

Phylogenetic estimates of species-level phenology improve ecological forecasting

immediate

June 15, 2024

Author list

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. Sodhi³ & E. M. Wolkovich^{3,4,5}

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 Umea, Sweden

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

*Corresponding author: Ignacio Morales-Castilla; e-mail: ignacio.moralesc@uah.es

Abstract

The ability to adapt to climate change requires accurate ecological forecasting. Current forecasts, however, have failed to capture important variability in biological responses, especially across species. Here, we present a novel method using Bayesian hierarchical phylogenetic models to overcome this challenge and estimate species-level responses. We apply our method to phenological experiments manipulating temperature and daylength and show species differences are larger than the average differences between cues. We illustrate how a focus on certain clades can bias prediction, but that predictions may be improved by integrating information on phylogeny. Our model demonstrates an underlying phylogenetic structure in plant phenological responses to temperature cues, whereas responses to photoperiod appear weaker, more uniform across species, and less phylogenetically constrained. Our approach provides an advance in ecological forecasting, with implications for predicting the impacts of climate change and other anthropogenic forces on ecosystems.

Methods

Phenological and Phylogenetic Data

Phenological data: To estimate phenological responses to chilling, forcing and photoperiod we used data from phenological experiments in controlled environments of temperate woody species, brought together in the Observed Spring Phenology Responses in Experimental Environments (OSPREE) database. In July 2019, we updated an earlier version of this database (1) by reviewing all papers found through searching ISI Web of Science and Google Scholar with the following terms:

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

We scraped data from all papers of woody species that tested for photoperiod and/or temperature effects on budburst, leafout, or flowering, resulting in data from 155 experiments across 97 papers in the updated database. (author?) (2), which used a portion (72 experiments across 49 papers) of the earlier OSPREE database, provides extensive methods on database creation and cleaning.

We focused on angiosperms (as gymnosperms are very poorly represented in spring phenology experiments), and included all budburst experiments where we could quantify chilling, forcing and photoperiod levels, resulting in 44 studies from 33 papers, and 2940 data points. In our dataset most studies come from Europe ($n=37$) and a few from North America ($n=7$). The same bias towards Europe is found across the full OSPREE dataset with less North American ($n=19$) than European ($n=60$) studies and only 3 studies located in the Southern Hemisphere. Given our need of daily gridded data for estimating chilling we only include studies from Europe and North America (see Fig. S3). Our final dataset is both geographically and taxonomically limited, mirroring the existing literature on phenological experiments and highlighting a critical need to expand this literature.

Across experiments chilling treatments were often fully or partially applied in the field, thus we estimated chilling—both in the field and applied in controlled environments—using Utah units with the `chillR` package. We estimated field chilling from 1 September to the date given for when samples were taken from the field using daily temperature data (converted to hourly) from both European (E-OBS, version 16, calculating the average of minimum and maximum daily temps, 3) and North American (v3, 4) gridded climate datasets. We also converted experimental chilling into Utah chill units, based on reported treatments (for studies with a mix of field and experimental treatments, we added field and experimentally applied Utah units). To avoid numerical instability in our models (from having predictor values on very different scales), we divided Utah units by 240 (roughly equivalent to 10 days of average chilling).

We report the Utah model because a small number of studies reported chilling only in Utah units, thus using this common metric allowed us to include the most data. The Utah model relies on the assumption that temperatures between 1.4 and 15.9°C affect endodormancy release differently, though recent findings show possibly similar effects for a wide range of temperatures (-2 to 10°C, see 5). Because chilling is a latent process an accurate model of it, especially for the 191 species in our dataset, is not currently possible (2). We found consistent results, however, using another common model of chilling—chill portions, suggesting our results are robust to the exact chilling metric used (see Tables S4 and S5 in Supplementary Information).

Forcing and photoperiod treatments occurred after chilling treatments; we report photoperiod as the length

of light and weighted these treatments by the reported photo- and thermo-periodicity (6). Most studies reported two temperatures per day across the whole experiment, one for day and night, but some had ramped temperatures and/or photoperiods (or other complexities). In these cases we built an hourly model of the full treatment period until budburst and took the mean value. For a phylogenetic tree, we pruned the megatree for seed plants (7) to extract the sub-tree containing only the species present in the OSPREE dataset, species not included in the megatree were added to the congeneric basal node age (using the function ‘congeneric.merge’ in 8), and assigned branch lengths to maintain tree ultrametricity. This addition of species to the tree can introduce polytomies (multifurcations) when many species are added to the same ancestral node, as was the case for *Acer*, for which several species were included in the OSPREE dataset but the megatree lacked species-level resolution within the genus. In total, our pruned tree had 8 polytomies affecting 46 out of 191 species. Our Bayesian hierarchical model, described below, is informed by the phylogenetic structure describing species evolutionary relationships. Errors in phylogenetic topology and branching times could thus impact model estimates, although if errors were large the contribution of phylogeny would simply be scaled to zero. To assess whether the inclusion of polytomies in our data biased model estimates, we ran sensitivity analyses excluding these species from models (see Table S8 in Supplementary Information). Our approach assumes a tree where branch lengths represent time, but it could be possible to change this assumption. For example, if the genes underlying plant responses to particular cues were known, branch lengths would directly represent mutational changes along gene sequences. In the absence of such detailed gene specific data, evolutionary time provides a useful proxy for species differences.

Bayesian hierarchical phylogenetic model

Commonly used phylogenetic regression methods today (e.g., phylogenetic generalized least squares models, PGLS, (9); phylogenetic mixed models, PMM, (10)) were originally conceived as statistical corrections for phylogenetic non-independence across observations—generally species—thus allowing multi-species studies to meet the assumptions of linear regression (9). These corrections incorporated phylogenetic structure by estimating the magnitude of a transformation of a variance-covariance (VCV) matrix whose elements were derived from the amount of evolutionary history (branch lengths) shared between species on a phylogeny. The most commonly used transformation was Pagel’s λ —a multiplier of the off-diagonal elements—where estimates of $\lambda = 1$ essentially left the VCV untransformed and suggested that the residuals of the regression had phylogenetic signal consistent with Brownian motion; estimates of $\lambda = 0$ suggested no phylogenetic signal. Because the original aim of these methods was to correct for statistical bias introduced by shared evolutionary history among species, the underlying assumption of phylogenetic regressions is that phylogenetic relatedness would only affect either model residuals (in PGLS approaches, 9) or model intercepts (e.g., in many PMM approaches, 10).

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

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

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

where

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

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

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

$$\begin{aligned} \boldsymbol{\alpha} &= [\alpha_1, \dots, \alpha_n]^T \text{ such that } \boldsymbol{\alpha} \sim \mathcal{N}(\mu_{\alpha}, \boldsymbol{\Sigma}_{\alpha}) \\ \boldsymbol{\beta}_{chill} &= [\beta_{chill,1}, \dots, \beta_{chill,n}]^T \text{ such that } \boldsymbol{\beta}_{chill} \sim \mathcal{N}(\mu_{\beta_{chill}}, \boldsymbol{\Sigma}_{\beta_{chill}}) \\ \boldsymbol{\beta}_{force} &= [\beta_{force,1}, \dots, \beta_{force,n}]^T \text{ such that } \boldsymbol{\beta}_{force} \sim \mathcal{N}(\mu_{\beta_{force}}, \boldsymbol{\Sigma}_{\beta_{force}}) \\ \boldsymbol{\beta}_{photo} &= [\beta_{photo,1}, \dots, \beta_{photo,n}]^T \text{ such that } \boldsymbol{\beta}_{photo} \sim \mathcal{N}(\mu_{\beta_{photo}}, \boldsymbol{\Sigma}_{\beta_{photo}}) \end{aligned} \quad (3)$$

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:

$$\begin{bmatrix} \sigma_i^2 & \lambda_i \times \sigma_i \times \rho_{12} & \dots & \lambda_i \times \sigma_i \times \rho_{1n} \\ \lambda_i \times \sigma_i \times \rho_{21} & \sigma_i^2 & \dots & \lambda_i \times \sigma_i \times \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_i \times \sigma_i \times \rho_{n1} & \lambda_i \times \sigma_i \times \rho_{n2} & \dots & \sigma_i^2 \end{bmatrix} \quad (4)$$

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

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

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

where the residual phylogenetic error terms (e.g., e_{α_j}) are elements of normal random vectors from multivariate normal distributions centered on 0 with the same phylogenetic variance-covariance matrices as in equation 4. Model code, including priors used here, are given in the supplement. We fit all models to our data using **RStan** using 4 chains of 4,000 iterations with a warmup of 2,000 each (resulting 8,000 posterior samples), and assessed fit via \hat{R} near 1 and adequate effective sample sizes (see Tables S2-S3).

To assess if the phylogenetic mixed model (PMM) presents any advantages with respect commonly used hierarchical mixed models (HMM; see e.g., 2), beyond fitting evolutionary parameters to model predictors,

we compare results of PMM and HMM. HMM is a simplified version of PMM where off-diagonal elements of the variance-covariance phylogenetic matrices are multiplied by zero ($\lambda = 0$). Both models account for differences in sample sizes and variances for each species, by partially pooling across all data while at the same time providing species-level estimates; however the PMM will pool more strongly to closely-related species when λ is high. Additional grouping factors beyond species could be added to these models. For example, similar approaches can be used to estimate study or location effects.

References

References

- [1] Wolkovich, E. M. *et al.* Observed Spring Phenology Responses in Experimental Environments (OS-PREE). doi:10.5063/F1CZ35KB (2019).
- [2] Ettinger, A. *et al.* Winter temperatures predominate in spring phenological responses to warming. *Nature Climate Change* 1–6 (2020).
- [3] Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. & Jones, P. D. An ensemble version of the E-OBS temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres* **123**, 9391–9409 (2018).
- [4] Sheffield, J., Goteti, G. & Wood, E. F. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate* **19**, 3088–3111 (2006).
- [5] Baumgarten, F., Zohner, C. M., Gessler, A. & Vitasse, Y. Chilled to be forced: the best dose to wake up buds from winter dormancy. *New Phytologist* **230**, 1366–1377 (2021).
- [6] Buonaiuto, D. M., Donahue, M. & Wolkovich, E. M. Experimental designs for testing the interactive effects of temperature and light in ecology: the problem of periodicity. *Functional Ecology* (2023).
- [7] Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *American journal of botany* **105**, 302–314 (2018).
- [8] Pearse, W. D. *et al.* Pez: Phylogenetics for the environmental sciences. *Bioinformatics* **31**, 2888–2890 (2015).
- [9] Freckleton, R. P., Harvey, P. H. & Pagel, M. Phylogenetic analysis and comparative data: a test and review of evidence. *The American Naturalist* **160**, 712–726 (2002).
- [10] Housworth, E. A., Martins, E. P. & Lynch, M. The phylogenetic mixed model. *The American Naturalist* **163**, 84–96 (2004).

