

Supplemental materials for Chilling outweighs photoperiod and forcing cues for temperate trees in experiments, but not in natural systems

A.K. Ettinger, C. Chamberlain, I. Morales-Castilla, D. Buonaiuto, D. Flynn, T. Savas,
J. Samaha & E. Wolkovich

June 14, 2019

May need to move back to main text ...

The paucity of studies directly manipulating chilling—which our results suggest has the greatest effect on budburst—suggests a major gap in current research. While many studies (X out of Y here) directly manipulated forcing, far fewer directly manipulated chilling (Z out of Y). Our model highlights how the choice of chill units can affect model estimates and associated forecasts (reference supplemental figure with utah units vs chill portions in main manuscript). Given the limited manipulation of chilling in studies, we estimated chilling for all studies by combining chilling from the field (i.e., chilling before plants were sampled) and experimental chilling into two widely used metrics of chilling: Utah and Chill portions (Dennis, 2003). We found the effects of chilling and other cues remain qualitatively consistent across the two chilling units, though chilling and photoperiod estimates were slightly lower using chill portions compared to Utah (cite supplemental table comparing estimates with both units).

An additional important limitation is the rarity of studies designed to test interactions. Interactions between these cues are widely expected (cites) and, when examined, often found (cites), we were unable to estimate interactions in our meta-analysis because very few studies design experiments to test for interactions between chilling, forcing, and photoperiod (cite table with number of interactions from coding challenge!). The few that do incorporate interactions generally use the Weinberger method, which is not designed to robustly tease out of the effects of multiple cues (cites, Tables, figs). Our estimated effects average over interactions (Gelman and Hill, 2006), but identifying them in future research will be critical to understanding and predicting budburst. For example, the most commonly observed interaction between chilling and forcing—that lower amounts of chilling increases forcing requirements for budburst (cite papers in the ospreo database that interact chilling and forcing)—is the hypothesized cause of declining sensitivities in European trees (Fu et al., 2015; ?). As more data become available, it would allow additional tests of important interactions, such as how responses vary across latitudes (ref latitude figure).

Things we'd like to discuss in main paper but don't have room for

Photothermoperiodicity, for example, is an ongoing challenge: chamber studies may seek to replicate patterns in nature, pairing daylength and temperature treatments such that night temperatures are always cooler than day temperatures (e.g., cite studies that do this). This results in daylength treatments that differ in temperature conditions (and therefore chilling and forcing treatments) as well, however.

Chilling: Current common models for chilling (i.e., Utah which was developed for XX species, chill portions which was developed for XX species) are *hypotheses* for how chilling may accumulate to affect the process of dormancy release, but are likely to be inaccurate for many species. Our model highlights how the choice of chill units can affect model estimates and associated forecasts (Figures 1,????, 1).

Supplemental Methods

1. Data: search terms, etc: Tim and Dan F.
2. **Species complex description:** Some species are only represented in one dataset in the OSPREE database, making it impossible to differentiate between species and treatment statistical effects for these taxa. To address this, we combined species found in only one study into “complexes” at the level of genera. Thus our taxonomic units of analysis are “species complexes”; either species represented in >1 dataset or generic complexes combining multiple species that are each singly represented in the dataset. Species represented in only one dataset with no congeners in other datasets were excluded from our analysis.
3. Equation of our model: Ailene and Lizzie

The models were fit using the programming languages **Stan** (Carpenter et al., 2017)(www.mc-stan.org), accessed via the *rstan* package (version XXX) in R (R Development Core Team, 2017), version XXXX. Stan provides efficient MCMC sampling via a No-U-Turn Hamiltonian Monte Carlo approach (more details can be found in Gelman et al. (2014) and in Carpenter et al. (2017)). We validated that our model code using test data, then fit the following model:

$$y_i = N(\alpha_{sp[i]} + \beta_{forcing_{sp[i]}} + \beta_{photoperiod_{sp[i]}} + \beta_{chilling_{sp[i]} + \epsilon_i}, \epsilon_i \sim N(0, \sigma_y^2)$$

The α and each of the three β coefficients were modeled at the species level, as follows:

$$\begin{aligned} \alpha_{sp} &\sim N(\mu_\alpha, \sigma_\alpha) \\ \beta_{forcing_{sp}} &\sim N(\mu_{forcing}, \sigma_{forcing}) \\ \beta_{photoperiod_{sp}} &\sim N(\mu_{photoperiod}, \sigma_{photoperiod}) \\ \beta_{chilling_{sp}} &\sim N(\mu_{chilling}, \sigma_{chilling}) \end{aligned}$$

We choose weakly informative priors; increasing the priors three-fold did not change the model results.

We ran four chains simultaneously, each with 1 500 warm-up iterations followed by 2 500 sampling iterations, yielding 4 000 posterior samples for each parameter. We assessed model performance through \hat{R} close to 1 and high n_{eff} (4 000 for most parameters, but as low as XXXX for several parameters) as well as visual consideration of chain convergence and posteriors (Gelman et al., 2014).

In our figures and tables we show means \pm XX% credible intervals from this model, because of our focus here is on the most likely value for each parameter (e.g., estimated response to forcing) and because they are computationally stable (Gelman et al., 2014; Carpenter et al., 2017). See tables XX other XX% credible intervals.

1. Quick review of what standardized units are (We used a standardized scale to allow comparisons of the three cues ...)

2. Forecasting with the OSPREE model: We selected sites in Germany where temperature and budburst have been monitored since the 1950s. We extracted mean temperature data from 1951 through 1961 (pre warming time period) and used these values as baseline data. We then investigated model predictions of budburst given different levels of warming (from 1-7 °C) above this baseline, including altered chilling and forcing as well as potential declines in photoperiod due to advancing phenology. We did this for one common European species: *Betula pendula* (silver birch) at all lat/longs included in the PEP database between 1951 and 1961. We also did this for another common European species, *Fagus sylvatica*, for a subset of sites where it occurred with *B. pendula*, in order to compare budburst responses of these two species when they experience the same baseline climate and warming levels.
3. To understand how experimental temperature, photoperiod, and budburst sensitivity compares to past and current conditions in nature, we used data from the PEP database (cite). We summarized forcing, chilling, and budburst day for two common species: *Betula pendula* (silver birch) and *Fagus sylvatica* (European beech) during a pre-warming time-period (1950-1980) and post-warming period (1981-2014?).

Supplemental Results/Discussion

1. Surprising species-specific responses:
 - (a) Positive Responses to chilling for *Tilia* and *Salix*. have positive response to chilling with chill portions model.
 - (b) Poisitive responses to forcing: *Fagus*, Acer-complex, *Fraxinus* complex, *Cornus alba*.

Supplemental Tables

1. Table 1S: Species included in OSPREE, with the reference(s) of studies they are included in, the range of forcing, chilling, and photoperiod treatments in these studies, and perhaps the range of latitudes.
2. Table 2S: model estimates (DONE)
3. Table 3S: a table that goes with Figure 3 (the 4-paneled 3D forecasting figure) that includes the mean budburst day of year, chilling estimates, mean winter temperature, and spring (forcing) temperatures for the 4 sites with no warming and with the 7 levels of warming.

Supplemental figures

1. Map of study locations, shading or symbol coding for number of cues (Lizzie)
2. Map of species forecasting to justify sites
3. Heat maps for the main data associated, including by actual study design and by calculated chilling (our calculations)
4. Photoperiod x latitude effects figure
5. Maps of current and forecasted budburst and chilling for PEP sites (heat maps)
6. Weinberger analysis figure

Understanding declines in temperature sensitivity in European long-term data

As our model results (when applied to realistic forcing, chilling and photoperiod conditions in Europe) do not predict a dramatic decline temperature sensitivity, we tested whether observed declines could instead be due to a statistical artifact of the method commonly used to estimate temperature sensitivity (a linear regression of annual budburst date versus mean or other aggregated metrics of spring temperature yielding estimates in days/ $^{\circ}\text{C}$). If warming produces systematically warmer daily temperatures this method will inherently estimate lower sensitivities, because the ‘days’ unit will effectively have increased in the thermal time it represents.

To test this hypothesis we compared observed trends with simulations. First, we collated PEP 725 (CITE) data for *Betula pendula* for all sites with leafout data each year from two 10-year time-periods: 1950-1960, which represented a period before significant anthropogenic warming (citeIPCC) and 2000-2010, a period with significant warming. We used leafout data (BBCH=11; which is defined as “leaf unfolding (first visible leaf stalk” in the PEP725 database)) instead of budburst (BBCH=7; defined as “Beginning of sprouting”) as leafout data are far more common in the PEP 725 database. Next, we simulated budburst data with constant cues. For this, we did not include any chilling or photoperiod cues, but simply assumed budburst occurred after a certain summation of heat, estimated via growing degree days with a base temperature of 0°C . We then estimated temperature sensitivity (days/ $^{\circ}\text{C}$) and the difference in these estimates given different levels of spring warming. For the simulations showed here we used a GDD requirement of 150, a base mean spring temperature of 6°C with a variance of 3°C and estimated temperature sensitivity for 10-year periods for 45 simulated sites (these values were chosen to best match the PEP 725 data, but note that the general findings are robust to other combinations of these parameter values).

As expected temperature sensitivity estimates for *Betula pendula* from PEP 725 declined across the two time periods in step with warming daily temperatures. Across the sites studied here we estimated a decline of 0.9 ± 0.3 days/ $^{\circ}\text{C}$ (comparing 2000-2010 and 1950-1960) and 1.1°C warming; this estimate was very similar to simulations given constant cues and 1°C warming (Fig. 7). Additionally, in both the real and simulated data variance in leafout date declined, as would be expected if plants use a thermal sum threshold of forcing to leaf out and warming produces systematically warmer days: in the PEP 725 data we found a decline in leafout variance of 57%, compared to a decline of 37% in the simulations. Additionally we found little change in accumulated chilling (1 September - 1 March of each year) in the PEP 725 data across the two time points (2247 ± 31 Utah units in 1950-1960, compared to 2236 ± 20 Utah units in 2000-2010), further suggesting that shifts in chilling do not explain the declining sensitivities. Simple plots of the chilling and forcing required for budburst suggest very low chilling is often required to dramatically increase the forcing required for budburst (Fig. fig:pepgddchill).

References

- Carpenter, B., A. Gelman, M. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. A. Brubaker, J. Guo, P. Li, and R. Allen. 2017. Stan: A probabilistic programming language. *Journal of Statistical Software* 76:10.18637/jss.v076.i01.
- Dennis, F. 2003. Problems in standardizing methods for evaluating the chilling requirements for the breaking of dormancy in buds of woody plants. *HortScience* 38:347–350.
- Fu, Y. S. H., H. F. Zhao, S. L. Piao, M. Peaucelle, S. S. Peng, G. Y. Zhou, P. Ciais, M. T. Huang, A. Menzel, J. P. Uelas, Y. Song, Y. Vitasse, Z. Z. Zeng, and I. A. Janssens. 2015. Declining global warming effects on the phenology of spring leaf unfolding. *Nature* 526:104–107.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2014. Bayesian Data Analysis. 3rd ed. CRC Press, New York.

Gelman, A., and J. Hill. 2006. Data analysis using regression and multilevel/hierarchical models. Cambridge university press.

R Development Core Team. 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Tables

Table 1: **Estimates from model fit with standardized predictors.** The model we present in the main text uses Utah units for chilling and includes studies that experimentally manipulated forcing and photoperiod. Using instead a model with chilling in Chill Portions (column two) results in quantitatively different species-level and overall estimates, though the results are qualitatively similar. We also present coefficients from a model including all species (i.e., with crops) with all treatment types. We present estimates and their 95 percent credible intervals from models in which the predictors have been standardized so that they are directly comparable,

parameter	utah.est	utah.ci	cp.est	cp.ci	allsp.est	allsp.ci
mu_a_sp	62.89	54.9-70.93	67.02	58.05-75.76	62.57	57.64-67.38
mu_b_force_sp	-0.79	-1.16-0.42	-0.85	-1.25-0.45	-1.02	-1.29-0.76
mu_b_photo_sp	-0.53	-0.93-0.16	-0.54	-0.9-0.18	-0.14	-0.35-0.06
mu_b_chill_sp	-2.86	-3.78-1.95	-0.25	-0.33-0.16	-2.49	-2.92-2.06
sigma_a_sp	19.17	14.76-25.04	21.93	16.9-28.18	17.7	15.31-20.42
sigma_b_force_sp	0.91	0.63-1.25	0.99	0.69-1.41	0.71	0.57-0.88
sigma_b_photo_sp	0.77	0.49-1.14	0.7	0.46-1.02	0.59	0.45-0.74
sigma_b_chill_sp	2.06	1.46-2.82	0.2	0.14-0.28	1.24	0.95-1.57
sigma_y	15.83	15.28-16.39	15.53	15.01-16.08	15.16	14.77-15.58
n_sp	39		39		205	

Table 2: **Estimates from models fit with predictors on their natural scales**, so that effect sizes can be readily interpreted in a meaningful way (e.g., change in days of budburst per degree C of warming for forcing temperature). The model we present in the main text uses Utah units for chilling. Here we also present coefficients from a model included all species, including crops, and all treatment types. We present estimates and their 95 percent credible intervals from models in which the predictors have been standardized so that they are directly comparable,

parameter	utah.est	utah.ci	cp.est	cp.ci	allsp.est	allsp.ci
mu_a_sp	29.92	26.41-33.37	30.77	27-34.46	30.86	28.61-33.17
mu_b_force_sp	-4.37	-6.7-2.12	-4.83	-7.12-2.46	-6.15	-8.56-3.72
mu_b_photo_sp	-3.13	-5.63-0.77	-3.17	-5.51-0.95	-1.03	-2.31-0.2
mu_b_chill_sp	-8.87	-12.06-5.77	-8.16	-11.19-5.16	-8.04	-9.65-6.42
sigma_a_sp	9.41	7.21-12.2	10.14	7.8-13.19	14.42	12.67-16.42
sigma_b_force_sp	5.68	4.01-7.77	6.04	4.33-8.27	8.51	6.53-10.88
sigma_b_photo_sp	5.22	3.36-7.85	4.46	2.92-6.62	3.69	2.83-4.67
sigma_b_chill_sp	7.38	5.27-10.1	7.88	5.67-10.59	6.27	4.8-7.85
sigma_y	15.77	15.24-16.32	15.47	14.93-16.02	14.93	14.55-15.32
n_sp	39		39		205	

Figures

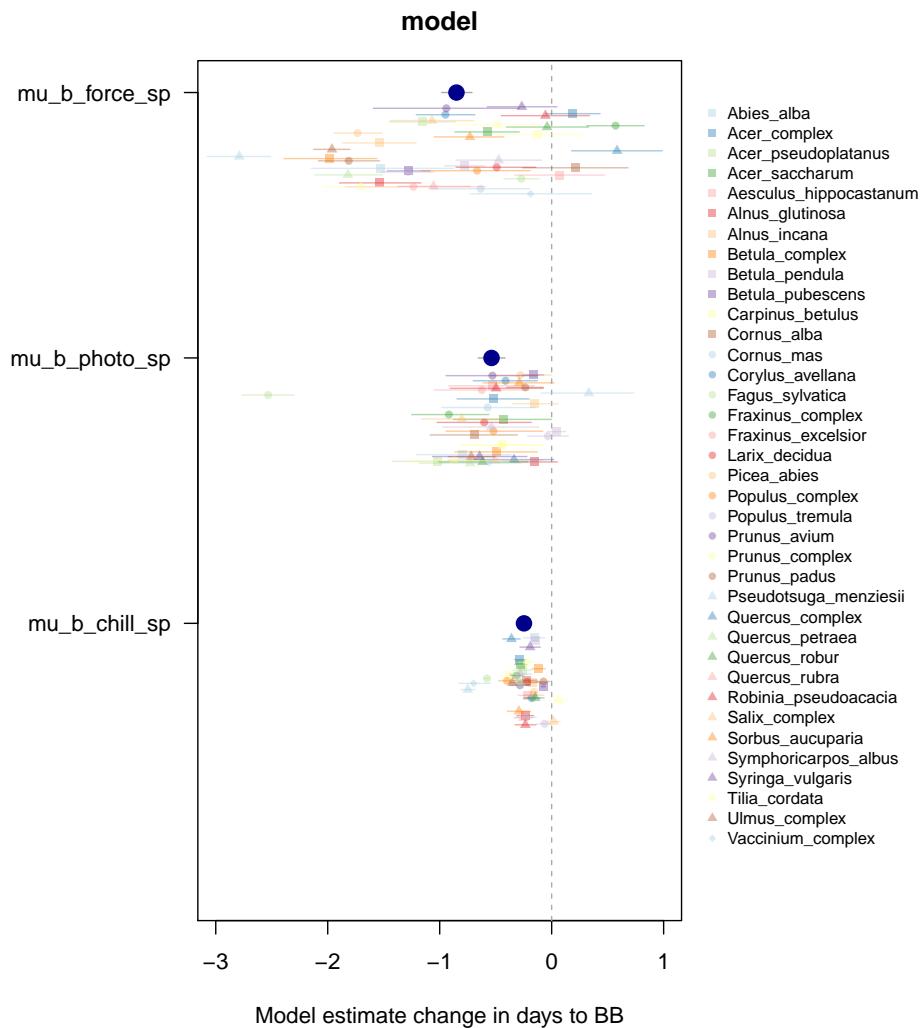


Figure 1: *Estimates for effects of chilling, forcing, and photoperiod from the models fit with uncentered predictors using Chill Portions. Convert this into a table and combine with above figure!*

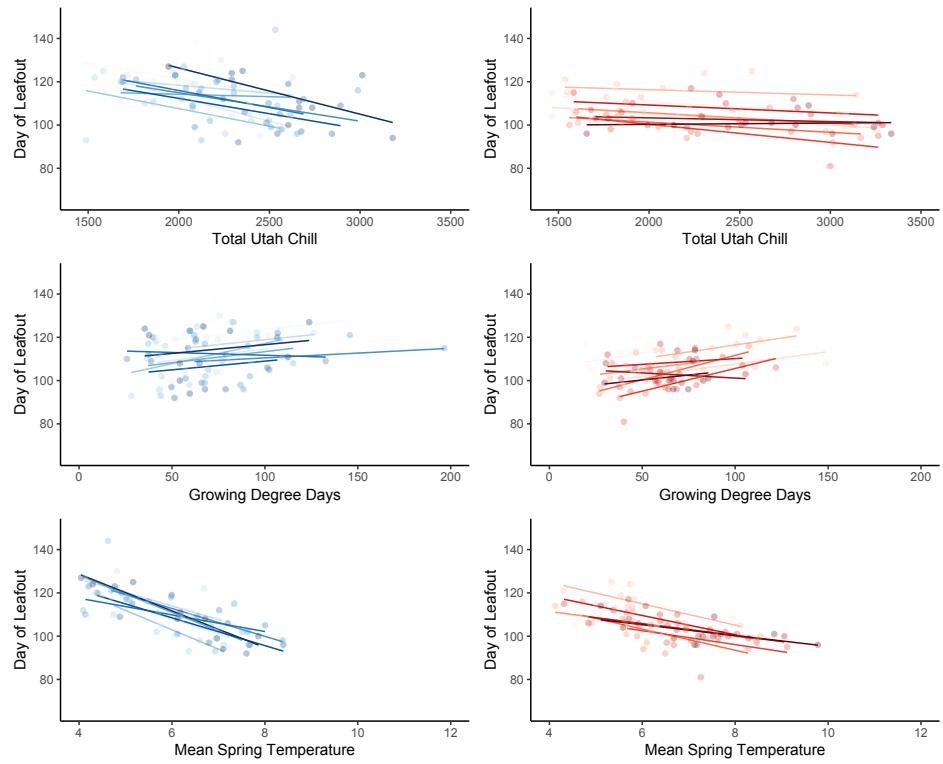


Figure 2: Day of leaf out versus chilling, growing degree-days, and mean spring temperature pre- (left panels, 1951-1961) and post- warming (right panels, 2000-2010) for PEP sites in Germany where *Betula pendula* phenology has been monitored for decades.

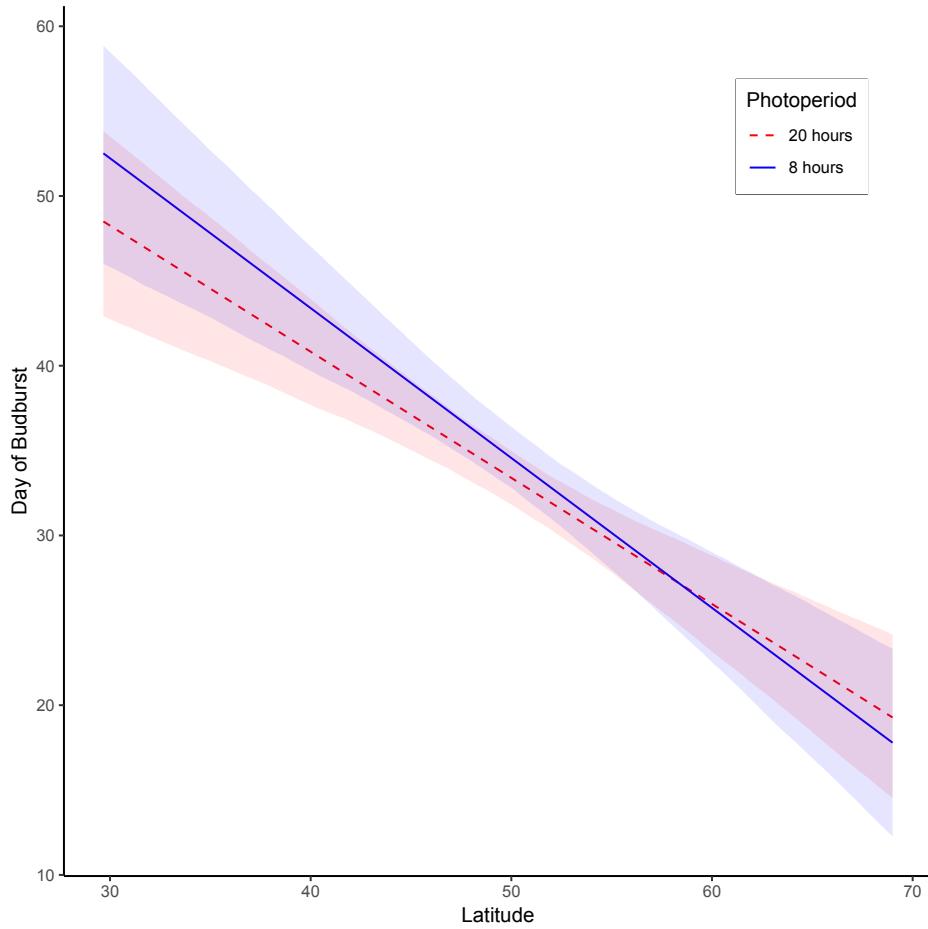


Figure 3: Latitude model. Cat: Could you make a muplot that is similar to those in the main text? Should we keep this plot, or something similar as a second panel?

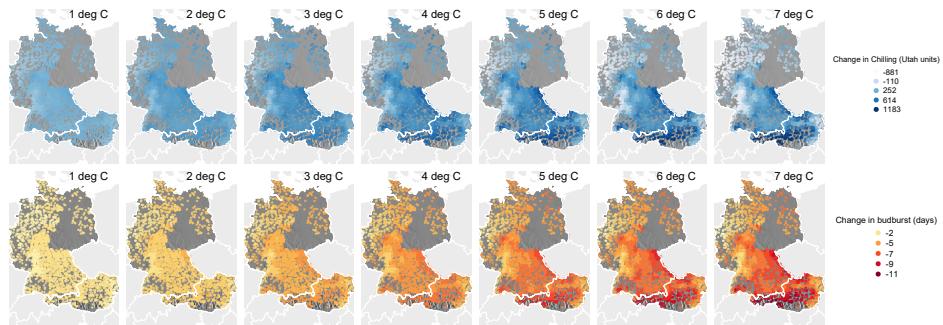


Figure 4: **Forecasted changes in chilling (top panel) and leafout for *Betula pendula* (bottom panel)**, in locations included in the PEP database, where phenology dates are known for the pre-warming time period (1951-1960). Changes in chilling and budburst are calculated relative to the mean chilling and budburst dates during this pre-warming timeperiod for each location.

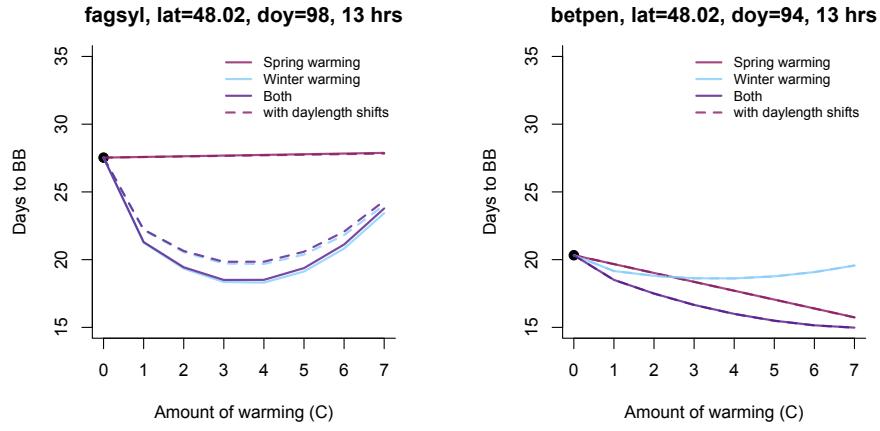


Figure 5: Forecasted budburst with shifts in daylength

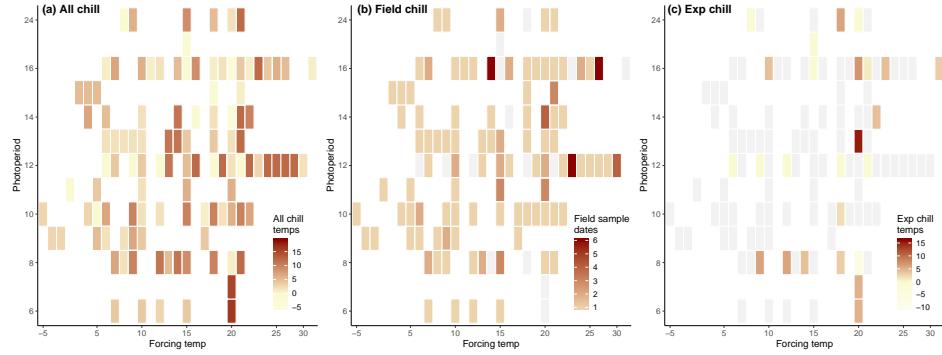


Figure 6: **Heatmaps of treatments ...** need to make sure this matches the data we have ended up using!

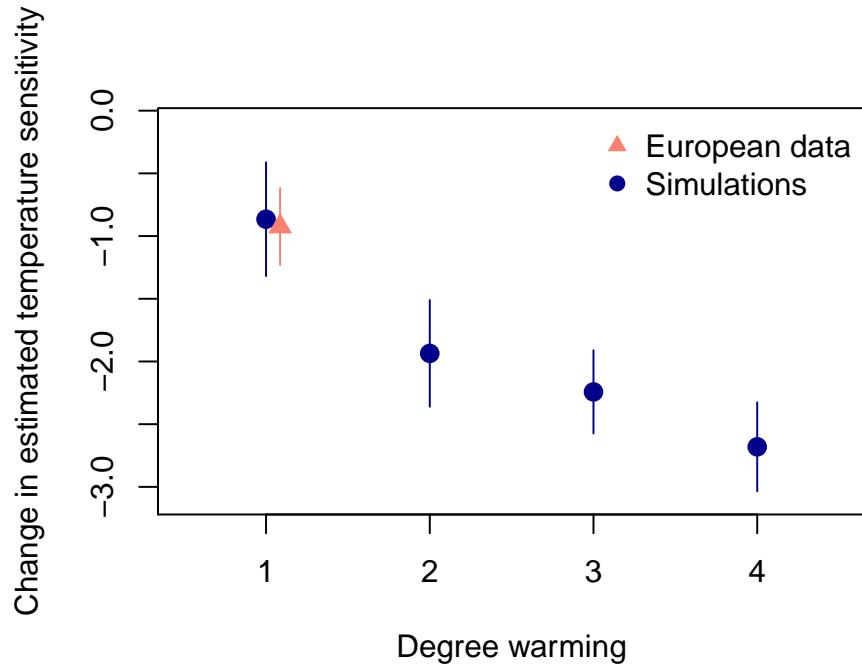


Figure 7: Declining sensitivities observed in long-term European data for a suite of common trees may be explained by a statistical artifact. We compared the sensitivity estimated from linear regressions of day of leafout versus mean spring temperature (estimated thus as days/ $^{\circ}\text{C}$) from PEP 725 data for *Betula pendula* from 45 sites ('European data') with estimated declines in simulations where the cues were held constant but spring temperatures warmed by 1-4 $^{\circ}\text{C}$ ('Simulations') and found the estimated temperature sensitivity measured as days/ $^{\circ}\text{C}$ declined even though the underlying cues had not changed, see *Understanding declines in temperature sensitivity in European long-term data* in Supplement for further details.

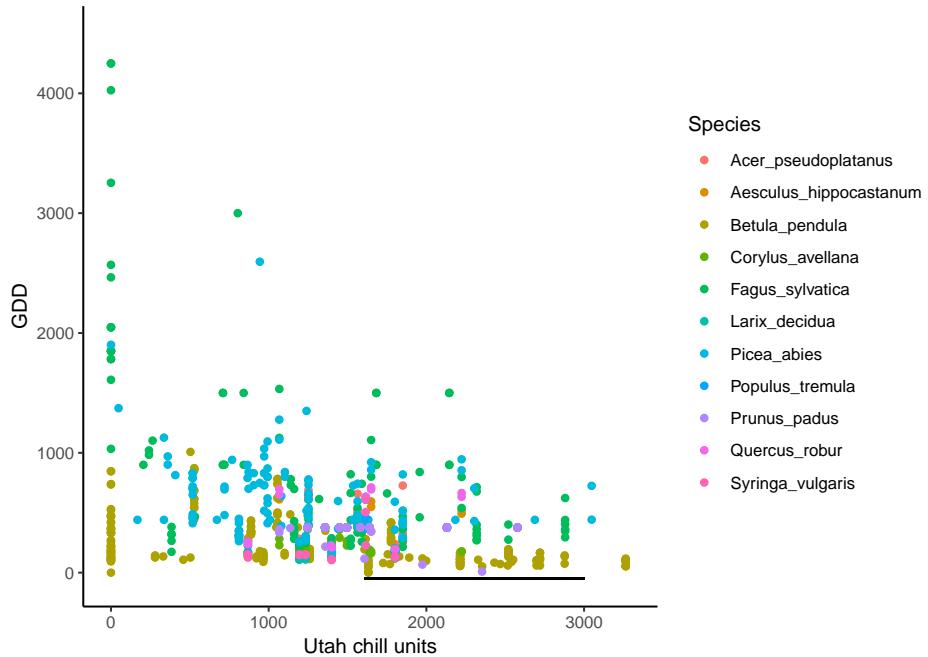


Figure 8: GDD (growing degree days) versus chill units at the time of budburst from the OSPREE database for common species in the PEP 725 long-term phenological database. The black line shows the range of chilling (10-90% quantiles) accumulated from 1 September to 1 March for 45 sites for *Betula pendula* (see also *Understanding declines in temperature sensitivity in European long-term data*). We calculated GDD here as the average daily forcing temperature multiplied by days to budburst.