

A new seasonal-deciduous spring phenology submodel in the Community Land Model 4.5: impacts on carbon and water cycling under future climate scenarios

MIN CHEN¹, ELI K. MELAAS², JOSH M. GRAY², MARK A. FRIEDL² and ANDREW D. RICHARDSON¹

¹Department of Organismic and Evolutionary Biology, Harvard University, 22 Divinity Avenue, Cambridge, MA 02138, USA,

²Department of Earth and Environment, Boston University, 675 Commonwealth Avenue, Boston, MA 02215, USA

Abstract

A spring phenology model that combines photoperiod with accumulated heating and chilling to predict spring leaf-out dates is optimized using PhenoCam observations and coupled into the Community Land Model (CLM) 4.5. In head-to-head comparison (using satellite data from 2003 to 2013 for validation) for model grid cells over the Northern Hemisphere deciduous broadleaf forests (5.5 million km²), we found that the revised model substantially outperformed the standard CLM seasonal-deciduous spring phenology submodel at both coarse (0.9 × 1.25°) and fine (1 km) scales. The revised model also does a better job of representing recent (decadal) phenological trends observed globally by MODIS, as well as long-term trends (1950–2014) in the PEP725 European phenology dataset. Moreover, forward model runs suggested a stronger advancement (up to 11 days) of spring leaf-out by the end of the 21st century for the revised model. Trends toward earlier advancement are predicted for deciduous forests across the whole Northern Hemisphere boreal and temperate deciduous forest region for the revised model, whereas the standard model predicts earlier leaf-out in colder regions, but later leaf-out in warmer regions, and no trend globally. The earlier spring leaf-out predicted by the revised model resulted in enhanced gross primary production (up to 0.6 Pg C yr⁻¹) and evapotranspiration (up to 24 mm yr⁻¹) when results were integrated across the study region. These results suggest that the standard seasonal-deciduous submodel in CLM should be reconsidered, otherwise substantial errors in predictions of key land–atmosphere interactions and feedbacks may result.

Keywords: carbon cycle, climate change, Community Land Model, ecosystem services, PhenoCam, phenology, water

Received 18 August 2015 and accepted 23 March 2016

Introduction

The vast boreal and temperate deciduous forests of the Northern Hemisphere are thought to account for a substantial fraction of the terrestrial carbon sink (Houghton, 2007; Luysaert *et al.*, 2007; Pan *et al.*, 2011). In these ecosystems, vegetation phenology controls numerous land surface characteristics including albedo (Hollinger *et al.*, 2010), microclimate (Richardson & O’Keefe, 2009), canopy roughness and conductance (Blanken & Black, 2004), and the exchanges of carbon and water between land and atmosphere (Richardson *et al.*, 2013). Phenology thus plays an important role in mediating vegetation feedbacks to the climate system (Peñuelas *et al.*, 2009). For deciduous trees, phenological transitions are usually modeled as a function of air temperature and photoperiod, and sometimes soil temperature and soil moisture (Richardson *et al.*, 2012).

However, existing land surface models generally employ poor phenological submodels for deciduous forests, which leads to biased estimates of forest–atmosphere fluxes and feedbacks (Keenan *et al.*, 2012; Richardson *et al.*, 2012).

The Community Land Model (CLM) simulates land surface processes in the Community Earth System Model (CESM) and is one of the most widely used land surface models for regional and global simulations of land–atmosphere exchanges. In CLM, vegetation phenology plays an essential role in almost all biophysical and biogeochemical processes on the land surface. In the most recently released version, CLM 4.5, three different vegetation phenology submodels are used for natural ecosystems (Lawrence *et al.*, 2011). First, the seasonal-deciduous submodel, which is used for boreal and temperate deciduous forests, has distinct growing and dormant seasons, and temperature and photoperiod determine the periods of leaf development and senescence, each of which occurs only once per year. Second, the stress-deciduous submodel, which applies to grasslands, shrublands, and tropical drought-deciduous forests, is similar to the seasonal-deciduous

Correspondence: Min Chen, Department of Global Ecology, Carnegie Institution for Science, 260 Panama Street, Stanford, CA 94305, USA, tel. +1 650-462-1047, fax +1 650-462-5968, e-mail: chenminbnu@gmail.com

model, but vegetation activity is limited by water availability and/or temperature, and there may be multiple growing cycles per year. Third, in the evergreen submodel, which is used for evergreen forests and shrubs, carbon allocation to new foliage occurs whenever photosynthesis occurs, while in parallel, a background rate of litterfall results in continuous shedding of foliage. Although limitations of the phenology submodels in CLM have been acknowledged for some time (Lawrence *et al.*, 2011), to date there have been only limited efforts to improve them (e.g. Dahlin *et al.*, 2015).

In this study, we aim to improve CLM's seasonal-deciduous spring phenology submodel using a new formulation derived for boreal and temperate deciduous broadleaf forests in the Northern Hemisphere, and more importantly, we focus on evaluating how the new spring phenology algorithm may influence carbon and water cycles in these forests under future climate scenarios. We conduct simulations and compare the results from the standard version of CLM 4.5 with results from a revised version that incorporates the improved phenology submodel. We investigate the following questions: (1) How do the two submodels compare in regard to predictions of spring leaf-out across Northern Hemisphere deciduous broadleaf forests? What are the spatial patterns, and which submodel agrees best with remotely sensed land surface phenology? (2) Does either model predict significant trends toward earlier or later dates of spring leaf-out under future climate scenarios? (3) How are model estimates of land-atmosphere carbon and water exchanges in the coming century affected by choice of phenology submodel?

Materials and methods

CLM seasonal-deciduous spring phenology

The seasonal-deciduous phenology algorithm in CLM 4.5 is directly adapted from an earlier ecosystem model, BIOME-BGC v. 4.1.2 (Thornton *et al.*, 2002). The timing of spring leaf-out is triggered when the accumulated growing degree days (GDD) exceed a threshold GDD_{crit} . The GDD temperature sum is calculated using the third-layer soil temperature (T_s , in K), with a base temperature equal to the water freezing temperature (T_f , 273.15 K). The accumulation of GDD begins at the winter solstice. Thus, for time step n , where Δt is the duration of the time step (in seconds), and DL is the day length (86 400 s):

$$GDD_n = GDD_{n-1} + \max(T_s - T_f, 0) \times (\Delta t / DL) \quad (1)$$

Once $GDD_n > GDD_{crit}$, leaf-out is triggered, effectively activating the growing season. Here, we refer to this date as the start of spring (SOS). In the CLM 4.5 seasonal-deciduous phenology algorithm, GDD_{crit} is calculated from the annual

average of 2 m air temperature ($T_{2m,avg,ann}$) in the preceding year:

$$GDD_{crit} = \exp[4.8 + 0.13(T_{2m,avg,ann} - T_f)] \quad (2)$$

PhenoCam spring phenology model

The PhenoCam network (<http://PhenoCam.sr.unh.edu>) was established to provide automated, near-surface remote sensing of vegetation phenology across North America using repeat digital photography (Sonnentag *et al.*, 2012). For a designated region of interest within each camera field of view, image time series are processed to the Green Chromatic Coordinate (GCC) index, from which estimates of SOS (in this context, defined as the spring date at which GCC reaches 50% of its seasonal amplitude) can be derived using curve-fitting methods, as described by Klosterman *et al.* (2014).

In a previous paper, Melaas *et al.* (2016) parameterized 13 different spring phenology models using Klosterman *et al.*'s (2014) PhenoCam dataset, which includes observations from 13 deciduous forest sites located across a 12 °C gradient in mean annual temperature and a more than 800 mm yr⁻¹ gradient in annual precipitation. The 'best' model was selected using the small-sample corrected Akaike information criterion (Burnham & Anderson, 2002) based on the residual sum of squared errors for observations (55 site-years of data). Here, we use the 'best' model from Melaas *et al.* (2016), but re-parameterized it using an additional 2 years of data (2010 and 2012) that were originally withheld for model testing. We refer to this model as the 'PhenoCam spring phenology model'. In total, 80 site-years of data were used to parameterize our model.

Similar to the CLM seasonal-deciduous spring phenology submodel, leaf-out in the PhenoCam model is predicted to occur when the accumulated GDD exceed a threshold GDD_{crit} . However, similar to the classic Alternating model (Cannell & Smith, 1983), in the PhenoCam spring phenology model, GDD_{crit} varies as a function of the accumulated chilling units (CU), and GDD and CU both accumulate beginning on date t_0 (optimized, conditional on the PhenoCam observations, to day of year 74, or March 14), relative to a single base temperature T_c (optimized to -3.32 °C). CU accumulate only if the daily mean air temperature $T_a < T_c$:

$$CU = CU + 1 \quad (3)$$

where GDD accumulate only if $T_a > T_c$:

$$GDD = GDD + (T_a - T_c) \quad (4)$$

and GDD_{crit} is calculated as:

$$GDD_{crit} = a + b \times \exp(c \times CU) \quad (5)$$

where $a = 207.87$, $b = 244.72$ and $c = -0.013$ are optimized parameters.

We used the PhenoCam spring phenology model to replace the seasonal-deciduous spring phenology submodel in CLM and used CLM's daily mean air temperature to drive the submodel. In this study, we refer to CLM runs using this new submodel as 'CLM-PhenoCam' runs.

Model evaluation and forward runs

We conducted two sets of model runs using the CLM spring phenology submodel and the PhenoCam spring phenology submodel. The first runs were for model evaluation and were conducted using a combination of coupled (i.e. with the phenology submodels embedded within CLM) and offline (i.e. just the phenology submodels on their own) runs. For the forward runs, from 2014 to 2100, the phenology submodels were embedded within CLM so that we could evaluate the impacts of future phenological change on global carbon and water cycling.

Model evaluation. For model evaluation, we conducted a number of different hindcast analyses. To evaluate the phenology submodel predictions against global grid-scale estimates of SOS, derived from satellite remote sensing, we ran CLM from 2000 to 2013, starting from initial conditions in 2000 provided by the standard release of CESM 1.2.0, and using the transient Climate Research Unit – National Centers for Environmental Prediction (CRUNCEP) meteorological forcing data (<http://dods.extra.cea.fr/data/p529viov/cruncep/readme.htm>). We then evaluated predicted SOS dates from both phenology models against SOS dates determined from Moderate Resolution Imaging Spectroradiometer (MODIS, using data from 2003 to 2013) satellite imagery. Previous work (e.g. Hufkens *et al.*, 2012; Klosterman *et al.*, 2014) has shown that SOS dates derived from visual inspection of PhenoCam images are highly correlated with SOS dates derived from PhenoCam GCC thresholds (with little or no bias), which are in turn in good agreement with start-of-spring dates derived from MODIS thresholds. We used OGI [onset of greenness increase, corresponding to 10% of the seasonal amplitude in the logistic-function-fitted enhanced vegetation index (EVI) time series] and MAT (greenness maturity, corresponding to 90% of the seasonal amplitude in the logistic-function-fitted EVI time series) dates from the MODIS Collection 5 Land Cover Dynamics product (MCD12Q2; 500 m resolution, calculated from the EVI; Zhang *et al.*, 2003; Ganguly *et al.*, 2010). To account for a small bias in OGI relative to surface observations (Klosterman *et al.*, 2014), we calculated on a pixel-by-pixel basis, by linear interpolation, MODIS-derived SOS as the date at which 20% of the seasonal amplitude was achieved (Melaas *et al.*, 2016). We excluded pixels where the interannual variation (as measured by the standard deviation of the SOS) was more than 20 days, which is more than double the typical standard deviation observed in budburst dates at either Harvard Forest (Richardson & O’Keefe, 2009), or in the PEP725 dataset (described below). We further excluded pixels where the mean SOS date was later than day of year 200, assuming that dates after this day would tend to suggest either bad MODIS retrievals, or that the pixels are not actually temperate deciduous forest.

Within each model grid cell, we calculated the mean SOS only using the pixels classified as deciduous broadleaf or mixed forest by the MODIS Collection 5 Land Cover product (MCD12Q1; Friedl *et al.*, 2010). However, we evaluated model predictions only for grid cells where at least 5% of the pixels were classified as deciduous broadleaf forest or mixed forest.

To address concerns about scale mismatch and uncertainties associated with aggregating MODIS data to coarse model grid cells, we also evaluated both phenological models at a much finer spatial resolution (1 km). For these offline (i.e. not embedded within CLM) runs conducted for the eastern United States, we used Daymet (Thornton *et al.*, 2014) meteorological forcing data, and we calculated the mean SOS for pixels classified as deciduous forest across a 3×3 MODIS pixel window centered on the Daymet grid cell. Following Melaas *et al.* (2016), we only evaluated the models for those grid cells where at least three of nine MODIS pixels were classified as deciduous broadleaf or mixed forest.

Finally, for evaluation against long-term ground observations, we conducted an analysis using the Pan-European Phenology Project dataset, known as PEP725 (<http://www.pep725.eu>). Because of the length of the available time series, this dataset is well suited to trend analysis. We used records of leaf unfolding (code 11) for the common tree genera of northern Europe (including *Acer*, maple; *Betula*, birch; *Fagus*, beech; *Fraxinus*, ash; *Juglans*, walnut; *Populus*, aspen; *Quercus*, oak; and *Tilia*, basswood), including those sites where data were available for 90% (or more) of the years between 1950 and 2014. Offline runs of both phenology submodels were driven by daily mean temperature from the E-OBS 0.25-degree gridded dataset (Haylock *et al.*, 2008), and we averaged all species within a model grid cell for further analysis.

The above analyses were conducted at various spatial scales therefore meteorological forcing data (i.e. CRUNCEP, Daymet and E-OBS) with different spatial resolution were employed. In each of the above analyses, we calculated the root mean square error (RMSE), mean bias error (MBE), and Pearson correlation coefficient (r) between data (i.e. MODIS or PEP725) and model predicted SOS at each model grid cell. For the grid-scale MODIS runs and the PEP725 runs, we also calculated the slopes of the linear trends of the SOS time series for both the observational data and for each of the two phenology submodels.

Forward runs. For the forward model runs (2014–2100), we used transient meteorological forcing and atmospheric CO₂ concentrations simulated by CCSM under Representative Concentration Pathways (RCP) scenarios 8.5 (high emissions) and 4.5 (medium emissions; Meehl *et al.*, 2013). Our objective with the forward runs was to investigate how future shifts in spring phenology might influence land–atmosphere exchanges of carbon and water under future climate regimes. We used the system state variables at the end of 2013 from our hindcast standard CLM run as the initial conditions for the forward runs.

Community Land Model runs were conducted at a spatial resolution of 0.9×1.25 degrees and a time step 30 min with the ‘BGC’ option turned on to incorporate the newest biogeochemistry developments in CLM 4.5 (Oleson *et al.*, 2013). Our analyses of model output focuses on SOS dates as well as carbon [gross primary production (GPP); net primary production (NPP); and autotrophic respiration (AR)] and water [evapotranspiration (ET)] fluxes integrated over Northern Hemisphere deciduous broadleaf forests (including both boreal and temperate broadleaf deciduous forests), the range of which is

based on the prescribed plant functional type distributions in CLM 4.5 default settings (Oleson *et al.*, 2013), as derived from the advanced very high resolution radiometer continuous fields tree cover dataset (Defries *et al.*, 2000). Grid cells that contain any fraction of deciduous forest are included into the CLM domain for further analysis. Our study domain is therefore larger than the spatial extent of deciduous broadleaf forests in MCD12Q1 (Fig. 1).

Results

Model evaluation

The spatial patterns of SOS predicted by the CLM and CLM-PhenoCam phenology submodels are broadly similar to MODIS-derived SOS dates (Fig. 1). Over the period from 2003 to 2013, SOS occurred later in mid- and high-latitude regions, and earlier at lower latitudes. This spatial pattern is largely driven by the strong temperature gradient from north to south. SOS from CLM was generally earlier than SOS from CLM-PhenoCam (0.7 ± 1.0 days, mean \pm 1 SD across years), with the models in closest agreement in middle latitudes between 30°N and 40°N and showing largest

disagreement in lower ($20\text{--}25^{\circ}\text{N}$) and higher latitudes ($65\text{--}70^{\circ}\text{N}$; Fig. 2a). More specifically, differences are small in mid-latitude regions of eastern North America and Eastern Europe, but are particularly large in boreal North America and subtropical regions of Southern Asia (Fig. 1d).

At the model grid scale ($0.9 \times 1.25^{\circ}$), SOS predictions from CLM are somewhat less consistent with MODIS-derived SOS than are SOS predictions from CLM-PhenoCam. Across all Northern Hemisphere deciduous broadleaf forests, the mode RMSE is 8 days (median, 10 days) for CLM, compared with a mode of 4 days (median, 6 days) for CLM-PhenoCam (Fig. 3a). SOS dates for CLM are also biased early (mode, -7 days; median, -3 days) compared to MODIS-derived SOS. SOS bias for CLM-PhenoCam is much smaller (mode, -2 ; median, -1 days; Fig. 3b). Finally, the correlation coefficient between predicted and MODIS-derived SOS dates is generally weaker for CLM (mode, $r = 0.7$) than CLM-PhenoCam (mode, $r = 0.9$; Fig. 3c).

We obtained similar results when the two phenology models were run at finer spatial resolution and evaluated at that scale. For example, forced with Daymet

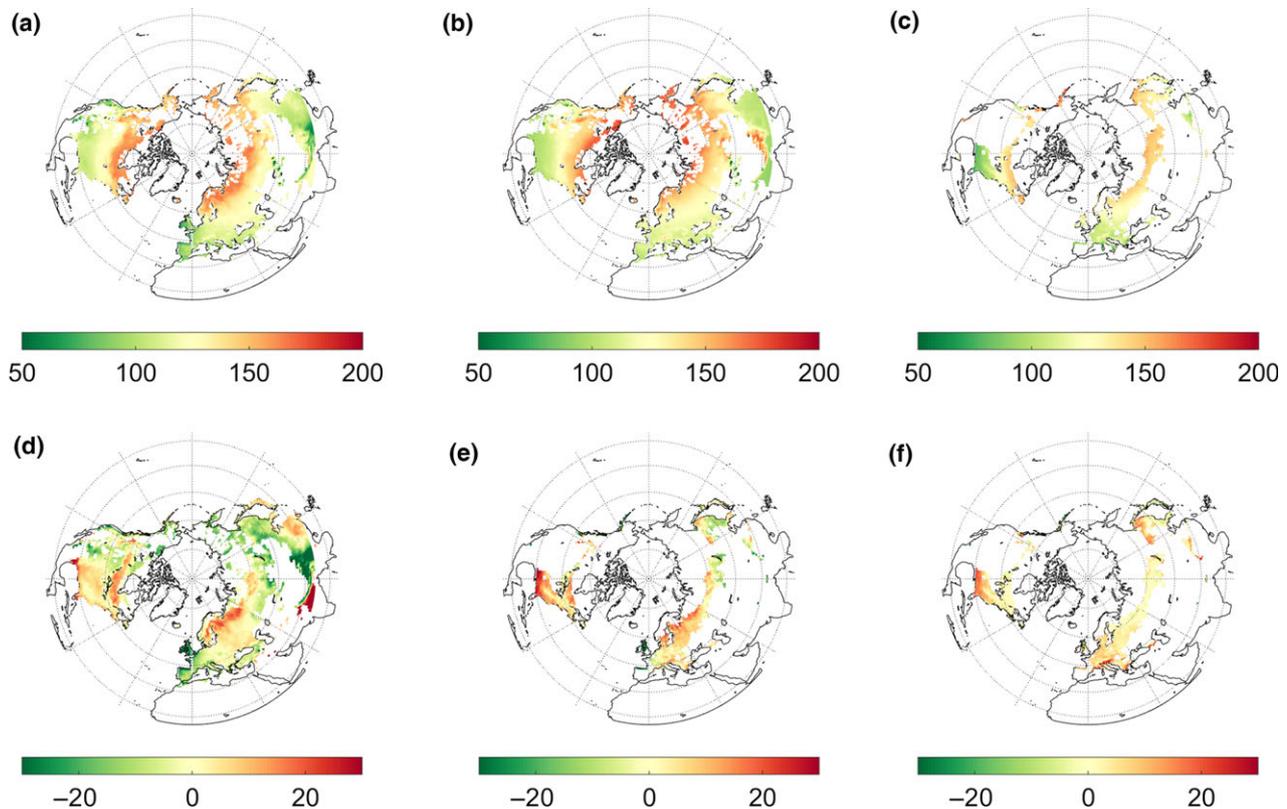


Fig. 1 Average start of spring (SOS) dates predicted by CLM and CLM-PhenoCam, compared with MODIS derived SOS, over the period 2003–2013: (a) CLM predicted SOS; (b) CLM-PhenoCam predicted SOS; (c) MODIS-derived SOS; (d) Differences between CLM and CLM-PhenoCam SOS (a–b); (e) Differences between CLM and MODIS SOS (a–c); (f) Differences between CLM-PhenoCam and MODIS SOS (b–c).

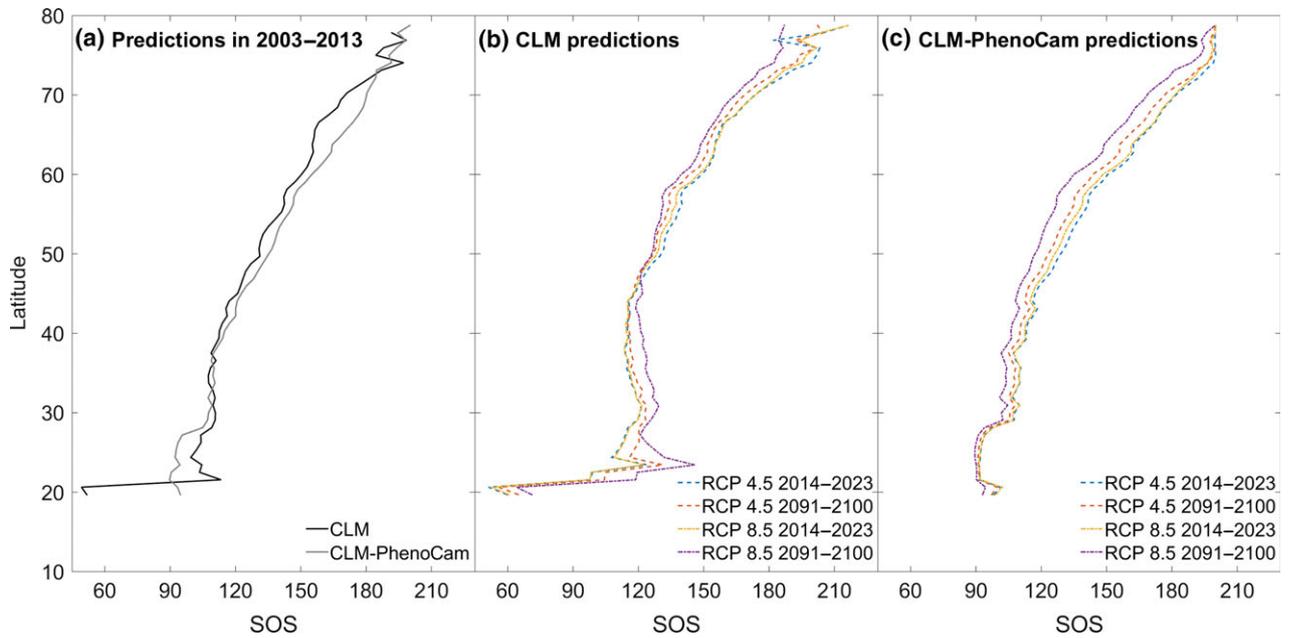


Fig. 2 Latitudinal mean of the start of spring (SOS) dates predicted by CLM and CLM-PhenoCam: (a) hindcast predictions, 2001–2013; (b) CLM predictions at the beginning (2014–2023) and end (2091–2100) of the forward runs under Representative Concentration Pathway (RCP) 4.5 and 8.5; and (c) CLM-PhenoCam predictions at the beginning (2014–2023) and end (2091–2100) of the forward runs under RCP 4.5 and 8.5.

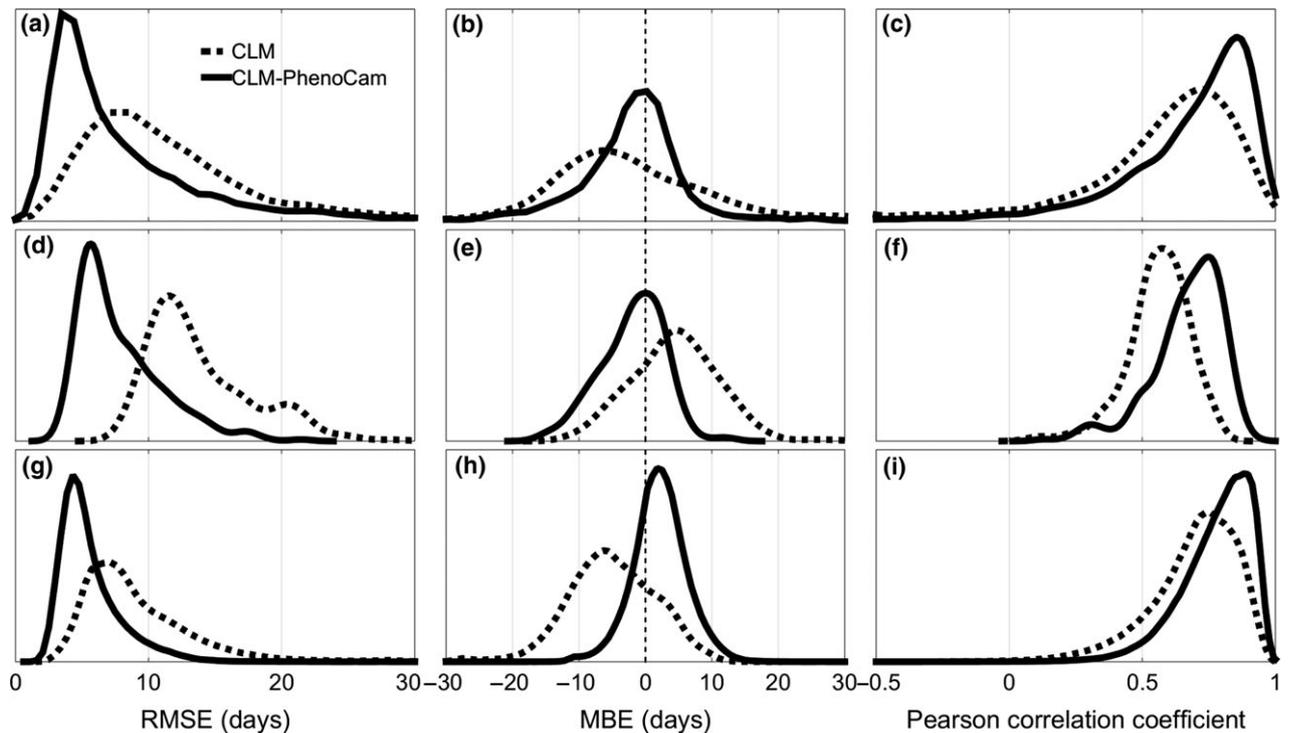


Fig. 3 Probability density estimates of the root mean square error (RMSE), mean bias error (MBE), and Pearson correlation coefficient (r), for start of spring (SOS) predicted by two models: CLM and CLMPhenoCam. (a–c): comparison against Moderate Resolution Imaging Spectroradiometer (MODIS)-derived SOS across the Northern Hemisphere deciduous broadleaf forest ($0.9 \times 1.25^\circ$ grid cells); (d–f): comparison against PEP725 data (0.25° grid cells); (g–i): comparison against MODIS-derived SOS across the eastern US deciduous forest (1 km grid cells).

data at 1 km resolution for the eastern United States, and evaluated against MODIS data aggregated to 3×3 pixel windows, the PhenoCam spring phenology model has a lower RMSE (mode, 4 days; median, 5 days) than the standard CLM spring phenology model (mode, 7 days; median, 8 days). The PhenoCam model also performs better in terms of smaller bias (mode MBE of 2 days vs. -5 days, for the two models, respectively) and higher correlation coefficient (mode r of 0.9 vs. 0.8, again for the two models, respectively; Fig. 3g–i). And, when run with E-OBS data at 0.25° resolution for northern Europe, and evaluated against PEP725 data, the PhenoCam spring phenology model again predicts.

Our trend analysis showed that the PhenoCam spring phenology model performed better than the standard CLM phenology model when evaluated against decadal trends in MODIS data, and multi-decadal trends in the PEP725 data. For example, using MODIS data upscaled to the CLM model grid as the reference, there is large spatial variability in the observed trends over the period 2003–2013. Overall, 60% of grid cells (68% on an area-weighted basis) show a trend toward earlier spring; the median rate of advancement is -0.13 day yr^{-1} (-0.24 day yr^{-1} on an area-weighted basis), but there is enormous variability among grid cells (interquartile range = 0.65). The PhenoCam spring phenology model does a better job in capturing the global variation in these trends than the standard CLM spring phenology model (Fig. S1; Pearson correlation between PhenoCam and MODIS, $r = 0.50$; Pearson correlation between CLM and MODIS, $r = 0.07$). Additionally, for 1563 of 2274 grid cells, the PhenoCam spring phenology model predicts the correct sign on the observed MODIS trend, whereas this was the case for 1467 of 2274 grid cells for the CLM spring phenology model.

Similarly, in the PEP725 data, there are significant ($P \leq 0.05$) phenological trends for 83% of the 308 model grid cells. Of these, 1 trend is positive (toward later leaf unfolding) and 254 are negative (toward earlier leaf unfolding). The standard CLM spring phenology model predicts significant trends that are of the same sign as the trend in the data for only 13% (39 of 308) of the model grid cells. By comparison, the PhenoCam spring phenology model predicts significant trends that are of the same sign as the trend in the data for 74% (224 of 308) of the model grid cells. And, the PhenoCam spring phenology model incorrectly predicts significant trends that are different in sign from the trends in the data for only 1% (4 of 308) of the model grid cells.

Aggregating the data and model predictions to a regional average, the PEP725 data indicate a trend toward earlier leaf unfolding of -0.18 ± 0.04 day yr^{-1}

(slope ± 1 SE) over the period 1950–2014. However, break-point analysis shows that this trend is not consistent over time (Fig. S2). Rather, there is a slight but not significant trend toward later leaf unfolding (0.03 ± 0.11 day yr^{-1}) in the PEP725 data from 1950 to 1982, and then a much stronger and more significant trend toward earlier spring (0.39 ± 0.08 day yr^{-1}) from 1983 to 2014. By comparison, the CLM spring phenology model shows a small but nonsignificant trend toward earlier SOS (-0.04 ± 0.23 day yr^{-1}) from 1950 to 1982 and then a stronger but still nonsignificant trend toward earlier SOS (-0.19 ± 0.26 day yr^{-1}) from 1983 to 2014. The main reason that the latter trend is insignificant is that the CLM spring phenology model predicts about twice as much interannual variability in SOS as is actually observed to occur, with modeled SOS varying by over 7 weeks from year-to-year. The PhenoCam spring phenology model is more consistent with the PEP725 data; it correctly predicts a slight but nonsignificant trend toward later spring (0.10 ± 0.08 day yr^{-1}) over the period 1950–1982, and a stronger and significant trend toward earlier spring (-0.21 ± 0.06 day yr^{-1}) from 1983 to 2014. While we acknowledge that over the period from 1983 to 2014 the PhenoCam spring phenology model trend is a little more than half that of in the PEP725 data, we note that the confidence intervals on these slopes overlap substantially, i.e., 0.55 to -0.23 day yr^{-1} for PEP725, -0.33 to -0.09 day yr^{-1} for PhenoCam spring phenology model.

Together, these results suggest that, across the Northern Hemisphere deciduous broadleaf forest, the revised phenology submodel incorporated into CLM-PhenoCam represent a substantial improvement over the standard seasonal-deciduous spring phenology submodel in CLM. By reducing errors and biases associated with SOS prediction under current climate regimes, and by doing a better job at reproducing both recent (decadal) and longer-term (multidecadal) phenological trends, the PhenoCam spring phenology submodel should, therefore, give us greater confidence in model predictions under future climate scenarios.

Forward runs

Comparing the end of the 21st century to the beginning of the 21st century, CLM predicts earlier SOS in colder, higher-latitude regions such as boreal North America, northeastern China, Siberia and Northern Europe, but later spring onset dates in warmer, lower-latitude regions (Fig. 2b) including the southeastern United States, Mediterranean Europe, southeastern China, and northern India (Fig. 4a, d). These patterns are most

readily apparent in the RCP 8.5 model runs. In contrast, CLM-PhenoCam generally predicts earlier SOS by the end of the 21st century across the entire Northern Hemisphere deciduous broadleaf forest (Figs 2c, 4b, e). At the end of the 21st century, differences between predictions from the two phenology submodels are greatest in low latitudes (Fig. 2b, c), including the southeastern United States, southeastern China and northern India (Fig. 4c, f).

For the current decade (2014–2023), the global mean (area-weighted) SOS predicted by CLM, roughly day 118 under both climate scenarios, is very similar to that predicted by CLM-PhenoCam (Fig. 5). However, model predictions clearly begin to diverge around 2050. Under RCP 4.5, CLM-PhenoCam predicts SOS just 4 days earlier than CLM by 2100, but under RCP 8.5 the difference is 11 days by 2100 (Fig. 5). Thus, the two models predict very different trends in SOS over the 85 years of our forward runs. CLM predicts little change in SOS under either RCP 4.5 (slightly earlier by $0.004 \text{ day yr}^{-1}$, $r = -0.04$, trend not significant) or RCP 8.5 (slightly later by 0.02 day yr^{-1} , $r = -0.21$, trend not significant).

In contrast, CLM-PhenoCam predicts a statistically significant trend toward earlier SOS under both RCP 4.5 (earlier by 0.05 day yr^{-1} , $r = -0.60$, $P < 0.001$) and RCP 8.5 (0.10 day yr^{-1} , $r = -0.87$, $P < 0.001$).

Differencing model predictions of CLM and CLM-PhenoCam shows that the earlier SOS predicted by CLM-PhenoCam results in increased carbon assimilation (GPP) and forest productivity (NPP), but also increases in autotrophic (plant) respiration (AR) across the Northern Hemisphere deciduous broadleaf forest biome (Fig. 6). With future climate warming and the associated advancing SOS predicted by CLM-PhenoCam, the phenologically driven enhancement of GPP reaches 0.28 ± 0.07 and $0.60 \pm 0.12 \text{ Pg C yr}^{-1}$ by the end of the 21st century for RCP 4.5 and 8.5, respectively. This is partially offset by enhanced AR, which reaches $0.14 \pm 0.03 \text{ Pg C yr}^{-1}$ for RCP 4.5 and $0.28 \pm 0.04 \text{ Pg C yr}^{-1}$ for RCP 8.5, by 2100. Thus by 2100, CLM-PhenoCam predicts about $0.14 \pm 0.04 \text{ Pg C yr}^{-1}$ more NPP under RCP 4.5, and $0.32 \pm 0.08 \text{ Pg C yr}^{-1}$ more NPP under RCP 8.5, compared to CLM run with the standard seasonal-deciduous submodel.

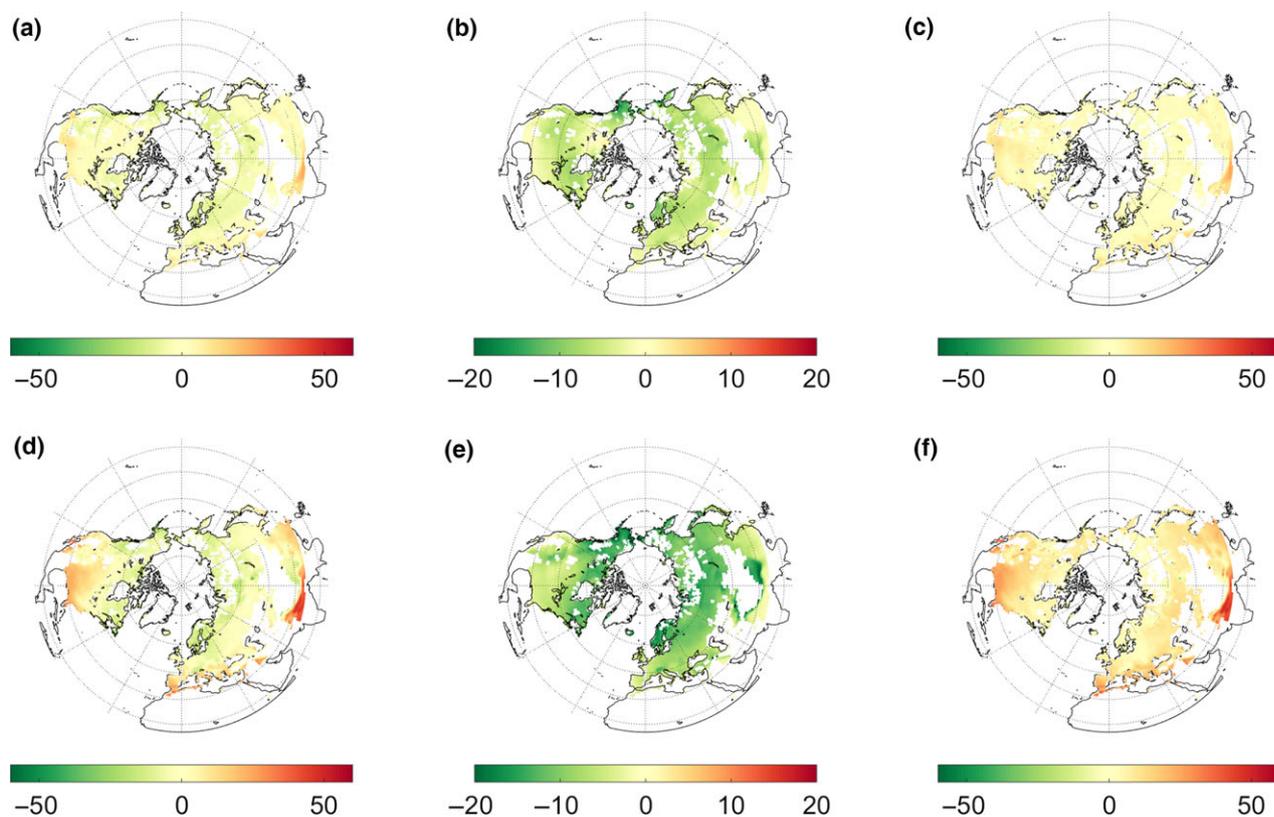


Fig. 4 Differences of start of spring (SOS) dates predicted by CLM and CLM-PhenoCam of forward model runs under the Representative Concentration Pathway (RCP) 4.5 (the first row) and 8.5 (the second row) scenarios. (a) Changes of SOS dates predicted by CLM between 2014–2023 and 2091–2100 under RCP 4.5 (calculated by using 2091–2100 results minus 2014–2023 results); (b) Same as (a) but from CLM-PhenoCam results; (c) Differences between (b) and (a); (d–f): Same as (a–c), respectively, but under RCP 8.5 scenario.

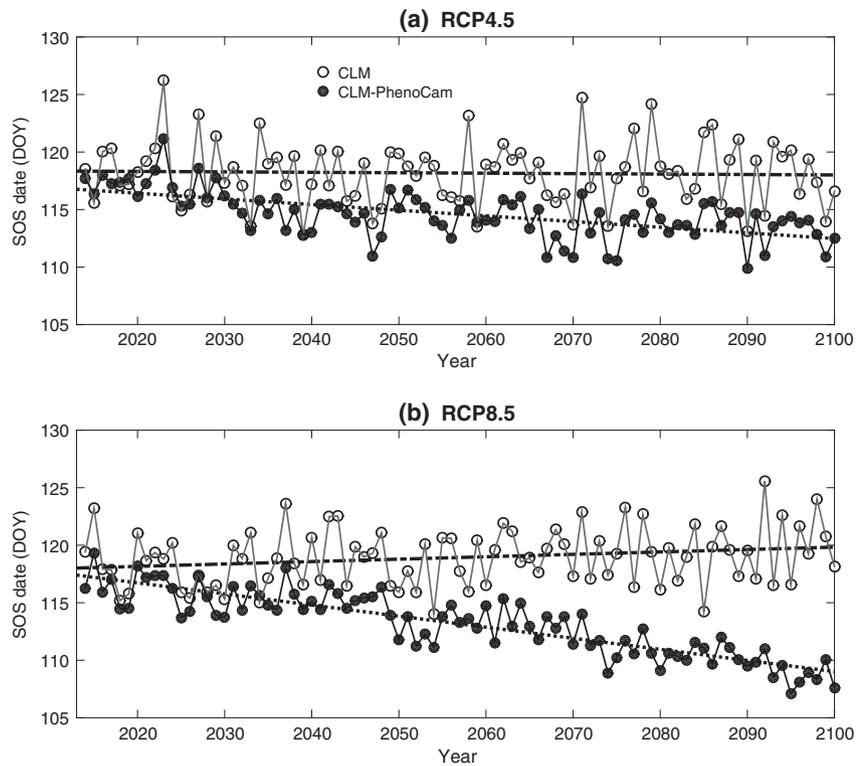


Fig. 5 Global area-weighted mean start of spring (SOS) date predicted from 2014 to 2100 under (a) Representative Concentration Pathway (RCP) 4.5 and (b) RCP 8.5 scenarios, for CLM's standard seasonaldeciduous phenology submodel and a revised phenology submodel calibrated to PhenoCam data.

Phenology also affects model predictions for ecosystem water fluxes. The earlier SOS predicted by CLM-PhenoCam results in more ET compared to the CLM run with the standard seasonal-deciduous submodel. At the end of the 21st century, both vegetation evaporation (QVEGE) and transpiration (QVEGT) are predicted by CLM-PhenoCam to be higher (by 3 ± 1 and 18 ± 3 mm yr⁻¹, respectively, under RCP 4.5; and by 7 ± 1 and 38 ± 5 mm yr⁻¹, respectively, under RCP 8.5) compared to CLM. At the same time, soil evaporation (QSOIL) is predicted by CLM-PhenoCam to be lower (by 9 ± 3 and 22 ± 4 mm yr⁻¹ under RCP 4.5 and 8.5, respectively) compared to CLM, because the longer duration canopy results in lower radiant energy fluxes incident on the soil surface. Together, the net effect of these changes in transpiration and evaporation leads to an overall increase (12 ± 2 and 24 ± 3 mm yr⁻¹ under RCP 4.5 and 8.5, respectively) in ecosystem ET by the end of the 21st century for CLM-PhenoCam compared to CLM, resulting in drier soils and reduced runoff in CLM-PhenoCam model runs.

Overall, compared to the standard CLM predictions for Northern Hemisphere deciduous broadleaf forests under RCP 8.5, CLM-PhenoCam predicts $9 \pm 2\%$ more

GPP, $8 \pm 2\%$ more NPP, $10 \pm 2\%$ more AR, $8 \pm 1\%$ less soil evaporation, $6 \pm 1\%$ more vegetation evaporation, and $12 \pm 2\%$ more transpiration, over the years 2090–2100.

Further, our analysis reveals evidence for both direct and indirect effects of earlier spring onset on processes related to ecosystem carbon and water cycling. We quantified direct effects by aggregating model differences strictly between the SOS day of year predicted by CLM-PhenoCam and the SOS day of year predicted by CLM. We quantified 'direct + indirect' effects by aggregating model differences over the entire year and indirect effects by the difference between direct and direct + indirect effects. The magnitudes of direct and the indirect effects of earlier spring onset were linearly correlated with the difference in predicted onset date, and these relationships were essentially invariant over the course of the 85 years of our forward runs. Thus, these 'phenological sensitivities' can be used for back-of-the-envelope estimates of how global deciduous forest carbon and water cycling would be altered under different assumptions (i.e. larger or smaller advance in phenology) about future phenological change. As shown in Table 1 and Fig. 7, a 1-day advancement of spring in CLM-PhenoCam (relative to CLM) was

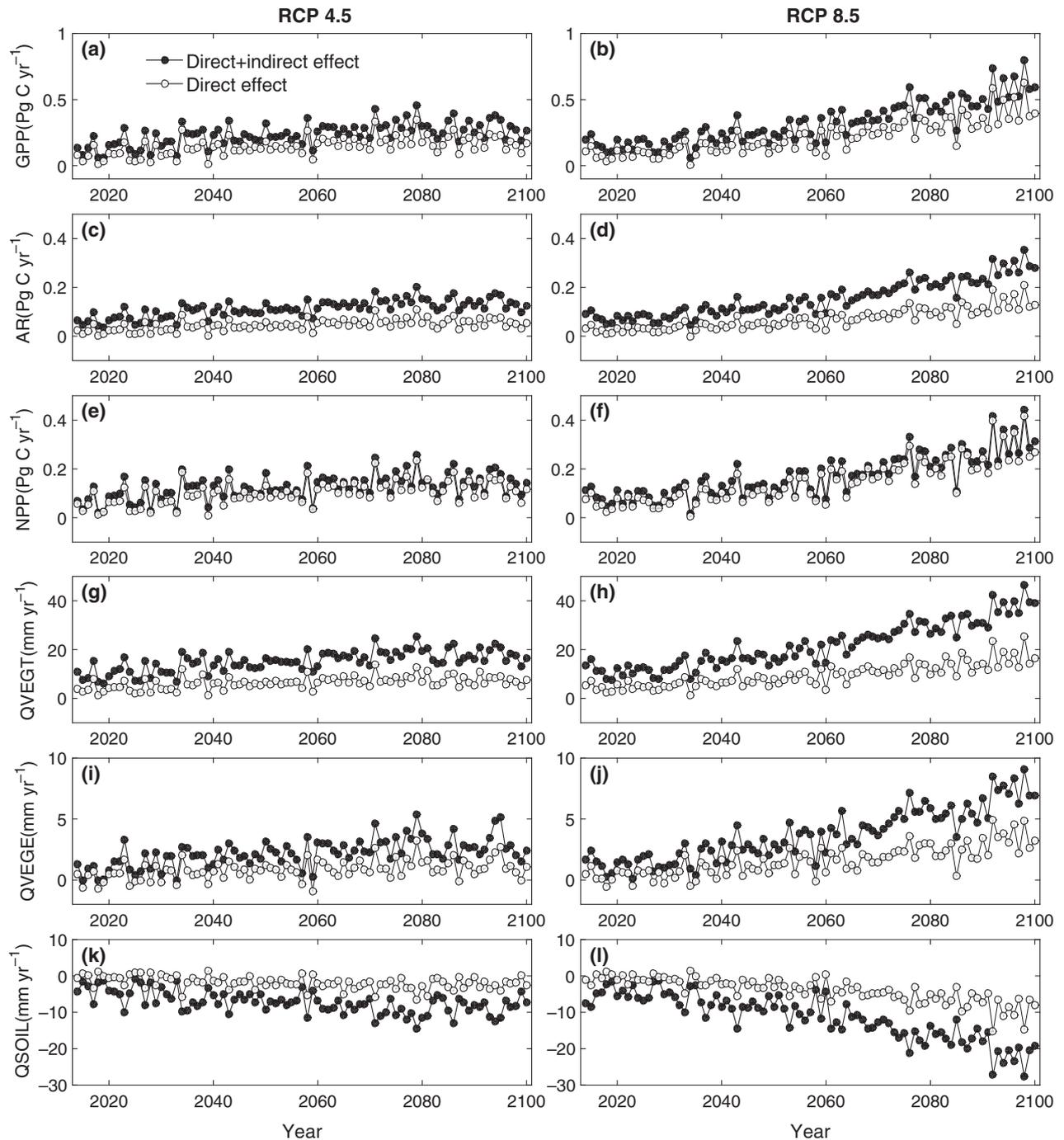


Fig. 6 Impacts of start of spring (SOS) submodel on globally-integrated carbon (GPP, gross primary production; AR, autotrophic respiration; NPP, net primary production) and water (QVEGT, vegetation transpiration; QVEGE, vegetation evaporation; QSOIL, soil evaporation) fluxes for 2014–2100 under Representative Concentration Pathway (RCP) 4.5 (left column) and RCP 8.5 (right column) scenarios. The direct effects are calculated from differences between CLM and CLM-PhenoCam model runs in springtime-integrated fluxes, while the direct + indirect effects are calculated from differences between CLM and CLM-PhenoCam model runs in annually-integrated fluxes.

overall associated with a direct $0.03 \text{ Pg C yr}^{-1} \text{ day}^{-1}$ increase in GPP, a $0.02 \text{ Pg C yr}^{-1} \text{ day}^{-1}$ increase in NPP, and a $0.01 \text{ Pg C yr}^{-1} \text{ day}^{-1}$ increase in AR.

Indirect effects of earlier spring were slightly smaller for GPP and larger for AR (both were about 0.01, but because these tended to cancel out, there was relatively

little indirect effect ($<0.01 \text{ Pg C yr}^{-1} \text{ day}^{-1}$) of earlier spring on NPP). Overall, a 1-day advancement of spring in CLM-PhenoCam was associated with a direct $1.25 \text{ mm yr}^{-1} \text{ day}^{-1}$ increase in vegetation transpiration, a $0.32 \text{ mm yr}^{-1} \text{ day}^{-1}$ increase in vegetation evaporation, and a 0.85 decrease in soil evaporation. Indirect effects of earlier spring were similar: a 1-day advancement of spring in CLM-PhenoCam was associated with an indirect $1.14 \text{ yr}^{-1} \text{ day}^{-1}$ increase in vegetation transpiration, a $0.23 \text{ mm yr}^{-1} \text{ day}^{-1}$ increase in vegetation evaporation, and a $0.75 \text{ mm yr}^{-1} \text{ day}^{-1}$ decrease in soil evaporation. Together, a 1-day advancement of spring in CLM-PhenoCam would result in a $1.34 \pm 0.05 \text{ mm yr}^{-1} \text{ day}^{-1}$ ($r = 0.90, P < 0.001$) increase in the ecosystem ET.

Discussion

Stimulated by concerns related to climate-change impacts on terrestrial ecosystems, there has been substantial effort devoted to improving the accuracy of many widely used land surface models (Williams *et al.*, 2009), including CLM (Bonan *et al.*, 2011; Levis *et al.*, 2012). However, it has also been demonstrated that phenology is one area where the performance of existing land surface models is particularly poor (Keenan *et al.*, 2012), and it has been argued that there is no reason to expect that performance will be improved under future climate scenarios (Richardson *et al.*, 2012). While there is broad consensus that future warming is likely to speed up plant developmental processes and advance spring phenology in temperate forests (Saxe *et al.*, 2001), accurately forecasting the impacts of these changes on carbon and water fluxes requires better phenology submodels to be integrated into large-scale land surface models. Our results – obtained using a

phenology model tuned to an extensive dataset derived from near-surface remote sensing and validated globally using MODIS observations – show that future shifts in phenology are likely to be smallest in warmer, lower-latitude temperate forests and largest in colder, higher-latitude boreal forests. By coupling vegetation phenology to carbon and water cycling processes, our analysis also shows that use of the standard spring phenology submodel in CLM is likely to substantially under-predict C uptake and ET across deciduous broadleaf forests in the Northern Hemisphere. This confirms that accurate prediction of spring phenological transitions is essential to reduce uncertainties in quantifying land–atmosphere exchanges of carbon and water under future climate scenarios. We note that our model runs were conducted ‘offline’ in that the forcing was prescribed and the biosphere does not feedback to the climate system. We expect that with fully coupled runs in the CESM, the increased carbon uptake and ET predicted by CLM-PhenoCam could have a substantial influence on the evolution of the climate system over the next 85 years.

We showed that when evaluated against MODIS-derived SOS for the entire Northern Hemisphere boreal and temperate deciduous broadleaf forest, the PhenoCam spring phenology submodel made considerably more accurate SOS predictions than the standard spring deciduous phenology submodel used in CLM. Against several different types of validation data, at different levels of spatial aggregation, the PhenoCam submodel had both lower RMSE and lower bias than the standard submodel (Fig. 3). This suggests that, for the most part, the PhenoCam sites (spanning 15° latitude, from 35 to 50°N , and almost two months in spring onset date, from day of year 80–140) we used for model development effectively capture the dominant patterns of

Table 1 Sensitivities of carbon and water fluxes in the Northern Hemisphere broadleaf deciduous forest to a 1-day advancement of start of spring (SOS), calculated by the difference in globally integrated carbon and water fluxes between CLM-PhenoCam and CLM using the standard winter deciduous spring phenology submodel. The sensitivities were calculated based on the linear slopes of differences in fluxes against differences in SOS date, under both RCP 4.5 and 8.5 scenarios

	Direct effect			Indirect effect		
	Mean	Standard error	<i>r</i>	Mean	Standard error	<i>r</i>
GPP (gross primary production)	0.03	<0.01	0.97	0.009	<0.001	0.82
AR (autotrophic respiration)	0.01	<0.01	0.97	0.007	<0.001	0.84
NPP (net primary production)	0.02	<0.01	0.97	0.001	<0.001	0.35
QVEGT (vegetation transpiration)	1.25	0.03	0.95	1.145	0.055	0.84
QVEGE (vegetation evaporation)	0.32	0.01	0.92	0.232	0.013	0.81
QSOIL (soil evaporation)	−0.85	0.03	−0.92	−0.754	0.035	−0.85

The statistical significance of these relationships is all $P < 0.001$. Units are $\text{Pg C yr}^{-1} \text{ day}^{-1}$ for GPP, AR, and NPP and are $\text{mm yr}^{-1} \text{ day}^{-1}$ for QVEGT, QVEGE, and QSOIL.

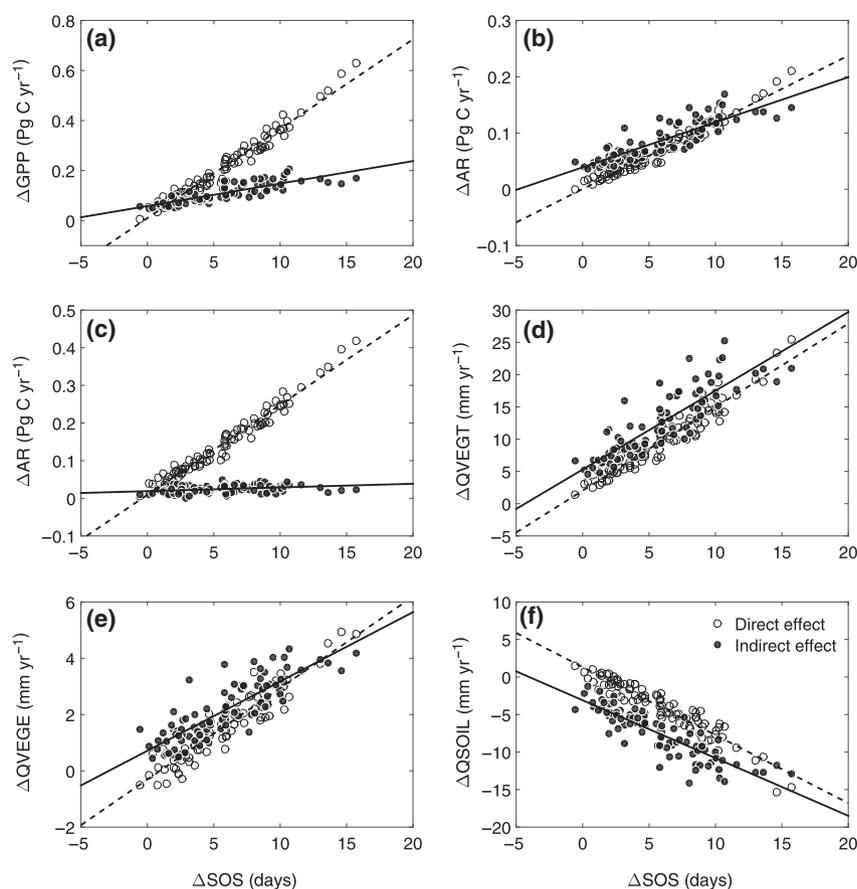


Fig. 7 Scatterplots of the direct and indirect differences of annual carbon and water fluxes from CLM and CLM-PhenoCam vs. the SOS differences from the forward model runs under both Representative Concentration Pathway (RCP) 8.5 scenario. The relationships are essentially identical for RCP 4.5.

spatial variation of spring phenology in Northern Hemisphere deciduous broadleaf forests. However, careful examination of the results clearly reveals that model performance was poorest in warm regions such as the southeastern United States. In these regions, PhenoCam model predictions showed significant bias toward later SOS compared to MODIS-derived SOS. One possible source of this bias is that there are large number of missing MODIS observations in these regions, which may lead to erroneous detection of seasonality metrics of OGI and MAT and therefore the MODIS SOS dates. Meanwhile, none of the PhenoCam sites that were used to calibrate the model were located in warm, low-latitude (22–35 °N) temperate forests, probably leading the model to be over-fit to cooler, northern sites. As more sites are added to the PhenoCam network in coming years (or as complementary data become available from other networks around the globe), it will be important to re-estimate this model using a more geographically representative sample. The increasing availability of other long-term, spatially

extensive phenological datasets, for example, from the USA-National Phenology Network (Jeong & Medvigy, 2014; Melaas *et al.*, 2016), should prove invaluable for the development and testing of new phenology models with better accuracy and improved generalizability (Richardson *et al.*, 2013).

A few previous studies have reported adaptation (implying genetic change from natural selection) or acclimation (implying reversible physiological adjustment) of plant spring phenology to warming temperatures (e.g. van Asch *et al.*, 2007; Bennie *et al.*, 2010; Keller *et al.*, 2011), suggesting an evolutionary tradeoff between advancing spring onset and avoiding catastrophic disturbance (e.g. spring frost and insect outbreaks). However, these responses may vary among species and also geographically. Thus, there is a lack of knowledge regarding the capacity for forest tree phenology to adjust to rising temperatures via either mechanism. We are not aware of any existing large-scale phenology model that explicitly considers adaptation to future climate change. The CLM model implicitly

assumes that acclimation will occur, as the GDD threshold to trigger bud burst depends on mean air temperature in the previous year (Eqn 2), but this is a short-term and reversible response. Notably, a surprising result from this model – which occurs precisely because of the acclimation effect – is the prediction that spring bud burst will actually be delayed for some warmer regions of the world by 2100, compared to present-day conditions (Fig. 4a, d). While there is some evidence that failure to meet chilling requirements may delay spring bud burst in a small selection of species (e.g. Heide, 1993; Orlandi *et al.*, 2004; Schwartz & Hanes, 2010), we are not aware of any observational studies which have yet demonstrated this kind of phenological delay, at a regional scale, in response to recent warming trends.

Similar to our results, many previous studies have found that land–atmosphere exchanges of carbon and water are sensitive to vegetation phenology (Richardson *et al.*, 2013). For example, Richardson *et al.* (2009) used ground observations of spring phenology, together with eddy covariance measurements of CO₂ exchange to estimate that earlier spring leaf-out increased annual GPP by about 10 g C m⁻² day⁻¹ in a temperate deciduous forest. By comparison, across Northern Hemisphere forests, Piao *et al.* (2007) used a model-based analysis to estimate that a 1-day extension in growing season length was associated with a 5.8 g C m⁻² day⁻¹ increase in GPP and a 2.8 g C m⁻² day⁻¹ increase in NPP. Yue *et al.* (2015) reported GPP in the north of 30°N has increased 0.01 Pg C yr⁻¹ day⁻¹ due to the phenological change during 1982–2011 based on global model simulations. Similarly, Zha *et al.* (2010) reported that warmer spring temperatures advanced spring leaf-out and enhanced both springtime and annual ET in Western Canadian ecosystems. By analyzing long-term flux data and ground observations of phenology during the past two decades, Keenan *et al.* (2014) suggested that a 1-day change in SOS would result in both more GPP (7.5 g C m⁻² day⁻¹) and more net C uptake (4.5 g C m⁻² day⁻¹). However, with observational studies, estimates of relationships between shifts in phenology and shifts in productivity may be confounded by processes that are simultaneously affected by the same factors driving variation in phenology, but which are not directly linked to phenology. A strength of our analysis is that it provides a defensible ‘model experiment’ framework, with both a ‘control’ (or null) model (in this case, the standard CLM spring phenology submodel) and a ‘treatment’ model (CLM-PhenoCam). By differencing the predictions of the two models, we can isolate effects of phenology ecosystem processes. This approach allows us to eliminate the

impact of climate change on other processes which are not phenologically mediated but that could be mistakenly attributed to phenology if only a single model was used (e.g. Piao *et al.*, 2007), because of their underlying covariation with temperature and hence SOS.

Changes in spring leaf-out dates have been hypothesized to have both direct (i.e. immediate) and indirect (i.e. lagged) effects on ecosystem processes (Richardson *et al.*, 2009). While the empirical evidence for the indirect effects of phenological anomalies is mixed (Richardson *et al.*, 2010), our model-based analysis allowed us to distinguish between the direct and the indirect effects. Our results show that the direct effects of changes in SOS account for the majority (about 66% for GPP, 44% for AR, and 86% for NPP) of the total (annual) differences in modeled carbon fluxes by the end of the 21st century (2090–2100). By comparison, for water fluxes, the direct effects of changes in SOS account for a smaller fraction (33% for soil evaporation, 43% for vegetation evaporation, and 43% for vegetation transpiration) of the differences, with the indirect effects being much larger. This may have to do with available soil water serving as both a constraint and a buffer on ET (Migliavacca *et al.*, 2012). However, empirical validation of these results using FLUXNET data is advised before these conclusions can be made with confidence. In summary, the results of this study show improved estimation of spring phenology for Northern Hemisphere boreal and deciduous broadleaf forests by incorporating a new and optimized submodel into the CLM version 4.5. The PhenoCam spring phenology submodel outperforms the standard CLM seasonal-deciduous spring phenology submodel in terms of both better accuracy and precision. Our analysis shows that with the standard seasonal-deciduous spring phenology submodel, errors in modeled SOS will propagate into modeled carbon and water fluxes (Richardson *et al.*, 2012) and that these errors are exacerbated under future climate change. Our results suggest that with the standard seasonal-deciduous spring submodel, CLM may underestimate GPP by 0.6 Pg C yr⁻¹ and NPP by 0.3 Pg C yr⁻¹ by the end of the 21st century. While the under-estimation of GPP is small relative to global terrestrial GPP (estimated at 123 ± 8 Pg C yr⁻¹ by Beer *et al.* (2010)), the under-estimation of NPP is considerable relative to the global terrestrial C sink, which for 2013 is estimated to be 2.5 ± 0.9 Pg C yr⁻¹ (Le Quéré *et al.*, 2015). Therefore, our results argue for a reconsideration of the standard seasonal-deciduous spring phenology submodel in CLM, as substantial errors in predictions of key land–atmosphere fluxes, as well as interactions and feedbacks between the biosphere and

the climate system, may otherwise result (Richardson *et al.*, 2013).

Acknowledgements

This research was supported by the National Science Foundation's Macrosystem Biology (EF-1065029) and LTER (DEB-1237491, DEB-1114804) programs, DOE Regional and Global Climate Modeling DE-SC0016011, and by NASA Grant number NNX14AJ35G. The authors acknowledge Dr. Danica Lombardozzi at NCAR and Dr. Ying Sun at UT Austin for their technical help on running CESM. The CESM project is supported by the National Science Foundation and the Office of Science (BER) of the U.S. Department of Energy. We would like to acknowledge high-performance computing support from Yellowstone (ark:/85065/d7wd3xhc) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. PEP725 data were provided by the members of the PEP725 project. We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (<http://ensembles-eu.metoffice.com>) and the data providers in the ECA&D project (<http://www.ecad.eu>).

References

- van Asch M, van Tiendren PH, Holleman LJM, Visser ME (2007) Predicting adaptation of phenology in response to climate change, an insect herbivore example. *Global Change Biology*, **13**, 1596–1604.
- Beer C, Reichstein M, Tomelleri E *et al.* (2010) Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science*, **329**, 834–838.
- Bennie J, Kubin E, Wiltshire A, Huntley B, Baxter R (2010) Predicting spatial and temporal patterns of bud-burst and spring frost risk in north-west Europe: the implications of local adaptation to climate. *Global Change Biology*, **16**, 1503–1514.
- Blanken PD, Black TA (2004) The canopy conductance of a boreal aspen forest, Prince Albert National Park, Canada. *Hydrological Processes*, **18**, 1561–1578.
- Bonan GB, Lawrence PJ, Oleson KW *et al.* (2011) Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. *Journal of Geophysical Research: Biogeosciences*, **116**, n/a–n/a.
- Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer Science & Business Media, New York.
- Cannell MGR, Smith RI (1983) Thermal time, chill days and prediction of budburst in *Picea sitchensis*. *Journal of Applied Ecology*, **20**, 951–963.
- Dahlin KM, Fisher RA, Lawrence PJ (2015) Environmental drivers of drought deciduous phenology in the Community Land Model. *Biogeosciences Discuss*, **12**, 5803–5839.
- Defries RS, Hansen MC, Townshend JRG, Janetos AC, Loveland TR (2000) A new global 1-km dataset of percentage tree cover derived from remote sensing. *Global Change Biology*, **6**, 247–254.
- Friedl MA, Sulla-Menashe D, Tan B, Schneider A, Ramankutty N, Sibley A, Huang X (2010) MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, **114**, 168–182.
- Ganguly S, Friedl MA, Tan B, Zhang X, Verma M (2010) Land surface phenology from MODIS: characterization of the Collection 5 global land cover dynamics product. *Remote Sensing of Environment*, **114**, 1805–1816.
- Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M (2008) A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research*, **113**, D20119.
- Heide OM (1993) Dormancy release in beech buds (*Fagus sylvatica*) requires both chilling and long days. *Physiologia Plantarum*, **89**, 187–191.
- Hollinger DY, Ollinger SV, Richardson AD *et al.* (2010) Albedo estimates for land surface models and support for a new paradigm based on foliage nitrogen concentration. *Global Change Biology*, **16**, 696–710.
- Houghton RA (2007) Balancing the global carbon budget. *Annual Review of Earth and Planetary Sciences*, **35**, 313–347.
- Hufkens K, Friedl M, Sonntag O, Braswell BH, Milliman T, Richardson AD (2012) Linking near-surface and satellite remote sensing measurements of deciduous broadleaf forest phenology. *Remote Sensing of Environment*, **117**, 307–321.
- Jeong S-J, Medvigy D (2014) Macroscale prediction of autumn leaf coloration throughout the continental United States. *Global Ecology and Biogeography*, **23**, 1245–1254.
- Keenan TF, Davidson E, Moffat AM, Munger W, Richardson AD (2012) Using model-data fusion to interpret past trends, and quantify uncertainties in future projections, of terrestrial ecosystem carbon cycling. *Global Change Biology*, **18**, 2555–2569.
- Keenan TF, Gray J, Friedl MA *et al.* (2014) Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nature Climate Change*, **4**, 598–604.
- Keller SR, Soolanayakanahally RY, Guy RD, Silim SN, Olson MS, Tiffin P (2011) Climate-driven local adaptation of ecophysiology and phenology in balsam poplar, *Populus balsamifera* L. (Salicaceae). *American Journal of Botany*, **98**, 99–108.
- Klosterman ST, Hufkens K, Gray JM *et al.* (2014) Evaluating remote sensing of deciduous forest phenology at multiple spatial scales using PhenoCam imagery. *Biogeosciences*, **11**, 4305–4320.
- Lawrence DM, Oleson KW, Flanner MG *et al.* (2011) Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. *Journal of Advances in Modeling Earth Systems*, **3**, M03001.
- Le Quéré C, Moriarty R, Andrew RM *et al.* (2015) Global carbon budget 2014. *Earth System Science Data*, **7**, 47–85.
- Levis S, Bonan GB, Kluzek E, Thornton PE, Jones A, Sacks WJ, Kucharik CJ (2012) Interactive crop management in the Community Earth System Model (CESM1): seasonal influences on land–atmosphere fluxes. *Journal of Climate*, **25**, 4839–4859.
- Luyssaert S, Inglima I, Jung M *et al.* (2007) CO₂ balance of boreal, temperate, and tropical forests derived from a global database. *Global Change Biology*, **13**, 2509–2537.
- Meehl GA, Washington WM, Arblaster JM *et al.* (2013) Climate change projections in CESM1(CAM5) compared to CCSM4. *Journal of Climate*, **26**, 6287–6308.
- Melaas EK, Friedl MA, Richardson AD (2016) Multiscale modeling of spring phenology across Deciduous Forests in the Eastern United States. *Global Change Biology*, **22**, 792–805.
- Migliavacca M, Sonntag O, Keenan TF, Cescatti A, O'Keefe J, Richardson AD (2012) On the uncertainty of phenological responses to climate change, and implications for a terrestrial biosphere model. *Biogeosciences*, **9**, 2063–2083.
- Oleson KW, Lawrence DM, Bonan GB *et al.* (2013) *Technical Description of Version 4.5 of the Community Land Model (CLM)*. NCAR Tech. Note NCAR/TN-503+STR, National Center for Atmospheric Research, Boulder, CO.
- Orlandi F, Garcia-Mozo H, Ezquerro LV, Romano B, Dominguez E, Galan C, Fornaciari M (2004) Phenological olive chilling requirements in Umbria (Italy) and Andalusia (Spain). *Plant Biosystems*, **138**, 111–116.
- Pan Y, Birdsey RA, Fang J *et al.* (2011) A large and persistent carbon sink in the world's forests. *Science*, **333**, 988–993.
- Peñuelas J, Rutishauser T, Filella I (2009) Phenology feedbacks on climate change. *Science*, **324**, 887–888.
- Piao S, Friedlingstein P, Ciais P, Viovy N, Demarty J (2007) Growing season extension and its impact on terrestrial carbon cycle in the Northern Hemisphere over the past 2 decades. *Global Biogeochemical Cycles*, **21**, n/a–n/a.
- Richardson A, O'Keefe J (2009) Phenological differences between understory and overstory. In: *Phenology of Ecosystem Processes SE – 4* (ed Noormets A), pp. 87–117. Springer, New York.
- Richardson AD, Hollinger DY, Dail DB, Lee JT, Munger JW, O'Keefe J (2009) Influence of spring phenology on seasonal and annual carbon balance in two contrasting New England forests. *Tree Physiology*, **29**, 321–331.
- Richardson AD, Andy Black T, Ciais P *et al.* (2010) Influence of spring and autumn phenological transitions on forest ecosystem productivity. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **365**, 3227–3246.
- Richardson AD, Anderson RS, Arain MA *et al.* (2012) Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis. *Global Change Biology*, **18**, 566–584.
- Richardson AD, Keenan TF, Migliavacca M, Ryu Y, Sonntag O, Toomey M (2013) Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agricultural and Forest Meteorology*, **169**, 156–173.
- Saxe H, Cannell MGR, Johnsen Ø, Ryan MG, Vourlitis G (2001) Tree and forest functioning in response to global warming. *New Phytologist*, **149**, 369–399.

- Schwartz MD, Hanes JM (2010) Continental-scale phenology: warming and chilling. *International Journal of Climatology*, **30**, 1595–1598.
- Sonnentag O, Hufkens K, Teshera-Sterne C *et al.* (2012) Digital repeat photography for phenological research in forest ecosystems. *Agricultural and Forest Meteorology*, **152**, 159–177.
- Thornton P, Law B, Gholz HL *et al.* (2002) Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests. *Agricultural and Forest Meteorology*, **113**, 185–222.
- Thornton PE, Thornton MM, Mayer BW, Wilhelmi N, Wei Y, Devarakonda R, Cook RB (2014) *Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 2*. Oak Ridge National Laboratory (ORNL), Oak Ridge, TN.
- Williams M, Richardson AD, Reichstein M *et al.* (2009) Improving land surface models with FLUXNET data. *Biogeosciences*, **6**, 1341–1359.
- Yue X, Unger N, Zheng Y (2015) Distinguishing the drivers of trends in land carbon fluxes and plant volatile emissions over the past 3 decades. *Atmospheric Chemistry and Physics*, **15**, 11931–11948.
- Zha T, Barr AG, van der Kamp G, Black TA, McCaughey JH, Flanagan LB (2010) Inter-annual variation of evapotranspiration from forest and grassland ecosystems in western Canada in relation to drought. *Agricultural and Forest Meteorology*, **150**, 1476–1484.
- Zhang X, Friedl MA, Schaaf CB *et al.* (2003) Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, **84**, 471–475.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Scatterplots of the trend in MODIS vs. the trend in the CLM (a) and CLM-PhenoCam (b) predicted SOS time series across all Northern Hemisphere boreal and temperate deciduous broadleaf forest grid cells.

Figure S2. Breakpoint trends of aggregated PEP725 data and model predicted SOS in Europe.