The problem and promise of
modeling plant leafout under warmer winters

E.M. Wolkovich[1], Justin Ngo[1], Jonathan Auerbach; plus others possible

September 16, 2025

[1] Forest & Conservation Sciences, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada

Abstract

Recent evidence of a slow-down in the advance of spring with increasing anthropogenic climate change has reinvigorated a debate over how warmer winters affect plant leafout. Plant biologists have long hypothesized that many woody temperate plants need sufficient winter cool temperatures—or 'chilling'—to budburst each spring; thus warmer winters could slow this process and slow advancing spring phenology. But models of this process are many and make diverse predictions for leafout, with different models given the same data often predicting everything from advanced to slowed leafout. Such variability highlights how little we understand the critical process of 'chilling,' which is generally never observed but often assumed. Here we show how current work, which routinely models 'chilling' with highly limited data has likely led to the proliferation of models—and thus forecasts—and suggest a path forward. We argue that integrating new insights and approaches from molecular biology could push the field to build new models that make falsifiable predictions. This approach could in turn greatly reduce the number of competing models of chilling that we have today——advancing our fundamental understanding of plant dormancy—and lead to better forecasts of leafout and, in turn, climate change.

Plant leafout in the spring has shifted weeks earlier in many regions due to warming from anthropogenic climate change with consequences for a suite of ecosystem services, including carbon storage (Keenan *et al.*, 2014; Pörtner *et al.*, 2022). What underlying processes drive this trend, however, has come under increasing debate, as recent research suggests winter warming may slow or stall this advance (Fu *et al.*, 2015; Piao *et al.*, 2017). 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. But this model of leafout—especially its chilling component—is only one of many models proposed since the concept was first introduced (at least 30 models are used today, see Basler, 2016; Hufkens *et al.*, 2018).

Current models of chilling predict a full suite of possible future leafout. Chilling in the same location under different models can easily forecast significantly increased or greatly reduced chilling and thus greatly advanced or slowed leafout (Guy, 2014; Chuine *et al.*, 2016). Such extreme variability

of spring phenology predicted from current models suggests fundamental and major gaps in our understanding. Indeed, new research suggests 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 (Wolkovich *et al.*, 2021; Gao *et al.*, 2024).

Now may be an opportune time to address the problems in models of leafout. Because accurate forecasts of spring phenology are critical for carbon storage, many crops and a number of other services important to humans, there is high interest in improving chilling models (Luedeling *et al.*, 2015; Chuine *et al.*, 2016). At the same time, new results from molecular and cellular studies of dormancy are providing new insights into when and how chilling works (Pan *et al.*, 2023; Zhu *et al.*, 2021), 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. In particular, we discuss how new insights from molecular and cellular studies combined with modern approaches to building biological models could revolutionize our understanding of chilling.

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., Lamb, 1948; Weinberger, 1950). Most work to date has focused on the idea that plants enter some form of dormancy in the fall, which is then released before spring warm temperatures 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.

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 (e.g., in Florida or Israel) often have extremely low fruitset because most flower buds do not burst (Weinberger, 1950; Overcash & Campbell, 1955; Erez *et al.*, 1971). Initial studies of this phenomenon with related experiments—where cut ends of dormant branches 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 (Weinberger, 1950; Ettinger *et al.*, 2020).

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. Focused on how chilling can accelerate events, researchers have calculated 'chilling' required for leafout of forest trees from satellite measures of greenup (Kaduk & Los, 2011), ground observations (Luedeling *et al.*, 2009b), and from similar cutting experiments to those used for peaches (reviewed in Ettinger *et al.*, 2020).

Alongside these more macro-scale studies of chilling, molecular approaches have also examined chilling, with a wealth of studies on vernalization in *Arabidopsis thaliana*—cool temperatures required for flowering (Kim *et al.*, 2009). These studies generally use controlled temperatures to vary the hypothesized amount of chilling—similar to cutting studies of forest trees and woody crops—then examine molecular and cellular responses (e.g., Pan *et al.*, 2021; Azeez *et al.*, 2021; Cai *et al.*, 2024).

Today these studies have led to over 30 basic models where accumulated chilling releases plants from dormancy and hundreds or more models considering variation on these models for different species and cultivars (Basler, 2016; Hufkens *et al.*, 2018). Though early debates considered whether plants were truly dormant or only growing slow ('dormancy' or 'rest' versus 'quiescent; Considine & Considine, 2016), today most research assumes a model with two phases of dormancy (Fig. 1).

In most models, chilling can only be accumulated under certain temperatures—often 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. This mirrors growing degree days in many ways, where a lower temperature (e.g., 5 or 10°C base temperature) is too cold for forcing units to accumulate and plants need some total sum of such units to leafout or flower, but has added complexity given the importance of both a lower and upper temperature threshold for 'chill units,' and 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. 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 that rarely warm much above zero before spring (Campbell & Sugano, 1979; Leinonen, 1996).

The problem with chilling

Experiments and models focused on chilling describe an unobserved process (Chuine *et al.*, 2016), with critical differences between the two approaches. In most models the term 'chilling' describes a process that occurs when plants are in a specific physiological phase and experiencing environmental temperatures known to induce progress towards the next physiological phase. In contrast, in most experiments the term 'chilling' often refers to a treatment where researchers do not know the physiological phase (Flynn & Wolkovich, 2018; Ettinger *et al.*, 2020). 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 without information on physiological phases. While this may seem like a small terminology issue it actually belies one of the major problems with our understanding—and thus modeling—of chilling today: measuring and estimating an unobserved process with so many unknowns is extremely challeinging (Ettinger *et al.*, 2020, , see also: Box: Why has progress on modeling stalled for decades?).

Most current models of chilling try to estimate a large number of unknowns—generally as model parameters—from data that yields limited inference. A simple model 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 for three total). Models then need to estimate when plants start and stop accumulating (two more parameters), but this is generally intractable (see Box: Why has progress on modeling stalled for decades?). To address this, models today assign the start date as known (e.g., starting 1 September in the Northern Hemisphere) and rely on assumptions to set the end date. A common assumption, developed by early work on peaches, is that high and rapid leaf or flowerburst is evidence that chilling has been met (Erez &

Lavee, 1971). While this assumption is widely used, it is rarely if ever tested beyond the early work on peaches. This approach of assigning some unknown parameters as known (without any variance) has the benefit of avoiding adding more unknown model parameters, but it also has led researchers to be overly confident in a model for leafout where more is actually unobserved and unknown than acknowledged.

Hidden assumptions and numerous parameters can easily drive diverging models. Even if we assume high and rapid percent leafout or flowerburst signals sufficient chilling most models today include parameters that cannot be uniquely identified with current data (see Box: Why has progress on modeling stalled for decades?). Given experiments and models have suggested many variants on a more complicated model of chilling—for example minimum, maximum and optimal temperatures, or high temperatures that reduce previous accumulation (Fig. 1 Luedeling *et al.*, 2011, 2012; Chuine *et al.*, 2016)—current data are relatively uninformative to try to estimate all the parameters 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 (Luedeling *et al.*, 2011, 2012; Gao *et al.*, 2024). Perhaps not surprisingly then, which model is deemed best varies strongly by method and approach (Caffarra *et al.*, 2011; Basler, 2016; Hufkens *et al.*, 2018), with no clear pattern.

New molecular insights could reshape the field and its models

Molecular insights have long helped crop and forest tree models of chilling (Chuine & Regniere, 2017). 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* (Fig. 2 Wilczek *et al.*, 2009; Kim *et al.*, 2009). Research has linked some of these pathways to similar ones in woody species, and have also highlighted the sugar callose (1,3-β-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 (van der Schoot *et al.*, 2014). Taken together, these results suggest the loss of callose—generally degraded through 1,3-β-glucanases (a group of enzymes)—may be an indicator of endo-dormancy release (though other factors, such as ABA, also often change at the same time, Tylewicz *et al.*, 2018; Pan *et al.*, 2021) and may provide a similar observable signal of endo-dormancy release. (Rinne *et al.*, 2018; André *et al.*, 2022).

If callose is functionally a major controller of endo-dormancy and its release it suggests chilling models could be limited to those that match the idea of glucanase degrading callose—meaning models that include a temperature range over which the enzyme is active. 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' (Zhao *et al.*, 2020) 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 (Antoniou-Kourounioti *et al.*, 2021; Satake *et al.*, 2022). The best models of leafout may thus need to integrate across multiple timescales. This could easily add complexity to models of spring phenology that are already challenged by too much complexity. But we argue that new insights from molecular biology could begin to rule out models by focusing on new experiments and modeling approaches that target the major problems facing models of chilling.

Building a better model of chilling

Redefining chilling through new experiments

One important way to leverage new molecular insights for modeling is through new experiments designed to redefine chilling. If experiments could find a way to observe and measure chilling it could quickly advance modeling progress. This means experiments designed to identify markers of the underlying physiological stage and shifts between stages. Experiments testing for evidence of callose loss using common temperature treatments of past studies (Ettinger *et al.*, 2020) currently seem very promising, but could be complemented by studies with other cellular and molecular indicators (Yu *et al.*, 2024). Testing these methods together 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 (Chuine *et al.*, 2016) or tracing water reactivation into cells (Faust *et al.*, 1991; Kalcsits *et al.*, 2009; Walde *et al.*, 2024)— could help align both new and old methods. It may be that the field's long reliance on rapid and high percent budburst under warm temperatures only aligns with physiological dormancy release for some species and/or conditions (Fouché *et al.*, 2023; Walde *et al.*, 2024).

Because chilling is functionally an unobserved process we argue that comparing methods should be a major priority for the field. This comparison will need to allow for the reality that different methods may measure different processes and, thus, terminology may need to adapt as well. As a first step, research could stop referring to treatments in experiments as 'chilling' or 'forcing,' or other terms that assume an underlying physiological state, and instead focus on the actual treatments (e.g., 'cool temperatures before warm').

Progress through new experiments could come from experiments that bridge across molecular and phenological methods in other aspects of their design, beyond simply what they measure. This includes more temperature levels for controlled studies to disentangle effects of temperature (including identifying which temperatures are optimal) and time, especially in molecular studies where cold treatments often vary only in duration, not temperature (e.g., Pan *et al.*, 2023). More efforts to test lab results in field conditions would especially challenge our current understanding and provide important data for models (see below, 'Model experimental and observational data together'). Molecular biologists have tested models through large-scale planting experiments (Wilczek *et al.*, 2009; Burghardt *et al.*, 2015) and more efforts to do this on both small and large-scales could speed progress (Satake *et al.*, 2022). 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. Similarly, phenology researchers could challenge their models more through more dramatic variation in biology, via experiments that include mutants or similar variants.

Current experimental designs are unlikely to radically challenge models of chilling, but small tweaks may still offer important insights. In particular, experiments considering multiple cool and warm temperature treatments could measure hardiness to improve the model of how hardiness and dormancy interact (Kovaleski, 2022). Cool treatments may also be useful if they tested for the effect of light regimes given the prevalence of cool temperature ('chilling') treatments in the dark to date (Ettinger *et al.*, 2020).

Highlight the unknown

Perhaps the largest problem with current models of chilling is that they cannot uniquely estimate the most important aspects of chilling—that is, current models are non-identifiable (Gelman & Hill, 2009). Even in some of the simplest models, estimates of what temperatures accumulate 'chilling' and how much chilling is required to shift physiological stages usually occupy two divergent options (see Box: Why has progress on modeling stalled for decades?), but researchers often only present one of the two (Chuine *et al.*, 2016). Highlighting this uncertainty instead of hiding it could quickly help compare models and guide experiments to reduce uncertainty. Thus we suggest research always include routine estimates of uncertainty and clearly state any modeling choices that may have limited insights into non-identifiably (e.g., limiting the parameter space the model can search when estimating parameter values). Given that non-identifiability in current models only grows with complexity, focusing more on simple models in the near-term may aid progress. While adding complexity makes all models fit better generally (McElreath, 2016), it also limits finding the major limits of current models and the best path forward if not approached carefully. Current models of chilling fit better when adding complexity to the shape of the optimal temperature range for chilling, making negation possible, shifting the start date of chilling and allowing chilling and forcing to act at once (Luedeling *et al.*, 2009a; Gusewell *et al.*, 2017; Hänninen, 1990; Kramer, 1994), but we are no closer to knowing which of these additions is most biologically relevant.

Making chilling models more biologically relevant could also come from improving statistical models of chilling. Most models of chilling used today for crops, land surface modeling etc. are process-based. These process-based models focus on the mechanistic model and challenge it with empirical data in limited ways. For example, many models infer parameters from other data or studies, then fit only certain parameters of their models to an empirical dataset at hand (Richardson, 1974) (maybe Victor can improve this description and add some better examples?). Process-based models are not routinely used to statistically fit empirical data from ground observations or new experiments. Instead, researchers fit new empirical data with so-called 'statistical models.' Statistical models are usually far simpler, and make a suite of unstated assumptions that contradict the current understanding of chilling (see Box: Why has progress on modeling stalled for decades?), including assumptions of linearity and that 'chilling' and 'forcing' treatments correspond to the physiological states that match these terms (endo- and eco-dormancy, respectively).

Improved statistical models could challenge some of these assumptions by inching closer to the biology. Log transformations better match the non-linear accumulation model of chilling and forcing (Wolkovich *et al.*, 2021). Models could also relax the assumption that 'chilling' and 'forcing' treatments correspond to different physiological states by testing whether the start date may vary with the treatments (see Box: Why has progress on modeling stalled for decades?). Both these

alternatives are simple to implement and thus could be included as alternative models that test alternative hypotheses in future studies. Developing new statistical approaches could also relax additional assumptions to provide insights into how chilling works biologically. For example, instrument variables are a statistical construct designed for when the treatment applied may not exactly match to the intended effect, and could be applied to studies attempting to manipulate chilling and forcing where the underlying state is not known. All these models, however, are likely to be limited in the inference they provide without substantial increases in data—especially in the diversity of climates considered.

Model experimental and observational data together

One way to increase the amount of data used to estimate chilling models is to include both experimental and observational data together in one model. This is rarely (if ever) done, likely because of the challenging diversity of environmental conditions across these two data types. For example, many experiments apply cold temperatures in the dark, while photoperiod shifts each day in observational data, or many experiments achieve extreme differences in cool and warm temperatures (e.g., very minimal cool temperatures, or very warm temperatures) while natural climate data lacks such extremes. Many of the original studies that led to the concept of chilling, however, were developed from datasets that created greater extremes in observational data—focusing on crop species planted well outside their natural range (e.g., peaches in Florida and Israel) and bridged across observational and experimental studies more often (Erez & Lavee, 1971; Richardson, 1974). Improved statistical approaches should push the field of chilling back towards these foundational studies and fit observational and experimental data together, ideally in one model.

Fitting experimental and observational data should drive coherency in what chilling models perform best and reduce support for some models. This would limit the growing number of studies that have compared different process-based models on observational phenology data—where 'natural' field conditions likely often satisfy chilling requirements for wild species (Basler, 2016; Hufkens *et al.*, 2018)—to find the simplest models perform best. By adding more extreme experimental conditions we expect more complex models will perform well. Even if models cannot fit both data types together we suggest that new research should should include comparisons of the model performance on observational data and experimental data—models that cannot fit both data types should be flagged as indicating a potential problem with the model. At the same time, datasets must provide the necessary information to fit chilling models (e.g., for experiments—what was the dormancy induction temperature, what as the thermo and photo-periodicity of each stage of the experiment).

Changing how we evaluate and compare models should make it possible to discard models more often. To date, models are often compared using model comparison statistics and using different datasets in different papers. We argue that these two approaches make it extremely difficult to advance the field, since one model may be preferred on some data for certain model statistics and another model given other data and statistics. Building standard datasets that are always tested alongside expected certain test statistics would alleviate some of this problem, and make it easier to identify why different datasets find different answers. But discarding models fully likely requires moving away from model comparisons statistics and towards designing models that make falsifiable

predictions.

Develop cross-disciplinary models with falsifiable predictions

Building models with falsifiable predictions is standard in molecular biology and highlights how working across these disciplines could especially advance both fields of research into chilling. For example, molecular biologists tested their vernalization model by through comparing predicted to observed flowering times in a common garden study across Europe (Wilczek *et al.*, 2009), and supported the temperature-dependent growth model by testing its predictions of what happens when growth is altered but temperature is held constant (Zhao *et al.*, 2020). Similar examples for challenging other models of chilling date back over 40 years to when many of the models used today were developed (Richardson, 1974; Chuine *et al.*, 2016; Ettinger *et al.*, 2020), 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). Similarly, synthesizing models—and their underlying biological understanding of chilling—across the many research fields developing chilling models today could eventually reduce the total number of models to compare.

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, highlighting a big problem, but also a major opportunity. Uniting these models, first by fully defining their assumptions and conditions (e.g., what species are they designed for, what phenological or dormancy phase do they start at?), then comparing their predictions and pushing them to make different testable predictions seems an obvious path to building a unified model of chilling, with implications for much better models of budburst, and cascading improvements in forecasts for crop yield to forest carbon sequestration.

Box: Why has progress on modeling stalled for decades?

Most major models of chilling were developed over 40 years ago (Richardson, 1974; Chuine *et al.*, 2016) following several decades of new work—especially on peaches and other temperate tree fruit crops—and much debate (Giertrych, 1974). Since then many more models have been proposed (Luedeling *et al.*, 2012; Chuine *et al.*, 2016), but 1970s models still often fit very well (Basler, 2016; Chuine *et al.*, 2016) and are widely used in major studies (e.g., Richardson, 1974; Chuine *et al.*, 2016; Ettinger *et al.*, 2020). While we outline many reasons for this, including a disconnect between experimental, observational data and modeling approaches (e.g., process-based versus statistical), one of the main reasons for this stalled progress could be that most models—from the old to new ones—cannot actually estimate the parameters in them. That is, the models are fundamentally non-identifiable.

Taking a simple example of a chilling model with three parameters—the minimum temperature for chilling, the maximum temperature, and the accumulation needed—shows that there are multiple solutions. Considering just two of these possible solutions highlights how the range trade-offs

with the accumulation: if the temperature range is wide (lower minimum, higher maximum) then the accumulation required will be higher, while if the range is smaller, then the accumulation needed is lower. The full suite of possible solutions is effectively endless (and the trade-off between range and accumulation is not linear, as it depends on the full width of the range, but also its placement relative to 0). Further, this model is not actually one of only three parameters as two additional parameters were set as known (start day of accumulation was set at 1 September, and the endodormancy break date at 30 January) so that the model could even be fit. Without setting these, the optimizing algorithm would not find solutions. This reality is present in every model of chilling and leafout, but it is rarely presented clearly. Being tacit about these issues likely has contributed to the expansion of models over the last few decades without any clear advances that have swept through the field. To date, new models of chilling add parameters (e.g., Luedeling *et al.*, 2009a), but there have been no additions to chilling models in 40 years that have been successful enough to be added to all models of chilling.

Being clear about non-identifiability in models, the full number of parameters—and how they were fit, and how well they fit—could rapidly advance model progress. First, it would help all researchers in the field recognize what is fundamentally unknown and thus hopefully drive research into these areas. Second, it would facilitate progress by highlighting what parameters are most often fixed (effectively assumptions in the model) versus fit to data, and to which type of data. With this, more research could easily compare across models and datasets to give better overviews of what is known, assumed, or most often studied (i.e., what parameter model studies try to fit). Given the extensive list of proposed complexities to chilling models (see Box: What we know matters to chilling, and what might matter), having simpler models (fewer parameters) that are routinely used to compare to more complex models, would likely help the field advance. Currently, very simple models often outcompete more complex models, suggesting how we advance from these simpler models is also a problem (e.g., Basler, 2016).

Being clear about what is unknown in experiments may aid modeling studies to be more upfront about assumptions and limitations. While experiments often assign treatments as 'chilling' and 'forcing,' most researchers know that the actual physiological transition may occur in the cool 'chilling' treatment or much later in the warm 'forcing' conditions, but fail to acknowledge it in their papers or statistical models (Ettinger *et al.*, 2020; Wolkovich *et al.*, 2022). While some experiments rely on a heuristic where budburst that is rapid and includes a high percentage of buds bursting indicates the physiological state associated with forcing (endo-dormancy) has started, the point when this threshold is crossed is rarely clear, and data showing it are often buried in supplements. These practices of hiding uncertainty in results and experimental designs that are effectively non-identifiable for the biological process under study likely contributes to confusion for researchers in the field and drives models with hidden assumptions. A coherent model requires researchers focused on experiments and modelers alike to embrace (and enforce) the current limited level of understanding in the field. Current methods use limited data to estimate a hypothesized accumulation that triggers an unobserved event—currently often described as the release of endo-dormancy—which then leads to another accumulation that leads to leafout or flowering.

BOX: WHAT WE KNOW MATTERS TO CHILLING, AND WHAT MIGHT MATTER

Studies on how cool winter temperatures affect budburst have identified numerous possible factors that appear to moderate 'chilling.' Certainly time and cool temperatures affect budburst timing. Teasing out the effect of time versus temperature has proven difficult, but plants and cuttings (clipped ends of tree branches with one or more bud) exposed for the same amount of time to different constant cool temperatures budburst at different rates and percentages (Ettinger *et al.*, 2020; Wolkovich *et al.*, 2022). Which temperatures are most effective at providing chilling—estimated as those that lead to either (or both) the highest percent or fastest budburst after exposure to warm (often 20°C) temperatures—is a common question addressed in experiments, with different experiments providing different answers CITES.

This has led to debate recently on whether sub-zero temperatures can provide chilling as one recent studied appeared to show (Baumgarten *et al.*, 2021), supporting previous research (Jones *et al.*, 2012; Sønsteby & Heide, 2014), but in contrast to other studies (Lamb, 1948; Cook *et al.*, 2005; Man & Lu, 2010). These studies, however, vary in a number of additional factors including the species they study, the full conditions under which cool temperatures were applied etc.. This debate highlights the complexity of identifying the dominant controllers on chilling when so many factors vary across studies, and so many factors appear to moderate effects. In addition to the temperature range, temperature variability also appears to play a role CITES. The role of sub-zero temperatures and temperature variability also matter strongly to plant hardiness—what level of cold temperatures plants can withstand without damage, which is induced alongside dormancy in the fall for most temperature species. Recent work by Kovaleski (2022) has reignited an old debate about whether hardiness influences dormancy CITES—or vice versa.

These debates—about the importance of sub-zero temperatures and plant hardiness to dormancy and chilling—are not new, and also highlight gaps in the approach many biologists have taken to studying chilling in an era where how well we understand it has major implications due to anthropogenic climate change. Some recent papers ask similar questions to those of decades ago, often failing to cite the earlier literature (e.g., Lamb, 1948; Man & Lu, 2010; Baumgarten *et al.*, 2021). While this highlights that such debates may not have been fully settled it can also slow progress. Because of this, we suggest a major need is to build and share databases of all relevant studies—across molecular, crop, and forest biology. Such databases should include information on factors that appear to effect chilling, including dormancy induction temperature, species, provenance location (as different populations may need different chilling Campbell & Sugano, 1979; Leinonen, 1996), thermo- and photo-periodicity of all applied treatments, and any hardiness information. We provide a start to this in the supplement.

1 References

André, D., Zambrano, J.A., Zhang, B., Lee, K.C., Rühl, M., Marcon, A. & Nilsson, O. (2022) Populus svl acts in leaves to modulate the timing of growth cessation and bud set. *Frontiers in Plant Science* **13**, 823019.

Antoniou-Kourounioti, R.L., Zhao, Y., Dean, C. & Howard, M. (2021) Feeling every bit of winter– distributed temperature sensitivity in vernalization. *Frontiers in Plant Science* **12**, 628726.

Azeez, A., Zhao, Y.C., Singh, R.K., Yordanov, Y.S., Dash, M., Miskolczi, P., Stojkovič, K., Strauss, S.H., Bhalerao, R.P. & Busov, V.B. (2021) Early bud-break 1 and early bud-break 3 control resumption of poplar growth after winter dormancy. *Nature communications* **12**, 1123.

Basler, D. (2016) Evaluating phenological models for the prediction of leaf-out dates in six temperate tree species across central europe. *Agricultural and Forest Meteorology* **217**, 10–21.

Baumgarten, F., Zohner, C.M., Gessler, A. & Vitasse, Y. (2021) Chilled to be forced: the best dose to wake up buds from winter dormancy. *New Phytologist* **230**, 1366–1377.

Burghardt, L.T., Metcalf, C.J.E., Wilczek, A.M., Schmitt, J. & Donohue, K. (2015) Modeling the influence of genetic and environmental variation on the expression of plant life cycles across landscapes. *American Naturalist* **185**, 212–227.

Caffarra, A., Donnelly, A. & Chuine, I. (2011) Modelling the timing of *Betula pubescens* budburst. II. Integrating complex effects of photoperiod into process-based models. *Climate Research* **46**, 159–170.

Cai, F., Jin, X., Tian, Y., Huang, Z., Wang, X., Zhang, Y., Sun, Y. & Shao, C. (2024) Molecular regulation of bud dormancy in perennial plants. *Plant Growth Regulation* **102**, 1–11.

Campbell, R.K. & Sugano, A.I. (1979) Genecology of bud-burst phenology in Douglas-fir - response to flushing temperature and chilling. *Botanical Gazette* **140**, 223–231.

Chuine, I., Bonhomme, M., Legave, J.M., de Cortazar-Atauri, I.G., Charrier, G., Lacointe, A. & Ameglio, T. (2016) Can phenological models predict tree phenology accurately in the future? the unrevealed hurdle of endodormancy break. *Global Change Biology* **22**, 3444–3460.

Chuine, I. & Regniere, J. (2017) Process-based models of phenology for plants and animals. *Annual Review of Ecology, Evolution, and Systematics* **48**, 159–182.

Considine, M.J. & Considine, J.A. (2016) On the language and physiology of dormancy and quiescence in plants. *Journal of Experimental Botany* **67**, 3189–3203.

Cook, N.C., Bellen, A., Cronjé, P.J., De Wit, I., Keulemans, W., Van den Putte, A. & Steyn, W. (2005) Freezing temperature treatment induces bud dormancy in 'granny smith'apple shoots. *Scientia horticulturae* **106**, 170–176.

Erez, A. & Lavee, S. (1971) Effect of climatic conditions on dormancy development of peach buds. i. temperature. *Journal of the American Society for Horticultural Science* .

Erez, A., Lavee, S. & Samish, R.M. (1971) Improved methods for breaking rest in the peach and other deciduous fruit species. .

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**, 1137–U119.

Faust, M., Liu, D., Millard, M.M. & Stutte, G. (1991) Bound versus free water in dormant apple buds—a theory for endodormancy. *HortScience* **26**, 887–890.

Flynn, D.F.B. & Wolkovich, E.M. (2018) Temperature and photoperiod drive spring phenology across all species in a temperate forest community. *New Phytologist* **29**, 1135–1492.

Fouché, M., Bonnet, H., Bonnet, D.M. & Wenden, B. (2023) Transport capacity is uncoupled with endodormancy breaking in sweet cherry buds: physiological and molecular insights. *Frontiers in Plant Science* **14**, 1240642.

Fu, Y.S.H., Zhao, H.F., Piao, S.L., Peaucelle, M., Peng, S.S., Zhou, G.Y., Ciais, P., Huang, M.T., Menzel, A., Uelas, J.P., Song, Y., Vitasse, Y., Zeng, Z.Z. & Janssens, I.A. (2015) Declining global warming effects on the phenology of spring leaf unfolding. *Nature* **526**, 104–107.

Gao, X., Richardson, A.D., Friedl, M.A., Moon, M. & Gray, J.M. (2024) Thermal forcing versus chilling? misspecification of temperature controls in spring phenology models. *Global Ecology and Biogeography* **33**, e13932, e13932 GEB-2024-0136.R2.

Gelman, A. & Hill, J. (2009) *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge, New York.

Giertrych, M. (ed.) (1974) *International Symposium on Dormancy in Trees Discussions: 5th to 9th September 1973.* Polish Academy of Sciences.

Gusewell, S., Furrer, R., Gehrig, R. & Pietragalla, B. (2017) Changes in temperature sensitivity of spring phenology with recent climate warming in switzerland are related to shifts of the preseason. *Global Change Biology* **23**, 5189–5202.

Guy, R.D. (2014) The early bud gets to warm. *New Phytologist* **202**, 7–9.

Hänninen, H. (1990) *Modelling bud dormancy release in trees from cool and temperate regions.* 213.

Hufkens, K., Basler, D., Milliman, T., Melaas, E.K. & Richardson, A.D. (2018) An integrated phenology modelling framework in r. *Methods in Ecology and Evolution* **9**, 1276–1285.

Jones, H., Hillis, R., Gordon, S. & Brennan, R. (2012) An approach to the determination of winter chill requirements for different *Ribes* cultivars. *Plant Biology* **15**, 18–27.

Kaduk, J.D. & Los, S.O. (2011) Predicting the time of green up in temperate and boreal biomes. *Climatic Change* **107**, 277–304.

Kalcsits, L., Kendall, E., Silim, S. & Tanino, K. (2009) Magnetic resonance microimaging indicates water diffusion correlates with dormancy induction in cultured hybrid poplar (populus spp.) buds. *Tree Physiology* **29**, 1269–1277.

Keenan, T.F., Gray, J., Friedl, M.A., Toomey, M., Bohrer, G., Hollinger, D.Y., Munger, J.W., O'Keefe, J., Schmid, H.P., Wing, I.S. *et al.* (2014) Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nature Climate Change* **4**, 598–604.

Kim, D.H., Doyle, M.R., Sung, S. & Amasino, R.M. (2009) Vernalization: winter and the timing of flowering in plants. *Annual Review of Cell and Developmental* **25**, 277–299.

Kovaleski, A. (2022) Woody species do not differ in dormancy progression: differences in time to budbreak due to forcing and cold hardiness. *Proceedings of the National Academy of Sciences* **119**, e2112250119.

Kramer, K. (1994) Selelecting a model to predict the onset of growth of *Fagus sylvatica*. *Journal of Applied Ecology* **31**, 172–181.

Lamb, R.C. (1948) Effect of temperatures above and below freezing on the breaking of rest in the Latham raspberry. *Proceedings of the American Society for Horticultural Science*, vol. 51, pp. 313–315, AMER SOC HORTICULTURAL SCIENCE 113 S WEST ST, STE 200, ALEXANDRIA, VA 22314 . . . .

Leinonen, I. (1996) Dependence of dormancy release on temperature in different origins of pinus sylvestris and betula pendula seedlings. *Scandinavian Journal of Forest Research* **11**, 122–128.

Luedeling, E., Blanke, M. & Gebauer, J. (2012) Chilling challenges in a warming world. *II International Symposium on Horticulture in Europe 1099*, pp. 901–907.

Luedeling, E., Blanke, M. & Gebauer, J. (2015) *Chilling Challenges in a Warming World*, vol. 1099 of *Acta Horticulturae*, pp. 901–907.

Luedeling, E., Girvetz, E.H., Semenov, M.A. & Brown, P.H. (2011) Climate change affects winter chill for temperate fruit and nut trees. *Plos One* **6**.

Luedeling, E., Zhang, M.H. & Girvetz, E.H. (2009a) Climatic changes lead to declining winter chill for fruit and nut trees in California during 1950-2099. *Plos One* **4**.

Luedeling, E., Zhang, M.H., McGranahan, G. & Leslie, C. (2009b) Validation of winter chill models using historic records of walnut phenology. *Agricultural and Forest Meteorology* **149**, 1854–1864.

Lundell, R., Hanninen, H., Saarinen, T., Astrom, H. & Zhang, R. (2020) Beyond rest and quiescence (endodormancy and ecodormancy): A novel model for quantifying plant-environment interaction in bud dormancy release. *Plant Cell and Environment* **43**, 40–54.

Man, R. & Lu, P. (2010) Effects of thermal model and base temperature on estimates of thermal time to bud break in white spruce seedlings. *Canadian Journal of Forest Research* **40**, 1815–1820.

McElreath, R. (2016) *Statistical Rethinking*, vol. 469 pp. CRC Press, New York.

Overcash, J. & Campbell, J. (1955) The effects of intermittent warm and cold periods on breaking the rest period of peach leaf buds. .

Pan, W., Li, J., Du, Y., Zhao, Y., Xin, Y., Wang, S., Liu, C., Lin, Z., Fang, S., Yang, Y. *et al.* (2023) Epigenetic silencing of callose synthase by vil1 promotes bud-growth transition in lily bulbs. *Nature Plants* **9**, 1451–1467.

Pan, W., Liang, J., Sui, J., Li, J., Liu, C., Xin, Y., Zhang, Y., Wang, S., Zhao, Y., Zhang, J. *et al.* (2021) Aba and bud dormancy in perennials: current knowledge and future perspective. *Genes* **12**, 1635.

Piao, S.L., Liu, Z., Wang, T., Peng, S.S., Ciais, P., Huang, M.T., Ahlstrom, A., Burkhart, J.F., Chevallier, F., Janssens, I.A., Jeong, S.J., Lin, X., Mao, J.F., Miller, J., Mohammat, A., Myneni, R.B., Penuelas, J., Shi, X.Y., Stohl, A., Yao, Y.T., Zhu, Z.C. & Tans, P.P. (2017) Weakening temperature control on the interannual variations of spring carbon uptake across northern lands. *Nature Climate Change* **7**, 359–+.

Pörtner, H.O., Roberts, D.C., Tignor, M., Poloczanska, E.S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A. & Rama, B. (2022) *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press.

Richardson, E. (1974) A model for estimating the completion of rest for 'Redhaven' and 'Elberta' peach trees. *HortScience* **9**, 331–332.

Rinne, P.L.H., Paul, L.K. & van der Schoot, C. (2018) Decoupling photo- and thermoperiod by projected climate change perturbs bud development, dormancy establishment and vernalization in the model tree populus. *BMC Plant Biology* **18**.

Satake, A., Nagahama, A. & Sasaki, E. (2022) A cross-scale approach to unravel the molecular basis of plant phenology in temperate and tropical climates. *New Phytologist* **233**, 2340–2353.

Sønsteby, A. & Heide, O.M. (2014) Chilling requirements of contrasting black currant (*Ribes nigrum* L.) cultivars and the induction of secondary bud dormancy. *Scientia Horticulturae* **179**, 256–265.

Tylewicz, S., Petterle, A., Marttila, S., Miskolczi, P., Azeez, A., Singh, R.K., Immanen, J., Mähler, N., Hvidsten, T.R., Eklund, D.M. *et al.* (2018) Photoperiodic control of seasonal growth is mediated by aba acting on cell-cell communication. *Science* **360**, 212–215.

van der Schoot, C., Paul, L.K. & Rinne, P.L.H. (2014) The embryonic shoot: a lifeline through winter. *Journal of Experimental Botany* **65**, 1699–1712.

Walde, M.G., Wenden, B., Chuine, I., Gessler, A., Saurer, M. & Vitasse, Y. (2024) Stable water isotopes reveal the onset of bud dormancy in temperate trees, whereas water content is a better proxy for dormancy release. *Tree Physiology* **44**, tpae028.

Weinberger, J.H. (1950) Chilling requirements of peach varieties. *Proceedings of the American Society for Horticultural Science* **56**, 122–128.

Wilczek, A.M., Roe, J.L., Knapp, M.C., Cooper, M.D., Lopez-Gallego, C., Martin, L.J., Muir, C.D., Sim, S., Walker, A., Anderson, J., Egan, J.F., Moyers, B.T., Petipas, R., Giakountis, A., Charbit, E., Coupland, G., Welch, S.M. & Schmitt, J. (2009) Effects of genetic perturbation on seasonal life history plasticity. *Science* **323**, 930–934.

Wolkovich, E.M., Auerbach, J., Chamberlain, C.J., Buonaiuto, D.M., Ettinger, A.K., Morales-Castilla, I. & Gelman, A. (2021) A simple explanation for declining temperature sensitivity with warming. *Global Change Biology* **27**, 4947–4949.

Wolkovich, E.M., Chamberlain, C.J., Buonaiuto, D.M., Ettinger, A.K. & Morales-Castilla, I. (2022) Integrating experiments to predict interactive cue effects on spring phenology with warming. *New Phytologist* **235**, 1719–1728.

Yu, S., Wenden, B., Ferguson, L. & Tian, L. (2024) Building climate resilient deciduous tree crops by deciphering winter dormancy.

Zhao, Y., Antoniou-Kourounioti, R.L., Calder, G., Dean, C. & Howard, M. (2020) Temperature-dependent growth contributes to long-term cold sensing. *Nature* **583**, 825–829.

Zhu, P., Lister, C. & Dean, C. (2021) Cold-induced arabidopsis frigida nuclear condensates for flc repression. *Nature* **599**, 657–661.
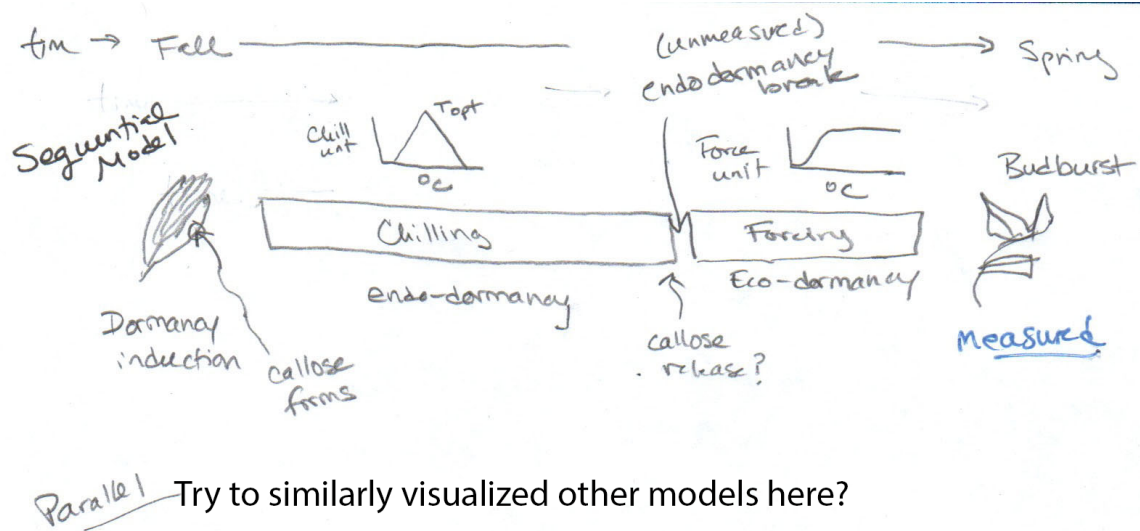
2   Figures

Figure 1: Draft idea of how to visualize simple sequential model but layering on physiological stages, some of the model parameters (not fully shown, but could go with small graphics above 'chilling' and 'forcing'), what is measured, and callose. I wanted to add other models, like the parallel, but was not sure how. I can work on that if we want to keep this figure idea.
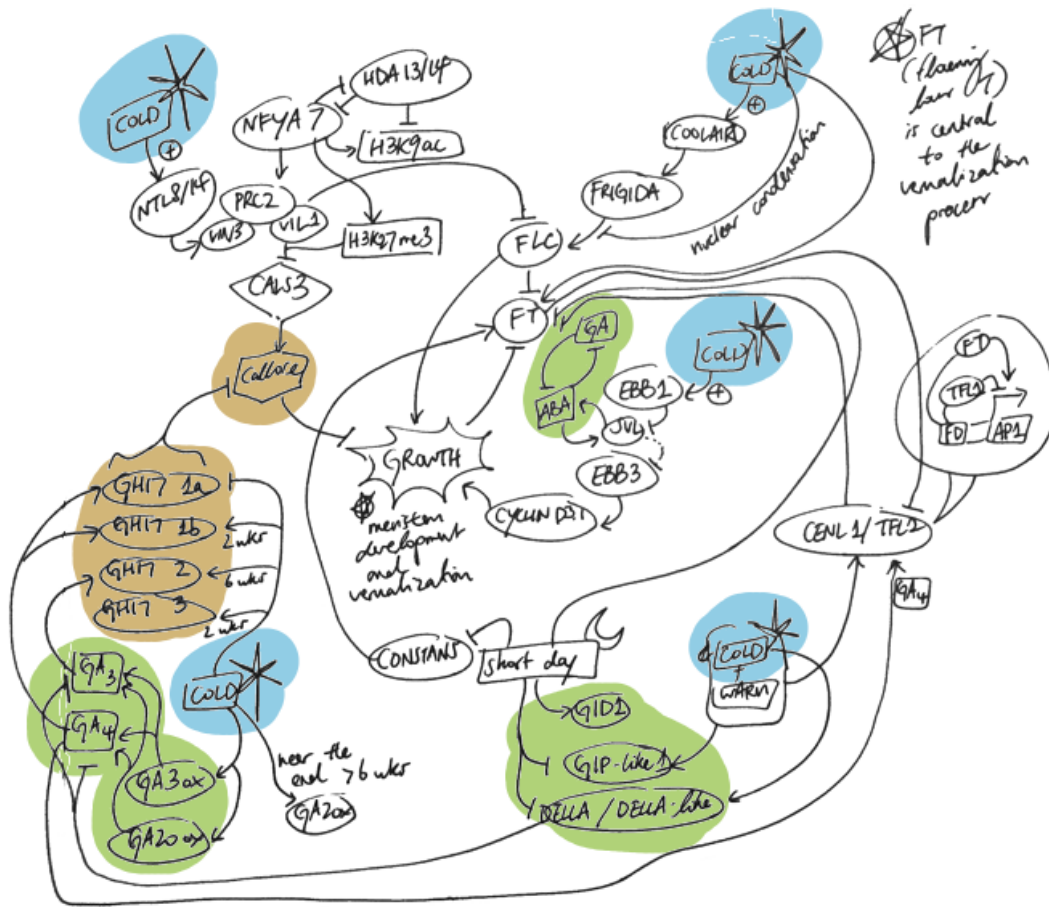
Figure 2: Draft visualization of molecular pathways identified in vernalization and chilling; if we want to keep this, we can update to clarify which have been found in woody species perhaps and match the text a little more. ADD to caption: In the first phase—endo-dormancy—plants are accumulating 'chilling' and cannot respond to external warm periods (thus the 'endo' part of the term, Chuine *et al.*, 2016; Lundell *et al.*, 2020). Once they have accumulated the appropriate amount of 'chilling' they are 'eco-dormant,' and will leaf or flower in response to a sufficient thermal sum (often called forcing, or 'growing degree days' in some models).