

# Patterns of spring-freeze risk for temperate trees contributes to phenological cue differences, but leaves much unexplained Other

Dan, Cat, Nacho, Ben Cook, Faith, Deidre, Mira Geoff and Lizzie

January 7, 2026

## Abstract

## Introduction

Phenology, the timing of annual life cycle events, allows for organisms to match critical life-cycle transitions with optimum environmental conditions. Through the phenology of spring budburst, temperate woody plants balance the resource and competition advantages of precocious leafout with the risk of damage from late season freezes (Savage and Cavender-Bares, 2013). To navigate this trade-off, woody plants have evolved complex physiological responses to sense environmental cues that signal the arrival of appropriate conditions for resuming growth (Polgar and Primack, 2011). Decades of research on phenology suggest that warming spring temperatures (forcing), cool winter temperatures (chilling) and day length (photoperiod) are the primary environmental cues for woody plant phenology in temperate regions (Ettinger et al., 2020; Forrest and Miller-Rushing, 2010). These studies also demonstrate that there are substantial cue-use differences among species, with some species relying more heavily on some cues over others (Laube et al., 2014; ?). Yet, our knowledge about why species differ in their cue responses is currently limited, and better understanding the ecological and evolutionary drivers that shape phenological cues is critical for our ability to predict the magnitude and impacts of phenological shifts with climate change.

The predictability of the arrival spring may strongly influence the evolution of phenological cues (Zohner et al., 2017; ?, ?). In regions where the start of spring is unpredictable, species should evolve stronger dependence on chilling and photoperiod cues to prevent premature leafout and exposure to frost damage. In contrast, in regions where the seasonal warming reliably indicates the start of spring, species should respond strongly to forcing and not chilling or photoperiod. This spring predictability hypothesis (hereafter:SPH) is intuitive and has found some recent support in the literature (Zohner et al., 2017). However, the SPH hinges on the assumption that species phenological responses are at a stable equilibrium with their environment, as assertion that is not well supported (??). It is also unclear the time scale at which season predictability would shape cue responses. Spring predictability could drive selective pressure to increase chilling and/or photoperiod sensitivity on an evolutionary time scale, or define species ranges based on their inherited cue sensitivities on an ecological time scale. Testing the predictions of SPH across multiple geographic scales can serve to evaluate this hypothesis, and offer an improved understanding of the drivers of biogeographic patterns of phenology cue sensitivity.

## Spatial predictions of the SPH

Global circulation patterns generate substantially different spring climatic conditions on either side of the North Atlantic (). In Eastern North America, the spring is marked by instability, while in Europe, the arrival of spring is generally more consistent (Figure 1). Given these contrasting climate regimes, the SPH predicts that North American species should have stronger sensitivity to chilling and photoperiod and weaker to forcing (?). We tested these predictions by comparing the cue sensitivities of North American and Europe species estimated from Bayesian hierarchical models.

It is also possible that that phenological responses to cues play a more important role in determining species range limits (?). The distributions of species that rely primarily on forcing, should be restricted geographic regions where spring predictability is high, while species that rely more heavily on chilling and photoperiod can persist in regions where spring predictability is low. If this is the case, the SPH predicts that within each continent, species with higher chilling and photoperiod cue responses should be associated with lower spring predictability across their ranges. We tested this prediction by regressing the species-level posterior estimates of forcing, chilling and photoperiod sensitivity from our previously described model against the two metrics of spring predictability, with separate models for each cue by continent.

Finally, it is possible local adaptation in phenological cue sensitivity overwhelms any relationship between species-level cue use and range-wide climate conditions. However, if cues are locally adapted, it follows that neither continental differences or range wide climate conditions would be strongly associated with cue use. To assess this possibility in our data, we designed a two-level, hierarchical model for studies in the OSPREE database that sampled species from multiple provenance locations to partition variance between populations and species.

Our study interrogated the SPH; one of the most fundamental biological assumptions relating to the evolution of phenological cue responses that is frequently asserted but rarely tested (but see Zohner et al., 2017). In our analyses, we found only mixed support for the SPH, highlighting the need to more holistically integrate multiple other kinds of drivers (community, phylogenetic) into the framework about the evolution of phenological cues.

## Materials and Methods:

### Species' range characteristics and climate data

We extracted climate data from daily gridded meteorological datasets for both Europe and North America. For Europe, we extracted minimum and maximum daily temperatures from the E-OBS dataset (<https://cds.climate.copernicus.eu/cdsapp#!/dataset/insitu-gridded-observations-europe?tab=overview>; last accessed on October 2021) corresponding to the period comprised between 1980 and 2016. Specifically, we used version 17 at a resolution of 0.5 latitudinal degrees. For North America, we extracted minimum and maximum daily temperatures from Justin Sheffield's Princeton Global Forcing dataset (<http://hydrology.princeton.edu/data/pgf/v3/0.25deg/daily/>; last accessed on October 2021) for the same period. We used version v3 at a resolution of 0.25 latitudinal degrees.

For 22 European and 16 North American tree species, we obtained published distributional range maps in shapefile format. European species ranges were downloaded from (<http://www.sciencedirect.com/science/article/pii/S2352340917301981?via%3Dihub#ec-research-data>; last accessed on XXX) (Caudullo et al., 2017) and North American ranges were obtained from <https://www.fs.fed.us/nrs/atlas/littlefia/#> (Prasad, 2003). For each species' range, we extracted climate data corresponding to all grid cells contained within the range.

We used minimum and maximum daily temperatures within species ranges were then used to compute Growing Degree Days (GDD), Growing Degree Days until the last frost and Spring Temperature Variability (STV). GDD was calculated as the summed temperatures above 10C recorded from January 1st until May 31st. GDD to the last frost was calculated as GDD but instead of summing temperatures above a threshold until a fixed date, the sum was performed until the date at which the latest minimum temperature below -5C was recorded. We then calculated the standard deviation (hereafter: Stdv in GDDs to last frost) to this date as one proxy for spring predictability. We then calculated STV, the standard deviation of mean minimum temperature from March 1st until May 31st (Zohner et al., 2017). Specifically, we computed these metrics variability in the within each location across years (1980 to 2016) and for each year across the grid cells comprised within each species' range.

## Statistical analyses

To estimate species-level responses to forcing, chilling and photoperiod we analyzed phenological observations from the Observed Spring Phenology Responses in Experimental Environments (OSPREE database (Wolkovich et al., 2019)) with a Bayesian Hierarchical model developed by ?. This model estimates forcing, chilling and photoperiod sensitivities through phylogenetically-structure partially pooling. The model was fit using Stan modeling language (R) ([www.mc-stan.org](http://www.mc-stan.org), and the data and code used to run this model are available at X.

To test whether European and North American taxa had systematically different cue sensitivities, we extracted the species-levels posterior estimates of forcing, chilling and photoperiod sensitivity and compared forcing, chilling and photoperiod cues using linear regression, with native continent as a main effect predictor.

To assess the relationship between spring predictability—estimated across species' ranges—and cue sensitivity, we regressed the posterior estimates of cue sensitivity against spring predictability. We used separate models for European and North American taxa, and modeled each cue sensitivity (forcing, chilling and photoperiod) separately. We also ran separate models using STV and Stdv in GDDs to last frost as proxies for spring predictability, as main effects. We ran these model using the R packages BRMS.

To assess variation within and across sites, we designed a two-level, hierarchical model using data from the OSPREE database. We subsetted the studies to include only those that had multiple provenance locations. Since chilling estimates were either from experimental chilling, from field chilling or a combination of both, we removed 'chilling' as a predictor for this model

since it correlated so strongly with provenance latitude and would result in nonidentifiability in our results. We used a Bayesian mixed-effects hierarchical model approach to analyze our data to best estimate the day of budburst. We fit a Gaussian distribution model using study, species and population as intercepts, forcing and photoperiod as predictors (fixed effects) and species nested within population (i.e., site) as modeled groups (random effects). The Bayesian model was fit using Stan modeling language ([www.mc-stan.org](http://www.mc-stan.org)), accessed via the *rstan* package (version 2.15.1), version 2.3.1, in R ([www.r-project.org/](http://www.r-project.org/)), version 3.3.1. We ran four chains, with 2,500 warm-up iterations followed by 3,000 sampling iterations, resulting in 12,000 posterior samples for each parameter.

For all models assessed good model performance through  $\hat{R}$  close to 1 and high  $n_{eff}$  as well as visual consideration of chain convergence and posteriors.

## Results

### Inter-continental Climatic Pattern

We found no substantial differences in cue sensitivity between continents (Figure: 2). Mean forcing sensitivity for European species was  $-6.76 \text{ UI}_{95}(-17.80, 2.08)$  and  $-7.94 \text{ UI}_{95}(-17.90, 1.93)$  for North American species. Mean photoperiod sensitivity was  $8.44 \text{ UI}_{95}(-22.60, 4.69)$  and  $-8.76 \text{ UI}_{95}(-26.90, 4.82)$  for North American species. Mean photoperiod sensitivity was  $1.36 \text{ UI}_{95}(-5.91, 2.89)$  and  $-1.35 \text{ UI}_{95}(-5.88, 2.98)$  for North American species.

### Within continent ranges

We found mixed support for the continent level predictions of the SPH. With STV as the metric for spring predictability, we found that in Europe—spring predictability across a species' range had no clear relationship to forcing or photoperiod cues (mean forcing effect:  $-0.59$ ,  $\text{UI}_{95}[-4.44, 3.01]$ , mean photoperiod effect:  $1.34$ ,  $\text{UI}_{95}[-0.49, 3.24]$ ). Contrary to the predictor of the SPH,  $6.37$ ,  $\text{UI}_{95}[3.99, 7.97]$ .

For North American species, spring predictability also had no clear relationship to forcing or photoperiod cue sensitivity (mean forcing effect:  $-0.05$ ,  $\text{UI}_{95}[-2.26, 2.06]$ , mean photoperiod effect:  $0.15$ ,  $\text{UI}_{95}[-0.93, 1.21]$ ). For chilling, decreasing spring predictability was associated with increased chilling ( $-3.45$ ,  $\text{UI}_{95}[-7.50, 0.69]$ ), supporting the prediction of the SPH.

With standard deviation in growing degree days to last frost as the metric for spring pre-

dictability, we observed different relationships between spring predictability and species-level cue sensitivity than we found using STV. For European species, we found decreasing spring predictability was associated with both higher forcing sensitivity (mean forcing effect: -3.61,  $UI_{95}[9.04, 1.64]$ ) and chilling sensitivity (mean chilling effect: -5.27,  $UI_{95}[11.45, 1.36]$ ), while there was no clear association between decreasing spring predictability and photoperiod sensitivity (mean forcing effect: -0.80,  $UI_{95}[-3.62, 2.09]$ ).

For North American species, decreased predictability was associated with weaker forcing sensitivity though the strength of this association was relatively weak (mean forcing effect: 0.09,  $UI_{95}[-0.04, 0.24]$ ). There was no clear relationship between spring predictability and chilling sensitivity (mean chilling effect: -0.02,  $UI_{95}[-0.34, 0.31]$ ). For photoperiod sensitivity, we observed a moderate association between decreasing spring predictability and photoperiod sensitivity (mean forcing effect: -0.04,  $UI_{95}[-0.11, 0.04]$ ).

## Intra-and Interspecific cue variation

While we detected limited population level variation in forcing and photoperiod cue sensitivity, though this within species variation was less substantial than among species variation (Fig. 6). Notably, we found the largest source of variation in phenological cue estimates was the study effect (Fig. 6). This result does not support the assertion that local adaptation is masking relationships between cue sensitivity and range-level climatic patterns.

## Discussion

In this study we compared estimates of cue sensitivity to climate conditions experienced by species across their range at multiple scales to test the predictions that spring predictability shapes cue sensitivities. Across scales, we found support for the SPH to be weak, highlighting a need to think more expansively about the factors that shape phenological cue response.

Contrary to expectations, we found no differences in cue sensitivities between North American and European species. This is not particularly surprising given that recent studies have found there to be strong phylogenetic conservatism in phenological cue responses, and that there are many closely related congeners found in both North America and Europe. It is therefore likely that patterns of cue use diverged among taxa well before the modern placement of continents, under different climate conditions than North America and Europe experience today. (wow say better). These results call into question the recent assertion that European plant species successfully invade North American ecosystems because their higher reliance on forcing cues

allows them to leafout earlier and gain a growth advantage over their competitors (?). While these kinds phenological priority effects have been documented as contributing to the success of invaders (??) our findings indicate that other mechanisms are likely more important for explaining the success of European woody plants in North America. Instead, this finding may help us understand why many European timber species have been successfully established in Northern America (and visa versa), without becoming aggressive on the landscape. We should note that when we subset this analysis to include only the 29 for which we could find well developed range maps, we did in-fact observe European species to have a weaker chilling sensitivity (Figure S1, need to make it), which may further explain the pervasiveness of the idea that European invaders are successful in North America due to weaker secondary phenological cues.

## 0.1 Continetal scale

For Europe, we found spring predictability measured was associated with increased chilling sensitivity and forcing sensitivity, especially when we measured spring predictability with variation in growing degree days to last frost. These patterns qualitatively match the previous findings from Zohner et al. (2017), that report stronger reliance on secondary cues (i.e. chilling) in North American compared to Europe based on the substantial differences in STV between the continents.

North America has substantially weaker spring predictability (i.e., higher STV ) than Europe, but despite this, we observed no clear relationship between

When considering the climate patterns experienced by species across their ranges, our analyses of the relationships between spring predictability and phenological sensitivity to forcing, chilling and photoperiod offer only marginal support for the SPH. For North American species, less spring predictability was associated with an increased reliance on chilling (i.e. chilling sensitivity) when we used STV as a metric of predictability and and increased reliance on photoperiod when we used Stdv in GDDs to last frost as a metric of spring predictability. For European species spring predictability was associated with increased reliance on chilling (i.e. stronger chilling sensitivity) when we used Stdv in GDDs to last frost as a metric of predictability. However, in interpreting these relationships, it is important to recognize that the amount of variation in GDDs to last frost in Europe is very small (0.62 to 2.97), especially when compared to North America (13.68 to 56.01). This suggest that the arrival of spring in Europe is always relatively predictably from a growing degree day perspective, and

we should exhibit caution in putting too much biological stock in the statistical association between European spring predictability and chilling cue sensitivity.

There is currently a debate about the extent to which local adaptation shapes phenological cues (?), and it is likely the strength of local adaptation varies among species, and phenological phases themselves (??).

While these three facets of the predictability/cue relationship support the expectations of the SPH, all other aspects of the predictability/cue relationship either show no relationship (i.e. forcing & photoperiod effects with STV in Europe & North America) or observed relationships contrary to the expectations of the SPH (i.e., chilling in Europe with STV and forcing in North America with Stdv. GDDs). We argue that these mixed results should not serve to refute the claims of the SPH—it is clear that spring predictability has a role to play in shaping where and how species successfully leaf out. However, our study indicates a more complicated picture of the factors that shape species’ level differences in phenological cues.

Our study indicates that predictability of spring species’ experience across their geographic ranges may play a role in shaping phenological cue sensitivities, there must be more to the story. When considering the fact that many species with divergent phenological cue sensitivities have highly overlapping ranges (Make Figure S2) and experience similar environmental conditions (i.e., spring predictability, late season frost risk, etc) across their ranges, it is worthwhile to consider phenological cues in a community context ().

**I am hoping Deirdre and Lizzie have capacity to write a paragraph or two here basically phenology should be considered like a functional trait in that they not only relate to resource aquisition/growth/ survival but also do niche stuff and therefore end up under stabilizing selection though competition and species interactions**

## 1 Conclusion:

In this study we found limited support for the assertion that the predictability of spring shaping their relative reliance on forcing, chilling and photoperiod cues for spring phenology. Our results suggest that climate variability may drive cue use pattern only when it is sufficiently high, like in contemporary North America. These results suggests that future studies of phenological cue-use would benefit from a holistic integration of these bio-climatic hypotheses with phylogenetic, functional trait, and climatic legacy hypotheses to fully understand the



evolution of phenological cues in woody plants, and how cue use patterns will impact species performance in the face of global change at across multiple spatial and temporal scales.

## References

Caudullo, G., E. Welk, and J. San-Miguel-Ayanz. 2017. Chorological maps for the main European woody species. *Data in Brief* 12:662–666.

Ettinger, A. K., C. J. Chamberlain, I. Morales-Castilla, D. M. Buonaiuto, D. F. B. Flynn, T. Savas, J. A. Samaha, and E. M. Wolkovich. 2020. Winter temperatures predominate in spring phenological responses to warming. *Nature Climate Change* 10:1137–1142.

Forrest, J., and A. J. Miller-Rushing. 2010. Toward a synthetic understanding of the role of phenology in ecology and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365:3101–3112.

Laube, J., T. H. Sparks, N. Estrella, J. Höfler, D. P. Ankerst, and A. Menzel. 2014. Chilling outweighs photoperiod in preventing precocious spring development. *Global Change Biology* 20:170–182.

Polgar, C., and R. Primack. 2011. Leaf-out phenology of temperate woody plants: From trees to ecosystems. *New Phytologist* 191:926–41.

Prasad, A. 2003. Little’s range and FIA importance value database for 135 eastern US tree species. [www.fs.fed.us/ne/delaware/4153/global/littlefia/index.html](http://www.fs.fed.us/ne/delaware/4153/global/littlefia/index.html).

Savage, J. A., and J. Cavender-Bares. 2013. Phenological cues drive an apparent trade-off between freezing tolerance and growth in the family salicaceae. *Ecology* 94:1708–1717.

Wolkovich, E. M., A. K. Ettinger, D. Flynn, T. Savas, C. Chamberlain, D. Buonaiuto, and J. Samaha. 2019. Observed spring phenology responses in experimental environments (OSPREE). doi:10.5063/F1CZ35KB.

Zohner, C. M., B. M. Benito, J. D. Fridley, J.-C. Svenning, and S. S. Renner. 2017. Spring predictability explains different leaf-out strategies in the woody floras of north america, europe and east asia. *Ecology Letters* 20:452–460.

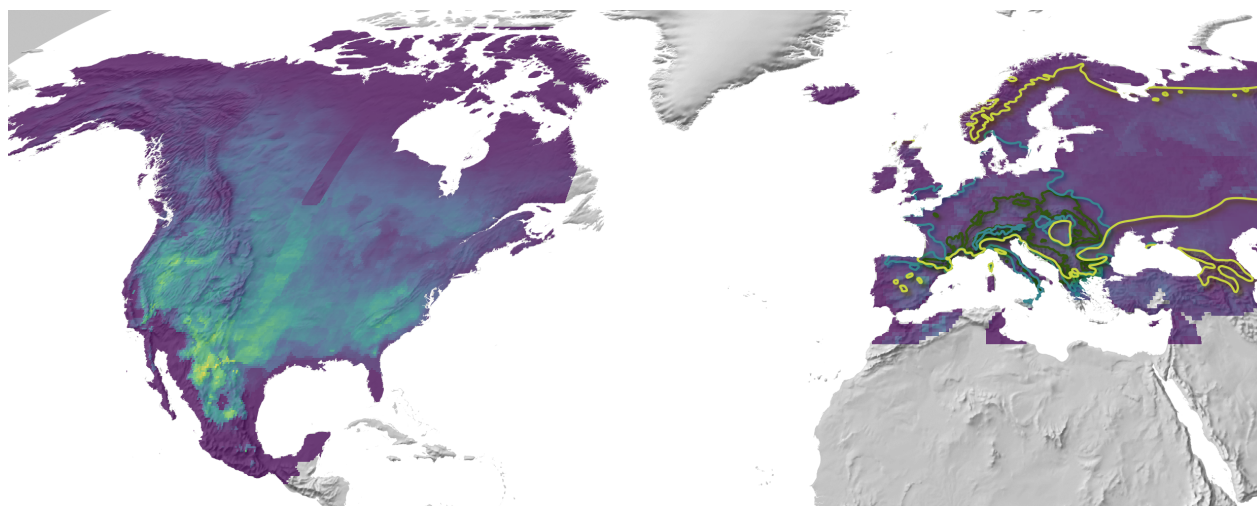


Figure 1: Contrasting patterns of spring predictability in North America and Europe.

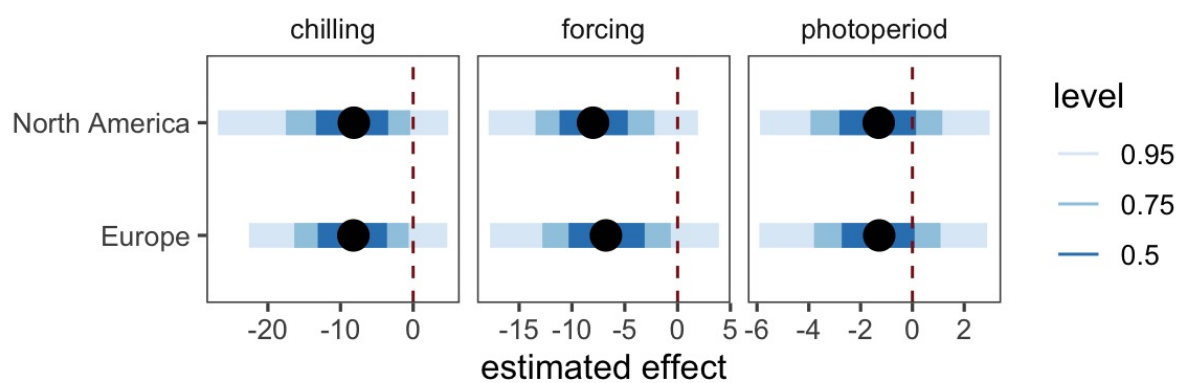


Figure 2: No difference between continents

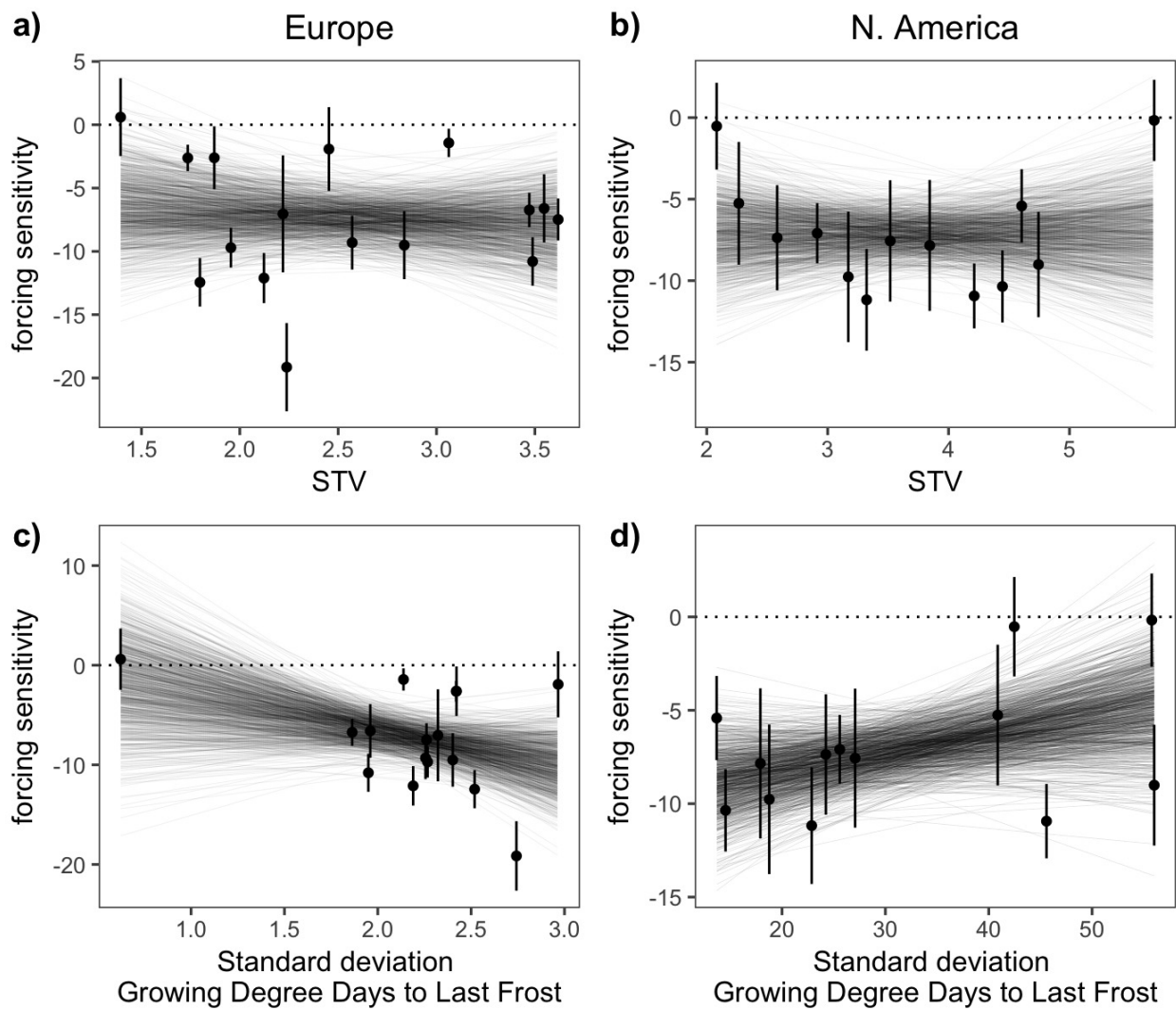


Figure 3: Forcing

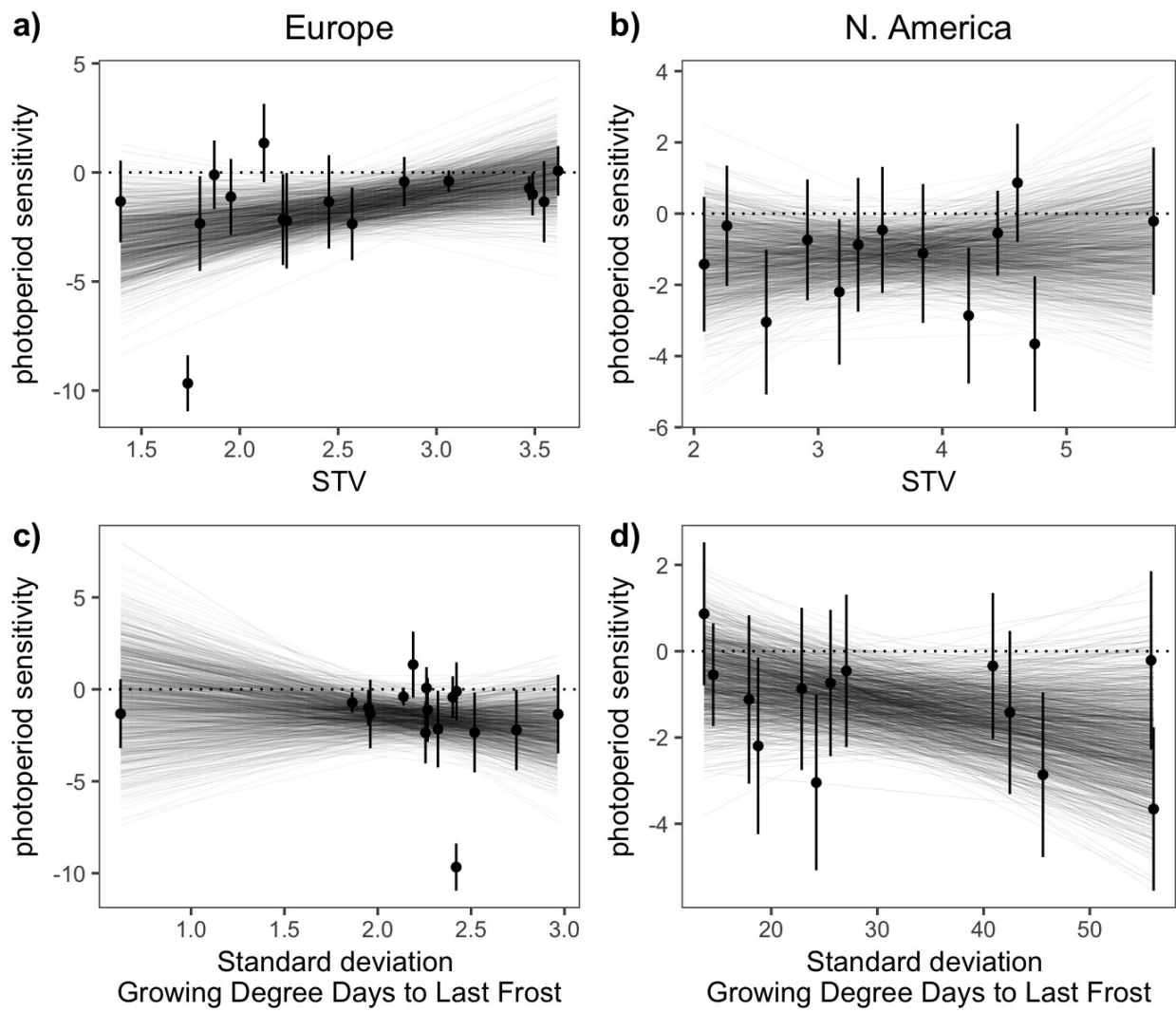


Figure 4: Photoperiod

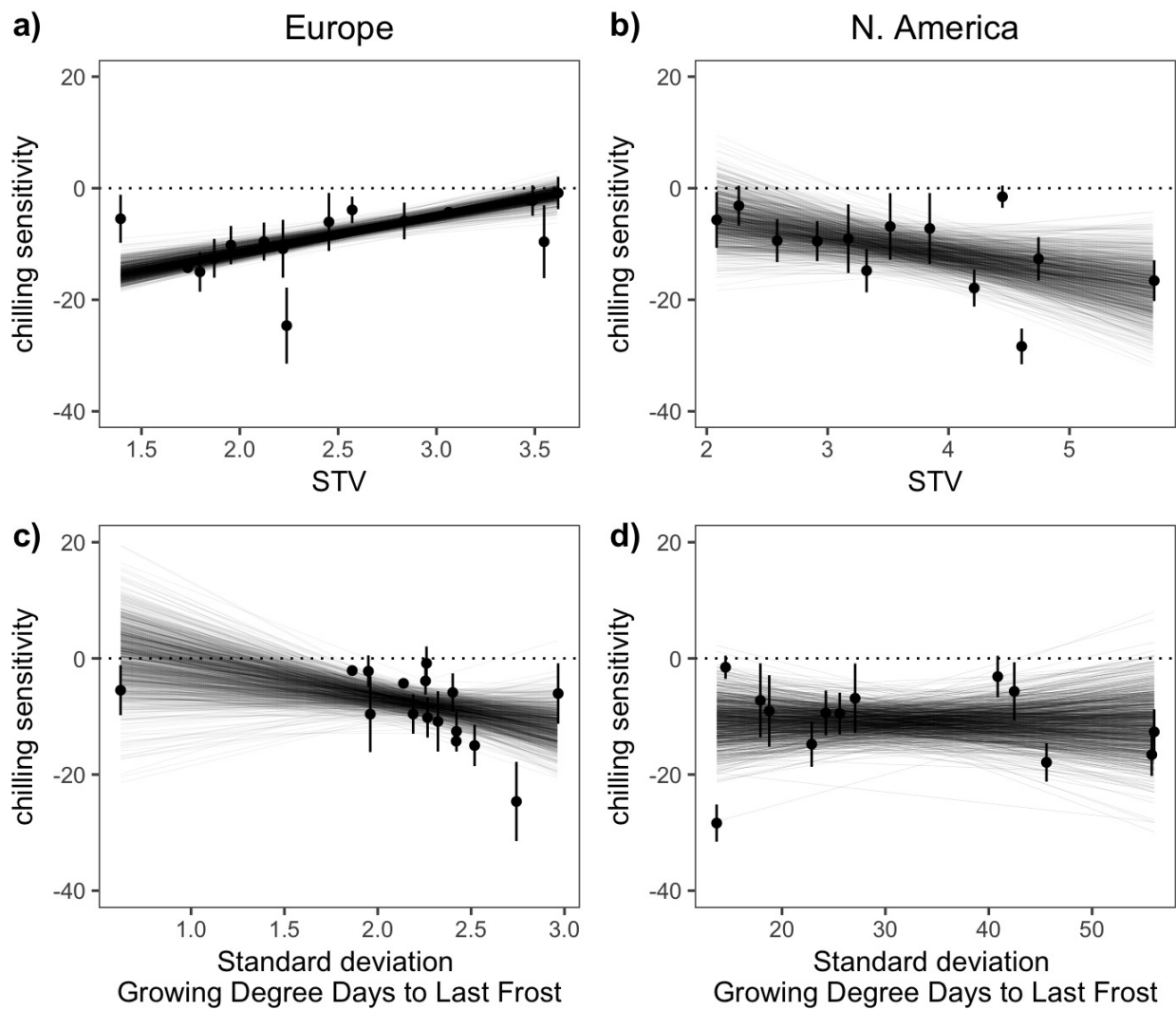


Figure 5: Chilling

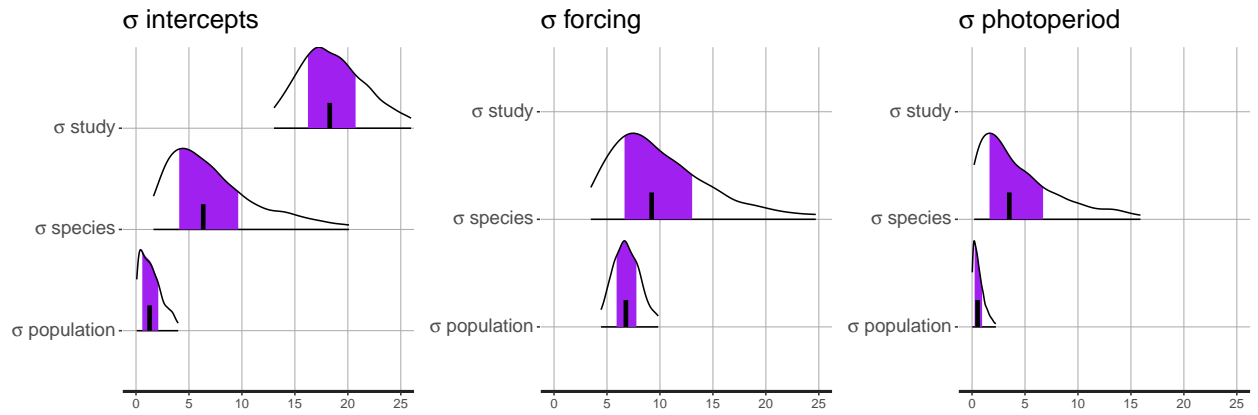


Figure 6: **Local adaptation model estimates of variation partitioning in the intercept and forcing and photoperiod predictors using the OSPREE dataset.** For both the forcing and photoperiod predictors, within species (intra-specific) variation is much smaller than across species (inter-specific) variation. Here we see that inter-specific variation exceeds intra-specific variation at the intercept-level as well but variation at the study level is largest, suggesting experimental design is driving the highest level of uncertainty.