

# Predicting pollination phenology in lodgepole pine

*Susannah Tysor*

*May 8, 2019*

```
forcingtype = "scaled_ristos"

## Loading required package: ggplot2

## Registered S3 methods overwritten by 'ggplot2':
##   method      from
##   [.quosures    rlang
##   c.quosures    rlang
##   print.quosures rlang

## Loading required package: StanHeaders

## rstan (Version 2.18.2, GitRev: 2e1f913d3ca3)

## For execution on a local, multicore CPU with excess RAM we recommend calling
## options(mc.cores = parallel::detectCores()).
## To avoid recompilation of unchanged Stan programs, we recommend calling
## rstan_options(auto_write = TRUE)

## Loading required package: parallel

## rthinking (Version 1.88)

## This is bayesplot version 1.7.0

## - Online documentation and vignettes at mc-stan.org/bayesplot
## - bayesplot theme set to bayesplot::theme_default()
##   * Does _not_ affect other ggplot2 plots
##   * See ?bayesplot_theme_set for details on theme setting

## -- Attaching packages ----- tidyverse 1.2.1 --
## v tibble  2.1.3     v purrr   0.3.2
## v tidyr   0.8.3     v dplyr   0.8.1
## v readr   1.3.1     v stringr 1.4.0
## v tibble  2.1.3     vforcats 0.4.0

## -- Conflicts ----- tidyverse_conflicts() --
## x tidyr::extract() masks rstan::extract()
## x dplyr::filter()  masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## x purrr::map()    masks rthinking::map()

# combined phenology and weather data
phenology_data <- read.csv("../data/phenology_heatsum_all.csv", header=TRUE, stringsAsFactors = FALSE) %>
  filter(forcing_type==forcingtype)
phenology_data_model <- read.csv("../data/phenology_heatsum.csv", header=TRUE, stringsAsFactors = FALSE)

# seed orchard provenances
SPU_dat <- read.csv("../research_phd/data/OrchardInfo/LodgepoleSPUs.csv",
  header=TRUE, stringsAsFactors = FALSE) %>%
  dplyr::select(SPU_Name, Orchard) %>%
```

```

  rename(Provenance=SPU_Name)

# join provenance and phenology data

dat <- phenology_data %>%
  na.omit() %>%
  left_join(SPU_dat) %>%
  unique()

## Joining, by = "Orchard"

```

## Abstract

The timing and duration of pollen shed and cone receptivity in lodgepole pine affects fecundity, gene flow, and adaptation in the species. Being able to predict flowering phenology in the past and present will help us understand patterns of local adaptation and population structure. Using a 14 year dataset collected at 7 sites, I built a multilevel Bayesian model of lodgepole pine flowering phenology and used it to determine the amount of forcing required to begin and end the flowering period for both male and female strobili. With this method, any record of mean daily temperature can be used to estimate lodgepole pine flowering phenology.

## Outline intro

Pollination phenology is important because - gene flow, dispersal - spatial and temporal variation - reproduction, demography - assisted migration, conservation - seed orchard operations, commercial/agricultural/breeding

While vegetative phenology has been the focus of much research, pollination phenology is relatively unexplored in conifers. The focus is usually on day of year and comparing provenances, looking for local adaptation. They aren't predictive. - examples

Benefits of predictive/mechanistic models - climate change - predict out of dataset

Predictive models are scarce because they rely on excellent phenological records with high quality temperature datasets. - counterexamples with strengths and weaknesses

Such a record does exist for lodgepole pine in British Columbia. - BC seed orchards, collection rationale, existing dataset

Understanding the precise relationship between temperature and pollination phenology is valuable in lodgepole because - it's an economically important species with lots of replanting and we need to understand how populations we're planting on the landscape will interact with local populations - lots of phenotypic and genomic data available in lodgepole pine. Local adaptation, population structure. Understanding temporal and spatial variation in pollination phenology could help disentangle effects on population differentiation (local adaptation, gene flow, etc) and explain population structure. - Pine pollen can be an allergen and appears to be made worse of one when it interacts with certain kinds of pollution. Knowing when pollen will be produced is very helpful for people with allergies, esp since medication works best if taken in advance.

In this paper, I build a predictive model of pollination phenology in lodgepole pine that estimates the amount of forcing required to cause lodgepole pine to begin and end flowering. Using these estimates, any record of mean daily temperature can be used to predict the timing and duration of flowering in lodgepole pine.

I confirm that pollination phenology in lodgepole pine is not strongly influenced by provenance, then used the model results to calculate variation in the timing and length of phenological period at 7 sites in British

Columbia between 1997 and 2011. Last, I considered how a business-as-usual climate change scenario may affect pollination phenology throughout this century.

## Introduction

As the climate changes, spring phenological events like budburst and flowering will advance, especially for plants active in rapid seasonal transitions and short growing seasons (Pau et al. 2011), like many high elevation and latitude conifers. This effect is already obvious in many species (Parmesan 2006; Franks, Weber, and Aitken 2014). Changes in pollination phenology can affect fecundity, gene flow, and even range size in a species and have effects [VAGUE] on dependent species (Inouye 2008; ???).

[POLISH THIS INTRO SENTENCE] Conifers are a big part of the enormous northern hemisphere forests and they have wide ranges with lots of local adaptation. Common garden experiments and genetic work reveal extensive local adaptation in many forest tree species, especially boreal and temperate conifers (reviewed in (???)). A locally adapted population only grows optimally in a subset of the range and may tolerate a more limited climatic range than the species as a whole. In northern hemisphere conifers, local adaptation often reflects strong trade-offs between avoidance of cold damage and competitive height growth (summarized in (???)).

Coniferous forest trees are wind pollinated with pollination possible over large distances. [EXAMPLES] Pollen is shed from male strobili and must arrive at receptive female strobili for successful pollination .rpe Shifts in the timing of pollen shed and cone receptivity (pollination phenology) in conifers could lead to gene flow changes that hinder or promote adaptation under climate change, decrease fitness, and even affect reforestation via seed production declines. [Also affects gene flow now and is important for understanding current spatial genetic structure and local adaptation]

Lodgepole pine is a good representative of these issues - an economically and ecologically important tree species facing multiple threats from climate change (Schneider et al. 2010; Sambaraju et al. 2012; Hamann and Wang 2006). Lodgepole pine has a very large geographical distribution (across 33° latitude and 31° longitude) encompassing a wide range of climates and soils (Figure 1) with widespread and significant local adaptation in many traits. For example, populations from both northern interior British Columbia and northern Idaho can survive in areas with mean annual temperatures between -4 and 6 °C, but the northern British Columbia population survives best where mean annual temperatures are ~ 1 °C and the Idaho population best at ~ 4 °C (PRUS 1971). Local adaptation in lodgepole pine can be observed even at relatively small spatial scales when topographic variability is high: in a reciprocal transplant experiment, growth declines were observed when moving high elevation populations just 100m in elevation (???. [Summarize/cite adaptree work]

Pollen is an important vector for identifying gene flow in lodgepole pine because outcrossing is common, pollen dispersal is extensive, and seed dispersal is relatively limited (Ennos 1994). There is evidence for spatially varying levels of gene flow in the species as populations from areas with higher regional climate heterogeneity have higher genetic variance (Yeaman and Jarvis 2006). Pollination phenology could control this.

Pollination phenology determines which populations can exchange genes, but predicting the timing of pollen shed and cone receptivity has not been done in lodgepole pine. Pollination phenology examples are uncommon. [Talk about Sarvas investigations for mechanistic background, why using temperature, etc.] Simple heat accumulation thresholds (*Pinus taeda*) (Boyer 1978) or elevation (*Pinus flexilis*) (Schuster, Alles, and Mitton 1989) were used previously to explain or predict pollen shed in limited spacial and temporal contexts. (???) reports that lodgepole pine pollen shed and cone receptivity occur when degree days reach about 500 at a threshold of 5 °C, but this is the only report of pollination phenology modeling for lodgepole pine and limited details are provided. Models of lodgepole pine vegetative phenology, on the other hand, are better represented in the literature (e.g. Chuine, Aitken, and Ying (2001)), and pollen shed and cone receptivity are not expected to have additional or more complex triggers or model forms than budburst (Chuine, Kramer, and Hanninen 2003).

Predicting pollination phenology will also have practical benefits. Seed orchard managers in British Columbia are particularly concerned about protandry, when all pollen in an area is shed before cones become receptive (*personal communication, Chris Walsh, Former Seed Orchard Manager, Kalamalka Seed Orchards, February 13, 2013*). Protandry occurs in particularly hot and dry years (Owens, Bennett, and L'Hirondelle 2005). If this pattern holds, protandry could become more common in natural populations, leaving some populations pollen limited and likely hampering local adaption. [Outcrossing! Inbreeding!]

This paper relies on 15 year pollination phenology data set collected in British Columbia lodgepole pine seed orchards. Seed orchards produce seed for reforestation. While not set up as common gardens, several provenances are typically represented at each site allowing testing for provenance effects and genetic x environmental interaction. and genotypes usually appear multiple times within a site and sometimes at multiple sites.

I will use pollination phenology and temperature data to fit a model predicting pollen shed and cone receptivity in *Pinus contorta* var. *latifolia*.

Specifically, I will answer 1) What is the relationship between temperature and pollen shed and temperature and cone receptivity timing and length? 2) Does provenance affect that relationship? 3) How does pollination phenology vary between cold and hot years? 4) Will protandry become more common?

## Methods

My aim was to model pollination phenology in lodgepole pine and calculate synchrony across the distribution. To determine the timing of pollen shed and cone receptivity, I modeled the forcing requirements for pollination phenology in lodgepole pine and investigated differences in forcing requirements between males and females and among provenances.

### Study species

Lodgepole pine is a wind-pollinated monoecious conifer with an extensive range and well documented local adaptation. While lodgepole pine has four subspecies, this work concerns only *Pinus contorta* subsp. *latifolia*. Pollen and seed cone buds differentiate in late summer and early fall, then go dormant. As temperatures warm the following year, buds resume development and strobili “flower”; receptive female strobili exude pollination drops between bracts that capture pollen shed from mature male strobili.

### Data

#### Phenology Data

This project takes advantage of an existing lodgepole pine pollination phenology dataset collected over a decade and a half by government and industry workers in seed orchards. Seed orchards in British Columbia produce large amounts of tree seed for reforestation from parent trees sourced from provenances around the province. To plan for future seed production and orchard establishment and management, seed orchard managers monitored pollination phenology and seed production. Pollination phenology data was collected at the Prince George Tree Improvement Station in British Columbia beginning in 1997 and collection at many other BC tree orchard sites began in 2006 under the Forest Genetics Council's Operational Tree Improvement Program 0722 (???). Collection continued intermittently through 2012.

#### Sites

Trees selected from across the British Columbia portion of the lodgepole pine range are grown in seed orchards as part of tree breeding and seed production programs. Between 1997 and 2011, flowering phenology of lodgepole pine was recorded at 7 seed orchard sites in the interior of British Columbia. I contacted Seed

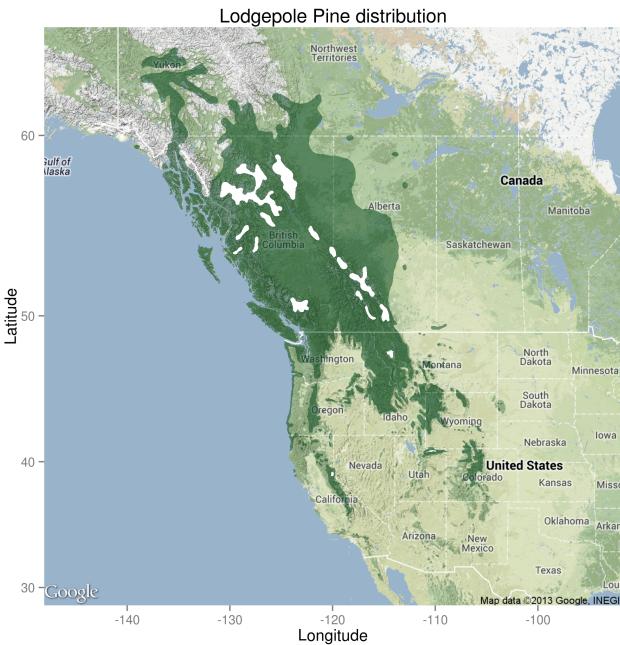


Figure 1: Rangemap of lodgepole pine. After Little E.L. (1971) .

Orchard Managers and other forestry professionals across British Columbia in 2012 and received pollination phenology data from C. Walsh, previously at Kalamalka Seed Orchards (now retired), R. Wagner at the Prince George Tree Improvement Station, and J.E. Webber previously at the Glyn Road Research Station (now retired). 4 of the sites are clustered near to one another around Vernon, BC. Sites span about 5 ° of latitude and are at elevations from 466 to 638 m.

### Provenances

Trees grown at the seed orchard sites were sourced from 7 Seed Planning Units (SPUs). SPUs are organizational units in BC forestry that are based on biogeoclimate regions. Trees with the same SPU provenance are grown together in an orchard at a given site. Genotypes (labelled with a Clone number in the data) in the orchards are represented by multiple ramets.

I obtained pollination phenology records from 17 of the 26 lodgepole pine seed orchards in British Columbia. Orchards include the offspring of trees from 5 of the 13 BC seed planning zones (SPZs) [SPZ REFERENCE OR FIGURE?] grown at 7 sites across BC. SPZs are biogeoclimatic and political units used for seed planting purposes by British Columbia. SPZs are divided into elevation bands called Seed Planning Units (SPUs), which form this project's 7 provenances 4.

Thirteen out of 17 orchards in my data set are first generation orchards and should faithfully represent their provenances. These first generation orchards represent 6 provenances at 5 sites. Second generation orchards have been selectively crossed and this may skew the mean or variance of phenology for a provenance if pollination phenology varies by provenance.

Most provenances are represented at 2 to 3 sites and have at least three years of data at a given site spanning 1997-2012 (Figure 3). The Prince George Tree Improvement Station (PGTIS) provides a continuous 15-year record of its three orchards' phenology. Most sites only observe one ramet per clone, PGTIS typically observed 2-4 ramets per clone.

[table with Site Columns and SPU rows with years of data as values] [table of number of trees/clones in a given year for a given site/spu combo]

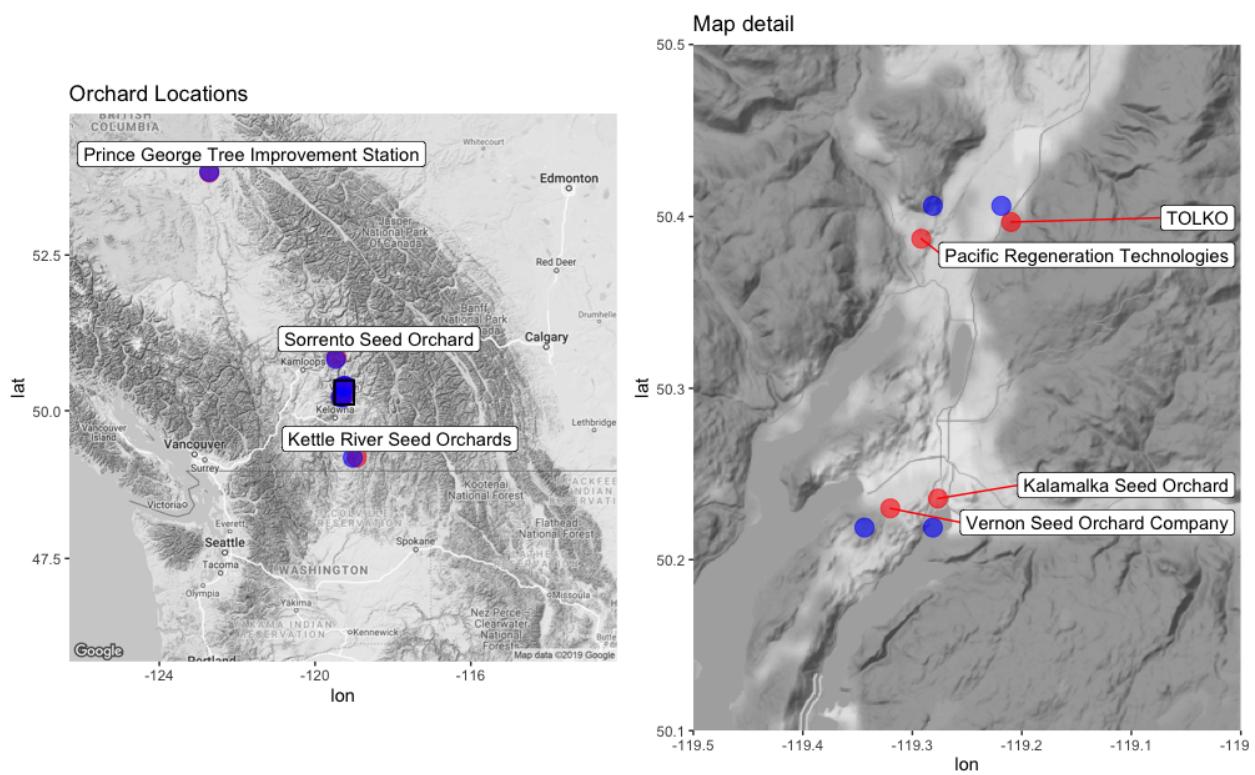


Figure 2: Map of seed orchard locations. Seed orchard locations are in red. Weather data gridpoints are in blue. Boxed area in map on left is shown in greater detail on the right.

	Kalamalka	KettleRiver	PGTIS	PRT	Sorrento	TOLKO	Vernon	Total
Bulkley Valley Low	5	0	17	0	2	0	3	27
Central Plateau Low	0	0	15	0	0	0	4	19
Nelson Low	5	0	0	3	0	0	0	8
Prince George Low	0	2	18	0	0	0	4	24
Thompson Okanagan High	0	0	0	0	0	3	0	3
Thompson Okanagan Low	0	0	0	6	0	0	0	6
Total	10	2	50	9	2	3	11	87

Figure 3: Contingency table of years of data for Seed Planning Units (rows) and Seed Orchard Sites (columns). Seed Planning Zones, used as provenances in this project, are usually represented at multiple years and multiple sites. There is particularly good representation at PGTIS.

### Phenology scoring protocol

Protocol C in (Woods, Stoehr, and Webber 1996) was used as the basis for collecting pollen shed and cone receptivity data, though operational constraints led to some modifications. Workers monitored seed orchards for the beginning of pollen shed and cone receptivity. ~15 clones, usually represented by 2 trees each, were selected for specific observations. When the active period seemed to be starting, workers went into the orchard every few days to make observations on the selected trees.

Stages of pollen and seed cone development are described by [Owens & Molder 1984 and updated in Owens2006] and were used as a general guide for determining the phenological state of pollen and seed cones. Pollen cones are flowering when tapping causes pollen to be released and seed cones are flowering when there are gaps between the bract-scale complexes. Pollination drops are also produced during flowering, though they recede midday if pollen is not present (???). q8

“Flowering” states were recorded for both pollen shed and cone receptivity at the level of each tree. Protocol C recommends marking the dates when 20% of the cones on a tree have begun flowering and when 80% of the cones on a tree have finished flowering. Operationally, there was some subjectivity and tree-level states for each cone type should be interpreted as “starting flowering” and “finished flowering.” There is some subjectivity here.

[DESCRIBE SURVEY PERIODS]

```
# Index by Site, Provenance, and Year
dat$spyind <- group_indices(dat, Site, Year, Provenance)

surveydf <- select(dat, spyind, Year, Site, Provenance, DoY) %>%
  distinct()

ggplot(surveydf, aes(x=DoY, y=as.factor(spyind), color=Provenance, shape=Site)) +
  geom_point() +
  geom_line() +
  facet_grid(rows=vars(Site,Year), scales="free_y") +
  scale_color_viridis_d(option="C") +
  scale_shape_manual(values=c(1:7)) +
```

## Lodgepole Pine Seed Planning Units

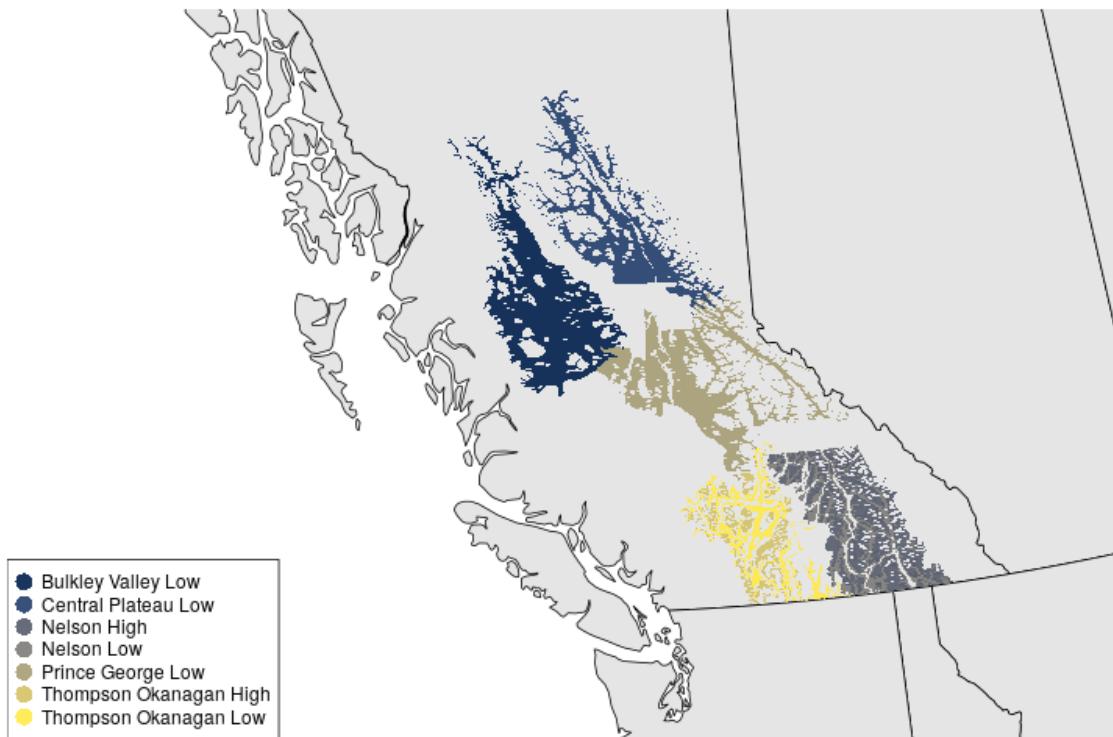


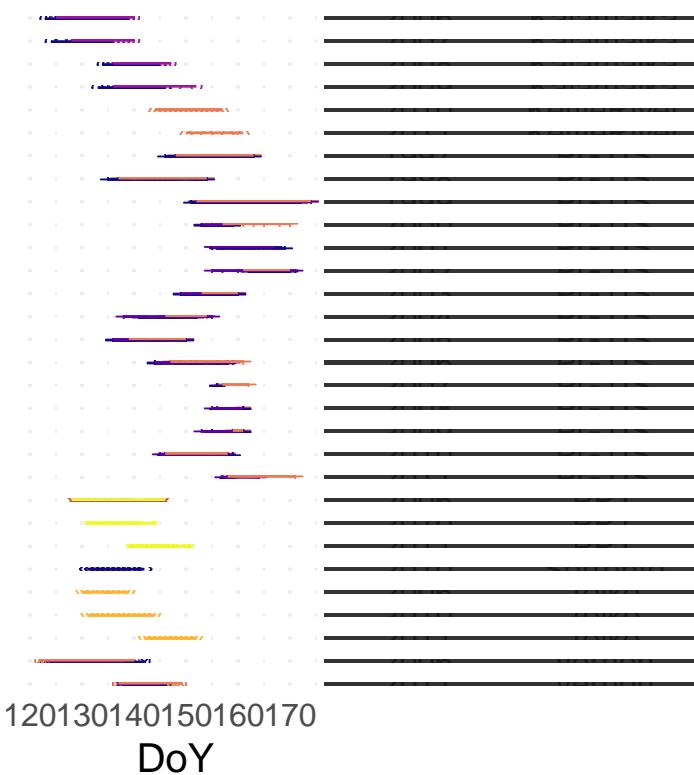
Figure 4: Map of Seed Planning Units (SPUs). Seed planning units are biogeoclimatic and political units used for seed planting purposes by British Columbia. Seed planning units form this project's provenances. High, Low, and Mid refer to elevational bands. Data is also available for East Kootenay Low, but will likely not be included in any analysis as it includes only one year at one site.

```

theme_bw(base_size = 15) +
theme(strip.text.y = element_text(angle = 0), axis.text.y = element_blank(), axis.ticks = element_blank())
ggtitle("Observation Dates")

```

## Observation Dates



### Provenance

- Bulkley Valley Low
- Central Plateau Low
- Nelson High
- Nelson Low
- Prince George Low
- Thompson Okanagan High
- Thompson Okanagan Low

### Site

- Kalamalka
- △ KettleRiver
- + PGTIS
- × PRT
- ◊ Sorrento
- ▽ Tolko
- ▣ Vernon

120130140150160170

DoY

Observations were not made every day and survey periods varied in length. At Prince George, not all trees have complete phenological records, *e.g.* if a tree is not flowering on the first day of observation, the start date is unknown.

### Harmonization and cleaning

[DESCRIBE DATA TRANSCRIPTION AND CLEANING].

There were some differences in how data was recorded at the Prince George Tree Improvement Station versus the other sites. At Prince George, trees were marked as flowering or not flowering on each day of observation. At other sites, only the first day observed flowering and the first day observed finished flowering were recorded. I inferred that trees were not yet flowering on observation days prior to their first recorded flowering date. I cleaned and harmonized the data for analysis in a single model using R scripts provided so that, for pollen cones and seed cones, stage 1 = not yet flowering, stage 2 = flowering, and stage 3 finished flowering.

There are three phenological stages of interest each for a tree's male and female cones

- 1. The cones have not yet flowered
- 2. The cones are flowering
- 3. The cones have finished flowering

Phenophases in the field were recorded using different symbol sets and resolutions. I assigned each symbol to one of the phenophases above. Trees that did not produce cones are assigned phenological stage 0.

In total, there were 21911 usable observations of phenological state.

Some observations at the Prince George Tree Improvement Station Site were redundant and were dropped for modeling. For example, if a tree's pollen cones were recorded as "finished flowering" for 3 days, only the first observation was kept - phenological states are not reversible. This improved the speed of model fitting.

Phenophase	Symbols	Male.Cones	Female.Cones
0	0	none produced	none produced
1	1, 2.5, -	not yet shedding	not yet receptive
2	3, 3.5, 4, 4.5, 5, pollensheded20, receptive20	shedding	receptive
3	-, receptive80, pollensheded80	finished shedding	no longer receptive

[Figure showing phenology data somehow]

### Weather data

Daily weather data at seed orchard sites was extracted from PNWNAmet, a daily gridded meteorological dataset at 1/16 ° (~ 6km) over northwest North America (Consortium 2014). The closest point in the PNWNAmet grid was used for each station 2. Mean daily temperature was calculated as the average of the minimum and maximum daily temperatures provided by PNWNAmet.

As the PNWNAmet gridpoints do not exactly align with the sites, a correction for elevation was applied using an adiabatic lapse rate of 6.4°C per 1000 m.

$$t_{site} = t_{gridpoint} + \left( \frac{6.4^\circ\text{C}}{1000\text{m}} \right) (elevation_{site} - elevation_{gridpoint})$$

### FIGURE SUMMARIZING RAW MEAN TEMPERATURES?

Figure X shows the corrected and uncorrected temperatures with temperatures collected by seed orchard workers at the sites using automated data loggers. In the figure, PNWNAmet temperatures have been translated to growing degree days with a 5°C threshold because only GDD calculated from the logged data was available from the sites.

### Forcing units

Flowering is a developmental process that speeds up and slows down according to the current temperature. Forcing units describe the relative effect of temperature on development. Observable phenological events occur only after a certain amount of forcing has accumulated. The mean daily temperature  $t$  was mapped to forcing units with a function that describes the relationship between temperature and development rate.

$$R(t) = \frac{1}{1 + e^{-0.185*t - 18.4}}$$

This equation for forcing units is based on experimental work in temperate forest tree species (Sarvas 1972, @hanninenModellingBudDormancy1990) and is verified to perform better than growing degree day models by Chuine (Chuine, Cour, and Rousseau 1999).

Accumulated forcing units on day  $d$  ( $f(d)$ ) are the sum of the relative temperature effect from January 1 (ordinal date 1) to day  $d$ .

$$f(d) = \sum_{i=1}^d R_d(x)$$

Cooler sites, like Prince George, accumulate forcing slower than warmer sites like Kalamalka.

[ACCUMULATED FORCING UNITS FIGURE]

## Phenology Modeling

I modeled phenological states as a function of accumulated forcing units with a Bayesian multilevel ordinal logistic model fit in Stan with the No-U-Turn Sampler (Team 2018). The model was fit separately to data for male strobili and female strobili. A logistic cumulative link function accounts for the ordering of phenological states and relates phenological states to a linear model. This type of model makes no assumption about the distance between phenological states. The model parameters describe a probability function for each state. Cutpoints  $\kappa$  separate each state while the slope parameter of the linear model influences the steepness of the curves. Transitions between phenological states happen rapidly when the slope is large and slowly when it is small.

The likelihood of being in state  $s \in \{1, 2, 3\}$  is

$$\Pr(S = s) = \text{OrderedLogistic}(s | \eta, \kappa) = \begin{cases} 1 - \text{logistic}(\eta - \kappa_1) & \text{if } s = 1 \\ \text{logistic}(\eta - \kappa_1) - \text{logistic}(\eta - \kappa_2) & \text{if } s = 2 \\ \text{logistic}(\eta - \kappa_2) - 0 & \text{if } s = 3 \end{cases}$$

where the cutpoints  $\kappa_s < \kappa_{s+1}$  and  $\eta$  is a linear model.

[CORRECT THIS TO MATCH STAN CODE]

$S_i$	$\sim \text{OrderedLogistic}(\eta, \kappa)$	probability of data
$\eta_i$	$= (\beta + \beta_{site} + \beta_{prov} + \beta_{clone} + \beta_{year})f_i$	linear model
$\kappa_k$	$\sim \text{gamma}(20, 1)$	fixed priors
$\beta$	$\sim \text{exponential}(2)$	
$\beta_{site}$	$\sim \text{normal}(0, \sigma_{site})$	adaptive priors
$\beta_{prov}$	$\sim \text{normal}(0, \sigma_{prov})$	
$\beta_{year}$	$\sim \text{normal}(0, \sigma_{year})$	
$\beta_{clone}$	$\sim \text{normal}(0, \sigma_{clone})$	
$\sigma_{site}$	$\sim \text{exponential}(5)$	hyperpriors
$\sigma_{prov}$	$\sim \text{exponential}(5)$	
$\sigma_{year}$	$\sim \text{exponential}(5)$	
$\sigma_{clone}$	$\sim \text{exponential}(5)$	

The linear model  $\eta$  includes a population mean slope  $\beta$  with deviations from that for site, provenance, clone and year. Priors on  $\kappa$ s and  $\beta$  force the correct sign (positive) and order of magnitude. Priors on the site, provenance, clone, and year effects are adaptive and effects are estimated from the data with partial pooling.

Low variation for provenances and years made sampling inefficient, so I used a non-centered parameterization in the Stan model code for these clusters. The model was run [Z] chains for [X] iterations of which [Y] were warmup [See code].

## Model Assumptions

I assume that chilling requirements are always met and that chilling and forcing periods do not overlap. Forcing is a function of temperature only, not photoperiod or chilling. Transitions between each state occur at the same rate, *i.e.*  $\beta$  parameters cannot vary by phenological state. This model does not account for how very cold temperatures and very warm temperatures damage tissues and cause abnormal development. [Cite Sarvas's weird hot stuff and some late spring phenology]

## Model diagnostics

### Flowering period timing and length

I defined the beginning of the flowering period as the point at which a tree was 20% likely to have passed out of stage 1 and the end of the flowering period as the probability at which a tree was 80% likely to have passed out of stage 2. This is equivalent to 20% of the trees in the population having reached stage 2 and 80% having reached stage 3.

### Begin and end of flowering period

The accumulated forcing units required to reach the beginning of the flowering period,  $fbegin_i$ , is

$$\Pr(s_i > 1) = 0.2 = \text{logistic}((\beta + \beta_{site} + \beta_{prov} + \beta_{clone} + \beta_{year})f_i)$$

$$fbegin_i = \frac{\text{logit}(0.2)}{\beta + \beta_{site} + \beta_{prov} + \beta_{clone} + \beta_{year}}$$

The accumulated forcing units required to reach the end of the flowering period,  $fend_i$ , is

$$\Pr(s_i > 2) = 0.8 = \text{logistic}((\beta + \beta_{site} + \beta_{prov} + \beta_{clone} + \beta_{year})f_i - \kappa_1)$$

$$fend_i = \frac{\text{logit}(0.8) + \kappa_1}{\beta + \beta_{site} + \beta_{prov} + \beta_{clone} + \beta_{year}}$$

I calculated  $fstart$  and  $fend$  for all combinations of clone, site, provenance, and year that occur in the dataset for all samples from the posterior. The range of forcing units that stage 2 occurs over, i.e. the flowering period, is

$$frange = fend - fbegin$$

### Translating between accumulated forcing units and days

While  $fbegin$  and  $fend$  are constant for a given combination of clone, site, provenance, (and year???),  $dbegin$  and  $dend$  differ from year to year and location to location. For years and locations of interest, accumulated forcing units were calculated from mean daily temperatures and used to translate between accumulated forcing units and day of year. The day of year that the accumulated forcing unit reached at least  $fbegin$  or  $fend$  was  $dbegin$  or  $dend$ . The length of the phenological period was then calculated as

$$length = dend - dbegin$$

???Because  $\beta_{year}$  is included as an effect,  $fstart$  and  $fend$  are specific to the year the data was collected.???

### Model checking/Posterior predictive checks

To check the usefulness of the model, the data was compared to model predictions. I calculated the point at which at least 50% of trees of a given sex in a given year at a given site from a given provenance had begun flowering ( $f50_{obs}$ ). Then, 30 phenological states were simulated from the model at [Z] accumulated forcing units between [X] and [Y] for each set of parameter values (model configurations) sampled from the posterior distribution .

$$s'_i \sim \text{OrderedLogistic}(\eta_i, \kappa | s)$$

[Draw a diagram to explain this]

The predicted accumulated forcing at which at least 50% of the trees have begun flowering  $f50_{pred}$  is the lowest accumulated forcing where at least half of simulations return stage 2 or greater.

Then I checked to see whether  $f50_{obs}$  was in the 90% highest density interval of  $f50_{pred}$ .

### Average predictive comparisons

#### Differences between provenances

The point at which a tree is 50% likely to have transitioned from one phenophase to another can be calculated by dividing the cutpoint for the transition by the slope.

The halfway point between any 2 consecutive phenological states is

$$\Pr(s > i) = 0.5 f_{half_i} = \frac{\kappa_i}{\beta}$$

$$\begin{aligned}\mathbf{H_1} &= \kappa_1 / (\beta_{site} + \beta_{prov} + \beta_{clone} + \beta_{year}) \\ \mathbf{H_2} &= \kappa_2 / (\beta_{site} + \beta_{prov} + \beta_{clone} + \beta_{year})\end{aligned}$$

Were calculated according to (Vander Mijnsbrugge, Onkelinx, and De Cuyper 2015) - “difference in days in which half the plants has reached the same score level”

$$H50_{prov_i} - H50_{prov_j} = \beta_{prov_j} - \beta_{prov_2} / \beta$$

## Results

### Model parameters

#### fstart and fend

#### Posterior predictive checks

[See p. 81 in lab notebook for figure idea]

### Average Predictive Comparison

#### Provenance differences

[FIGURE SHOWING THE PROPORTION OF TREES THAT HAVE REACHED STAGE 2 AND STAGE 3 AT EACH SITE/PROV COMBINATION over forcing units]

- 1) How strong are the genetic vs. environmental effects on pollination phenology?

To answer this question I will compare the parameters of models fit by provenance. If different provenances have significantly different model parameters - that is, have different heat sum requirements for pollen shed and cone receptivity by provenance, then the rest of my project will be limited to the provenances I have data from. If not, I can use the same models across the entire range when predicting phenological events.

The budburst and spring pollination phenology of temperate tree species is determined by temperature, and there may be some regional variation in the temperatures required for pollen shed and cone receptivity; in

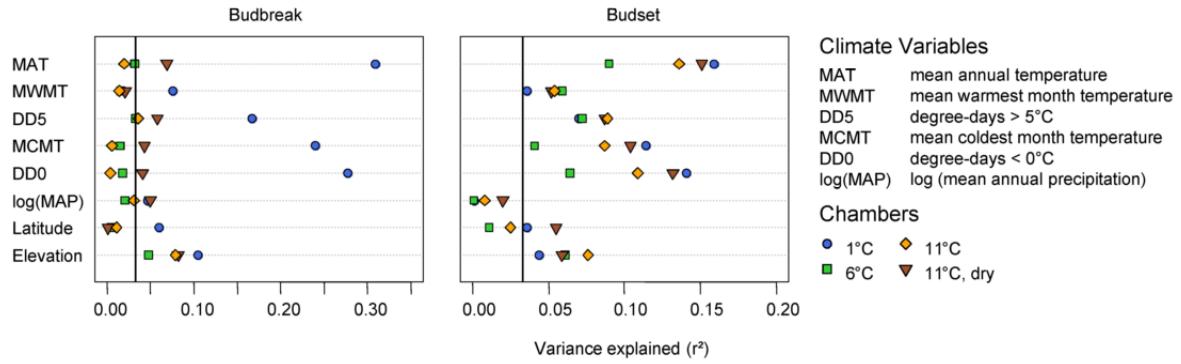


Figure 5: Variance in budbreak and budset for lodgepole pine explained by provenance climate. Seedlings from 281 locations across lodgepole's BC range were grown in growth chambers under four treatments: approximated seasonal and daily variation in temperature for geographic locations with a MAT of 1, 6, and 11 °C, with the warmest treatment having a dry and wet version. Pearson's correlation coefficient between phenotypic traits in lodgepole pine and provenance climates are presented. The vertical line represents the critical  $r^2$  value after an adjustment for multiple inferences. From Liepe (2014) (submitted).

lodgepole pine, the threshold for shoot elongation was 5.1 °C for coastal and interior provenances, but 6.5 °C for northern provenances (Chuine, Aitken, and Ying 2001).

However, seedling spring vegetative phenology data from four growth chamber temperature and moisture treatments in the AdapTree project suggests that budburst shows substantial clinal variation (> 10% of phenotypic variance explained) only in the coldest treatment tested (MAT of 1 °C), and much colder than the seed orchard locations (Figure 6, Figure 5, Liepe (2014) (submitted)). Analysis of phenotypic data from an outdoor common garden in Vancouver (*personal communication, Ian MacLachlan, AdapTree Graduate Student, January, 1 2015*) also shows that very little variation exists among provenances for spring phenology, regardless of whether seedlings originate from wild-stand or selectively bred seed orchard seedlots.

- 2) How does faster spring warm-up like we expect under climate change affect within population mating success and frequency of outcrossing?

I expect that faster spring warm-up will condense as well as advance growth initiation dates and also shorten pollination phenological periods by speeding development prior to and during the phenological period. This should further synchronize populations in similar climates and act as a barrier to breeding with populations in more different climates. On the other hand, it could favor outcrossing over selfing by causing protandry - pollen shed in advance of cone receptivity. Based on anecdotal reports from seed orchards managers, I expect protandry (and thus outcrossing) to increase under climate change.

I will examine the variability of the start date across the range ( $t_1$ ) to see if it shrinks in warmer years and as the climate changes. I will also calculate overlap between within population cone receptivity and pollen shed under current conditions and with climate warming and see if the heat sum requirements for cone receptivity and pollen shed are different.

## Model summary

```
# ffit.stan <- readRDS("../FEMALE_slopes_scaled.rds")
# mfit.stan <- readRDS("../MALE_slopes_scaled.rds")
#
# fsum <- rstan::summary(ffit.stan)$summary
# fsum <- as.data.frame(fsum)
```

```

# fsum$par <- rownames(fsum)
# fpost <- as.data.frame(ffit.stan)
# fpost_re <- extract.samples(ffit.stan)
# fparam_names <- colnames(fpost)
#
# flp <- log_posterior(ffit.stan)
# fnp <- nuts_params(ffit.stan) #nuts params
#
# msum <- rstan::summary(mfit.stan)$summary
# msum <- as.data.frame(msum)
# msum$par <- rownames(msum)
# mpost <- as.data.frame(mfit.stan)
# mpost_re <- extract.samples(mfit.stan)
# mparam_names <- colnames(mpost)
#
# mlp <- log_posterior(mfit.stan)
# mnp <- nuts_params(mfit.stan) #nuts params
#
# #mf
# fpost$sex <- 0
# mpost$sex <- 1
#
#
# post <- full_join(mpost, fpost) %>%
#   select(-`b_clone[260]`)

```

## Parameter estimates

Clones not included

*Female*

```

# kable(
# fsum[str_detect(fsum$par, "b_clone", negate=TRUE) , ], digits=3)

```

*Male*

```

# kable(
# msum[str_detect(msum$par, "b_clone", negate=TRUE) , ], digits=3)

# compare_fm <- function(femplot, mplot, nrow = 2, ...) {
#   bayesplot_grid(
#     femplot, mplot,
#     grid_args = list(nrow = nrow),
#     subtitles = c("Female",
#                  "Male"),
#     ...
#   )
# }
#
# compare_fm_wide <- function(femplot, mplot, ncol = 2, ...) {
#   bayesplot_grid(
#     femplot, mplot,
#     grid_args = list(ncol = ncol),

```

```

#           subtitles = c("Female",
#                               "Male"),
#           ...
#       )
# }

# ggplot(post, aes(x=beta, fill=as.factor(sex))) +
#     geom_density(alpha=0.5) +
#     ggtitle("male and female speed of transition estimates")
#
# fint_beta <- mcmc_areas(fpost, regex_pars = "beta") + ggtitle("beta")
# mint_beta <- mcmc_areas(mpost, regex_pars = "beta")
#
# compare_fm(fint_beta, mint_beta, xlim=c(0.05, 0.10))
#
# fkap <- mcmc_areas(fpost, regex_pars = "kappa")
# mkap <- mcmc_areas(mpost, regex_pars = "kappa")
#
# compare_fm_wide(fkap, mkap, xlim=c(18, 29))
#
# color_scheme_set("blue")
# fint_site <- mcmc_intervals(fpost, regex_pars = c("site", "prov"))
# mint_site <- mcmc_intervals(mpost, regex_pars = c("site", "prov"))
#
# compare_fm_wide(fint_site, mint_site)
#
# fa_sigma <- mcmc_areas(fpost, regex_pars= "sigma") + ggtitle("variance")
# ma_sigma <- mcmc_areas(mpost, regex_pars= "sigma")
#
# compare_fm_wide(fa_sigma, ma_sigma)
#
# fa_y <- mcmc_areas(fpost, regex_pars = "year") + ggtitle("year")
# ma_y <- mcmc_areas(mpost, regex_pars = "year")
#
# compare_fm_wide(fa_y, ma_y)

# fint_c <- mcmc_intervals(fpost, regex_pars = c("clone"), point_est="none") + ggtitle("Clone effects")
# mint_c <- mcmc_intervals(mpost, regex_pars = c("clone"), point_est="none")
#
# compare_fm_wide(fint_c, mint_c)
# ``
#
# # ### Sex diff
#
# ``{r half transitions - no effects}
# fpost$f1 <- fpost$kappa[1]/fpost$beta
# fpost$f2 <- fpost$kappa[2]/fpost$beta
# mpost$f1 <- mpost$kappa[1]/mpost$beta
# mpost$f2 <- mpost$kappa[2]/mpost$beta
#
# ff <- mcmc_areas(fpost, pars = c("f1", "f2")) + ggtitle("half transitions")
# mf <- mcmc_areas(mpost, pars = c("f1", "f2")) + ggtitle("half transitions")
#
# compare_fm(ff, mf)

```

# Discussion

## Limitations of the dataset

Benefits of this dataset include the large number of trees, long time series, multiple provenances grown at multiple sites giving a semi-common garden design, and the inclusion of clones. Limitations include interval and end censoring, especially at Prince George, irregular scoring systems, subjective scoring, non-random clone selection, and selective breeding.

## Grafting bias?

## Problems from breeding

There are two types of orchards: first generation and advanced. Each orchard contains trees that are the descendants of seeds and scions collected from mother trees in one particular provenance. First generation orchards have not undergone selective breeding but may be subject to selection through the process of choosing mother trees (called “plus” trees by tree breeders), testing and selection of mother tree offspring, or culling of unfavorable trees in the orchard. First generation orchard trees are clones of mother trees and their offspring grafted onto rootstock. There are three types of first generation orchards subject to different levels of selection: 1.0, 1.5, and 1.75 generation orchards. All orchards are subject to selection based on mother tree selection. In 1.5 generation orchards, superior offspring of mother trees are selected for the orchard or poor trees are culled from the orchard. This selection may be strong; more than 50% of offspring can be rejected in this process due to poor growth form (Ukrainetz 2011). In 1.75 generation orchards, data from 10 year tests of mother tree offspring are used to select 40 of the best genotypes for a given provenance. Offspring from controlled crosses between trees in first generation orchards are tested in field trials and the best trees are then cloned and grafted onto root stock in advanced generation orchards.

```
orchards_in_data <- read.csv("~/Documents/research_phd/data/PhenologyAndPollenCounts/data_formatted_and_cleaned.csv")
  select(Orchard, Year, ) %>%
  group_by(Orchard) %>%
  dplyr::summarise(Years = n_distinct(Year))
orchard_gen <- read.csv("~/Documents/research_phd/data/OrrchardInfo/OrrchardGen.csv", header=TRUE, stringsAsFactors=FALSE)

orchard_meta <- dplyr::left_join(orchards_in_data, orchard_gen) %>%
  unique()

## Joining, by = "Orchard"
knitr::kable(orchard_meta)
```

Orchard	Years	Generation
218	2	1.5
219	1	1.5
220	15	1.75
222	2	1.5
223	14	1.75
228	15	1.75
230	4	1.5
234	2	Advanced
237	3	Advanced
240	1	Unknown
307	4	1.75
308	1	1.75
310	1	1.5

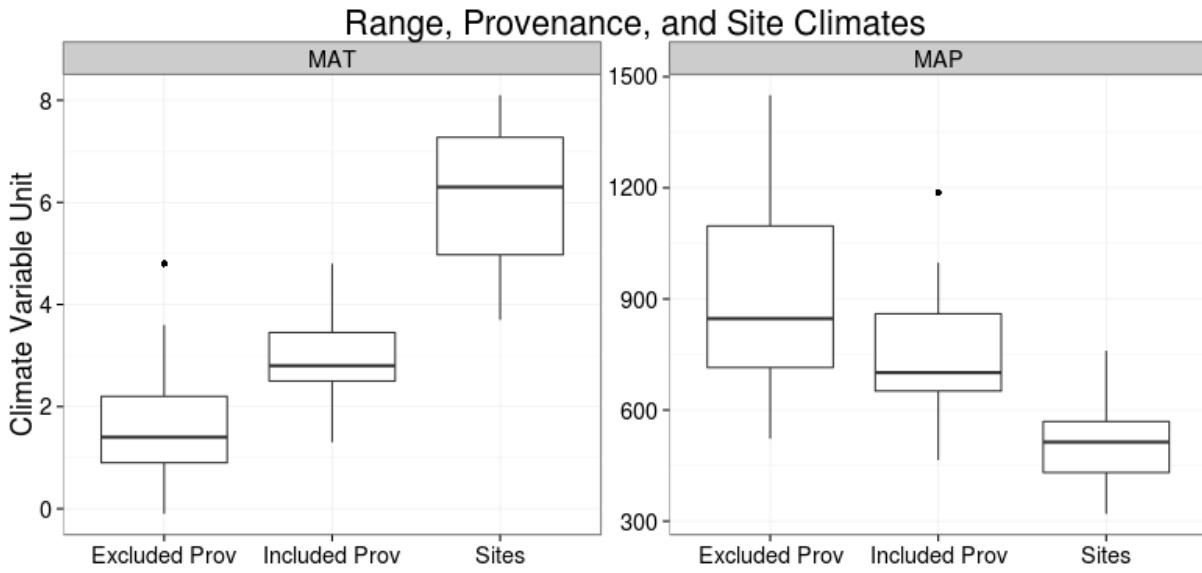


Figure 6: Provenance (SPU) climates of data included in this proposal, provenances not included in this proposal, and sites where provenances were grown. MAT is mean annual temperature and MAP is mean annual precipitation.

Orchard	Years	Generation
311	1	1.5
313	1	1.5
338	2	Advanced
339	2	1

#### selection only from a portion of the range

Provenances and sites included in my data exclude the coldest and wettest parts of the range. Figure 6 shows the 1961-1990 climate normal mean annual temperature (MAT) and mean annual precipitation (MAP) for data from all lodgepole pine provenances in BC, provenances included in my data set, and sites where provenances were grown. Orchard sites are generally much warmer and drier than many locations where lodgepole pine grow. I may have difficulty extrapolating into the coolest and wettest parts of the lodgepole pine range, depending on actual yearly conditions at sites for which I have data. However, this may be advantageous when it comes to predicting phenology under climate change.

#### no chilling

Previous models for lodgepole pine have fitted spring growth phenology models without considering chilling (Chuine, Aitken, and Ying 2001) and, in general, budburst models for boreal tree species have not required chilling to give accurate predictions (Linkosalo 2000). For lodgepole pine growing in natural environments, chilling is always met and should continue to be met under the next century of climate change; in the AdapTree project, lodgepole pine seedlings grown for three to four weeks at 4 °C met their chilling requirements. Thus, I expect to be able to disregard the state of chilling and fit only Equations ?? and ??). Once the model is parameterized I will be able to calculate the timing of pollen shed and cone receptivity across the lodgepole pine species range using location specific climate data from seed orchards and ClimateWNA (Wang et al.

2012). I will validate the model by comparing predictions to phenological measurements from the National Phenology Network.

## Acknowledgements

Thanks to R. Johnstone for help with transcription. GySgt T. Kerr, Ret. provided invaluable commentary on an earlier draft.

Boyer, WD. 1978. "Heat Accumulation: An Easy Way to Anticipate the Flowering of Southern Pines." *Journal of Forestry*, no. January: 20–23.

Chuine, I, K Kramer, and H Hanninen. 2003. "Plant Development Models." In *Phenology: An Integrative Environmental Science*, edited by Mark Schwartz, 2nd ed., 217–35. Springer.

Chuine, Isabelle, Sally N. Aitken, and Cheng C. Ying. 2001. "Temperature Thresholds of Shoot Elongation in Provenances of &lt;I>Pinus Contorta</I>;" *Canadian Journal of Forest Research* 31 (8): 1444–55. <https://doi.org/10.1139/cjfr-31-8-1444>.

Chuine, Isabelle, P. Cour, and D. D. Rousseau. 1999. "Selecting Models to Predict the Timing of Flowering of Temperate Trees: Implications for Tree Phenology Modelling." *Plant, Cell and Environment* 22 (1): 1–13. <https://doi.org/10.1046/j.1365-3040.1999.00395.x>.

Consortium, Pacific Climate Impacts. 2014. "PNWNAmet (1945-2012)."

Ennos, R A. 1994. "Estimating the Relative Rates of Pollen and Seed Migration Among Plant Populations." *Heredity* 72 (August 1993): 250–59.

Franks, S J, J J Weber, and Sally N Aitken. 2014. "Evolutionary and Plastic Responses to Climate Change in Terrestrial Plant Populations." *Evolutionary Applications* 7: 123–39.

Hamann, Andreas, and Tongli Wang. 2006. "Potential Effects of Climate Change on Ecosystem and Tree Species Distribution in British Columbia." *Ecology* 87 (11): 2773–86. [https://doi.org/10.1890/0012-9658\(2006\)87\[2773:PEOCOC\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[2773:PEOCOC]2.0.CO;2).

Hänninen, Heikki. 1990. "Modelling Bud Dormancy Release in Trees from Cool and Temperate Regions." <https://doi.org/10/gfxh7x>.

Inouye, David W. 2008. "Effects of Climate Change on Phenology, Frost Damage, and Floral Abundance of Montane EFFECTS OF CLIMATE CHANGE ON PHENOLOGY, FROST DAMAGE, AND FLORAL ABUNDANCE OF MONTANE WILDFLOWERS." *Ecology* 89 (2): 353–62. <https://doi.org/10.1890/06-2128.1>.

Liepe, Katharina J. 2014. "Genetic Variation in Lodgepole Pine and Interior Spruce: Adaptation to Climate and Implications for Seed Transfer." PhD thesis, University of Alberta.

Linkosalo, T. 2000. "Mutual Regularity of Spring Phenology of Some Boreal Tree Species: Predicting with Other Species and Phenological Models." *Canadian Journal of Forest Research* 30 (5): 667–73. <https://doi.org/10.1139/x99-243>.

Little E.L., Jr. 1971. *Atlas of United States Trees, Volume 1, Conifer and Imprtant Hardwoods*. U.S. Department of Agriculture.

Owens, John N, Jordan Bennett, and Sylvia L'Hirondelle. 2005. "Pollination and Cone Morphology Affect Cone and Seed Production in Lodgepole Pine Seed Orchards." *Canadian Journal of Forest Research* 35 (2): 383–400. <https://doi.org/10.1139/x04-176>.

Parmesan, Camille. 2006. "Ecological and Evolutionary Responses to Recent Climate Change." *Annual Review of Ecology, Evolution, and Systematics* 37 (1): 637–69. <https://doi.org/10.1146/annurev.ecolsys.37.091305.110100>.

- Pau, Stephanie, Elizabeth M. Wolkovich, Benjamin I. Cook, T. Jonathan Davies, Nathan J.B. Kraft, Kjell Bolmgren, Julio L. Betancourt, and Elsa E. Cleland. 2011. "Predicting Phenology by Integrating Ecology, Evolution and Climate Science." *Global Change Biology* 17 (12): 3633–43. <https://doi.org/10.1111/j.1365-2486.2011.02515.x>.
- PRUS, TADEUSZ. 1971. "The Assimilation Efficiency of *Asellus Aquaticus* L. (Crustacea, Isopoda)." *Freshwater Biology* 1 (3): 287–305. <https://doi.org/10.1111/j.1365-2427.1971.tb01564.x>.
- Sambaraju, Kishan R., Allan L. Carroll, Jun Zhu, Kerstin Stahl, R. Dan Moore, and Brian H. Aukema. 2012. "Climate Change Could Alter the Distribution of Mountain Pine Beetle Outbreaks in Western Canada." *Ecography* 35 (3): 211–23. <https://doi.org/10.1111/j.1600-0587.2011.06847.x>.
- Sarvas, Risto. 1972. "Investigations on the Annual Cycle of Development of Forest Trees. Active Period." *Metsantutkimuslaitoksen Julkaisuja* 76 (3): 1–110.
- Schneider, Richard R., Maria Cecilia Latham, Brad Stelfox, Dan Farr, and Stan Boutin. 2010. "Effects of a Severe Mountain Pine Beetle Epidemic in Western Alberta, Canada Under Two Forest Management Scenarios." *International Journal of Forestry Research* 2010: 1–7. <https://doi.org/10.1155/2010/417595>.
- Schuster, William S., David L. Alles, and Jeffry B. Mitton. 1989. "Gene Flow in Limber Pine: Evidence from Pollination Phenology and Genetic Differentiation Along an Elevational Transect." *American Journal of Botany* 76 (9): 1395. <https://doi.org/10.2307/2444563>.
- Team, Stan Development. 2018. "RStan: The R Interface to Stan."
- Ukrainetz, Nicholas. 2011. "Program History: Lodgepole Pine Program Review."
- Vander Mijnsbrugge, Kristine, Thierry Onkelinx, and Bart De Cuyper. 2015. "Variation in Bud Burst and Flower Opening Responses of Local Versus Non-Local Provenances of Hawthorn (*Crataegus Monogyna* Jacq.) in Belgium." *Plant Systematics and Evolution* 301 (4): 1171–9. <https://doi.org/10.f66qzm>.
- Wang, Tongli, Andreas Hamann, David L. Spittlehouse, and Trevor Q. Murdock. 2012. "ClimateWNA-High-Resolution Spatial Climate Data for Western North America." *Journal of Applied Meteorology and Climatology* 51 (1): 16–29. <https://doi.org/10.1175/JAMC-D-11-043.1>.
- Woods, JH, MU Stoehr, and JE Webber. 1996. "Protocols for Rating Seed Orchard Seedlots in British Columbia. Research Report 06." Victoria, BC: Province of British Columbia, Ministry of Forests Research Program.
- Yeaman, Sam, and Andy Jarvis. 2006. "Regional Heterogeneity and Gene Flow Maintain Variance in a Quantitative Trait Within Populations of Lodgepole Pine." *Proceedings of the Royal Society B: Biological Sciences* 273 (1594): 1587–93. <https://doi.org/10.1098/rspb.2006.3498>.