Biogeosciences

Probing the past 30-year phenology trend of US deciduous forests

X. Yue[1], N. Unger[1], T. F. Keenan[2], X. Zhang[3], and C. S. Vogel[4]

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

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

Abstract. Phenology is experiencing dramatic changes over deciduous forests in the USA. 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 USA through integrating observations with models. Here, we construct an extensive database of ground measurements of phenological events across the USA, 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 likely the dominant driver of the observed trend during the past 3 decades. Our results show that phenological trends are not uniform over the contiguous USA, 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. Additional sensitivity tests show that intraspecific variations may not influence the predicted phenological trends. 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 temperature variation (Körner and Basler, 2010; Polgar and Primack, 2011; Richardson et al., 2013) and thus exhibits a long-term trend with the changing climate (Badeck et al., 2004; Cleland et al., 2007; Gordo and Sanz, 2009; Jeong et al., 2011). Long-term changes in phenology may affect 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 springs and delayed autumns over the Northern Hemisphere, especially in Europe, during the past several decades (Menzel and Fabian, 1999; Fitter and Fitter, 2002; Menzel et al., 2006; Gordo and Sanz, 2009). However, the extent of regional phenological trends in the USA 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 temperatures before budburst (called chilling requirement), 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; Vitasse, 2013; Caldararu et al., 2014; Laube et al., 2014). Second, most deciduous forests in the USA are found at the mid-latitudes, where temperature increases have not been uniform, and are not as strong as those at high latitudes (Hartmann et al., 2013). 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 the 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) and the European Phenology Network (www.pep725.eu), contain many more species but typically do not have long-term data records (with some exceptions). Remote sensing provides a way to examine phenological changes over large scales but is inherently limited by short timescales or infrequent retrieval times and must be validated using ground measurements. Most 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 data set of phenological observations to calibrate and evaluate 13 models (9 for spring and 4 for autumn) of deciduous tree phenology across the USA. We first calibrate each model using derived phenological dates based on the long-term ground observations of leaf area index (LAI) 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 data set of ground measurements was compiled and used to validate the performance of these models. In total, phenological observations from 1151 sites were used for model validation. 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

Decadal measurements of 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 and Fig. 1). 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 top panel, in the Supplement). Since the measurements are discrete, we estimate the budburst dates (D1, marked on Fig. S1a), 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

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

	Studies	Period	Data sets	Results
A	Reed (2006)	1982–2003	NDVI (AVHRR)	Spring: scattered trends towards advance and delay. Autumn: significant delay in the northeast and northwest
B	Zhang et al. (2007)	1982–2005	NDVI (AVHRR)	Spring: advance in center and east but delay in the north and southeast.
C	Julien and Sobrino (2009)	1981–2003	NDVI (GIMMS)	Spring: advance in the west and east, no trend in the northeast and southeast, delay in the north. Autumn: advance almost everywhere.
D	White et al. (2009)	1982–2006	NDVI (AVHRR)	Spring: no evidence for time trends for most areas with significant delay in the north.
E	Jeong et al. (2011)	1982–2008	NDVI (AVHRR)	Spring: no evidence for time trends for most areas. Autumn: delay in the west, north, northeast, and southeast (except center).
F	Dragoni and Rahman (2012)	1989–2008	NDVI (AVHRR)	Autumn: significant delay in the northeast but insignificant changes in the east and north.
G	Zhu et al. (2012)	1982–2006	NDVI (GIMMS)	Spring: significant delay in the center and east. Autumn: significant delay in the west but almost no changes in the east
H	Buitenwerf et al. (2015)	1981–2012	NDVI (GIMMS)	Spring: advance in the east and southeast, but almost no changes in the northeast, north, and west. Autumn: delay in the east, southeast, northeast, and north, but almost no changes in the west.

NDVI: normalized difference vegetation index; AVHRR: Advanced Very High-Resolution Radiometers; GIMMS: Global Inventory Mapping and Monitoring Studies

panel):

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (O_i - P_i)^2}, \tag{1}$$

where O_i is the observation, P_i is the regression or prediction, and n is the number of samples. The measurements of LAI are not evenly distributed from year to year, and data at some years are too sparse to form the full annual cycle. As a result, we derive the decadal average phenological dates by regressing against all available LAI records at one site. The derived phenological dates are presented in Table 3. The average budburst date at US-MMS and US-MOz is earlier by 3 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 (Table 4 and S1–S2). Data from three out of the four calibration sites (US-Ha1, US-UMB, US-MMS, Table 2 and Fig. 1) are also used for validation; however, we use them in different ways. For calibration, we use the decadal 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. Most of the phenological records are discrete and evaluation of the annual cycle of tree phenology is difficult. Following definitions in earlier literatures (e.g., Zhu et al., 2012; Richardson et al., 2013), we validate spring budburst date (or the onset of growing season, the dates D1 in Fig. S1) and dormancy onset date (or the end of leaf fall period, the dates D2 plus falling length L2 in Fig. S1) from phenology models with the site-level records. The dormancy onset date defined here is based on the canopy level instead of the bud dormancy examined in a recent review paper by Delpierre et al. (2015).

The two New England sites, Harvard Forest (http://harvardforest.fas.harvard.edu/) and Hubbard Brook (US-HB1) Forest (http://www.hubbardbrook.org/), have decadal 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 Supplement). Otherwise, we qualitatively estimate phenological dates based

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.

Site	Name	Latitude	Longitude	Years	n^*	Reference
US-Ha1	Harvard Forest	42.54° N	72.17° W	1998–2008	68	Urbanski et al. (2007)
US-UMB	Univ. of Michigan Biological Station	45.56° N	84.71° W	1999–2007	116	Gough et al. (2008)
US-MMS	Morgan–Monroe State Forest	39.32° N	86.41° W	1999–2010	207	Schmid et al. (2000)
US-MOz	Missouri Ozark	38.74° N	92.2° W	2006–2012	149	Gu et al. (2006)

* n denotes the number of records.
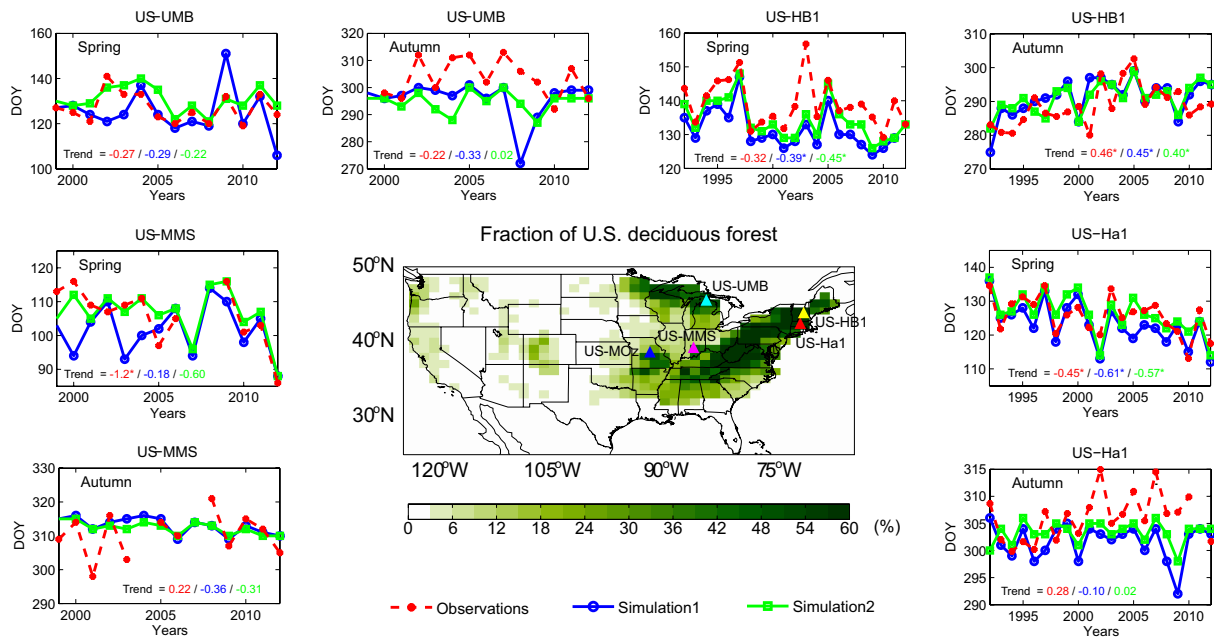
Figure 1. Simulation of spring and autumn phenology at four US deciduous broadleaf forest (DBF) sites. The map shows the fraction of US DBF derived from the Advanced Very High Resolution Radiometer (AVHRR). The area with > 3 % coverage is the domain for this study. Five triangles indicate the locations 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. Derived phenological dates at US-MOz are used for model calibration but not validation and are not shown here. At each site, two simulations are performed with the spring model S9 and autumn model A4 (refer to Fig. 2), driven by temperatures from either the in situ measurements (blue) or the Modern-Era Retrospective Analysis for Research and Applications (MERRA) reanalysis (green). Trend of each time series (units in 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.

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). Based on the changes of tree color in these photos, we can easily identify the phase changes in phenology. For example, changes from gray to light green in spring could occur within several days. We select the middle of these few days as the budburst date. Similarly, changes from brown to gray in late autumn may happen within 1 week and the middle day of the week is selected as the dormancy onset date. An example of autumn dormancy at US-UMB is shown in Fig. S3. The dates derived

from photos may have comparable precision as the observations from site-level phenological records (e.g., Fig. S2c), because the latter are also reported weekly or half-weekly.

Data from ground networks were 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 was significantly enriched thereafter. We select observations during 2011–2012 for 52 deciduous tree species that are most common in the

Table 3. Phenological and climatological parameters for four deciduous forest sites predicted by segmented regressions (Fig. S1) and the selected phenology models (S9+A4, refer to Table 5).

Sites	Annual Temp (°C)	Budburst (day of year)		Grow length (days)		Offset start (day of year)		Offset length (days)	
		LAI-derived	Model	LAI-derived	Model	LAI-derived	Model	LAI-derived	Model
US-Ha1	8.0	125	122 ± 5	30	47 ± 7	271	270 ± 10	39	33 ± 6
US-UMB	7.2	124	125 ± 4	42	45 ± 5	273	265 ± 21	34	34 ± 16
US-MMS	12.3	100	103 ± 5	51	39 ± 4	276	275 ± 6	35	40 ± 6
US-MOz	13.3	103	102 ± 4	41	35 ± 3	270	275 ± 5	45	42 ± 4

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

Site/network	Category	Duration	Sites	Species	Trees	n	Reference/Link
Harvard Forest	Dates	1990–2012	1	16	56	32393	O'Keefe (2000)
Hubbard Brook	Dates	1989–2012	1	3	27	1081	Bailey (2014)
US-UMB	LAI	1999–2012	1	N/A	N/A	171	http://ameriflux.ornl.gov/
	Dates	1999–2012	1	5	66	259	Gough et al. (2008)
	Photos	2005–2012	1	N/A	N/A	1265	Sonnentag et al. (2012)
US-MMS	LAI	1999–2012	1	N/A	N/A	207	http://ameriflux.ornl.gov/
	Dates	2000–2004	1	N/A	N/A	4	http://ameriflux.ornl.gov/
	Photos	2008–2012	1	N/A	N/A	1480	Sonnentag et al. (2012)
National Phenology Network 1[a]	Dates	2011–2012	588	52	1986	29280	https://www.usanpn.org/
National Phenology Network 2[b]	Dates	2004–2012	167	7	195	4231	https://www.usanpn.org/
North American Lilac Network	Dates	1982–2003	392	2	N/A	5072	Schwartz and Reiter (2000)

[a] Data used to evaluate spatial distribution of simulated phenology. Detailed species information is listed in Table S1. [b] Data used to evaluate temporal variation of simulated phenology. Detailed species information is listed in Table S2.

USA (Table S1 in the Supplement). The derived phenological dates for individual trees are averaged if they are observed at the same location (see Supplement). 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. 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 can be di-

vided into two categories, spring warming (or one-phase) and chilling (or two-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 one-phase models are as efficient as two-phase models at the site level (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; Sarvas 1972) 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^*) must 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:

$$S_c(t) = \sum_{t_1}^{t} R_c(x_t), \qquad (2)$$

$$S_f(t) = \sum_{t_2}^{t} R_f(x_t), \qquad (3)$$

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).

$$\text{CF1}: \quad R_c(x_t) = \begin{cases} 0, & x_t \geq T_c \\ 1, & x_t < T_c \end{cases} \qquad (4)$$

$$\text{CF1}: \quad R_f(x_t) = \begin{cases} x_t - T_f, & x_t \geq T_f \\ 0, & x_t < T_f \end{cases} \qquad (5)$$

$$\text{CF2}: \quad R_c(x_t) = \begin{cases} 0, & x_t \leq -3.4 \text{ or } x_t \geq 10.4 \\ \frac{x_t + 3.4}{T_c + 3.4}, & -3.4 < x_t < T_c \\ \frac{x_t - 10.4}{T_c - 10.4}, & T_c < x_t < 10.4 \end{cases} \qquad (6)$$

$$\text{CF2}: \quad R_f(x_t) = \begin{cases} 0, & x_t \leq 0 \\ \frac{28.4}{1 + e^{-0.185(x_t - 18.4)}}, & x_t > 0 \end{cases} \qquad (7)$$

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 $S_c > C^*$. We apply the same fixed thresholds (e.g., 3.4 and 10.4) for Eqs. (6)–(7) as in Chuine et al. (1999); however, we re-calibrate other parameters (e.g., T_c and C^*) so that these functions adapt to the phenological changes in US deciduous forest.

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 the 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 the 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 apply the exhaustive enumeration method to evaluate all combinations of the discrete parameters. We 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:

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

where f_S is spring 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 (Delpierre et al., 2009; Richardson et al., 2013). 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–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):

$$C_a(t) = \sum_{t_3}^{t} R_a(x_t), \qquad (9)$$

$$R_a(x_t) = \begin{cases} T_b - x_t, & x_t < T_b \\ 0, & x_t \geq T_b \end{cases}, \qquad (10)$$

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 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

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. Optimized parameters for other models are summarized in Table S4.

ID	Model Name	Category	Fixed Parameters	Fit Parameters
S1	Sequential CF1	Spring	0	5 (t_1, T_f, T_c, C^*, F^*)
S2	Sequential CF2	Spring	0	4 (t_1, T_c, C^*, F^*)
S3	Parallel1 CF1	Spring	0	6 (t_1, T_f, T_c, C^*, a, b)
S4	Parallel1 CF2	Spring	0	5 (t_1, T_c, C^*, a, b)
S5	Parallel2 CF1	Spring	0	6 (t_1, T_f, T_c, C^*, a, b); $t_2 = t_1$
S6	Parallel2 CF2	Spring	0	5 (t_1, T_c, C^*, a, b); $t_2 = t_1$
S7	Alternating CF1	Spring	0	4 (t_1, T_c, a, b); $T_f = T_c$; $t_2 = t_1$
S8	Alternating CF1 t1 fixed	Spring	1 (t_1)	3 (T_c, a, b); $T_f = T_c$; $t_2 = t_1$
S9	Alternating CF1 modified	Spring	3 (t_1, T_c, r)	2 (a, b); $T_f = \quad T_c$; $t_2 = t_1$
A1	Jolly-2005 Origin	Autumn	4 (T_i, T_x, P_i, P_x)	0
A2	Jolly-2005 Adjusted	Autumn	0	4 (T_i, T_x, P_i, P_x)
A3	Delpierre-2009	Autumn	0	6 (P_{start}, T_b, x, y, Y_{crit}, L_f)
A4	CDD-photoperiod	Autumn	2 (t_3, T_b)	4 (F_s, L_f, P_i, P_x)

falling period is determined by L_f as follows:

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

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),

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

where P is the day length in minutes. P_i and P_x are the lower and upper limits of day length during the period of leaf fall. Following Jolly et al. (2005), the final autumn phenology f_A 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 stand-alone 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),

$$\text{AIC} = n \cdot \log \sigma^2 + 2p + \frac{2p(p+1)}{n - p - 1} \quad (13)$$

where n is the number of samples, p is the number of fit parameters for the model, and σ^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 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 and calibrated forcing threshold F^* of 50 degree days. As a result, forcing value S_f begins accumulation from the winter solstice and budburst occurs if only $S_f > F^*$. The whole process is not dependent on the value of chilling units S_c. Meanwhile, we lift the photoperiod cap for leaf senescence by setting $f_P = 1$, so that the autumn phenology is only determined by temperature ($f_A = f_T$).

We analyze species-specific temperature sensitivity of tree phenology at Harvard Forest (Sect. 4.3). Based on these re-

sults, we perform two additional sensitivity tests to evaluate modeling uncertainties from the intraspecific variations. In the first run (simulation 4), phenological parameters are derived based on records of species with the lowest temperature sensitivity for both spring (sweet birch, *Betula lenta*) and autumn (paper birch, *Betula papyrifera*). In the other run (simulation 5), parameters are derived using records of species with the highest temperature sensitivity for spring (striped maple, *Acer pensylvanicum*) and autumn (black oak, *Quercus velutina*). We applied the derived parameters for the whole domain of the USA by ignoring the realistic fractional coverage of specific species, so as to estimate the maximum uncertainty of prediction due to the intraspecific variations. 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 in the eastern US, where > 90 % deciduous forests are located (Fig. 1). The site-level evaluations for nine spring models and four autumn models are shown in Fig. 2 and summarized in Table S5. 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 five parameters, increasing 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, such as accidental frost, strong wind and rainfall, may be required to improve the current model structures (Richardson et al., 2006; Schuster et al., 2014). 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
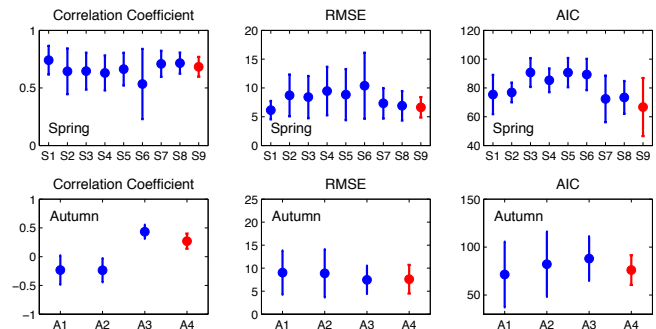
Figure 2. Comparison of model performance in the prediction of phenological dates at four US DBF sites among (top) nine spring phenology models and (bottom) four autumn phenology models. The statistical metrics are correlation coefficient, root mean square error (RMSE), and the Akaike information criterion (AIC). 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. The values of correlation coefficients, RMSE, and AIC are summarized in Table S5.

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 ~ 0.5 day yr^{-1} during 1992–2012. However, at US-HB1, the observed trend of -0.3 day yr^{-1} is not significant due to large interannual variations. In contrast, the dormancy start dates remains almost constant at US-Ha1, similar to that reported by Lee et al. (2003), but exhibits a significant delay of ~ 0.5 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 (~ 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 to 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 west of 105° W but earlier in the low latitudes of the east (Fig. 3a).
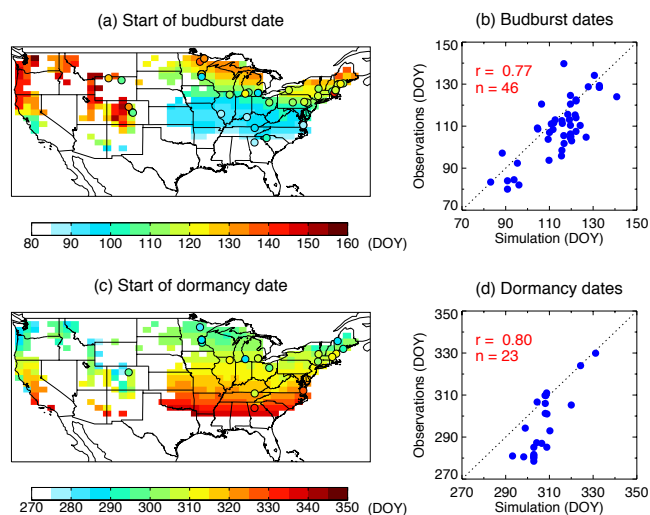
Figure 3. Comparison of the simulated **(a, b)** budburst and **(c, d)** dormancy dates with in situ observations (colored circles) from the USA National Phenology Network for 2011–2012. Simulations are performed with the spring model S9 and autumn model A4. The number of the sites and the correlation coefficients are shown in the scatter plots. The separate evaluations in 2011 and 2012 are shown in Figs. S12 and S13. The coverage of colored patches in **(a)** and **(c)** differs from that in Fig. 1 because values at and beyond the low end of color scales have been shown in white. The number of the sites and the correlation coefficients are shown in the scatter plots. The number of sites shown in the plot is much fewer than the total of 588 because only a small portion of the sites have continuous records for both years and a stringent screening process is applied to derive phenological dates (see Supplement). The separate evaluations in 2011 and 2012 are shown in Figs. S12–S13.

The area-weighted (based on cover fraction of deciduous forest) budburst date for the western USA is 4 May or the 124th day of the year (DOY), with higher values of > DOY 140 over the ridge of Rocky Mountains. In contrast, the mean budburst date is 15 April (DOY 105) east of 105° W, with even earlier dates of < DOY 100 south of 40° N. At higher latitudes, such as the forests over New England and the Great Lakes, spring usually begins after DOY 125 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). However, the predicted budburst date at these sites is 5 days on average later than the observations. Causes of such bias are unclear but might be related to the uneven spatial distribution of network sites, the distinct nature of the protocols between calibration and validation data, and the incompatibility between model parameters derived at regional scale versus those for the continental scale. The distribution of autumn phenology shows an almost opposite pattern to 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 vali-
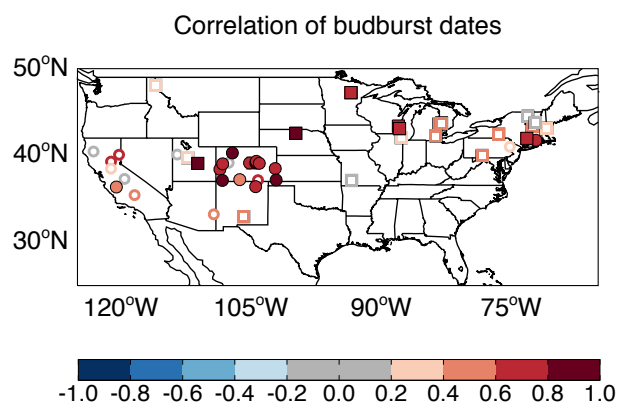
Figure 4. Correlations (circles) between the predicted budburst dates and observed first-bloom dates from the North American Lilac Network (circle) and first-leaf dates from the USA National Phenology Network (squares). Simulations are performed with the spring model S9. The correlation coefficients are calculated for individual trees with at least 6 years of observations during 1982–2012. Correlations with $p < 0.1$ are denoted with filled symbols.

dation against observations from 23 USA-NPN sites yields a significant correlation coefficient of 0.80 for the simulated autumn phenology (Fig. 3d). Similar to the spring budburst, predicted dormancy onset is later by 11 days than the observations. The S9 model also reproduces year-to-year changes in spring phenology. In 2011, the area-weighted budburst date is DOY 117 (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 networks (Fig. 4). The correlations between modeled and observed budburst dates are larger than 0.3 for 47 out of 59 sites, among which 26 are significant ($p < 0.1$), suggesting that the predicted interannual variation and long-term trend of spring phenology are generally reasonable on the continental scale. The insignificant correlations at 33 sites are in part attributed to the deviations in species between model (plant functional type (PFT)-level) and measurements (species level). Furthermore, the record length may also contribute to these biases because 28 out of the 33 sites with insignificant correlations have records shorter than 8 years. The large interannual variability in the spring phenology (see Sect. 4.1) may affect the correlations especially for time series with short record length. On the other hand, no long-term records are available to evaluate the temporal variation of simulated autumn phenology on the continental scale.
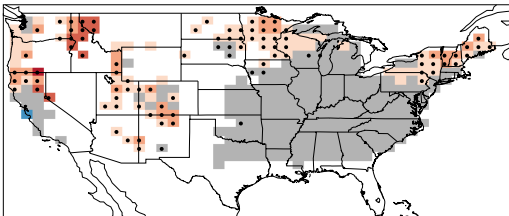
3.2 Phenological change in US deciduous forests

Driven with the MERRA forcing, the model simulates a significant advance of spring budburst dates in the central east-

(a) Trend in start of budburst date for 1982-2012
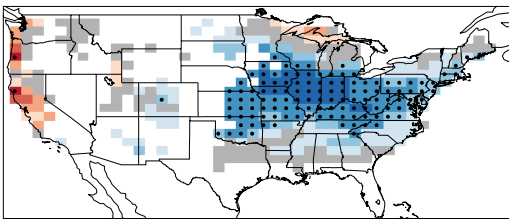
(b) Trend in start of dormancy date for 1982-2012

-0.6　-0.4　-0.2　0.0　0.2　0.4　0.6　(day/yr)

Figure 5. Trend in the simulated **(a)** budburst and **(b)** dormancy dates for deciduous forests in the USA 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.

ern USA 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 the central and southern USA 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. S14). Both the reanalysis data and ground records show a significant spring warming of 0.75 °C decade^{-1} over the central and eastern USA while insignificant changes in the other portion of deciduous forest (Fig. S14c–d). Meanwhile, the warmer winter may delay the spring budburst by reduc-

ing chilling days, especially for forests in the northern USA (Fig. S14a–b). On the other hand, autumn warming in the northern, northeastern, and western forests (Fig. S14e–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. S14e–f) or regulation through photoperiod. Based on the synchronous phenological responses to temperature changes, we estimate long-term temperature sensitivities of −3.3 day °C^{-1} for spring budburst date and 2.2 days °C^{-1} for dormancy start date over US deciduous forests. These values are close to the estimates of −2.8 ± 0.3 day °C^{-1} (spring) and 1.8 ± 0.8 day °C^{-1} (autumn) based on observations from five US deciduous sites (Keenan et al., 2014). The temperature sensitivity of spring budburst date is also within the range from −2.05 to −7.48 days °C^{-1} for different species based on the field experiments performed by Vitasse et al. (2009).

Advanced spring and delayed autumn together increased the length of the growing season across the USA (Fig. 6). 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. 6). 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).

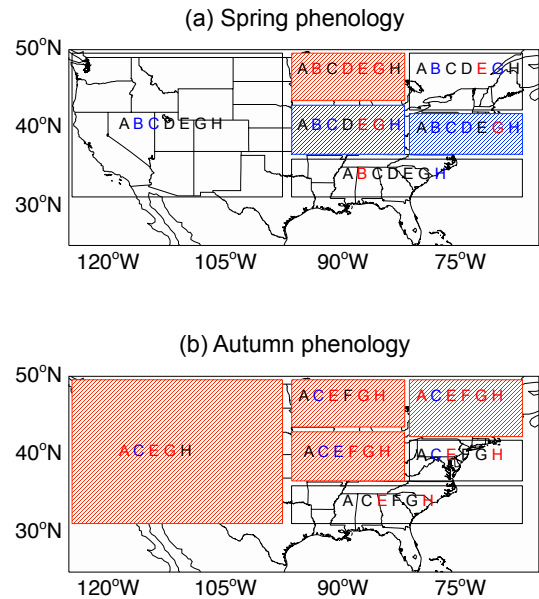### 3.3　Comparison with results from remote sensing

Most up-to-date estimates of the changes in US forest phenology are performed with remote sensing data. In a recent study, Buitenwerf et al. (2015) found an overall extension of the growing season over boreal and temperate forests during 1981–2012 based on the normalized difference vegetation index (NDVI) from satellite data. However, the exact phenological changes that underlie such overall greening differ among regions. For US forests, the longer growing season is primarily driven by later leaf-off dates, though regional advance of spring is also observed. Our results are generally consistent with their conclusions but with some deviations. For example, they observed later autumn in almost all the eastern US, where we predict delays only in the north and northeast (Fig. 5b). Such discrepancies reflect prediction biases, and may also be a consequence of satellite retrieval uncertainties (Table 1).

We further compare our results to recent reports from the literature, selecting all studies that examine phenological trends across the USA for at least 20 years (Table 1). All selected studies use the NDVI, 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
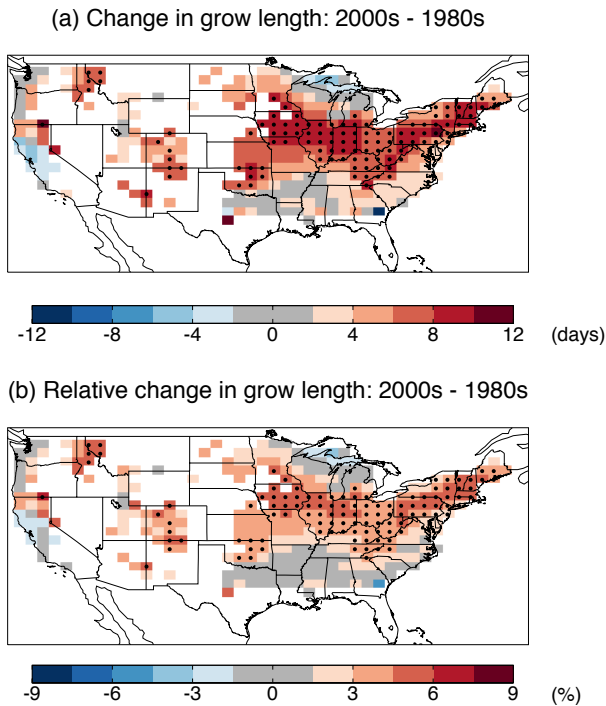
Figure 6. The (a) difference and (b) its relative change in the growth length for US deciduous forests between 2000s and 1980s. Significant changes ($p < 0.05$) are denoted with dots.

Figure 7. Comparison of phenology trend over the USA 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.

in the extraction of the dates (Keenan et al., 2014). Here, we summarize their results on Fig. 7 so as to conclude the most robust changes for US forest phenology in the past 2–3 decades. Since the definition of phenological events varies among different studies (White et al., 2009), we qualitatively compare the simulations with the remote sensing retrievals so as to evaluate the ensemble spatial distribution of phenological changes in the past decades. For spring phenology, four out of seven studies predict advanced budburst or green-up dates in the east, while four predict delayed dates in the north (Fig. 7a). There are no evident phenological changes in the west, northeast, and southeast. Our results show similar changes in spring phenology to 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, north, and northeast (Figs. 5b and 7b). However, the examined studies also exhibit significant delays in the central states, in contrast to our results. In other areas, the trends are insignificant (southeast and east).

4 Discussion

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 timescales. 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 day yr^{-1} between 2000 and 2012 over the eastern USA (Fig. S15a), 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–S13). If we exclude this specific year, we achieve an average trend of only -0.05 day yr^{-1} for 2000–2011, with delayed budburst dates in central and southern states (Fig. S15b). In addition, interannual variability may affect the significance of the derived trend. As shown in Fig. S15a, the advance of spring phenology is
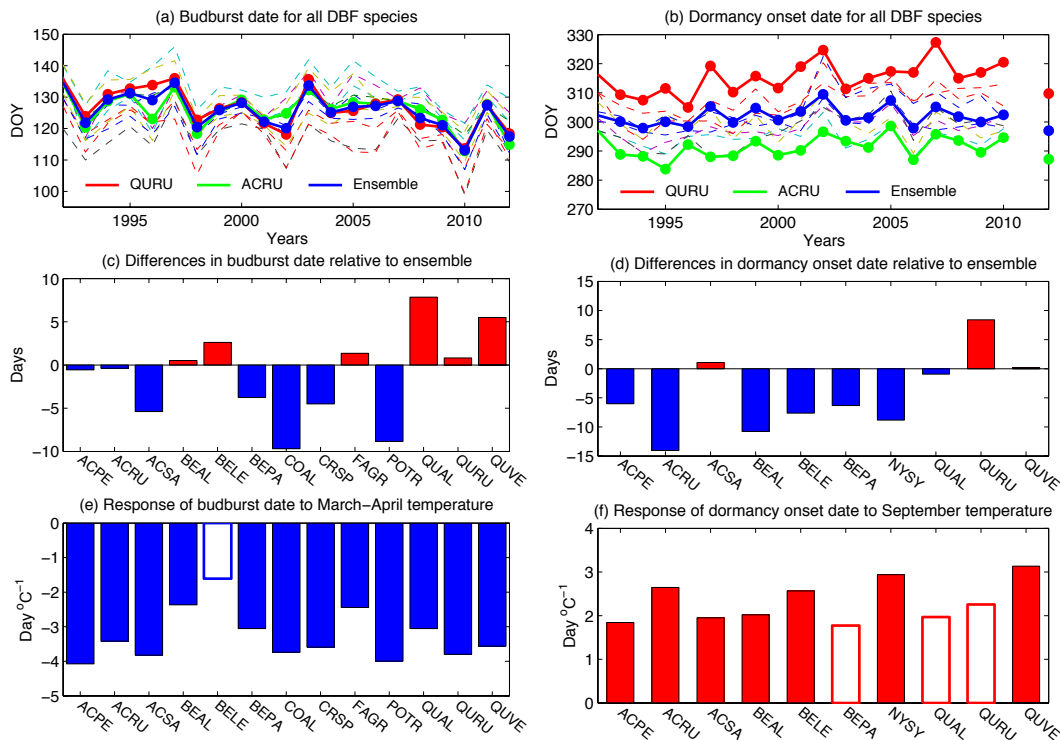
Figure 8. Interannual variations of phenological dates and their responses to temperature changes during 1992–2011 for each DBF species at Harvard Forest. The year-to-year **(a)** budburst and **(b)** dormancy onset dates are presented for species with observations available for at least 20 years. Values 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 alternate-leaved dogwood (*Cornus alternifolia* short as COAL) and Hawthorne (*Crataegus* spp, abbreviated as CRSP).

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. S16). A similar result is shown for autumn phenology (Fig. S15d). 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. S15c and f), suggesting that the estimate of long-term trend is more robust compared to the short-term trend.

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 (one-phase) models are as efficient as chilling (two-phase) models (Vitasse et al., 2011; Fu et al., 2012a; Migliavacca et al., 2012). In 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, simulation 3 (Fig. S17a) predicts a similar spatial pattern for the trend of budburst date in the USA to that in simulation 2 (Fig. 5a), although the former estimates larger advances in the central (0.52 day yr^{-1}) and eastern USA (0.43 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. S14a–b). In 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. S17b) results in similar trends in autumn phenology to that with photoperiod (Fig. 5b), suggesting that the response to temperature dominates the phenological change in US deciduous forests.

Our investigation of the roles of chilling and photoperiod is sensitive to the model structure, climate variability, and data availability. First, the similar performance between spring warming and chilling models might also result from the inaccurate representation of chilling/photoperiod mechanisms. For example, the chilling units used in our parameterization are calculated based on daily average temperatures, while Piao et al. (2015) suggested that leaf unfolding dates during 1982–2011 are triggered by daytime more than by nighttime temperature. The up-to-date autumn phenology model fails to capture interannual variability of dormancy onset (Fig. 2), suggesting that unknown processes are involved in the autumn leaf fall (Keenan and Richardson, 2015). It is unclear whether these processes are related to the variations of photoperiod. Second, the decadal changes in temperature may mask the role of chilling. The trend of winter warming is not significant for most areas in the USA (Fig. S14a), suggesting that chilling requirements have been fulfilled in the past 3 decades. However, it is unclear whether the winter warming will intensify in the future, which may slow the advancement of spring budburst. Third, we choose to calibrate the phenological parameterization at the level of PFT because species-specific measurements are usually incomplete in time and uneven in space. Such incompleteness may influence the accuracy of derived decadal phenological records used for both model calibration and validation. At the same time, PFT-level parameterization may be too broad for the vegetation modeling because it fails to capture intraspecific variations (Van Bodegom et al., 2012; Reichstein et al., 2014). Observations at the community level suggest that the budburst of some species is sensitive to autumn/winter and spring warming but with opposite signs (Cook et al., 2012). In the next subsection, we examine the records of 13 deciduous tree species at Harvard Forest.

4.3 Impact of species aggregation

Tree phenology and its responses to temperature changes have been shown to vary among species (Vitasse et al., 2009; Fu et al., 2012a; Archetti et al., 2013). In this study, however, we calibrate model parameters based on the long-term phenological cycle derived from LAI, which represents the mean growing seasonality averaged among species. We do not perform the species-specific simulation for the following three reasons. First, the species-level measurements are usually not available on the continental scale, which influences both model calibration and validation. Second, species-level modeling increases the complexity and computational costs while decreasing predictive reliability (Prentice et al., 2015). Third, investigations at both site level and continental scale show similar temperature sensitivity of tree phenology between the species-specific and species-aggregation approaches. We analyze the temperature sensitivity of tree phenology for 13 deciduous broadleaf forest (DBF) species at Harvard Forest (Fig. 8). We 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 alternate-leaved dogwood (*Cornus alternifolia*) and the latest white oak (*Quercus alba*; Fig. 8c). 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. 8a). 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 day $^\circ$C^{-1}) and read maple (-3.4 day $^\circ$C^{-1}; Fig. 8c). Such similarity also provides us the foundation to validate 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. 8d), leading to medium ensemble values (Fig. 8b). 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 day $^\circ$C^{-1} for red maple and 2.3 day $^\circ$C^{-1} for red oak (Fig. 8f), 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 USA 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 sensitivities 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).

We perform two sensitivity runs to evaluate the modeling uncertainties due to intraspecific variations at the continental scale (Fig. S18). Simulations with either the lowest (simulation 4) or the highest (simulation 5) temperature sensitivity yield very similar phenological trends to that in the control simulation (simulation 2). In the east, simulation 4 predicts a spring advance by 0.33 day yr^{-1} while simulation 5 predicts an advance by 0.35 day yr^{-1}, both of which are close to the 0.34 day yr^{-1} from the control run. In the west and northeast, both sensitivity runs predict autumn delay by 0.13–0.15 day yr^{-1}, lower than value of 0.14–0.16 day yr^{-1} from control run, suggesting that site-level responses may not be necessarily consistent with responses at the continental scale. Both the similar temperature sensitivity at site level (Fig. 8) and the predicted phenological trends at continental

scale (Fig. S18) support the concept of phenological modeling at the forest and PFT level, and corroborates the further investigation of phenology–climate interactions at the continental and global scale.

5 Conclusions

We performed a model inter-comparison to identify a 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 forests 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 during 1982–2012. However, no significant changes were found over the western, northern, northeastern, and southeastern USA. On the other hand, the autumn dormancy onset dates are 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 are 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 data set 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. 3–4). However, due to the data scarcity, all the selected sites are located in temperate areas ranging from 38 to 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: Vitasse, 2013; Caldararu et al., 2014; and species: Vitasse et al., 2009) and abiotic (e.g. water stress: Jones et al., 2014; accidental frost: Schuster et al., 2014; strong wind and air pollution: Gallinat et al., 2015; and timing of spring flushing: Fu et al., 2014; Keenan and Richardson, 2015) factors, may jointly affect leaf fall in a process that is currently not well understood.

Given these uncertainties, our results show a significant advance of 0.34 day yr^{-1} for spring budburst dates in the east of the USA 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/bg-12-4693-2015-supplement.

References

Akaike, H.: Information theory and an extension of the maximum likelihood principle, in: Proceedings of the Second International Symposium on Information Theory, edited by: Petrov, B. N. and Csaki, F., Akademiai Kiado, Budapest, 267–281, 1973.

Archetti, M., Richardson, A. D., O'Keefe, J., and Delpierre, N.: Predicting Climate Change Impacts on the Amount and Duration of Autumn Colors in a New England Forest, Plos One, 8, e57373, doi:10.1371/journal.pone.0057373, 2013.

Badeck, F. W., Bondeau, A., Bottcher, K., Doktor, D., Lucht, W., Schaber, J., and Sitch, S.: Responses of spring phenology to climate change, New Phytol., 162, 295–309, doi:10.1111/J.1469-8137.2004.01059.X, 2004.

Bailey, A.: Routine Phenology Measurements. Hubbard Brook Data Archive, available at: http://hubbardbrook.org/data/dataset.php?id=51 (last access: 13 October 2014), Durham, NH, 2014.

Basler, D. and Korner, C.: Photoperiod sensitivity of bud burst in 14 temperate forest tree species, Agr. Forest Meteorol., 165, 73–81, doi:10.1016/J.Agrformet.2012.06.001, 2012.

Buitenwerf, R., Rose, L., and Higgins, S. I.: Three decades of multi-dimensional change in global leaf phenology, Nature Climate Change, 5, 364–368, 2015.

Burnham, K. P. and Anderson, D. R.: Model selection and multimodel infere: A practical information-theoretic approach 2nd edition, Springer-Verlag, New York, 488 pp., 2002.

Caldararu, S., Purves, D. W., and Palmer, P. I.: Phenology as a strategy for carbon optimality: a global model, Biogeosciences, 11, 763–778, doi:10.5194/bg-11-763-2014, 2014.

Chuine, I., Cour, P., and Rousseau, D. D.: Selecting models to predict the timing of flowering of temperate trees: implications for tree phenology modelling, Plant Cell Environ., 22, 1–13, doi:10.1046/J.1365-3040.1999.00395.X, 1999.

Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A., and Schwartz, M. D.: Shifting plant phenology in response to global change, Trends Ecol. Evol., 22, 357–365, doi:10.1016/J.Tree.2007.04.003, 2007.

Cook, B. I., Wolkovich, E. M., and Parmesan, C.: Divergent responses to spring and winter warming drive community level flowering trends, P. Natl. Acad. Sci. USA, 109, 9000–9005, doi:10.1073/Pnas.1118364109, 2012.

Delpierre, N., Dufrene, E., Soudani, K., Ulrich, E., Cecchini, S., Boe, J., and Francois, C.: Modelling interannual and spatial variability of leaf senescence for three deciduous tree species in France, Agr. Forest Meteorol., 149, 938–948, doi:10.1016/J.Agrformet.2008.11.014, 2009.

Delpierre, N., Vitasse, Y., Chuine, I., Guillemot, J., Bazot, S., Rutishauser, T., and Rathgeber, C. B. K.: Temperate and boreal forest tree phenology: from organ-scale processes to terrestrial ecosystem models, Ann. For. Sci., doi:10.1007/s13595-015-0477-6, 2015.

Dragoni, D. and Rahman, A. F.: Trends in fall phenology across the deciduous forests of the Eastern USA, Agr. Forest Meteorol., 157, 96–105, doi:10.1016/J.Agrformet.2012.01.019, 2012.

Dufrene, E., Davi, H., Francois, C., le Maire, G., Le Dantec, V., and Granier, A.: Modelling carbon and water cycles in a beech forest Part I: Model description and uncertainty analysis on modelled NEE, Ecol. Model., 185, 407–436, doi:10.1016/J.Ecolmodel.2005.01.004, 2005.

Easterling, D. R., Karl, T. R., Mason, E. H., Hughes, P. Y., and Bowman, D. P.: United States Historical Climatology Network (US HCN) Monthly Temperature and Precipitation Data, ORNL/CDIAC-87, NDP-019/R3, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee, 214 pp., 1996.

Fitter, A. H. and Fitter, R. S. R.: Rapid changes in flowering time in British plants, Science, 296, 1689–1691, doi:10.1126/Science.1071617, 2002.

Forrest, J. and Miller-Rushing, A. J.: Toward a synthetic understanding of the role of phenology in ecology and evolution, Philos. T. R. Soc. B, 365, 3101–3112, doi:10.1098/Rstb.2010.0145, 2010.

Fu, Y., Campioli, M., Deckmyn, G., and Janssens, I. A.: The Impact of Winter and Spring Temperatures on Temperate Tree Budburst Dates: Results from an Experimental Climate Manipulation, Plos One, 7, e47324, doi:10.1371/journal.pone.0047324, 2012a.

Fu, Y., Campioli, M., Van Oijen, M., Deckmyn, G., and Janssens, I. A.: Bayesian comparison of six different temperature-based budburst models for four temperate tree species, Ecol. Model., 230, 92–100, doi:10.1016/J.Ecolmodel.2012.01.010, 2012b.

Fu, Y. S. H., Campioli, M., Vitasse, Y., De Boeck, H. J., Van den Berge, J., AbdElgawad, H., Asard, H., Piao, S. L., Deckmyn, G., and Janssens, I. A.: Variation in leaf flushing date influences autumnal senescence and next year's flushing date in two temperate tree species, P. Natl. Acad. Sci. USA, 111, 7355–7360, doi:10.1073/Pnas.1321727111, 2014.

Gallinat, A. S., Primack, R. B., and Wagner, D. L.: Autumn, the neglected season in climate change research, Trends Ecol. Evol., 30, 169–176, doi:10.1016/J.Tree.2015.01.004, 2015.

Gordo, O. and Sanz, J. J.: Long-term temporal changes of plant phenology in the Western Mediterranean, Glob. Change Biol., 15, 1930–1948, doi:10.1111/J.1365-2486.2009.01851.X, 2009.

Gough, C. M., Vogel, C. S., Schmid, H. P., Su, H. B., and Curtis, P. S.: Multi-year convergence of biometric and meteorological estimates of forest carbon storage, Agr. Forest Meteorol., 148, 158–170, doi:10.1016/J.Agrformet.2007.08.004, 2008.

Gu, L. H., Meyers, T., Pallardy, S. G., Hanson, P. J., Yang, B., Heuer, M., Hosman, K. P., Riggs, J. S., Sluss, D., and Wullschleger, S. D.: Direct and indirect effects of atmospheric conditions and soil moisture on surface energy partitioning revealed by a prolonged drought at a temperate forest site, J. Geophys. Res., 111, D16102, doi:10.1029/2006jd007161, 2006.

Hartmann, D. L., Tank, A. M. G. K., Rusticucci, M., Alexander, L. V., Brönnimann, S., Charabi, Y. A.-R., Dentener, F. J., Dlugokencky, E. J., Easterling, D. R., Kaplan, A., Soden, B. J., Thorne, P. W., Wild, M., and Zhai, P.: Observations: Atmosphere and Surface, in: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Hurrell, J., Marengo, J., Tangang, F., and Viterbo, P., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., 159–254, 2013.

Iler, A. M., Hoye, T. T., Inouye, D. W., and Schmidt, N. M.: Long-Term Trends Mask Variation in the Direction and Magnitude of Short-Term Phenological Shifts, Am. J. Bot., 100, 1398–1406, doi:10.3732/Ajb.1200490, 2013.

Jeong, S. J., Ho, C. H., Gim, H. J., and Brown, M. E.: Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982–2008, Glob. Change Biol., 17, 2385–2399, doi:10.1111/J.1365-2486.2011.02397.X, 2011.

Jolly, W. M., Nemani, R., and Running, S. W.: A generalized, bioclimatic index to predict foliar phenology in response to climate, Glob. Change Biol., 11, 619–632, doi:10.1111/J.1365-2486.2005.00930.X, 2005.

Jones, M. O., Kimball, J. S., and Nemani, R. R.: Asynchronous Amazon forest canopy phenology indicates adaptation to both water and light availability, Environ. Res. Lett., 9, 124021, doi:10.1088/1748-9326/9/12/124021, 2014.

Julien, Y. and Sobrino, J. A.: Global land surface phenology trends from GIMMS database, Int. J. Remote Sens., 30, 3495–3513, doi:10.1080/01431160802562255, 2009.

Kaduk, J. and Heimann, M.: A prognostic phenology scheme for global terrestrial carbon cycle models, Clim. Res., 6, 1–19, 1996.

Keenan, T. F. and Richardson, A. D.: The timing of autumn senescence is affected by the time of spring phenology: implications for predictive models, Glob. Change Biol., 21, 2634–2641, doi:10.1111/gcb.12890, 2015.

Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., Munger, J. W., O'Keefe, J., Schmid, H. P., SueWing, I., Yang, B., and Richardson, A. D.: Net carbon uptake has increased through warming-induced changes in temperate forest phenology, Nature Climate Change, 4, 598–604, doi:10.1038/Nclimate2253, 2014.

Körner, C. and Basler, D.: Phenology Under Global Warming, Science, 327, 1461–1462, doi:10.1126/Science.1186473, 2010.

Laube, J., Sparks, T. H., Estrella, N., and Menzel, A.: Does humidity trigger tree phenology? Proposal for an air humidity based framework for bud development in spring, New Phytol., 202, 350–355, doi:10.1111/Nph.12680, 2014.

Lee, D. W., O'Keefe, J., Holbrook, N. M., and Feild, T. S.: Pigment dynamics and autumn leaf senescence in a New England deciduous forest, eastern USA, Ecol. Res., 18, 677–694, doi:10.1111/J.1440-1703.2003.00588.X, 2003.

Linkosalo, T., Lappalainen, H. K., and Hari, P.: A comparison of phenological models of leaf bud burst and flowering of boreal trees using independent observations, Tree Physiol., 28, 1873–1882, 2008.

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 springtime vegetation activity onset in forest ecosystems, Agr. Forest Meteorol., 171, 46–56, doi:10.1016/J.Agrformet.2012.11.018, 2013.

Menzel, A. and Fabian, P.: Growing season extended in Europe, Nature, 397, 659–659, doi:10.1038/17709, 1999.

Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aasa, A., Ahas, R., Alm-Kubler, K., Bissolli, P., Braslavska, O., Briede, A., Chmielewski, F. M., Crepinsek, Z., Curnel, Y., Dahl, A., Defila, C., Donnelly, A., Filella, Y., Jatcza, K., Mage, F., Mestre, A., Nordli, O., Penuelas, J., Pirinen, P., Remisova, V., Scheifinger, H., Striz, M., Susnik, A., Van Vliet, A. J. H., Wielgolaski, F. E., Zach, S., and Zust, A.: European phenological response to climate change matches the warming pattern, Glob. Change Biol., 12, 1969–1976, doi:10.1111/J.1365-2486.2006.01193.X, 2006.

Menzel, A., Estrella, N., Heitland, W., Susnik, A., Schleip, C., and Dose, V.: Bayesian analysis of the species-specific lengthening of the growing season in two European countries and the influence of an insect pest, Int. J. Biometeorol., 52, 209–218, doi:10.1007/S00484-007-0113-8, 2008.

Migliavacca, M., Sonnentag, O., Keenan, T. F., Cescatti, A., O'Keefe, J., and Richardson, A. D.: On the uncertainty of phenological responses to climate change, and implications for a terrestrial biosphere model, Biogeosciences, 9, 2063–2083, doi:10.5194/bg-9-2063-2012, 2012.

Morin, X., Lechowicz, M. J., Augspurger, C., O' Keefe, J., Viner, D., and Chuine, I.: Leaf phenology in 22 North American tree species during the 21st century, Glob. Change Biol., 15, 961–975, doi:10.1111/J.1365-2486.2008.01735.X, 2009.

Murray, M. B., Cannell, M. G. R., and Smith, R. I.: Date of Budburst of fifteen Tree Species in Britain Following Climatic Warming, J. Appl. Ecol., 26, 693–700, doi:10.2307/2404093, 1989.

O'Keefe, J.: Phenology of Woody Species at Harvard Forest since 1990, Long Term Ecological Research Network, doi:10.6073/pasta/b151c3eb552433a2a94c6f8de489740b, 2000.

Piao, S. L., Cui, M. D., Chen, A. P., Wang, X. H., Ciais, P., Liu, J., and Tang, Y. H.: Altitude and temperature dependence of change in the spring vegetation green-up date from 1982 to 2006 in the Qinghai-Xizang Plateau, Agr. Forest Meteorol., 151, 1599–1608, doi:10.1016/J.Agrformet.2011.06.016, 2011.

Piao, S. L., Tan, J. G., Chen, A. P., Fu, Y. H., Ciais, P., Liu, Q., Janssens, I. A., Vicca, S., Zeng, Z. Z., Jeong, S. J., Li, Y., Myneni, R. B., Peng, S. S., Shen, M. G., and Pennuelas, J.: Leaf onset in the northern hemisphere triggered by daytime temperature, Nat. Commun., 6, 6911, doi:10.1038/Ncomms7911, 2015.

Polgar, C. A. and Primack, R. B.: Leaf-out phenology of temperate woody plants: from trees to ecosystems, New Phytol., 191, 926–941, doi:10.1111/J.1469-8137.2011.03803.X, 2011.

Prentice, I. C., Liang, X., Medlyn, B. E., and Wang, Y.-P.: Reliable, robust and realistic: the three R's of next-generation land-surface modelling, Atmos. Chem. Phys., 15, 5987–6005, doi:10.5194/acp-15-5987-2015, 2015.

Reed, B. C.: Trend analysis of time-series phenology of North America derived from satellite data, Gisci. Remote Sens., 43, 24–38, 2006.

Reichle, R. H., Koster, R. D., De Lannoy, G. J. M., Forman, B. A., Liu, Q., Mahanama, S. P. P., and Toure, A.: Assessment and Enhancement of MERRA Land Surface Hydrology Estimates, J. Climate, 24, 6322–6338, doi:10.1175/Jcli-D-10-05033.1, 2011.

Reichstein, M., Bahn, M., Mahecha, M. D., Kattge, J., and Baldocchi, D. D.: Linking plant and ecosystem functional biogeography, P. Natl. Acad. Sci. USA, 111, 13697–13702, doi:10.1073/Pnas.1216065111, 2014.

Richardson, A. D., Bailey, A. S., Denny, E. G., Martin, C. W., and O'Keefe, J.: Phenology of a northern hardwood forest canopy, Glob. Change Biol., 12, 1174–1188, doi:10.1111/j.1365-2486.2006.01164.x, 2006.

Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G. S., Chen, J. M., Ciais, P., Davis, K. J., Desai, A. R., Dietze, M. C., Dragoni, D., Garrity, S. R., Gough, C. M., Grant, R., Hollinger, D. Y., Margolis, H. A., McCaughey, H., Migliavacca, M., Monson, R. K., Munger, J. W., Poulter, B., Raczka, B. M., Ricciuto, D. M., Sahoo, A. K., Schaefer, K., Tian, H. Q., Vargas, R., Verbeeck, H., Xiao, J. F., and Xue, Y. K.: Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis, Glob. Change Biol., 18, 566–584, doi:10.1111/J.1365-2486.2011.02562.X, 2012.

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., and Toomey, M.: Climate change, phenology, and phenological control of vegetation feedbacks to the climate system, Agr. Forest Meteorol., 169, 156–173, 2013.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G. K., Bloom, S., Chen, J. Y., Collins, D., Conaty, A., Da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and

Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, J. Climate, 24, 3624–3648, doi:10.1175/Jcli-D-11-00015.1, 2011.

Sarvas, R.: Investigations on the Annual Cycle of Development of Forest Trees: II, Autumn dormancy and winter dormancy, Communicationes Instituti Forestalis Fenniae, 211 pp., 1972.

Schmid, H. P., Grimmond, C. S. B., Cropley, F., Offerle, B., and Su, H. B.: Measurements of CO_2 and energy fluxes over a mixed hardwood forest in the mid-western United States, Agr. Forest Meteorol., 103, 357–374, doi:10.1016/S0168-1923(00)00140-4, 2000.

Schuster, C., Kirchner, M., Jakobi, G., and Menzel, A.: Frequency of inversions affects senescence phenology of Acer pseudoplatanus and Fagus sylvatica, Int. J. Biometeorol., 58, 485–498, doi:10.1007/S00484-013-0709-0, 2014.

Schwartz, M. D. and Crawford, T. M.: Detecting energy-balance modifications at the onset of spring, Phys. Geogr., 22, 394–409, 2001.

Schwartz, M. D. and Hanes, J. M.: Intercomparing multiple measures of the onset of spring in eastern North America, Int. J. Climatol., 30, 1614–1626, doi:10.1002/Joc.2008, 2010.

Schwartz, M. D. and Reiter, B. E.: Changes in North American spring, Int. J. Climatol., 20, 929–932, doi:10.1002/1097-0088(20000630)20:8< 929::Aid-Joc557> 3.0.Co;2-5, 2000.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, Glob. Change Biol., 9, 161–185, doi:10.1046/J.1365-2486.2003.00569.X, 2003.

Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A. M., Friedl, M., Braswell, B. H., Milliman, T., O'Keefe, J., and Richardson, A. D.: Digital repeat photography for phenological research in forest ecosystems, Agr. Forest Meteorol., 152, 159–177, doi::10.1016/J.Agrformet.2011.09.009, 2012.

Sparks, T. H. and Menzel, A.: Observed changes in seasons: An overview, Int. J. Climatol., 22, 1715–1725, doi:10.1002/Joc.821, 2002.

Urbanski, S., Barford, C., Wofsy, S., Kucharik, C., Pyle, E., Budney, J., McKain, K., Fitzjarrald, D., Czikowsky, M., and Munger, J. W.: Factors controlling CO_2 exchange on timescales from hourly to decadal at Harvard Forest, J. Geophys. Res., 112, G02020, doi:10.1029/2006JG000293, 2007.

Van Bodegom, P. M., Douma, J. C., Witte, J. P. M., Ordonez, J. C., Bartholomeus, R. P., and Aerts, R.: Going beyond limitations of plant functional types when predicting global ecosystem-atmosphere fluxes: exploring the merits of traits-based approaches, Global Ecol. Biogeogr., 21, 625–636, doi:10.1111/J.1466-8238.2011.00717.X, 2012.

Vitasse, Y.: Ontogenic changes rather than difference in temperature cause understory trees to leaf out earlier, New Phytol., 198, 149-155, doi:10.1111/Nph.12130, 2013.

Vitasse, Y., Delzon, S., Dufrene, E., Pontailler, J. Y., Louvet, J. M., Kremer, A., and Michalet, R.: Leaf phenology sensitivity to temperature in European trees: Do within-species populations exhibit similar responses?, Agr. Forest Meteorol., 149, 735–744, doi:10.1016/J.Agrformet.2008.10.019, 2009.

Vitasse, Y., Francois, C., Delpierre, N., Dufrene, E., Kremer, A., Chuine, I., and Delzon, S.: Assessing the effects of climate change on the phenology of European temperate trees, Agr. Forest Meteorol., 151, 969–980, doi:10.1016/J.Agrformet.2011.03.003, 2011.

White, M. A., Thornton, P. E., and Running, S. W.: A continental phenology model for monitoring vegetation responses to interannual climatic variability, Global Biogeochem. Cy., 11, 217–234, doi:10.1029/97gb00330, 1997.

White, M. A., de Beurs, K. M., Didan, K., Inouye, D. W., Richardson, A. D., Jensen, O. P., O'Keefe, J., Zhang, G., Nemani, R. R., van Leeuwen, W. J. D., Brown, J. F., de Wit, A., Schaepman, M., Lin, X. M., Dettinger, M., Bailey, A. S., Kimball, J., Schwartz, M. D., Baldocchi, D. D., Lee, J. T., and Lauenroth, W. K.: Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982-2006, Glob. Change Biol., 15, 2335–2359, doi:10.1111/J.1365-2486.2009.01910.X, 2009.

Zhang, X. Y., Tarpley, D., and Sullivan, J. T.: Diverse responses of vegetation phenology to a warming climate, Geophys. Res. Lett., 34, L19405, doi:10.1029/2007gl031447, 2007.

Zhao, M. F., Peng, C. H., Xiang, W. H., Deng, X. W., Tian, D. L., Zhou, X. L., Yu, G. R., He, H. L., and Zhao, Z. H.: Plant phenological modeling and its application in global climate change research: overview and future challenges, Environ. Rev., 21, 1–14, doi:10.1139/Er-2012-0036, 2013.

Zhu, W. Q., Tian, H. Q., Xu, X. F., Pan, Y. Z., Chen, G. S., and Lin, W. P.: Extension of the growing season due to delayed autumn over mid and high latitudes in North America during 1982–2006, Global Ecol. Biogeogr., 21, 260–271, doi:10.1111/J.1466-8238.2011.00675.X, 2012.