Steve spearheaded the data wrangling and analysis, while Parker took lead on the write-up and formatting. Both people contributed to the interpretation, discussion, and critique.

0.7   References

0.8   Appendix

- Must include a compiled RMarkdown with all of our results