Received: 4 August 2023 | Accepted: 21 March 2024

DOI: 10.1111/1365-2745.14317

Journal of Ecology | BRITISH ECOLOGICAL SOCIETY

FORUM

Phenological mismatch between trees and wildflowers: Reconciling divergent findings in two recent analyses

Benjamin R. Lee[1,2,3] | Evelyn F. Alecrim[4] | Tara K. Miller[5] | Jessica R. K. Forrest[4] | J. Mason Heberling[1,2] | Richard B. Primack[6] | Risa D. Sargent[7]

[1]Carnegie Museum of Natural History, Section of Botany, Pittsburgh, Pennsylvania, USA; [2]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; [3]Holden Arboretum, Kirtland, Ohio, USA; [4]Department of Biology, University of Ottawa, Ottawa, Ontario, Canada; [5]Repair Lab, University of Virginia, Charlottesville, Virginia, USA; [6]Biology Department, Boston University, Boston, Massachusetts, USA and [7]Faculty of Land and Food Systems, University of British Columbia, Vancouver, British Columbia, Canada

Correspondence
Benjamin R. Lee
Email: benrlee@umich.edu

Funding information
Natural Sciences and Engineering Research Council of Canada; National Science Foundation Division of Graduate Education, Grant/Award Number: DGE-1735087; Division of Environmental Biology, Grant/Award Number: 1936877 and 1936971; Ontario Ministry of Research, Innovation and Science

Handling Editor: Amy Iler

Abstract

1. Recent evidence suggests that community science and herbarium datasets yield similar estimates of species' phenological sensitivities to temperature. Despite this, two recent studies by Alecrim et al. (2023) and Miller et al. (2022) found very different results when using different data sources (community science and herbarium specimens, respectively) to investigate whether warming threatens wildflowers with phenological mismatch in relation to shading by deciduous trees.
2. Here, we investigated whether differences between the two studies' results could be reconciled by testing four hypotheses related to model design, species, spatiotemporal data extent and phenophase.
3. Hybrid model structures brought results from the two datasets closer together but did not fully reconcile the differences between the studies. Neither the species nor the phenophase selected for analysis seemed to be responsible for differences in results. Cropping the datasets to match spatial and temporal extents appeared to reconcile most differences but only at the cost of much higher uncertainty associated with reduced sample size.
4. *Synthesis*: Our analysis suggests that although species-level estimates of phenological sensitivity may be similar between community science and herbarium datasets, inherent differences in the types and extent of data may lead to contradictory inference about complex biotic interactions. We conclude that, until community science data repositories expand to match the range of climate conditions present in herbarium collections or until herbarium collections match the spatial extent and temporal frequency of community science repositories, ecological studies should ideally be evaluated using both datasets to test the possibility of biased results from either.

KEYWORDS
citizen science, climate change, community science, global change ecology, herbarium collections, nature's notebook, phenological escape, phenology, spring ephemeral wildflowers, spring-blooming wildflowers

1 | INTRODUCTION

A trend towards earlier flowering and leafing-out of plants in the temperate zone of the Northern Hemisphere in response to warmer spring temperatures has emerged as one of the most sensitive and well-documented biological indicators of climate change (Chmielewski & Rötzer, 2001; Inouye, 2022; Parmesan, 2007). This shift in spring phenology has wide-ranging implications for ecosystem services and processes (Kim et al., 2018; Piao et al., 2019; Richardson et al., 2013) as well as for ecological interactions involving pollinators, insect herbivores, and seed dispersers (e.g. Freimuth et al., 2022; Ren et al., 2020; Simmonds et al., 2020). Co-evolved, interacting species may react differently to warming, potentially altering their phenological overlap, a situation that has been termed "phenological mismatch" (Renner & Zohner, 2018). Ecologists are actively investigating this phenomenon and discussing what constitutes proof of its existence (Iler et al., 2021; Kharouba & Wolkovich, 2023; Samplonius et al., 2021). Phenological mismatch is challenging to measure, and evidence to support it is limited, with most studies focusing on mismatches between trophic levels (i.e. "trophic mismatch"), especially with respect to plant–pollinator interactions (Renner & Zohner, 2018).

Recently, a paper by Heberling, McDonough, et al. (2019), using field observational records that date back to the 1850s for 15 tree species and 14 herbaceous wildflower species at one location, concluded that trees and spring-blooming wildflowers in eastern North America are increasingly exhibiting phenological mismatch because tree phenology is more responsive to spring temperature compared to that of wildflowers. Spring-blooming wildflowers rely on access to early seasonal light (i.e. "phenological escape"; Jacques et al., 2015; Lee & Ibáñez, 2021a) to assimilate often more than half—and up to 100%—of their annual carbon budget (Heberling, Cassidy, et al., 2019; Lapointe, 2001). However, when tree phenology is more responsive to climate change than wildflower phenology, wildflowers experience less spring light than in the past, resulting in a phenological mismatch. Indeed, one of the key findings of the Heberling, McDonough, et al. study is that wildflowers are experiencing less spring light than they did in the past because of lower phenological responsiveness to spring temperature in comparison with trees. Further, this window of spring sunlight is expected to become even shorter in coming decades, with projected reductions in energy budgets, survival, and reproductive success of wildflower species (Heberling, Cassidy, et al., 2019; Lee et al., 2022; Lee & Ibáñez, 2021b). Because species in the herbaceous understory make up most of the plant diversity in North American forests (Gilliam, 2007; Spicer et al., 2020, 2022), with spring ephemerals comprising about a fifth of all understory herbaceous species (Yancy et al., 2023), projected declines in energy budgets could have serious consequences for the conservation of biodiversity in these forests.

While the Heberling, McDonough, et al. (2019) study was innovative, one limitation was that all the phenological observations were obtained from a single location: the town of Concord, Massachusetts, USA. Would the pattern of phenological mismatch between trees and wildflowers shown for this one place be shared across larger geographical ranges encompassing wider variation in environmental conditions? Recently, two articles written by independent research teams investigated this very question across eastern North America (Alecrim et al., 2023; Miller et al., 2022). Surprisingly, they reached opposing conclusions (Figure 1a). Miller et al. (2022), using data from herbarium specimens of 12 species (six trees and six wildflowers) collected between 1870 and 2019, concluded that eastern North American trees were more responsive to a warming climate than were spring-blooming wildflowers. These findings echoed the single-site results of the Heberling, McDonough, et al. (2019) study. In contrast, Alecrim et al. (2023), using phenological data on 11 wildflower and 10 tree species collected by community scientists (USA National Phenology Network; https://www.usanpn.org) between 2009 and 2021, found that spring-blooming wildflowers were more responsive to a warming climate than were trees—suggesting that wildflowers would instead experience a longer window of spring sunlight as temperatures rise. While data from community science projects and herbaria can be correlated and reveal similar patterns (e.g. Iwanycki Ahlstrand et al., 2022; Ramirez-Parada et al., 2022; Spellman & Mulder, 2016), the studies by Miller et al. (2022) and Alecrim et al. (2023) suggest that these different data sources may sometimes lead to starkly different conclusions when applied to complex ecological interactions.

For the present article, several authors of these two papers, and of a related paper (Lee et al., 2022), came together to discuss what factors (summarized in Table 1) may have contributed to the contradictory results and to test whether controlling for some of these factors could lead to better agreement between the two papers' analyses and conclusions. Ultimately, we aimed to reconcile the contrasting results from these two studies, which would also help to inform the use and interpretation of community science and herbarium data in future phenological studies.

In our comparison of these two studies, we first identified several hypotheses which may explain the incongruencies in results: (1) differences in how models were structured and analysed (including how data were binned, how spring temperatures were calculated, how models were coded, and what statistical packages were used); (2) differences in the species for which data were available; (3) differences in the spatiotemporal extents of the data used (and consequent differences in the temperature ranges considered); and (4) differences in the phenophase measured (i.e. flowering versus first leaf-out of the wildflowers). Here, we address these points using a comparative modelling approach and quantitative analyses. We then discuss the implications of our findings for other studies that use long-term data to assess phenological mismatch.

2 | METHODS AND ANALYSIS

2.1 | H1. Differences in model structure and analysis

Studies of how climate change affects phenological mismatch typically compare the reaction norms (phenological sensitivities) of interacting species to the same climatic driver (often

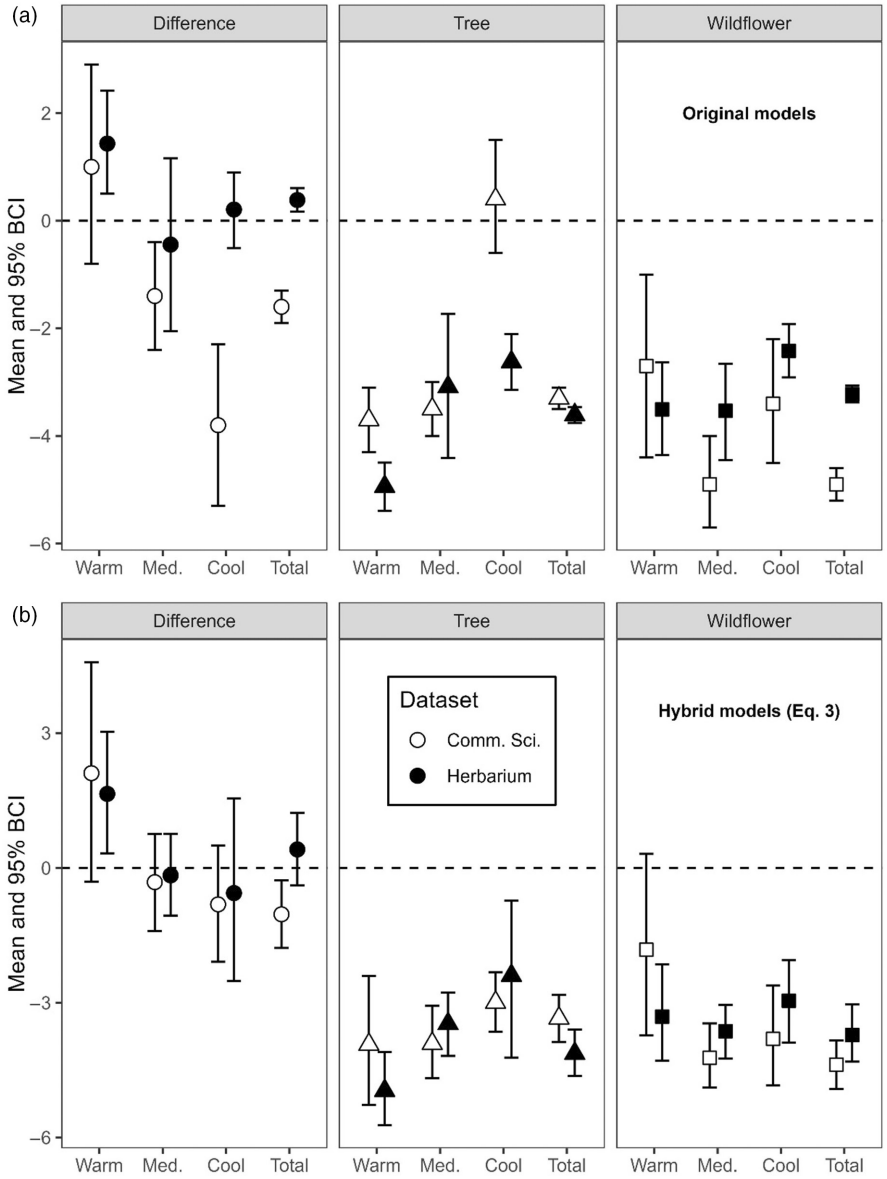

FIGURE 1 Posterior estimated means (points) and 95% Bayesian credible intervals (BCI; whiskers) of phenological sensitivities (β; in days °C^{-1}) of wildflowers (squares), trees (triangles), and the difference between them (circles), (a) as reported by Alecrim et al. (2023), using Community Science (Comm. Sci.) data; open symbols) and Miller et al. (2022), using Herbarium data; filled symbols), and (b) controlling for temperature window (Mar–May), historical temperature data estimates (PRISM), model structure (see Equation 3), binning approach, and statistical program. A negative difference indicates that wildflowers have higher sensitivity to spring temperature than trees; a positive difference indicates the opposite.

temperature). Indeed, both Miller et al. (2022; using herbarium collections data) and Alecrim et al. (2023; using community science observations) used similar statistical approaches to quantify the sensitivity of spring wildflower and canopy tree phenology to average spring temperature, which is known to be a major driver of plant phenology in spring in eastern North America (Flynn & Wolkovich, 2018; Lee & Ibáñez, 2021a; Polgar et al., 2014; Sevenello et al., 2020). From here on, we refer to these studies by their data sources (i.e. the 'herbarium' and 'community science' studies, respectively).

In addition to the differences in the data used, there were nuanced differences in model structure. Specifically, the community science model structure estimated responses of Day of Year (DOY) to spring temperature (spanning March–May) with hierarchical group interactions (j; species × site level; equivalent to random effects in mixed modelling approaches) shaping estimates of model slope and intercept. The community science analysis also added a year random effect (k) to the intercept to account for non-independence of observations from the same year:

$$\mathrm{DOY}_{[i]} = \alpha_{jk[i]} + \beta_{j[i]} \times \mathrm{temperature}_{[i]} \quad (1)$$

The herbarium model used a different metric of spring temperature (spanning March–April) than the community science model. Further, these authors used a common intercept (α_0) for all species within each functional group (i.e. spring-blooming wildflowers vs. trees), did not include random effects in the slope estimates, and included random intercepts only for species (s; note: not the same as the j species × site term used in Equation 1 since herbarium collections and measurements are not typically repeated at the site level) and year (k):

$$\mathrm{DOY}_{[i]} = \alpha_0 + \alpha_{sk[i]} + \beta_{[i]} \times \mathrm{temperature}_{[i]} \quad (2)$$

These differences in model structure and in the definition of spring temperature between the two studies could have contributed

TABLE 1 Differences between Alecrim et al. (2023) and Miller et al. (2022) studies. "Hypothesis" refers to the numbered hypotheses discussed in the text.

Hypothesis	Factor	Alecrim et al. (2023); community science data	Miller et al. (2022); herbarium data	Explanation
N/A	Nature of data	Community science records (USA–National Phenology Network)	Herbarium records	Herbarium specimens were most likely to be collected at peak of phenophase, with limited repeat sampling of same population; community science records were filtered to obtain the earliest observation at a locality, with more common repeat observations of same population (or even individual)
H1	Model structure	Random intercepts for year; random intercepts and slopes for species × site	Random intercepts for species and year; no random slopes	Complex model structures could affect the results in subtle ways that are difficult to predict
H1	Modelling software	STAN (using *brms* package; Bürkner, 2017, 2021)	JAGS (Plummer, 2003; using *R2jags* package, Su & Yajima, 2015)	Differences resulting from using different statistical software are unlikely, but technically possible
H1	Temperature metric used	Mean of March 1 to May 31 temperature from PRISM (but other date intervals also considered)	Mean of March 1 to April 30 temperature from nearby NOAA weather stations	Temperature ranges used might affect phenological sensitivity estimates if species are affected more by temperatures early or later in the season (Keenan et al., 2020)
H1	Binning approach	Binned data by latitude	Binned data by temperature	Different bins could create artefacts in the analysis that result in different slopes based on how the data are partitioned. However, this should not affect overall (range-wide) conclusions
H2	Species considered	10 native tree species, 11 native wildflower species	6 native tree species, 6 native wildflower species	Only 4 tree and 2 wildflower species in common between the two datasets; species likely vary in phenological response to temperature, so different species suites could cause differences in overall signal
H3	Geographic extent	35.0–48.2° N, 67.4–90.9° W; northeastern USA only	27.2–48.6° N, 68.0–96.4° W; eastern USA and Canada	Larger geographic range may encompass broader range of temperature and phenological response; different range extents may not equally account for spatial autocorrelative effects previously shown to affect phenological sensitivity (Wang et al., 2023)
H3	Temporal extent	2009–2021	1870–2019	Longer timespan may encompass broader range of temperature and phenological response, but longer time series can also reduce likelihood of detecting temporal trends. Shorter datasets can show fluctuations that can be misconstrued as trends (e.g. Iler et al., 2013; Schmidt et al., 2023; Thomson, 2019)
H4	Phenophase considered	Wildflowers: leaf-out date ("above ground buds with green tips") Trees: leaf-out date ("fully unfolded leaves")	Wildflowers: presence of open flowers Trees: presence of young leaves	Different phenophases may respond differently and may be more or less sensitive to changing temperature (e.g. Ettinger et al., 2018; Geng et al., 2022)

to the observed differences in overall estimates of phenological sensitivity (Table 1). In both original studies, tree phenological sensitivity was modelled separately from wildflower phenological sensitivity. Therefore, the α_0 term in the herbarium model (Equation 2) can be interpreted as the effect of functional group (trees versus wildflowers). With the community science model (Equation 1), the effect of functional group can be calculated from the averages of the hierarchical $\alpha_{j,k}$ and β_j terms for all species within each functional group.

Here, we modelled phenological sensitivity as a function of spring temperature using a hybrid model structure, adopting the

March–May period used in the original community science analysis. We downloaded spring temperature data for all observations in both datasets from the PRISM database (PRISM Climate Group, Oregon State University, 2022, https://www.prism.oregonstate.edu), following the methodology described in Alecrim et al. (2023).

Importantly, both studies included additional ("binned") analyses aimed at quantifying differences in phenological sensitivity at southern, mid, and northern latitudes. The community science paper separated observations by latitude (prioritizing spatial congruency), while the herbarium paper instead binned observations based on their associated March–April mean temperatures (prioritizing environmental similarity). These different binning approaches could lead to between-study differences in trends observed within latitudinal or climatic zones; however, they would not on their own explain why the community science paper found range-wide greater temperature sensitivity in wildflowers while the herbarium paper found greater temperature sensitivity in trees. We controlled for binning strategy in the present hybrid analysis by adopting a latitudinal binning strategy similar to the one used in the community science paper, but modified to better encapsulate the differences in latitudinal variation across both datasets (Figure 2a). Furthermore, we present results for all hypotheses in terms of overall model results (considering all data together without binning) and in terms of results for individual bins.

Finally, while both analyses used similar hierarchical Bayesian modelling approaches, the herbarium analysis used JAGS software, which approximates the posterior distribution using a Gibbs sampler (Plummer, 2003), whereas the community science paper used STAN, which approximates the posterior distribution using a Hamilton Monte Carlo sampler (Bürkner, 2017, 2021). Previous studies suggest it is unlikely that the different programs and samplers would yield substantially different results (Monnahan et al., 2017), but it is still possible that they could be responsible for the discrepancies in posterior estimates noted above and thus should be controlled for.

To test the influence of these differences in model structure, programming environment, and spring temperature window, we reran the analyses using a single common spring temperature window (average March–May temperatures extracted from PRISM), and statistical program (STAN). We used a hybrid model structure that primarily reflects the herbarium model structure (Equation 2), but with the addition of random slope effects for species:

$$\mathrm{DOY}_{[i]} = \alpha_{sk[i]} + \beta_{s[i]} \times \mathrm{temperature}_{[i]} \quad (3)$$

2.2 | H2. Differences in species investigated

Phenological sensitivity has been shown to vary systematically across different groups of plants based on shared traits, growth strategies and genetics (Panchen et al., 2014; Rafferty & Nabity, 2017). While phenological sensitivity is commonly aggregated by groups of species (as with the two papers analysed here), the most common level at which phenological sensitivity is reported is at the species level, and sensitivity can vary substantially across species within the same functional group. For example, although the community science study found an average phenological sensitivity of −4.9 days °C^{-1}, species-specific sensitivities ranged between −5.4 and −3.0 days °C^{-1}, suggesting the potential for large differences in outcomes with respect to changes in spring light availability. Some species, like *Sanguinaria canadensis*, were extremely sensitive (and therefore more likely to gain access to light with warming), while others, such as *Trillium grandiflorum*, were less sensitive and thus more likely to experience reduced access to spring light.

The community science and herbarium analyses each derived the overall sensitivity of wildflower and tree phenology by averaging across different groups of species in these functional groups. The community science analysis collated observations for more species (11 wildflower species and 10 tree species) than did the herbarium analysis (six of each group). Four tree species were included in both

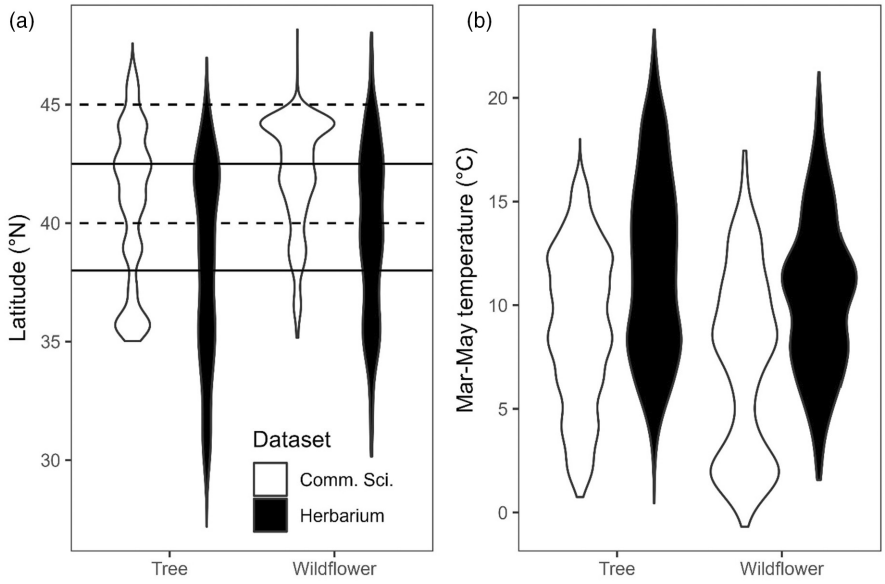

FIGURE 2 Violin plots showing the distributions of observations of trees and wildflowers in the Community Science (Comm. Sci.; open shapes) and Herbarium (filled shapes) datasets by (a) latitude and (b) average spring (March–May) temperature. Horizontal lines in Panel (a) represent the latitudinal bin breaks used in the original community science analysis (dashed lines at 40° N and 45° N) and those we used for model comparisons (solid lines at 38° N and 42.5° N).

studies (*Acer rubrum*, *Acer saccharum*, *Fagus grandifolia* and *Quercus rubra*), but only two wildflower species were common between them (*Erythronium americanum* and *S. canadensis*). Because the herbarium analysis did not include species-level random slopes, we are unable to compare species-level sensitivities from the original models. Instead, we compare the species-level reaction norms for the hybrid binned model design (Equation 3) for the wildflowers and trees in each study while maintaining the same spring temperature windows, latitudinal bins and programming environment described in H1.

2.3 | H3. Differences in spatiotemporal extent

Shifts in plant phenology are dependent on environmental variation in drivers such as temperature and precipitation. While many global change studies focus on how environmental conditions vary over time (Fu et al., 2012; Kudo et al., 2004; Zohner et al., 2018), there is also substantial literature showing how geospatial variation in environmental conditions affects the timing of flowering and leaf-out both within a given growing season and in terms of long-term temperature sensitivity (Peaucelle et al., 2019; Wang et al., 2023; Willems et al., 2022). Moreover, several of these studies focus specifically on how temporal and spatial environmental variation interact to influence the phenology of both plants and animals (Kharouba et al., 2018; Kudo & Ida, 2013; Wann et al., 2019).

There are substantial differences between the two studies in both the spatial and temporal extent of the data, with the difference in temporal extent being larger due to the data sources involved. The community science study collated all observations from the USA–National Phenology Network database (https://data.usanpn.org/observations/get-started; hereafter 'NPN') from 1 January to 31 July in the years 2009–2021, limiting their dataset to locations from latitudes 35° to 48.2° N (based on data availability) and longitudes east of 91° W (reflecting the geographical range of eastern hardwood forests). This data source only includes observations made in the United States, so the authors could not include any Canadian observations. In contrast, the herbarium study included observations dating from 1870 to 2019, with more observations in the south than in the north, ranging from 27.2° to 48.6° N. Their dataset included observations from southern Canada and from as far west as 96.4° W. The latitudinal distribution of observations in both datasets is presented in Figure 2a.

Unsurprisingly, the herbarium dataset extends much further back in time than the community science dataset. Although earlier years experienced cooler springs compared to more recent years, the southern skew of the herbarium dataset (Figure 2a) led to it having more observations collected during warmer springs compared to the more recent community science dataset (Figure 2b). Thus, the two studies capture largely overlapping but distinct spatial portions of the overall spring phenology–temperature relationship for each plant functional group, with the herbarium dataset skewing more towards warmer temperatures than the community science dataset. Further, the functional groups within each dataset also did not completely overlap—wildflowers in the community science dataset were skewed towards colder temperatures than the trees, and trees in the herbarium dataset extended to both warmer and colder temperatures compared to the wildflowers (Figures 2 and 3). If relationships between phenology and spring temperature are in fact non-linear (e.g. Iler et al., 2013; Kopp et al., 2020; Willems et al., 2022), but were modelled as linear by each team of researchers, the different slope estimates could simply be the result of quantification over different portions of the observed environmental variation—and would consequently be valid only over the range of the data each group considered (as pictured in Figure 3).

To evaluate the extent to which differences in spatiotemporal range influenced the results of these two studies, we cropped the herbarium dataset to match the community science dataset in (1) spatial, (2) temporal or (3) spatial and temporal extent. That is, we re-ran the original models described in Equation 3 using either (1) only the data from the geographic range delimited in the original community science analysis, (2) only the data from the temporal range delimited in the original community science analysis, or (3) only the data from the same temporal *and* geographic range as in the community science analysis. The original community science dataset comprised a subset of both the temporal and spatial extents of the herbarium dataset, so it was not cropped for any of these comparisons. However, we note that the community science dataset still extends beyond the thermal range of the herbarium dataset into cooler climates, especially for wildflowers (Figure 3). As before, we held the spring temperature window and programming environment constant between the two datasets.

When cropping the herbarium dataset by year, we used the same latitudinal bins as in H1. However, in the two comparisons where we cropped the herbarium dataset by latitude (1 and 3), we elected to use the latitudinal bins employed in the original community science analysis to better represent the geographic distribution of the data (Figure 2a).

2.4 | H4. Differences in phenophase

Previous studies have shown that different phenophases have different temperature sensitivities (Buonaiuto et al., 2021; Geng et al., 2022). The herbarium and community science studies examined here differed regarding the specific phenophases used, which could have contributed to the differences in results (Table 1). Specifically, while both original studies used some measure of canopy tree leaf-out phenology, the herbarium study measured phenological sensitivity of wildflower flowering (for species that leaf out at about the same time as flowering) whereas the community science study measured changes in wildflower leaf-out. To test whether these differences led to the different results, we ran the community science model (following the model structure in Equation 3) but with the phenophase described by NPN as "open flowers" (Denny et al., 2014) for wildflowers (as opposed to the original "leaves" phenophase used). We kept the phenophase "leaf-out" for trees and we

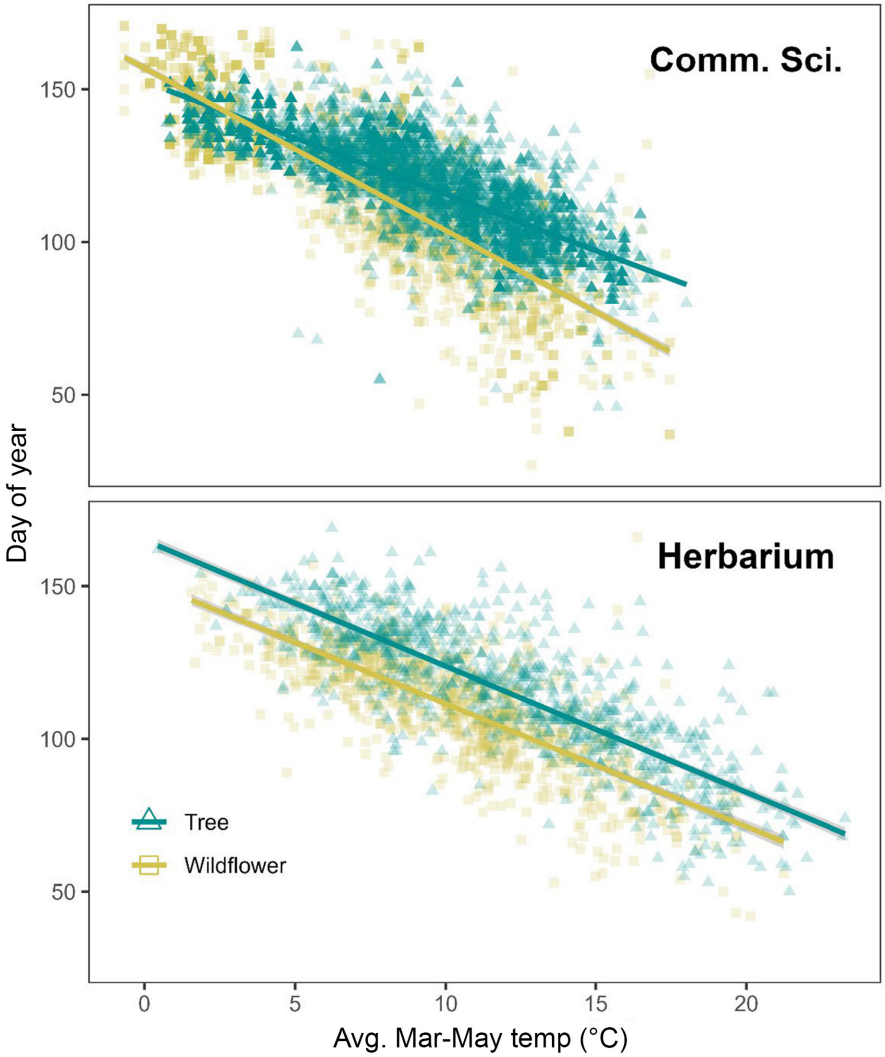

FIGURE 3 Spring phenology vs. springtime (March–May) temperatures based on the Alecrim et al. (2023); Community Science (Comm. Sci.)) and Miller et al. (2022); Herbarium) datasets. Green triangles represent observations of tree leaf-out (Community Science dataset) or presence of young leaves on trees (Herbarium dataset); yellow squares represent observations of wildflower leaf-out (Community Science dataset) or presence of open flowers on herbaceous plants (Herbarium dataset). Lines are linear fits with 95% confidence; they do not account for effects of species, year, or locality.

used the same March–May spring temperature window as described above for the hybrid model. We note that data for the "open flowers" phenophase are available for only a 10-year timespan (vs. the 13 years of leaf-out data used in the original analysis). We could not do the converse analysis (i.e. changing the wildflower phenophase used in the herbarium dataset) because of the heavy bias in herbarium collections towards wildflowers that are in bloom.

3 | RESULTS

3.1 | H1: Differences in model structure and analysis

Within the community science dataset, controlling for model structure, latitudinal bins, programming environment, and spring temperature window makes the overall temperature sensitivities (i.e. non-binned phenological reaction norms) of spring-blooming wildflowers and trees more similar to one another (Figure 1b). However, the results still indicate that spring-blooming wildflowers have greater temperature sensitivity than trees (95% BCI: −1.78 to −0.28; does not overlap zero). For the herbarium dataset, the estimated difference between functional groups now overlaps zero (95% BCI: −0.39 to 1.23). These findings indicate that the different modelling approaches and temperature calculations explain some, but not all, of the differences between the two original studies. Specifically, the overall (non-binned) estimates from the community science analysis were more robust to changes in model structure, but in both cases the estimates for trees and wildflowers moved closer to each other and the differences between datasets were reduced.

Importantly, however, slope parameters (phenological sensitivities) estimated using the latitudinal bins showed a different pattern (Figure 1). Here, results from the herbarium dataset tended to be more robust than those from the community science dataset. The estimated difference in phenological sensitivity did not differ in the warmest (southernmost) bins between the original and hybrid model structure for either dataset: the estimated difference in the community science dataset still overlapped zero and the estimated difference in the herbarium dataset remained positive (Figure 1). However, while the herbarium results were the same (with difference estimates overlapping zero) in both the original and hybrid model structures for the central and northern bins, the difference

estimates from the community science dataset moved from strongly negative to overlapping zero, with mean values very close to zero. Together, this shows that the community science range-wide (non-binned) trends are more robust to changes in model structure whereas the herbarium results tend to be more robust when broken down into latitudinal bins.

3.2 | H2: Differences in species investigated

For species common to the two datasets, phenological sensitivities (slopes of phenology vs. spring temperature) were generally similar between datasets when analysed using the hybrid modelling framework (Figure 4; 95% Bayesian credible intervals overlap for 4 of the 6 shared species). However, the community science dataset consistently showed higher temperature sensitivity for wildflowers (i.e. more negative values; Figure 4a) and lower temperature sensitivity for trees (values closer to zero; Figure 4b) compared to the herbarium dataset. This pattern mirrors the overall difference between the original studies (Figure 1) and suggests that the difference in original results cannot be explained by the inclusion of different species in each study.

A similar pattern can be seen for species not shared between the two datasets: the community science dataset shows wildflowers to be overall more sensitive (Figure 4a) and trees to be less sensitive (Figure 4b), compared to the herbarium dataset. Thus, the difference in overall (non-binned) results noted in the original studies does not seem to be due to species differences.

When evaluating differences among the latitudinal bins for each species, there was again broad overlap within the species shared across both datasets. Fifteen out of the 18 species×bin combinations showed similar estimates of phenological sensitivity across the two datasets (Figure 5; i.e. 95% BCIs overlap). The only meaningful differences in binned sensitivity between the two datasets were for the wildflower *E. americanum* (which was estimated to be more

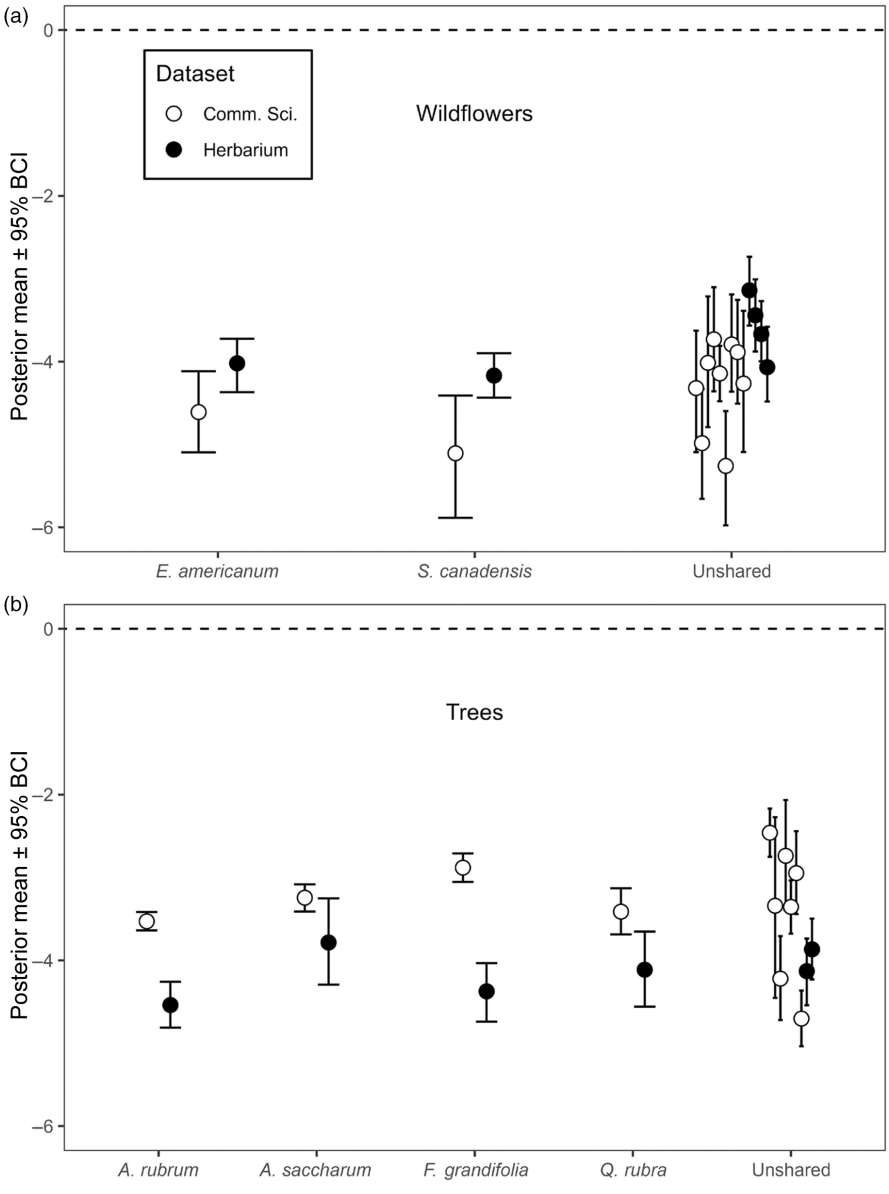

FIGURE 4 Species-level phenological sensitivities (β; in days °C^{-1}) for (a) wildflower species and (b) tree species in the Alecrim et al. (2023), Comm. Sci.; open symbols) and Miller et al. (2022), Herbarium; filled symbols) datasets. Named species were included in both papers' species pools while all species unique to each paper are grouped in the "Unshared" category. All reaction norms in this figure are estimated across the entirety of the geographic range present in the dataset. Binned parameter estimates for the shared species are shown in Figure 5 and those for the unshared species are shown in Figure S1.

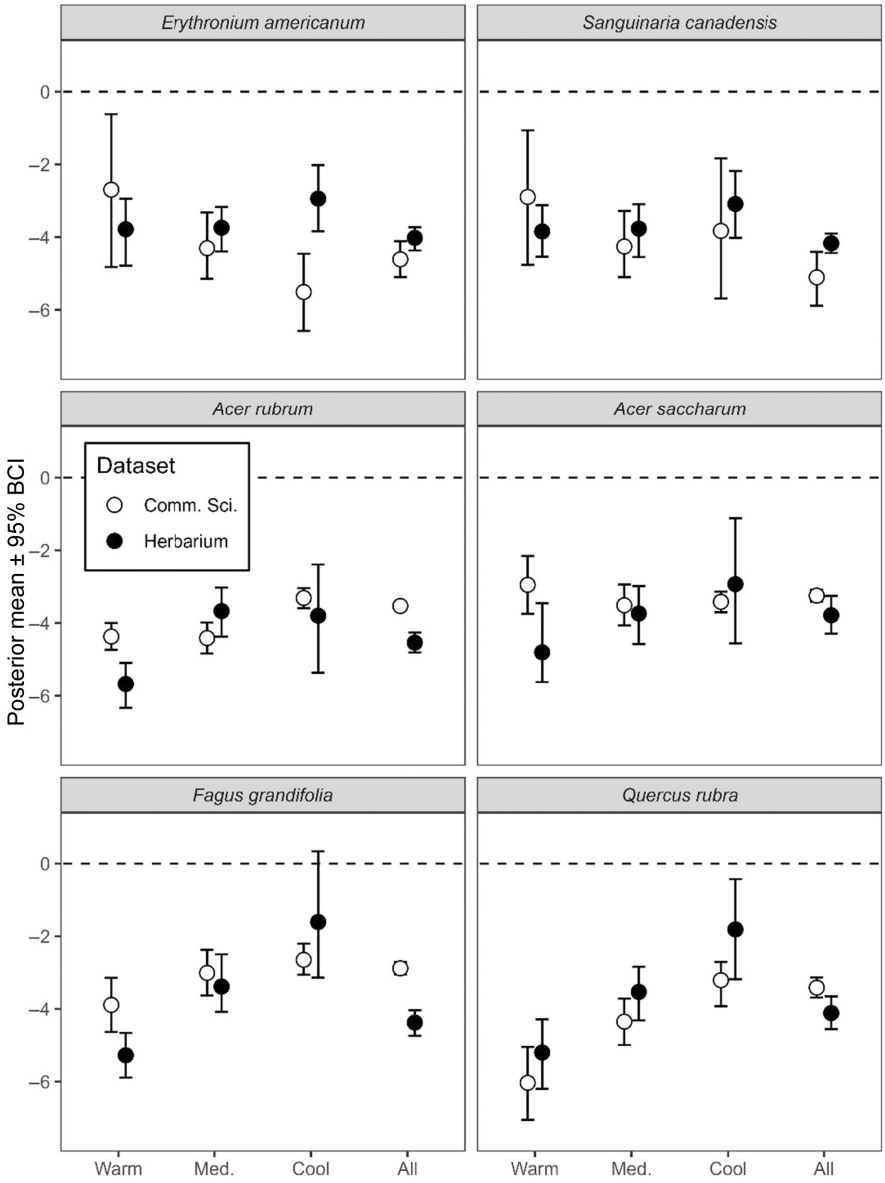

FIGURE 5 Estimated means and 95% Bayesian Credible Intervals (BCIs) for the binned species-level reaction norms (in days $°C^{-1}$) for the two wildflower species (top row) and four tree species (bottom two rows) shared across both datasets. Open symbols indicate estimates for the Alecrim et al. (2023), Comm. Sci.) dataset and filled symbols represent estimates for the Miller et al. (2022), Herbarium) dataset.

sensitive in the northern, cool bin when using the community science dataset instead of the herbarium dataset) and the trees *Acer rubrum* and *Fagus grandifolia* (which were both estimated to be more sensitive to temperature in the southern, warm bin when using the herbarium dataset instead of the community science dataset) (Figure 5). Thus, the few species-level binned differences were consistent with the overall trends, with wildflower species relatively more sensitive in the community science dataset and tree species relatively more sensitive in the herbarium dataset. Lastly, there was broad overlap between datasets in the 95% BCIs among the species that were unique to each study, although there were some pairwise differences (Figure S1).

3.3 | H3: Differences in spatiotemporal extent

Cropping the herbarium dataset to match the spatial and temporal ranges of the community science dataset reduced the extent to which the two datasets predicted opposite trends (Figure 6; compare to Figure 1b). Again, we note that because the years and the geographic extent of the community science dataset were each a subset of those encompassed by the herbarium dataset, the community science parameter estimates reported in Figure 6 are the same as those reported in Figure 1b ($n = 8045$, no change from cropping). The estimated non-binned mean difference in phenological sensitivity in the herbarium dataset moved from positive to near zero when data were cropped by latitude (Figure 6a) and moved below zero into negative values (i.e. trees less sensitive than wildflowers) when cropped by year or by both latitude and year (Figure 6b,c). In addition to these shifting mean values, 95% BCIs also became much wider and overlapped zero when compared to the results from the original herbarium dataset (Equation 3, Figure 1b), especially when cropped by year.

The increased uncertainty was primarily due to substantial reductions in sample size: the herbarium dataset was reduced

from $n = 2016$ to $n = 1582$ (21.5% reduction) when cropped by latitude, from $n = 2016$ to $n = 67$ (96.7% reduction) when cropped only by year, and to $n = 47$ (97.7% reduction) when cropped to the same spatial and temporal extents as the community science dataset. Cropping thus led to drastic limitations in the model's ability to converge in the binned analysis (see below) and this severely limits the interpretation of the overall model results presented in Figure 6. While these results suggest that differences in spatiotemporal extent of the datasets may be at least partially responsible for the differences in initially-reported model results, the reduction in sample size prevents us from making firm conclusions.

Sample size issues were most apparent with the binned results. When cropped by latitude, results from the herbarium dataset showed sharp increases in uncertainty in the northernmost cool bin (Figure S2), likely due to the strong reduction in sample size (Figure S5). Cropping by year led to strong increases in uncertainty in the warm and medium bins for the herbarium dataset (Figure S3) and left zero observations in the northernmost cool bin for both trees and wildflowers. This prevented us from making any inference about phenological escape sensitivity at northern latitudes when cropping by year. Cropping by year and latitude led to even wider credible intervals (Figure S4) and similarly prevented us from making any inference at northern latitudes. Although cropping the herbarium dataset yielded parameter estimates indistinguishable from those produced by the community science dataset, we cannot determine whether this reflects convergence on a single "true" signal or simply the high uncertainty associated with lack of data.

3.4 | H4: Differences in phenophase

When we changed the phenophase used in the community science dataset to "flowering" (as opposed to the leaf-out phenophase used in the original paper), the difference in phenological response between wildflowers and trees was not meaningfully different from either the estimates in the herbarium model or from zero. In part, this was because the posterior estimated mean difference in sensitivity moved from −1.03 in the hybrid model with wildflower leaf phenology (Figure 1b) to −0.64 when using wildflower flowering phenology instead (Figure 7). However, this shift was also affected by the greater uncertainty surrounding our posterior estimates of difference between wildflower and tree sensitivity, with wider confidence intervals when using the community science flowering data (BCIs change from (−1.78, −0.28) when using the leaf-out data to (−2.07, 0.65) when using the flowering data, an increase in uncertainty of ~1.2 days $C^{\circ -1}$). This increase in uncertainty is probably at least partially associated with reduced sample size: The wildflower sample size declined from $n = 2005$ for the leaf-out phenophase to $n = 1285$ for the flowering phenophase, a 35.9% reduction. Still, the mean signal from the non-binned community science data remained negative, meaning that the two datasets still show overall opposite trends for phenological escape duration.

Only one aspect of the binned results changed in our estimates of wildflower sensitivity when using the flowering phenophase instead of leaf-out in the community science dataset: Wildflower sensitivity in the southern, warm region changed from

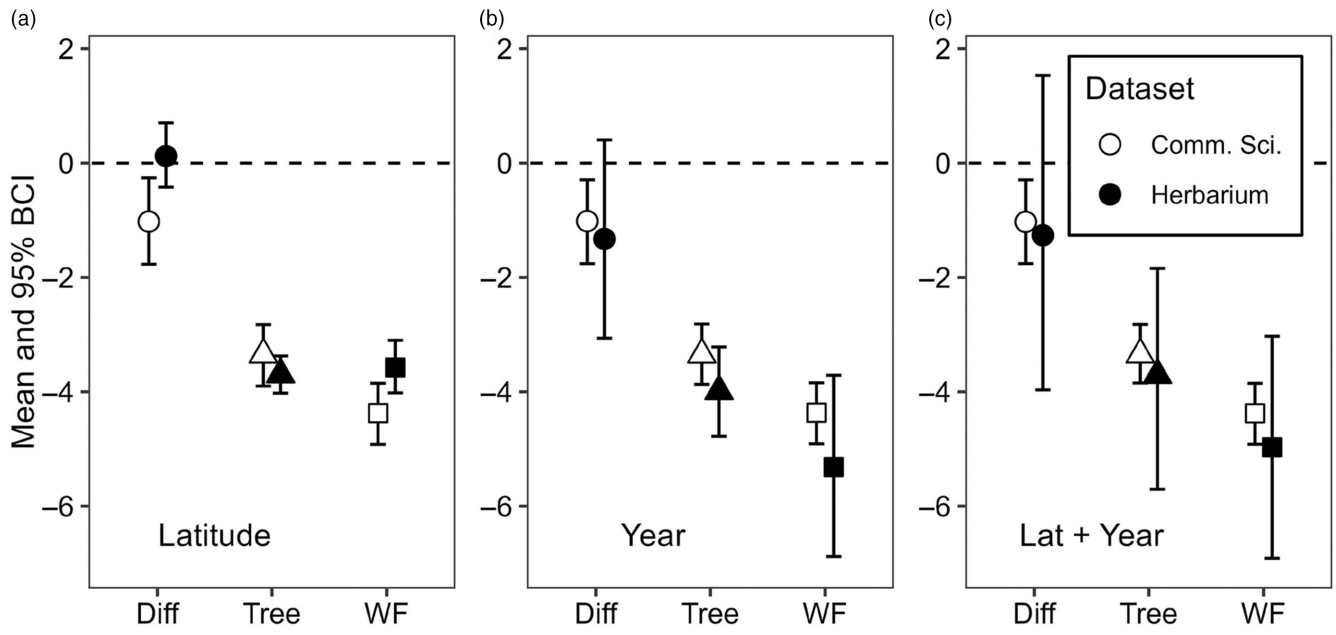

FIGURE 6 Posterior estimated means (points) and 95% Bayesian credible intervals (whiskers) of phenological sensitivities (β, days $°C^{-1}$) of wildflowers (WF, squares), trees (triangles), and the difference between them (circles) using the Community Science (Comm. Sci.; open symbols) and Herbarium (filled symbols) datasets. Panels differ in whether the Herbarium dataset was cropped by (a) latitude, (b) year, or (c) latitude and year to match the spatial and/or temporal extents present in the Community Science dataset. Posterior parameter estimates for latitudinal/temperature bins for versions of the cropped models are provided in Figures S2–S4.

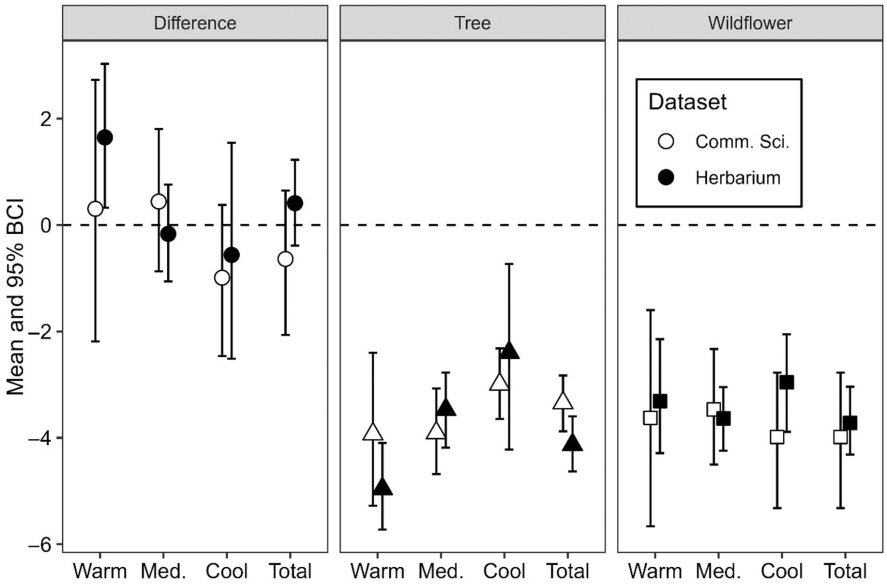

FIGURE 7 Posterior estimated means (points) and 95% Bayesian credible intervals (whiskers) of phenological sensitivities of wildflowers (squares), trees (triangles), and the difference between them (circles) for the Community Science (Comm. Sci.; open symbols) and Herbarium (filled symbols) datasets. Here, both datasets use a flowering phenophase for wildflowers instead of the leaf-out phase used in the original and hybrid models of the community science dataset (Figure 1).

not different from zero in the hybrid model (Figure 1b) to negative (Figure 7). This caused the estimated difference between tree and wildflower sensitivity in the warm bin to move from strongly positive to near zero, although in both cases the credible intervals overlapped zero.

4 | DISCUSSION

In this paper, we presented a series of analyses to test defined hypotheses with the aim of reconciling contradictory results from two recent studies—one using community science (Alecrim et al., 2023), the other using herbarium specimens (Miller et al., 2022)—describing the sensitivity of canopy tree phenology and understory wildflower phenology to warmer spring temperatures. Overall, we find that differences in model structure (H1)—specifically, how data were binned, how spring temperatures were calculated, how models were coded, and what statistical packages were used—appear to be a major factor in the different conclusions reached by the two studies. The hybrid model structure (Equation 3) brought the two results closer together, but discrepancies remained that were not explainable by differences in model structure and analysis. Below we discuss the roles that each of the hypothesized factors (model structure and analysis, species, spatiotemporal extent and phenophase) appear to have played in the difference in results of the two studies.

4.1 | H1: Differences in model structure and analysis

The original herbarium paper result, that trees are more responsive to temperature change than wildflowers, was no longer supported when random slope estimates were included to match the model structure of the original community science paper approach (Figure 1). Merging the model structures also moved the reaction-norm difference described in the community science paper closer to zero, although it remained negative (indicating higher sensitivity in spring wildflowers; Figure 1b).

Surprisingly, though, the datasets yielded different results when data were binned by latitude and temperature. The binned results from the herbarium dataset did not change with respect to overall signal compared to the original herbarium study. Trees were more sensitive than wildflowers in the southern/warm bin, with no meaningful differences between functional groups in the central and northern/cool bins (Miller et al., 2022). In contrast, the results from the community science dataset changed such that estimated differences in sensitivity between wildflowers and trees now overlapped zero in all bins (Figure 1b).

Whether data were binned or not, using the same hybrid model brought the results closer together. Interestingly, the overall (non-binned) signal for the herbarium dataset was less robust to model changes than was the community science dataset, whereas the opposite was true when evaluating the results for latitudinal/thermal bins. This difference between binned and non-binned results may be caused (at least in part) by the similarity of the overall model structure (random slope estimates, temperature window, and statistical software) to what was used in the original community science analysis (Alecrim et al., 2023), and the similarity of the binning to what was used in the original herbarium analysis (Miller et al., 2022). Therefore, the observed changes in herbarium results when non-binned were likely due to the inclusion of random slopes, whereas changes in community science results in the central and northern bins are likely associated with the changes we made to the latitudinal bins (shifting them south to better represent the latitudinal distribution of both datasets; Figure 2a).

Random slopes, which were absent in the original analysis of the herbarium dataset, are typically used when analysing data that involves repeated measurements over time and where individual subjects (or, in this case, species) might be expected to differ in their response to a predictor. They are also helpful when analysing data

with a hierarchical structure, such as when measurements are taken within nested groups (e.g. species within a functional group). In these cases, random slopes can account for within-group variation and improve the accuracy of the analysis (Oberauer, 2022). For the original community science study, a random slope for the interaction between species and site was included to account for the variability in the response to temperature caused by differences between species and geographical locations. The decision to use nonhierarchical, shared slopes in the original herbarium study was based largely on precedent from previous models of phenological sensitivity (e.g. Heberling, McDonough, et al., 2019). This approach did not allow the authors to evaluate species-level sensitivities; it instead placed the focus on community or functional-group signals. Overall, the decision to use random slopes in Bayesian analysis should be based on the specific characteristics of the data and research question at hand, but it's important to note that our analysis here indicates that the decision can affect model results and inference.

4.2 | H2: Differences in species investigated

Differences in species selection did not explain much, if any, of the difference in results between the community science and herbarium studies, suggesting that group-level signals are relatively strong (at least within each dataset). The differences in sensitivities between the datasets remain when looking at only species that were shared across both analyses (6 of 28 species; Figure 4). Species unique to each study also appeared to have sensitivities in line with other members of their functional group ('Unshared' species in Figure 4). When binned by latitude, most species-level responses overlapped between datasets for all three bins (Figure 5, Figure S1), likely as a result of the reduced sample sizes for these comparisons. The consistency with the original analyses of the species-level signals in both datasets suggests that the different species pools for each dataset are not responsible for the original differences between studies (shown in Figure 1a).

4.3 | H3: Differences in spatiotemporal extent

Differences in the spatial, temporal and spatiotemporal extents of the community science and herbarium datasets seemed to explain some of the discrepancy in results. If the herbarium dataset is constrained to only include latitudes and/or years used in the community science paper, the estimated differences between the two datasets are either maintained (Figure 6a, when cropped only by latitude) or reduced (Figure 6b,c, when cropped by year or year and latitude). This was the only hypothesis we tested where the sign changed for the overall signal for one of the datasets: When cropping by year, or by year and latitude (Figure 6b,c), the overall signal from the herbarium dataset moved from positive (indicating a reduction in phenological escape duration under warmer temperatures) to negative (indicating an increase in duration and agreeing with the original conclusions of the community science paper). However, cropping the herbarium dataset by year alone reduced the combined sample size of wildflowers and trees from $n = 2016$ to $n = 67$, a nearly 97% reduction. Cropping by both year and latitude further reduced the sample size to $n = 47$.

When we binned the data, data attrition from cropping the herbarium dataset prevented meaningful statistical inference and, in some cases, prevented us from making any inference at all. For example, cropping the dataset by year alone removed all observations of tree and wildflower phenology in the northern, cool bin, meaning that we were unable to estimate phenological sensitivities (or the difference between them) in that part of their range (Figures S2–S4). Even in the central and southern bins, cropping by year left no more than 19 individual observations in any of the four remaining functional group × bin combinations (Figure S5).

It is also possible that differences in results between the original articles are due in part to changes in phenological sensitivity over time or differences in phenological sensitivity across space (especially at the southern range limits for some of the species studied). Specifically, differences between the two datasets could be partially attributable to non-linear phenological responses to environmental variation which are differently captured by the herbarium and community science datasets. Such differences could imply variation in local adaptation across the plants' ranges (Kharouba & Vellend, 2015; Zettlemoyer & Peterson, 2021) or saturating phenological responses to environmental cues (Iler et al., 2013). But these results could also be artefacts of severe data limitations; analyses of phenological changes are known to be sensitive to sample size, time period and geographic extent (e.g. Bolmgren et al., 2013; Primack et al., 2009).

4.4 | H4: Differences in phenophase

Differences between the two studies were not fully explained by choice of phenophase. The community science dataset, using wildflower flowering instead of leaf-out, still points towards an overall trend of wildflowers being more sensitive to temperature than trees (Figure 7), although there was increased overlap in parameter estimates from the two datasets. The 95% credible intervals for the overall difference between wildflower and tree sensitivity also changed to overlap with zero, meaning we would no longer expect a change in phenological escape duration with warming. When data were binned by latitude and temperature, there were still no differences in community science estimates between the leaf-out and flowering model estimates (other than a small change in statistical significance for wildflower sensitivity in the southern, warm bin).

On one hand, this is not that surprising because early spring wildflowers (and spring ephemeral wildflowers especially) often flower immediately following emergence in spring (Neufeld & Young, 2014). The tight relationship between flowering and emergence for these species is often cited in herbarium studies as a justification for using

the flowering phenophase as a proxy for leaf expansion (Heberling, McDonough, et al., 2019; Lee et al., 2022). On the other hand, however, flowering and leaf-out in temperate tree species have been found to be differently sensitive to changes in spring temperature (Buonaiuto & Wolkovich, 2021), with the order of flowering and leaf bud burst potentially changing with ongoing climate change (Buonaiuto et al., 2021). It is thus not unreasonable to think that a similar dynamic could occur for understory wildflower species, in which case we would have expected there to be a difference in the signal depending on the phenophase used. Regardless, our results suggest that the findings from the community science dataset were generally robust in the choice of phenophase used in analysis.

5 | CONCLUSIONS

Where does this leave us with reconciling the contradictory results in the two original analyses? The community science and herbarium datasets seem to consistently indicate that the phenology of wildflowers in the southern/warmer portion of eastern North American temperate deciduous forests is less sensitive to changes in temperature than that of co-occurring tree species (represented by positive differences in the left-hand panels of Figure 1). This result appears to be relatively robust. In addition, both datasets suggest strikingly similar estimates of the sensitivity of wildflowers and trees to spring temperature at both the functional group level (Figure 1b; 95% Bayesian credible intervals for each dataset generally overlap for each bin × functional group combination) and the species level (Figure 5, Figure S1). Further, phenological activity was consistently found to be earlier in warmer years and warmer locations. That is, for all hypotheses we tested, for all species and both functional groups (wildflowers and trees), and for both the binned and non-binned results, phenological sensitivity was found to be fairly consistent and negative even if the difference between functional groups varied. This largely agrees with previous studies that have found similar parallels when comparing herbarium and community-generated observational datasets (Iwanycki Ahlstrand et al., 2022; Ramirez-Parada et al., 2022; Spellman & Mulder, 2016).

However, full reconciliation of model differences remains elusive. The results associated with both datasets were generally consistent (and consistently different from each other) across the hypotheses we tested. While merging model structure (including how data were binned, how spring temperatures were calculated, how models were coded, and what statistical packages were used) tended to move the results closer together, the datasets still produced results that suggested different ecological trends. Specifically, the community science dataset consistently suggests that wildflowers are generally more sensitive to spring temperature than trees, whereas the herbarium dataset suggests that wildflowers are either as sensitive as or less sensitive than tree species. Remaining differences in results could be attributable to other factors not accounted for in our analysis. For example, it is possible that regional or local microclimatic differences, which may differ with forest characteristics, influence phenology in ways that are also profound but difficult to disentangle (De Frenne et al., 2013; Willems et al., 2021; Zellweger et al., 2020), especially using herbarium and community science datasets that almost never include environmental data collected at the time of sampling.

In sum, our analyses demonstrate that choices of model structure can greatly impact study conclusions, a point that has been raised previously in ecology (e.g. Arnqvist, 2020). In addition, although our results seem to support recent findings that community science data and herbarium collection data yield analogous estimates of phenological sensitivity in North America at the species level (Iwanycki Ahlstrand et al., 2022; Ramirez-Parada et al., 2022; Spellman & Mulder, 2016), they also suggest that inferences about complex ecological mechanisms such as phenological mismatch in species interactions may still be significantly affected by the type and spatiotemporal extent of data. In light of this sensitivity, along with the fact that the effect sizes we have calculated are on the order of only a 1- to 2-day change in phenological escape duration per degree C, ecologists should be cautious about inferring that climate change will lead to population declines in understory wildflowers due to loss of phenological escape. Finally, until community science data repositories grow to match (or at least approach) the coverage present in herbarium collections or until herbarium collections expand to capture the same spatial extent and temporal frequency of community science repositories, we conclude that ecological studies should ideally incorporate both datasets in analysis as much as possible to rule out the possibility that one or the other biases inference.

AUTHOR CONTRIBUTIONS

Benjamin R. Lee and Evelyn F. Alecrim led the statistical analysis. All authors contributed equally to the writing and interpretation of results.

ACKNOWLEDGEMENTS

We thank Sara Kuebbing and Abe Miller-Rushing for help with this paper. We thank the PRISM climate group and the National Phenology Network for making their data freely available for use in this study. This work was supported by the Ontario Ministry of Research, Innovation and Science through an Early Researcher Award to J.R.K.F., a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant to R.D.S., and National Science Foundation grants no. DEB 1936971 to J.M.H., no. DGE 1735087 to T.K.M., and 1936877 to R.B.P. and supporting B.R.L.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/1365-2745.14317.

DATA AVAILABILITY STATEMENT

Datasets used in this analysis are available in their entirety with the original manuscripts that we compare. Data for Miller et al. (https://doi.org/10.1111/1365-2745.14021) can be found on Open Science Framework: https://doi.org/10.17605/OSF.IO/YMHJR (Miller, 2022). R code for Alecrim et al. (https://doi.org/10.1111/1365-2745.14023) can be found on Zenodo: https://doi.org/10.5281/zenodo.6525428 (Alecrim et al., 2022).

ORCID

Benjamin R. Lee https://orcid.org/0000-0002-5256-0515
Evelyn F. Alecrim https://orcid.org/0000-0002-5365-3887
Tara K. Miller https://orcid.org/0000-0002-2827-4578
Jessica R. K. Forrest https://orcid.org/0000-0002-5273-9339
J. Mason Heberling https://orcid.org/0000-0003-0756-5090
Risa D. Sargent https://orcid.org/0000-0001-8736-5712

REFERENCES

Alecrim, E. F., Sargent, R. D., & Forrest, J. R. K. (2022). *Spring-flowering herbs in North American temperate forests advance their phenology more than trees with warming temperatures* [Data set]. Zenodo. https://doi.org/10.5281/zenodo.6525428

Alecrim, E. F., Sargent, R. D., & Forrest, J. R. K. (2023). Higher-latitude spring-flowering herbs advance their phenology more than trees with warming temperatures. *Journal of Ecology*, *111*(1), 156–169. https://doi.org/10.1111/1365-2745.14023

Arnqvist, G. (2020). Mixed models offer no freedom from degrees of freedom. *Trends in Ecology & Evolution*, *35*(4), 329–335. https://doi.org/10.1016/j.tree.2019.12.004

Bolmgren, K., Vanhoenacker, D., & Miller-Rushing, A. J. (2013). One man, 73 years, and 25 species. Evaluating phenological responses using a lifelong study of first flowering dates. *International Journal of Biometeorology*, *57*(3), 367–375. https://doi.org/10.1007/s00484-012-0560-8

Buonaiuto, D. M., Morales-Castilla, I., & Wolkovich, E. M. (2021). Reconciling competing hypotheses regarding flower–leaf sequences in temperate forests for fundamental and global change biology. *New Phytologist*, *229*(3), 1206–1214. https://doi.org/10.1111/nph.16848

Buonaiuto, D. M., & Wolkovich, E. M. (2021). Differences between flower and leaf phenological responses to environmental variation drive shifts in spring phenological sequences of temperate woody plants. *Journal of Ecology*, *109*(8), 2922–2933. https://doi.org/10.1111/1365-2745.13708

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P. C. (2021). Bayesian item response modeling in *R* with brms and *Stan*. *Journal of Statistical Software*, *100*(5), 1–54. https://doi.org/10.18637/jss.v100.i05

Chmielewski, F.-M., & Rötzer, T. (2001). Response of tree phenology to climate change across Europe. *Agricultural and Forest Meteorology*, *108*(2), 101–112. https://doi.org/10.1016/S0168-1923(01)00233-7

De Frenne, P., Rodríguez-Sánchez, F., Coomes, D. A., Baeten, L., Verstraeten, G., Vellend, M., Bernhardt-Römermann, M., Brown, C. D., Brunet, J., Cornelis, J., Decocq, G. M., Dierschke, H., Eriksson, O., Gilliam, F. S., Hédl, R., Heinken, T., Hermy, M., Hommel, P., Jenkins, M. A., … Verheyen, K. (2013). Microclimate moderates plant responses to macroclimate warming. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(46), 18561–18565. https://doi.org/10.1073/pnas.1311190110

Denny, E. G., Gerst, K. L., Miller-Rushing, A. J., Tierney, G. L., Crimmins, T. M., Enquist, C. A. F., Guertin, P., Rosemartin, A. H., Schwartz, M. D., Thomas, K. A., & Weltzin, J. F. (2014). Standardized phenology monitoring methods to track plant and animal activity for science and resource management applications. *International Journal of Biometeorology*, *58*(4), 591–601. https://doi.org/10.1007/s00484-014-0789-5

Ettinger, A. K., Gee, S., & Wolkovich, E. M. (2018). Phenological sequences: How early-season events define those that follow. *American Journal of Botany*, *105*(10), 1771–1780. https://doi.org/10.1002/ajb2.1174

Flynn, D. F. B., & Wolkovich, E. M. (2018). Temperature and photoperiod drive spring phenology across all species in a temperate forest community. *New Phytologist*, *219*(4), 1353–1362. https://doi.org/10.1111/nph.15232

Freimuth, J., Bossdorf, O., Scheepens, J. F., & Willems, F. M. (2022). Climate warming changes synchrony of plants and pollinators. *Proceedings of the Royal Society B: Biological Sciences*, *289*(1971), 20212142. https://doi.org/10.1098/rspb.2021.2142

Fu, Y. H., Campioli, M., Deckmyn, G., & Janssens, I. A. (2012). The impact of winter and spring temperatures on temperate tree budburst dates: Results from an experimental climate manipulation. *PLoS One*, *7*(10), e47324. https://doi.org/10.1371/journal.pone.0047324

Geng, X., Fu, Y. H., Piao, S., Hao, F., De Boeck, H. J., Zhang, X., Chen, S., Guo, Y., Prevéy, J. S., Vitasse, Y., Peñuelas, J., Janssens, I. A., & Stenseth, N. C. (2022). Higher temperature sensitivity of flowering than leaf-out alters the time between phenophases across temperate tree species. *Global Ecology and Biogeography*, *31*(5), 901–911. https://doi.org/10.1111/geb.13463

Gilliam, F. S. (2007). The ecological significance of the herbaceous layer in temperate forest ecosystems. *BioScience*, *57*(10), 845–858. https://doi.org/10.1641/B571007

Heberling, J. M., Cassidy, S. T., Fridley, J. D., & Kalisz, S. (2019). Carbon gain phenologies of spring-flowering perennials in a deciduous forest indicate a novel niche for a widespread invader. *New Phytologist*, *221*(2), 778–788. https://doi.org/10.1111/nph.15404

Heberling, J. M., McDonough MacKenzie, C., Fridley, J. D., Kalisz, S., & Primack, R. B. (2019). Phenological mismatch with trees reduces wildflower carbon budgets. *Ecology Letters*, *22*(4), 616–623. https://doi.org/10.1111/ele.13224

Iler, A. M., CaraDonna, P. J., Forrest, J. R. K., & Post, E. (2021). Demographic consequences of phenological shifts in response to climate change. *Annual Review of Ecology, Evolution, and Systematics*, *52*(1), 221–245. https://doi.org/10.1146/annurev-ecolsys-011921-032939

Iler, A. M., Høye, T. T., Inouye, D. W., & Schmidt, N. M. (2013). Nonlinear phenological responses to climate: Are species approaching their limits of phenological change? *Philosophical Transactions of the Royal Society B*, *368*, 20120489. https://doi.org/10.1098/rstb.2012.0489

Inouye, D. W. (2022). Climate change and phenology. *WIREs Climate Change*, *13*(3), e764. https://doi.org/10.1002/wcc.764

Iwanycki Ahlstrand, N., Primack, R. B., & Tøttrup, A. P. (2022). A comparison of herbarium and citizen science phenology datasets for detecting response of flowering time to climate change in Denmark. *International Journal of Biometeorology*, *66*(5), 849–862. https://doi.org/10.1007/s00484-022-02238-w

Jacques, M. H., Lapointe, L., Rice, K., Montgomery, R. A., Stefanski, A., & Reich, P. B. (2015). Responses of two understory herbs, *Maianthemum canadense* and *Eurybia macrophylla*, to experimental forest warming: Early emergence is the key to enhanced reproductive output. *American Journal of Botany*, *102*(10), 1610–1624. https://doi.org/10.3732/ajb.1500046

Keenan, T. F., Richardson, A. D., & Hufkens, K. (2020). On quantifying the apparent temperature sensitivity of plant phenology. *New Phytologist*, *225*(2), 1033–1040. https://doi.org/10.1111/nph.16114

Kharouba, H. M., Ehrlén, J., Gelman, A., Bolmgren, K., Allen, J. M., Travers, S. E., & Wolkovich, E. M. (2018). Global shifts in the phenological synchrony of species interactions over recent decades. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(20), 5211–5216. https://doi.org/10.1073/pnas.1714511115

Kharouba, H. M., & Vellend, M. (2015). Flowering time of butterfly nectar food plants is more sensitive to temperature than the timing of butterfly adult flight. *Journal of Animal Ecology*, *84*(5), 1311–1321. https://doi.org/10.1111/1365-2656.12373

Kharouba, H. M., & Wolkovich, E. M. (2023). Lack of evidence for the match-mismatch hypothesis across terrestrial trophic interactions. *Ecology Letters*, *26*(6), 955–964. https://doi.org/10.1111/ele.14185

Kim, J. H., Hwang, T., Yang, Y., Schaaf, C. L., Boose, E., & Munger, J. W. (2018). Warming-induced earlier Greenup leads to reduced stream discharge in a temperate mixed forest catchment. *Journal of Geophysical Research: Biogeosciences*, *123*(6), 1960–1975. https://doi.org/10.1029/2018JG004438

Kopp, C. W., Neto-Bradley, B. M., Lipsen, L. P. J., Sandhar, J., & Smith, S. (2020). Herbarium records indicate variation in bloom-time sensitivity to temperature across a geographically diverse region. *International Journal of Biometeorology*, *64*(5), 873–880. https://doi.org/10.1007/s00484-020-01877-1

Kudo, G., & Ida, T. Y. (2013). Early onset of spring increases the phenological mismatch between plants and pollinators. *Ecology*, *94*(10), 2311–2320. https://doi.org/10.1890/12-2003.1

Kudo, G., Nishikawa, Y., Kasagi, T., & Kosuge, S. (2004). Does seed production of spring ephemerals decrease when spring comes early? Seed-set of spring ephemerals. *Ecological Research*, *19*(2), 255–259. https://doi.org/10.1111/j.1440-1703.2003.00630.x

Lapointe, L. (2001). How phenology influences physiology in deciduous forest spring ephemerals. *Physiologia Plantarum*, *113*(2), 151–157. https://doi.org/10.1034/j.1399-3054.2001.1130201.x

Lee, B. R., & Ibáñez, I. (2021a). Improved phenological escape can help temperate tree seedlings maintain demographic performance under climate change conditions. *Global Change Biology*, *27*, 3883–3897. https://doi.org/10.1111/gcb.15678

Lee, B. R., & Ibáñez, I. (2021b). Spring phenological escape is critical for the survival of temperate tree seedlings. *Functional Ecology*, *35*, 1848–1861. https://doi.org/10.1111/1365-2435.13821

Lee, B. R., Miller, T. K., Rosche, C., Yang, Y., Heberling, J. M., Kuebbing, S. E., & Primack, R. B. (2022). Wildflower phenological escape differs by continent and spring temperature. *Nature Communications*, *13*(1), 7157. https://doi.org/10.1038/s41467-022-34936-9

Miller, T. K. (2022). *Phenological mismatch among forest plants* [Data set]. OSF. https://doi.org/10.17605/OSF.IO/YMHJR

Miller, T. K., Heberling, J. M., Kuebbing, S. E., & Primack, R. B. (2022). Warming temperatures are linked to phenological mismatch among native and non-native trees, shrubs, and wildflowers across the Eastern Deciduous Forest of North America. *Journal of Ecology*, *111*, 356–371. https://doi.org/10.1111/1365-2745.14021

Monnahan, C. C., Thorson, J. T., & Branch, T. A. (2017). Faster estimation of Bayesian models in ecology using Hamiltonian Monte Carlo. *Methods in Ecology and Evolution*, *8*, 339–348. https://doi.org/10.1111/2041-210X.12681

Neufeld, H. S., & Young, D. R. (2014). Ecophysiology of the herbaceous layer in temperate deciduous forests. In F. S. Gilliam (Ed.), *The herbaceous layer in forests of eastern North America* (pp. 35–95). Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199837656.003.0003

Oberauer, K. (2022). The importance of random slopes in mixed models for Bayesian hypothesis testing. *Psychological Science*, *33*(4), 648–665. https://doi.org/10.1177/09567976211046884

Panchen, Z. A., Primack, R. B., Nordt, B., Ellwood, E. R., Stevens, A., Renner, S. S., Willis, C. G., Fahey, R., Whittemore, A., Du, Y., & Davis, C. C. (2014). Leaf out times of temperate woody plants are related to phylogeny, deciduousness, growth habit and wood anatomy. *New Phytologist*, *203*(4), 1208–1219. https://doi.org/10.1111/nph.12892

Parmesan, C. (2007). Influences of species, latitudes and methodologies on estimates of phenological response to global warming. *Global Change Biology*, *13*(9), 1860–1872. https://doi.org/10.1111/j.1365-2486.2007.01404.x

Peaucelle, M., Janssens, I. A., Stocker, B. D., Descals Ferrando, A., Fu, Y. H., Molowny-Horas, R., Ciais, P., & Peñuelas, J. (2019). Spatial variance of spring phenology in temperate deciduous forests is constrained by background climatic conditions. *Nature Communications*, *10*(1), 5388. https://doi.org/10.1038/s41467-019-13365-1

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., & Zhu, X. (2019). Plant phenology and global climate change: Current progresses and challenges. *Global Change Biology*, *25*(6), 1922–1940. https://doi.org/10.1111/gcb.14619

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, Vienna, 20–22 March 2003, 1–10.

Polgar, C., Gallinat, A., & Primack, R. B. (2014). Drivers of leaf-out phenology and their implications for species invasions: Insights from Thoreau's Concord. *New Phytologist*, *202*(1), 106–115. https://doi.org/10.1111/nph.12647

Primack, R. B., Ibáñez, I., Higuchi, H., Lee, S. D., Miller-Rushing, A. J., Wilson, A. M., & Silander, J. A. (2009). Spatial and interspecific variability in phenological responses to warming temperatures. *Biological Conservation*, *142*(11), 2569–2577. https://doi.org/10.1016/j.biocon.2009.06.003

Rafferty, N. E., & Nabity, P. D. (2017). A global test for phylogenetic signal in shifts in flowering time under climate change. *Journal of Ecology*, *105*(3), 627–633. https://doi.org/10.1111/1365-2745.12701

Ramirez-Parada, T. H., Park, I. W., & Mazer, S. J. (2022). Herbarium specimens provide reliable estimates of phenological responses to climate at unparalleled taxonomic and spatiotemporal scales. *Ecography*, *2022*(10), e06173. https://doi.org/10.1111/ecog.06173

Ren, P., Néron, V., Rossi, S., Liang, E., Bouchard, M., & Deslauriers, A. (2020). Warming counteracts defoliation-induced mismatch by increasing herbivore-plant phenological synchrony. *Global Change Biology*, *26*(4), 2072–2080. https://doi.org/10.1111/gcb.14991

Renner, S. S., & Zohner, C. M. (2018). Climate change and phenological mismatch in trophic interactions among plants, insects, and vertebrates. *Annual Review of Ecology, Evolution, and Systematics*, *49*(1), 165–182. https://doi.org/10.1146/annurev-ecolsys-110617-062535

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., & Toomey, M. (2013). Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agricultural and Forest Meteorology*, *169*, 156–173. https://doi.org/10.1016/j.agrformet.2012.09.012

Samplonius, J. M., Atkinson, A., Hassall, C., Keogan, K., Thackeray, S. J., Assmann, J. J., Burgess, M. D., Johansson, J., Macphie, K. H., Pearce-Higgins, J. W., Simmonds, E. G., Varpe, Ø., Weir, J. C., Childs, D. Z., Cole, E. F., Daunt, F., Hart, T., Lewis, O. T., Pettorelli, N., … Phillimore, A. B. (2021). Strengthening the evidence base for temperature-mediated phenological asynchrony and its impacts. *Nature Ecology & Evolution*, *5*(2), 155–164. https://doi.org/10.1038/s41559-020-01357-0

Schmidt, N. M., Kankaanpää, T., Tiusanen, M., Reneerkens, J., Versluijs, T. S. L., Hansen, L. H., Hansen, J., Gerlich, H. S., Høye, T. T., Cirtwill, A. R., Zhemchuzhnikov, M. K., Peña-Aguilera, P., & Roslin, T. (2023). Little directional change in the timing of Arctic spring phenology over the past 25 years. *Current Biology*, *33*(15), 3244–3249.e3. https://doi.org/10.1016/j.cub.2023.06.038

Sevenello, M., Sargent, R. D., & Forrest, J. R. K. (2020). Spring wildflower phenology and pollinator activity respond similarly to climatic variation in an eastern hardwood forest. *Oecologia*, *193*(2), 475–488. https://doi.org/10.1007/s00442-020-04670-4

Simmonds, E. G., Cole, E. F., Sheldon, B. C., & Coulson, T. (2020). Phenological asynchrony: A ticking time-bomb for seemingly stable populations? *Ecology Letters*, *23*(12), 1766–1775. https://doi.org/10.1111/ele.13603

Spellman, K. V., & Mulder, C. P. H. (2016). Validating herbarium-based phenology models using citizen-science data. *BioScience*, *66*(10), 897–906. https://doi.org/10.1093/biosci/biw116

Spicer, M. E., Mellor, H., & Carson, W. P. (2020). Seeing beyond the trees: A comparison of tropical and temperate plant growth forms and their vertical distribution. *Ecology*, *101*(4), 1–9. https://doi.org/10.1002/ecy.2974

Spicer, M. E., Radhamoni, H. V. N., Duguid, M. C., Queenborough, S. A., & Comita, L. S. (2022). Herbaceous plant diversity in forest ecosystems: Patterns, mechanisms, and threats. *Plant Ecology*, *223*(2), 117–129. https://doi.org/10.1007/s11258-021-01202-9

Su, Y.-S., & Yajima, M. (2015). Package "*R2jags*" [Computer software]. http://cran.r-project.org/package=R2jags

Thomson, J. D. (2019). Progressive deterioration of pollination service detected in a 17-year study vanishes in a 26-year study. *New Phytologist*, *224*(3), 1151–1159. https://doi.org/10.1111/nph.16078

Wang, X., Morin, X., Zhang, J., Chen, G., Mao, L., Chen, Y., Song, Z., Du, Y., & Ma, K. (2023). Geographical patterns and determinants in plant reproductive phenology duration. *Frontiers in Plant Science*, *14*, 1199316. https://doi.org/10.3389/fpls.2023.1199316

Wann, G. T., Aldridge, C. L., Seglund, A. E., Oyler-McCance, S. J., Kondratieff, B. C., & Braun, C. E. (2019). Mismatches between breeding phenology and resource abundance of resident alpine ptarmigan negatively affect chick survival. *Ecology and Evolution*, *9*(12), 7200–7212. https://doi.org/10.1002/ece3.5290

Willems, F. M., Scheepens, J. F., Ammer, C., Block, S., Bucharova, A., Schall, P., Sehrt, M., & Bossdorf, O. (2021). Spring understory herbs flower later in intensively managed forests. *Ecological Applications*, *31*(5), e02332. https://doi.org/10.1002/eap.2332

Willems, F. M., Scheepens, J. F., & Bossdorf, O. (2022). Forest wildflowers bloom earlier as Europe warms: Lessons from herbaria and spatial modelling. *New Phytologist*, *235*(1), 52–65. https://doi.org/10.1111/nph.18124

Yancy, A. J., Lee, B. R., Kuebbing, S. E., Neufeld, H. S., Spicer, M. E., & Heberling, J. M. (2023). Evaluating the definition and distribution of spring ephemeral wildflowers in eastern North America. *bioRxiv* https://doi.org/10.1101/2023.10.04.560873

Zellweger, F., De Frenne, P., Lenoir, J., Vangansbeke, P., Verheyen, K., Bernhardt-Römermann, M., Baeten, L., Hédl, R., Berki, I., Brunet, J., Van Calster, H., Chudomelová, M., Decocq, G., Dirnböck, T., Durak, T., Heinken, T., Jaroszewicz, B., Kopecký, M., Máliš, F., … Coomes, D. (2020). Forest microclimate dynamics drive plant responses to warming. *Science*, *368*(6492), 772–775. https://doi.org/10.1126/science.aba6880

Zettlemoyer, M. A., & Peterson, M. L. (2021). Does phenological plasticity help or hinder range shifts under climate change? *Frontiers in Ecology and Evolution*, *9*, 689192. https://doi.org/10.3389/fevo.2021.689192

Zohner, C. M., Mo, L., & Renner, S. S. (2018). Global warming reduces leaf-out and flowering synchrony among individuals. *eLife*, *7*, e40214. https://doi.org/10.7554/eLife.40214

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Binned posterior estimates of the mean (points) and 95% Bayesian Credible Intervals (BCI, whiskers) of phenological sensitivity for the 13 wildflower (top panel) and nine canopy tree species (bottom panel) unique to the Community Science (open symbols) and Herbarium (closed symbols) datasets.

Figure S2. Posterior means (points) and 95% Bayesian credible intervals (whiskers) for the estimated difference in phenological sensitivity (circles) between wildflower (squares) and canopy tree (triangles) sensitivity when the Herbarium dataset (closed symbols) is cropped to the same spatial extent as the Community Science dataset (open symbols).

Figure S3. Posterior means (points) and 95% Bayesian credible intervals (whiskers) for the estimated difference in phenological sensitivity (circles) between wildflower (squares) and canopy tree (triangles) sensitivity when the Herbarium dataset (closed symbols) is cropped to the same temporal extent as the Community Science dataset (open symbols).

Figure S4. Posterior means (points) and 95% Bayesian credible intervals (whiskers) for the estimated difference in phenological sensitivity (circles) between wildflower (squares) and canopy tree (triangles) sensitivity when the Herbarium dataset (closed symbols) is cropped to the same spatial and temporal extent as the Community Science dataset (open symbols).

Figure S5. Distributions of observations of wildflower (WF, yellow volumes) and tree (blue volumes) phenology in (A) the original Community Science dataset and (B) the Herbarium dataset when cropped by latitude, year, or latitude and year to match the Community Science dataset (as described in H3 Methods).

How to cite this article: Lee, B. R., Alecrim, E. F., Miller, T. K., Forrest, J. R. K., Heberling, J. M., Primack, R. B., & Sargent, R. D. (2024). Phenological mismatch between trees and wildflowers: Reconciling divergent findings in two recent analyses. *Journal of Ecology*, *112*, 1184–1199. https://doi.org/10.1111/1365-2745.14317