Adapting a process based cold hardiness model to conifers

Miro Stuke

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2022

Committee:

Soo-Hyung Kim

Gregory Ettl

Constance Harrington

Program Authorized to Offer Degree:

School of Environmental and Forest Sciences

University of Washington

Abstract

Adapting a process based cold hardiness model to conifers

Miro Stuke

Chair of the Supervisory Committee:

Soo-Hyung Kim

School of Environmental and Forest Sciences

Plant phenology has been and continues to be impacted by climate change. Process-based modeling of phenology reveals biological characteristics through interpretation of model results and parameter values. This paper aims to implement a cold hardiness model using historical Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) data, determine if model results improve with data clustering by seed source or growing location, and interpret model results into biological meaning. These interpretations have applications for reforestation and studies of plant phenological response to climate change.

Cold hardiness data were compiled through literature review. A process-based cold hardiness model using daily temperature inputs was applied to multiple data clustering scenarios. Model results were analyzed for goodness of fit to determine error, efficiency, and bias. A sensitivity analysis and cross validation were performed to determine parameter sensitivity, model bias, and variance.

Data clustering by seed source improved fit compared to clustering by growing location or no clustering when applied to the full dataset. Using only temperature inputs, model results had low error when data modeled were similar. Results show that for cold hardiness acclimation a linear growing degree function with a threshold of 10°C was adequate across testing data, as was a maximum cold hardiness temperature of -3°C.

Interpretation of model results show that both acclimation to growing sites and seed source genetics impact cold hardiness response, though clustering by seed source improved model performance. This model can be applied to mitigate cold related risks to seedlings during production and establishment, and can be a template for predicting phenological responses in simulated future climate scenarios.

Keywords: Cold Hardiness, Phenology, Douglas-fir, Process-based model

Background

Plant phenology is the timing of cyclical life events such as budburst and flowering. The timing of phenological events is important in determining growing season duration and coordination of tissue development and tolerance to seasonal changes in the growing environment. Changes in plant phenology have been a method of observing ecosystem and biological effects of climate change (Field & Barros, 2014; Peñuelas et al., 2009). Major phenological features of tree species, cold hardiness in winter, and bud burst timing in spring, are critical factors in forest health, productivity, and ecosystem resilience, and are impacted by warming temperatures (Bansal et al., 2015; Ettinger et al., 2020). The ability of trees and seedlings to survive cold exposure and to initiate vegetative growth in the spring can determine success or failure of reforestation, restoration, production, and survival (Krasowski & Simpson, 2001).

Plant cold hardiness indicates the ability of a plant tissue to survive exposure to cold and freezing temperatures without damage. In some conifers maximum cold hardiness has been documented to be below -196°C, the temperature of liquid nitrogen (Sakai & Weiser, 1973). The maximum cold hardiness of conifer species is correlated with the USDA climatic zones and average minimum temperature (Bannister & Neuner, 2001), indicating that climate factors impact the development of cold hardiness, through either evolutionary genetic adaptation over millennia, or site based phenotypic acclimation. The seasonal cycle of cold hardiness can be visualized by the general curve pattern in Figure 1, which shows cold hardiness as the lethal temperature of 50% of the population (LT_{50}). The seasonal cycle begins in the fall with acclimation, where tissues become progressively more cold hardy based on day length and

chilling temperatures. Peak cold hardiness occurs during the coldest winter months, followed by deacclimation, when cold hardiness is rapidly reduced in relation to warming temperatures which also lead to bud burst and shoot elongation in the spring. (Bigras et al., 2001; Harrington et al., 2010). The cold hardiness acclimation period in conifers is controlled by temperature and photoperiod (G. Zhang et al., 2003), while deacclimation in the spring is primarily controlled by temperature (Harrington & Gould, 2015).

Spring bud burst and shoot elongation in conifers signal the initiation of vegetative growth during the growing season. Bud burst is a key determinant of growing season length which impacts net primary productivity (Richardson et al., 2013), an important consideration for both ecosystem services and carbon relations. The progression towards bud burst, similar to cold hardiness, is driven by a combination of chilling and forcing temperatures (Harrington & Gould, 2015). Temperature changes related to climate change have modified the timing and duration of these transitions, with reductions in chilling being the most important driver across temperate tree species (Ettinger et al., 2020). Seasonality is predicted to change more in the future, including shorter winters, springs, and autumns (J. Wang et al., 2021), further impacting plant phenology and limiting chilling potential. As bud burst and shoot elongation occur, cold hardiness is at a minimum. Thus, a mismatch in the timing of budburst and seasonal climate, due to rapid changes in climate, can result in potential negative outcomes. These include increased likelihood of cold damage due to early deacclimation (Arora & Taulavuori, 2016; Wisniewski et al., 2018), or delayed budburst due to lower accumulation of chilling temperatures (Ettinger et al., 2020; R. Zhang et al., 2021). The likelihood of these impacts varies by population geography, continentality, and elevation (Ford et al., 2016).

Climate change, especially spring warming, has accelerated bud burst timing, but this rate of acceleration has declined recently due to winter warming (Ettinger et al., 2020; Fu et al., 2015). Modified phenology has additional feedbacks on climate systems which impact ecosystem productivity (Fu et al., 2015; Peñuelas et al., 2009). Terrestrial carbon sinks are the primary source of uncertainty in global carbon models, due to the complexity of direct effects and feedbacks including CO_2 fertilization, nitrogen deposition, increased respiration rates, and changes in growing season and phenology (Eyring et al., 2016; Tharammal et al., 2019; Zhou et al., 2018). Terrestrial carbon models could be improved by better understanding of phenological mechanisms (Piao et al., 2019).

Researching the mechanisms of cold hardiness and bud burst will improve our understanding of the impact of climate change on these phenological events. Additionally, improved quantification of chilling and forcing requirements can inform material selection for assisted migration (Boiffin et al., 2017; Chakraborty et al., 2019; Isaac-Renton et al., 2014; Malmqvist et al., 2018). Assisted migration of temperate tree species is intended to improve success and productivity with changing climate but may create additional risks of cold damage both during seedling production and after outplanting as less cold hardy populations shift north or up elevation.

Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) cold hardiness during deacclimation is highly correlated with the timing of spring bud burst (Aitken & Adams, 1997), and several functions have been proposed to model cold hardiness (Greer et al., 2001; Leinonen, 1996; Leinonen et al., 1995; Timmis et al., 1994) and bud burst (Hänninen & Kramer, 2007; Harrington et al., 2010; Harrington & Gould, 2015). Many of the models developed for conifer cold hardiness were based on controlled environment data without natural system verification

(Leinonen et al., 1995; G. Zhang et al., 2003). Refinement and utilization of these functions into a cohesive model for both cold hardiness and bud burst would provide a better framework for understanding the impact of climate change on phenology for the productive forests found in the Pacific Northwest. Model based understanding of dormancy phenology, including cold hardiness, can be a tool for regeneration decision making and to analyze the potential impacts of climate change these processes.

Douglas-fir provides an ideal model species for Pacific Northwest conifers because it is a dominant forest species in natural and managed forest systems, has a broad range, has been used for assisted migration trials, has high economic and social value, and is a species that has a depth of historical research (Bansal et al., 2015; Boiffin et al., 2017; Wisniewski et al., 2018). Across the range of Douglas-fir, some populations are at increased risk of early deacclimation and frost exposure, and at risk of delayed growth, reducing total growth during months with greater soil moisture availability and lower vapor pressure deficit, though this is regionally specific (Ford et al., 2016). Forest management practices may contribute to this risk by reduced biodiversity of artificial regeneration, species exclusion, monoculture, and seed source selection, vegetation management, and reduced canopy cover during seedling establishment which can contribute to lower temperatures at the groundline (Krasowski & Simpson, 2001; Pawson et al., 2013). Post wildfire regeneration can also provide challenges to regeneration success and seedling survival, with heightened risks due to topography driven microclimates and environmental extremes which may include cold related damage (Krasowski & Simpson, 2001; Marsh et al., 2022).

In broad terms, this work aims to fill a gap in cold hardiness modeling by applying and testing a model of cold hardiness in Douglas-fir using diverse data collected in situ throughout the full dormancy season. This is a novel approach because previous phenological modeling of

Douglas-fir has been tested primarily on controlled environment data, individual trials have had limited range in terms of genetics, there has been little to no temporal and spatial replication in previous model testing, or observations have not included the full season of cold hardiness phenology with a greater availability of cold hardiness acclimation data than deacclimation data. The intent of this approach is to derive a model that can be applied to the full dormancy season by practitioners and can also be used for in silico simulations of phenological changes under future climate scenarios RCMP 4 and RCMP 8.5 (Allan et al., 2021).

The relevance of this work in an applied context can clearly be demonstrated by the recent passing of the REPLANT act by the United States federal legislature. This act will invest $123 million per year over the next 10 years into artificial regeneration of 1.2 billion trees on degraded federal forest lands (Stabenow, 2021). Two outcomes of this increased demand in seedlings may enhance the necessity of understanding seedling cold hardiness. First is greater planting volume in a biologically limited window, which may result in planting outside of optimal conditions, including increased fall planting which can have greater risk of cold damage in nursery cultivated seedlings if they have not received adequate hardening (Grossnickle & MacDonald, 2021). Second is production of seedlings geographically and climatically distant from the seed source origin. This issue is already observed within the current seedling demand (Fargione et al., 2021). Both scenarios run increased risk of issues relating to cold hardiness and dormancy phenology.

From a modeling perspective, the timing of this work is also relevant. Changes in bud burst phenology due to anthropogenic climate change has been observable for several decades, but in recent years advancement has slowed due to reduced chilling (Ettinger et al., 2020). Additionally, fire exclusion management practices on North American forested land of the past

two centuries has resulted in increased fire severity. As a result, we are seeing an increase in regeneration, both artificial and natural, on exposed sites (Marsh et al., 2022). Understanding all risk to regeneration success in real time and future risks is critical to decision making around seed source selection, especially in cases of assisted migration, and prioritization of management intervention. Development and testing of a universal process-based cold hardiness model will create a tool applicable to both current and future climates.

The specific aim of this project is to calibrate and test a process-based model of cold hardiness for conifers, using data collected from historical literature, to determine the efficacy of phenological modeling of cold hardiness in a model species, Douglas-fir. This model can provide a template for further development of a unified cold hardiness and bud burst phenology model for conifers and show that a model of cold hardiness that uses temperature as the only predictor, can simulate Douglas-fir cold hardiness. Applying this model to data clustered by seed source origin, or cultivation site, can reveal whether genetic traits related to seed source or phenotypically plastic traits driven by site acclimation provide better predictions of cold hardiness phenology. Testing data clustered by seed source or growing site characteristics can increase the size of the modeling data set while also revealing whether clustering by these characteristics will improve model performance. Determining the model sensitivity to the chilling and forcing temperature parameters can identify uncertainty in cold-hardiness prediction. The predictive capability of the cold hardiness model will improve when data are clustered into more similar groups. Inferences made by model performance in these aspects can be used to determine sources of uncertainty in phenological development under warming climate conditions, and predict risks associated with ongoing and future changes in cold hardiness phenology.

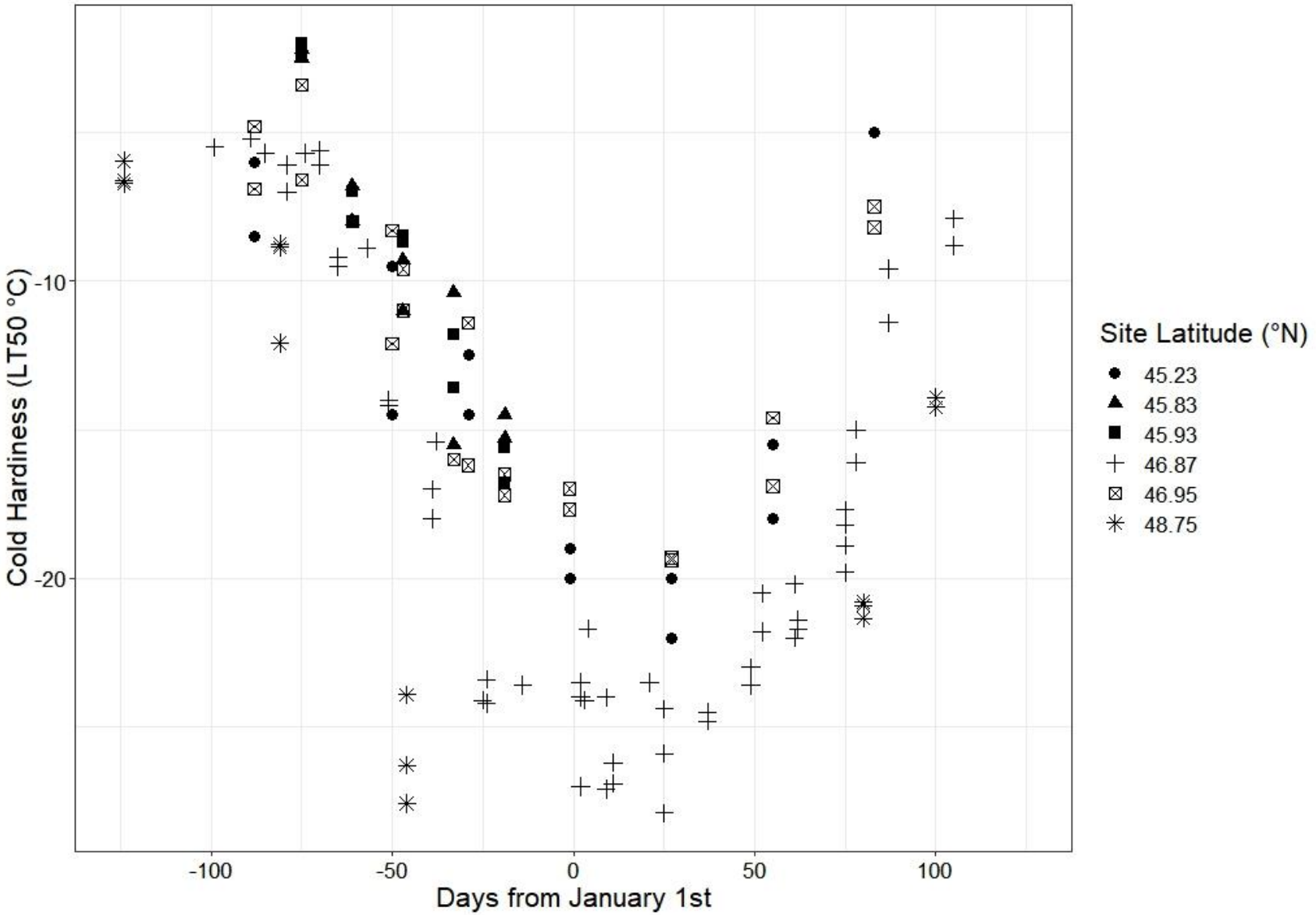

Figure 1. Cold hardiness data compiled from multiple sources (x-axis). Y-axis shows date depicted as days from January first, to allow plotting of many years. Point shape indicates latitude of observation growing locations. Negative slope on the left shows the acclimation phase and positive slope on the right shows the deacclimation phase of cold hardiness phenology.

Materials and Methods

Cold Hardiness Data Collection and Clustering

Cold hardiness observational data were collected from published articles, and publicly available thesis reports. In total, twenty-one potential sources were identified for data extraction (Supplementary Table S1). Of the these, data were extracted from five sources (Fisker, 1992; Haase et al., 2016; Leinonen et al., 1995; Stevenson et al., 1999; Timmis et al., 1994), based on replicated cold hardiness testing dates and populations, utilization of standard units for cold hardiness, and data availability within published works (Table 1). Several papers only tested cold hardiness on limited dates and were not ideal for full season model calibration. Some papers reported cold hardiness as damage, or in raw terms of chlorophyll fluorescence and could not be conformed to the phenology model based on an extrapolated metric, lethal temperature of 50% of samples (LT_{50}). The total number of cold hardiness observations compiled was 150 (Fig. 1).

Cold hardiness observations were extracted from data tables or figures. Data extraction from figures was performed with the ImageJ (Schneider et al., 2012) 'Figure Calibration' plugin (Miller, 2011). Compiled data from four of the sources includes eleven distinct seed sources cultivated at one of six sites representing populations from the west of the Cascades between 43° and 51°N (Table 1, Fig. 2). One source utilized an interior seed source from New Mexico which was tested under controlled environments (Leinonen et al., 1995) and was excluded for being dissimilar to other data sources. Of the remaining four sources, three sampled second year seedlings that were cultivated at nurseries. The other data source sampled seedlings outplanted in a realized gain trial. These seedlings were cultivated in containers for one year and outplanted three years before sampling occurred (Stevenson et al., 1999). In all sources cold hardiness was determined using artificial freeze tests and calculation of the LT_{50}. The method of assessing

Table 1. List of data sources that cold hardiness data used for modeling were extracted from. For Fisker (1993) each seed source was cultivated at both growing sites for a single season. For Timmis et al. (1994) two seed sources were cultivated at one site over four seasons, three seasons of data were used. For Haase et al. (2016) each seed source was site specific, two seed sources were cultivated per site. For Stevenson et al. (1999) one population with three orchard production methods was grown in the field for three seasons prior to testing and the data used for modeling encompasses different seed production methods. See supplemental table S1 for complete list of sources reviewed for data availability.

Data Source	Growing Site Latitude (°N)	Growing Site Longitude (°W)	Seed Source Latitude (°N)	Seed Source Longitude (°W)	Seed Source Elevation (feet)	Number of Data Points
Fisker, Sue. (1993). Chlorophyll fluorescence as a measure of cold hardiness and freezing stress in 1 + 1 Douglas-fir seedlings: response to seasonal changes in the nursery [Master's Thesis, Oregon State University]. ScholarsArchive@OSU.*	46.95	-122.95	46.75	-123.33	1500	28
	45.23	-122.76	43.33	-124.31	50	
Timmis, R., Flewelling, J., & Talbert, C. (1994). Frost injury prediction model for Douglas-fir seedlings in the Pacific Northwest. Tree Physiology, 14(7-8–9), 855–869.	46.87	-123.08	46.11	-122.54	2070	53
			46.93	-123.81	315	
Haase, D. L., Khadduri, N., Mason, E., & Dumroese, R. K. (2016). Relationships among chilling hours, photoperiod, calendar date, cold hardiness, seed source, and storage of Douglas-fir seedlings. Tree Planters' Notes, 59(1), 12.	46.95	-122.95	46.79	-120.72	3500	30
			45.77	-122.40	1000	
	45.83	-122.74	45.23	-122.16	2200	
			45.55	-123.85	1000	
	45.93	-122.64	44.05	-122.74	2000	
			44.01	-122.98	500	
Stevenson, J. F., Hawkins, B. J., & Woods, J. H. (1999). Spring and Fall Cold Hardiness in Wild and Selected Seed Sources of Coastal Douglas-fir. Silvae Genetica, 48(1), 29-34.	48.75	-123.68	49.49	-123.81	350	12

*1+1 seedling refers to the production method.

sample damage to determine LT_{50} varied between sources including visual assessment, and chlorophyll fluorescence. LT_{50} is calculated by scoring damage of subsamples after exposure to multiple freezing temperatures and estimating the 50^{th} percentile by fitting damage score results (Anekonda et al., 2000). The method for fitting damage score results can vary between publications.

Model calibration, to select optimized parameters, is ideally performed using three years of replicated data for each population (Ferguson et al., 2011, 2014), before any data are tested. With the historical data compiled for this project, that was not possible since seed source replication was limited. To overcome this limitation, data sets were aggregated to increase the total number of data points for model calibration and reduced the total number of calibrations and resulting parameter sets. The number of clusters per aggregation scenario was limited to two, to maintain data set size and reduce number of model calibrations. Three separate approaches were tested for data aggregation.

1. All data sets were combined into a single set. The geographic range of seed sources included in this data set stretched from 43° to 51°N and from elevations from 15 to over 1000 meters. In addition to the geographic data, twenty-five climate normal factors from each seed source origin site were compiled from ClimateNA (T. Wang et al., 2016). Literature shows that there are differences in cold hardiness response based on seed source origin (Aitken et al., 1996; Aitken & Adams, 1997; Bansal et al., 2015; St. Clair, 2006), suggesting this is a coarse approach for aggregation since different populations tested at the same time and location are not differentiated.

2. Data sets were aggregated by bioclimatic factors from the origin of seed sources into two data sets. For each seed source origin site, the location, elevation, and twenty-five climate factors were used in a cluster analysis in R (R Core Team, 2022). All variables were scaled to standardize them. Clusters were made by reducing the sum of squares within groups using the kmeans function and the Hartigan and Wong algorithm (Hartigan & Wong, 1979). Creating two clusters reduced within group sum of squares from 280 to 147 (Fig. 2). Of the two clusters formed, one included four seed sources, all from an elevation of 450 meters or above, and the second included seven seed sources, six of which were from elevations of 300 meters or below. Though the groups were not completely split by elevation, it appears to be an important factor in clustering the data.

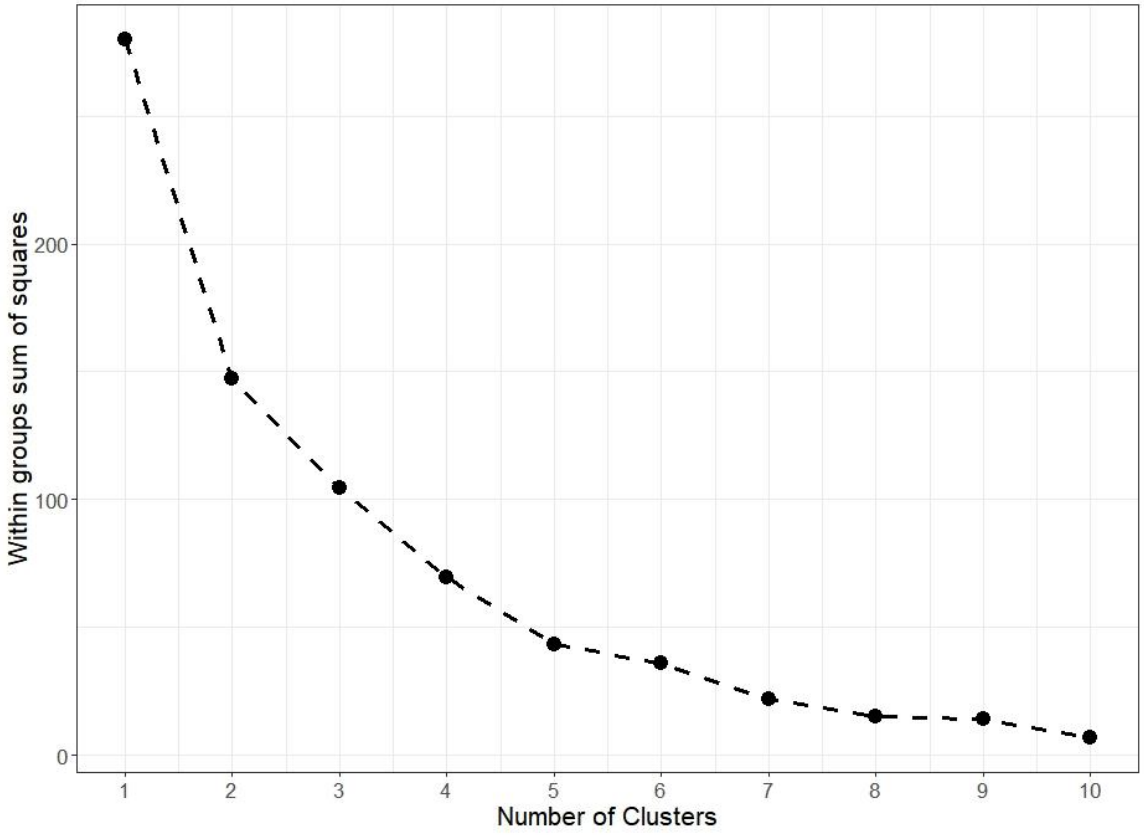

Figure 2. Effect of number of data clusters based on seed source climate factors on sum of squares within each cluster. Two clusters were formed for the seed source clustering scenario. The largest decrease in sum of squares was between 1 cluster (all data combined) and two clusters. Though decrease continued with more clusters, two clusters were used to maintain adequate dataset size.

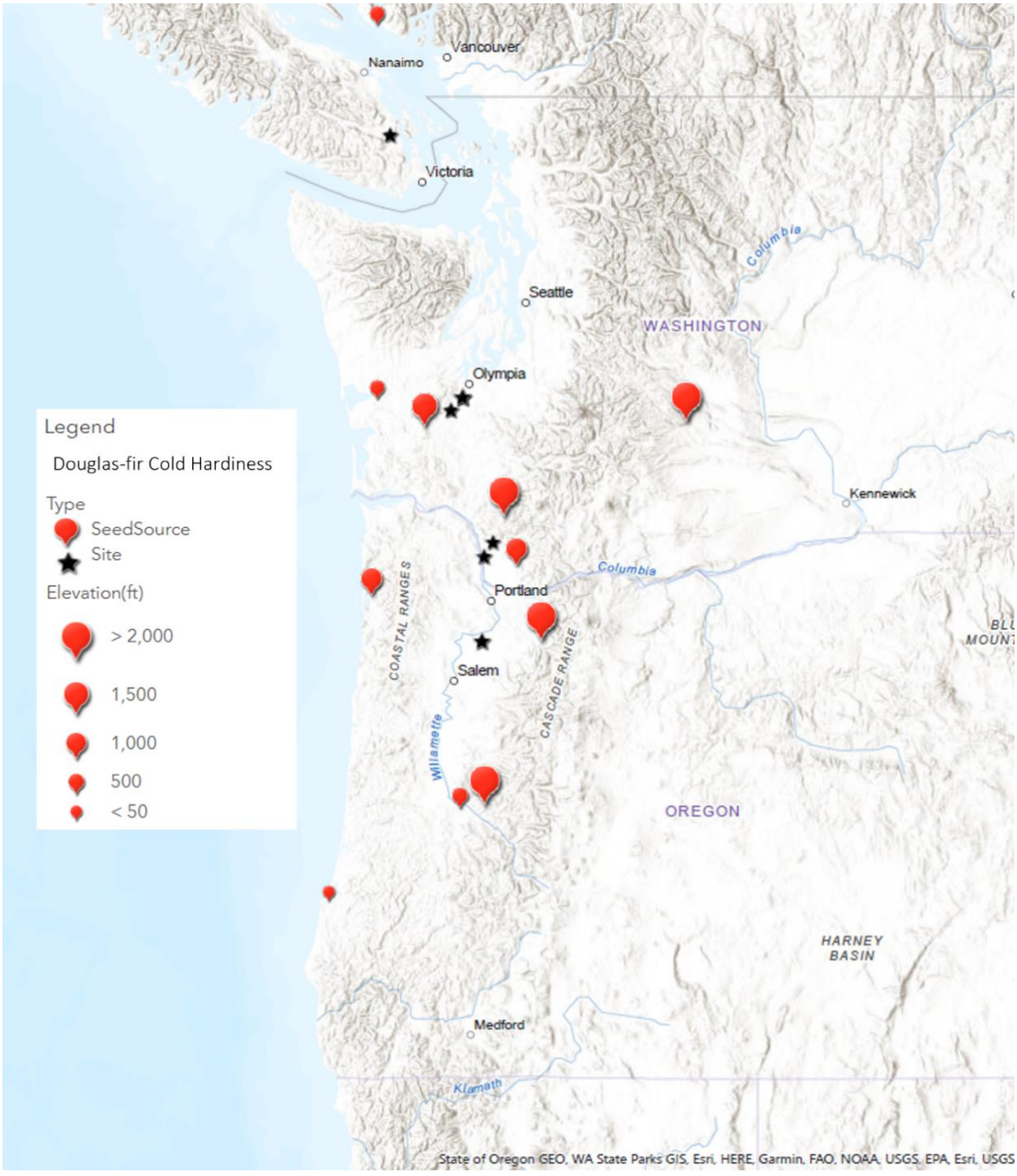

Figure 3. A map representing the cold hardiness data sources. Pins show seed source approximate origin, with size indicating elevation. Stars show site of cultivation. Seed sources include populations from southern British Columbia to southern Oregon and from 15 to 1000 m elevation. Sites range from southern British Columbia to central Oregon.

3. Data sets were split into two groups based on cultivation site location. These two groups were determined based on a latitudinal split at 46°N. This arbitrary line was drawn to create two groups will an equal number of sites. This differentiation is driven by modeling logistics rather than specific site weather or other biologically relevant features. The two groups consisted of the northern group, including sites near Olympia, WA and on Vancouver Island, and a southern group, including sites in Southern Washington and Oregon (Fig. 3).

Environmental Data Collection

Temperature data for United States sites were acquired from climate data online (CDO) hosted by the National Oceanic and Atmospheric Administration (NOAA) (NOAA, 2022). This service allows weather station searches by state, county, city, or zip code, among other search options. For each cultivation site and year in cold hardiness test data, daily temperature minimum and maximum were downloaded from CDO. Weather stations closest to the site with complete daily time series data were used. This excludes one source (Haase et al., 2016), for which daily temperature data were provided by the author, except for the data October 1st through 12th and May 1st, for which weather data was collected from CDO. For two cultivation sites, the Washington State Department of Natural Resources' Webster Forest Nursery, and Weyerhaeuser's Mima Nursery, the closest publicly available weather station was Olympia Airport, station ID USW00024227. There were two additional sites in Washington that the authors provided most weather data for, except for the few dates mentioned above. The gaps in these data sets were filled in with data from the Longview, WA weather station, USC00454769. For the Aurora, OR nursery site, weather data from the Silverton, OR station, USC00357823, were used. Additionally, there was one field site at Holt Creek, on Vancouver Island, British Columbia. The government of Canada hosts historical climate records online (Canada, 2011),

and the Duncan Glenora weather station ID 1022571 was used to extract climate data for this data set.

For each seed source, elevation was determined based on the coordinates of the collection site. Elevation was determined using the USGS's national map online app (USGS, 2022). Similarly, photoperiod data for each cultivation site was determined using geographic coordinates and the photoperiod function in the meteor package in R (R Core Team, 2022). Across the sites the photoperiod only varied by approximately 0.5 hours or less per day, at the date when photoperiod was most disparate.

In addition to the clustering scenarios, a single data source with repetition over time was modeled. The cold hardiness model has previously been parameterized using a minimum of three years of data for a single grape cultivar collected from a single location (Ferguson et al., 2011, 2014). In the case of wild type conifers, the closest analog to cultivar is seed source or population. None of the data sets extracted here provide three years of training data for a single seed source. However, Timmis et al. (1994) tested two seed sources for three full seasons. The Timmis data were modeled independently to provide a more uniform comparison to the clustering scenarios using the compiled data set from multiple sources. A random 70% of the Timmis data were used to train the cold hardiness model.

Model Description

The model used in this project was adapted from a grape cold hardiness extension testing program (Ferguson et al., 2011, 2014). This model was selected because it is used for real time predictions of cold hardiness for dozens of different grape cultivars, and inputs are limited to daily minimum and maximum temperature. Though there are numerous differences between grapes and conifers, this model has several benefits including utilizing inputs of daily average

temperature for cold hardiness prediction, testing with many years of climate data from a broad geographic range, biological interpretable parameters, inclusion of chilling and forcing temperature accumulation requirements, and parameters specific to the acclimation and deacclimation periods of cold hardiness phenology. These aspects of this model make it a desirable candidate to be used for operational decisions as daily summary temperature data that is available for many areas through public weather resources such as AgWeatherNet (WSU, 2022), NOAA (NOAA, 2022), and the Government of Canada (Canada, 2011). The drawbacks of this model are that it does not include photoperiod, which has been shown to impact cold hardiness acclimation in conifers (Leinonen et al., 1995; G. Zhang et al., 2003), and it uses a linear degree day function to determine chilling and forcing temperature exposure, which is less optimum than other temperature response functions proposed for conifer phenology in the literature (Harrington et al., 2010; Harrington & Gould, 2015). Though photoperiod is not included, effectively modeling cold hardiness without this value would be more parsimonious than other potential models. Differences in photoperiod are limited across the geographic range of Douglas-fir and within the season of cold hardiness phenology. Advantages to including photoperiod as a model predictor have only been shown for controlled environment growing conditions with daylength differences larger than those seen across temporal and spatial range used here (Greer et al., 2001; Leinonen et al., 1995).

Below are the equations used in the cold hardiness model adapted from grape (Ferguson et al., 2011, 2014). Average daily temperature is determined by the model inputs T_{max} and T_{min} (Eq. 1).

$$T_{mean} = \frac{(T_{max} + T_{min})}{2} \quad (1)$$

Degree days for chilling (Eq. 2) or forcing (Eq. 3) are determined for each time step by the difference between the average daily temperature and the temperature threshold for acclimation or deacclimation. Chilling is accumulated only when DD_c is less than 0 and forcing accumulates only when DD_f is greater than 0.

$$DD_c = T_{mean} - T_{th,a} \quad (2)$$

$$DD_f = T_{mean} - T_{th,d} \quad (3)$$

A logistic equation for either the acclimation (Eq. 6) or deacclimation phase (Eq. 7) is multiplied by the associated degree days (DDc or DD_f) and a rate constant (k_a or k_d) to calculate the daily change in cold hardiness (ΔH_c) (Eq. 6&7).

$$c_{log,a} = \frac{H_{c,min} - H_{c,i-1}}{H_{c,min} - H_{c,max}} \quad (4)$$

$$c_{log,d} = \frac{H_{c,i-1} - H_{c,max}}{H_{c,min} - H_{c,max}} \quad (5)$$

$$\Delta H_c = (DD_c * k_a * c_{log,a}) \quad (6)$$

$$\Delta H_c = (DD_f * k_d * c_{log,d}) \quad (7)$$

Daily cold hardiness ($H_{c,i}$) is calculated by adding the daily change in cold hardiness (ΔH_c) to the previous cold hardiness value ($H_{c,i-1}$).

$$H_{c,i} = H_{c,i-1} + \Delta H_c \quad (8)$$

Table 2. Parameter and variable definitions and symbols for the cold hardiness model. Adapted from Ferguson et al., 2011 and 2014. Use of each line item is demonstrated in equations 1 through 8.

Symbol	Description	Units
T_{max}	Maximum Daily Temperature	°C
T_{min}	Minimum Daily Temperature	°C
T_{mean}	Mean Daily Temperature	°C
DD_c	Chilling Degree-Days	°C
DD_f	Forcing Degree-Days	°C
H_c	Cold Hardiness	°C
$H_{c,min}$	Minimum Cold Hardiness Value	°C
$H_{c,max}$	Maximum Cold Hardiness Value	°C
$H_{c,0}$	Initial Cold Hardiness Value	°C
$T_{th,a}$	Threshold Temperature of Acclimation	°C
$T_{th,d}$	Threshold Temperature of Deacclimation	°C
k_a	Acclimation Rate	No Unit
k_d	Deacclimation Rate	No Unit
R_f	Forcing Requirement	°C
R_c	Chilling Requirement	°C

Table 2 shows the description of each parameter used in the model. Most parameters have a directly interpretable biological meaning. The rate constants k_a and k_d control the shape of the acclimation and deacclimation curves respectively. Cold hardiness, in units of LT_{50} °C, is represented as H_c and the starting value for each time step is $H_{c,i}$. $H_{c,min}$ and $H_{c,max}$ limit the range of cold hardiness values that can be predicted by the model. Chilling and forcing accumulate beyond temperature thresholds, $T_{th,a}$ and $T_{th,d}$, which temperature must be below or above, respectively. Required chilling and forcing accumulation necessary to transition from acclimation to deacclimation and from deacclimation to dormancy release are represented by R_c and R_f. Average daily temperature (T_{mean}) is calculated from the minimum and maximum daily temperatures (T_{min} and T_{max}). Daily average temperature is used to calculate degree days for chilling and forcing (DD_c and DD_f) based on the respective temperature thresholds ($T_{th,a}$ and $T_{th,d}$).

Model Calibration and Testing

Modeling was performed with the Cropbox framework, a declarative modeling framework (Yun & Kim, 2022). Cropbox is a framework developed for the Julia programming language that allows for the implementation of crop models, as well as model calibration, evaluation, and visualization. For this project Cropbox was used in a Julia (Bezanson et al., 2017) Jupyter notebook.

The cold hardiness model for each of the three data aggregation scenarios was trained with a randomly selected 70% of all data points and tested with the remaining 30%. These data points varied in site, seed source, and year of testing, and were selected randomly in R (R Core Team, 2022). The total number of data points was 123 and the size of the 30% testing data subsets ranged from 34 to 37 depending on the scenario and varying due to how data were subset into clusters (Table 1 & 5).

The separation of data into testing and training subsets limited the amount of data clusters per scenario to two. Limiting the number of clusters to two allowed each cluster to retain ample enough data points for training and testing. For each training data set optimal parameter sets were selected for the cold hardiness model using the calibrate function in the Cropbox framework (Yun & Kim, 2022). This model calibration method determines the best fit parameters using a global optimization method that is restricted to user defined parameter ranges (Table 3). The calibration was performed with up to five-thousand steps to determine the best fit parameters to model the training data set. Estimated parameter ranges based on biological assumption were input to enhance the calibration process (Table 3).

All model iterations were run starting on October 1st and ending on May 1st of the following year. The calibrated model was then tested with the remaining 30% test data associated with the training dataset to determine model performance. Predicted cold hardiness values for

each time and temperature series were extracted and compared to observed values from the test data set using the HydroGOF package in R (R Core Team, 2022)**.** The HydroGOF package provides several goodness of fit (gof) calculations and visualization functions. The gof function was used to generate a list of model fit metrics to compare between aggregated data sets.

Sensitivity Analysis

A sensitivity analysis was performed by varying each parameter independently of other parameters. Parameters were adjusted within a biological plausible range. The sensitivity analysis was performed on the model of all data to include the most data points in this analysis. Each parameter was adjusted by a percentage of the best fit value determined during model calibration and other parameters were fixed to the best fit value (Table 4). Adjustments were plus and minus 10, 30, and 50 percent. For negative parameter values, R_c, $H_{c,min}$, $H_{c,max}$ and $H_{c,0}$, the plus percentages were negative so that absolute value increased in the plus adjustments, and absolute value decreased in the minus adjustments. Results of the sensitivity analysis were summarized by comparing residuals from each parameter set. Mean absolute error was used to compare sensitivity of model results over datasets and dates.

Cross Validation

Due to the limitation of data set size, a modified 19-fold cross validation was performed to examine model bias and variance. A k-fold cross validation separates a dataset into k number of folds, and reruns model training and testing for each fold, where training excludes one fold used for testing in each iteration (Roberts et al., 2017). In this cross validation each subset of data for a single seed source, growing site, and observation date range was divided into a fold. In a traditional k-fold cross validation the n of each fold is equal, in this modified version n varied from 4 to 13 with a median of 7 and a standard deviation of 2.2. This created training data sets ranging from 89% to 97% of total datapoints. This modification was made to avoid creating folds

that were not representative of the complete time series of cold hardiness phenology. As an added benefit, each fold represented a distinct subset of data, so that model predictions could be tested for each of the folds and model fit for each data subset could be visualized. Cross validation was performed manually following the same modeling methods described previously.

Results

Data Clustering

The model output included estimated values for each parameter based on the best fit during calibration. These parameter estimates for each of the data scenarios can be seen in Table 4. Parameter values were limited in the calibration process to the ranges listed in Table 3. Though the limit for $T_{th,a}$ was set to a minimum of zero and a maximum of 10, all 5 scenarios modeled returned parameter estimates for $T_{th,a}$ between 9.4 and 10.0 °C. Another parameter that performed similarly across scenarios was the rate constant during acclimation, k_a, which ranged between 0.10 and 0.16. Other parameters had a more disparate range of estimations between scenarios. $H_{c,max}$, which represents the least cold hardy temperature since all H_c values are negative, had a range of -0.8 to -4.6. Previous work shows Douglas-fir maintains a low level of cold hardiness around -5 °C (Leinonen et al., 1995), so $H_{c,max}$ values may be overestimated.

To compare model performance, summary statistics for goodness of fit were calculated from the test data subsets (Table 5). The no clusters scenario included calibrating all compiled cold hardiness data as a single data set. The two additional scenarios included separate calibrations for two subsets of data. The seed source cluster scenario grouped data into clusters based on the similarity between twenty-five bioclimatic factors related to the seed source provenance, and geographic location. An alternate scenario, the growing site cluster, grouped data into two clusters based on geographic location of the growing site with the two subsets

divided at 46°N. These scenarios calibrated the cold hardiness model for two subsets of the full data set. The best model fit was achieved in the seed source cluster scenario based on goodness of fit metrics and model residuals (Fig. 4 & 5). Several statistics comparing the goodness of fit for all scenarios are listed in Table 5. Mean absolute error (MAE) and root mean square error (RMSE) for the seed source cluster scenario show reduced error and therefore improved fit when compared to the full data set or the growing site scenarios. Model efficiency (NSE) had the best results for all the seed source clustering scenarios, individual clusters and combined. The index of agreement (d), where one would be a perfect match between the observed and simulated values, also supports the seed source clustering showing improved values compared to the other two scenarios. The better fit of the seed source scenario is further supported by Pearson's correlation coefficient (r), and the coefficient of determination (R^2), which both show higher values for this scenario when compared to the other scenarios. The mean error (ME) and percent

Table 3: Parameter limits for model calibration. Parameter calibration took place within the limits listed in this table.

Parameter	**Lower Limit**	**Upper Limit**	**Unit**
$H_{c,min}$	-40	-15	°C
$H_{c,max}$	-5	0	°C
$H_{c,0}$	-7	0	°C
$T_{th,a}$	0	10	°C
$T_{th,d}$	0	15	°C
k_a	0	1	No unit
k_d	0	1	No unit
R_f	100	500	Degree Days
R_c	-500	-100	Degree Days

Table 4: Parameter estimates for each data scenario modeled for cold hardiness. Parameter definitions are listed in Table 1. *Chilling degree days are given in negative values. For the site and seed source clustering scenarios there is one set of parameters for each cluster, two per scenario. Parameter definitions are listed in table 2.

Parameter	No Clusters (All Data)	Sites North of 46°N	Sites South of 46°N	Seed Source Cluster 1	Seed Source Cluster 2	Units
$T_{th,a}$	10.0	9.9	9.9	9.4	9.9	°C
$T_{th,d}$	7.4	4.3	9.2	7.9	8.6	°C
R_c	-396.5	-474.8	-363.2	-291.6	-349.0	Degree Days*
R_f	277.8	104.5	437.0	265.5	290.5	Degree Days
$H_{c,min}$	-23.6	-28.6	-18.0	-26.7	-30.6	°C
$H_{c,max}$	-3.5	-4.6	-0.8	-1.3	-3.4	°C
$H_{c,0}$	-4.7	-6.4	-3.2	-6.9	-5.0	°C
k_a	0.10	0.11	0.16	0.16	0.10	No unit
k_d	0.35	0.14	0.51	0.35	0.66	No unit

Table 5: Summary statistics for goodness of fit for all scenarios tested. Statistics were calculated based on model residuals. The first three columns represent different scenarios for the full set of compiled data. The last four columns show results for each cluster in the two grouped data scenarios. ME is mean error; MAE mean absolute error; MSE is mean squared error; RMSE is root mean square error; d is index of agreement; r is Pearson's correlation coefficient; R^2 is the coefficient of determination; PBIAS is the percent bias; and n is the number of data points in each scenario represented by the column headers.

	No Clusters (All Data Combined)	Seed Source Cluster Combined	Growing Site Clusters Combined	Seed Source Cluster 1	Seed Source Cluster 2	Growing Site North of 46°N	Growing Site South of 46°N
ME	-1.82	-0.21	0.62	0.8	-0.63	0.33	1.33
MAE	3.33	2.87	3.72	2	3.23	4.39	2.12
MSE	18	14.79	25.5	5.19	18.76	33.46	6.42
RMSE	4.24	3.85	5.05	2.28	4.33	5.78	2.53
NSE	0.29	0.72	0.42	0.86	0.64	0.11	0.79
d	0.86	0.92	0.84	0.95	0.9	0.74	0.94
r	0.81	0.85	0.7	0.95	0.81	0.52	0.92
R^2	0.65	0.72	0.5	0.91	0.66	0.27	0.85
PBIAS	13.4	1.4	-3.8	-4.3	4.5	-1.8	-11
n	36	37	34	8	29	24	10

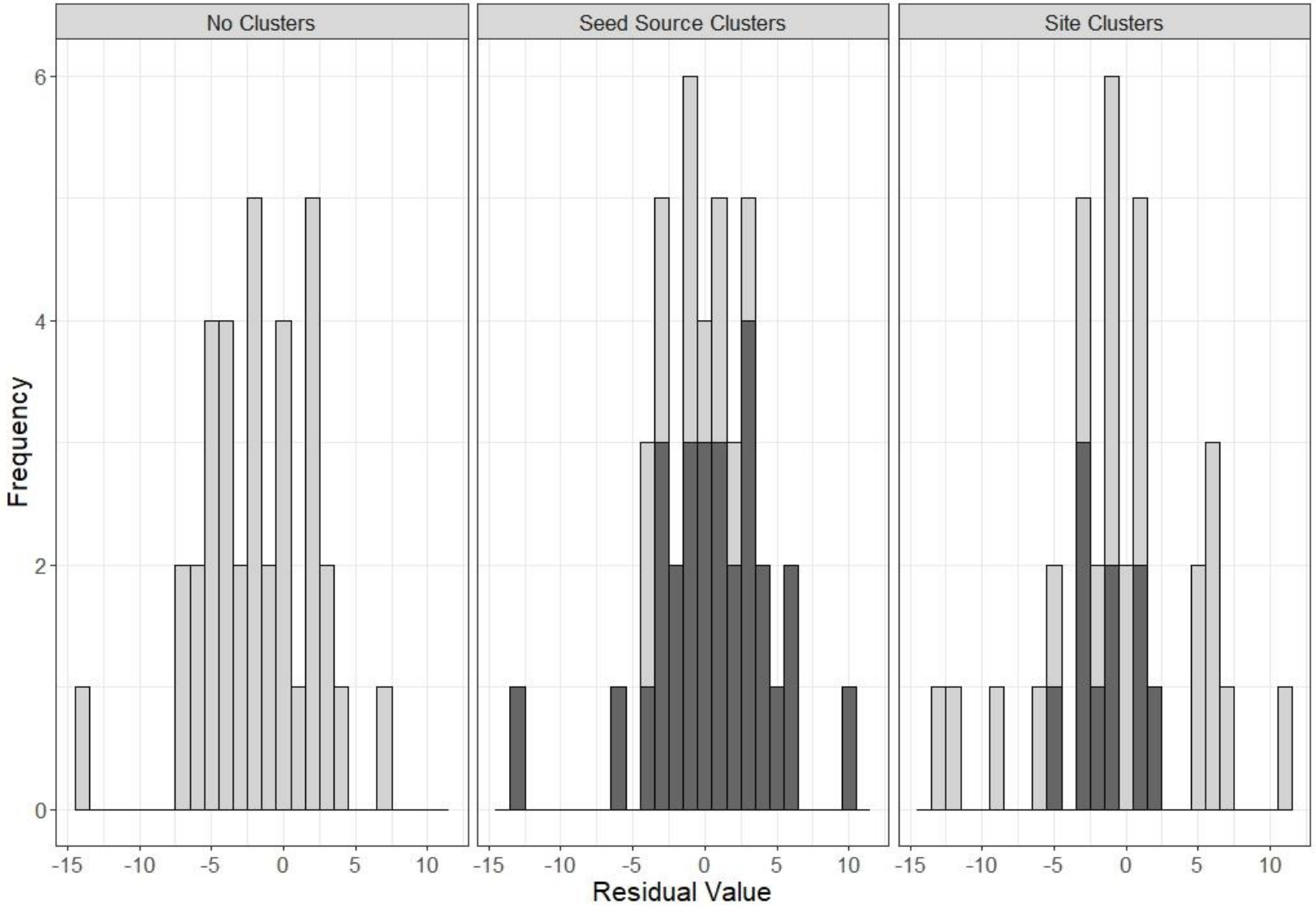

Figure 4. Histogram of residuals by scenario showing the frequency for residual values (bin=1). Dark grey color represents Seed Source Cluster 2 and Growing Sites South of 46°N. Light grey color represents all data in the no clustering scenario, as well as Seed Source Cluster 1 and Growing Sites North of 46°N. The northern site cluster shows greater spread in residuals compared to the southern cluster.

bias (PBIAS) show that there is model overestimation bias in the single parameter set model, but in the two clustering scenarios there is little or no bias in the residuals. There was some improvement in MAE and RMSE in the growing site cluster combined data set, the cluster containing sites north of 46°N performed less well when compared to the no clustering scenario (Table 5). In both clustering scenarios, the residuals from the larger cluster have a greater spread, and thus a worse fit, compared to the smaller cluster within the same scenario (Fig. 4).

In addition to the clustering scenario comparison, model results from the Timmis et al. (1994) data set were analyzed. This data set represents only a subset of the compiled data used in the clustering scenarios and is limited to a single cultivation site and two seed sources. The Timmis et al. (1994) data are unique compared to other data sources because the same seed sources and growing site were replicated over several years. The goodness of fit statistics for this subset of data are listed in Table 6. Compared to all clustering scenarios for the full data set, seed source clustering, cultivation site clustering, and no clustering, the Timmis data subset showed improvement in most goodness of fit metrics. Statistics that are sensitive to bias, such as mean error (ME) and percent bias (PBIAS), had similar results for the two clustering scenarios compared to the Timmis data subset. However, the full data set parameterization showed -20% bias when looking only at the Timmis data, a much greater degree of bias than all other scenarios. This bias can be visualized in Figure 5, which shows the relationship between observed and predicted cold hardiness values. In the no cluster scenario, the data points from the Timmis subset display this bias by primarily falling above the 1:1 line, which represents a perfect fit, showing the bias.

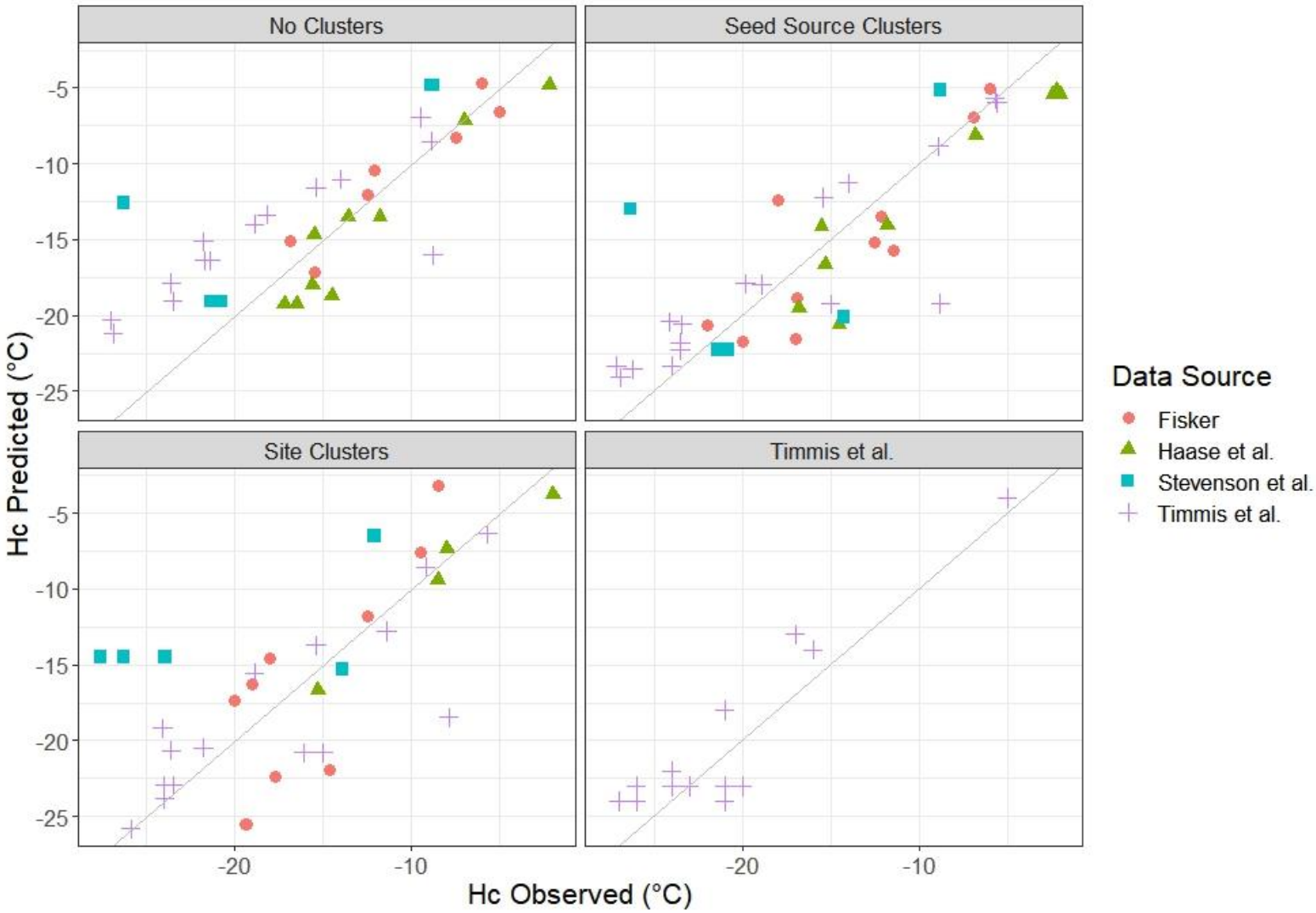

Figure 5. Relationship between observed and predicted cold hardiness values for each model scenario. Colors and symbol shapes indicate the publication data were derived from. Gray lines show a 1:1 relationship between observations and predictions. Wider spread of points indicates greater error. Uneven distribution of points to either side of the line indicates bias. Only datapoints from the testing subset are shown (30% of all datapoints).

Table 6. Summary statistics for goodness of fit all model scenarios for only the Timmis et al. (1994) subset. Each column shows the results for only the testing data from this subset, but using the models calibrated for each scenario. The Timmis calibration was performed for both seed sources included in the data from this source. See Table 4 for statistic abbreviation definitions.

	Timmis et al. (1994)	No Clusters	Seed Source Clusters Combined	Growing Site Clusters Combined
ME	1.0	3.7	0.77	-0.42
MAE	2.0	4.73	2.54	2.66
MSE	5.5	25.55	12.04	14.7
RMSE	2.35	5.05	3.47	3.83
NSE	0.8	0.31	0.77	0.66
d	0.95	0.78	0.93	0.89
r	0.92	0.84	0.89	0.81
R^2	0.85	0.71	0.79	0.66
PBIAS	-4.7	-20	-4.2	2.4
n	16	14	17	15

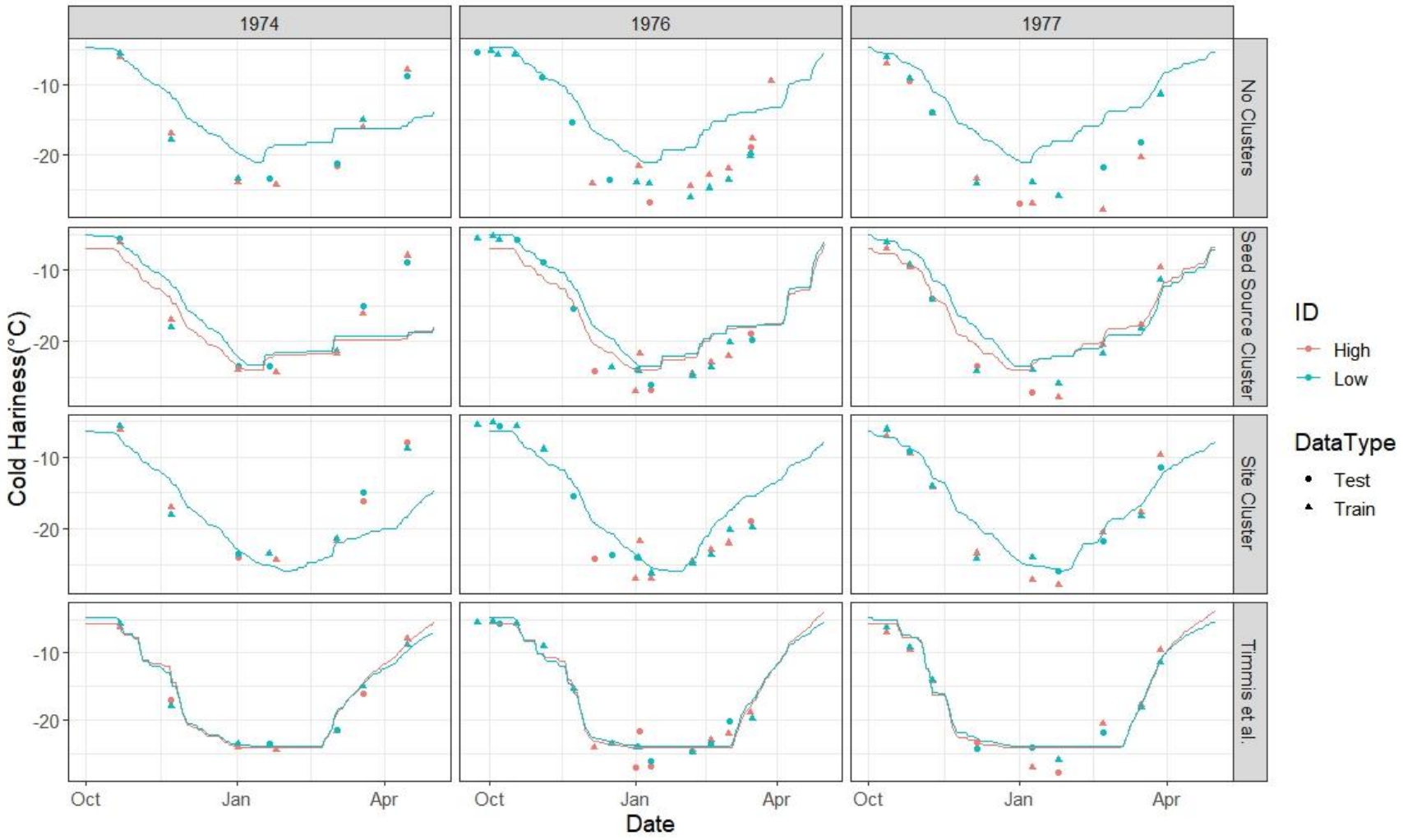

Figure 6. Model fits for the Timmis et al. (1994) data subset for all model parameterization scenarios. Colors indicate seed source and point shape indicates whether observations were used for model testing or trainings. Lines show model predictions. For scenarios that models did not account for seed source the depicted line was the predicted cold hardiness for both seed sources. The model calibration with Timmis data shows the best fit, and the no cluster scenario shows the worst fit, and frequent overestimation.

For the Timmis data subset the no clusters scenario had larger error (MAE) and performed worse than the seed source cluster scenario or the Timmis data model parameterization. There was approximately a 45% decrease in MAE (Table 6) for both clustering scenarios when compared to the no cluster scenario.

Figure 6 shows the model fit for the Timmis et al. data subset for each model parameterization scenario. For the no clustering and site cluster 1 scenarios only one prediction was produced per site and year, since there was only one temperature data set. The Timmis et al. observations represent two seed sources observed at one site over several years and the seed source clustering and Timmes et al. model fit scenarios have predictions for the individual seed sources. It is apparent that the no clustering scenario did not achieve a low enough minimum cold hardiness value, which explains the high level of bias observed (Table 6). All 3 scenarios that used multiple data source for calibration fit the deacclimation curve of the Timmis data

Table 7. Parameter value comparison for deacclimation response parameters for each scenario compared to the Timmis et al. (1994) data set in figure 6. The calibrations with Timmis data for high and low seed sources have much lower threshold temperature ($T_{th,d}$) for deacclimation, allowing deacclimation to occur at a greater temperature range.

Scenario	R_c	$T_{th,d}$	k_d
No Clustering	-396.5	7.4	0.35
Seed Source Cluster 1	-291.6	7.9	0.35
Seed Source Cluster 2	-349	8.6	0.66
Site Cluster 1	-474.8	4.3	0.14
Timmis et al. high	-337.7	**0.069**	0.095
Timmis et al. low	-394.3	**1.3**	0.12

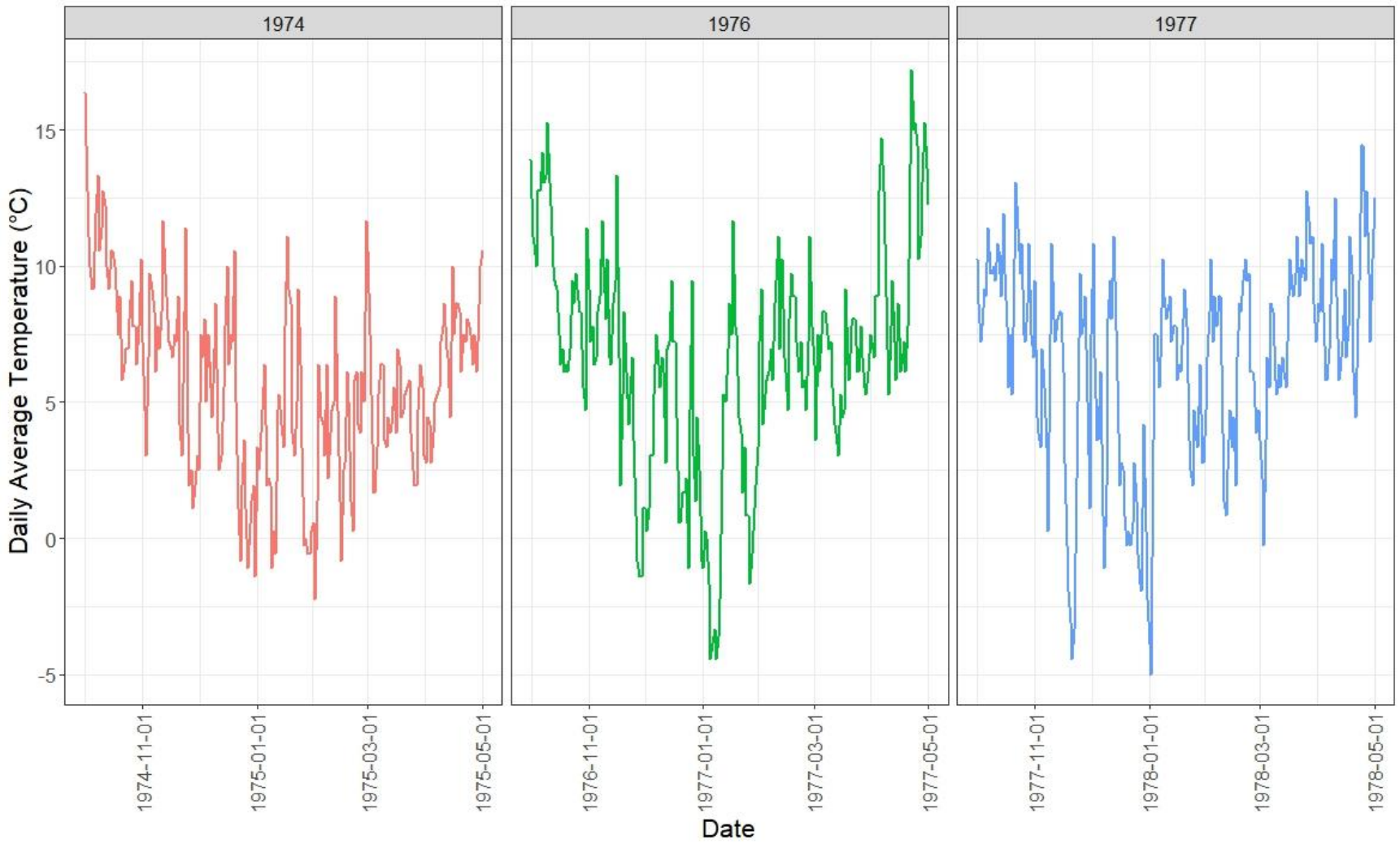

Figure 7. Daily average temperature from Timmis et al. (1994) data for each season modeled. Grey box indicates year of dormancy onset. Lower late winter and spring temperatures can be observed in the dormancy season starting in 1974.

poorly for the 1974 samples. When the model was calibrated with only the Timmis data the fit of this section of the curve improved noticeably. This issue is not observed in deacclimation curves from the following seasons. One likely cause for this effect in the different model scenarios is the low temperature for the $T_{th,d}$, which the threshold that controls which temperatures will count towards the progression of deacclimation. In the model trained with only a subset of the Timmis data set the value of $T_{th,d}$ is lower than any other scenarios (Table 7), which will allow deacclimation to progress over a greater temperature range. This becomes most obvious in the poor fit of the deacclimation curve by other model scenarios in the dormancy season starting in 1974. In early 1975 average daily temperature remained low as compared to the late winter and spring of the other years represented in the Timmis data (Fig. 7), which is a likely driver in the low value of $T_{th,d}$. This example demonstrates the importance of replication for model parameter estimation.

Sensitivity Analysis

A parameter sensitivity analysis was performed for the model fit to all data (Fig. 8). This is comparable to the no clustering scenario, which had the greatest model error, and the most data points. Parameter sensitivity was ranked based on comparisons between model residuals for each parameter set. Some parameters had little to no impact on model mean absolute error (MAE), demonstrating low sensitivity. These parameters include required forcing (R_f), deacclimation rate (k_d), initial cold hardiness temperature ($H_{c,0}$), and maximum temperature of cold hardiness ($H_{c,max}$). These parameters showed very little change in mean and median absolute error and very little increase in spread in absolute residual distribution. R_f and k_d are features of the deacclimation portion of the cold hardiness curve. R_f has no impact on model performance, as the model is currently limited to specific calendar dates, not dormancy release.

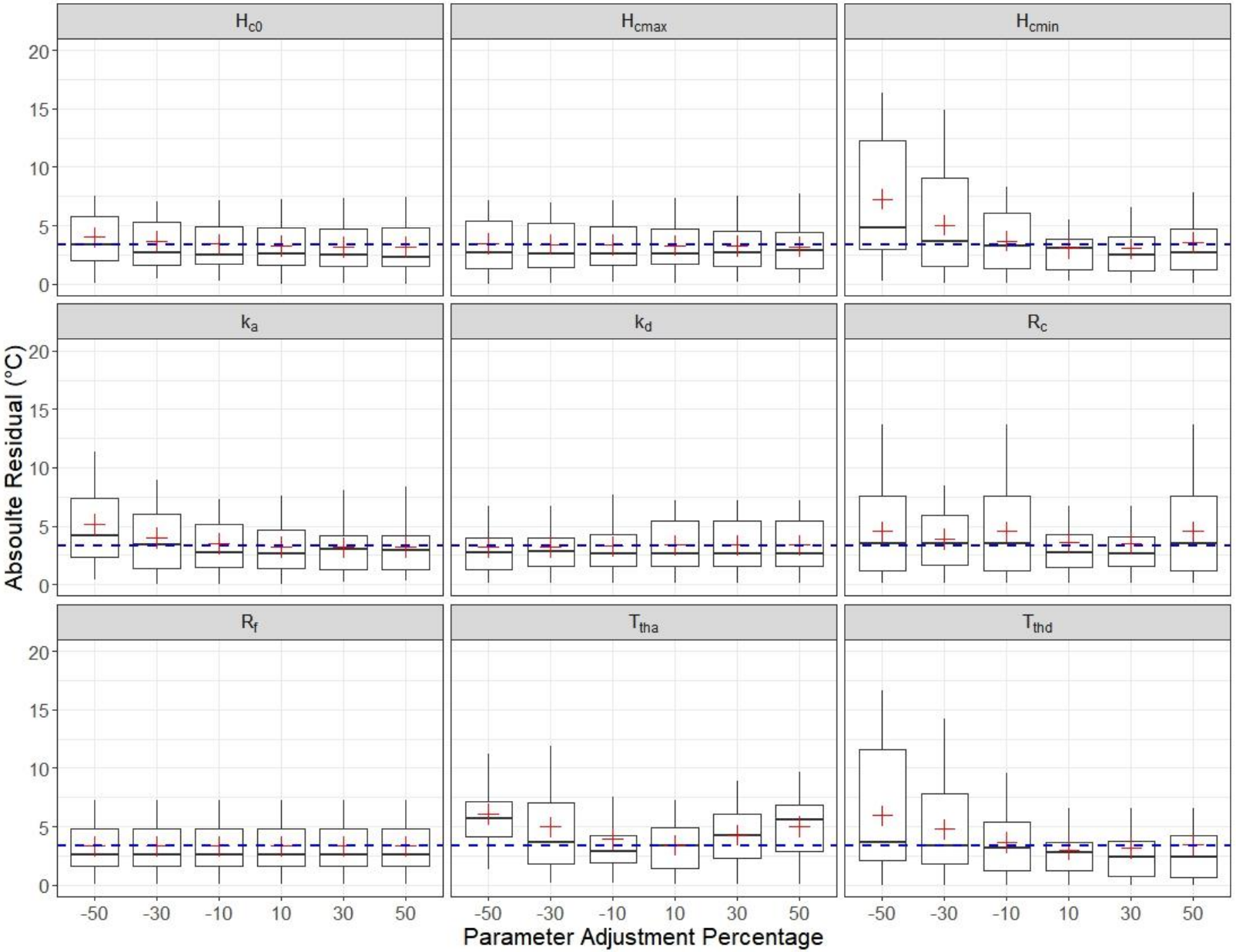

Figure 8. Absolute residuals from parameter sensitivity analysis for each parameter in the cold hardiness model. Boxes indicate the two middle quartiles of absolute residuals for each adjustment, and whiskers show the two outer quartiles. Total size of box and whiskers for each adjustment shows residual spread. Crosses show mean absolute error and black solid lines show median absolute error. Dashed red lines show mean absolute error from the best fit parameters in table 3.

However, k_d may impact the curve shape during deacclimation. In equation 8, k_d is multiplied by the degree days for forcing and the deacclimation logistic equation to determine the daily incremental change in cold hardiness during deacclimation. $H_{c,0}$ and $H_{c,max}$ are parameters that may be fixed based on their low sensitivity and small range of biologically viable values. Notably $H_{c,0}$ and $H_{c,max}$ are negative parameters, so adjustments made were based on absolute value, not sign.

Parameters with moderate sensitivity include acclimation rate (k_a), and chilling required (R_c). These two parameters showed changes in the distribution of residuals for some parameter adjustments, and small or no changes in the mean and median of the absolute residuals. For both k_a and R_c residual spread increased when they were reduced in absolute value in the minus percentage adjustments. R_c only showed a change in the most extreme increase in absolute value, and k_a did not have increased error when increased. This suggests that these parameters, which both relate to the acclimation portion of the cold hardiness curve, have a range of higher values that perform similarly. Similar to k_d, k_a may impact the shape of the cold hardiness curve but during the acclimation period. However, k_a had a small range of values (0.10 to 0.16) in models for the different scenarios. Reducing k_a by 30 and 50 percent showed an increase in MAE and increased the spread of residuals. The low and moderate sensitivity of these shape parameters, k_a and k_d, is a beneficial outcome in terms of model interpretation because these parameters are the least directly explained biologically. Additionally, the small spread and biased sensitivity of k_a demonstrates a minimum value for this parameter in the model and provides justification for reducing range of this parameter during calibration or fixing the parameter or narrowing in on a fixed value with additional testing.

Three parameters that show high sensitivity: the minimum cold hardiness temperature ($H_{c,min}$), and the two temperature threshold parameters for degree day accumulation, ($T_{th,a}$ and $T_{th,d}$). Of these, the $H_{c,min}$ and $T_{th,d}$ sensitivities are largely skewed to the parameter adjustments that decreased absolute parameter value. Increasing absolute value of these parameters did not show high sensitivity. The increased error at -10, -30, and -50 percent adjustment suggest that these adjusted values did not achieve a low enough temperature for peak cold hardiness, overestimating the maximum cold hardiness had little negative impact on model results, and lower values for $H_{c,min}$ actually showed slight improvement in MAE, but greater residual spread. $T_{th,a}$ shows increased mean absolute error and residual spread at all adjusted parameter values except for increasing $T_{th,a}$ by 10 percent. The sensitivity of $T_{th,a}$ highlights the importance of quantifying plant temperature response in phenology modeling. In this case there is a very limited range of values that perform well for $T_{th,a}$ as shown in the sensitivity analysis and in this parameter value across model scenarios (Table 4). Chilling degree day threshold ($T_{th,a}$) for Douglas-fir may also be a fixed parameter value of 10°C when chilling degree days (DD_c) are calculated using daily average temperature as depicted in equation 2.

Cross Validation

Results of the 19-fold cross validation demonstrate model bias and variance. With a 19-fold cross validation both bias and variance can be assessed moderately. In general, the more folds in an k-fold cross validation the more bias is revealed in relation to variance, in an effect known as the bias-variance tradeoff (Nowak, 1997). The results of the 19-fold cross validation can be seen in detail in Figure 9 and Tables 8 and 9. Table 8 shows parameter estimates, fold size, and model goodness of fit metrics for each of the 19 folds. Table 9 shows summary statistics for the results shown in Table 8. Plot 9 provides a visualization of individual fold

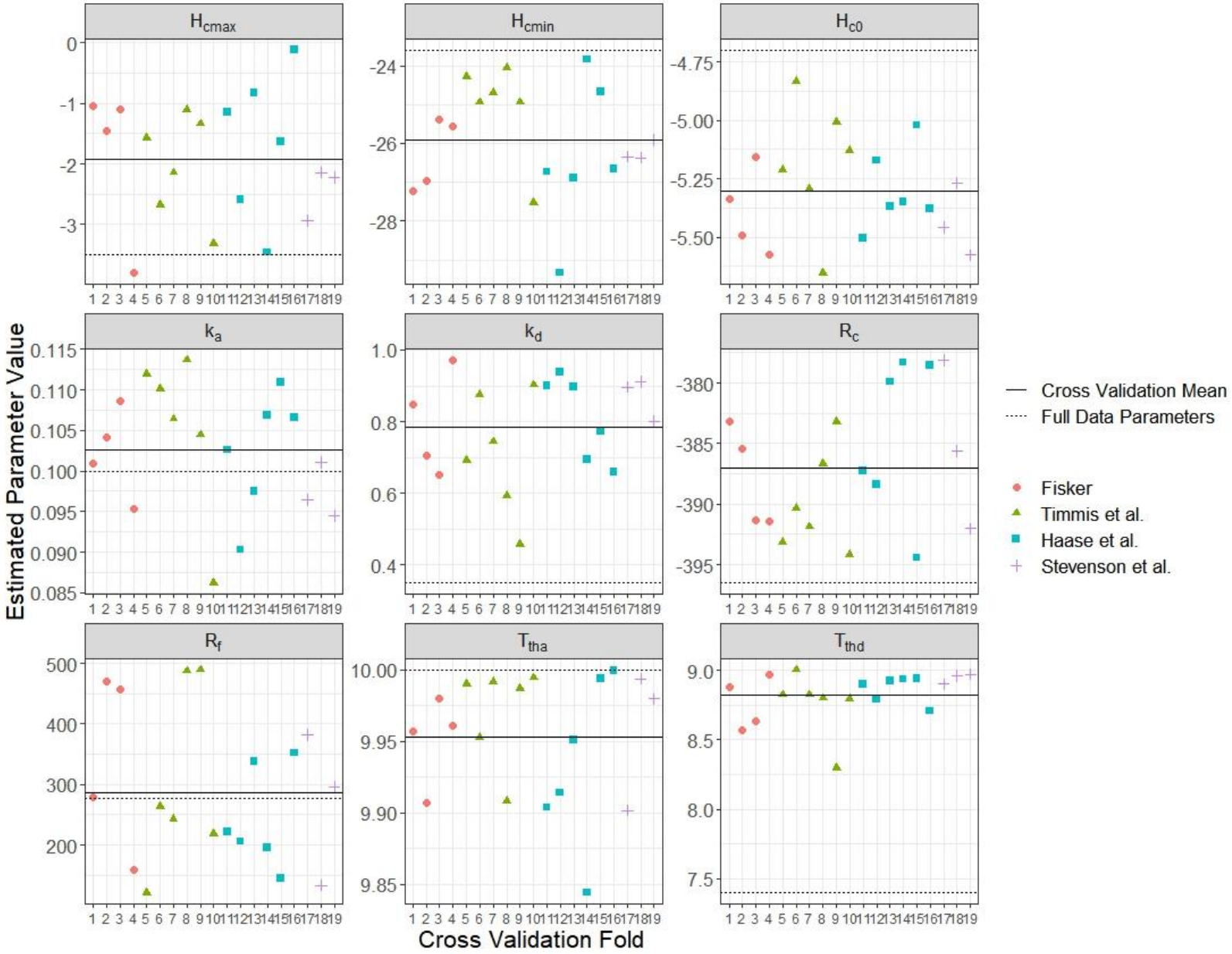

Figure 9. 19-fold cross validation results for each parameter (plot), fold (points), mean cross validation results (solid lines), and comparison to parameter estimates from modeling the full data set (dotted lines). Point color and shape indicate which publication data were derived from. Each fold represents one data subset for seed source, growing site, and year of observations.

Table 8. Complete results of modified 19-fold cross validation, including model goodness of fit metrics and parameter estimates. For each data source there are multiple rows to encompass different seed sources, sites, and date ranges. For definitions see Table 2.

Fold	Source	RMSE	NSE	n	$H_{c,max}$	$H_{c,min}$	$H_{c,0}$	k_a	k_d	R_c	R_f	$T_{th,a}$	$T_{th,d}$
1	Fisker, 1993	1.90	0.81	7	-1.05	-27.22	-5.33	0.10	0.85	-383.18	278.74	9.96	8.88
2	Fisker, 1993	2.49	0.77	7	-1.45	-26.96	-5.49	0.10	0.70	-385.44	471.13	9.91	8.57
3	Fisker, 1993	2.76	0.75	7	-1.09	-25.38	-5.16	0.11	0.65	-391.34	456.73	9.98	8.63
4	Fisker, 1993	1.13	0.96	7	-3.80	-25.57	-5.57	0.10	0.97	-391.39	159.64	9.96	8.97
5	Timmis et al., 1994	3.35	0.76	10	-1.57	-24.29	-5.21	0.11	0.69	-393.16	121.14	9.99	8.82
6	Timmis et al., 1994	3.54	0.80	13	-2.68	-24.94	-4.83	0.11	0.88	-390.30	262.84	9.95	9.00
7	Timmis et al., 1994	4.65	0.50	7	-2.15	-24.69	-5.30	0.11	0.74	-391.87	242.39	9.99	8.82
8	Timmis et al., 1994	4.11	0.69	10	-1.11	-24.05	-5.65	0.11	0.59	-386.68	487.85	9.91	8.80
9	Timmis et al., 1994	5.12	-1.00	7	-1.34	-24.94	-5.01	0.10	0.46	-383.23	489.40	9.99	8.29
10	Timmis et al., 1994	4.98	0.47	7	-3.32	-27.53	-5.13	0.09	0.90	-394.19	218.32	9.99	8.80
11	Haase et al., 2016	2.27	0.74	5	-1.15	-26.73	-5.51	0.10	0.90	-387.24	221.65	9.90	8.90
12	Haase et al., 2016	2.25	0.79	5	-2.59	-29.33	-5.17	0.09	0.94	-388.40	205.73	9.91	8.79
13	Haase et al., 2016	1.19	0.92	5	-0.83	-26.88	-5.37	0.10	0.90	-379.89	338.53	9.95	8.93
14	Haase et al., 2016	2.40	0.78	5	-3.46	-23.84	-5.35	0.11	0.69	-378.30	194.94	9.84	8.94
15	Haase et al., 2016	3.45	0.28	5	-1.63	-24.68	-5.02	0.11	0.77	-394.42	144.61	9.99	8.94
16	Haase et al., 2016	2.33	0.78	5	-0.12	-26.67	-5.38	0.11	0.66	-378.56	352.45	10.00	8.71
17	Stevenson et al., 1999	7.28	-0.19	4	-2.94	-26.34	-5.46	0.10	0.90	-378.11	382.33	9.90	8.90
18	Stevenson et al., 1999	6.23	-0.14	6	-2.14	-26.37	-5.27	0.10	0.91	-385.61	133.40	9.99	8.96
19	Stevenson et al., 1999	8.49	-0.92	5	-2.22	-25.91	-5.58	0.09	0.80	-392.02	296.01	9.98	8.97

Table 9. Summary results for 19-fold cross validation fold size and parameter estimates.

	n	$H_{c,max}$	$H_{c,min}$	$H_{c,0}$	k_a	k_d	R_c	R_f	$T_{th,a}$	$T_{th,d}$
Minimum	4.00	-3.80	-29.33	-5.65	0.09	0.46	-394.42	121.14	9.84	8.29
Mean	6.68	-1.93	-25.91	-5.30	0.10	0.78	-387.02	287.25	9.95	8.82
Median	7.00	-1.63	-25.91	-5.33	0.10	0.80	-387.24	262.84	9.96	8.88
Maximum	13.00	-0.12	-23.84	-4.83	0.11	0.97	-378.11	489.40	10.00	9.00
Standard Deviation	2.24	1.00	1.40	0.22	0.01	0.14	5.55	123.48	0.04	0.17
Coefficient of Variance	0.33	-0.52	-0.05	-0.04	0.07	0.18	-0.01	0.43	0.00	0.02

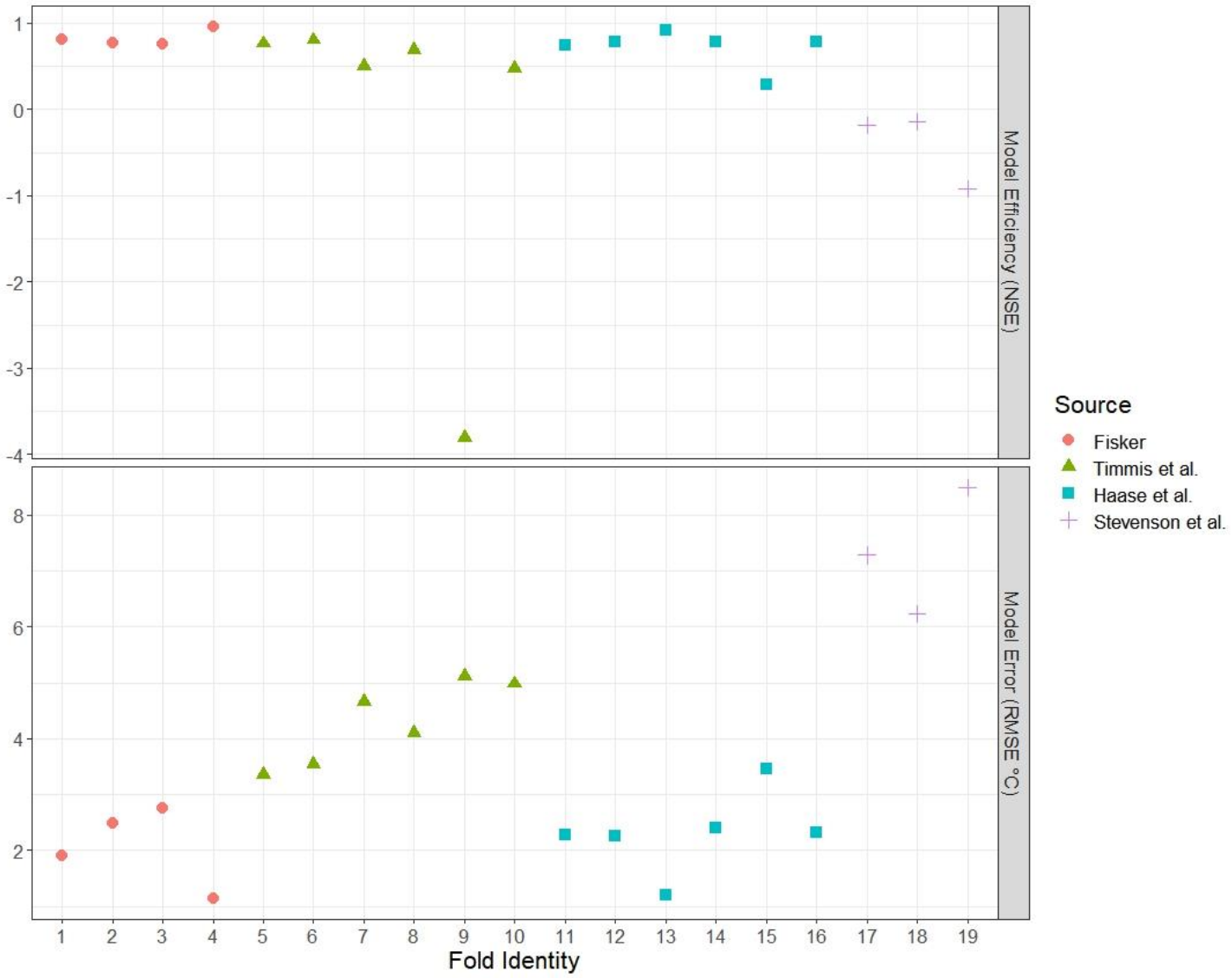

Figure 10. Cross validation model error and efficiency results for each fold. Point shape and color show which data source each fold was derived from. Each fold represents one data subset encompassing all dates for a unique combination of seed source, growing site, and year of observations.

parameter values and horizontal line to compare individual fold results to mean values for all folds and to the original model trained with 70% of all data points, with 30% reserved for testing (Table 4). Cross validation summary statistics include the minimum, mean, median, maximum, standard deviation, and coefficient of variance values for each model parameter. Comparison of point range, and difference between cross validation means and parameter estimates from the full results (Fig. 9), show which parameters were most biased in the original model estimates, and an estimate of variation of model parameter estimates for this data set. Comparison of the coefficient of variance between parameters (Table 9) reveals which parameters are the most variable for this model and dataset. Based on the coefficient of variance $H_{c,max}$ and R_f show relatively high variance, k_d moderate variance, and all other parameters low variance.

The cross validation was performed on subsets of data that represented observations for one dormancy season, seed source, and site. This structure also allows for comparison of model fit for each data source (Fig 10). Figure 10 shows the model error (RMSE) and efficiency (NSE) for each fold of the cross validation. RMSE is a measurement of error that is biased to account for larger error in residuals when compared to mean absolute error. NSE shows residual error on a normalized scale of 1 to $-\infty$, where 1 is a perfect fit to observation data, 0 shows that the model predictions are equal to the mean of observed data, and negative values are worse than the mean. Overall, model fit was grouped by data source, showing that data from Stevenson et al. did not perform well when the model was trained with the full data set, had the largest error (RMSE) and had negative NSE values. This subset of data was the furthest north (Table 1) and also consisted of seedlings that had been established in a realized gain trial for three years before observations were made (Stevenson et al., 1999). The other data sources show overlapping ranges of error and efficiency, except for one data subset from Timmis et al. (1994) that showed a negative NSE.

Analysis of model residuals for this subset, fold 9 from the cross validation, shows that LT_{50} temperature was overestimated at all timepoints, and this was emphasized by the lack of data points during acclimation within this subset, explaining the low NSE. These results show no distinct grouping within a single data source by seed source elevation, the primary grouping characteristic in the seed source cluster scenario, or by growing site location, which include sites in Washington and Oregon (Fig. 3).

Discussion

Model Performance

The results shown here support the two hypotheses proposed in the aims. The cold hardiness model using only temperature as a predictor was able to simulate Douglas-fir cold hardiness. Model predictions improved in goodness of fit as data modeled were more similar. Clustering by seed source reduced model error for the full compiled data set when cluster residuals were combined. However, there was variability within the individual seed source clusters, which demonstrated inconsistency in this improvement. Though error was variable, model efficiency was improved for the seed source clustering scenarios for both combined residuals and within each cluster. The results from the growing site clustering also showed variability between clusters. Error (RMSE) and efficiency (NSE) showed improvement compared to no clustering for sites south of 46°N but not for sites north of 46°N. Overall, all the poor fit for the northern cluster resulted in combined residuals showing greater error than no clustering but also greater, and thereby improved, efficiency. Though there was variability within the performance of individual clusters, clustering data still improved model performance over the no cluster scenario, grouping all data.

This improvement is further exemplified in the examination of model performance when applied only to the Timmis data subset. This subset includes one growing site and two seed sources that were observed over multiple dormancy seasons. For these data, parameters that were calibrated using a training data set just from this data subset performed the best (Fig. 5 & 6, Table 6). The results show a clear model performance improvement in the parameters from both the seed source and growing site clustering scenarios when compared to the parameters from the no clustering scenario. Additionally, the goodness of fit metrics shows improvement for both clustering scenarios applied to the Timmis data than to the full data set. This supports that model fit is improved with more similar data. Since the Timmis data includes only one location, only a single parameter set from the growing location clusters was used to fit the model for this data subset. This is notable because this parameter set is for the growing site north of 46°N, which performed worse when applied to all the compiled data in the no cluster scenario. This suggests that the poor performance of the north of 46°N cluster may be improved by further division of the cluster. This cluster includes sites in the south Puget Sound region of Washington state and a site located on Vancouver Island, British Columbia, which may be an outlier and contribute to model error within this cluster when applied to all data. The disparate performance of this data set can be further seen in the cross validation results which show higher model error and low model efficiency for each data set derived from this source (Figure 10).

The parameter sensitivity analysis revealed parameters that are candidates for fixed values: $T_{th,a}$, $H_{c,0}$, and $H_{c,max}$. Low sensitivity was observed in $H_{c,0}$ and $H_{c,max}$. Fixing these values within the range tested will not have negative impacts on model performance. Parameter $T_{th,a}$ shows high sensitivity, with model error increasing as the parameter is adjusted farther from the best fit value. Though there is high sensitivity for this parameter, the fit value across data sets

and scenarios tested ranged only between 9.4 and 10.0°C. This is a small range of values that are appropriate for this threshold, and when the no cluster scenario was tested with values above or below 10.0°C mean absolute error and residual spread increased. Thus, a fixed value of 10.0°C would be appropriate for $T_{th,a}$. Fixing these parameters will reduce the total number of parameters that need to be calibrated for each population cluster, which will increase model parsimony and decrease the likelihood of overfitting the model.

Fixing $H_{c,max}$ can be informed by the results of the sensitivity analysis and cross validation. The low sensitivity of $H_{c,max}$ implies little impact on model fit if a fixed value is selected within an appropriate range. For the data used in this modeling exercise, this range may be selected by the results of the cross validation. The cross validation showed a mean value close to -2°C and the no clustering scenario showed a value of -3.5 °C. A selection of -3°C for this parameter would represent this range well and would be unlikely to negatively impact model fit, based on the sensitivity analysis.

The potential fixed parameters also provide insight into thresholds on biological processes. $T_{th,a}$ shows a fixed value for degree day accumulation during chilling for average daily temperatures below 10.0 °C. The consistency of this result across model scenarios suggests that this is representative of the populations and environmental conditions tested here. $H_{c,max}$, in the unit of lethal temperature of 50% (LT_{50}), can further support estimations of this value that have been made through the use of controlled environment trials. It's important to note that LT_{50} is a population level metric, and absolute maximum temperature of an individual tissue does not necessarily represent the variation within a population, so this metric may be an interesting comparison to measurements made of absolute maximums. $H_{c,0}$ represents the starting value for

LT_{50} in the model. This is primarily a logistical parameter for model implementation, but could show an opportunity for model simplification by basing this value on $H_{c,max.}$

One parameter that has received little discussion thus far is the forcing requirement (R_f). This parameter is not currently utilized in the cold hardiness model, as the model is based on calendar date, not a biological derived end point. R_f remains in the model as a method to tie in budburst prediction, a goal for future iterations of this model. Though cold hardiness and dormancy phenology are not identical, they are correlated, and this correlation increases during spring deacclimation (Aitken & Adams, 1997; Stevenson et al., 1999). Thus, R_f can be used to toggle from the cold hardiness model to a bud burst sub-model that may be running simultaneously based on chilling and forcing accumulation. In addition to applied uses of a unified cold hardiness and bud burst model, it could be used along with future climate simulated data to determine risk of cold damage and changes in bud burst timing based on population specific parameters.

The results of a 19-fold cross validation identified two parameters with high variance, and one with moderate variance (Table 9). Both parameters that had high variance, $H_{c,max}$ and R_f, had low sensitivity, suggesting that this variance is unlikely to impact model fit. The variance of $H_{c,max}$ was high in relative terms, but this variance is over emphasized by the small parameter range from 0 to -4 °C in the cross validation scenario (Fig. 9), which can be reduced further by the biological assumption of low underlying cold hardiness. R_f is a parameter in the model with low sensitivity and no impact on model performance in the current model implementation, as mentioned in the previous paragraph. The only other parameter with an absolute coefficient of variance greater than 0.10, or in other words has a standard deviation that is greater than 10% of the mean, is k_d, which also showed low sensitivity (Fig. 8). This parameter has the greatest

uncertainty in the model, but with low sensitivity, is unlikely to impact model error and efficiency.

Parameter values from the no cluster scenario fall on the edge or slightly outside of parameter values estimated for each of the 19 folds of the cross validation. This can be explained by the larger percentage of training data compared to testing data in the cross validation. Since the cross validation did not have an equal number of data points per fold, the range of training data was from 97% to 89% of all data points, compared to 70% in the no cluster scenario. This shows some bias in parameter estimation in the no cluster scenario, but the direction of the bias varies by parameter and can be visualized in Figure 9.

Broader Impacts

Simplifying the calibration of this model will by fixing some parameters will increase the useability for application in conifer production and regeneration by reducing the chance of overfitting. Conifer seedling production is on the rise as reforestation and wildfire response are receiving federal attention (Fargione et al., 2021). Increased demand creates more opportunity for implementation of phenology models that can aid in timing of production and establishment workflows and cultural decisions. Knowledge of seedling cold hardiness status, and population specific response to temperatures, can be used to prioritize planting order, planting seedlings with greater risk of rapid deacclimation or cold damage due to lower hardiness levels can be delayed over hardier seedlings. Model parameterization for specific populations can reveal these traits through parameter estimates k_d and $H_{c,max}$ or cold hardiness of a certain day ($H_{c,i}$). In addition to knowledge of a populations estimated cold hardiness at a point in time, understanding temperature response through population specific parameters, such as chilling requirement (R_c) and deacclimation rate constant (k_d), can be leveraged to rate or rank seedlings in terms of

vulnerability. This vulnerability rating can be used as a proxy for planting order prioritization when estimated cold hardiness level is unknown. In addition to outplanting order, populations in nursery production can also be ranked in a similar fashion. In the Pacific-Northwest, weather can be a major limiting factor of seedling lifting timing in bareroot production. Rain can saturate soils and exclude equipment from accessing fields, or frost protection through irrigation can create artificially saturated soil conditions. In these cases, prioritizing seedlings that may deacclimate sooner can prevent spring cold damage, or budburst before lifting can occur.

Additionally, increased understanding of seedling cold hardiness status can be applied to reduce regeneration mortality due to cold events after planting and before seedlings acclimate to site climate. This may be especially true for fall planted seedlings, a regionally variable component of regeneration planting in North America (Grossnickle & MacDonald, 2021). If fall planted seedlings are not properly conditioned through nursery culture and chilling temperature exposure, they may be damaged on planting sites that are much colder than nursery conditions. This issue can be compounded by site microclimates, and little or no protection from canopy on many regeneration sites (Krasowski & Simpson, 2001; Marsh et al., 2022).

Reforestation for timber production can see stem defect at mid rotation as a result of early cold damage. This is demonstrated by observation of increased stem canker with cold damage (Reich & Kamp, 1993). As we continue to see changing climate conditions, we may see increased risk of cold damage in certain populations, regions, or microclimatic zones (Ford et al., 2016). This may be due to changes in climate related to temperature or related to precipitation and snowpack. Outplanted seedlings that may have been insulated by snowpack may now be exposed to damaging cold temperatures, especially during the acclimation and deacclimation shoulder seasons (Hankin et al., 2019).

Climate change is presenting novel risks to tree populations that are unable to migrate as rapidly as these risks are being realized. Population movement to mitigate the impact of climate change on forests, otherwise known as assisted migration, is proposed for Douglas-fir (Bansal et al., 2015; Chmura et al., 2011). Though there is evidence that this will mitigate some challenges associated with high temperature and drought, there may also be risks of increased cold related damage, and attempts to move Douglas-fir seed sources to Eastern Europe have experienced this risk (Boiffin et al., 2017; Isaac-Renton et al., 2014; Malmqvist et al., 2018). Thus, the establishment of forests that will have both an advantage in future climates over native populations and will tolerate low temperature extremes is a balance that requires improved understanding of plant phenology. The results of this model demonstrate that through calibration and parameter selection, understanding of population specific temperature response can be built. This understanding can be applied by foresters and restoration ecologists to decision making around assisted migration and population selection. Determining accurate population specific parameter ranges for the acclimation and deacclimation phases as well as maximum cold hardiness can be applied to determine which populations will be at greatest risks to cold damage during the fall, spring, or winter, respectively.

Cold damage can be particularly catastrophic during forest regeneration or seedling production, as seedlings are more vulnerable than mature trees (Duryea & McClain, 1984). Successful forest regeneration in the future will require the production of hardy seedlings and the ability for planted and naturally regenerated seedlings to withstand site conditions. Additionally, populations used for assisted migration must be able to withstand the current conditions for forest establishment to be successful (Isaac-Renton et al., 2014). Thus, implementing cold hardiness and phenology models, such as the one demonstrated here, in both production and

reforestation efforts, has the potential to be a useful tool as continued climate change threatens forest success.

Collectively, the results suggest that this model provides adequate predictions when data modeled are similar. Modeling of data from too broad of a range, or for seedlings with different life histories, exemplified by the Stevenson et al. data, creates large errors and reduces model efficiency. Several parameters are candidates for fixed values, which will improve model calibration in future uses, and reveal biological thresholds for cold hardiness development.

Future Direction

Unified Dormancy Phenology Model: The future direction of this body of work will be to unify the cold hardiness model used here with proposed bud burst models that also require only temperature inputs. A unified model can be used to predict future risks due to changes in chilling and forcing. The unified model will provide the benefit of predicting two major risks associated with the dormant season of temperate trees: cold damage, and delayed bud burst. This model will be used to identify populations vulnerable to changes in phenology, and the associated risks. Machine learning methods may be utilized to improve model performance and applicability to large scale data.

Comparative Phenology Modeling: Additionally, future work will compare this model to other proposed models that include photoperiod as predictors. The model used here requires less input than previously proposed models for temperate conifer cold hardiness (Greer et al., 2001; Leinonen, 1996; Leinonen et al., 1995). Previous work has shown that day length impacts conifer cold hardiness development in controlled environments (Leinonen et al., 1995; G. Zhang et al., 2003), but the results of the model used here suggest that the natural day length difference across

the range of a population may only have minor effects during historical and current climate conditions.

Expanded Modeling of Taxa: The next phase of model testing may be expanded to include other temperate woody species. This work will aim to determine if a temperature driven process-based cold hardiness model is adequate for predicting cold hardiness across temperate wood plants.

Expanded Modeling of Future Climates: Future objectives include testing model iterations with historical and projected future simulated climate data. As seasonality and phenology change due to climate change, photoperiod may become a more impactful predictor of phenology. Future divergence of photoperiod and chilling-forcing dynamics may impact cold hardiness and other phenological events. These continued efforts will also be leveraged to answer questions about plasticity of phenological responses through acclimation and the limitation of within and across population genetic adaptation.

Works Cited

Aitken, S. N., & Adams, W. T. (1997). Spring cold hardiness under strong genetic control in Oregon populations of Pseudotsuga menziesii var. Menziesii. *Canadian Journal of Forest Research*, *27*(11), 1773–1780. https://doi.org/10.1139/x97-151

Aitken, S. N., Adams, W. T., Schermann, N., & Fuchigami, L. H. (1996). Family variation for fall cold hardiness in two Washington populations of coastal Douglas-fir (Pseudotsuga menziesii var. Menziesii (Mirb.) Franco). *Forest Ecology and Management*, *80*(1), 187–195. https://doi.org/10.1016/0378-1127(95)03609-1

Allan, R. P., Hawkins, E., Bellouin, N., & Collins, B. (2021). *IPCC, 2021: Summary for Policymakers* (V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou, Eds.). Cambridge University Press. https://centaur.reading.ac.uk/101317/

Anekonda, T. S., Adams, W. T., & Aitken, S. N. (2000). Cold Hardiness Testing for Douglas-Fir Tree Improvement Programs: Guidelines for a Simple, Robust, and Inexpensive Screening Method. *Western Journal of Applied Forestry*, *15*(3), 129–136. https://doi.org/10.1093/wjaf/15.3.129

Arora, R., & Taulavuori, K. (2016). Increased Risk of Freeze Damage in Woody Perennials VIS-À-VIS Climate Change: Importance of Deacclimation and Dormancy Response. *Frontiers in Environmental Science*, *4*. https://www.frontiersin.org/article/10.3389/fenvs.2016.00044

Bannister, P., & Neuner, G. (2001). Frost Resistance and the Distribution of Conifers. In F. J. Bigras & S. J. Colombo (Eds.), *Conifer Cold Hardiness* (pp. 3–21). Springer Netherlands. https://doi.org/10.1007/978-94-015-9650-3_1

Bansal, S., St. Clair, J. B., Harrington, C. A., & Gould, P. J. (2015). Impact of climate change on cold hardiness of Douglas-fir (*Pseudotsuga menziesii*): Environmental and genetic considerations. *Global Change Biology*, *21*(10), 3814–3826. https://doi.org/10.1111/gcb.12958

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A Fresh Approach to Numerical Computing. *SIAM Review*, *59*(1), 65–98. https://doi.org/10.1137/141000671

Bigras, F. J., Ryyppö, A., Lindström, A., & Stattin, E. (2001). Cold Acclimation and Deacclimation of Shoots and Roots of Conifer Seedlings. In F. J. Bigras & S. J. Colombo (Eds.), *Conifer Cold Hardiness* (pp. 57–88). Springer Netherlands. https://doi.org/10.1007/978-94-015-9650-3_3

Boiffin, J., Badeau, V., & Bréda, N. (2017). Species distribution models may misdirect assisted migration: Insights from the introduction of Douglas-fir to Europe. *Ecological Applications*, *27*(2), 446–457. https://doi.org/10.1002/eap.1448

Canada, E. and C. C. (2011, October 31). *Historical Data—Climate—Environment and Climate Change Canada*. https://climate.weather.gc.ca/historical_data/search_historic_data_e.html

Chakraborty, D., Matulla, C., Andre, K., Weissenbacher, L., & Schueler, S. (2019). Survival of Douglas-fir provenances in Austria: Site-specific late and early frost events are more important than provenance origin. *Annals of Forest Science*, *76*(4), 100. https://doi.org/10.1007/s13595-019-0883-2

Chmura, D. J., Anderson, P. D., Howe, G. T., Harrington, C. A., Halofsky, J. E., Peterson, D. L., Shaw, D. C., & Brad St.Clair, J. (2011). Forest responses to climate change in the northwestern United States: Ecophysiological foundations for adaptive management. *Forest Ecology and Management*, *261*(7), 1121–1142. https://doi.org/10.1016/j.foreco.2010.12.040

Duryea, M. L., & McClain, K. M. (1984). Altering Seedling Physiology to Improve Reforestation Success. In M. L. Duryea & G. N. Brown (Eds.), *Seedling physiology and reforestation success: Proceedings of*

the Physiology Working Group Technical Session (pp. 77–114). Springer Netherlands. https://doi.org/10.1007/978-94-009-6137-1_5

Ettinger, A. K., Chamberlain, C. J., Morales-Castilla, I., Buonaiuto, D. M., Flynn, D. F. B., Savas, T., Samaha, J. A., & Wolkovich, E. M. (2020). Winter temperatures predominate in spring phenological responses to warming. *Nature Climate Change*, *10*(12), Article 12. https://doi.org/10.1038/s41558-020-00917-3

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Fargione, J., Haase, D. L., Burney, O. T., Kildisheva, O. A., Edge, G., Cook-Patton, S. C., Chapman, T., Rempel, A., Hurteau, M. D., Davis, K. T., Dobrowski, S., Enebak, S., De La Torre, R., Bhuta, A. A. R., Cubbage, F., Kittler, B., Zhang, D., & Guldin, R. W. (2021). Challenges to the Reforestation Pipeline in the United States. *Frontiers in Forests and Global Change*, *4*. https://www.frontiersin.org/articles/10.3389/ffgc.2021.629198

Ferguson, J. C., Moyer, M. M., Mills, L. J., Hoogenboom, G., & Keller, M. (2014). Modeling Dormant Bud Cold Hardiness and Budbreak in Twenty-Three Vitis Genotypes Reveals Variation by Region of Origin. *American Journal of Enology and Viticulture*, *65*(1), 59–71. https://doi.org/10.5344/ajev.2013.13098

Ferguson, J. C., Tarara, J. M., Mills, L. J., Grove, G. G., & Keller, M. (2011). Dynamic thermal time model of cold hardiness for dormant grapevine buds. *Annals of Botany*, *107*(3), 389–396. https://doi.org/10.1093/aob/mcq263

Field, C. B., & Barros, V. R. (2014). *Climate Change 2014 – Impacts, Adaptation and Vulnerability: Regional Aspects*. Cambridge University Press.

Fisker, S. E. (1992). *Chlorophyll fluorescence as a measure of cold hardiness and freezing stress in 1 1 Douglas-fir seedlings: Response to seasonal changes in the nursery. [Master's Thesis, Oregon State University]*. ScholarsArchive@OSU.

Ford, K. R., Harrington, C. A., Bansal, S., Gould, P. J., & St. Clair, J. B. (2016). Will changes in phenology track climate change? A study of growth initiation timing in coast Douglas-fir. *Global Change Biology*, *22*(11), 3712–3723. https://doi.org/10.1111/gcb.13328

Fu, Y. H., Zhao, H., Piao, S., Peaucelle, M., Peng, S., Zhou, G., Ciais, P., Huang, M., Menzel, A., Peñuelas, J., Song, Y., Vitasse, Y., Zeng, Z., & Janssens, I. A. (2015). Declining global warming effects on the phenology of spring leaf unfolding. *Nature*, *526*(7571), Article 7571. https://doi.org/10.1038/nature15402

Greer, D., Leinonen, I., & Repo, T. (2001). Modelling Cold Hardiness Development and Loss in Conifers. In *Conifer Cold Hardiness* (Vol. 1, pp. 437–460). https://doi.org/10.1007/978-94-015-9650-3_16

Grossnickle, S., & MacDonald, J. (2021). Fall Planting in Northern Forests as a Reforestation Option: Rewards, Risks, and Biological Considerations. *Tree Planters' Notes*, *64*, 57.

Haase, D. L., Khadduri, N., Mason, E., & Dumroese, R. K. (2016). Relationships among chilling hours, photoperiod, calendar date, cold hardiness, seed source, and storage of Douglas-fir seedlings. *Tree Planters' Notes*, *59*(1), 12.

Hankin, L. E., Higuera, P. E., Davis, K. T., & Dobrowski, S. Z. (2019). Impacts of growing-season climate on tree growth and post-fire regeneration in ponderosa pine and Douglas-fir forests. *Ecosphere*, *10*(4), e02679. https://doi.org/10.1002/ecs2.2679

Hänninen, H., & Kramer, K. (2007). A framework for modelling the annual cycle of trees in boreal and temperate regions. *Silva Fennica*, *41*(1). https://doi.org/10.14214/sf.313

Harrington, C. A., & Gould, P. J. (2015). Tradeoffs between chilling and forcing in satisfying dormancy requirements for Pacific Northwest tree species. *Frontiers in Plant Science*, *6*. https://www.frontiersin.org/article/10.3389/fpls.2015.00120

Harrington, C. A., Gould, P. J., & St.Clair, J. B. (2010). Modeling the effects of winter environment on dormancy release of Douglas-fir. *Forest Ecology and Management*, *259*(4), 798–808. https://doi.org/10.1016/j.foreco.2009.06.018

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *28*(1), 100–108. https://doi.org/10.2307/2346830

Isaac-Renton, M. G., Roberts, D. R., Hamann, A., & Spiecker, H. (2014). Douglas-fir plantations in Europe: A retrospective test of assisted migration to address climate change. *Global Change Biology*, *20*(8), 2607–2617. https://doi.org/10.1111/gcb.12604

Krasowski, M. J., & Simpson, D. G. (2001). Frost-Related Problems in the Establishment of Coniferous Forests. In F. J. Bigras & S. J. Colombo (Eds.), *Conifer Cold Hardiness* (pp. 253–285). Springer Netherlands. https://doi.org/10.1007/978-94-015-9650-3_10

Leinonen, I. (1996). A Simulation Model for the Annual Frost Hardiness and Freeze Damage of Scots Pine. *Annals of Botany*, *78*(6), 687–693. https://doi.org/10.1006/anbo.1996.0178

Leinonen, I., Repo, T., Hänninen, H., & Burr, K. E. (1995). A Second-order Dynamic Model for the Frost Hardiness of Trees. *Annals of Botany*, *76*(1), 89–95. https://doi.org/10.1006/anbo.1995.1082

Malmqvist, C., Wallertz, K., & Johansson, U. (2018). Survival, early growth and impact of damage by late-spring frost and winter desiccation on Douglas-fir seedlings in southern Sweden. *New Forests*, *49*(6), 723–736. https://doi.org/10.1007/s11056-018-9635-7

Marsh, C., Crockett, J. L., Krofcheck, D., Keyser, A., Allen, C. D., Litvak, M., & Hurteau, M. D. (2022). Planted seedling survival in a post-wildfire landscape: From experimental planting to predictive

probabilistic surfaces. *Forest Ecology and Management*, *525*, 120524. https://doi.org/10.1016/j.foreco.2022.120524

Miller, L. (2011, September 7). *Digitizing data from old figures with ImageJ*. https://lukemiller.org/index.php/2011/09/digitizing-data-from-old-figures-with-imagej/

NOAA. (2022). *Climate Data Online (CDO)—The National Climatic Data Center's (NCDC) Climate Data Online (CDO) provides free access to NCDC's archive of historical weather and climate data in addition to station history information. | National Climatic Data Center (NCDC)*. https://www.ncei.noaa.gov/cdo-web/

Nowak, R. D. (1997). Optimal signal estimation using cross-validation. *IEEE Signal Processing Letters*, *4*(1), 23–25. https://doi.org/10.1109/97.551692

Pawson, S. M., Brin, A., Brockerhoff, E. G., Lamb, D., Payn, T. W., Paquette, A., & Parrotta, J. A. (2013). Plantation forests, climate change and biodiversity. *Biodiversity and Conservation*, *22*(5), 1203–1227. https://doi.org/10.1007/s10531-013-0458-8

Peñuelas, J., Rutishauser, T., & Filella, I. (2009). Phenology Feedbacks on Climate Change. *Science*, *324*(5929), 887–888. https://doi.org/10.1126/science.1173004

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., & Zhu, X. (2019). Plant phenology and global climate change: Current progresses and challenges. *Global Change Biology*, *25*(6), 1922–1940. https://doi.org/10.1111/gcb.14619

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Reich, R. W., & Kamp, B. J. van der. (1993). Frost, canker, and dieback of Douglas-fir in the central interior of British Columbia. *Canadian Journal of Forest Research*, *23*(3), 373–379. https://doi.org/10.1139/x93-054

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., & Toomey, M. (2013). Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agricultural and Forest Meteorology*, *169*, 156–173. https://doi.org/10.1016/j.agrformet.2012.09.012

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929. https://doi.org/10.1111/ecog.02881

Sakai, A., & Weiser, C. J. (1973). Freezing Resistance of Trees in North America with Reference to Tree Regions. *Ecology*, *54*(1), 118–126. https://doi.org/10.2307/1934380

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), Article 7. https://doi.org/10.1038/nmeth.2089

St. Clair, J. B. (2006). Genetic variation in fall cold hardiness in coastal Douglas-fir in western Oregon and Washington. *Canadian Journal of Botany*, *84*(7), 1110–1121. https://doi.org/10.1139/b06-084

Stabenow, D. (2021, March 18). *S.866 - 117th Congress (2021-2022): Repairing Existing Public Land by Adding Necessary Trees Act* (2021/2022) [Legislation]. http://www.congress.gov/

Stevenson, J. F., Hawkins, B. J., & Woods, J. H. (1999). Spring and Fall Cold Hardiness in Wild and Selected Seed Sources of Coastal Douglas-fir. *Silvae Genetica*, *48*(1), 6.

Tharammal, T., Bala, G., Devaraju, N., & Nemani, R. (2019). A review of the major drivers of the terrestrial carbon uptake: Model-based assessments, consensus, and uncertainties. *Environmental Research Letters*, *14*(9), 093005. https://doi.org/10.1088/1748-9326/ab3012

Timmis, R., Flewelling, J., & Talbert, C. (1994). Frost injury prediction model for Douglas-fir seedlings in the Pacific Northwest. *Tree Physiology*, *14*(7-8–9), 855–869. https://doi.org/10.1093/treephys/14.7-8-9.855

USGS. (2022, October 9). *The National Map—Data Delivery | U.S. Geological Survey*. https://www.usgs.gov/the-national-map-data-delivery

Wang, J., Guan, Y., Wu, L., Guan, X., Cai, W., Huang, J., Dong, W., & Zhang, B. (2021). Changing Lengths of the Four Seasons by Global Warming. *Geophysical Research Letters*, *48*(6), e2020GL091753. https://doi.org/10.1029/2020GL091753

Wang, T., Hamann, A., Spittlehouse, D., & Carroll, C. (2016). Locally Downscaled and Spatially Customizable Climate Data for Historical and Future Periods for North America. *PLOS ONE*, *11*(6), e0156720. https://doi.org/10.1371/journal.pone.0156720

Wisniewski, M., Nassuth, A., & Arora, R. (2018). Cold Hardiness in Trees: A Mini-Review. *Frontiers in Plant Science*, *9*. https://www.frontiersin.org/article/10.3389/fpls.2018.01394

WSU. (2022). *Home | AgWeatherNet at Washington State University*. https://weather.wsu.edu/

Yun, K., & Kim, S.-H. (2022). *Cropbox: A declarative crop modeling framework* (p. 2022.10.10.511649). bioRxiv. https://doi.org/10.1101/2022.10.10.511649

Zhang, G., Ryyppö, A., Vapaavuori, E., & Repo, T. (2003). Quantification of additive response and stationarity of frost hardiness by photoperiod and temperature in Scots pine. *Canadian Journal of Forest Research*, *33*(9), 1772–1784. https://doi.org/10.1139/x03-100

Zhang, R., Lin, J., Wang, F., Shen, S., Wang, X., Rao, Y., Wu, J., & Hänninen, H. (2021). The chilling requirement of subtropical trees is fulfilled by high temperatures: A generalized hypothesis for tree endodormancy release and a method for testing it. *Agricultural and Forest Meteorology*, *298–299*, 108296. https://doi.org/10.1016/j.agrformet.2020.108296

Zhou, S., Liang, J., Lu, X., Li, Q., Jiang, L., Zhang, Y., Schwalm, C. R., Fisher, J. B., Tjiputra, J., Sitch, S., Ahlström, A., Huntzinger, D. N., Huang, Y., Wang, G., & Luo, Y. (2018). Sources of Uncertainty in Modeled Land Carbon Storage within and across Three MIPs: Diagnosis with Three New Techniques. *Journal of Climate*, *31*(7), 2833–2851. https://doi.org/10.1175/JCLI-D-17-0357.1

Supplementary Table S1. Supplemental table showing data sources reviewed for Douglas-fir cold hardiness and budburst data. The top four rows were included in the modeling data set. Columns two through six indicate the presence or absence of various data characteristics. Phase indicates what phenological phase of cold hardiness was reported, either the full dormancy season, just acclimation (Acc), just deacclimation (DeAcc), or both Acc and DeAcc but excluding peak cold hardiness. Method indicates the cold hardiness testing method, either visual, freeze induced electrolyte leakage (FIEL), or chlorophyll fluorescence (CFL). For data not used for modeling the reason for excluding is listed in the last column.

Citation	Model	Seed Source ID	Site Id	Cold Hardiness Data	Budburst Data	Temperature Data	Phase	Method	Reason for Excluding
Fisker, Sue. (1993). Chlorophyll fluorescence as a measure of cold hardiness and freezing stress in 1 + 1 Douglas-fir seedlings: response to seasonal changes in the nursery [Master's Thesis, Oregon State University]. ScholarsArchive@OSU.	Yes	Yes	Yes	Yes	No	No	Full	Visual	NA
Haase, D. L., Khadduri, N., Mason, E., & Dumroese, R. K. (2016). Relationships among chilling hours, photoperiod, calendar date, cold hardiness, seed source, and storage of Douglas-fir seedlings. 59(1), 12.	Yes	Yes	Yes	Yes	No	Yes	Acc	Visual	NA
Stevenson, J. F., Hawkins, B. J., & Woods, J. H. (1999). Spring and Fall Cold Hardiness in Wild and Selected Seed Sources of Coastal Douglas-fir. Silvae Genetica, 48(1), 29-34.	Yes	Yes	No	Yes	Yes	Yes	Full	FIEL	NA
Timmis, R., Flewelling, J., & Talbert, C. (1994). Frost injury prediction model for Douglas-fir seedlings in the Pacific Northwest. Tree Physiology, 14(7-8–9), 855–869. https://doi.org/10.1093/treephys/14.7-8-9.855	Yes	Yes	Yes	Yes	No	No	Full	Visual	NA
Aitken, S. N., & Adams, W. T. (1997). Spring cold hardiness under strong genetic control in Oregon populations of Pseudotsuga menziesii var. Menziesii. Canadian Journal of Forest Research, 27(11), 1773–1780. https://doi.org/10.1139/x97-151	No	Yes	Yes	Yes	Yes	No	DeAcc	Visual	Limited temporal replication; Data in units other than LT_{50}
Aitken, S. N., Adams, W. T., Schermann, N., & Fuchigami, L. H. (1996). Family variation for fall cold hardiness in two Washington populations of coastal Douglas-fir (Pseudotsuga menziesii var. Menziesii (Mirb.) Franco). Forest Ecology and Management, 80(1), 187–195. https://doi.org/10.1016/0378-1127(95)03609-1	No	Yes	Yes	Yes	Yes	Yes	Full	Visual	Cold hardiness data not present in publication
Alden, John; 1971; Freezing resistance of tissues of the twig in Douglas-fir [Doctoral Dissertation, Oregon State University]. ScholarsArchive@OSU.	No	No	No	Yes	No	No	Full	Visual	Data in units other than LT_{50}
Bailey, J. D., & Harrington, C. A. (2006). Temperature regulation of bud-burst phenology within and among years in a young Douglas-fir (Pseudotsuga menziesii) plantation in western Washington, USA. Tree Physiology, 26(4), 421–430. https://doi.org/10.1093/treephys/26.4.421	No	Yes	Yes	No	Yes	Yes	DeAcc	NA	Cold hardiness not tested
Bansal, S., Harrington, C. A., & St. Clair, J. B. (2016). Tolerance to multiple climate stressors: A case study of Douglas-fir drought and cold hardiness. Ecology and Evolution, 6(7), 2074–2083. https://doi.org/10.1002/ece3.2007	No	Yes	Yes	Yes	No	Yes	Acc	Visual	Limited temporal replication
Bansal, S., St. Clair, J. B., Harrington, C. A., & Gould, P. J. (2015). Impact of climate change on cold hardiness of Douglas-fir (Pseudotsuga menziesii): Environmental and genetic considerations. Global Change Biology, 21(10), 3814–3826. https://doi.org/10.1111/gcb.12958	No	Yes	Yes	Yes	No	Yes	Acc	Visual	Limited temporal replication
Campbell, R. K., & Sorensen, F. C. (1973). Cold-Acclimation in Seedling Douglas-Fir Related to Phenology and Provenance. Ecology, 54(5), 1148–1151. https://doi.org/10.2307/1935582	No	Yes	Yes	No	Yes	No	DeAcc	Other	Cold hardiness determined by occurrence of natural frost damage

Supplementary Table S1 Continued. Supplemental table showing data sources reviewed for Douglas-fir cold hardiness and budburst data. The top four rows were included in the modeling data set. Columns two through six indicate the presence or absence of various data characteristics. Phase indicates what phenological phase of cold hardiness was reported, either the full dormancy season, just acclimation (Acc), just deacclimation (Deacc), or both Acc and Deacc but excluding peak cold hardiness. Method indicates the cold hardiness testing method, either visual, freeze induced electrolyte leakage (FIEL), or chlorophyll fluorescence (CFL). For data not used for modeling the reason for excluding is listed in the last column.

Citation	Model	Seed Source ID	Site Id	Cold Hardiness Data	Budburst Data	Temperature Data	Phase	Method	Reason for Excluding
Campbell, R. K., & Sugano, A. I. (1979). Genecology of Bud-Burst Phenology in Douglas-Fir: Response to Flushing Temperature and Chilling. Botanical Gazette, 140(2), 223–231. https://doi.org/10.1086/337079	No	Yes	No	No	Yes	Yes	DeAcc	NA	Cold hardiness not tested
Guak, S., Olsyzk, D. M., Fuchigami, L. H., & Tingey, D. T. (1998). Effects of elevated CO2 and temperature on cold hardiness and spring bud burst and growth in Douglas-fir (Pseudotsuga menziesii). Tree Physiology, 18(10), 671–679. https://doi.org/10.1093/treephys/18.10.671	No	Yes	Yes	Yes	Yes	Yes	Full	Visual	Cold hardiness data averaged by site or seed source; Controlled environment
Hawkins, B. J., & Stoehr, M. (2009). Growth, phenology, and cold hardiness of 32 Douglas-fir full-sib families. Canadian Journal of Forest Research, 39(10), 1821–1834. https://doi.org/10.1139/X09-092	No	Yes	Yes	Yes	Yes	No	Acc& DeAcc	CFL	Data in units other than LT_{50}
Hawkins, B. J., & Stoehr, M. (2009). Growth, phenology, and cold hardiness of 32 Douglas-fir full-sib families. Canadian Journal of Forest Research, 39(10), 1821–1834. https://doi.org/10.1139/X09-092	No	Yes	Yes	Yes	Yes	Yes	Full	FIEL	Cold hardiness data not present in publication
Hawkins, B. J., Guest, H. J., & Kolotelo, D. (2003). Freezing tolerance of conifer seeds and germinants. Tree Physiology, 23(18), 1237–1246. https://doi.org/10.1093/treephys/23.18.1237	No	Yes	Yes	Yes	No	Yes	NA	Visual	Plant ontogeny different from other sources
Leinonen, I., Repo, T., Hänninen, H., & Burr, K. E. (1995). A Second-order Dynamic Model for the Frost Hardiness of Trees. Annals of Botany, 76(1), 89–95. https://doi.org/10.1006/anbo.1995.1082	No	Yes	Yes	Yes	Yes	Yes	Full	FIEL	Southwestern seed source; Controlled environment
Malmqvist, C., Wallin, E., Lindström, A., & Säll, H. (2017). Differences in bud burst timing and bud freezing tolerance among interior and coastal seed sources of Douglas fir. Trees, 31(6), 1987–1998. https://doi.org/10.1007/s00468-017-1603-x	No	Yes	Yes	Yes	Yes	Yes	DeAcc	FIEL	Cold hardiness data not present in publication; Single date tested
Schuch, Ursula; 1987; Frost hardiness of Douglas-fir (Pseudotsuga menziesii (Mirb.) Franco) seedlings raised in three nurseries [Master's Thesis Oregon State University]. ScholarsArchive@OSU.	No	Yes	Yes	Yes	Yes	No	Acc&DeAcc	Visual	Cold hardiness data averaged by site or seed source
St. Clair, J. B. (2006). Genetic variation in fall cold hardiness in coastal Douglas-fir in western Oregon and Washington. Canadian Journal of Botany, 84(7), 1110–1121. https://doi.org/10.1139/b06-084	No	Yes	Yes	Yes	No	Yes	Acc	Visual	Limited temporal replication
Stevenson, J. F., Hawkins, B. J., & Woods, J. H. (n.d.). Spring and Fall Cold Hardiness in Wild and Selected Seed Sources of Coastal Douglas-fir. 6.	No	No	Yes	Yes	Yes	No	Acc&DeAcc	Visual	Cold hardiness data not present in publication