

Phylogenetic estimates of species-level phenology improve ecological forecasting

June 3, 2023

Authors:

Ignacio Morales-Castilla,¹ T. J. Davies,^{2,3} Geoffrey Legault,³ D. M. Buonaiuto,^{4,5,6} Catherine J. Chamberlain,^{4,5,7} Ailene K. Ettinger,^{5,8} Mira Garner,³ Faith A. M. Jones,^{3,10} Deirdre Loughnan,³ William D. Pearse,¹¹ Darwin S. Sodhi,³ & E. M. Wolkovich^{3,4,5}

Author affiliations:

¹GloCEE - Global Change Ecology and Evolution Group, Department of Life Sciences, University of Alcalá, Alcalá de Henares, Spain

²Botany, Faculty of Sciences, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada

³Forest & Conservation Sciences, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada

⁴Organismic & Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, Massachusetts, USA

⁵Arnold Arboretum of Harvard University, 1300 Centre Street, Boston, Massachusetts, USA

⁶Department of Environmental Conservation, University of Massachusetts-Amherst, 160 Holdsworth Way, Amherst, MA, USA

⁷The Nature Conservancy, 334 Blackwell St Ste 300, Durham, NC, USA

⁸The Nature Conservancy of Washington, 74 Wall Street, Seattle, WA USA

¹⁰Department of Wildlife, Fish and Environmental Studies, Swedish University of Agricultural Sciences, 901 83 Umeå, Sweden

¹¹Department of Life Sciences, Imperial College London, Silwood Park, Ascot, Berkshire, SL5 7PY, UK

*Corresponding author: ignacio.moralesc@uah.es

Abstract

Adaptation to ongoing climate change hinges on accurate ecological forecasting to predict shifts in key ecosystem services, such as carbon storage and biodiversity maintenance. Current forecasts, however, have generally failed to capture the important variability in biological responses, especially observed across species. Here, using novel Bayesian hierarchical phylogenetic models, we overcome this challenge to estimate species-level responses to two major environmental cues of spring phenology: temperature and daylength. We find that variation is greater across species than across cues, suggesting the current focus on identifying which cue is the strongest may be less important than predicting how each species responds to a combination of cues. Further, our model provides insights on how evolutionary history has shaped responses to cues, suggesting stronger evolutionary dynamics in responses to temperature—especially cool winter temperatures—than daylength. Our approach provides a major advance in ecological forecasting, with implications for predicting the impacts of climate change and other anthropogenic forces on species interactions and ecosystem functioning.

¹ Introduction

² The biological impacts of climate change will have major implications for ecosystem functioning
³ and stability. With rising global temperatures many species have shifted their geographic distri-
⁴ butions northward in space and recurring life-history events—their phenology—earlier in time
⁵ (IPCC, 2014; Parmesan and Yohe, 2003), against a background of high variability. These shifts
⁶ have cascading consequences on many ecosystem services including carbon storage, making both
⁷ mitigation and human adaptation to future warming dependent on accurate ecological forecasts
⁸ (Richardson et al., 2013).

⁹ While ecological forecasting has improved over recent years (Dietze, 2017; Lewis et al., 2022), it
¹⁰ remains a challenge to reproduce the high variability observed in responses (IPCC, 2014). Some
¹¹ of this variability results from the complexity of climate change itself, including regional and
¹² seasonal variation in warming that underlies average trends alongside shifts in other climate axes
¹³ (e.g. precipitation). Much of it, however, is driven by species-specific variation, reflecting evolved
¹⁴ differences in species' sensitivities to underlying environmental cues and their interactions, which
¹⁵ we know for only a few well-studied species (Chuine and Regniere, 2017; Ettinger et al., 2020).
¹⁶ In the absence of detailed data on individual species, species groupings (e.g., functional groups)
¹⁷ have improved ecosystem models (Moorcroft et al., 2001; Griffith et al., 2020), but still capture
¹⁸ only a fraction of the important variability (Fuccillo Battle et al., 2022). Improving forecasts,
¹⁹ will require models that accurately predict species-level differences in responses to complex
²⁰ environmental change.

²¹ Recent efforts that have attempted to model species-specific responses to the environment (Diez
²² et al., 2012) are often confounded by data availability—especially the common problem that data
²³ are often prevalent for some species and sparse across others. The rise of Bayesian hierarchical
²⁴ models can allow inference across species in such cases. However, underlying most hierarchical
²⁵ models is an implicit assumption that species are exchangeable (Gelman and Hill, 2006), and they
²⁶ thus partially pool ('shrink') towards estimates for species with the most data and least variable
²⁷ responses, making inference at the species-level unreliable (Ettinger et al., 2020). More reliable
²⁸ estimates of species-level responses would allow us to better incorporate species differences into
²⁹ models of ecosystem change.

³⁰ Including the evolutionary history of species relationships in models of species responses could
³¹ at once provide more robust species-level estimates than current approaches and a better un-
³² derstanding of the evolutionary constraints that might limit future adaptation to change. For
³³ example, strong phylogenetic niche conservatism (Wiens et al., 2010) could potentially inhibit
³⁴ adaptive responses by drawing species back to an evolutionary conserved optimum, which is sub-
³⁵ optimal under new conditions. Whereas, traditionally, incorporating such evolutionary history
³⁶ is seen as necessary either as a statistical correction or to better understand species' evolutionary
³⁷ history, the use of such phylogenetic information should also improve model fitting and forecasts
³⁸ (Freckleton et al., 2002).

39 Research using long-term observational data has especially highlighted the role that evolutionary
40 history may play in structuring plant phenological responses—which are critical to accurate
41 forecasts of carbon storage. Phylogenetic signal in plant phenology (Kochmer and Handel, 1986;
42 Willis et al., 2008; Davies et al., 2013) suggests species responses to cues have diverged over
43 macro-evolutionary timescales, helping explain species present day differences. However, approaches
44 using traditional phylogenetic comparative methods, have produced conflicting results
45 (e.g., Willis et al., 2008; CaraDonna and Inouye, 2014; Yang et al., 2021). In addition, evidence
46 for phylogenetic conservatism of phenological responses appears to depend on method
47 and species, even varying between sites with overlapping species sets (e.g., Rafferty and Nabity,
48 2017), which violates the fundamental idea of shared evolutionary history (the common ancestor
49 of two sets of species cannot possess two separate evolutionary histories for the same trait). A
50 first challenge is thus how to better integrate evolutionary history into multi-species models of
51 plant phenological responses.

52 Generating robust ecological forecasts requires addressing a second major hurdle—underlying
53 environmental cues that are complex and interacting. Decades of research have informed our
54 understanding of how species use environmental cues to time their phenotypic responses with
55 the temporal distribution of key resources and to avoid periods of high abiotic or biotic stress
56 (Larcher, 1980; Bonamour et al., 2019). Commonly, however, responses to environmental cues,
57 and their evolution, are studied individually, linking a given phenotypic response to a single cue,
58 for example, time of leafout responding to summed heat during early spring (Davies et al., 2013).
59 These efforts fail to capture the more likely scenario for most phenotypic traits in which multiple
60 cues interacting along evolutionary history have shaped species' responses (Ackerly, 2009). For
61 many plant species, phenological events are determined by a combination of temperature and
62 light (Chuine and Regnieri, 2017), with additional factors (e.g., other cues—like humidity, or
63 species physiology—vasculature or SLA) likely further mediating species responses. Although
64 these mediating factors are often less well understood (Chuine and Regnieri, 2017), they can be
65 accounted for in models either as latent processes or by allowing non-stationarity in responses
66 across species (Davies et al., 2019).

67 Spring plant phenology may represent our best opportunity to improve forecasts of species'
68 responses to interacting environmental cues. Beyond being the most studied biological impact
69 of climate change, the primary cue system is well established (Chuine and Regnieri, 2017),
70 especially for temperate woody species where phenology is generally thought to be determined
71 by two components of temperature—chilling (cool temperatures during dormancy period over
72 winter) and forcing (warm temperatures, generally in the spring)—and photoperiod (Ettinger
73 et al., 2021). Plant phenology is also one of few phenotypic traits with extensive experimental
74 data on responses to multiple environmental cues across species. Recent multi-species analyses
75 considering forcing, chilling and photoperiod have shown that chilling and forcing together often
76 determine complex non-linear responses to warming, but cannot forecast beyond several well-
77 studied species (Ettinger et al., 2020).

78 Here we present a novel Bayesian framework that extends upon phylogenetic mixed models
79 (Housworth et al., 2004) to examine how chilling, forcing (both metrics of temperature) and
80 photoperiod together determine plant phenology. By allowing non-stationarity in species re-
81 sponds across the phylogeny (Davies et al., 2019), our model departs from most previous work
82 and assumptions of traditional phylogenetic comparative methods (e.g. Freckleton et al., 2002;
83 Ives and Helmus, 2011; Hadfield, 2010), and moves towards integrating evolutionary history in
84 models of phenological responses to environmental change. To understand how evolution has
85 shaped the cues underlying shifting phenology with climate change (Uyeda et al., 2017), we ex-
86 plicitly incorporate phylogenetic structure across model intercepts and slopes (that is, allowing
87 a separate model of evolutionary history for chilling, forcing and photoperiod, see full Methods
88 & Materials for a complete description).

89 We illustrate our method with an unprecedented dataset on phenological responses to environ-
90 mental cues (chilling, forcing and photoperiod) determined experimentally for 191 deciduous
91 woody species (by far the most studied group of species in phenology experiments, see Ettinger
92 et al., 2020), in an updated version of the Observed Spring Phenology Responses in Experi-
93 mental Environments (OSPRE) database (Wolkovich et al., 2019). These data combined with the
94 Smith and Brown (2018) megatree adjusted to our species and our modeling approach allows
95 us to address the common question of which cue has the largest effect on budburst and at the
96 same time provide robust estimates of how cues vary across species. Using spring phenology, we
97 identify historical regime shifts (Uyeda et al., 2017) in phenological responses, and highlight how
98 our approach could advance forecasting of other critical responses to ongoing global change.

99 Results & Discussion

100 Most species respond to all three primary cues—forcing, chilling, and photoperiod (Fig. 1,
101 Supporting Table S1)—with responses to chilling approximately five-fold greater than to pho-
102 toperiod (phenological advances of 6.9 days per standardized unit vs 1.2 days, for chilling and
103 photoperiod, respectively; see Table S1). We estimated lower average responses to temperature
104 compared to a model without phylogeny (model slopes for forcing and chilling decreased by 18%
105 and 22%, respectively); responses to chilling and forcing were also more similar when including
106 phylogeny (though chilling was still greater: 6.9 vs. 6.1 per standard unit), which contrasts with
107 previous results suggesting chilling responses are much greater than forcing (Laube et al., 2014;
108 Ettinger et al., 2020).

109 These average estimates, however, fail to capture the large differences in species' responses to
110 both chilling and forcing (Fig. 1, Supporting Table S3). By allowing species responses to
111 vary, based on a model including their shared evolutionary history, we found species differences
112 dwarfed the mean differences between cues, especially temperature cues (Fig. 1). The largest cue
113 in magnitude—chilling—varied 24-fold between species, while variation to forcing varied 7-fold.
114 This variation indicates large differences between chilling and forcing occur at the species-level

and not across species (as previously suggested Laube et al., 2014; Ettinger et al., 2020). These results highlight why robust phenological forecasts must account for both the complexity of multiple cues and species-level variation in responses to them.

118 *Differences across clades & cues*

The large differences across species produced striking differences between clades. For example, several groups—oaks and beeches (Fagaceae), elms (Ulmaceae) and buckthorns (Rhamnaceae)—are highly sensitive to chilling while others—rhododendrons (Ericaceae), butterfly bushes (Scrophulariaceae) and spindles (Celastraceae)—show little to no response to chilling (Fig. 1a). Similar clade-level variation was observed for forcing, where some of these clades—e.g., Ericaceae, Rhamnaceae, Ulmaceae, or Fagaceae—were particularly sensitive (advancing their budburst more than 10 days per standardized unit of forcing) and others, such as the Sapindaceae, Cornaceae or Juglandaceae, show little response (Fig. 1b).

Some species responded strongly to both temperature cues, which could suggest the existence of syndromes where the genetic basis for responses to one cue—e.g. forcing—has been selected for alongside responses to another cue—e.g. chilling. This could occur if selection operates jointly on responses to both cues; for example, if sensitivity to multiple cues provides greater insurance against leafing out before the last frost (Bonamour et al., 2019; Wolkovich and Donahue, 2021). Additionally, linkage or pleiotropism among loci associated with different cues (Nakagawa et al., 2005) could induce these correlations. However, the correlation in species responses across cues was generally weak ($r = 0.31$; between forcing and chilling) and some genera, such as *Tilia* and *Rhododendron* (Ericaceae), displayed strong responses to forcing but weak responses to chilling, while others, such as *Acer* (Sapindaceae), show moderately strong responses to chilling but weak responses to forcing (Fig. 1). Species sensitivity to one cue, thus, does not constrain sensitivity to another cue, and it seems selection can operate independently on responses to different cues (Bonamour et al., 2019).

In contrast to temperature cues (chilling and forcing), species-level responses to photoperiod were almost uniform across species. This consistency provides novel insight on a large debate over the prevalence of photoperiod cues in temperate trees, where previous experiments (Basler and Körner, 2012; Zohner et al., 2016) and models (e.g., Hunter and Lechowicz, 1992; Schaber and Badeck, 2003) suggested important variability across species that may constrain the responses of certain species to warming (Way and Montgomery, 2015). Our results indicate variability is limited to a handful of species in Fagaceae, which have been particularly well studied, especially *Fagus sylvatica* (e.g., Basler and Körner, 2012; Zohner et al., 2016; Kramer et al., 2017). As *Fagus sylvatica* is nearly five times more sensitive to photoperiod than most other measured tree species, our results caution against using it to draw inferences of photoperiod responses more widely. These same few species are also where most evidence of local adaptation in photoperiod cues for spring phenology comes from (e.g., Kramer et al., 2017), in contrast with common garden studies of other species, which found little evidence of local adaptation in spring (but not

fall) phenology (Aitken and Bemmels, 2016). The uniformity of response to photoperiod in our results supports this latter view of generally low local adaptation in photoperiod cues for spring phenology (i.e., if local adaptation were high in photoperiod cues we would have expected more variability across species).

157 *Phylogenetic structure of phenological cues*

158 Variation—or lack thereof—in cues across species and clades provides possible insights into
159 the evolution of cues across the phylogeny. While responses to each cue were phylogenetically
160 structured, with closely related species exhibiting more similar sensitivities than distantly related
161 species, the strength of phylogenetic conservatism in response differed between cues (Fig. 2).
162 Responses to temperature (forcing and chilling) were moderately structured ($\lambda = 0.65$ and
163 $\lambda = 0.54$, for forcing and chilling, respectively). Phylogenetic structure in species responses to
164 photoperiod was comparatively weak ($\lambda = 0.4$) (see Fig. 2, Table S1).

165 Differences among species in their temperature responses represent shifts in the slope of the
166 relationship between the observed phenology and the cue. The observed phylogenetic structure
167 in temperature responses (forcing and chilling) would be consistent with an interaction with a
168 latent trait that moderates responses, and which also covaries with phylogeny (Davies et al.,
169 2019). This fits fundamentally with the idea that early-season phenology plays a critical role in
170 shaping species temporal niches (Gotelli and Graves, 1996) and thus should covary with a suite
171 of life-history traits, including whether species are early-active with rapid return on investment
172 traits, or start later in the season and have traits associated with higher competitive abilities
173 (e.g., Grime, 1977; Wolkovich and Donahue, 2021).

174 Weak phylogenetic signal in photoperiod sensitivity (Fig. 2) might seem at odds with the uniformity
175 of species response—i.e., there is very little variation in the responses to photoperiod across
176 species. However, somewhat counterintuitively, both uniform and random responses can manifest
177 as low phylogenetic signal when indexed by Brownian motion expectations (Wiens et al.,
178 2010). While rapid local adaptation within species might erase the phylogenetic structure in
179 photoperiod responses, it does not agree with the uniformity we find in species' responses. How-
180 ever, if responses to photoperiod evolved early in plants, as seems likely (Serrano-Bueno et al.,
181 2017), and subsequent selection on photoperiod sensitivity was constrained by stabilizing selec-
182 tion operating on other life-history attributes sensitive to photoperiod (e.g., Rinne et al., 1994;
183 Wilczek et al., 2014; Azeez and Sane, 2015), we would predict both low interspecific variation
184 and weak phylogenetic signal in responses, matching observations. This latter interpretation is
185 also consistent with our estimates of lower σ for photoperiod responses (Fig. 2). Here, as in
186 more traditional phylogenetic comparative methods, σ represents the rate of evolution, and thus
187 our results suggest photoperiod responses are also evolving slower than temperature responses
188 (see Fig. S8).

189 Phylogenetic conservatism (high λ) and slow evolutionary rates (low σ) in traits has sometimes
190 been interpreted as indicative of evolutionary constraints to adaptive change (Wiens et al., 2010;

191 Bennett et al., 2021). If this were the case, we might then conclude that species where responses
192 are dominated by forcing cues might be more vulnerable to future warming because phylogenetic
193 conservatism (λ) in forcing is higher compared to other cues and its evolutionary rate (σ) is lower
194 than that estimated for chilling. This is misleading, however, as estimates of λ are independent
195 from the rate of evolution, and macroevolutionary rates are estimated on phylogenetic trees that
196 integrate across millions of years of evolutionary history, and thus do not necessarily inform us of
197 maximum possible rates of evolution over much shorter timescales. Indeed, there is accumulating
198 evidence for rapid evolution to shifting climates (Bradshaw and Holzapfel, 2006; Franks et al.,
199 2014). Our estimates are thus more useful in providing unique insights into the evolutionary
200 history of phenological cues, and emphasize the critical importance of incorporating species-level
201 differences in ecological forecasts

202 *Forecasting species-level responses*

203 Our results highlight that species-level variability can be extremely high—when properly esti-
204 mated. Our approach, which partially pooled species responses based on their shared evolution-
205 ary history, estimated substantially higher variation across species compared with more widely
206 used hierarchical models. This was especially noticeable in temperature responses (for chilling
207 variance across species means, $var(\beta_{chill,j}$ from eqn. 2), was estimated as 23.55 in the phyloge-
208 netic model, versus 17.47 in the non-phylogenetic model; variance across means, $var(\beta_{force,j})$:
209 8.75 compared to 5.01) while photoperiod, which had low phylogenetic structure, was more
210 similar across approaches (variance across means, $var(\beta_{photo,j})$: 0.83 compared to 0.64).

211 The consequences of including shared evolutionary history in forecasting are most apparent for
212 poorly sampled species nested within more well-sampled clades. For example, forecasts for *Acer*
213 *campestre*, which has only 6 observations, shift by up to 35% in our phylogenetically informed
214 model compared to one without phylogeny (Fig. 4). In contrast, forecasts for *Betula pendula*,
215 which is one of the most sampled species, are nearly identical across models (Fig. 4). This occurs
216 because cue estimates for *Acer campestre* in the phylogenetically informed model are strongly
217 influenced by other *Acer* species, which diverge from other clades. In the non-phylogenetically
218 informed model all species are equally exchangeable and thus *Acer campestre* is pulled strongly
219 towards well-sampled species, such as *Betula pendula* ($n = 311$), leading to forecasted shifts that
220 are more similar across all species (Fig. 4).

221 The increase in variability across species in our model with phylogenetic structure also decreased
222 the uncertainty in estimates for each individual species temperature responses (Fig. S3). Thus,
223 traditional (non-phylogenetically informed) approaches that partially pool across species (most
224 hierarchical models in ecology, e.g., Flynn and Wolkovich, 2018; Ettinger et al., 2020) may also
225 lead to less precise predictions and forecasts of phenology for individual species, although overall
226 model accuracy might still appear reasonable (see Fig. S4).

227 The contrasts between temperature and photoperiod responses—in both their variability across
228 species and phylogenetic structure—have important implications for generating multi-species

229 forecasts. Notably, responses to photoperiod appear weaker, more uniform across species, and
230 less phylogenetically constrained compared to temperature. For temperature responses, the
231 large variability among species makes predicting species-level responses challenging, but the
232 phylogenetic structure in responses lets us borrow information from close relatives to improve
233 our predictions. However, given that Brownian motion (our assumed model of evolution) is an
234 extremely noisy process, we recommend imputation only for missing taxa that are closely related
235 to other well sampled species or clades (Molina-Venegas et al., 2018, 2023).

236 While we focused on spring phenology here, our approach suggests a path forward for more
237 general forecasting of species-level climate change responses. Our results show how including
238 the phylogenetic relationship of species in a mechanistic model of underlying cues can overcome
239 major limitations of most current hierarchical models—correcting biased model estimates, prop-
240 erly estimating the full variability across species and reducing uncertainty around individual
241 species estimates—while at once providing insight into the evolutionary history of biological
242 responses. Using this approach improved forecasts of phenological responses to climate change
243 and could help anticipate impacts on critical ecosystem services from species-level shifts and
244 thus aid mitigation and human adaption to warming.

245 **Methods & Materials**

246 **Phenological and Phylogenetic Data**

247 *Phenological data:* To estimate phenological responses to chilling, forcing and photoperiod we
248 used data from phenological experiments in controlled environments of temperate woody species,
249 brought together in the Observed Spring Phenology Responses in Experimental Environments
250 (OSPREE) database. In July 2019, we updated an earlier version of this database (Wolkovich
251 et al., 2019) by reviewing all papers found through searching ISI Web of Science and Google
252 Scholar with the following terms:

- 253 1. TOPIC = (budburst OR leaf-out) AND (photoperiod OR daylength) AND temperature*,
254 which yielded 623 publications
- 255 2. TOPIC = (budburst OR leaf-out) AND dorman*, which yielded 270 publications

256 We scraped data from all papers of woody species that tested for photoperiod and/or tempera-
257 ture effects on budburst, leafout, or flowering, resulting in data from 155 experiments across 97
258 papers in the updated database. Ettinger et al. (2020) used a portion (72 experiments across
259 49 papers) of the earlier OSPREE database and provides extensive methods on the database
260 creation and cleaning. For our analysis here, we focused on angiosperms (as gymnosperms are
261 very poorly represented in spring phenology experiments), and included all budburst experi-
262 ments where we could quantify chilling, forcing and photoperiod levels, resulting in 44 studies
263 from 33 papers.

264 Across experiments chilling treatments were often fully or partially applied in the field, thus we
265 estimated field chilling ourselves in Utah units with the `chillR` package combined with daily
266 temperature data (converted to hourly) from both European (E-OBS, version 16, calculating
267 the average of minimum and maximum daily temps, Cornes et al., 2018) and North American
268 (v3, Sheffield et al., 2006) gridded climate datasets. Ettinger et al. (2020) provides additional
269 details on these calculations (however, to have climate data through all study years given the
270 updated OSPREE database, we used a different climate dataset here for North America).

271 For a phylogenetic tree, we pruned the phylogenetic megatree for seed plants (Smith and Brown,
272 2018) to extract a subset of the phylogenetic tree containing only the angiosperm species in the
273 OSPREE dataset, then added species that were not present in the megatree as polytomies at
274 the generic level (using the function ‘congeneric.merge’ in Pearse et al., 2015), with a branch
275 length of zero. Our pruned tree had 8 Polytomies affecting 46 out of 191 species. To test for
276 the ability of polytomies to bias our results we run sensitivity analyses excluding these species
277 from models (see Supporting Information).

278

279 **Bayesian hierarchical phylogenetic model**

280 Commonly used phylogenetic regression methods today (e.g., phylogenetic generalized least
281 squares models, PGLS, Freckleton et al. (2002); phylogenetic mixed models, PMM, Hous-
282 worth et al. (2004)) were originally conceived as statistical corrections for phylogenetic non-
283 independence across observations—generally species—thus allowing multi-species studies to meet
284 the assumptions of linear regression (Freckleton et al., 2002). These corrections incorporated
285 phylogenetic structure by estimating the magnitude of a transformation of a variance-covariance
286 (VCV) matrix whose elements were derived from the amount of evolutionary history (branch
287 lengths) shared between species on a phylogeny. The most commonly used transformation was
288 Pagel’s λ —a multiplier of the off-diagonal elements—where estimates of $\lambda = 1$ essentially left the
289 VCV untransformed and suggested that the residuals of the regression had phylogenetic signal.
290 Consistent with Brownian motion, estimates of $\lambda = 0$ suggested no phylogenetic signal. Because
291 the original aim of these methods was to correct for statistical bias introduced by shared evo-
292 lutionary history among species, the underlying assumption of phylogenetic regressions is that
293 phylogenetic relatedness would only affect either model residuals (in PGLS approaches, Freck-
294 leton et al., 2002) or the model intercepts (e.g., in many PMM approaches, Housworth et al.,
295 2004).

296 Because our aim is to understand how evolution may have imprinted biological responses to
297 multiple interactive cues, our approach expands the above methods by explicitly incorporating
298 phylogenetic structure across model intercepts and slopes. Doing so allows explicitly estimating
299 the amount of phylogenetic relatedness in species’ sensitivities to each cue, when these sensitiv-
300 ities are modelled in a multi-predictor regression setting.

301 For each observation i of species j , we assumed that the timing of phenological events were
302 generated from the following sampling distribution:

$$y_{i,j} \sim \mathcal{N}(\mu_j, \sigma_e^2) \quad (1)$$

where

$$\mu_j = \alpha_j + \beta_{chill,j} X_{chill} + \beta_{force,j} X_{force} + \beta_{photo,j} X_{photo} \quad (2)$$

303 and σ_e^2 represents random error unrelated to the phylogeny.

304 Predictors X_{chill} , X_{force} , X_{photo} are standardized chilling, forcing, and photoperiod, and their
305 effects on the phenology of species j are determined by parameters $\beta_{chill,j}$, $\beta_{force,j}$, $\beta_{photo,j}$,
306 representing species’ responses (or sensitivities) to each of the cues. These responses, including
307 the species-specific intercept α_j , are elements of the following normal random vectors:

$$\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_n]^T \text{ such that } \boldsymbol{\alpha} \sim \mathcal{N}(\mu_{\boldsymbol{\alpha}}, \boldsymbol{\Sigma}_{\boldsymbol{\alpha}}) \quad (3)$$

$$\begin{aligned}\boldsymbol{\beta}_{chill} &= [\beta_{chill,1}, \dots, \beta_{chill,n}]^T \text{ such that } \boldsymbol{\beta}_{chill} \sim \mathcal{N}(\mu_{\boldsymbol{\beta}_{chill}}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}_{chill}}) \\ \boldsymbol{\beta}_{force} &= [\beta_{force,1}, \dots, \beta_{force,n}]^T \text{ such that } \boldsymbol{\beta}_{force} \sim \mathcal{N}(\mu_{\boldsymbol{\beta}_{force}}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}_{force}}) \\ \boldsymbol{\beta}_{photo} &= [\beta_{photo,1}, \dots, \beta_{photo,n}]^T \text{ such that } \boldsymbol{\beta}_{photo} \sim \mathcal{N}(\mu_{\boldsymbol{\beta}_{photo}}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}_{photo}})\end{aligned}$$

where the means of the multivariate normal distributions are root trait values (i.e., values of cue responses prior to evolving across a phylogenetic tree) and $\boldsymbol{\Sigma}_i$ are $n \times n$ phylogenetic variance-covariance matrices of the form:

$$\left[\begin{array}{cccc} \sigma_i^2 & \lambda_i \times \sigma_i \times \rho_{12} & \dots & \lambda_i \times \sigma_i \times \rho_{1n} \\ \lambda_i \times \sigma_i \times \rho_{21} & \sigma_i^2 & \dots & \lambda_i \times \sigma_i \times \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_i \times \sigma_i \times \rho_{n1} & \lambda_i \times \sigma_i \times \rho_{n2} & \dots & \sigma_i^2 \end{array} \right] \quad (4)$$

308 where σ_i^2 is the rate of evolution across a tree for trait i (here assumed to be constant along
 309 all branches), λ_i scales branch lengths and therefore is a measure of the phylogenetic signal or
 310 extent of phylogenetic relatedness on each model parameter (i.e., α_j , $\beta_{force,j}$, $\beta_{force,j}$, $\beta_{photo,j}$),
 311 and ρ_{xy} is the phylogenetic correlation between species x and y , or the fraction of the tree shared
 312 by the two species.

313 The above specification is equivalent to writing equation 2 in terms of root trait values and
 314 residuals, such that:

$$\mu_j = \mu_{\alpha} + \mu_{\beta_{chill}} X_{chill} + \mu_{\beta_{force}} X_{force} + \mu_{\beta_{photo}} X_{photo} + e_{\alpha_j} + e_{\beta_{force,j}} + e_{\beta_{chill,j}} + e_{\beta_{photo,j}} \quad (5)$$

315 where the residual phylogenetic error terms (e.g., e_{α_j}) are elements of normal random vec-
 316 tors from multivariate normal distributions centered on 0 with the same phylogenetic variance-
 317 covariance matrices as in equation 4. Model code, including priors used here, are given in the
 318 supplement. We fit all models to our data using RStan using 4 chains of 4,000 iterations with
 319 a warmup of 2,000 each (resulting 8,000 posterior samples), and assessed fit via \hat{R} near 1 and
 320 adequate effective sample sizes (see Tables S1-S2).

321 References

322 Ackerly, D. 2009. Conservatism and diversification of plant functional traits: Evolutionary rates
 323 versus phylogenetic signal. Proceedings of the National Academy of Sciences of the United
 324 States of America 106:19699–19706.

- 325 Aitken, S. N., and J. B. Bemmels. 2016. Time to get moving: assisted gene flow of forest trees.
326 *Evol Appl* 9:271–90.
- 327 Azeez, A., and A. P. Sane. 2015. Photoperiodic growth control in perennial trees. *Plant signaling
328 & behavior* 10:e1087631.
- 329 Basler, D., and C. Körner. 2012. Photoperiod sensitivity of bud burst in 14 temperate forest
330 tree species. *Agricultural and Forest Meteorology* 165:73–81.
- 331 Bennett, J. M., J. Sunday, P. Calosi, F. Villalobos, B. Martínez, R. Molina-Venegas, M. B.
332 Araújo, A. C. Algar, S. Clusella-Trullas, B. A. Hawkins, et al. 2021. The evolution of critical
333 thermal limits of life on earth. *Nature communications* 12:1–9.
- 334 Bonamour, S., L. M. Chevin, A. Charmantier, and C. Teplitsky. 2019. Phenotypic plasticity in
335 response to climate change: the importance of cue variation. *Philosophical Transactions of
336 the Royal Society B-Biological Sciences* 374.
- 337 Bradshaw, W. E., and C. M. Holzapfel. 2006. Evolutionary response to rapid climate change.
338 *Science* 312:1477–1478.
- 339 CaraDonna, P. J., and D. W. Inouye. 2014. Phenological responses to climate change do not
340 exhibit phylogenetic signal in a subalpine plant community. *Ecology* 96:355–361.
- 341 Chuine, I., and J. Regniere. 2017. Process-based models of phenology for plants and animals.
342 *Annual Review of Ecology, Evolution, and Systematics* 48:159–182.
- 343 Cornes, R. C., G. van der Schrier, E. J. van den Besselaar, and P. D. Jones. 2018. An ensemble
344 version of the E-OBS temperature and precipitation data sets. *Journal of Geophysical
345 Research: Atmospheres* 123:9391–9409.
- 346 Davies, T., E. Wolkovich, N. Kraft, N. Salamin, and S. E. Travers. 2013. Phylogenetic conservatism
347 in plant phenology. *Journal of Ecology* 101:1520–1530.
- 348 Davies, T. J., J. Regetz, E. M. Wolkovich, and B. J. McGill. 2019. Phylogenetically weighted
349 regression: A method for modelling non-stationarity on evolutionary trees. *Global ecology
350 and biogeography* 28:275–285.
- 351 Dietze, M. 2017. Ecological forecasting. *In Ecological Forecasting*. Princeton University Press.
- 352 Diez, J. M., I. Ibáñez, A. J. Miller-Rushing, S. J. Mazer, T. M. Crimmins, M. A. Crimmins,
353 C. D. Bertelsen, and D. W. Inouye. 2012. Forecasting phenology: from species variability to
354 community patterns. *Ecology letters* 15:545–553.
- 355 Ettinger, A., C. Chamberlain, I. Morales-Castilla, D. Buonaiuto, D. Flynn, T. Savas, J. Samaha,
356 and E. Wolkovich. 2020. Winter temperatures predominate in spring phenological responses
357 to warming. *Nature Climate Change* pages 1–6.

- 358 Ettinger, A. K., D. M. Buonaiuto, C. J. Chamberlain, I. Morales-Castilla, and E. M. Wolkovich.
359 2021. Spatial and temporal shifts in photoperiod with climate change. *New Phytologist*
360 230:462–474.
- 361 Flynn, D. F. B., and E. M. Wolkovich. 2018. Temperature and photoperiod drive spring phe-
362 nology across all species in a temperate forest community. *New Phytologist* 219:1353–1362.
- 363 Franks, S. J., J. J. Weber, and S. N. Aitken. 2014. Evolutionary and plastic responses to climate
364 change in terrestrial plant populations. *Evolutionary Applications* 7:123–139.
- 365 Freckleton, R. P., P. H. Harvey, and M. Pagel. 2002. Phylogenetic analysis and comparative
366 data: a test and review of evidence. *The American Naturalist* 160:712–726.
- 367 Fuccillo Battle, K., A. Duhon, C. R. Vispo, T. M. Crimmins, T. N. Rosenstiel, L. L. Armstrong-
368 Davies, and C. E. de Rivera. 2022. Citizen science across two centuries reveals phenological
369 change among plant species and functional groups in the northeastern us. *Journal of Ecology*
370 110:1757–1774.
- 371 Gelman, A., and J. Hill. 2006. Data analysis using regression and multilevel/hierarchical models.
372 Cambridge University Press.
- 373 Gotelli, N. J., and G. R. Graves. 1996. The temporal niche. Pages 95–111 in N. J. Gotelli and
374 G. R. Graves, eds. Null Models in Ecology. Smithsonian Institution, Washington, DC.
- 375 Griffith, D. M., C. P. Osborne, E. J. Edwards, S. Bachle, D. J. Beerling, W. J. Bond, T. J. Gal-
376 laher, B. R. Helliker, C. E. R. Lehmann, L. Leatherman, J. B. Nippert, S. Pau, F. Qiu, W. J.
377 Riley, M. D. Smith, C. A. E. Stromberg, L. Taylor, M. Ungerer, and C. J. Still. 2020. Lineage-
378 based functional types: characterising functional diversity to enhance the representation of
379 ecological behaviour in land surface models. *New Phytologist* 228:15–23.
- 380 Grime, J. P. 1977. Evidence for existence of 3 primary strategies in plants and its relevance to
381 ecological and evolutionary theory. *American Naturalist* 111:1169–1194.
- 382 Hadfield, J. D. 2010. Mcmc methods for multi-response generalized linear mixed models: the
383 mcmcglmm r package. *Journal of statistical software* 33:1–22.
- 384 Housworth, E. A., E. P. Martins, and M. Lynch. 2004. The phylogenetic mixed model. *The*
385 *American Naturalist* 163:84–96.
- 386 Hunter, A. F., and M. J. Lechowicz. 1992. Predicting the timing of budburst in temperate trees.
387 *Journal of Applied Ecology* 29:597–604.
- 388 IPCC. 2014. Climate Change 2014: Impacts, Adaptation, and Vulnerability. Cambridge Uni-
389 versity Press, Cambridge, United Kingdom and New York, NY, USA.

- 390 Ives, A. R., and M. R. Helmus. 2011. Generalized linear mixed models for phylogenetic analyses
391 of community structure. *Ecological Monographs* 81:511–525.
- 392 Kochmer, J. P., and S. N. Handel. 1986. Constraints and competition in the evolution of flowering
393 phenology. *Ecological monographs* 56:303–325.
- 394 Kramer, K., A. Ducousoo, D. Gomory, J. K. Hansen, L. Ionita, M. Liesebach, A. Lorent,
395 S. Schueler, M. Sulkowska, S. de Vries, and G. von Wuehlisch. 2017. Chilling and forcing
396 requirements for foliage bud burst of European beech (*Fagus sylvatica* L.) differ between
397 provenances and are phenotypically plastic. *Agricultural and Forest Meteorology* 234:172–181.
- 398 Larcher, W. 1980. *Plant Physiological Ecology*. Springer-Verlag.
- 399 Laube, J., T. H. Sparks, N. Estrella, J. Höfler, D. P. Ankerst, and A. Menzel. 2014. Chilling
400 outweighs photoperiod in preventing precocious spring development. *Global Change Biology*
401 20:170–182.
- 402 Lewis, A. S., C. R. Rollinson, A. J. Allyn, J. Ashander, S. Brodie, C. B. Brookson, E. Collins,
403 M. C. Dietze, A. S. Gallinat, N. Juvigny-Khenafou, et al. 2022. The power of forecasts to
404 advance ecological theory. *Methods in Ecology and Evolution* .
- 405 Molina-Venegas, R., I. Morales-Castilla, and M. Á. Rodríguez. 2023. Unreliable prediction of
406 b-vitamin source species. *Nature Plants* 9:31–33.
- 407 Molina-Venegas, R., J. C. Moreno-Saiz, I. Castro Parga, T. J. Davies, P. R. Peres-Neto, and
408 M. Á. Rodríguez. 2018. Assessing among-lineage variability in phylogenetic imputation of
409 functional trait datasets. *Ecography* 41:1740–1749.
- 410 Moorcroft, P., G. Hurtt, and S. Pacala. 2001. A method for scaling vegetation dynamics: The
411 ecosystem demography model (ed). *Ecological Monographs* 71:557–585.
- 412 Nakagawa, H., J. Yamagishi, N. Miyamoto, M. Motoyama, M. Yano, and K. Nemoto. 2005.
413 Flowering response of rice to photoperiod and temperature: a qtl analysis using a phenological
414 model. *Theoretical and Applied Genetics* 110:778–786.
- 415 Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts
416 across natural systems. *Nature* 421:37.
- 417 Pearse, W. D., M. W. Cadotte, J. Cavender-Bares, A. R. Ives, C. M. Tucker, S. C. Walker,
418 and M. R. Helmus. 2015. Pez: Phylogenetics for the environmental sciences. *Bioinformatics*
419 31:2888–2890.
- 420 Rafferty, N. E., and P. D. Nabity. 2017. A global test for phylogenetic signal in shifts in flowering
421 time under climate change. *Journal of Ecology* 105:627–633.

- 422 Richardson, A. D., T. F. Keenan, M. Migliavacca, Y. Ryu, O. Sonnentag, and M. Toomey. 2013.
423 Climate change, phenology, and phenological control of vegetation feedbacks to the climate
424 system. Agricultural and Forest Meteorology 169:156–173.
- 425 Rinne, P., A. Saarelainen, and O. Junntila. 1994. Growth cessation and bud dormancy in
426 relation to ABA level in seedlings and coppice shoots of *Betula pubescens* as affected by a
427 short photoperiod, water stress and chilling. Physiologia Plantarum 90:451–458.
- 428 Schaber, J., and F. Badeck. 2003. Physiology-based phenology models for forest tree species in
429 germany. International Journal of Biometeorology 47:193–201. International Conference on
430 the Times they are a Changin, WAGENINGEN, NETHERLANDS, DEC 05-07, 2001.
- 431 Serrano-Bueno, G., F. J. Romero-Campero, E. Lucas-Reina, J. M. Romero, and F. Valverde.
432 2017. Evolution of photoperiod sensing in plants and algae. Current Opinion in Plant Biology
433 37:10–17.
- 434 Sheffield, J., G. Goteti, and E. F. Wood. 2006. Development of a 50-year high-resolution global
435 dataset of meteorological forcings for land surface modeling. Journal of Climate 19:3088–3111.
- 436 Smith, S. A., and J. W. Brown. 2018. Constructing a broadly inclusive seed plant phylogeny.
437 American journal of botany 105:302–314.
- 438 Uyeda, J. C., M. W. Pennell, E. T. Miller, R. Maia, and C. R. McClain. 2017. The evolution of
439 energetic scaling across the vertebrate tree of life. The American Naturalist 190:185–199.
- 440 Way, D. A., and R. A. Montgomery. 2015. Photoperiod constraints on tree phenology, perfor-
441 mance and migration in a warming world. Plant, Cell & Environment 38:1725–1736.
- 442 Wiens, J. J., D. D. Ackerly, A. P. Allen, B. L. Anacker, L. B. Buckley, H. V. Cornell, E. I.
443 Damschen, T. Jonathan Davies, J.-A. Grytnes, S. P. Harrison, et al. 2010. Niche conservatism
444 as an emerging principle in ecology and conservation biology. Ecology letters 13:1310–1324.
- 445 Wilczek, A. M., M. D. Cooper, T. M. Korves, and J. Schmitt. 2014. Lagging adaptation to
446 warming climate in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences
447 111:7906–7913.
- 448 Willis, C. G., B. Ruhfel, R. B. Primack, A. J. Miller-Rushing, and C. C. Davis. 2008. Phyloge-
449 netic patterns of species loss in thoreau’s woods are driven by climate change. Proceedings of
450 the National Academy of Sciences 105:17029–17033.
- 451 Wolkovich, E. M., and M. J. Donahue. 2021. How phenological tracking shapes species and
452 communities in non-stationary environments. Biological Reviews .
- 453 Wolkovich, E. M., A. K. Ettinger, D. Flynn, T. Savas, C. Chamberlain, D. Buonaiuto, and
454 J. Samaha. 2019. Observed Spring Phenology Responses in Experimental Environments (OS-
455 PREE). doi:10.5063/F1CZ35KB.

- 456 Yang, Z., Y. Du, M. Shen, N. Jiang, E. Liang, W. Zhu, Y. Wang, and W. Zhao. 2021. Phy-
457 logenetic conservatism in heat requirement of leaf-out phenology, rather than temperature
458 sensitivity, in tibetan plateau. Agricultural and Forest Meteorology 304.
- 459 Zohner, C. M., B. M. Benito, J. C. Svenning, and S. S. Renner. 2016. Day length unlikely to
460 constrain climate-driven shifts in leaf-out times of northern woody plants. Nature Climate
461 Change 6:1120–1123.

⁴⁶² **Tables and Figures**

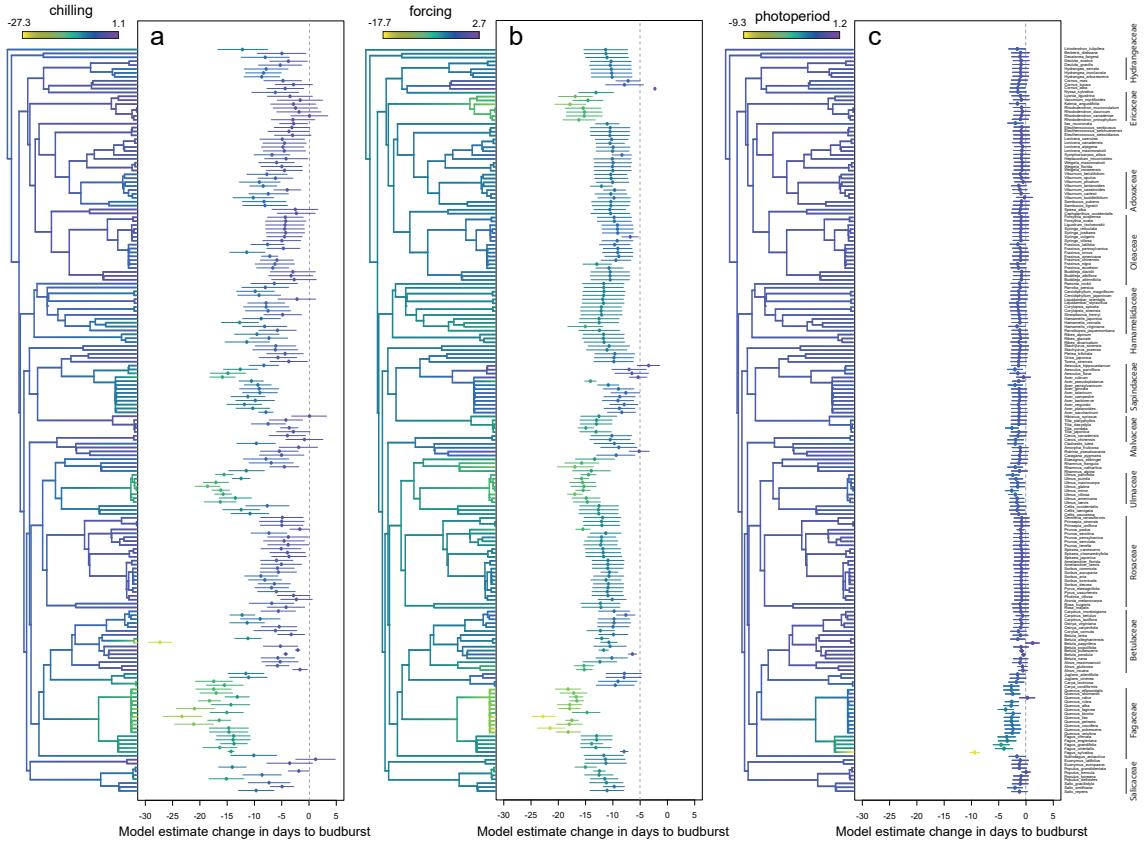


Figure 1: Phenological sensitivity to three environmental cues, chilling (a), forcing (b) and photoperiod (c) measured as change in days to budburst per standardized unit (z-transformation) of the cues across 191 tree species. The same phylogenetic tree is shown in each panel, colored according to an estimation of ancestral character states, being the states at the tips the species' sensitivities to a cue, as estimated by our hierarchical phylogenetic model. Species sensitivities are shown along with 50% uncertainty intervals in the diagrams. Note that the color scale varies in each panel. Total tree depth is 81 My.

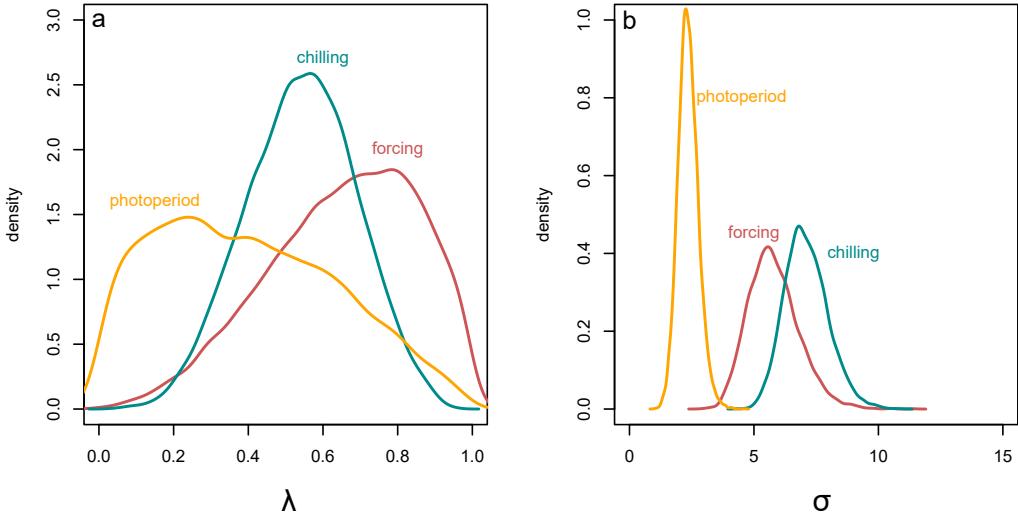


Figure 2: Density plots comparing the posterior distributions of phylogenetic parameters λ and σ estimated for each cue in the model: chilling (blue), forcing (red), and photoperiod (orange). Panels correspond to λ (a) and σ (b) from the phylogenetic model.

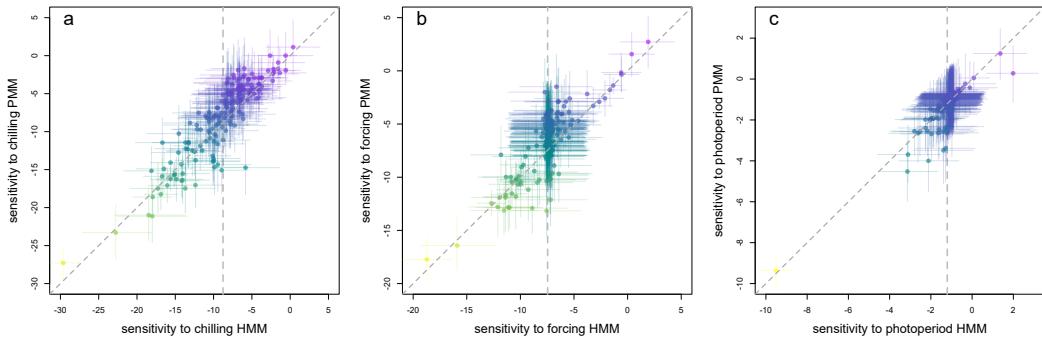


Figure 3: Correlations between model parameters as estimated by the model including phylogenetic structure on each phenological cue (y -axis), and the more commonly used hierarchical model where species are exchangeable (where λ is constrained to be equal to zero, x -axis). While species with large amounts of data may be estimated similarly by both models, in the more commonly used hierarchical model (x -axis) many species are pulled towards the overall average (shown by dashed grey vertical lines). The strength and prevalence of pulling across species is particularly obvious for forcing (b). Panels correspond to sensitivity to chilling (a), forcing (b), and photoperiod (c). Dashed grey 1:1 lines also shown. Estimate colors are in the same scale as in 2, and error lines correspond to 50% uncertainty intervals.

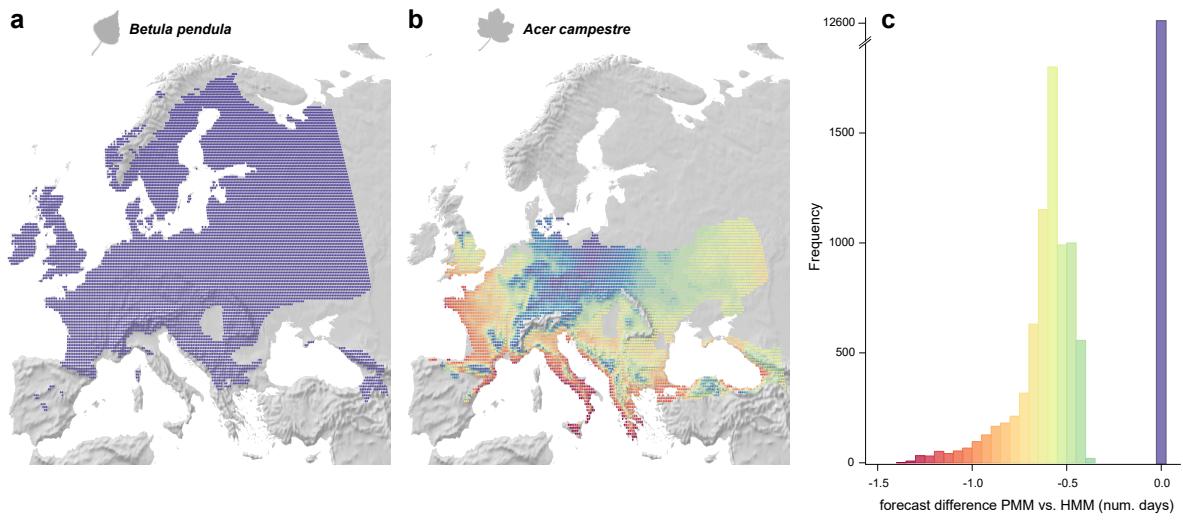


Figure 4: Comparison of forecasts of phenological shifts (i.e., computed as the difference between predictions under current climate vs. a 2°C warmer climate) resulting from a phylogenetic (PMM) and a non-phylogenetic (HMM) approach. Differences in forecasted shifts are negligible for well sampled species (*Betula pendula*, $n = 311$, a), but can be substantially different for poorly sampled species in well-sampled clades (*Acer campestre*, $n = 6$, b). The maps show the difference in number of days between the shifts predicted by PMM and HMM, with values colored according to histograms in panel c (days here are relative to start of forcing conditions, not calendar days). See Supplementary Material for details on forecast calculation.