

# **Understanding temperate tree and shrub phenology under climate change**

A dissertation presented

by

Catherine J Chamberlain

to

The Department of Organismic and Evolutionary Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biology

Harvard University

Cambridge, Massachusetts

June 2021

© 2021 Catherine J Chamberlain

All rights reserved.

## **Understanding temperate tree and shrub phenology under climate change**

### **Abstract**

An abstract should be less than 350 words. Here's some filler text. Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Temperate tree and shrub species are at risk of damage from late spring freezing events, however the extent of damage and the frequency and intensity of these events is still largely unknown. Individuals that initiate budburst before the last spring freeze are at risk of leaf tissue loss, damage to the xylem, and slowed, or even stalled, canopy development. These damaging events are called false springs and have the potential to detrimentally affect forest growth and sustainability, which can result in highly adverse ecological and economic consequences. It is crucial for scientists and management teams to have a better understanding of false spring future trends in order to better conserve our forest ecosystems.

Climate change has brought renewed interest to a major factor that shapes the life history of many non-tropical plant species: false spring events. While increased interest has led to a growing number of studies, much of the research takes a simplified view of these events, which I argued can lead to incorrect estimates and forecasting. Combining theory from ecology, climatology, physiology, biogeography and crop science we examined the effects of false

springs, and the complexity of factors that drive plants risk to frost damage.

Here, I asked which climatic and geographic factors are the strongest predictors of false springs across six tree species, and how these predictors have shifted with recent climate change. By investigating leafout observations of six deciduous tree species from Europe, we unraveled the effects of species, spring temperature, elevation, distance from the coast and NAO index on false spring risk with climate change. We found that recent warming has reshaped the influence of these factors and magnified species-level variation in false spring risk.

For this experiment, I investigate the interplay of false spring events and warmer winters (generally expected to reduce chilling) across eight temperate deciduous tree species examining a suite of phenological, growth and leaf tissue traits. I found that false springs increased tissue damage, decreased leaf toughness and leaf thickness, and slowed budburst to leafout timing—extending the period of maximum freezing risk. Chilling, however, shortened this period of maximum risk, even under false spring conditions, thus compensating for some of the more adverse phenological effects of false springs. Despite major shifts in phenology from false springs and chilling I did not find evidence of phenological reordering within the community of species we studied. The results instead suggest climate change will reshape forest communities through impacts on growth and leaf traits from the coupled effects of false springs and warmer winters under future climate change.

Often we use mixed models to answer ecological questions, though we do not always understand the intricacies of the model output, nor do we investigate what is missing from the model output. Here, I work to understand mixed models using simulation data and test myriad hypotheses through these simulations. These methods can be applied to many ecological questions investigating climate data across global habitats but I specifically investigate spring plant phenology and how using different methods to measure climate can impact predictions. Understanding and predicting spring plant phenology is essential for determining growing season length and predicting and individuals risk of false spring under climate change.

## **Contents**

## **List of Tables**

## **List of Figures**

## Acknowledgments

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

To my parents

# Introduction: Rethinking False Spring Risk

Authors:

C. J. Chamberlain <sup>1,2</sup>, B. I. Cook <sup>3</sup>, I. Garcia de Cortazar Atauri <sup>4</sup> & E. M. Wolkovich <sup>1,2,5</sup>

*Author affiliations:*

<sup>1</sup>Arnold Arboretum of Harvard University, 1300 Centre Street, Boston, Massachusetts, USA;

<sup>2</sup>Organismic & Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, Massachusetts, USA;

<sup>3</sup>NASA Goddard Institute for Space Studies, New York, New York, USA;

<sup>4</sup>French National Institute for Agricultural Research, INRA, US1116 AgroClim, F-84914 Avignon, France

<sup>5</sup>Forest & Conservation Sciences, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4

\*Corresponding author: 248.953.0189; cchamberlain@g.harvard.edu

*Keywords:* false spring, phenology, freezing tolerance, climate change, forest communities

*Paper type:* Opinion

## Abstract

Temperate plants are at risk of being exposed to late spring freezes. These freeze events—often called false springs—are one of the strongest factors determining temperate plants species range limits and can impose high ecological and economic damage. As climate change may alter the prevalence and severity of false springs, our ability to forecast such events has become more critical, and it has led to a growing body of research. Many false spring studies largely simplify the myriad complexities involved in assessing false spring risks and damage. While these studies have helped advance the field and may provide useful estimates at large scales, studies at the individual to community levels must integrate more complexity for accurate predictions of plant damage from late spring freezes. Here we review current metrics of false spring, and how, when and where plants are most at risk of freeze damage. We highlight how life stage, functional group, species differences in morphology and phenology, and regional climatic differences contribute to the damage potential of false springs. More studies aimed at understanding relationships among species tolerance and avoidance strategies, climatic regimes, and the environmental cues that underlie spring phenology would improve predictions at all biological levels. An integrated approach to assessing past and future spring freeze damage would provide novel insights into fundamental plant biology, and offer more robust predictions as climate change progresses, which is essential for mitigating the adverse ecological and economic effects of false springs.

## Introduction

Plants from temperate environments time their growth each spring to follow rising temperatures alongside the increasing availability of light and soil resources. During this time, individuals that budburst before the last freeze date are at risk of leaf loss, damaged wood tissue, and slowed canopy development (??). These damaging late-spring freezes are also known as false springs, and are widely documented to result in adverse ecological and economic consequences (??).

Climate change is expected to cause an increase in damage from false spring events due to earlier spring onset and potentially greater fluctuations in temperature in some regions (??). In recent years multiple studies have documented false springs (????) and some have linked these events to climate change (?????). This interest in false springs has led to a growing body of research investigating the effects across ecosystems. Such work builds on decades of research across the fields of ecophysiology, climatology, ecosystem and alpine ecology examining how spring frosts have shaped the life history strategies of diverse species and determine the dynamics of many ecosystems, especially in temperate and boreal systems where frost is a common obstacle to plant growth. While this literature has highlighted the complexity of factors that underlie false springs, many current estimates of false spring risk and damage seek to simplify the process.

Current metrics for estimating false springs events often require only two pieces of information: an estimate for the start of biological ‘spring’ (i.e., budburst) and whether temperatures below a particular threshold occurred in the following week. Such estimates provide a basic understanding of potential false spring damage. However, they inherently assume consistency of damage across functional groups, species, life stages, and regional climates, ignoring that such factors can greatly impact plants’ false spring risk. As a result, such indices may lead to inaccurate estimates and predictions, slowing our progress in understanding false spring events and how they may shift with climate change. To produce accurate predictions, researchers need improved methods that can properly evaluate the effects of false springs across diverse species and climate regimes.

In this paper we highlight the complexity of factors driving a plant’s false spring risk and provide a road map for improved metrics. We show how freeze temperature thresholds (?), location within a forest or canopy (?), interspecific variation in tolerance and avoidance strategies (??), and regional effects (?) unhinge simple metrics of false spring. We argue that while current simplified metrics have advanced the field and offer further advances at large scales, greater progress can come from new approaches. In particular, approaches that integrate the major factors shaping false spring risk would help accurately determine current

false spring damage and improve predictions of spring freeze risk under a changing climate — while potentially providing novel insights to how plants respond to and are shaped by spring frost. We focus on temperate forests, where much recent and foundational research has been conducted, but our approaches can be extended to other ecosystems shaped by spring frost events.

## Defining false springs

### When are plants vulnerable to frost damage?

At the level of an individual plant, vulnerability to frost damage varies across tissues and seasonally with plant development. Different tissues are often more or less sensitive to low temperatures. Flower and fruit tissues are often easily damaged by freezing temperatures (????), while wood and bark tissues can survive lower temperatures through various methods (?). Similar to wood and bark, leaf and bud tissues can often survive lower temperatures without damage (?). However, for most tissues, tolerance of low temperatures varies seasonally with the environment through the development of cold hardiness (i.e. freezing tolerance), which allows plants to survive colder winter temperatures through various physiological mechanisms (e.g., deep supercooling, increased solute concentration, and an increase in dehydrins and other proteins, ??).

Cold hardiness is an essential process for temperate plants to survive cold winters and hard freezes (?), especially in allowing bud tissue to overwinter without damage. Much cold hardiness research focuses on vegetative and floral buds, especially in the agricultural literature, where buds greatly determine crop success each season.

The actual temperatures that plants can tolerate vary strongly by species (Figure 1) and by a tissue's degree of cold hardiness. During the cold acclimation phase — which is generally triggered by shorter photoperiods (????) and, in some species, cold nights (??) — cold hardiness increases rapidly as temperate plants begin to enter dormancy. At maximum cold hardiness, vegetative tissues can generally sustain temperatures from -25°C to -40°C (???)

or sometimes even lower temperatures (to -60°C in extreme cases, ?). Freezing tolerance diminishes again during the cold deacclimation phase, when metabolism and development start to increase, and plant tissues become especially vulnerable.

Once buds begin to swell and deharden, freezing tolerance greatly declines and is lowest between budburst to leafout (i.e., -2 to -4°C for most species), then generally increases slightly once the leaves fully mature (i.e., at this stage most species can sustain temperatures at least 1-4°C lower than they can between budburst to leafout, ??). Thus, plants that have initiated budburst but have not fully leafed out are more likely to sustain damage from a false spring than individuals past the leafout phase (?). This timing is also most critical when compared to the fall onset of cold hardiness: as plants generally senesce as they gain cold hardiness, tissue damage during the fall is far less common and less critical (??).

Temperate forest plants, therefore, experience elevated risk of frost damage during the spring due both to the stochastic timing of frosts and the rapid decrease in freezing tolerance, which can have important consequences for individual plants all the way up to the ecosystem-level. Freezing temperatures following a warm spell can result in plant damage or even death (??). It can take 16-38 days for trees to refoliate after a spring freeze (????), which can detrimentally affect crucial processes such as carbon uptake and nutrient cycling (???). Additionally, plants can suffer greater long-term effects from the loss of photosynthetic tissue through impacts on multiple years of growth, reproduction, and canopy development (??). For these reasons, we focus primarily on spring freeze risk for the vegetative phases, specifically between budburst and leafout, when vegetative tissues are most at risk of damage.

**Figure 1:** A comparison of damaging spring freezing temperature thresholds across ecological and agronomic studies. Each study is listed on the vertical axis along with the taxonomic group of focus. Next to the species name is the freezing definition used within that study (e.g., 100% is 100% whole plant lethality). Each point is the best estimate recorded for the temperature threshold with standard deviation if indicated in the study.

## Current metrics of false spring

Currently researchers use several methods to define a false spring. A common definition is fundamentally empirical and describes a false spring as having two phases: rapid vegetative growth prior to a freeze and a post-freeze setback (?). However, as data on tissue damage is often lacking, most definitions do not require it. Other definitions focus on temperatures in the spring that are specific to certain regions (e.g., in ?, false spring for the Midwestern United States is defined as a warmer than average March, a freezing April, and enough growing degree days between budburst and the last freeze date). A widely used definition integrates a mathematical equation to quantify a false spring event. This equation, known as a False Spring Index (FSI), signifies the likelihood of damage to occur from a late spring freeze. Currently, FSI is evaluated annually by the day of budburst and the day of last spring freeze (often calculated at -2.2°C, ?) through the simple equation (?):

$$FSI = \text{Day of Year}(\text{LastSpringFreeze}) - \text{Day of Year}(\text{Budburst}) \quad (1)$$

Negative values indicate no-risk situations, whereas a damaging FSI is currently defined to be seven or more days between budburst and the last freeze date (Equation 1) (?). This index builds off our fundamental understanding that cold hardiness is low following budburst (i.e., the seven-day threshold attempts to capture that leaf tissue is at high risk of damage from frost in the period after budburst but before full leafout), and, by requiring only data on budburst and temperatures, this index can estimate where and when false springs occurred (or will occur) without any data on tissue damage.

## Measuring false spring in one temperate plant community

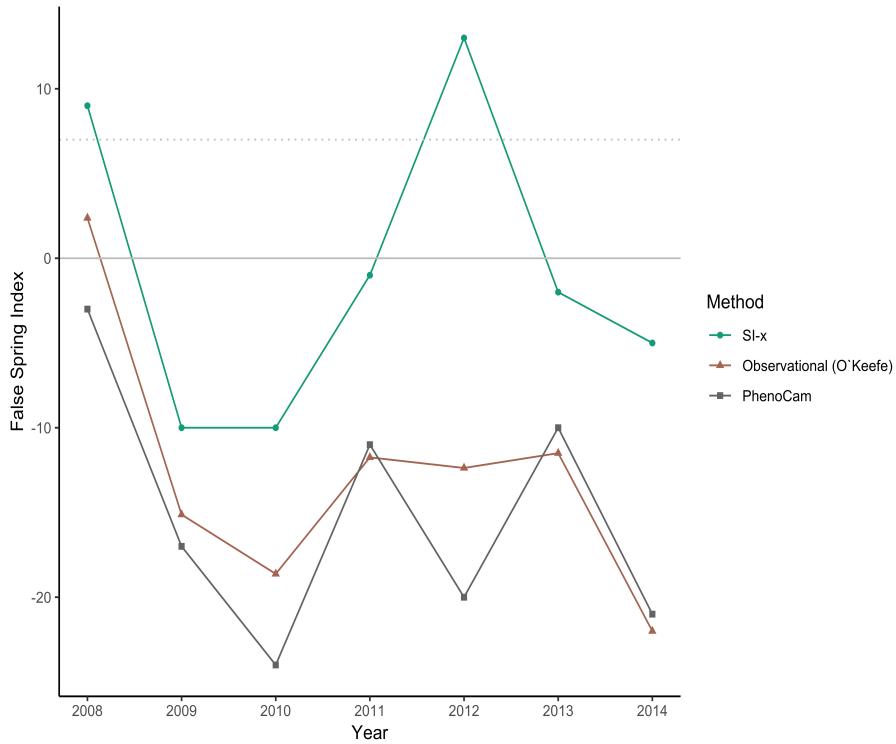
To demonstrate how the FSI definition works—and is often used—we applied it to data from the Harvard Forest Long-term Ecological Research program in Massachusetts. We selected this site as it has been well monitored for spring phenology through multiple methods for several years. While at the physiological level, frost damage is most likely to occur between budburst and leafout, data on the exact timing of these two events are rarely available and

surrogate data are often used to capture ‘spring onset’ (i.e., initial green-up) at the community level. We applied three commonly used methods to calculate spring onset: long-term ground observational data (?), PhenoCam data (?), and USA National Phenology Network’s (USA-NPN) Extended Spring Index (SI-x) “First Leaf - Spring Onset” data (?). These three methods for spring onset values require different levels of effort and are—thus—variably available for other sites. The local ground observational data (?)—available at few sites—require many hours of personal observation, but comes the closest to estimating budburst and leafout dates. PhenoCam data require only the hours to install and maintain a camera observing the canopy, then process the camera data to determine canopy color dynamics over seasons and years. Finally, SI-x data can be calculated for most temperate sites, as the index was specifically designed to provide an available, comparable estimate of spring onset across sites. Once calculated for this particular site we inputted our three estimates of spring onset into the FSI equation (Equation 1) to determine the FSI from 2008 to 2014 (Figure 2).

Each methodology rendered different FSI values, suggesting different false spring damage for the same site over the same years. For most years, the observational FSI and PhenoCam FSI are about 10-15 days lower than the SI-x data. This is especially important for 2008, when the SI-x data and observational data indicate a false spring year, whereas the PhenoCam data do not. In 2012, the observational data and PhenoCam data diverge slightly and the PhenoCam FSI is over 30 days less than the SI-x value.

The reason for these discrepancies is that each method effectively evaluates spring onset by integrating different attributes such as age, species or functional group. Spring phenology in temperate forests typically progresses by functional group: understory species and younger trees tend to initiate budburst first, whereas larger canopy species start later in the season (??). The different FSI values determined in Figure 2 exemplify the differences in functional group spring onset dates and illustrate variations in forest demography and phenology. While the SI-x data (based on observations of early-active shrub species, especially including the—non-native to Massachusetts—species lilac, *Syringa vulgaris*) may best capture understory dynamics, the PhenoCam and observational FSI data integrate over larger canopy species,

which burst bud later and thus are at generally lower risk of false springs. Such differences are visible each year, as the canopy-related metrics show lower risk, but are especially apparent in 2012. In 2012, a false spring event was reported through many regions of the US due to warm temperatures occurring in March (?). These high temperatures would most likely have been too early for larger canopy species to burst bud but they would have affected smaller understory species, as is seen by the high risk of the SI-x FSI in Figure 2.



**Figure 2:** False Spring Index (FSI) values from 2008 to 2014 vary across methods. To calculate spring onset, we used the USA-NPN Extended Spring Index tool for the USA-NPN FSI values, which are the circles (USA-NPN, 2016), long-term ground observational data for the observed FSI values, which are the triangles (O'Keefe, 2014), and near-surface remote-sensing canopy data for the PhenoCam FSI values, which are the squares (Richardson, 2015). See the Supplement for extended details. The solid grey line at  $FSI=0$  indicates a boundary between a likely false spring event or not, with positive numbers indicating a false spring likely occurred and negative numbers indicating a false spring most likely did not occur. The dotted grey line at  $FSI=7$  indicates the seven-day threshold frequently used in false spring definitions, which suggests years with FSI values greater than seven very likely had false spring events.

Differing FSI estimated from our three metrics of spring onset for the same site and years

highlight variation across functional groups, which FSI work currently ignores — instead using one metric of spring onset (often from SI-x data, which is widely available) and assuming it applies to the whole community of plants (????). As the risk of a false spring varies across habitats and functional groups (?) one spring onset date cannot be used as an effective proxy for all species and researchers should more clearly align their study questions and methods. FSI using such estimates as the SI-x may discern large-scale basic trends across space or years, but require validation with ground observations to be applied to any particular location or functional group of species.

Ideally researchers should first assess the forest demographics and functional groups relevant to their study question, then select the most appropriate method to estimate the date of budburst to determine if a false spring could have occurred. This, however, still ignores variation in the date of leafout (when cold tolerance increases slightly). Further, considering different functional groups is unlikely to be enough for robust predictions in regards to level of damage from a false spring, especially for ecological questions that operate at finer spatial and temporal scales. For many research questions—as we outline below—it will be important to develop false spring metrics that integrate species differences within functional groups, by considering the tolerance and avoidance strategies that species have evolved to mitigate false spring effects.

## Improving false spring definitions

### Integrating avoidance and tolerance strategies

While most temperate woody species use cold hardiness to tolerate low winter temperatures, species vary in how they minimize spring freeze damage through two major strategies: tolerance and avoidance. Many temperate forest plants employ various morphological or physiological traits to be more frost tolerant. Some species have increased ‘packability’ of leaf primordia in winter buds which may permit more rapid leafout (?) and thus shorten the exposure time of less resistant tissues. Other species have young leaves with more trichomes,

which protect leaf tissue from herbivory and additionally may act as a buffer against hard or radiative frosts (??). Species living in habitats with drier winters develop shoots and buds with decreased water content, which makes the buds more tolerant to drought and also to false spring events (??????). These strategies are probably only a few of the many ways plants avoid certain types of spring frost damage, thus more studies are needed to investigate the interplay between morphological and physiological traits and false spring tolerance.

Rather than being more tolerant of spring freezing temperatures, many species have evolved to avoid frosts by bursting bud later in the spring, well past the last frost event. Such species may lose out on early access to resources, but benefit from rarely, if ever, losing tissue to false spring events. They may further benefit from not needing traits related to frost tolerance (?).

The difference in budburst timing across temperate deciduous woody species—which effectively allows some species to avoid false springs—is determined by their responses to three environmental cues that initiate budburst: low winter temperatures (chilling), warm spring temperatures (forcing), and increasing photoperiods (?). The evolution of these three cues and their interactions have permitted temperate plant species to occupy more northern ecological niches (?) and decrease the risk of false spring damage for all species (?). Species that burst bud late are expected to have high requirements of chilling, forcing and/or photoperiod. For example, the combination of a high chilling and a spring forcing requirement (that is, a species that requires long periods of cool temperatures to satisfy a chilling requirement before responding to any forcing conditions) will avoid bursting bud during periods of warm temperatures too early due to insufficient chilling (?). An additional photoperiod requirement for budburst can also allow species to avoid false springs. Species with strong photoperiod cues have limited responses to spring forcing until a critical daylength is met, and thus are unlikely to have large advances in budburst with warming. Thus, as long as the critical daylength is past freeze events, these species will evade false spring events (?).

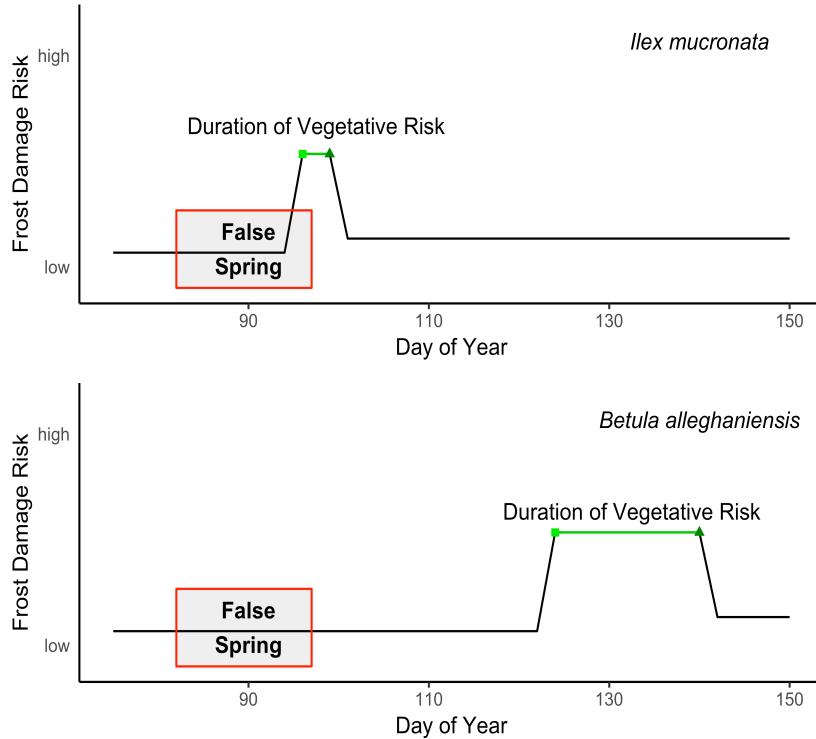
Given the diverse array of spring freezing defense mechanisms, improved metrics of false spring events would benefit from a greater understanding of avoidance and tolerance strategies

across species, especially under a changing climate. If research could build a framework to help classify species into what strategy they employ, estimates of false spring could quickly identify some species that effectively are never at risk of false spring events versus those that more commonly experience false springs. Of this latter group, specific strategies or traits may then help define which species will see the greatest changes in false spring events with climate change. For example, species that currently avoid false springs through high chilling requirements may see the effectiveness of this strategy erode with warming winters (?). Alternatively, for species that tolerate false spring through a rapid budburst to leafout phase, climate change may alter the rate of this phase and thus make some species more or less vulnerable.

### **Integrating phenological cues to predict vegetative risk**

Understanding what determines the timing of budburst and the length of time between budburst and leafout is essential for predicting the level of damage from a false spring event. The timing between these phenophases (budburst to leafout), which we refer to as the duration of vegetative risk (Figure 3), is a critical area of future research. Currently research shows there is significant variation across species in their durations of vegetative risk, but basic information, such as whether early-budburst species and/or those with fewer morphological traits to avoid freeze damage have shorter durations of vegetative risk compared to other species, is largely unknown, but important for improved forecasting. With spring advancing, species that have shorter durations of vegetative risk would avoid more false springs compared to those that have much longer durations of vegetative risk, especially among species that burst bud early. This hypothesis, however, assumes the duration of vegetative risk will be constant with climate change, which seems unlikely as both phenophases are shaped by environmental cues. The duration of vegetative risk is therefore best thought of as a species-level trait with potentially high variation determined by environmental conditions. Understanding the various physiological and phenological mechanisms that determine budburst and leafout across species will be important for improved metrics of false spring, especially for species- and/or

site-specific studies.



**Figure 3:** Differences in spring phenology and false spring risk across two species: *Ilex mucronata* (L.) and *Betula alleghaniensis* (Marsh.). We mapped a hypothetical false spring event based on historical weather data and long-term observational phenological data collected at Harvard Forest (O’Keefe, 2014). In this scenario, *Ilex mucronata*, which bursts bud early and generally has a short period between budburst (squares) and leafout (triangles), would be exposed to a false spring event during its duration of vegetative risk (i.e., from budburst to leafout), whereas *Betula alleghaniensis* would avoid it entirely (even though it has a longer duration of vegetative risk), due to later budburst.

Decades of research on phenology provide a starting point to understand how the environment controls the duration of vegetative risk across species. As reviewed above, the three major cues that control budburst (e.g., low winter temperatures, warm spring temperatures, and increasing photoperiods, ?) play a dominant role. Comparatively fewer studies have examined all three cues for leafout, but work to date suggests both forcing and photoperiod play major roles (??). The most useful research though would examine both budburst and leafout at once. Instead, most phenological studies currently focus on one phenophase (i.e., budburst or leafout) making it difficult to test how the three phenological cues, and their

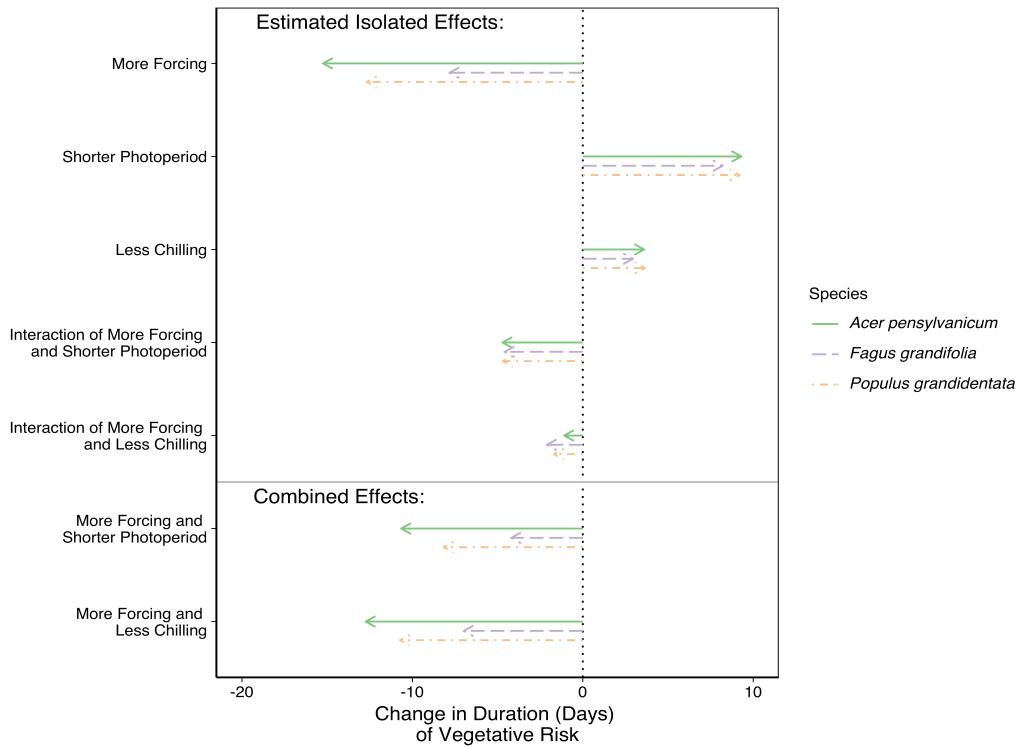
interactions, affect the duration of vegetative risk.

With data in hand, phenological cues can provide a major starting point for predicting how climate change will alter the duration of vegetative risk. Robust predictions will require more information, especially the emissions scenario realized over coming decades (?), but some outcomes with warming are more expected than others. For example, higher temperatures are generally expected to increase the total forcing and decrease the total chilling over the course of the fall to spring in many locations, as well as to trigger budburst at times of the year when daylength is shorter. Using data from a recent study that manipulated all three cues and measured budburst and leafout (?) shows that any one of these effects alone can have a large impact on the duration of vegetative risk (Figure 4): more forcing shortens it substantially (-15 to -8 days), while shorter photoperiods and less chilling increase it to a lesser extent (+3 to 9 days). Together, however, the expected shifts generally shorten the duration of vegetative risk by 4-13 days, both due to the large effect of forcing and the combined effects of multiple cues. How shortened the risk period is, however, varies strongly by species and highlights how climate change may speed some species through this high risk period, but not others. Additionally, as our results are for a small set of species we expect other species may have more diverse responses, as has already been seen in shifts in phenology with warming (???).

These findings highlight the need for further studies on the interplay among chilling, forcing, and photoperiod cues and the duration of vegetative risk across species. This is especially true for species occupying ecological niches more susceptible to false spring events; even if warming causes a shortened duration of vegetative risk for such species, the related earlier budburst dates could still lead to greater risk of false spring exposure.

Studies aiming to predict species shifts across populations (e.g., across a species' range) will also need much more information on how a single species' budburst and leafout timing vary across space. Research to date has studied only a handful of species and yielded no patterns that can be easily extrapolated to other species or functional groups. Some studies have investigated how phenological cues for budburst vary across space, including variation

across populations, by using latitudinal gradients (????), which indicates that more southern populations tend to rely on photoperiod more than northern populations. Other studies have examined distance from the coast (see ???), and some have found that it is a stronger indicator of budburst timing than latitude (?), with populations further inland initiating budburst first, whereas those closer to the coast burst bud later in the season. Changes in chilling requirements for budburst have been repeatedly documented to vary with distance from the coast, and appear predictable based on local climate variation (??).



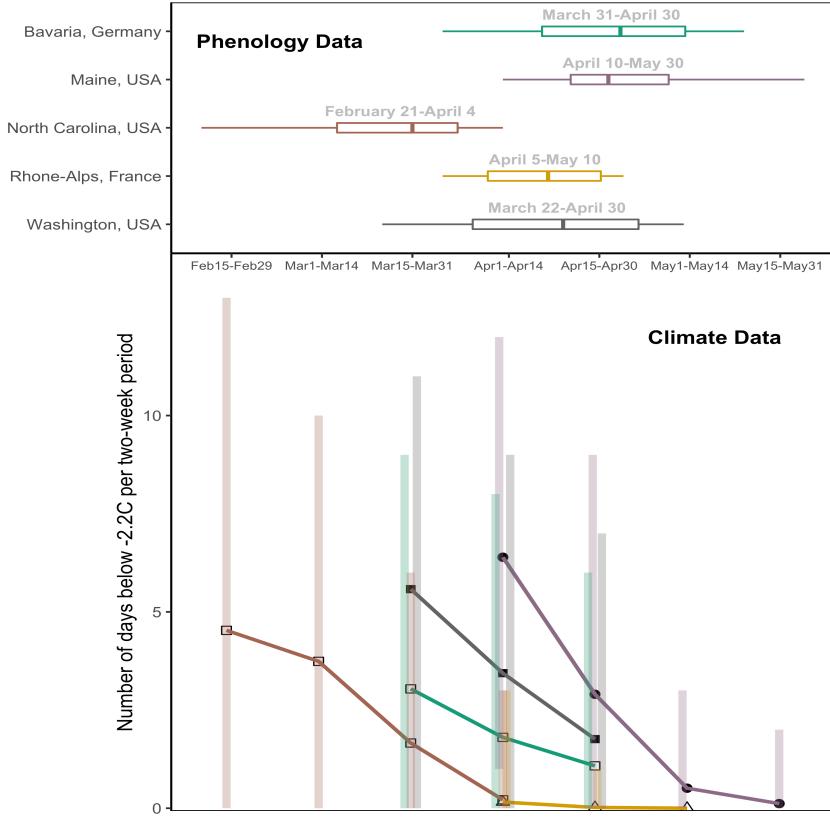
**Figure 4:** Effects of phenological cues on the duration of vegetative risk across three species: *Acer pensylvanicum*, *Fagus grandifolia*, and *Populus grandidentata* (see the Supplement for further details). ‘More Forcing’ is a  $5^{\circ}\text{C}$  increase in spring warming temperatures, ‘Shorter Photoperiod’ is a 4-hour decrease in photoperiod and ‘Less Chilling’ is a 30-day decrease in over-winter chilling. Along with the estimated isolated effects, we show the combined predicted shifts in phenological cues with potential climate change (i.e., more forcing with shorter photoperiod and more forcing with less chilling) and the subsequent shifts in duration of vegetative risk across species. To calculate the combined effects, we added the estimated isolated effects of each cue alone with the interaction effects for the relevant cues for each species.

## Integrating predictable regional differences in false spring risk

Understanding the environmental cues that determine the timing and duration of vegetative risk would provide a major step forward in improving metrics of false spring, but then must be combined with a nuanced appreciation of climate. Research to date (???) highlights the interplay of species cues with a specific location's climate, especially its extremes (??). Climate regime extremes (e.g., seasonal trends, annual minima and annual maxima) vary across regions and are expected to shift dynamically in the future: as climatic regimes are altered by climate change, false spring risk could vary in intensity across regions and time (i.e., regions currently at high risk of false spring damage could become low-risk regions in the future and vice versa). To highlight this, we analyzed five archetypal regions across North America and Europe. Through the use of both phenology (????) and climate data (from the NOAA Climate Data Online tool ?) we determined the number of false springs (i.e., temperatures at -2.2°C or below) for each region. Here, we used the FSI equation, which can help understand the interplay of varying climate regimes and phenology at a cross-regional scale; we tallied the number of years when FSI was positive. We found that some regions experienced harsher winters and greater temperature variability throughout the year (Figure 5, e.g., Maine, USA), and these more variable regions often have a much higher risk of false spring than others (Figure 5, e.g., Lyon, France). Here FSI was a valuable resource to elucidate the regional differences in false spring risk, but for useful projections these estimates should be followed up with more refined data (see *The future of false spring research* below).

Understanding and integrating spatiotemporal effects and regional differences when investigating false spring risk—especially for studies at regional or larger spatial scales—would improve predictions as climate change progresses. As we have discussed above, such differences depend both on the local climate, the local species and the cues for each species at that location. Both single- and multi-species studies will need to integrate these multiple layers of variation, as different species, within the same location can exhibit different sensitivities to the three cues (??), and as a single species may have varying cues across space. Based on cues alone then, different regions may have different durations of vegetative risk for the

same species (???), and accurate predictions will need to integrate cue and climatic variation across space.



**Figure 5:** False spring risk can vary dramatically across regions. Here we show the period when plants are most at risk to tissue loss – between budburst and leafout (upper, lines represent the range with the thicker line representing the interquartile range) and the variation in the number of freeze days ( $-2.2^{\circ}C$ ) (Schwartz, 1993) that occurred on average over the past 50 years for five different sites (lower, bars represent the range, points represent the mean). Data come from USA-NPN SI-x tool (1981-2016), NDVI and remote-sensing, and observational studies (1950-2016) for phenology (Schaber & Badeck, 2005; Soudani et al., 2012; USA-NPN, 2016; White et al., 2009) and NOAA Climate Data Online tool for climate (from 1950-2016). See the Supplement for further details on methods.

## The future of false spring research

With climate change, more researchers across diverse fields and perspectives are studying false springs. Simplified metrics, such as the FSI, have helped to understand how climate change may alter false springs now and in the future. They have helped estimate potential

damage and, when combined with methods that can assess tissue loss (e.g., PhenoCam images can capture initial greenup, defoliation due to frost or herbivory, then refoliation, ?), have documented the prevalence of changes to date. Related work has shown that duration of vegetative risk can be extended if a freezing event occurs during the phenophases between budburst and full leafout (?), which could result in exposure to multiple frost events in one season. Altogether they have provided an important way to meld phenology and climate data to understand impacts on plant growth and advance the field (????). As research in this area grows, however, the use of simple metrics to estimate when and where plants experience damage may slow progress in many fields.

As we have outlined above, current false spring metrics depend on the phenological data used, and thus often ignore important variation across functional groups, species, populations, and life-stages—variation that is critical for many types of studies. Many studies in particular use gridded spring-onset data (e.g., SI-x). Studies aiming to forecast false spring risk across a species' range using SI-x data may do well for species similar to lilac (*Syringa vulgaris*), such as other closely-related shrub species distributed across or near lilac's native southwestern European range. But we expect predictions would be poor for less-similar species. No matter the species, current metrics ignore variation in cues underlying the duration of vegetative risk across space (and, similarly, climate) and assume a single threshold temperature and 7-day window. These deficiencies, however, highlight the simple ways that metrics such as FSI can be adapted for improved predictions. For example, researchers interested in false spring risk across a species range can gather data on freezing tolerance, the environmental cues that drive the variation in the duration of vegetative risk and whether those cues vary across populations, then adjust the FSI or similar metrics. Indeed, given the growing use of the SI-x for false spring estimates research into the temperature thresholds and cues for budburst and leafout timing of *Syringa vulgaris* could refine FSI estimates using SI-x.

Related to range studies, studies of plant life history will benefit from more-specialized metrics of false spring. Estimates of fitness consequences of false springs at the individual-population- or species-levels must integrate over important population and life-stage variation.

In such cases, careful field observational and lab experimental data will be key. Through such data, researchers can capture the variations in temperature thresholds, species- and lifestage-specific tolerance and avoidance strategies and climatic effects, and more accurately measure the level of damage.

Though time-consuming, we suggest research to discover species  $x$  life-stage  $x$  phenophase specific freezing tolerances and related cues determining the duration of vegetative risk will make major advances in fundamental and applied science. Such studies can help determine at which life stages and phenophases false springs have important fitness consequences, and whether tissue damage from frost for some species  $x$  life stages actually scales up to minimal fitness effects. As more data are gathered, researchers can test whether there are predictable patterns across functional groups, clades, life history strategies, or related morphological traits. Further, such work would form the basis to predict how future plant communities may be reshaped by changes in false spring events with climate change. False spring events could have large-scale consequences on forest recruitment, and potentially impact juvenile growth and forest diversity, but predicting this is another research area that requires far more and improved species-specific data.

We suggest most studies at the individual to community levels need far more complex metrics of false spring to make major progress, however, simple metrics of false spring may be appropriate for a suite of studies at ecosystem-level scales. Single-metric approaches, such as the FSI, are better than not including spring frost risk in relevant studies. Thus, these metrics could help improve many ecosystem models, including land surface models (????). In such models, SI- $x$  combined with FSI could provide researchers with predicted shifts in frequency of false springs under emission scenarios. Some models, such as the Ecosystem Demography (ED) and the BIOME-BGC models, already integrate phenology data by functional group (????), by adding last freeze date information, FSI could then be evaluated to predict false spring occurrence with predicted shifts in climate. By including even a simple proxy for false spring risk, models, including ED and BIOME-BGC, could better inform predicted range shifts. As such models often form a piece of global climate models (?), incorporating false

spring metrics could refine estimates of future carbon budgets and related shifts in climate. As more data help to refine our understanding of false spring damage for different functional groups, species and populations, these new insights can in turn help improve false spring metrics used for ecosystem models. Eventually earth system models could include feedbacks between how climate shifts alter false spring events, which may reshape forest demography and, in turn, alter the climate itself.

## Acknowledgments

We thank D. Buonaiuto, W. Daly, A. Ettinger, I. Morales-Castilla and three reviewers for comments and insights that improved the manuscript.