Final Project

RJ Cass

Table of contents I

Abstract

Tulips form a significant portion of the exports from the Netherlands. We want to understand what changes can be made to ensure tulip growth is not impacted by the changing climate.

We used a XXXXX model and found:

- Effect of chilling times on germination is dependent on species: some improve with longer chilling, some get worse with chilling
- Ideal chilling time for each population (for most, range of 8-10 weeks)
- Predicted impact of the chilling time decreasing from 10 weeks to 9 (some species do worse, some very much worse, and one does a little better)

Context

Tulip Production:

- Tulip products form 25% of agricultural exports from the Netherlands
- Changing climate puts the tulip industry at risk
- Want to understand how to adapt to these changes and protect the industry

Dataset of sample tulip growth populations:

- Year they were grown
- Number weeks the bulbs were chilled
- Whether or not the bulb germinated
- Indices (can be removed from dataset)

Questions of Interest

We want to use the provided data to answer the following questions:

1. What is the effect of chilling time for the different species of tulips? Is it the same across the species? Which species are the same/different?
2. Is there an ideal chilling time for each species? If so, is it the same for all species?
3. Given climate change conditions, winters are expected to decrease from 10 to 9 weeks in the coming few years. What effect will this decrease in chilling time have on the probability of germination for each species? Is it the same for all species?

EDA - Population 12

None of population 12 germinated. We removed it from the dataset to not dilute the rest of the data

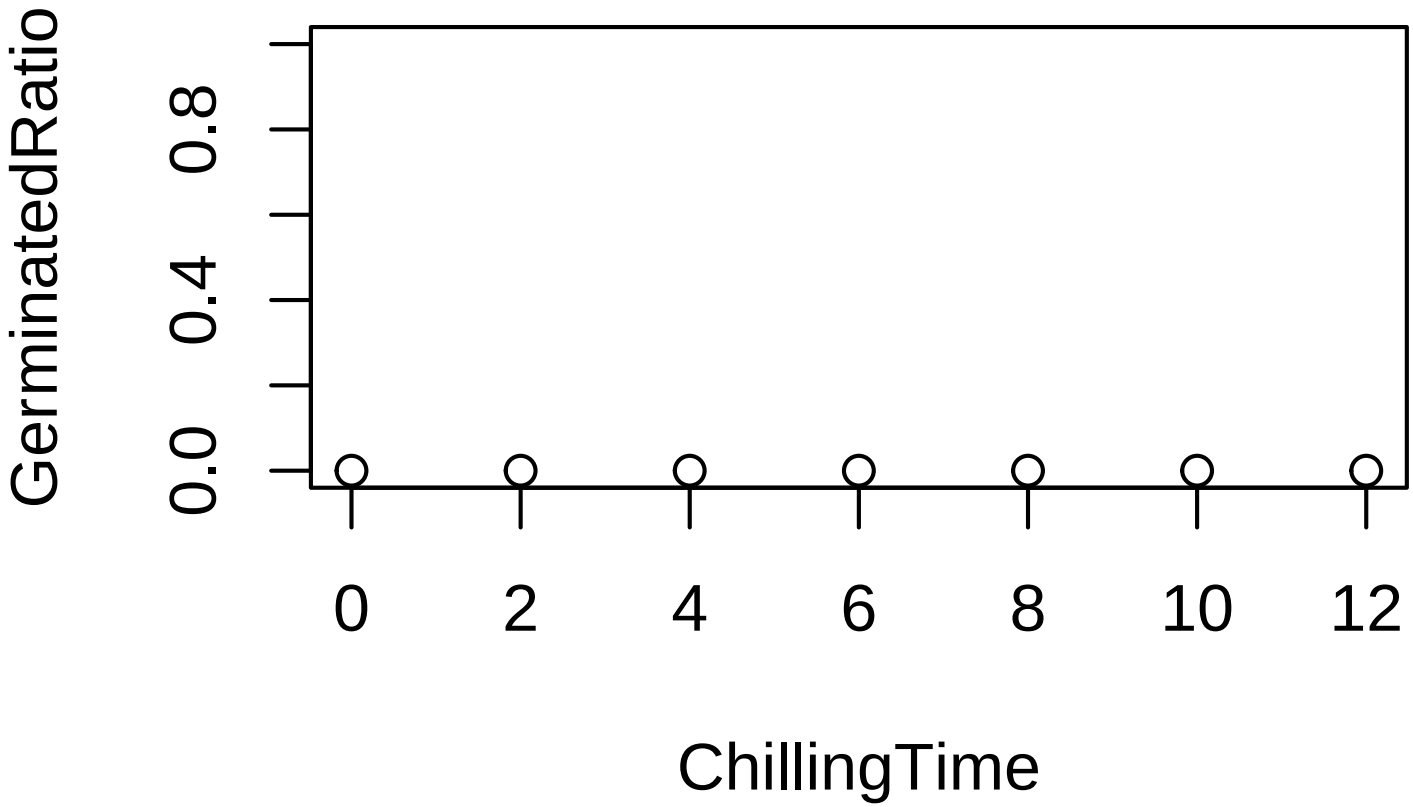


Figure 1: *Population 12 had a 0% germination rate across all chilling times*

EDA - Year

- Each population was tested in only 1 year (ie. no crossing with different years having an effect on one population)
- Physically, given testing conditions, we don't expect year to have an impact on germination
- In variable selection, Year was not important ($p > .05$)
- Removing year from models

EDA - Interactions - Population vs. Chilling

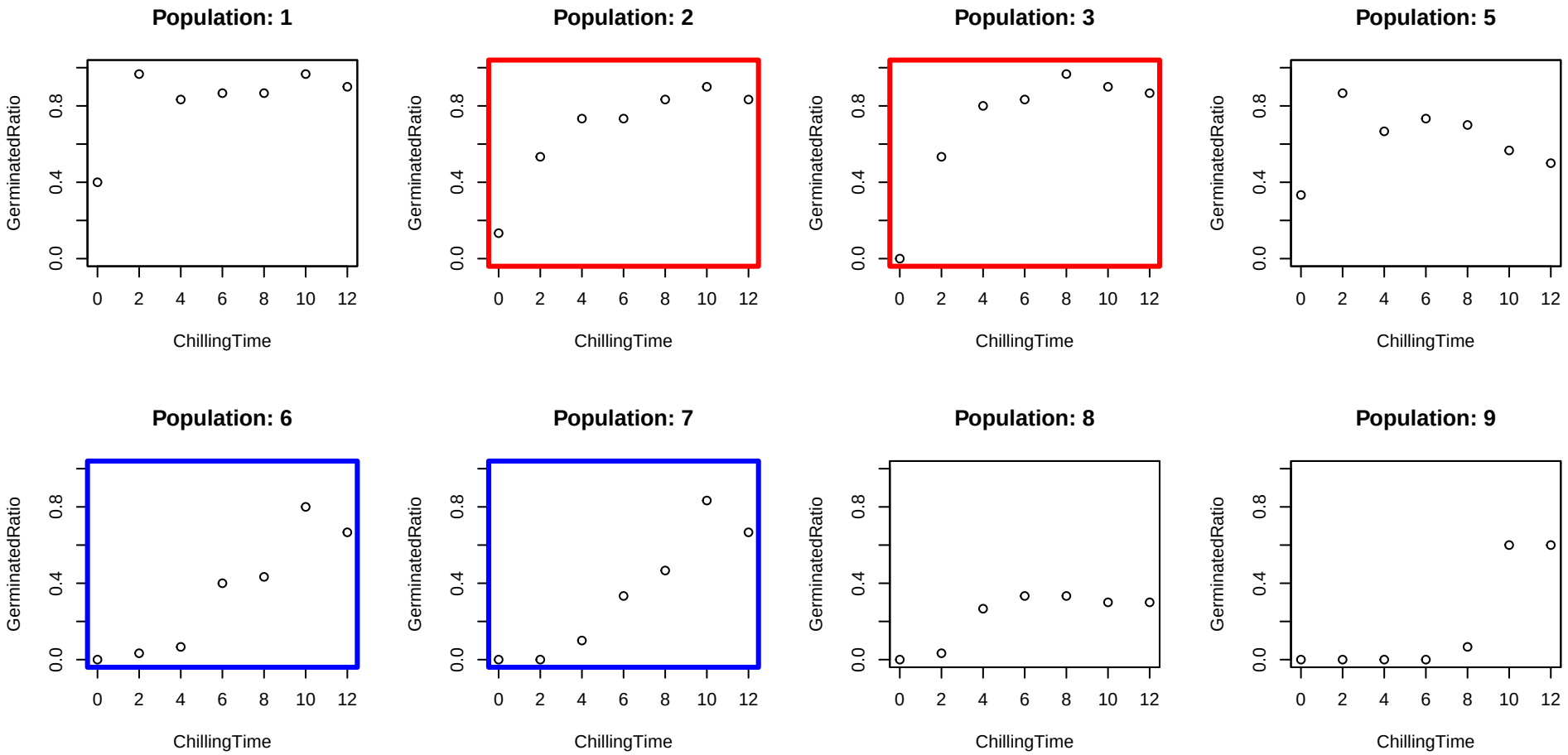


Figure 2: *Sample interaction plots of population and chilling. Some behave similarly, others are vastly different. Note the non-linearity of some populations.*

EDA - Summary

Removing from dataset:

- Population 12
- 'Year' variate

Need to account for the following:

- Interaction between Chilling Time and Population
  - Will need to ensure model includes interactions (either manually, or use a model that explores interactions)
  - If not included, resulting model will not capture the full impact of each variate
- Non-linearity of relationship between chilling time and germination rate
  - Will need to ensure the model handles non-linearity
  - If not included, model will not represent the correct relationship of this variate

Proposed Models - 1

Logistic Model - Interactions and spline

$$\begin{aligned} Y_n =& ln(\frac{p}{1-p}) = \beta_0 + \beta_1 X_{pop} + \beta_2 bs(ChillingTime) \\ &+ \beta_3 (X_{pop} * ChillingTime) \end{aligned} \tag{1}$$

- Accounts for interactions of population and year on ChillingTime
- Accounts for non-linearity of ChillingTime

Strengths:

- Captures interactions
- The concept of logistic (change in log-odds) is realtively interpretable

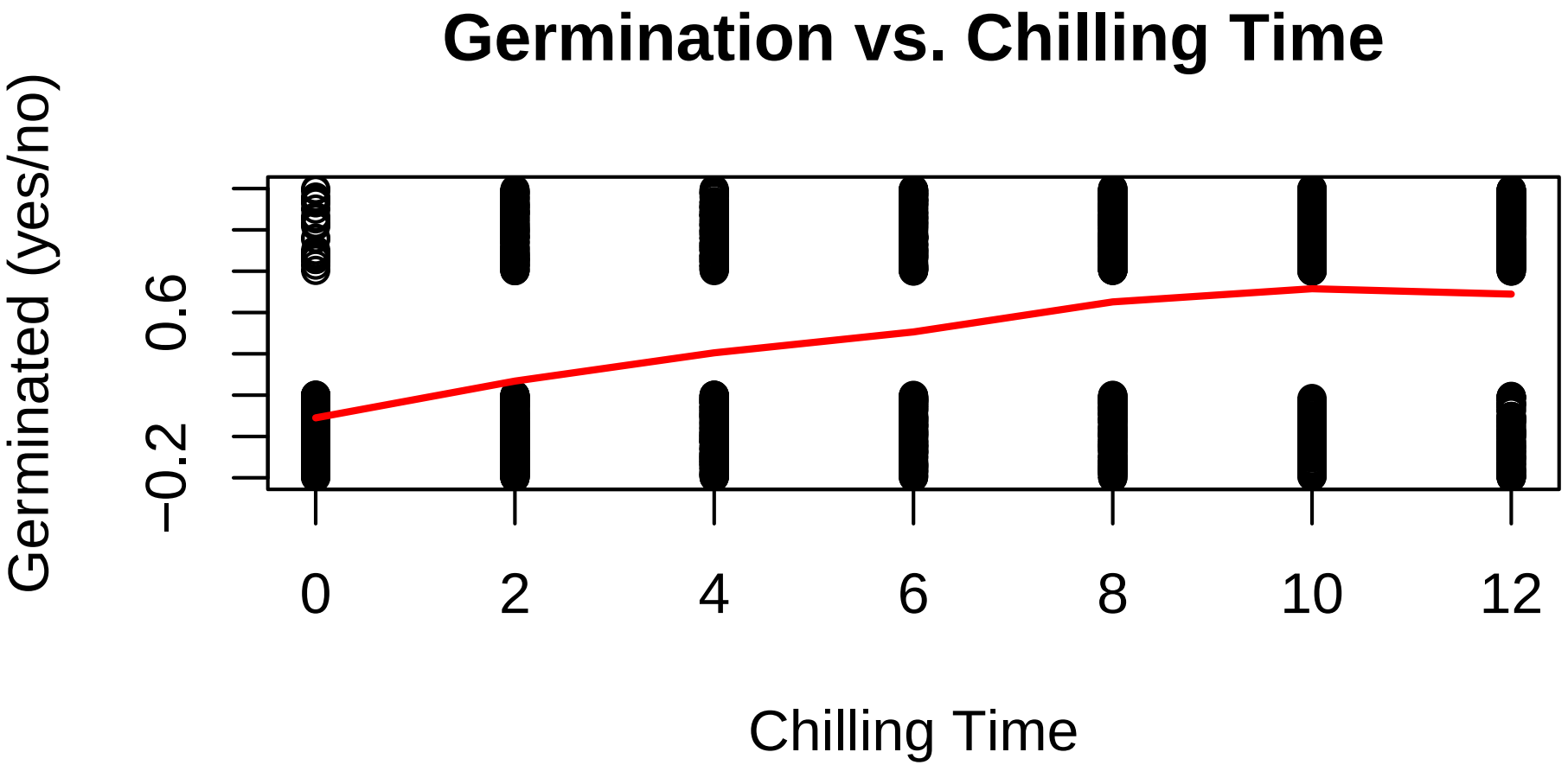
Weaknesses:

- Using splines loses interpretability

Proposed Models - 1 - Cont'd

Assumptions - Indepndence, Monotonocity

- Independence: Assumed due to the design of the experiment
- Monotonicty
  - Shown in below graph, Germination rate is acceptably monotonic



Proposed Models - 2

Random Forest

SAMPLE TREE IMAGE

Strengths:

- Relatively explainable (lots of trees, each tree gets a vote, average the votes, compare to cutoff)

Weaknesses:

- Can be prone to overfitting

Model Evaluation/Selection

COMPARE IN SAMPLE AND OUT OF SAMPLE PREDICTION
ESPECIALLY LOOK AT INTERPRETABILITY

Effect of Chilling Time

Ideal Chilling Time

Effect of Decrease in Chilling Time

Summary

Next Steps

STUFF IF NEEDED