

# Sensitivity of Spring Phenology to Warming Across Temporal and Spatial Climate Gradients in Two Independent Databases

Benjamin I. Cook,<sup>1,2\*</sup> Elizabeth M. Wolkovich,<sup>3,4</sup> T. Jonathan Davies,<sup>5</sup> Toby R. Ault,<sup>6</sup> Julio L. Betancourt,<sup>7</sup> Jenica M. Allen,<sup>8</sup> Kjell Bolmgren,<sup>9,20</sup> Elsa E. Cleland,<sup>3</sup> Theresa M. Crimmins,<sup>10</sup> Nathan J. B. Kraft,<sup>4</sup> Lesley T. Lancaster,<sup>11</sup> Susan J. Mazer,<sup>12</sup> Gregory J. McCabe,<sup>13</sup> Brian J. McGill,<sup>14</sup> Camille Parmesan,<sup>15,16</sup> Stephanie Pau,<sup>11</sup> James Regetz,<sup>11</sup> Nicolas Salamin,<sup>17</sup> Mark D. Schwartz,<sup>18</sup> and Steven E. Travers<sup>19</sup>

<sup>1</sup>NASA Goddard Institute for Space Studies, New York, New York, USA; <sup>2</sup>Ocean and Climate Physics, Lamont-Doherty Earth Observatory, Palisades, New York, USA; <sup>3</sup>Division of Biological Sciences, University of California-San Diego, La Jolla, California, USA; <sup>4</sup>Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, Canada; <sup>5</sup>Department of Biology, McGill University, Montreal, Quebec, Canada; <sup>6</sup>National Center for Atmospheric Research, Boulder, Colorado, USA; <sup>7</sup>U.S. Geological Survey, Tucson, Arizona, USA; <sup>8</sup>Department of Ecology & Evolutionary Biology, University of Connecticut, Storrs, Connecticut, USA; <sup>9</sup>Department of Biology, Theoretical Population Ecology and Evolution Group, Lund University, Lund, Sweden; <sup>10</sup>USA National Phenology Network, Tucson, Arizona, USA; <sup>11</sup>National Center for Ecological Analysis and Synthesis, Santa Barbara, California, USA; <sup>12</sup>Department of Ecology, Evolution and Marine Biology, University of California-Santa Barbara, Santa Barbara, California, USA; <sup>13</sup>U.S. Geological Survey, Denver Federal Center, Denver, Colorado, USA; <sup>14</sup>Ecology and Environmental Science, University of Maine, Orono, Maine, USA; <sup>15</sup>Integrative Biology, University of Texas, Austin, Texas, USA; <sup>16</sup>Marine Sciences Institute Portland Square, University of Plymouth, Plymouth, UK; <sup>17</sup>Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland; <sup>18</sup>Department of Geography, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin, USA; <sup>19</sup>Department of Biological Sciences, North Dakota State University, Fargo, North Dakota, USA; <sup>20</sup>Swedish National Phenology Network, Swedish University of Agricultural Sciences, Ås, Sweden

Received 30 December 2011; accepted 3 July 2012;  
published online 5 September 2012

**Electronic supplementary material:** The online version of this article (doi:10.1007/s10021-012-9584-5) contains supplementary material, which is available to authorized users.

**Author Contributions:** All authors contributed to the study design and offered comments on this manuscript as part of the “Forecasting Phenology” working group funded by the National Center for Ecological Analysis and Synthesis. Author Cook conceived of and designed the study, analyzed the data, and wrote the paper. Wolkovich and Davies contributed significantly to the refinement of the ideas and analyses and assisted with the writing. Ault contributed significantly to the analyses, including processing of the GHCN climate data. Betancourt contributed significantly to the writing and study design. The first five authors are listed in order of their contributions; all other authors are listed alphabetically.

\*Corresponding author; e-mail: benjamin.i.cook@nasa.gov

## ABSTRACT

Disparate ecological datasets are often organized into databases post hoc and then analyzed and interpreted in ways that may diverge from the purposes of the original data collections. Few studies, however, have attempted to quantify how biases inherent in these data (for example, species richness, replication, climate) affect their suitability for addressing broad scientific questions, especially in under-represented systems (for example, deserts, tropical forests) and wild communities. Here, we quantitatively compare the sensitivity of species first flowering and leafing dates to spring warmth in two phenological databases from the Northern Hemisphere. One—PEP725—has high replication within and across sites, but has low species diver-

sity and spans a limited climate gradient. The other—NECTAR—includes many more species and a wider range of climates, but has fewer sites and low replication of species across sites. PEP725, despite low species diversity and relatively low seasonality, accurately captures the magnitude and seasonality of warming responses at climatically similar NECTAR sites, with most species showing earlier phenological events in response to warming. In NECTAR, the prevalence of temperature responders significantly declines with increasing mean annual temperature, a pattern that cannot be detected across the limited climate gradient spanned by the PEP725 flowering and leafing data. Our results showcase broad areas of agreement between

the two databases, despite significant differences in species richness and geographic coverage, while also noting areas where including data across broader climate gradients may provide added value. Such comparisons help to identify gaps in our observations and knowledge base that can be addressed by ongoing monitoring and research efforts. Resolving these issues will be critical for improving predictions in understudied and under-sampled systems outside of the temperature seasonal mid-latitudes.

**Key words:** phenology; climate responders; NECTAR; PEP725; sensitivity; climate change.

## INTRODUCTION

Developing robust predictions of ecological responses to climate change requires long-term data from diverse species, ecosystems, and climates. Most of the recent advances, however, are based on analyses of large observational databases that primarily represent mesic, temperature seasonal systems in the mid-latitudes (Parmesan and Yohe 2003; Root and others 2003; Rosenzweig and others 2008; Schwartz and others 2006; Schwartz and Reiter 2000; Walther and others 2002). Although analyses of these databases are still quite valuable, any results will be heavily skewed by the disproportionate sampling of temperate forests and meadows that are ubiquitous throughout the mid-latitudes. This may mean that any conclusions and interpretations cannot be generalized to under-represented and poorly sampled systems, such as deserts, semi-arid grasslands, and tropical forests.

One such database that has figured prominently in these studies is the Pan European Phenological Database (PEP725, built upon the earlier COST725 database: PEP725 2010; Koch and others 2010a, b; Menzel and others 2006). The PEP725 network represents an effort to standardize and integrate, into a common database, wild and cultivated species observations from a suite of phenological observation networks distributed across Europe, including clonal species from the International Phenological Gardens (IPG) (Menzel 2000). Analyses of PEP725 have detected shifts in spring and fall phenology due to climate change and variability (Menzel and others 2006; Scheffinger and others 2002, 2003). Notably, studies using PEP725 were heavily cited by the Working Group II report on climate change impacts for the Intergovern-

mental Panel on Climate Change Fourth Assessment Report (Parry 2007).

PEP725, however, was not developed with the specific intention of serving such global analyses, but rather to simply “promote and facilitate phenological research by delivering a pan European phenological database with an open, unrestricted data access for science, research and education” (PEP725 2010). And although the value of PEP725 is unequivocal, it has many of the limitations and biases alluded to earlier, including a relatively small geographic area (continental Europe), narrow climate gradient, and low species richness compared to many other natural communities and climate zones. Understanding how well inferences from PEP725 can be generalized to wild plant communities and systems outside of PEP725 may help to identify critical gaps in our observations and knowledge and to inform the efficacy of current monitoring and analysis efforts.

To investigate these issues, we developed a new phenological database that unites a suite of twelve phenological datasets independent of PEP725, synthesized under the name Network of Ecological and Climatological Timings Across Regions (NECTAR) (a detailed description of NECTAR can be found in the Electronic Supplementary Material; ESM). NECTAR draws from datasets across North America and Europe, has much higher species richness than PEP725, and includes sites outside of the relatively narrow climate space occupied by PEP725. We do not single out PEP725 because of any specific critique or criticism; rather, we draw from this database as an example of a network widely used and cited in global change analyses (see previous references). We focus our analysis on two major axes which have been shown to affect spring phenological sensitivities to temperature:

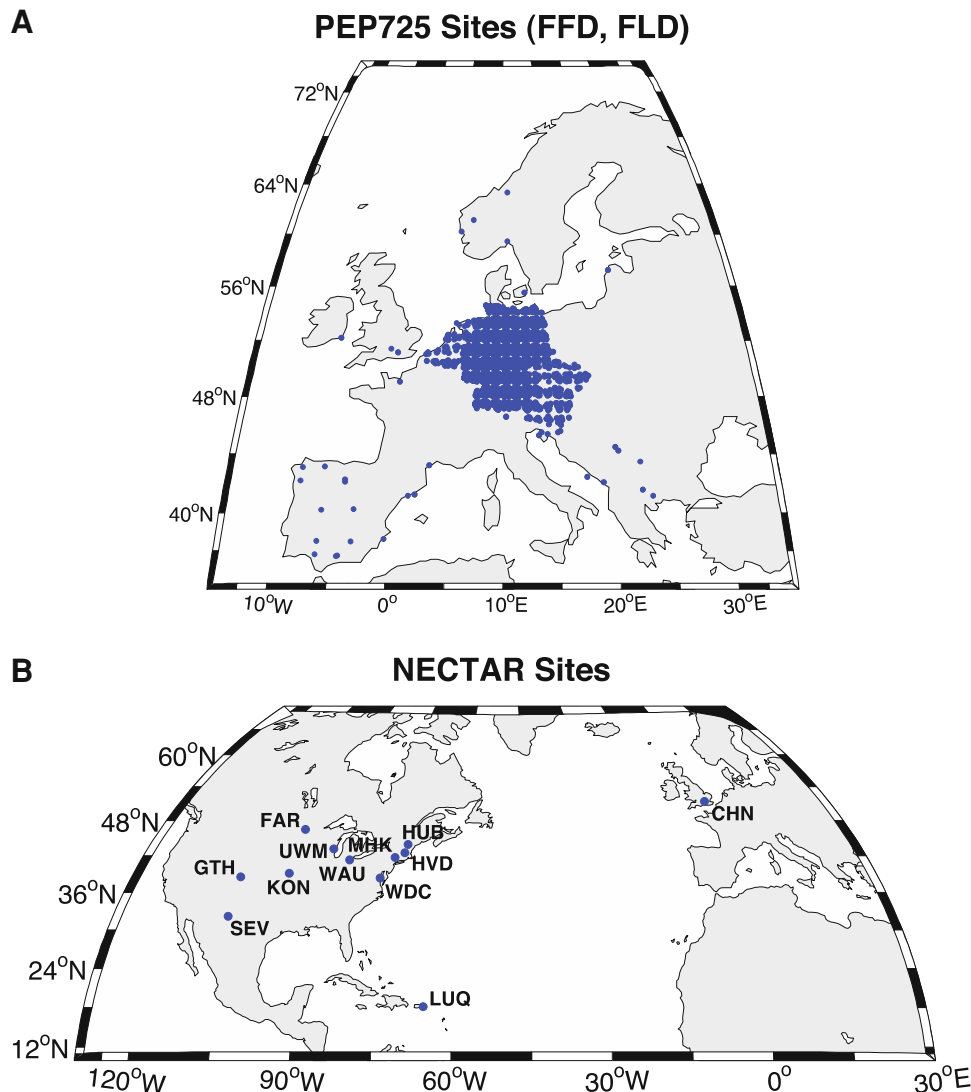
the growing season (for example, Menzel and others 2006; Sherry and others 2007) and climate space (for example, Olsson and Ågren 2002; Rötzer and Chmielewski 2001). Specifically, we compare the sensitivity of species level first flowering (FFD) and leafing (FLD) dates in each database to spring warming (1) across the growing season and (2) across spatial climate gradients.

## MATERIALS AND METHODS

### Phenological Data

The PEP725 and NECTAR databases include thousands of phenological time series, representing multiple sites, species, and phenological phases. We selected data from PEP725 that documented events similar to FFD (ESM Table 1)

and FLD (ESM Table 2) to maximize comparability with the NECTAR database, which is composed almost entirely of FFD and FLD time series. These phases in PEP725 primarily represent observations of first observable open flowers and first visible leaf stalks of the season. Site locations with FLD and FFD data in PEP725 span continental Europe, but are heavily biased towards the north-central region of the continent, especially Germany, Austria, and Switzerland (Figure 1a). For FFD, there were 91,808 individual time series (median number of observations per time series per site = 20), including 42 unique taxa at 5,562 sites. For FLD, there were 24,700 individual time series (median number of observations per time series per site = 20), including 14 unique taxa at 5,074 sites. Clonal records from the International Phenological Gardens make up only a small



**Figure 1.** Station and dataset locations for FFD and FLD data in the **A** PEP725 database and **B** new NECTAR database. NECTAR includes many more species and covers a wider geographic range and temporal span than PEP725. PEP725 has higher replication of individual phenophases across sites and includes many more sites.

fraction of the data used from the PEP725 database: 0.20% (187 records) of the FFD data and 1.05% (259 records) of the FLD data. The FFD and FLD data we use were collected from 1951 through 2008.

The NECTAR sites (Table 1) cover diverse climates including temperature seasonal mid-latitude sites similar to those found in PEP725 (CHN, HUB, HVD, MHK, UWM, WDC, WAU), continental grassland sites (FAR, KON), one semi-arid site (SEV), one alpine site (GTH), and one tropical site (LUQ) (Figure 1b). All datasets comprise observations of natural, non-cultivated populations and all sites are located in North America except for one European site (CHN). These data represent both amateur observations and focused monitoring efforts. All data were either recorded as FFD or FLD, or converted to first-events. The UWM FLD dates represent the dates when the 50th percentile of full spring leaf flush was achieved and the LUQ FFD are based on litter basket collections. We used a total of 1,475 phenological time series from NECTAR, representing 1,200 unique taxa. Further details on the NECTAR database, including time intervals over which the data were collected, are contained in the ESM. We did not conduct any outlier screening of the phenological data prior to our analyses. However, repeating our main analyses with outlier screenings based on 3-sigma and 2-sigma standard deviations from the mean yielded similar results, indicating that any outliers have a negligible effect on our main results and conclusions (ESM).

## Climate Data

Many studies investigating the sensitivity of phenological events to climate use seasonal or monthly mean climate predictors (Willis and others 2008). This approach is analytically tractable and takes advantage of the many high-quality and homogenized monthly climate data products available. However, these analyses overlook the importance of integrative forcing for phenology, such as growing degree day or heating summations (Leon and others 2001; McMaster and Wilhelm 1997). To specifically address this gap, we restricted our climate data choices to daily data (minimum and maximum temperatures, and precipitation) so as to account for integrative forcing (that is, growing degree days).

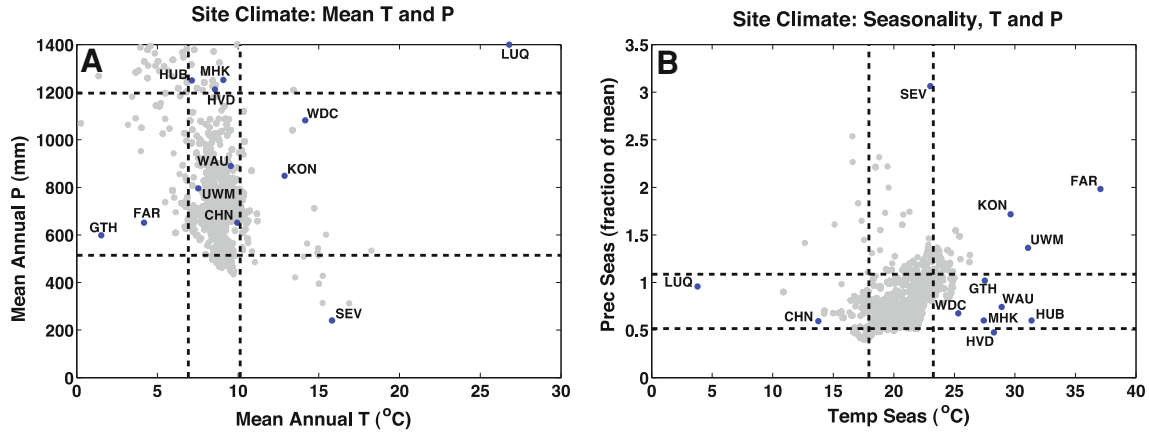
Climate data for comparison with the PEP725 database were taken from the European Climate Assessment (ECA) (Haylock and others 2008; Hofstra and others 2009; Klok and Klein Tank 2009). In the ECA data, daily temperature and precipitation for 1950–2009 are interpolated from meteorological stations to a continuous 0.25° grid. We matched the PEP725 phenology records to daily ECA temperature and precipitation at the geographically closest grid cell.

For NECTAR, we used daily temperature and precipitation data from meteorological stations in the Global Historical Climatology Network (GHCN) (Peterson and Vose 1997). We based our selection of the appropriate meteorological station data on (1) proximity to the NECTAR sites and (2) temporal overlap with the NECTAR phenological records.

**Table 1.** Phenological Datasets Assembled into the NECTAR Database

| Dataset                 | Code | Phase  | # spp | Median <i>n</i> | References   |
|-------------------------|------|--------|-------|-----------------|--|
| Chinnor, UK             | CHN  | FFD    | 384   | 37              | Fitter and Fitter (2002)                                 |
| Fargo, ND, USA          | FAR  | FFD    | 214   | 12.5            | Dunnell and Travers (2011),<br>Travers and others (2009) |
| Gothic, CO, USA         | GTH  | FFD    | 79    | 30              | Inouye (2008)  |
| Hubbard Brook, MA, USA  | HUB  | FLD    | 3     | 20              | Richardson and others (2006)                             |
| Harvard Forest, MA, USA | HVD  | FFD    | 33    | 12              | Richardson and others (2006)                             |
| Harvard Forest, MA, USA | HVD  | FLD    | 24    | 12              | Richardson and others (2006)                             |
| Konza, KS, USA          | KON  | FFD    | 152   | 8               | Wright (2001)  |
| Luquillo, PR, USA       | LUQ  | FFDBSK | 63    | 16              | Zimmerman and others (2007)                              |
| Mohonk, NY, USA         | MHK  | FFD    | 18    | 35              | Cook and others (2008)                                   |
| Sevilleta, NM, USA      | SEV  | FLD    | 71    | 9               | Wetherill (2000)   |
| UW-Milwaukee, WI, USA   | UWM  | FLD50  | 25    | 10              | Schwartz and Hanes (2010)                                |
| Washington, DC, USA     | WDC  | FFD    | 385   | 17              | Abu-Asab and others (2001)                               |
| Wauseon, OH, USA        | WAU  | FLD    | 24    | 11              | Smith (1915)   |

*All data are observations of wild, non-cultivated, and non-clonal populations. Locations for each dataset are shown in Figure 1. FFD, first flowering date; FLD, first leafing date; FFDBSK, first flowering date from litter basket collection; FLD50, fiftieth percentile of full leaf flush. Median *n* refers to the median number of observations (years) of all phenological time series at that site.*



**Figure 2.** We defined the climate characteristics for each site in NECTAR and PEP725 based on the mean state (**A**) and the seasonality (**B**), using the same temperature and precipitation data used to develop our climate predictors. Mean state is defined as mean annual temperature (°C) and mean annual precipitation (mm). Temperature seasonality is the mean annual temperature range (°C) and precipitation seasonality is the mean annual precipitation range (wettest minus driest pentads) divided by the mean annual precipitation (expressed as fraction of the annual mean). PEP725 sites are in *gray*, with 5th and 95th percentiles marked by the *dashed lines*. NECTAR sites are in *blue*, marked with the associated three letter code (see Table 1).

Details on the meteorological data used in the NECTAR analyses are contained in the ESM.

Using these temperature and precipitation data, we can characterize the climate of the PEP725 and NECTAR sites in terms of the mean annual climate (Figure 2a) and seasonality (Figure 2b). Overall, sites in the PEP725 database (gray dots) occupy a much narrower climatic range relative to NECTAR. The PEP725 sites have mean climates typical of the temperate mid-latitudes, and these sites are well represented in the NECTAR database. But NECTAR also includes warm tropical and semi-arid sites (LUQ, SEV) and much colder continental sites (FAR, GTH) (Figure 2a). From a seasonality perspective, the difference between the two databases is even more striking. Temperature and precipitation seasonality in PEP725 is tightly constrained, and every NECTAR site has a temperature or precipitation seasonality that falls outside the 5–95% range from PEP725 (Figure 2b). Both the reduced seasonality and the limited temperature range for the PEP725 sites likely reflect the moderating influence of westerly winds originating from the Atlantic during all seasons, a dominant feature of mainland European climate.

## Analyses

For each site, we identified a suite of climate predictors based on 3-month moving window growing degree day (GDD) summations. First, we composited the daily temperature data for each year into 3-month moving windows, beginning in the

preceding fall and running through the end of the concurrent year (for example, August–September–October, September–October–November, October–November–December, November–December–January, December–January–February, and so on). Daily temperatures in each 3-month window were converted to daily GDDs:

$$\text{GDD} = \max(T_{\text{mean}} - \text{GDD}_{\text{thresh}}, 0) \quad (1)$$

where  $T_{\text{mean}}$  is daily mean temperature and  $\text{GDD}_{\text{thresh}}$  is the temperature threshold for a day to qualify as a GDD (in our case, 0°C). The daily GDD values were then summed for each 3-month window for each year:

$$\text{GDD}_{\text{sum}}(\text{years}) = \sum \text{GDD} \quad (2)$$

where  $\text{GDD}_{\text{sum}}$  are the GDD summations calculated over all days for a specific 3-month window for each year that data were available. Second, we standardized each 3-month  $\text{GDD}_{\text{sum}}$  time series to zero mean and unit standard deviation:

$$Z_t = \frac{\text{GDD}_{\text{sum}} - \text{mean}(\text{GDD}_{\text{sum}})}{\text{std}(\text{GDD}_{\text{sum}})} \quad (3)$$

Each 3-month  $\text{GDD}_{\text{sum}}$  was standardized separately, using the mean and interannual standard deviation calculated across the entire temporal length of the  $\text{GDD}_{\text{sum}}$  time series. This ensures all the climate predictors at a given site are standardized uniformly regardless of potential missing values or varying time spans in the associated phenological records. The standardization step



normalizes the variance in climate across the year and across sites, which is important because (1) climate in winter and spring tends to be more variable from year to year than summer or fall and (2) inter-annual variability in climate varies across sites. With the standardization of the climate predictors, variations in the species sensitivities across seasons and sites are thus disentangled from variations in climate variability.

Because we used a single unified temperature dataset (ECA) for the PEP725 analysis, we were able to standardize temperatures across all sites over a common time interval (1950–2009). This was impossible with the NECTAR analysis because of the varying lengths, serial completion, and temporal overlap of the phenological time series and accompanying meteorological data. To test the robustness of our results to the standardization periods used, we repeated our main analyses, but restricted the temporal range to the last 30 years of the twentieth century (1970–1999), a common interval for 8 of the NECTAR sites (CHN, GTH, HUB, HVD, LUQ, MHK, and WDC). Results for these sites were similar between this analysis and the main analysis using all available data, indicating our analysis is largely insensitive to the choice of standardization period (ESM).

We used these standardized, seasonally composited temperature summations as predictor variables for determining temperature sensitivities of the phenological time series in NECTAR and PEP725. For our analyses, we required at least 8 overlapping years between the climate predictors and each phenological time series. The choice of an  $n \geq 8$  threshold was based on a trade off between a desire for increased statistical robustness (which is improved with a higher  $n$ ) versus increased inclusion of NECTAR sites with many short phenological time series (for example, SEV, KON). To assess the statistical robustness of our results, we repeated our main analyses for select NECTAR sites using stricter thresholds of  $n \geq 15$  and  $n \geq 20$ . In terms of the percentage of significant responders and the mean site level sensitivities, our results were largely insensitive to the minimum  $n$  threshold that we chose (ESM).

For each FFD or FLD series, we conducted separate model selection procedures for the  $Z_t$  predictors using standard least squares linear regression. In each case, a separate regression was conducted between all climate predictors that occurred before the mean occurrence date of the phenological time series. The beta coefficient from the best fit (highest  $R^2$ ) was retained as our metric of sensitivity to climate, defined as the change in the date of a

phenological event (for example, FFD) of a given taxon, at a given site, in response to one standard deviation of the standardized GDD<sub>sum</sub> index (units of days  $Z_t^{-1}$ ). Given this framework, negative beta coefficients ('sensitivities') indicated earlier flowering or leafing with increased warming and positive beta slopes indicated delayed phenology with warming.

Our methodology represented a tradeoff between selecting a single global predictor versus multiple species-specific models. A single global predictor (for example,  $Z_t$  during one static season for all species) would allow for the best comparison of sensitivities across species, but would likely fit poorly in many cases. Alternatively, fitting species-specific models (for example, using model-determined optimal window lengths for the GDD<sub>sum</sub> calculations, instead of stable 3-month windows) might have improved individual model fits, but would have made it difficult to compare sensitivities across regions and species.

## RESULTS

### Sensitivity to Temperature in PEP725 and NECTAR

Significant temperature sensitivities are ubiquitous throughout the PEP725 database, occurring in 83.6% of all time series. The sign of these sensitivities is overwhelmingly negative (97.6% of the significant temperature sensitivities), and the mean sensitivity across all time series is  $-8.19$  days  $Z_t^{-1}$ . Across the NECTAR sites, 67.0% of all species are significant temperature responders with a mean sensitivity of  $-6.08$  days  $Z_t^{-1}$ . Of the significant temperature responders, 95.5% have negative temperature sensitivities. The distribution and magnitude of these significant temperature sensitivities varies widely from site to site, however (Table 2). Cooler, mesic sites with high temperature seasonality have the highest percentages ( $\geq 70\%$ ) of significant temperature responders: CHN, FAR, GTH, HUB, HVD, MHK, UWM, and WAU. The absolute magnitude of the sensitivities themselves is generally greater at these sites, with the coldest site (GTH) having the highest average sensitivity. Significant temperature responders are least prevalent at SEV and LUQ, our two warmest sites.

### Seasonal Sensitivity Patterns

On average, earlier flowering and leafing species in PEP725 have larger absolute magnitude sensitivity

**Table 2.** For the NECTAR Sites: Mean Site Averaged Temperature Sensitivity (days  $Z_t^{-1}$ ) for Significant Responders and the Percent of Significant Responders at Each Site (Relative to All Species at that Site)

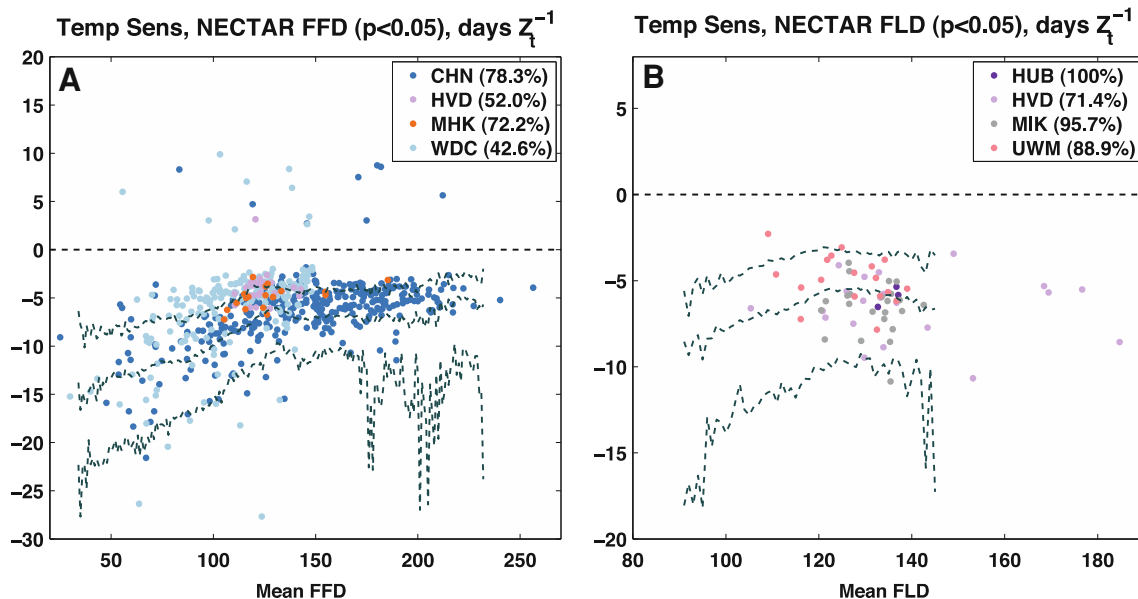
| Site     | Mean sensitivity | % Significant |
|----------|------------------|---------------|
| CHN      | −6.41            | 86.2          |
| FAR      | −5.28            | 79.4          |
| GTH      | −9.9             | 91.1          |
| HUB      | −5.91            | 100           |
| HVD, FFD | −4               | 75.8          |
| HVD, FLD | −6.5             | 87.5          |
| KON      | −5.3             | 58.6          |
| LUQ      | −9.04            | 14.3          |
| MHK      | −4.93            | 100           |
| SEV      | −0.07            | 19.7          |
| UWM      | −4.96            | 72            |
| WDC      | −5.85            | 50.6          |
| WAU      | −6.59            | 95.8          |

Site codes and other details are described in Table 1.

ties than those that flower or produce leaves later in the season (Figure 3). This implies that these earlier taxa will be more responsive to warming, consistent with other analyses (Menzel and Fabian 1999; Menzel 2000). Dashed green lines represent the fifth, fiftieth, and ninety-fifth percentiles of the

negative temperature sensitivities, plotted as a function of mean FFD or FLD (the percentiles provide a tractable way to communicate information from the >90,000 FFD and >20,000 FLD time series in PEP725). Both FFD and FLD in PEP725 have a significant change in the slope that characterizes the relationship between the negative temperature sensitivities and mean FFD and FLD, with a Chow test identifying a significant breakpoint at DOY 119 for FFD and DOY 125 for FLD. Prior to these breakpoints, the regression slopes are strongly positive (FFD  $\beta = +0.088$ ,  $R^2 = 0.231$ ; FLD  $\beta = +0.108$ ,  $R^2 = 0.096$ ), but are near zero after, with  $R^2$  values explaining less than 5% of the underlying variance. These results suggest relatively unchanging sensitivities across the late spring and into the summer, a pattern replicated in other studies (for example, Moeller 2004; Palmer and others 2003; Sherry and others 2007; Sparks and others 2010).

Temperature sensitivities at climatically similar sites in NECTAR follow the same general seasonal pattern as the PEP725 sensitivities (Figure 3A, B: CHN, HVD, MHK, WDC, HUB, WAU, UWM), with a tendency for earlier FFD species to have larger magnitude sensitivities. Except for WDC and HVD (FFD), all of these sites have at least 70% of their species' sensitivities for FFD fall within the PEP725 sensitivities range, defined by the fifth and ninety-



**Figure 3.** Significant temperature sensitivities (days  $Z_t^{-1}$ ) for FFD (left panel) and FLD (right panel) from seven NECTAR sites with similar climate to the PEP725 sites: CHN, HUB, HVD, MHK, UWM, WAU, and WDC. For reference, the 5th, 50th, and 95th percentiles for the negative sensitivities from the PEP725 data are marked by the dashed green lines. Percentages in the panel legends indicate the percentage of significant NECTAR temperature sensitivities for each site that fall within the 5th–95th percentile sensitivity range from PEP725.

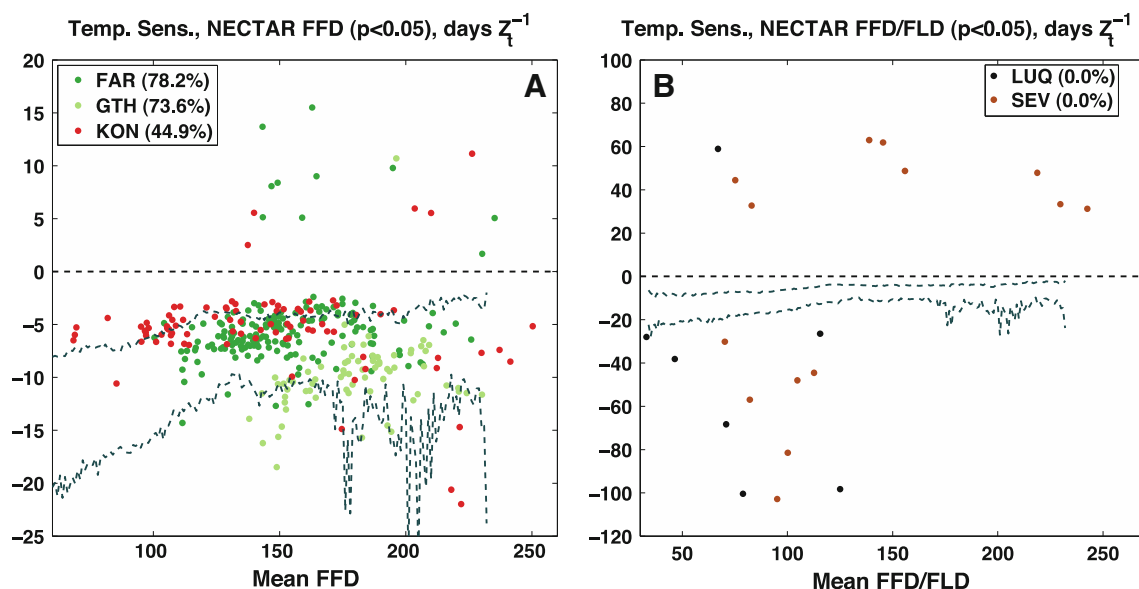


Figure 4. Same as Figure 3, but for FFD and FLD data from NECTAR sites largely outside of the PEP725 climate space.

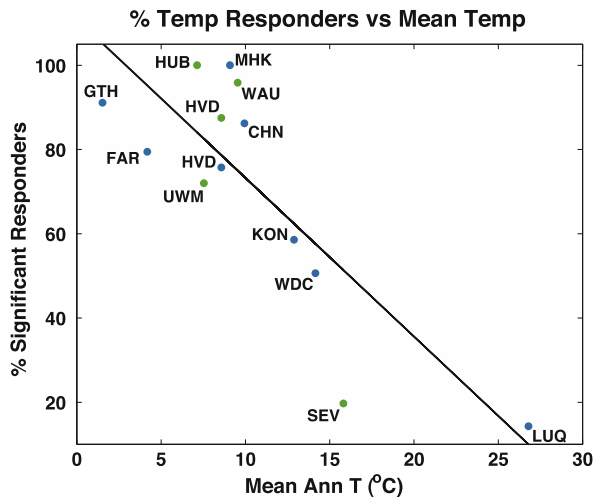
fifth percentiles. The warmest site, WDC, has the highest proportion of sensitivities falling outside the PEP725 range, consistent with other indicators, including a weaker site average sensitivity and lower percentage of responders, that suggest a divergence in climate sensitivity from PEP725. In the case of FFD for species at HVD, the reasons for the divergence from PEP725 are less clear, although it may be a result of the relatively small sample size compared to other sites. Pooled together, the negative temperature sensitivities at these NECTAR sites have a break point at DOY 149, with results similar to the PEP725 analysis before ( $\beta = +0.088$ ,  $R^2 = 0.303$ ) and after ( $\beta = +0.016$ ,  $R^2 = 0.045$ ) the breakpoint. This pattern is driven primarily by WDC and CHN, the sites with the highest total number of species available and the best sampling across the growing season. There is no significant breakpoint in the FLD data, although for this phenophase we are limited by relatively few observations early in the growing season. Sensitivities at other temperature seasonal NECTAR sites that fall outside the range of PEP725 climate (FAR, GTH, KON; Figure 4A) show a high prevalence of negative temperature responders, with a similar tendency towards increased sensitivity in the earlier flowering species, but without significant breakpoints identified above. About half the sensitivities at KON fall outside the PEP725 range; similar to DC, this site has relatively high temperature seasonality but also higher mean temperatures than many of the other NECTAR sites. Our tropical (LUQ) and semi-arid sites (SEV) diverge

dramatically from the PEP725 pattern (Figure 4B); none of the temperature sensitivities for species at these sites fall within the bounds defined by PEP725.

### Sensitivities Across Climate Space

Despite their high temperature seasonality, only about half of the species at KON and WDC are significant temperature responders. Both sites are among the warmest in the NECTAR database, pointing to the possibility of changing climatic constraints across the wide climatic gradient covered by NECTAR. Across the NECTAR sites, the percentage of significant temperature responders (including data from both FFD and FLD) significantly declines with increasing mean annual temperature (Figure 5;  $\beta = -3.8$  percent of species per  $^{\circ}\text{C}$ ,  $R^2 = 0.680$ ,  $P < 0.001$ ). Even if we exclude SEV and LUQ, two sites that may have a disproportionate influence on the regression, our results are similar and borderline significant ( $\beta = -2.5$  percent of species  $^{\circ}\text{C}$ ,  $R^2 = 0.304$ ,  $P < 0.08$ ). By contrast, a similar regression (the percentage of significant responders vs. mean annual temperature) across the thousands of PEP725 sites is unable to detect an effect of mean annual temperature on the percent of responders. For FFD in PEP725, the regression is significant, but the slope is shallow and the low  $R^2$  indicates little explanatory power ( $\beta = -0.8$  percent of species per  $^{\circ}\text{C}$ ,  $R^2 = 0.0036$ ,  $P < 0.001$ ). Results for FLD in PEP725 are similar and non-significant ( $\beta = -0.1$  percent of species





**Figure 5.** For the NECTAR sites, percentage of significant temperature responders as a function of mean annual site temperature. *Blue points* are FFD and *green points* are FLD. The *black line* represents the best fit linear least squares regression.

per °C,  $R^2 \leq 0.001$ ,  $P = 0.7271$ ). Because of the limited climate space covered by PEP725 (Figure 2), results from this database may be less useful for generalizing about species responses across broad climate gradients.

## DISCUSSION

Biases in ecological data, and how those biases skew interpretations of species responses to climate change, have been widely acknowledged but only rarely explicitly tested (Loiselle and others 2008; Trivedi and others 2008). This is especially true for phenological data, where large-scale analyses have focused primarily on observational networks that have relatively low species diversity and large geographic and climatic biases. Although the importance of networks in the field of global change ecology is well established (Menzel and Fabian 1999; Menzel 2000), moving from synthesis to large-scale prediction will require integrative assessments of how well these data represent wild populations and speciose communities across broad climate gradients. We compared the climate sensitivity of two extensive plant phenological databases that differ in important and ecologically relevant ways (species richness, climate space, and so on): the well known PEP725 network, and a new database of wild species observations, NECTAR. Our methodology is designed to facilitate the comparison of species sensitivities across spatial climate gradients and the growing season using seasonal heat sums; our results will thus be

relevant primarily for informing how future climate trends will impact phenology. Ecological responses to climate change, however, will not be solely driven by changes in the mean climate, but also through shifts in climate variability and extreme events (for example, Augspurger 2009; Jackson and others 2009; Walther 2009). The impact of such events (for example, droughts, early frosts, cold snaps, heat waves) on phenology cannot be fully resolved in our analysis, but will likely modulate the overall response of phenology to climate trends into the future.

## Predictability of Temperate Species' Sensitivities from PEP725

Results from the seasonal comparisons between PEP725 and NECTAR suggest that the PEP725 database succeeds quite well in predicting the temperature sensitivity of natural populations and communities at climatically similar sites. This includes matching the seasonal distribution of sensitivities, the prevalence of significant responders, and the magnitudes of the sensitivities themselves (Table 1; Figure 3). This implies that the relatively low seasonality and cool temperatures inherent in PEP725, rather than the poor species diversity, may be the major limitation on the ability to generalize from PEP725 to many other natural systems, including temperature seasonal systems with warmer mean annual temperatures (for example, WDC). It may be that PEP725 accurately captures the prevalence and magnitude of significant temperature sensitivities by virtue of at least some limited exposure to variations in climate across space (that is, the different sites) and time (that is, the growing season). Notably, PEP725 includes many widely distributed species with broad ecological tolerances which may have phenologies that are phenotypically or genotypically plastic. Indeed, single species or even genotypes (Stinchcombe and others 2004) can have dramatically different flowering times and specific climate responses with varied environmental conditions.

## Climate Sensitivities Beyond the Temperate Mid-latitudes

With increasing mean annual temperature—diverging from PEP725 climate space—the percentages of significant temperature responders declines across the NECTAR sites, even at locations expected to have significant seasonal temperature stress, such as KON and WDC. This result is difficult to discern from an analysis across PEP725 sites,

given their limited climatic range, and points directly to (1) the value of integrative analyses across systems (as we conducted with NECTAR), and (2) the need for more data from under-represented systems, including relatively warm, but still temperature seasonal systems such as WDC and KON. Other cold sites in NECTAR show strong temperature responses, but with seasonal patterns distinctly different from PEP725. For example, our cool grassland and alpine meadow sites (FAR, GTH, KON) have strong temperature sensitivities and generally increase in sensitivity earlier in the season, but do not have the pronounced seasonal breakpoint in sensitivity that other sites have.

Sensitivities at LUQ and SEV, sites well beyond the typical range of climates encompassed by PEP725, diverge even more widely from the PEP725 sensitivity patterns. This may result, at least partially, from limitations inherent in our analytical framework, which was designed to generate temperature-based sensitivities for phenological time series in such a way as to maximize comparability across species, sites, and the growing season. In addition, phenology in these systems may reflect different climate or environmental limitations, including moisture stress (SEV) or day length and seasonal insolation (LUQ). Because of the temperature focus of our analyses, and the low representation of data from arid and tropical regions, any general conclusions drawn about these systems should be considered cautiously. Beyond the influence of changing climatic constraints, differences in taxonomic composition within and across sites may also contribute to differences in their mean sensitivities, in the frequency distribution of sensitivities, and in the relationship between temperature and FFD or FLD.

### Ecological Networks: Strengths and Weaknesses

PEP725 exemplifies the great strengths of current phenological monitoring in the temperate mid-latitudes, including high replication of species and phenophases across sites. Our comparison with NECTAR confirms that these databases can provide estimates of plant phenology responses to temperature in wild communities across similar climate gradients, despite low species diversity. This suggests that adequate sampling across the growing season is important for characterizing the community response, and may be able to compensate for low diversity in species sampling. Further, the robustness of the results between PEP725 and NECTAR means we may be able to move beyond

basic statistical diagnosis of trends and towards explicit predictions into the future, at least for these systems.

Our database comparison also highlights the need for increased sampling of understudied systems, where responses can be expected to significantly diverge from models based on databases such as PEP725. Aside from the value of increasing our understanding of poorly sampled systems, observations across broad spatial environmental gradients also allow for unified investigations of how species and ecosystem responses to climate change vary as a function of changing environmental limitations. For example, our analysis of NECTAR shows that the proportion of species within a community that are sensitive to temperature generally declines with increasing mean annual temperature. This is information that could not be gleaned from the limited climate space covered by the PEP725 sites, and shows that there is still significant value in integrating observations of wild populations and communities across broad areas and climates, even without high cross-site species replication.

Understanding and predicting ecological responses to climate change is an active area of research, but one that has proved difficult to bring together into a general theoretical framework. Synthesis efforts that target databases with different limitations and biases can contribute to this effort, helping identify and quantify the strengths and weaknesses inherent in various monitoring and analysis approaches. Identifying and resolving the limitations inherent in the data will be especially important for determining whether increased monitoring and sampling is needed, or if the data is sufficient for forecasting and prediction applications.

### ACKNOWLEDGMENTS

Primary data collections were made possible through the support of many granting agencies; please see ESM for complete information and grant numbers. This work was conducted as a part of the "Forecasting Phenology" Working Group supported by the National Center for Ecological Analysis and Synthesis, a center funded by the National Science Foundation (NSF) (Grant #EF- 0553768), the University of California, Santa Barbara, and the State of California. Special thanks to the many data holders and data managers who assisted us throughout the process including K. Vanderbilt and K. Wetherill (SEV), Chris Nytch and Jess Zimmerman (LUQ), George Aldridge and David Inouye (GTH), John O'Keefe (HVD), and Paul Huth,

Shanan Smiley, and John Thompson from the Mohonk Preserve (MHK). Some data used in this publication were obtained by scientists of the Hubbard Brook Ecosystem Study; this publication has not been reviewed by those scientists. The Hubbard Brook Experimental Forest is operated and maintained by the Northeastern Research Station, U.S. Department of Agriculture, Newtown Square, Pennsylvania. Additional support was also provided by the USA National Phenology Research Coordination Network, supported by NSF grant #IOS-0639794. NJBK was supported by the NSERC CREATE Training Program in Biodiversity Research. PEP725 data were provided by the members of the PEP725 project. Special thanks to E. Koch and W. Lipa for providing the PEP725 data and the accompanying climate data. Support for EMW came from the NSF Postdoctoral Fellow program (Grant #DBI-0905806). Thanks to the editors, two anonymous reviewers, Jonathan Hanes, and David Inouye for providing valuable comments that significantly improved the quality of this manuscript. LDEO contribution number #7580.

## REFERENCES

- Abu-Asab MS, Peterson PM, Shetler SG, Orli SS. 2001. Earlier plant flowering in spring as a response to global warming in the Washington, DC, area. *Biodivers Conserv* 10:597–612.
- Augsburger CK. 2009. Spring 2007 warmth and frost: phenology, damage, and refoliation in a temperate deciduous forest. *Funct Ecol* 23:1031–9.
- Cook BI, Cook ER, Huth PC, Thompson JE, Forster A, Smiley D. 2008. A cross-taxa phenological dataset from Mohonk Lake, NY and its relationship to climate. *Int J Climatol* 28:1369–83.
- Dunnell KL, Travers SE. 2011. Shifts in the flowering phenology of the Northern Great Plains: patterns over 100 years. *Am J Bot* 98:935–45.
- Fitter AH, Fitter RSR. 2002. Rapid changes in flowering time in British plants. *Science* 296:1689.
- Haylock MR, Hofstra N, Tank AMGK, Klok EJ, Jones PD, New M. 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *J Geophys Res* 113:D20119.
- Hofstra N, Haylock M, New M, Jones PD. 2009. Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature. *J Geophys Res* 114:D21101.
- Inouye DW. 2008. Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. *Ecology* 89:353–62.
- Jackson ST, Betancourt JL, Booth RK, Gray ST. 2009. Ecology and the ratchet of events: climate variability, niche dimensions, and species distributions. *Proc Natl Acad Sci* 106:19685–92.
- Klok EJ, Klein Tank AMG. 2009. Updated and extended European dataset of daily climate observations. *Int J Climatol* 29:1182–91.
- Koch E, Adler S, Lipa M, Ungersbock M, Zach-Hermann S (2010a) The Pan European phenological database PEP725. *Berichte des Meteorologischen Institutes der Albrecht-Ludwigs-Universität Freiburg*. In Proceedings of the 7th conference on Biometeorology.
- Koch E, Adler S, Ungersbock M, Zach-Hermann S (2010b) PEP725 Pan European Phenological Database. *Geophys Res Abst* 12.
- Leon AJ, Lee M, Andrade FH. 2001. Quantitative trait loci for growing degree days to flowering and photoperiod response in sunflower (*Helianthus annuus* L.). *Theor Appl Genet* 102:497–503.
- Loiselle BA, Jørgensen PM, Consiglio T, Jiménez I, Blake JG, Lohmann LG, Montiel OM. 2008. Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? *J Biogeogr* 35:105–16.
- McMaster GS, Wilhelm WW. 1997. Growing degree-days: one equation, two interpretations. *Agric For Meteorol* 87:291–300.
- Menzel A. 2000. Trends in phenological phases in Europe between 1951 and 1996. *Int J Biometeorol* 44:76–81.
- Menzel A, Fabian P. 1999. Growing season extended in Europe. *Nature* 397:659.
- Menzel A, Sparks TH, Estrella N et al. 2006. European phenological response to climate change matches the warming pattern. *Glob Change Biol* 12:1969–76.
- Moeller DA. 2004. Facilitative interactions among plants via shared pollinators. *Ecology* 85:3289–301.
- Olsson K, Ågren J. 2002. Latitudinal population differentiation in phenology, life history and flower morphology in the perennial herb *Lythrum salicaria*. *J Evol Biol* 15:983–96.
- Palmer TM, Stanton ML, Young TP. 2003. Competition and coexistence: exploring mechanisms that restrict and maintain diversity within mutualist guilds. *Am Nat* 162:S63–79.
- Parmesan C, Yohe G. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421:37–42.
- Parry ML. 2007. Climate Change 2007: Impacts, adaptation and vulnerability: Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press.
- PEP725 (2010) PEP725: Pan European Phenology Data. <http://www.zamg.ac.at/pep725/>.
- Peterson TC, Vose RS. 1997. An overview of the Global Historical Climatology Network temperature database. *Bull Am Meteorol Soc* 78:2837–49.
- Richardson AD, Bailey AS, Denny EG, Martin CW, O'Keefe J. 2006. Phenology of a northern hard-wood forest canopy. *Glob Change Biol* 12:1174–88.
- Root TL, Price JT, Hall KR, Schneider SH, Rosenzweig C, Pounds JA. 2003. Fingerprints of global warming on wild animals and plants. *Nature* 421:57–60.
- Rosenzweig C, Karoly D, Vicarelli M et al. 2008. Attributing physical and biological impacts to anthropogenic climate change. *Nature* 453:353–7.
- Rötzer T, Chmielewski FM. 2001. Phenological maps of Europe. *Clim Res* 18:249–57.
- Scheffinger H, Menzel A, Koch E, Peter C, Ahas R. 2002. Atmospheric mechanisms governing the spatial and temporal variability of phenological phases in central Europe. *Int J Climatol* 22:1739–55.
- Scheffinger H, Menzel A, Koch E, Peter C. 2003. Trends of spring time frost events and phenological dates in Central Europe. *Theor Appl Climatol* 74:41–51.

- Schwartz MD, Hanes JM. 2010. Intercomparing multiple measures of the onset of spring in eastern North America. *Int J Climatol* 30:1614–26.
- Schwartz MD, Reiter BE. 2000. Changes in North American spring. *Int J Climatol* 20:929–32.
- Schwartz MD, Ahas R, Aasa A. 2006. Onset of spring starting earlier across the Northern Hemisphere. *Glob Change Biol* 12:343–51.
- Sherry RA, Zhou X, Gu S et al. 2007. Divergence of reproductive phenology under climate warming. *Proc Natl Acad Sci* 104:198.
- Smith JW. 1915. Phenological dates and meteorological data recorded by Thomas Mikesell at Wauseon, Fulton County, Ohio. *Mon Weather Rev Suppl* 2:21–93.
- Sparks TH, Górski-Zajackowska M, Wojtowicz W, Tryjanowski P. (2010) Phenological changes and reduced seasonal synchrony in western Poland. *Int J Biometeorol* 55:1–7.
- Stinchcombe JR, Weinig C, Ungerer M et al. 2004. A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene FRIGIDA. *Proc Natl Acad Sci* 101:4712.
- Travers SE, Dunnell K. 2009. First-flowering dates of plants in the Northern Great Plains. *Ecology* 90:2332.
- Trivedi MR, Berry PM, Morecroft MD, Dawson TP. 2008. Spatial scale affects bioclimate model projections of climate change impacts on mountain plants. *Glob Change Biol* 14:1089–103.
- Walther GR. 2009. Two steps forward, one step back. *Funct Ecol* 23:1029–30.
- Walther GR, Post E, Convey P et al. 2002. Ecological responses to recent climate change. *Nature* 416:389–95.
- Wetherill K (2000) Plant phenology transects. Tech. rep., Sevilleta Long Term Ecological Research Site.
- Willis CG, Ruhfel B, Primack RB, Miller-Rushing AJ, Davis CC. 2008. Phylogenetic patterns of species loss in Thoreaus woods are driven by climate change. *Proc Natl Acad Sci* 105:17029.
- Wright V (2001) Prairies phenology. Tech. rep., Konza Environmental Education Program.
- Zimmerman JK, Wright SJ, Calderon O, Pagan MA, Paton S. 2007. Flowering and fruiting phenologies of seasonal and aseasonal neotropical forests: the role of annual changes in irradiance. *J Trop Ecol* 23:231–51.