Probing the past 30 year phenology trend of US deciduous forests

X. Yue¹, N. Unger¹, T. F. Keenan², X. Zhang³, and C. S. Vogel⁴

¹School of Forestry and Environmental Studies, Yale University, 195 Prospect Street, New Haven, CT 06511, USA
²Department of Biological Sciences, Macquarie University, Sydney, New South Wales 2109, Australia
³Geospatial Sciences Center of Excellence, South Dakota State University, Brookings, SD 57007, USA
⁴The University of Michigan Biological Station, Pellston, MI 49769, USA

Received: 20 March 2015 – Accepted: 6 April 2015 – Published: 24 April 2015

Correspondence to: X. Yue (xuyueseas@gmail.com)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Phenology is experiencing dramatic changes over deciduous forests in the US. Estimates of trends in phenology on the continental scale are uncertain, however, with studies failing to agree on both the magnitude and spatial distribution of trends in spring and autumn. This is due to the sparsity of in situ records, uncertainties associated with remote sensing data, and the regional focus of many studies. It has been suggested that reported trends are a result of recent temperature changes, though multiple processes are thought to be involved and the nature of the temperature forcing remains unknown. To date, no study has directly attributed long-term phenological trends to individual forcings across the US through integrating observations with models. Here, we construct an extensive database of ground measurements of phenological events across the US, and use it to calibrate and evaluate a suite of phenology models. The models use variations of the accumulative temperature summation, with additional chilling requirements for spring phenology and photoperiod limitation for autumn. Including a chilling requirement or photoperiod limitation does not improve model performance, suggesting that temperature change, especially in spring and autumn, is the dominant driver of the observed trend during the past 3 decades. Our results show that phenological trends are not uniform over the contiguous US, with a significant advance of 0.34 day yr\(^{-1}\) for the spring budburst in the East, a delay of 0.15 day yr\(^{-1}\) for the autumn dormancy onset in the Northeast and West, but no evidence of change elsewhere. Relative to the 1980s, the growing season in the 2000s is extended by about 1 week (3–4%) in the East, New England, and the upper Rocky Mountains forests. These results help reconcile conflicting reports of phenological trends in the literature, and directly attribute observed trends to long-term changes in temperature.
1 Introduction

Plant phenology, such as the timing of spring budburst and autumn leaf fall, is sensitive to climatic conditions (Korner and Basler, 2010; Richardson et al., 2013) and is thus exhibiting a long-term trend with the changing climate (Badeck et al., 2004; Gordo and Sanz, 2009; Jeong et al., 2011). Long-term changes in phenology may be affecting ecosystem carbon assimilation (Keenan et al., 2014), surface water and energy balance (Schwartz and Crawford, 2001), and forest composition and evolution (Forrest and Miller-Rushing, 2010). Emerging observations have shown advanced spring and delayed autumn over the Northern Hemisphere, especially in Europe, during the past several decades (Menzel and Fabian, 1999; Fitter and Fitter, 2002; Gordo and Sanz, 2009). However, the extent of regional phenological trends in US remains uncertain as different studies present inconsistent and even opposite results (Table 1).

The uncertainty of the phenological changes in US forests could be attributed to genetic, geographic, and temporal factors. First, experiments have suggested that different species may have different phenological sensitivity to temperature (Vitasse et al., 2009). Some species may also require cold winter temperatures before budburst, leading to divergent responses of US plants to spring and winter warming at the community level (Cook et al., 2012) and the continental scale (Zhang et al., 2007). In addition, it is not clear whether other biotic and/or abiotic factors (e.g. humidity, photoperiod, tree age, and tree species) may play a role (Morin et al., 2009; Basler and Korner, 2012; Caldararu et al., 2014; Laube et al., 2014). Second, most deciduous forests in the US are found at mid-latitudes, where temperature increases have not been uniform, and are not as strong as those at high latitudes (Hartmann et al., 2013). The high elevation in the western US also induces additional complexity for the phenological responses, as temperature sensitivity to altitudinal trends also differs among species (Vitasse et al., 2009). Third, differences in the time frames used in different studies may lead to apparently inconsistent trends (Badeck et al., 2004).
There are generally three approaches for estimating phenology at regional and continental scales: ground networks, remote sensing, and numerical modeling. Ground-based measurements can provide the most accurate phenological dates, such as budburst, flowering, and leaf fall. Some records last for decades and even centuries (Sparks and Menzel, 2002), making it possible to study long-term phenological change. However, such measurements usually have very limited spatial coverage. Ground-based networks, such as North American Lilac Network (Schwartz and Reiter, 2000), improve the spatial coverage but focuses only on 1–2 species, which may not represent the average phenological status of local plants. More extensive networks, such as the North American Phenology Network (www.usanpn.org) or the European Phenology Network (www.pep725.eu), contain many more species but typically do not have long data records (with exceptions of course). Remote sensing provides a way to examine phenological changes over large scales but is inherently limited by short time scales or infrequent retrieval times and must be validated using ground measurements. Most of the recent estimates of phenological changes on the continental scale are performed using satellite retrievals (e.g., Zhang et al., 2007; Jeong et al., 2011; Piao et al., 2011). The lack of a universally accepted definition of phenological status for this method may lead to discrepancies up to 60 days for the timing of events among different algorithms and products (White et al., 2009). Moreover, date retrieval is often hampered, e.g., by cloud cover, which can lead to poor correlations with ground observations (Badeck et al., 2004; Schwartz and Hanes, 2010).

Phenological models are useful tools for diagnosing causes of phenological changes and also for understanding the feedback of those changes to the Earth system (Richardson et al., 2013; Zhao et al., 2013). Evaluations of well-calibrated phenological models have shown high correlations between predictions and observations (e.g., White et al., 1997; Richardson et al., 2006; Delpierre et al., 2009; Vitasse et al., 2011). However, most of these state-of-art schemes are not evaluated at continental or even larger scales, thus limiting their applicability in dynamic vegetation models and climate models. Recent model-data comparisons have shown that the bias in the prediction of...
vegetation phenology is a large source of uncertainty in models of ecosystem carbon uptake (Richardson et al., 2012). This necessitates the development and evaluation of continental scale phenology models with continental scale observations.

In this study, we use an extensive dataset of phenological observations to calibrate and evaluate 13 models (9 for spring and 4 for autumn) of deciduous tree phenology across the US. We first calibrate each model using long-term ground observations of phenology at four deciduous forests. We then examine modeled interannual variability and trends, along with regional phenological differences, using an extensive network of phenological observations. The phenology model best supported by the observations is then applied to: (1) estimate the trend of both spring and autumn phenology of US deciduous forests over the last three decades; (2) compare our results with other approaches (ground network, remote sensing, and model based) to identify robust changes and assess discrepancies; and (3) examine the underlying drivers of both the observed trends and interannual variability.

2 Materials and methods

We assembled and compared a suite of published models of spring and autumn phenology. Most of these models are built using cumulative thermal summations with constraining processes, such as chilling requirements and photoperiod limits. Model parameters were calibrated using long-term observations at four deciduous forest sites, with some model constants estimated based on literature values. An independent dataset of ground measurements was compiled and used to validate the performance of these models. In total, phenological observations from ~1000 sites were used. In this section we first present the observations used for calibration and validation, followed by a description of the various model formulations tested and simulations performed.
2.1 Ground measurements for calibration

Long-term measurements of leaf area index (LAI) from four US deciduous broadleaf sites are collected from the Ameriflux network (http://ameriflux.ornl.gov/) to calibrate parameters of the phenology model (Table 2). We derive annual cycles of phenology by normalizing individual LAI values to the maximum and minimum LAI in each year for each site (Fig. S1 in the Supplement, top panel). Since the measurements are discrete, we estimate the long-term average budburst dates (D1), growing length (L1), offset start dates (D2), and falling length (L2) based on segmented regressions, which yield the minimum root mean square error (RMSE) against observations (Fig. S1 middle and bottom panel):

\[ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (O_i - P_i)^2} \]  

where \( O_i \) is the observation, \( P_i \) is the regression or prediction, and \( n \) is the number of samples. The derived phenological dates are presented in Table 3. The average budburst date at US-MMS and US-MOz is earlier by three weeks than that at US-Ha1 and US-UMB, probably because the former sites are \( \sim 5 \) °C warmer than the latter. However, the start of leaf senescence is similar at all four sites, suggesting that photoperiod may also play an important role in regulating the autumn phenology, especially at the two warmer sites.

2.2 Ground measurements for validation

We use > 75,000 records for deciduous trees to evaluate the temporal variation and spatial distribution of simulated phenology (Tables 4, S1 and S2 in the Supplement). Although data at some sites, such as US-Ha1, US-UMB, and US-MMS (Table 2), are also used for calibration, we use them in different ways. For calibration, we use the long-term average phenology derived from the multiple-year LAI measurements, so...
that every calibrated model can capture the spatial pattern of phenology events on the continental scale. For validation, we use year-to-year phenological dates estimated from date records, photos, and LAI at each year, so as to identify the model that best captures the temporal variations.

The two New England sites, Harvard Forest (http://harvardforest.fas.harvard.edu/) and Hubbard Brook (US-HB1) Forest (http://www.hubbardbrook.org/), have long-term measurements back to 1990. The full records at Harvard Forest include 34 species, 16 of which are deciduous trees. The forest within the tower footprint is dominated by red oak (Quercus rubra, 60% basal area), red maple (Acer rubrum, 23% basal area), and secondary deciduous species. Hubbard Brook has three species, namely sugar maple (Acer saccharum), American beech (Fagus grandifolia), and yellow birch (Betula alleghaniensis). We average over all trees and species at each site to generate average phenological dates for each year. Phenological observations are incomplete at two of the Ameriflux sites, US-UMB and US-MMS. We derive the missing phenological dates based on LAI data from Ameriflux and images from the PhenoCam project (http://phenocam.sr.unh.edu/webcam/). If the year-round LAI data are available at one site, we estimate budburst and dormancy start dates as the days when the interpolated or extrapolated LAI is equal to a selected threshold (see descriptions in the Supplement). Otherwise, we qualitatively estimate phenological dates based on photos from PhenoCam, which is a near-surface remote sensing network that observes phenology changes with high-resolution digital cameras (Sonnentag et al., 2012). We define the budburst date as the middle of the few days when tree colors change rapidly from gray to light green. On the contrary, a dormancy start date is defined as the middle of days with rapid color changes from brown to gray. An example of autumn dormancy at US-UMB is shown in Fig. S3.

Data from ground networks was used to evaluate the model performance on the continental scale. The USA National Phenology Network (USA-NPN) is a nationwide project collecting standardized ground phenology observations by researchers, students, and volunteers. The network has limited records before 2009 but is significantly
enriched thereafter. We select observations during 2011–2012 for 52 deciduous tree species that are most common in the US (Table S1). The derived phenological dates for individual trees are averaged if they are observed at the same location (or site). We also used observations from The North American Lilac Network (NALN), which provides records of the first leaf and first bloom dates of two lilac species, Common Lilac (*Syringa vulgaris*) and Red Rothomagensis lilac (*Syringa chinensis*), for the period of 1956–2003 (Schwartz and Reiter, 2000). As we shown in Sect. 3.4.3, the phenology of individual species may vary by up to 3 weeks, however, the responses of phenology to temperature changes are relatively similar across species. We calculate correlations of budburst dates between observations and simulations at the available sites of NALN to validate the simulated temporal variations of phenology. We also adopt the limited long-term records from USA-NPN (Table S2) to evaluate the model over regions not covered by NALN.

### 2.3 Spring phenology models

Dozens of spring phenology models have been evaluated and inter-compared in the past two decades (Chuine et al., 1999; Linkosalo et al., 2008; Vitasse et al., 2011; Fu et al., 2012a, b; Migliavacca et al., 2012; Melaas et al., 2013). These models may have different formats and parameters, but are generally dependent on temperature and photoperiod and could be divided into two categories, spring warming (or 1-phase) and chilling (or 2-phase), based on their assumptions of how warm and cold temperatures control the phenology development (Migliavacca et al., 2012). Although regional studies have demonstrated that the 1-phase models are as efficient as 2-phase models for most species (e.g., Vitasse et al., 2011; Fu et al., 2012a; Migliavacca et al., 2012), we consider that chilling requirement may be necessary for the phenology at the continental and global scales where divergent phenological responses are observed (Zhang et al., 2007; Cook et al., 2012).

The chilling models have different formulations based on the sequences (sequential, parallel, or alternating) and forms (thermal summation or the Sarvas function) of
chilling and forcing (Chuine et al., 1999). According to these differences, Migliavacca et al. (2012) summarized and compared eight models, listed as S1–S8 in Table 5, to fit phenology data at Harvard Forest. The sequential models require that a chilling threshold ($C^*$) much be achieved before the forcing ($S_f$) is effective. The parallel and alternating models calculate chilling units ($S_c$) and $S_f$ at the same time, however, the increases in $S_c$ can reduce the budburst threshold ($F^*$) for $S_f$ following an exponential relationship $F^* = a \exp(b \times S_c)$. The functions of $S_c$ and $S_f$ are calculated as the cumulative thermal unit as follows:

$$S_c(t) = \sum_{t_1}^{t} R_c(x_t)$$  \hspace{1cm} (2)

$$S_f(t) = \sum_{t_2}^{t} R_f(x_t)$$  \hspace{1cm} (3)

where $x_t$ is the daily temperature. The thermal unit may have two different formats. In the thermal summation approach (CF1, Eqs. 4 and 5), $S_c$ is the number of chilling days (<$T_c$) from a starting day $t_1$ and $S_f$ is the cumulative temperature higher than $T_f$ (commonly named growing degree day, GDD) from day $t_2$. In the other approach (CF2, Eqs. 6 and 7), both $R_c$ and $R_f$ are functions of daily temperature (Chuine et al., 1999).

**CF1:** $R_c(x_t) = \begin{cases} 0, & x_t \geq T_c \\ 1, & x_t < T_c \end{cases}$  \hspace{1cm} (4)

**CF1:** $R_f(x_t) = \begin{cases} x_t - T_f, & x_t \geq T_f \\ 0, & x_t < T_f \end{cases}$  \hspace{1cm} (5)

**CF2:** $R_c(x_t) = \begin{cases} 0, & x_t \leq -3.4 \text{ or } x_t \geq 10.4 \\ \frac{x_t + 3.4}{T_c + 3.4}, & -3.4 < x_t < T_c \\ \frac{x_t - 10.4}{T_c - 10.4}, & T_c < x_t < 10.4 \end{cases}$  \hspace{1cm} (6)
CF2: \[ R_f(x_t) = \begin{cases} 0, & x_t \leq 0 \\ \frac{28.4}{1 + e^{-0.185(x_t-18.4)}}, & x_t > 0 \end{cases} \] (7)

For both parallel and alternating models, \( t_1 \) is equal to \( t_2 \), and for the latter, \( T_c \) is equal to \( T_f \). For sequential models, \( t_2 \) is the first day when \( S_c > C^* \). In a modified alternating scheme (S9), we decrease model complexity by fixing some parameters based on literature values. First, we fix \( t_1 \) as the winter solstice (22 December in Northern Hemisphere, NH), after which photoperiod increases gradually. Second, we set \( T_c \) to 5°C, a value widely used for woody species (Murray et al., 1989; Kaduk and Heimann, 1996; Sitch et al., 2003). Third, we redefine the format of forcing threshold as \( F^* = a + b \exp(r \times S_c) \) following Murray et al. (1989) and set \( r = -0.01 \), a value used for temperate trees (e.g. beech and black locust). For each model in Table 5, we adjust all unfixed parameters step by step and select the optimized parameters that jointly predict the lowest RMSE for the long-term budburst dates at the four calibration sites.

We assume the green up process is linearly dependent on forcing \( S_f \) as follows,

\[
f_T = \begin{cases} 0, & S_f < F^* \\ \frac{S_f - F^*}{L_g}, & F^* \leq S_f \leq F^* + L_g \\ 1, & S_f > F^* + L_g \end{cases}
\] (8)

where \( f_T \) is a temperature-dependent phenology ranging from 0 to 1. The parameter \( L_g \) is a growing length constraint calibrated based on the cycle of forest phenology (Fig. S1).

### 2.4 Autumn phenology models

Autumn phenology is more uncertain than budburst because it is affected by both temperature and photoperiod. Three models have been developed to predict leaf fall with constraint from temperature and photoperiod, namely the continental phenology model...
by White et al. (1997), the growing season index (GSI) by Jolly et al. (2005) and the cold-degree day photoperiod-dependent model by Delpierre et al. (2009). The White et al. (1997) scheme is not compared in this study as it depends on soil temperature, which is not available at some sites. Jolly et al. (2005) calculated global phenology as the product of three segmented functions, which depend on the upper and lower limits in temperature ($T_x$ and $T_i$), vapor pressure deficit (VPD) ($V_x$ and $V_i$), and photoperiod ($P_x$ and $P_i$), respectively. The value of VPD function is set to constant of 1 for temperate forests with no water stress. Delpierre et al. (2009) calculated the cumulative products of the functions of temperature and photoperiod. Those functions may have power indexes ranging from 0 to 2, suggesting that autumn phenology could be unrelated, linearly related, or exponentially related with the constraints from temperature and photoperiod. We calibrate all model parameters based on the observations at US deciduous forests (A2 and A3 in Table 5). We also use the original parameters from Jolly et al. (2005), which have been validated based on remote sensing data on the global scale (A1 in Table 5).

We also construct a simple scheme based on cumulative cold degree-days. The scheme, named “CDD-photoperiod” (A4 in Table 5), calculates cold degree days (CDD) $C_a$ following Richardson et al. (2006):

$$C_a(t) = \sum_{t_3}^{t} R_a(x_t)$$

$$R_a(x_t) = \begin{cases} T_b - x_t, & x_t < T_b \\ 0, & x_t \geq T_b \end{cases}$$

where $t_3$ is the starting day set to summer solstice (22 June in NH), and $T_b$ is a base temperature of 20°C as that in Dufrene et al. (2005) and Richardson et al. (2006). The leaf fall is triggered if $C_a$ is higher than a threshold $F_s$ and the length of falling period is
determined by $L_f$ as follows,

$$f_T = \begin{cases} 
0, & C_a \leq F_s \\
\frac{C_a - F_s}{L_f}, & F_s < C_a < F_s + L_f \\
1, & C_a \geq F_s + L_f 
\end{cases} \tag{11}$$

here $f_T$ is the temperature-dependent phenology ranging from 0 to 1. We also define a photoperiod-limited phenology following Jolly et al. (2005),

$$f_P = \begin{cases} 
0, & P \leq P_i \\
\frac{P - P_i}{P_x - P_i}, & P_i < P < P_x \\
1, & P \geq P_x 
\end{cases} \tag{12}$$

where $P$ is the daylength in minutes. $P_i$ and $P_x$ are the lower and upper limits of daylength during the period of leaf fall. Following Jolly et al. (2005), the final autumn phenology of deciduous forest is determined as the product of $f_T$ (Eq. 11) and $f_P$ (Eq. 12).

### 2.5 Simulations

We perform both site-level and continental-scale simulations. For standalone simulations (simulation 1), phenology models are driven with daily surface air temperature sampled at each site (http://ameriflux.ornl.gov/). We gap-filled in situ temperature with daily reanalysis data from the Modern-Era Retrospective Analysis for Research and Applications (MERRA, Reichle et al., 2011; Rienecker et al., 2011), which is interpolated to each site based on the site location. The time span of each site-level simulation varies depending on the availability of the phenology observations. We perform a model inter-comparison to determine which model is most supported by observations. The statistical metrics we used for evaluations include correlations, RMSE (Eq. 1)
and the Akaike Information Criterion (AIC), a measure of the trade-off between model predictability and model complexity (Akaike, 1973; Burnham and Anderson, 2002),

\[
\text{AIC} = n \cdot \log \sigma^2 + 2p + \frac{2p(p+1)}{n-p-1}
\]  

(13)

where \( n \) is the number of samples, \( p \) is the number of fit parameters for the model, and \( \sigma^2 \) is the square of RMSE between prediction and observations. A good prediction usually has high correlation coefficients but low RMSE and AIC values with observations.

For the regional simulation (simulation 2), we utilize daily surface air temperature from MERRA to drive the selected model on a resolution of 1° by latitude and 1.33° by longitude for 1982–2012. The uncertainty of predicted phenology is very sensitive to that of drivers (Migliavacca et al., 2012), as a result, we compare the MERRA forcing with ground observations from the United States Historical Climatology Network (USHCN, Easterling et al., 1996), which provides a high quality data set of daily and monthly temperature from 1218 observing stations across the contiguous United States. We analyze the phenological trend for different time periods so as to understand how the selected time frame and interannual variability may influence our conclusions.

We also perform a sensitivity analysis (simulation 3) to evaluate the uncertainty due to phenological schemes. In this run, we do not include chilling constraint for the spring phenology by using a fixed forcing threshold \( F^* \). Meanwhile, we lift the photoperiod cap for autumn phenology by setting \( f_p = 1 \). We consider a change, trend, or correlation is significant if \( p < 0.05 \), unless otherwise stated.
3 Results

3.1 Model evaluation

3.1.1 Site-level evaluation

The five sites we select to calibrate and evaluate models are all located at Eastern US, where >90% deciduous forests are located (Fig. 1). The site-level evaluations for 9 spring models and 4 autumn models are shown in Fig. 2. For the spring phenology, the alternating approach (S7–S9) has higher correlations and lower RMSE compared to parallel models (S3–S6). The sequential approach with thermal summation (S1) shows the largest correlations and lowest biases. However, it requires fitting 5 parameters, decreasing its AIC value relative to the alternating models. The three alternating models have comparable correlations and RMSE. However, the modified alternating model (S9) has the lowest AIC, suggesting that fixing some parameters based on literature does not weaken the performance but can reduce model complexity. For the autumn phenology, no models predict correlations higher than 0.5, indicating that missing mechanisms, especially some physiological processes (e.g. synthesis, viscosity, and diffusion), may be required to improve the current model structures (Richardson et al., 2006; Linkosalo et al., 2008). The “CDD-photoperiod” scheme (A4) has comparable performance with that from Delpierre et al. (2009) (A3) based on correlation and RMSE, and has lower AIC than the latter due to the lower number of fit parameters (Table 5). As a result of the site-level evaluations, we select the spring model S9 and autumn model A4 (parameters listed in Table S3) as the state-of-art schemes for the regional simulations.

Site-level simulations with models S9 and A4 capture both the interannual variations and temporal trends of phenology at the validation sites (Fig. 1). Sites US-Ha1 and US-HB1 provide >20 years of phenology records. The observation-simulation correlations for budburst dates are 0.7–0.8 at these sites. Model performance is poor for autumn phenology, with correlation coefficients between 0.2–0.4. Both observed and predicted
budburst dates at US-Ha1 show significant advances of $\sim 0.5 \text{ day yr}^{-1}$ during 1992–2012. However, at US-HB1, the observed trend of $-0.3 \text{ day yr}^{-1}$ is not significant due to large interannual variations. In contrast, the dormancy start dates remains almost constant at US-Ha1, similar as that reported by Lee et al. (2003), but exhibits a significant delay of $\sim 0.5 \text{ day yr}^{-1}$ at US-HB1 in the past two decades, as reported by Keenan et al. (2014).

Sites US-UMB and US-MMS have relatively short observations for 1999–2012. Missing in situ forcing values limit the model's spring phenology performance compared to that using MERRA reanalysis. With MERRA forcing, the model shows high correlations ($\sim 0.8$) and low biases (2–4 days) in the prediction of budburst dates. The simulated autumn phenology again has lower correlations with observations at these sites. The predicted dormancy start dates at US-UMB match the observed interannual variation before 2010 but fail to capture the perturbations thereafter. The prediction at US-MMS shows similar year-to-year variations as observations but with smaller magnitude. The spring budburst dates show moderate changes at US-UMB but a significant advance at US-MMS in the past decade. For the autumn phenology, both observations and simulations show insignificant changes.

### 3.1.2 Continental-scale evaluation

Phenology has a distinctive spatial distribution over US deciduous forests (Fig. 3). Budburst occurs relatively later in the west of 105° W but earlier in the low latitudes of the East (Fig. 3a). The area-weighted (based on cover fraction of deciduous forest) budburst date for the western US is 4 May or 124 day of the year (DOY), with higher values of $> 140 \text{ DOY}$ over the ridge of Rocky Mountains. In contrast, the mean budburst date is 15 April (105 DOY) for the east of 105° W, with even earlier dates of $< 100 \text{ DOY}$ at south of 40° N. At higher latitudes, such as the forests over New England and Great Lakes, spring usually begins after 125 DOY due to the colder spring temperatures. The simulated spatial pattern is consistent with phenology records from the USA-NPN network,
with a correlation coefficient of 0.77 over 46 sites (Fig. 3b). The distribution of autumn phenology shows almost opposite pattern as that of spring phenology (Fig. 3c). At high latitudes and/or altitudes, autumn phenology is sensitive to cold temperatures and as a result exhibits an early dormancy onset. The validation against observations from 23 USA-NPN sites yields a significant correlation coefficient of 0.80 for the simulated autumn phenology (Fig. 3d). The S9 model also reproduces year-to-year changes in spring phenology. In 2011, the area-weighted budburst date is 117 DOY (Fig. S12), which is advanced by 13 days in 2012 (Fig. S13). Such change follows the continental warming of spring (March–May) temperature by ~ 3 °C in the latter year (not shown).

We further evaluate the simulated year-to-year budburst dates with available long-term records from NALN and USA-NPN network (Fig. 4). The model-observation correlations are significantly positive (p < 0.2) for most sites, suggesting that the predicted interannual variation and long-term trend of spring phenology are reasonable on the continental scale. However, no long-term records are available to evaluate the temporal variation of simulated autumn phenology on the continental scale.

### 3.2 Phenological change in US deciduous forests

Driven with the MERRA forcing, the model simulates a significant advance of spring budburst dates in central eastern US during 1982–2012 (Fig. 5a). The largest advance of 0.42 day yr\(^{-1}\) is predicted in the states of Illinois and Indiana. For eastern states covered with > 50% deciduous forests, such as Pennsylvania, West Virginia, and Virginia, the budburst date is advanced by 0.34 day yr\(^{-1}\). However, for deciduous forests in the western, northern, northeastern, and southeastern US, the changes are either small or insignificant. Two New England sites, Harvard Forest and Hubbard Brook, are located within the same region but have different trends of spring phenology (Fig. 5a), consistent with site-level evaluations for 1992–2012 (Fig. 1). On the other hand, the dormancy start date is delayed by 0.20 day yr\(^{-1}\) in the northern (Minnesota), 0.14 day yr\(^{-1}\) in the northeastern, and 0.16 day yr\(^{-1}\) in the western forests (Fig. 5b). However, the autumn
Phenology in central and southern US does not show significant changes, consistent with site-level evaluations at US-UMB and US-MMS (Fig. 1).

The spatial pattern of the trend in forest phenology follows spatial patterns of temperature changes in the past 3 decades (Fig. 6). Both the reanalysis data and ground records show a significant spring warming of $0.75 \, ^\circ \text{C} \text{decade}^{-1}$ over central and eastern US while insignificant changes in the other portion of deciduous forest (Fig. 6c and d). Meanwhile, the warmer winter may delay the spring budburst by reducing chilling days, especially for forest in the northern US (Fig. 6a and b). On the other hand, autumn warming in the northern, northeastern, and western forests (Fig. 6e and f) results in delayed dormancy dates in those regions (Fig. 5b). However, autumn phenology in central, eastern, and southern forests shows no significant change, due to either moderate changes in temperature (Fig. 6e and f) or regulation through photoperiod. Based on the synchronous phenological responses to temperature changes, we estimate long-term temperature sensitivities of $-3.3 \, \text{days} \, ^\circ \text{C}^{-1}$ for spring budburst date and $2.2 \, \text{days} \, ^\circ \text{C}^{-1}$ for dormancy start date over US deciduous forests. These values are close to the estimates of $-2.8 \pm 0.3 \, \text{days} \, ^\circ \text{C}^{-1}$ (spring) and $1.8 \pm 0.8 \, \text{days} \, ^\circ \text{C}^{-1}$ (autumn) based on observations from five US deciduous sites (Keenan et al., 2014).

Advanced spring and delayed autumn together increased the length of the growing season across the US (Fig. 7). Relative to the 1980s, the growing season in the 2000s extends by 5.5 days (3.0 %) in the eastern states with dense forest coverage (fraction > 50 %). The model predicts larger extension of 6.4 days (3.9 %) in New England, 7.0 days (3.6 %) in states Illinois and Indiana, and 6.0 days (4.3 %) in the upper Rocky Mountains forests (Fig. 7). This magnitude is comparable to the trend of 2.1–4.2 days per decade in Eurasian and North American temperate forest estimated by other studies (Menzel et al., 2008; Jeong et al., 2011).

### 3.3 Comparison with results from remote sensing

Most of up-to-date estimates of the changes in US forest phenology are performed with remote sensing data. We compare our results to recent reports from the literature,
selecting all studies that exam phenological trends across the US for at least 20 years (Table 1). All selected studies use the Normalized Difference Vegetation Index (NDVI) from satellite data, however, they report different and even opposite trends. Such discrepancies may be attributed to the differences in the definitions of phenological dates (White et al., 2009) or the statistical algorithms in the extraction of the dates (Keenan et al., 2014). Here, we summarize their results on Fig. 8 so as to conclude the most robust changes for US forest phenology in the past 2–3 decades.

For spring phenology, three out of six studies predict advanced budburst or greenup dates in the east, while four predict delayed dates in the north (Fig. 8a). There are no evident phenological changes in the west, northeast, and southeast. Our results show similar changes in spring phenology as the ensemble of the remote sensing studies, except that we predict smaller delays in the northern states (Fig. 5a). In addition, our data-informed model simulates significant spring advances in the central US, while remote sensing studies largely disagree over this area. On the other hand, both the remote sensing studies and our results show that autumn phenology is significantly delayed in the West and Northeast (Figs. 5b and 8b). However, the examined studies do not exhibit significant delays in the northern states, in contrast to our results. In other areas, the trends are either insignificant (southeast and east) or uncertain (center).

3.4 Modeling uncertainties

3.4.1 Impact of interannual variability

Estimates of trends in phenology are sensitive to the length of the examined time frame due to relatively large internal climate variability (Badeck et al., 2004; Iler et al., 2013). Our analyses show that interannual variations may also cause large uncertainties in the estimated phenology trend, especially on short decadal time scales. For example, Keenan et al. (2014) estimated a large advance of 0.48 day yr$^{-1}$ in the spring phenology in both the Harvard Forest and Hubbard Brook sites between 1990 and 2012, and across the Eastern US temperate forest for 2000–2012. Our data-informed modeling
approach estimated a similar change of $0.42 \text{day yr}^{-1}$ between 2000 and 2012 over the eastern US (Fig. S14a), but the trend was largely affected by the record-breaking advance of spring in 2012 (Jolly et al., 2005), especially over the central and eastern states (Figs. S12 and S13). If we exclude this specific year, we achieve an average trend of only $-0.05 \text{day yr}^{-1}$ for 2000–2011, with delayed budburst dates in central and southern states (Fig. S14b). In addition, interannual variability may affect the significance of the derived trend. As shown in Fig. S14a, the advance of spring phenology is not significant for 2000–2012, based on the linear regression, possibly because of the large year-to-year variations and the insignificant changes in air temperature (Fig. S15). A similar result is shown for autumn phenology (Fig. S14d). However, if we extend the analysis period to 1982–2011, the estimated trends and their significance are not affected by the anomalous phenology change in the year 2012 (Fig. S14c and f), suggesting that the estimate of long-term trend is more robust compared to the short-term trend.

3.4.2 Impact of chilling requirement and photoperiod limit

We perform an additional sensitivity experiment (simulation 3) to examine the impact of model structure on the phenology prediction. For spring phenology, model validations have shown that the spring warming (1-phase) models are as efficient as chilling (2-phase) models (Vitasse et al., 2011; Fu et al., 2012a; Migliavacca et al., 2012). In the simulation 3, we remove the limit of chilling requirement on the forcing threshold $F^*$ by defining a fixed forcing threshold. The site-level evaluation shows that this simulation has higher correlations at three out of four sites compared to that with chilling requirement (not shown). Driven with MERRA temperature, the simulation 3 (Fig. S16a) predicts a similar spatial pattern for the trend of budburst date in the US as that in simulation 2 (Fig. 5a), although the former estimates larger advances in central ($0.52 \text{day yr}^{-1}$) and eastern US ($0.43 \text{day yr}^{-1}$). Such stronger signal in the trend of spring phenology could be attributed to the omission of offset effects from the winter
warming (Fig. 6a and b). In the simulation 3, we also remove the cap of photoperiod for autumn dormancy and achieve better correlations between simulations and observations at all sites, though this method tends to generate later dormancy, especially at warm sites (up to 20 days, not shown). Continental-scale simulation without photoperiod limit (Fig. S16b) results in similar trend in autumn phenology as that with photoperiod (Fig. 5b), suggesting that the response to temperature dominates the phenological change in the US deciduous forest.

### 3.4.3 Impact of species aggregation

Tree phenology and its responses to temperature changes are thought to vary among species (Vitasse et al., 2009; Fu et al., 2012a). In this study, however, we do not perform species-specific model calibration and validation, because we focus on phenology on the continental scale. We calibrate model parameters based on the long-term phenological cycle derived from LAI, which represents the mean growing seasonality averaged among species. As a check, we analyze the temperature sensitivity of tree phenology for 13 DBF species at Harvard Forest (Fig. 9). We also calculate the ensemble phenology based on the basal area of each species (the dominant species are red oak (*Quercus rubra*, 60% basal area) and red maple (*Acer rubrum*, 23% basal area)) in order to represent the average phenology at Harvard Forest, which has been used in the site-level evaluation (Fig. 1).

For spring phenology, the mean budburst dates vary by up to 3 weeks among different species, with the earliest being alternated-leaved dogwood (*Cornus alterniflora*) and the latest white oak (*Quercus alba*) (Fig. 9c). Two dominant species, red oak and red maple, have similar year-to-year variations, leading to a similar magnitude of ensemble phenology and the long-term trend (Fig. 9a). The 21 year average of the ensemble budburst date is DOY 126, very close to the DOY 125 derived from LAI (Table 3). Regressions against mean March and April temperature show similar sensitivity of budburst date for most species, especially for red oak (−3.8 days°C⁻¹) and red maple (−3.4 days°C⁻¹) (Fig. 9c). Such similarity also provides us the foundation to val-
idate the simulated interannual variation of spring phenology with the lilac data (Fig. 4). For autumn phenology, the averaged dormancy onset date of red maple is 23 days earlier than that of red oak (Fig. 9d), leading to medium ensemble values (Fig. 9b). The 21 year average of the ensemble dormancy onset date is DOY 306, again close to the estimate of DOY 310 based on LAI (Table 3). The temperature sensitivity of autumn phenology is positive for all species, including similar magnitude of 2.6 days°C⁻¹ for red maple and 2.3 days°C⁻¹ for red oak (Fig. 9f), though the latter is insignificant due to the large year-to-year variations. The species-specific analyses show that calibration based on LAI may capture the representative phenology at deciduous forests, and is not affected by the large deviations among species. Since the eastern US is dominated by oak and maple trees (http://www.nrs.fs.fed.us/atlas/tree/curr_fortypes.html), which we show have very similar temperature sensitivity for both the spring and autumn phenology, we expect that the species aggregation applied in this study may reasonably capture the temperature sensitivity of forest phenology on the continental scale, given that temperature is likely the dominant driver of phenology change for such deciduous forests (Fig. 5).

4 Conclusions and discussion

We performed model inter-comparison to identify the state-of-art scheme for predicting tree phenology of US deciduous forests. An extensive database of ground measurements, including long-term records of phenological events at the site level and short-term records widely scattered on the national scale, was compiled to evaluate the models. The selected models with the lowest AIC values utilized the accumulative temperature summation, with additional constraints of winter chilling on spring phenology and photoperiod on autumn phenology. The 30 year phenology trend of US deciduous forest was explored using the selected models. Consistent with an ensemble of remote-sensing studies, the continental simulation showed a significant advance of 0.34 day yr⁻¹ for spring budburst dates in the East with > 50% coverage of deciduous forests.
ous forests during 1982–2012. However, no significant changes were found over the western, northern, northeastern, and southeastern US. On the other hand, the autumn dormancy onset dates is delayed by 0.20 day yr$^{-1}$ in the northern, 0.14 day yr$^{-1}$ in the northeastern, and 0.16 day yr$^{-1}$ in the western forests, but is not significant elsewhere.

Uncertainties in phenological predictions originate from drivers, parameters, and model structures (Migliavacca et al., 2012). In this study, we minimize uncertainties from meteorological forcings by utilizing an updated reanalysis product and validate the gridded forcings with site-based observations. For the model parameters, we calibrate model parameters with long-term average phenology at four deciduous sites with diverse spatial distribution. This approach was chosen because a well-calibrated phenology model based on a single dataset may have poor performance against external data sets (Chuine et al., 1999; Richardson et al., 2006). The validation shows that the predicted spatial pattern is reasonable and the long-term average matches observations within sampling uncertainty (Figs. 3 and 4). However, due to the data scarcity, all the selected sites are located in temperate areas ranging from 38–46° N, suggesting that the model should be used cautiously at other latitudes and parameters may require re-calibration. For model structure, we perform sensitivity tests both with and without chilling requirements and photoperiod limit and find that the predicted phenology and its change is not sensitive to these constraints at least for the US domain.

Our model inter-comparison does not show a distinct advantage for a specific spring model, suggesting that the model formulation, such as sequential, parallel, and alternating, is not a dominant source of uncertainty for estimates of spring phenology. On the other hand, the evaluation of autumn phenology shows that models with cumulative cold summation and photoperiod limits may better capture the trend of the dormancy onset dates. However, the state-of-art autumn models still have large biases in capturing year-to-year variations. Missing mechanisms, potentially including biotic (e.g. tree age (Caldararu et al., 2014) and species, Vitasse et al., 2009), abiotic (e.g. water stress; Jones et al., 2014), environmental (e.g. accidental frost (Schuster et al., 2014), strong wind, and air pollution; Gallinat et al., 2015) and physiological (e.g. rate
of synthesis (Schaber and Badeck, 2003) and timing of spring flushing; Keenan and Richardson, 2015) factors, may jointly affect leaf fall in a process that is currently not well understood.

Given these uncertainties, our results showed a significant advance of 0.34 day yr\(^{-1}\) for spring budburst dates in the East of US during 1982–2012, while a delay of 0.15 day yr\(^{-1}\) for autumn dormancy onset dates in the Northeast and West. Such long-term changes in phenology are mainly attributed to the trends in temperature, as simulations without chilling requirement and photoperiod limit showed similar phenological changes. Due to either the advances in spring or delays in autumn, tree growth period extends by about 1 week (3–4%) at the 2000s relative to the 1980s, indicating prominent influences of climate change on the carbon cycle and ecological evolution of the US deciduous forests.

The Supplement related to this article is available online at doi:10.5194/bgd-12-6037-2015-supplement.

Acknowledgements. Hubbard Brook phenology data were provided by A. Bailey at the USDA Forest Service, Northern Research Station, Hubbard Brook Experimental Forest. This project was supported in part by the facilities and staff of the Yale University Faculty of Arts and Sciences High Performance Computing Center. T. F. Keenan acknowledges funding from a Macquarie University Research Fellowship.

References


Melaas, E. K., Richardson, A. D., Friedl, M. A., Dragoni, D., Gough, C. M., Herbst, M., Montagnani, L., and Moors, E.: Using FLUXNET data to improve models of spring-


Table 1. Summary of studies estimating phenology trend in US for at least 20 years.

<table>
<thead>
<tr>
<th>Studies</th>
<th>Period</th>
<th>Data sets</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Autumn: significant delay in the Northeast.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Autumn: advance almost everywhere.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Autumn: delay in the West, North, Northeast, and Southeast (except Center).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Autumn: significant delay in the West but almost no changes in the East</td>
</tr>
</tbody>
</table>

NDVI: Normalized Difference Vegetation Index.
AVHRR: Advanced Very High-Resolution Radiometers.
GIMMS: Global Inventory Mapping and Monitoring Studies.
Table 2. Ground measurements of leaf area index (LAI) used to calibrate the phenology model. The location of these sites is denoted on Fig. 1.

<table>
<thead>
<tr>
<th>Site</th>
<th>Name</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Years</th>
<th>n</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>US-MOz</td>
<td>Missouri Ozark</td>
<td>38.74° N</td>
<td>92.2° W</td>
<td>2006–2012</td>
<td>149</td>
<td>Gu et al. (2006)</td>
</tr>
</tbody>
</table>
Table 3. Phenological and climatological parameters for four deciduous forest sites predicted by segmented regressions (Fig. S1) and the selected phenology models (S9 and A4, refer to Table 5).

<table>
<thead>
<tr>
<th>Sites</th>
<th>Annual Temp (°C)</th>
<th>Budburst (day of year)</th>
<th>Grow length (days)</th>
<th>Offset start (day of year)</th>
<th>Offset length (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>US-Ha1</td>
<td>8.0</td>
<td>125</td>
<td>122</td>
<td>30</td>
<td>47</td>
</tr>
<tr>
<td>US-UMB</td>
<td>7.2</td>
<td>124</td>
<td>125</td>
<td>42</td>
<td>45</td>
</tr>
<tr>
<td>US-MMS</td>
<td>12.3</td>
<td>100</td>
<td>103</td>
<td>51</td>
<td>39</td>
</tr>
<tr>
<td>US-MOz</td>
<td>13.3</td>
<td>103</td>
<td>102</td>
<td>41</td>
<td>35</td>
</tr>
</tbody>
</table>

6069
### Table 4. Ground phenology measurements of deciduous trees used to validate the model.

<table>
<thead>
<tr>
<th>Site/network</th>
<th>Category</th>
<th>Duration</th>
<th>Sites</th>
<th>Species</th>
<th>Trees</th>
<th>n</th>
<th>Reference/Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>US-UMB</td>
<td>LAI</td>
<td>1999–2012</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>171</td>
<td><a href="http://ameriflux.ornl.gov/">http://ameriflux.ornl.gov/</a></td>
</tr>
<tr>
<td></td>
<td>Photos</td>
<td>2005–2012</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1265</td>
<td>Sonnentag et al. (2012)</td>
</tr>
<tr>
<td>US-MMS</td>
<td>LAI</td>
<td>1999–2012</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>207</td>
<td><a href="http://ameriflux.ornl.gov/">http://ameriflux.ornl.gov/</a></td>
</tr>
<tr>
<td></td>
<td>Dates</td>
<td>2000–2004</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>4</td>
<td><a href="http://ameriflux.ornl.gov/">http://ameriflux.ornl.gov/</a></td>
</tr>
<tr>
<td></td>
<td>Photos</td>
<td>2008–2012</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
<td>1480</td>
<td>Sonnentag et al. (2012)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Data used to evaluate spatial distribution of simulated phenology. Detailed species information is listed in Table S1.

<sup>b</sup> Data used to evaluate temporal variation of simulated phenology. Detailed species information is listed in Table S2.
Table 5. Summary of phenology models with fit parameters calibrated against the long-term phenology at four US deciduous sites. The detailed parameters for the selected models, S9 and A4, are summarized in Table S3.

<table>
<thead>
<tr>
<th>ID</th>
<th>Model Name</th>
<th>Category</th>
<th>Fixed Parameters</th>
<th>Fit Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Sequential CF1</td>
<td>Spring</td>
<td>0</td>
<td>5 ( (t_1, T_f, T_c, C^<em>, F^</em>) )</td>
</tr>
<tr>
<td>S2</td>
<td>Sequential CF2</td>
<td>Spring</td>
<td>0</td>
<td>4 ( (t_1, T_c, C^<em>, F^</em>) )</td>
</tr>
<tr>
<td>S3</td>
<td>Parallel1 CF1</td>
<td>Spring</td>
<td>0</td>
<td>6 ( (t_1, T_f, T_c, C^*, a, b) )</td>
</tr>
<tr>
<td>S4</td>
<td>Parallel1 CF2</td>
<td>Spring</td>
<td>0</td>
<td>5 ( (t_1, T_c, C^*, a, b) )</td>
</tr>
<tr>
<td>S5</td>
<td>Parallel2 CF1</td>
<td>Spring</td>
<td>0</td>
<td>6 ( (t_1, T_f, T_c, C^*, a, b); t_2 = t_1 )</td>
</tr>
<tr>
<td>S6</td>
<td>Parallel2 CF2</td>
<td>Spring</td>
<td>0</td>
<td>5 ( (t_1, T_c, C^*, a, b); t_2 = t_1 )</td>
</tr>
<tr>
<td>S7</td>
<td>Alternating CF1</td>
<td>Spring</td>
<td>0</td>
<td>4 ( (t_1, T_c, a, b); T_f = T_c; t_2 = t_1 )</td>
</tr>
<tr>
<td>S8</td>
<td>Alternating CF1 t1 fixed</td>
<td>Spring</td>
<td>1 ( t_1 )</td>
<td>3 ( (T_c, a, b); T_f = T_c; t_2 = t_1 )</td>
</tr>
<tr>
<td>S9</td>
<td>Alternating CF1 modified</td>
<td>Spring</td>
<td>3 ( t_1, T_c, r )</td>
<td>2 ( (a, b); T_f = T_c; t_2 = t_1 )</td>
</tr>
<tr>
<td>A1</td>
<td>Jolly-2005 Origin</td>
<td>Autumn</td>
<td>4 ( T_i, T_x, P_i, P_x )</td>
<td>0</td>
</tr>
<tr>
<td>A2</td>
<td>Jolly-2005 Adjusted</td>
<td>Autumn</td>
<td>0</td>
<td>4 ( T_i, T_x, P_i, P_x )</td>
</tr>
<tr>
<td>A3</td>
<td>Delpierre-2009</td>
<td>Autumn</td>
<td>0</td>
<td>6 ( P_{start}, T_b, x, y, Y_{crit}, L_f )</td>
</tr>
<tr>
<td>A4</td>
<td>CDD-photoperiod</td>
<td>Autumn</td>
<td>2 ( t_3, T_b )</td>
<td>4 ( F_b, T_f, P_i, P_x )</td>
</tr>
</tbody>
</table>
Figure 1. Simulation of spring and autumn phenology at four US deciduous broadleaf forest (DBF) sites. The map shows the fraction of US DBF derived from the Advanced Very High Resolution Radiometer (AVHRR). The area with >3% coverage is the domain for this study. Five triangles indicate the locations of sites whose long-term measurements of meteorology and phenology are used for the calibration and/or validation of the model: Harvard Forest (US-Ha1), Hubbard Brook Forest (US-HB1), Morgan–Monroe State Forest (US-MMS), University of Michigan Biological Station Forest (US-UMB), and Missouri Ozark Forest (US-MOz). Phenological dates are recorded at US-Ha1 and US-HB1 during 1992–2012. Measurements of leaf area index (LAI) and photos are used to derive phenology at US-UMB and US-MMS for 1999–2012. Two simulations are performed with the literature-based phenology model, driven by temperatures from either the in situ measurements (blue) or the Modern Era Retrospective-Analysis (MERRA) reanalysis (green). Trend of each time series (units: day yr$^{-1}$) is shown with colors indicating results from observations (red) and simulations (blue or green). Significant trends ($p < 0.05$) are marked with asterisks.
**Figure 2.** Comparison of model performance in the prediction of phenological dates at four US DBF sites among (top) nine spring phenology models and (bottom) four autumn phenology models. The statistical metrics are correlation coefficient, root-mean-square error, and the Akaike Information Criterion. Each point represents the mean values of the statistical metrics at four sites for one model. The error bar represents the range of the metrics. Each model uses the optimized parameters as summarized in Table 5 for the prediction. The red ones are the models used for the continental predictions. Detailed predictions at each site are shown in Figs. S4–S11.
Figure 3. Comparison of the simulated (a, b) budburst and (c, d) dormancy dates with in situ observations (colored circles) from the USA National Phenology Network for 2011–2012. Simulations are performed with the spring model S9 and autumn model A4. The number of the sites and the correlation coefficients are shown in the scatter plots. The separate evaluations in 2011 and 2012 are shown in Figs. S12 and S13.
Figure 4. Correlations (circles) between the predicted budburst dates and observed first-bloom dates from the North American Lilac Network (circle) and first-leaf dates from the USA National Phenology Network (squares). Simulations are performed with the spring model S9 and autumn model A4. The correlation coefficients are calculated for individual trees with at least 6 years of observations during 1982–2012. Correlations with \( p < 0.2 \) are denoted with filled symbols.
Figure 5. Trend in the simulated (a) budburst and (b) dormancy dates for deciduous forests in the US during 1982–2012. Simulations are performed with the spring model S9 and autumn model A4. The results are shown only for the grid squares where the fraction of deciduous forest is larger than 3%. Significant trends ($p < 0.05$) are denoted with dots.
Figure 6. Trend of surface air temperature for (a, b) January, (c, d) April, and (e, f) September over deciduous forest during 1982–2012. The temperature data are from (a, c, e) MERRA reanalyses and (b, d, f) USHCN Network. Significant trends ($p < 0.05$) are denoted with dots (a, c, e) or filled circles (b, d, f). The trends for 2000–2012 are presented in Fig. S15.
Figure 7. The (a) difference and (b) its relative change in the growth length for US deciduous forests between 2000s and 1980s. Significant changes ($p < 0.05$) are denoted with dots.
Figure 8. Comparison of phenology trend over US for (a) spring and (b) autumn estimated by different studies. The US domain is divided into six patches to represent different geographic areas: west, north, northeast, center, east, and southeast. In each patch, different characters represent estimates from different studies over that area. A summary of all studies used for comparison is listed in Table 1. The color of a character indicate the sign of a trend as follows: red is positive, blue is negative, and black is zero or insignificant. A patch is hatched with the same color as the dominant trend if it is non-zero.
Figure 9. Interannual variations of phenological dates and their responses to temperature changes during 1992–2011 for each DBF species at Harvard Forest. The year-to-year (a) budburst and (b) dormancy onset dates are presented for species with observations available for at least 20 years. Values for red oak (*Quercus rubra*, in red), red maple (*Acer rubrum*, in green), and the ensemble average (in blue) based on basal area are highlighted in bold. The differences of species-specific dates relative to the ensembles are presented in (c) for budburst and (d) for dormancy onset. Temperature sensitivity of (e) budburst is calculated as the regressions between year-to-year budburst dates and March–April temperature. Similarly, regressions between dormancy onset dates and September temperature is calculated as the temperature sensitivity of (f) autumn phenology. For the middle and bottom panels, positive values are marked as red while negative ones are in blue. Significant ($p < 0.05$) temperature sensitivity in bottom panel is denoted with filled bar. Full names of species abbreviations are listed in Table S1, except for alternated-leaved dogwood (*Cornus alterniflora* short as COAL) and Hawthorne (*Crataegus* sp. short as CRSP).