

# The problem and promise of modeling plant leafout under warmer winters

E.M. Wolkovich<sup>1</sup>, Justin Ngo<sup>1</sup>, Victor van der Meersch<sup>1</sup>, Jonathan Auerbach<sup>2</sup>

December 12, 2025

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

<sup>2</sup> Department of Statistics, George Mason University, 4511 Patriot Circle, Fairfax, VA 22030, United States

## Abstract

How warmer winters affect plant leafout has major implications for carbon storage and ecosystem ~~change~~ **stability**. Recent evidence **shows** the advance of spring leafout **has slowed as** anthropogenic climate change **has increased, and attempts to explain the slowed advance has reignited** a decades-old debate over winter ‘chilling’—the ~~concept of a critical but unobserved process where~~ **theory that** woody temperate plants require sufficient winter cool temperatures to budburst each spring. **The problem with chilling is that it is largely used as a placeholder theory: that does not lead to unique predictions and thus has stalled it does not correspond with an observable event, is not falsifiable, and thus is of limited use for advancing scientific progress understanding.** In this paper, we review how current ~~work~~ **research is built on the theory of chilling, which which routinely models ‘chilling’ with highly** **due in part to** limited data, has likely led to the proliferation of **untestable models—and thus forecasts—that yield that are consistent with a wide range of forecasts and provide little additional insight into actual observable plant behavior.** We argue that **the idea of chilling as a scientific theory must be reevaluated like past theoretical constructs of limited scientific value,** such as ‘dormin’ or the ‘ether,’ **until which time researchers will be slow to incorporate** new insights from molecular biology ~~that could better define the mechanism of chilling will likely lead to only limited progress.~~ **By revisiting the limits of chilling as a theory, we can both** **This can be done by** 1- building better models by looking for the necessary insight from molecular biology to fill in the details; and 2- designing better experiments to test those theories.

Plant leafout in the spring has shifted weeks earlier in many regions due to warming from anthropogenic climate change with consequences for a ~~suite~~ **wide range** of ecosystem services (??). Which underlying processes drive this trend, however, is debated, as recent research suggests winter warming may slow or stall further advances (??). Such reports focus on a two-step model of leafout where plants first require cool winter temperatures, often called ‘chilling,’ before they can accumulate enough warm temperatures—‘forcing’—to leafout each year. This model of chilling is only one of many models proposed since the concept was first introduced (??). But, while forcing ends

in an observable event such as leafout, the day plants achieve their chilling requirements coincides with no known measurable event.

Current models of chilling are consistent with a diverse suite of possible future leafout. Chilling in the same location for the same plant under different models can easily forecast significantly increased or greatly reduced chilling and thus greatly advanced or slowed leafout (??). This variability often occurs within any one model of chilling, given different parameter values (Figure ??). Such extreme variability of spring phenology predicted from current models highlights fundamental gaps in our understanding. While decades of experimental evidence suggest that winter temperatures impact leafout (??) new research has highlighted major flaws in these models as currently applied, including multiple papers now suggesting estimates of ‘chilling’ could easily be artifacts of poor models or correlated observational climate data (??).

Now appears an opportune time to address the problems in our concept of chilling in models of leafout. Because accurate forecasts of spring phenology are critical for carbon sequestration, many crops and a number of other services important to humans, there is widespread interest in improving chilling models (??). At the same time, new results from molecular and cellular studies of dormancy are providing new insights into when and how chilling works (??), which could rule in—or out—some of the myriad current models. Here, we review the concept of chilling, its origins and potential problems, as well as new opportunities for major advances and how shifting current practices could accelerate progress. We argue that revisiting the chilling mechanism could help build better models and design better experiments.

## What is chilling?

How plants in temperature-limited systems avoid leafout during warm spells in the winter has long been debated by plant biologists (e.g., ??). Most work to date has focused on the idea that plants enter some form of dormancy in the fall, which is then released before warm temperatures in the spring begin. This idea hypothesizes that the slow accumulation of cool temperatures—or chilling—over the winter extends through periods of short warm spells in the winter and thus prevents leafout before spring. The term ‘chilling’ is now used across numerous fields in plant biology to refer to a process where dormant buds exposed to cool temperatures accelerate a phenological event that later occurs after warm temperatures.

Much of our fundamental understanding of chilling comes from studies on temperate woody fruit crops where chilling can be critical to yield. Peach trees planted into warmer climates well outside their range often have extremely low fruitset because most flower buds do not burst (???). Initial studies of this phenomenon with related experiments—where cut ends of dormant branches (cuttings) exposed to cooler temperatures in chambers burst more fully and more quickly—underlies most of the models of chilling used today for crops and wild tree species (??).

Focused on how chilling can accelerate events, researchers have calculated ‘chilling’ required for leafout of forest trees from cutting experiments similar to those used for peaches (reviewed in ?), ground observations of budburst (?), and satellite measures of greenup (?). These estimates rely on phenological models that have become critical across a suite of fields, from climatology, where

modeling how vegetation on the land surface responds to anthropogenic climate change affects carbon storage, to crop biology, where estimated chill units guide growers in which specific cultivar to plant, and has led to the cross-disciplinary field of phenology (??).

Alongside these more macro-scale studies of chilling, molecular approaches have also examined chilling. Many studies have focused on vernalization—cool temperatures required for flowering (?)—in *Arabidopsis thaliana*, with studies in woody species, especially *Populus* examining chilling before budbreak (??). These studies generally use controlled temperatures to vary the hypothesized amount of chilling then examine molecular and cellular responses (e.g., ???).

Today these studies have led to over 30 basic models where accumulated chilling releases plants from dormancy and hundreds more when considering different species and cultivars (??). Though early debates considered whether plants were truly dormant or only growing slow (‘dormancy’ or ‘rest’ versus ‘quiescence’; ?), today most research assumes a model with two phases of dormancy: endodormancy—where chilling occurs—and a period after but before the observed event, called ecodormancy (Figure ??). In most models, chilling can only be accumulated under certain temperatures—traditionally above zero but below 10°C—with certain temperatures being optimal for the most rapid accumulation of ‘chill units,’ where some unknown sum of chill units breaks endo-dormancy. Which temperatures are most effective at providing chilling is a common question addressed in experiments, with different experiments providing different answers (??). Within species, populations may require different sums of chill units, with populations in more mild climates—where warm interruptions in the winter are more common—requiring more chill units than those in areas with cold winters, where temperatures rarely rise above zero before spring (??). Further complexity comes from the hypothesized diversity of these temperature thresholds and sums across species and populations. Most assume different species require different sums of chill units, and may have different lower, upper and optimal chill temperatures.

## The problem with chilling

Chilling is a latent, unobserved process. Typically, ‘chilling’ describes the physiological phase (endodormancy) in which a plant experiences environmental temperatures that induce progress towards the next physiological phase (eco-dormancy) that ends in budburst (Figure ??). The problem is that the transition from endo- to eco-dormancy corresponds with no clearly measurable phenomenon, and thus the properties of each period cannot be determined from the observed leafout, even under experimental conditions in which the temperature can be manipulated. The parameters of a model governing this system are said to be underdetermined or unidentified.

For example, consider a simple model in which one would need to estimate the minimum and maximum temperatures that allow chilling to accumulate (two parameters) and the total sum of those temperature units needed to trigger a shift into the next physiological phase (often called, ‘endo-dormancy break,’ for one additional parameter). Models then need to estimate when plants start and stop accumulating (two more parameters). An experiment that raises temperatures and observes an earlier leafout could be explained by more rapid chilling accumulation leading to earlier endo-dormancy break, more rapid ecodormancy break, or an almost limitless mix of the two (Figure ??).

To address this, models often assign the start date of the endo-dormancy as known (e.g., starting 1 September in the northern hemisphere) and rely on assumptions to set the end date (Figure ??). A common and widely-used assumption, developed by early work on peaches, is that high and rapid budburst (leaf or flower) is evidence that chilling has been met (i.e. endo-dormancy has ended ?), but is rarely evaluated for other taxa or populations. This approach of assigning some unknown parameters as known has the benefit of avoiding adding more unknown model parameters, but has led researchers to be overly confident in a model of chilling where more is actually unobserved and unknown than acknowledged.

Hidden assumptions and numerous parameters can easily drive diverging models. Even if we assume that high and rapid percent budburst signals sufficient chilling, most models today include parameters that cannot be uniquely identified with current data, and explains why one model can often predict a huge variety of leafout possibilities (Figure ??). Researchers tacit approach to these major issues likely has contributed to the expansion of models over the last few decades without any clear advances. While various models have added complexity via the shape of the optimal temperature range for chilling, allowing accumulated chilling to be reduced, shifting the start date of chilling, and/or allowing chilling and forcing to act at once (????), none of these have swept through the field. These new models of chilling have all added parameters, but none of the parameters added to chilling models in 40 years that have been successful enough to be added to all models of chilling.

As experiments and models have suggested a more complicated model of chilling, data have remained relatively uninformative to try to estimate all the complexities the models include. Further, recent models have often relied on even less data (?). Many current methods use only observational data of the timing of leafout (or flowering) to attempt to estimate a model of chilling for different species or locations and project it forward to understand effects of anthropogenic climate change (???). Perhaps not surprisingly then, which model is deemed best varies strongly by method and approach (???), with no clear pattern.

## New molecular insights could reshape the field and its models

JA: In this section, it sounds to me like you're saying how can molecular biology refine the theory of chilling such that science can progress. Logically, that seems to me to fall in the next section, "Overcoming the chilling problem." In this section, I was expecting that you might bridge to this idea by point out that if the goal is simply to predict when the peach trees will bloom given observed chilling and forcing, the incorrectly specified models are probably fine. If the goal is biological understanding and then extrapolation/anticipation of how ecosystems will be affected, the parameterization matters. Molecular biology (and machine learning/AI for that matter) may provide new ways to test these theories, but they also add more parameters/levers and do not automatically solve the problem unless chilling becomes observable/testable. i.e. we have a chicken and egg problem: if we observed chill, we could identify the genes that code for it, and if we knew the genes that code for chill, we could measure/observe chill. Progress requires that we move beyond chilling to something more scientific, e.g. your callose example, which again seems like "overcoming."

New molecular research provides hope that the transition from endo- to eco-dormancy is measurable. Molecular insights have long contributed to crop and forest tree models of chilling (?). Decades of work on vernalization have outlined the pathways—and genes—that lead to flowering only after winter’s cool temperatures in biennial (herbaceous) populations of *Arabidopsis thaliana* (Figure ??, ??). Research has linked some of these pathways to similar ones in woody species, and have also highlighted the sugar callose (1,3- $\beta$ -D-glucan) as potentially pivotal for chilling (??). Multiple studies across multiple species have now shown that (1) lower temperatures appear to degrade callose and (2) the release/loss of callose appears to re-start cell-to-cell communication before budburst (?). Taken together, these results suggest the loss of callose—generally degraded through 1,3- $\beta$ -glucanases (a group of enzymes)—may be an indicator of endo-dormancy release, though other factors, such as phytohormones, also often change at the same time (??), and may provide a similar observable signal of endo-dormancy release (??).

If callose is a major controller of endo-dormancy and its release, chilling models could be limited to those that match the idea of callose degradation—meaning models that include a temperature range over which the enzyme is active (Figure ??). In contrast, models using simple temperature thresholds (e.g., all hours below -5°C equally allow chilling) would appear less biologically accurate, as enzymes generally do not work over such a wide range of temperatures.

Other new molecular insights similarly suggest that such simplified temperature metrics used in many chilling models may not map to molecular realities. For example, new work on how slow growth itself may act a ‘long-term thermosensor’ (?) adds to an increasing number of molecular studies that suggest plants integrate long-term thermo-sensing in the winter alongside responses to short-term temperatures (??). The best models of leafout may thus need to integrate across multiple timescales. While this could easily complicate models already burdened with complexity, new insights from molecular biology could rule out models by focusing on new experiments and modeling approaches that target the major problems facing models of chilling—if the field is more open about those issues and open to new approaches.

## Overcoming the chilling problem

Richer, more informative data from molecular biology studies and other approaches (??) hold the promise to shift chilling from an unobserved complex process to something we understand and can robustly forecast. Taking full advantage of this opportunity, however, would benefit from re-examining the concept our chilling across different disciplines to combine observational and experimental evidence, while providing the opportunity to question the future utility of ‘chilling.’

### Model experimental and observational data together

Research on chilling could accelerate by working across what today are three fairly separated silos: model building for crops and other forecasting needs (usually called process-based models), experiments at the branch or whole-plant level, and molecular research. Currently, crop biologists, phenological process-based modelers of forest trees, molecular biologists and hardiness modelers all develop unique and rarely compared models of dormancy and budburst (but see ?), highlighting a major problem, but also a major opportunity. Synthesizing and benchmarking models—and their

underlying biological understanding of chilling—across the many research fields developing chilling models today would help identify models that are equivalent and/or perform especially poorly and could begin to discard some models.

Comparing models more directly would hopefully drive new approaches that estimate chilling by fitting experimental and observational data together in one model. This is rarely (if ever) done, in part because of how differently they may be observed, including the challenging diversity of environmental conditions across these two data types (for example, many experiments apply cold temperatures in the dark and include extreme temperature differences, while in observational data photoperiod shifts each day and temperatures are more similar) but also because of separate modeling approaches. Researchers rarely if ever fit process-based models to new empirical data; instead they use so-called ‘statistical models’ that often follow canonical treatment designs (e.g., ANOVA). Statistical models are usually far simpler, and make a suite of unstated assumptions that contradict the current understanding of chilling (see Figure ??). Many of the original studies that led to the concept of chilling, however, bridged across observational and experimental studies more often and leveraged datasets that created greater extremes in observational data—by focusing on crops planted well outside their natural range (e.g., peaches in Florida and Israel, ??). **JA: I think the issue is that the mapping from statistical models to biological processes are not made explicit so it’s not exactly clear what’s being tested and thus how to refine biological theory accordingly.**

Experiments bridging across methods may have the greatest opportunity to provide data that would truly challenge current models of chilling. Testing models across large environmental gradients in the field is one of the best ways to find out where models work—and where they fail. For example, molecular biologists tested vernalization models by through comparing predicted to observed flowering times in a common garden study across Europe (?), and supported the temperature-dependent growth model by testing its predictions of what happens when growth is altered but temperature is held constant (?). Similar examples for challenging other models of chilling date back over 40 years to when many of the models used today were developed (???), but could take place now. Models of chilling can make predictions under lower field chilling then test them using individuals planted beyond the range (either planting those individuals now or identifying such cases in forestry provenance trials or similar). Process-based modelers could also challenge their models more through more dramatic variation in biology, via experiments that include genetic mutants or similar variants.

### **Is a theory of ‘chilling’ necessary for scientific progress?**

Results from new molecular experiments designed to identify novel ways to directly observe and measure chilling could upend the current concept of ‘chilling.’ As increasing evidence suggests the concept aligns with multiple shifts at the cellular and molecular level (Figure ??), researchers across disciplines will need to consider how to continue using the term ‘chilling.’ One approach is to try to align previous and current definitions through new studies, for example, by testing for callose loss (?) alongside previous-used markers of dormancy shifts—including the often-used bioassay of high and rapid budburst at higher temperatures (indicating endo-dormancy release), and additional methods, such as weighing flower buds (?) or tracing water reactivation into cells (???). Another approach is to work towards new terms that more clearly align with new measurements and insights.

Because chilling is an unobserved process, it may link eventually to one measurable compound or regulator (similar to florigen or auxin) or to a more complicated biology, making the term less useful (?). Current results suggests the latter, and already the term ‘chilling’ is used to mean different things. For example, while we focus here on the term as an accumulation process during an unobserved physiological phase (endo-dormancy), many experiments currently use the term ‘chilling’ to refer to a treatment where researchers do not know the physiological phase (for example, cuttings or buds from woody plants are often chilled at 5°C for 6 weeks in the dark in a ‘chilling’ treatment, then transferred to warming ‘forcing’ conditions, ??).

Diverging definitions suggest that the best path for ‘chilling’ may be to follow that of other terms modern plant biologists may not remember, such as dormin or anthesin (hypothesized compounds to induce dormancy and floral anthesis, respectively, ??), and let the term fade into obsolescence. Models for forecasting and new experiments at the branch and whole plant level would then focus on modeling and measuring genes, regulators, hormones and other compounds that emerged from a history of research on ‘chilling,’ but use other terms and concepts. Chill unit estimates for widely planted forest trees and specific varieties of crops would likely hold on longer. They too, however, may shift to new estimates or units based on better measures, especially as increasing anthropogenic climate change makes getting the mechanistic biology of plant development right for human adaptation more important.