



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

## Citation

Chen, Min, Eli K. Melaas, Josh M. Gray, Mark A. Friedl, and Andrew D. Richardson. 2016. "A New Seasonal-Deciduous Spring Phenology Submodel in the Community Land Model 4.5: Impacts on Carbon and Water Cycling Under Future Climate Scenarios." *Global Change Biology* 22 (11) (May 14): 3675–3688. Portico. doi:10.1111/gcb.13326.

## Published Version

10.1111/gcb.13326

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:33373337>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

1 **Title: A new seasonal-deciduous spring phenology submodel in the**  
2 **Community Land Model 4.5: Impacts on carbon and water cycling under**  
3 **future climate scenarios**

4 **Running title: Phenology impacts land-atmosphere exchanges**

5 Min Chen<sup>1,2</sup>, Eli K. Melaas<sup>3</sup>, Josh M. Gray<sup>3</sup>, Mark A. Friedl<sup>3</sup>, Andrew D. Richardson<sup>1</sup>

6 **Affiliations**

7 1. Harvard University, Department of Organismic and Evolutionary Biology, 22 Divinity Ave.,  
8 Cambridge MA 02138 USA

9 2. Now at Carnegie Institution for Science, Department of Global Ecology. 260 Panama Street,  
10 Stanford, CA 94305, USA

11 3. Boston University, Department of Earth and Environment, 675 Commonwealth Ave., Boston,  
12 MA 02215, USA

13 **Corresponding author:** Min Chen (chenminbnu@gmail.com)

14 Phone: (617) 496-0825 Fax: 617-495-9484

15

16 **Key words:** Community Land Model; Ecosystem services; Phenology; PhenoCam; Carbon cycle;  
17 Water; Climate change

18

19 **Primary Research Article**

**(Submitted to *Global Change Biology*)**

**20 Abstract**

21 A spring phenology model that combines photoperiod with accumulated heating and  
22 chilling to predict spring leaf out dates is optimized using PhenoCam observations and coupled  
23 into the Community Land Model (CLM) 4.5. In head-to-head comparison (using satellite data  
24 from 2003-2013 for validation) for model grid cells over the Northern Hemisphere deciduous  
25 broadleaf forests (5.5 million km<sup>2</sup>), we found that the revised model substantially out-performed  
26 the standard CLM seasonal-deciduous spring phenology sub-model at both coarse (0.9×1.25  
27 degree) and fine (1km) scales. The revised model also does a better job of representing recent  
28 (decadal) phenological trends observed globally by MODIS, as well as long-term trends (1950-  
29 2014) in the PEP725 European phenology dataset. Moreover, forward model runs suggested a  
30 stronger advancement (up to 11 days) of spring leaf out by the end of the 21<sup>st</sup> century for the  
31 revised model. Trends towards earlier advancement are predicted for deciduous forests across the  
32 whole northern hemisphere boreal and temperate deciduous forest region for the revised model,  
33 whereas the standard model predicts earlier leaf out in colder regions, but later leaf out in  
34 warmer regions, and no trend globally. The earlier spring leaf out predicted by the revised model  
35 resulted in enhanced gross primary production (up to 0.6 Pg C yr<sup>-1</sup>) and evapotranspiration (up to  
36 24 mm yr<sup>-1</sup>) when results were integrated across the study region. These results suggest that the  
37 standard seasonal deciduous submodel in CLM should be reconsidered, otherwise substantial  
38 errors in predictions of key land-atmosphere interactions and feedbacks may result.

39

## 40 1. Introduction

41 The vast boreal and temperate deciduous forests of the Northern Hemisphere are thought  
42 to account for a substantial fraction of the terrestrial carbon sink (Houghton, 2007; Luysaert *et*  
43 *al.*, 2007; Pan *et al.*, 2011). In these ecosystems, vegetation phenology controls numerous land  
44 surface characteristics including albedo (Hollinger *et al.*, 2010), microclimate (Richardson &  
45 O'Keefe, 2009), canopy roughness and conductance (Blanken & Black, 2004), and the  
46 exchanges of carbon and water between land and atmosphere (Richardson *et al.*, 2013).  
47 Phenology thus plays an important role in mediating vegetation feedbacks to the climate system  
48 (Peñuelas *et al.*, 2009). For deciduous trees, phenological transitions are usually modeled as a  
49 function of air temperature and photoperiod, and sometimes soil temperature and soil moisture  
50 (Richardson *et al.*, 2012). However, existing land surface models generally employ poor  
51 phenological sub-models for deciduous forests, which leads to biased estimates of forest-  
52 atmosphere fluxes and feedbacks (Keenan *et al.*, 2012; Richardson *et al.*, 2012).

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

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

70         In this study we aim to improve CLM's seasonal-deciduous spring phenology submodel  
71 using a new formulation derived for boreal and temperate deciduous broadleaf forests in the  
72 Northern Hemisphere, and more importantly, we focus on evaluating how the new spring  
73 phenology algorithm may influence carbon and water cycles in these forests under future climate  
74 scenarios. We conduct simulations and compare the results from the standard version of CLM  
75 4.5 with results from a revised version that incorporates the improved phenology submodel. We  
76 investigate the following questions: (1) How do the two submodels compare in regard to  
77 predictions of spring leaf out across Northern Hemisphere deciduous broadleaf forests? What are  
78 the spatial patterns, and which submodel agrees best with remotely sensed land surface  
79 phenology? (2) Does either model predict significant trends towards earlier or later dates of  
80 spring leaf-out under future climate scenarios? (3) How are model estimates of land-atmosphere  
81 carbon and water exchanges in the coming century affected by choice of phenology submodel?

82

## 83 **2. Materials and Methods**

### 84 **2.1 CLM seasonal-deciduous spring phenology**

85 The seasonal-deciduous phenology algorithm in CLM 4.5 is directly adapted from an  
 86 earlier ecosystem model, Biome-BGC v. 4.1.2 (Thornton *et al.*, 2002). The timing of spring leaf  
 87 out is triggered when the accumulated Growing Degree Days (GDD) exceed a threshold  $GDD_{crit}$ .  
 88 The GDD temperature sum is calculated using the 3<sup>rd</sup>-layer soil temperature ( $T_s$ , in K), with a  
 89 base temperature equal to the water freezing temperature ( $T_f$ , 273.15 K). The accumulation of  
 90 GDD begins at the winter solstice. Thus for time step  $n$ , where  $\Delta t$  is the duration of the time step  
 91 (in seconds), and  $DL$  is the day length (86400 seconds):

$$92 \quad GDD_n = GDD_{n-1} + \max(T_s - T_f, 0) \times (\Delta t / DL) \quad (1)$$

93 Once  $GDD_n > GDD_{crit}$ , leaf out is triggered, effectively activating the growing season. Here, we  
 94 refer to this date as the start of spring (SOS). In the CLM 4.5 seasonal-deciduous phenology  
 95 algorithm,  $GDD_{crit}$  is calculated from the annual average of 2 m air temperature ( $T_{2mavg,ann}$ ) in the  
 96 preceding year:

$$97 \quad GDD_{crit} = \exp[4.8 + 0.13(T_{2mavg,ann} - T_f)] \quad (2)$$

98

## 99 **2.2 PhenoCam spring phenology model**

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

107 In a previous paper, Melaas *et al.* (2016) parameterized 13 different spring phenology  
 108 models using Klosterman *et al.*'s (2014) PhenoCam dataset, which includes observations from  
 109 13 deciduous forest sites located across a 12°C gradient in mean annual temperature and a more  
 110 than 800 mm year<sup>-1</sup> gradient in annual precipitation. The “best” model was selected using the  
 111 small-sample corrected Akaike Information Criterion (AIC<sub>c</sub>) (Burnham & Anderson, 2002)  
 112 based on the residual sum of squared errors for observations (55 site-years of data). Here, we use  
 113 the “best” model from Melaas *et al.* (2016), but re-parameterized it using an additional two years  
 114 of data (2010 and 2012) that were originally withheld for model testing. We refer to this model  
 115 as the “PhenoCam spring phenology model”. In total, 80 site-years of data were used to  
 116 parameterize our model.

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

$$124 \quad CU = CU + 1 \quad (3)$$

125 where GDD accumulate only if  $T_a > T_c$ :

$$126 \quad GDD = GDD + (T_a - T_c) \quad (4)$$

127 and  $GDD_{crit}$  is calculated as:

$$128 \quad GDD_{crit} = a + b \times \exp(c \times CU) \quad (5)$$

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

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

133

### 134 **2.3 Model evaluation and forward runs**

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

#### 141 *Model evaluation*

142 For model evaluation we conducted a number of different hindcast analyses. To evaluate  
143 the phenology submodel predictions against global grid-scale estimates of SOS, derived from  
144 satellite remote sensing, we ran CLM from 2000 to 2013, starting from initial conditions in 2000  
145 provided by the standard release of CESM 1.2.0, and using the transient Climate Research Unit –  
146 National Centers for Environmental Prediction (CRUNCEP) meteorological forcing data  
147 (<http://dods.extra.cea.fr/data/p529viov/cruncep/readme.htm>). We then evaluated predicted SOS  
148 dates from both phenology models against SOS dates determined from Moderate Resolution  
149 Imaging Spectroradiometer (MODIS, using data from 2003-2013) satellite imagery. Previous  
150 work (e.g. Hufkens *et al.*, 2012; Klosterman *et al.*, 2014), has shown that SOS dates derived  
151 from visual inspection of PhenoCam images are highly correlated with SOS dates derived from  
152 PhenoCam GCC thresholds (with little or no bias), which are in turn in good agreement with

153 start-of-spring dates derived from MODIS thresholds. We used OGI (onset of greenness increase,  
154 corresponding to 10% of the seasonal amplitude in the logistic-function-fitted EVI time series)  
155 and MAT (greenness maturity, corresponding to 90% of the seasonal amplitude in the logistic-  
156 function-fitted EVI time series) dates from the MODIS Collection 5 Land Cover Dynamics  
157 product (MCD12Q2) (500 m resolution, calculated from the Enhanced Vegetation Index (EVI))  
158 (Zhang *et al.*, 2003; Ganguly *et al.*, 2010). To account for a small bias in OGI relative to surface  
159 observations (Klosterman *et al.*, 2014), we calculated on a pixel-by-pixel basis, by linear  
160 interpolation, MODIS-derived SOS as the date at which 20% of the seasonal amplitude was  
161 achieved (Melaas *et al.*, 2016). We excluded pixels where the interannual variation (as measured  
162 by the standard deviation of the SOS) was more than 20 days, which is more than double the  
163 typical standard deviation observed in budburst dates at either Harvard Forest (Richardson &  
164 O’Keefe, 2009), or in the PEP725 dataset (described below). We further excluded pixels where  
165 the mean SOS date was later than day of year 200, assuming that dates after this day would tend  
166 to suggest either bad MODIS retrievals, or that the pixels are not actually temperate deciduous  
167 forest.

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

172         To address concerns about scale mismatch and uncertainties associated with aggregating  
173 MODIS data to coarse model grid cells, we also evaluated both phenological models at a much  
174 finer spatial resolution (1km). For these offline (*i.e.* not embedded within CLM) runs, conducted  
175 for the eastern United States, we used Daymet (Thornton *et al.*, 2014) meteorological forcing

176 data, and we calculated the mean SOS for pixels classified as deciduous forest across a 3×3  
177 MODIS pixel window centered on the Daymet grid cell. Following Melaas *et al.* (2016) we only  
178 evaluated the models for those grid cells where at least 3 of 9 MODIS pixels were classified as  
179 deciduous broadleaf or mixed forest.

180 Finally, for evaluation against long-term ground observations, we conducted an analysis  
181 using the Pan-European Phenology Project dataset, known as PEP725 (<http://www.pep725.eu>).  
182 Because of the length of the available time series, this dataset is well suited to trend analysis. We  
183 used records of leaf unfolding (code 11) for the common tree genera of northern Europe  
184 (including *Acer*, maple; *Betula*, birch; *Fagus*, beech; *Fraxinus*, ash; *Juglans*, walnut; *Populus*,  
185 aspen; *Quercus*, oak; and *Tilia*, basswood), including those sites where data were available for  
186 90% (or more) of the years between 1950 and 2014. Offline runs of both phenology submodels  
187 were driven by daily mean temperature from the E-OBS 0.25-degree gridded dataset (Haylock *et*  
188 *al.*, 2008), and we averaged all species within a model grid cell for further analysis.

189 The above analyses were conducted at various spatial scales therefore meteorological  
190 forcing data (*i.e.*, CRUNCEP, Daymet and E-OBS) with different spatial resolution were  
191 employed. In each of the above analyses, we calculated the Root Mean Square Error (RMSE),  
192 Mean Bias Error (MBE), and Pearson correlation coefficient (*r*) between data (*i.e.* MODIS or  
193 PEP725) and model predicted SOS at each model grid cell. For the grid-scale MODIS runs and  
194 the PEP725 runs we also calculated the slopes of the linear trends of the SOS time series for both  
195 the observational data and for each of the two phenology submodels.

#### 196 *Forward runs*

197 For the forward model runs (2014-2100), we used transient meteorological forcing and  
198 atmospheric CO<sub>2</sub> concentrations simulated by CCSM under Representative Concentration

199 Pathways (RCP) scenarios 8.5 (high emissions) and 4.5 (medium emissions) (Meehl *et al.*, 2013).  
200 Our objective with the forward runs was to investigate how future shifts in spring phenology  
201 might influence land-atmosphere exchanges of carbon and water under future climate regimes.  
202 We used the system state variables at the end of 2013 from our hindcast standard CLM run as the  
203 initial conditions for the forward runs.

204 CLM runs were conducted at a spatial resolution of  $0.9 \times 1.25$  degrees and a time step 30  
205 minutes with the “BGC” option turned on to incorporate the newest biogeochemistry  
206 developments in CLM 4.5 (Oleson *et al.*, 2013). Our analyses of model output focuses on SOS  
207 dates as well as carbon (gross primary production, GPP; net primary production, NPP; and  
208 autotrophic respiration, AR) and water (evapotranspiration, ET) fluxes integrated over Northern  
209 Hemisphere deciduous broadleaf forests (including both boreal and temperate broadleaf  
210 deciduous forests), the range of which is based on the prescribed plant functional type (PFT)  
211 distributions in CLM 4.5 default settings (Oleson *et al.*, 2013), as derived from the Advanced  
212 Very High Resolution Radiometer (AVHRR) continuous fields tree cover dataset (Defries *et al.*,  
213 2000). Grid cells that contain any fraction of deciduous forest are included into the CLM domain  
214 for further analysis. Our study domain is therefore larger than the spatial extent of deciduous  
215 broadleaf forests in MCD12Q1 (Fig. 1).

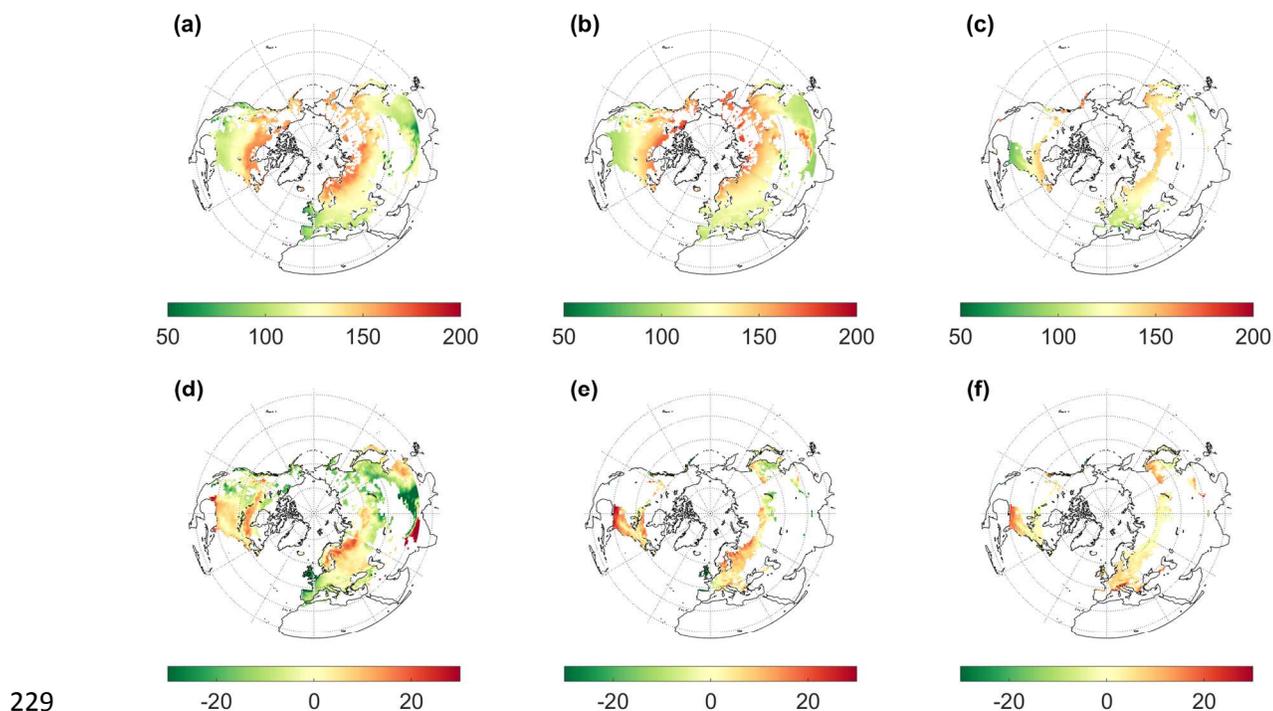
216

### 217 **3. Results**

#### 218 *Model evaluation*

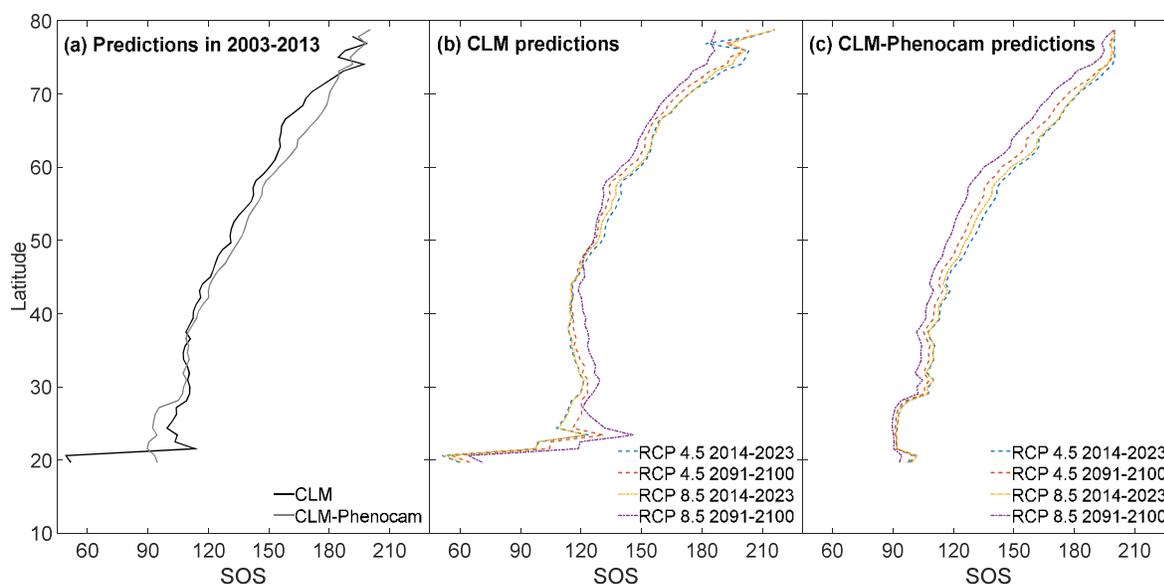
219 The spatial patterns of SOS predicted by the CLM and CLM-PhenoCam phenology  
220 submodels are broadly similar to MODIS-derived SOS dates (Fig. 1). Over the period from 2003  
221 to 2013, SOS occurred later in mid- and high-latitude regions, and earlier at lower latitudes. This

222 spatial pattern is largely driven by the strong temperature gradient from north to south. SOS from  
 223 CLM was generally earlier than SOS from CLM-PhenoCam ( $0.7 \pm 1.0$  days, mean  $\pm$  1 SD across  
 224 years), with the models in closest agreement in middle latitudes between  $30^\circ\text{N}$  and  $40^\circ\text{N}$  and  
 225 showing largest disagreement in lower ( $20^\circ\text{N}$  to  $25^\circ\text{N}$ ) and higher latitudes ( $65^\circ\text{N}$  to  $70^\circ\text{N}$ )  
 226 (Fig. 2a). More specifically, differences are small in mid-latitude regions of eastern North  
 227 America and Eastern Europe, but are particularly large in boreal North America and subtropical  
 228 regions of Southern Asia (Fig. 1d).



230 **Figure 1.** Average SOS (start of spring) dates predicted by CLM and CLM-PhenoCam, compared with  
 231 MODIS-derived SOS, over the period 2003-2013: (a) CLM predicted SOS; (b) CLM-PhenoCam  
 232 predicted SOS; (c) MODIS-derived SOS; (d) Differences between CLM and CLM-PhenoCam SOS [(a)-  
 233 (b)]; (e) Differences between CLM and MODIS SOS [(a)-(c)]; (f) Differences between CLM-PhenoCam  
 234 and MODIS SOS [(b)-(c)].

235

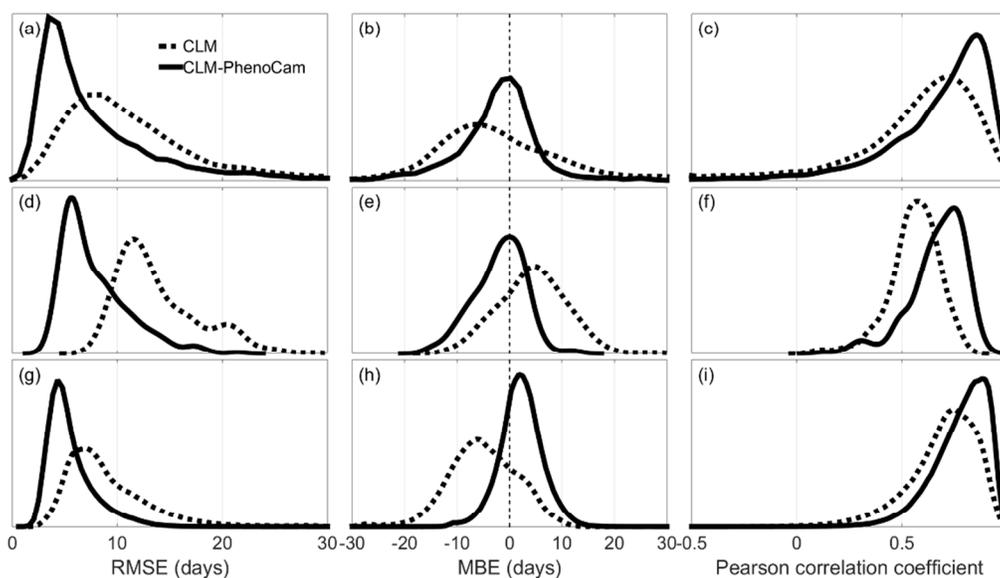


236  
 237 **Figure 2.** Latitudinal mean of the SOS dates predicted by CLM and CLM-PhenoCam: (a) hindcast  
 238 predictions, 2001-2013; (b) CLM predictions at the beginning (2014-2023) and end (2091-2100) of the  
 239 forward runs under RCP 4.5 and 8.5; and (c) CLM-PhenoCam predictions at the beginning (2014-2023)  
 240 and end (2091-2100) of the forward runs under RCP 4.5 and 8.5.

241

242 At the model grid scale ( $0.9 \times 1.25$  degree), SOS predictions from CLM are somewhat less  
 243 consistent with MODIS-derived SOS than are SOS predictions from CLM-PhenoCam. Across all  
 244 Northern Hemisphere deciduous broadleaf forests, the mode RMSE is 8 days (median, 10 days)  
 245 for CLM, compared with a mode of 4 days (median, 6 days) for CLM-PhenoCam (Fig. 3a). SOS  
 246 dates for CLM are also biased early (mode, -7 days; median, -3 days) compared to MODIS-  
 247 derived SOS. SOS bias for CLM-PhenoCam is much smaller (mode, -2; median, -1 days) (Fig.  
 248 3b). Finally, the correlation coefficient between predicted and MODIS-derived SOS dates is  
 249 generally weaker for CLM (mode,  $r = 0.7$ ) than CLM-PhenoCam (mode,  $r = 0.9$ ) (Fig. 3c).

250



251

252 **Figure 3.** Probability density estimates of the root mean square error (RMSE),  
 253 and Pearson correlation coefficient ( $r$ ), for start of spring (SOS) predicted by two models: CLM and  
 254 CLM-PhenoCam. (a)-(c): comparison against Moderate Resolution Imaging Spectroradiometer  
 255 (MODIS)-derived SOS across the Northern Hemisphere deciduous broadleaf forest ( $0.9 \times 1.25$  degree grid  
 256 cells); (d)-(f): comparison against PEP725 data ( $0.25$  degree grid cells); (g)-(i): comparison against  
 257 MODIS-derived SOS across the eastern US deciduous forest ( $1\text{km}$  grid cells).

258

259 We obtained similar results when the two phenology models were run at finer spatial  
 260 resolution and evaluated at that scale. For example, forced with Daymet data at  $1\text{km}$  resolution  
 261 for the eastern US, and evaluated against MODIS data aggregated to  $3 \times 3$  pixel windows, the  
 262 PhenoCam spring phenology model has a lower RMSE (mode, 4 days; median, 5 days) than the  
 263 standard CLM spring phenology model (mode, 7 days; median, 8 days). The PhenoCam model  
 264 also performs better in terms of smaller bias (mode MBE of 2 days vs. -5 days, for the two  
 265 models respectively) and higher correlation coefficient (mode  $r$  of 0.9 vs. 0.8, again for the two  
 266 models respectively) (Fig. 3g-i). And, when run with E-OBS data at  $0.25^\circ$  resolution for northern

267 Europe, and evaluated against PEP725 data, the PhenoCam spring phenology model again  
268 predicts substantially more accurate SOS, in terms of lower RMSE, smaller bias, and higher  
269 correlation coefficient (Fig. 3d-f) than the standard CLM spring phenology model.

270 Our trend analysis showed that the PhenoCam spring phenology model performed better  
271 than the standard CLM phenology model when evaluated against decadal trends in MODIS data,  
272 and multi-decadal trends in the PEP725 data. For example, using MODIS data upscaled to the  
273 CLM model grid as the reference, there is large spatial variability in the observed trends over the  
274 period 2003-2013. Overall, 60% of grid cells (68% on an area-weighted basis) show a trend  
275 towards earlier spring; the median rate of advancement is  $-0.13 \text{ d yr}^{-1}$  ( $-0.24 \text{ d yr}^{-1}$  on an area-  
276 weighted basis) but there is enormous variability among grid cells (interquartile range = 0.65).  
277 The PhenoCam spring phenology model does a better job in capturing the global variation in  
278 these trends than the standard CLM spring phenology model (Fig. S1; Pearson correlation  
279 between PhenoCam and MODIS,  $r = 0.50$ ; Pearson correlation between CLM and MODIS,  $r =$   
280  $0.07$ ). Additionally, for 1563 of 2274 grid cells, the PhenoCam spring phenology model predicts  
281 the correct sign on the observed MODIS trend, whereas this was the case for 1467 of 2274 grid  
282 cells for the CLM spring phenology model.

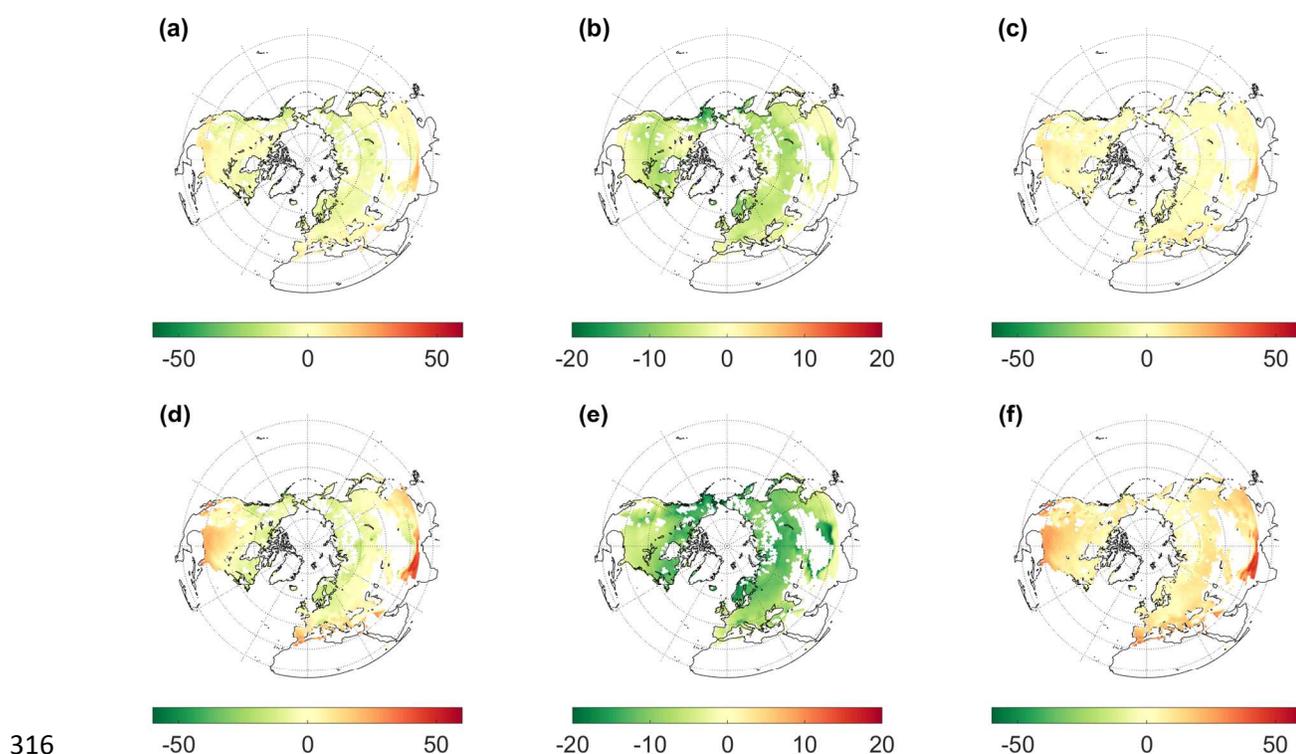
283 Similarly, in the PEP725 data, there are significant ( $p \leq 0.05$ ) phenological trends for 83%  
284 of the 308 model grid cells. Of these, 1 trend is positive (toward later leaf unfolding) and 254 are  
285 negative (toward earlier leaf unfolding). The standard CLM spring phenology model predicts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

289 cells. And, the PhenoCam spring phenology model incorrectly predicts significant trends that are  
290 different in sign from the trends in the data for only 1% (4 of 308) of the model grid cells.

291         Aggregating the data and model predictions to a regional average, the PEP725 data  
292 indicate a trend towards earlier leaf unfolding of  $-0.18 \pm 0.04 \text{ d yr}^{-1}$  (slope  $\pm 1 \text{ SE}$ ) over the  
293 period 1950 to 2014. However, break-point analysis shows that this trend is not consistent over  
294 time (Fig. S2). Rather, there is a slight but not significant trend towards later leaf unfolding ( $0.03$   
295  $\pm 0.11 \text{ d yr}^{-1}$ ) in the PEP725 data from 1950 to 1982, and then a much stronger and more  
296 significant trend towards earlier spring ( $0.39 \pm 0.08 \text{ d yr}^{-1}$ ) from 1983 to 2014. By comparison,  
297 the CLM spring phenology model shows a small but non-significant trend towards earlier SOS  
298 ( $-0.04 \pm 0.23 \text{ d yr}^{-1}$ ) from 1950 to 1982, and then a stronger but still non-significant trend  
299 towards earlier SOS ( $-0.19 \pm 0.26 \text{ d yr}^{-1}$ ) from 1983 to 2014. The main reason that the latter trend  
300 is insignificant is that the CLM spring phenology model predicts about twice as much  
301 interannual variability in SOS as is actually observed to occur, with modeled SOS varying by  
302 over 7 weeks from year-to-year. The PhenoCam spring phenology model is more consistent with  
303 the PEP725 data; it correctly predicts a slight but non-significant trend towards later spring ( $0.10$   
304  $\pm 0.08 \text{ d yr}^{-1}$ ) over the period 1950 to 1982, and a stronger and significant trend towards earlier  
305 spring ( $-0.21 \pm 0.06 \text{ d yr}^{-1}$ ) from 1983 to 2014. While we acknowledge that over the period from  
306 1983 to 2014 the PhenoCam spring phenology model trend is a little more than half that of in the  
307 PEP725 data, we note that the confidence intervals on these slopes overlap substantially, *i.e.*  $0.55$   
308 to  $-0.23 \text{ d yr}^{-1}$  for PEP725,  $-0.33$  to  $-0.09 \text{ d yr}^{-1}$  for PhenoCam spring phenology model.

309         Together, these results suggest that, across the Northern Hemisphere deciduous broadleaf  
310 forest, the revised phenology submodel incorporated into CLM-PhenoCam represent a  
311 substantial improvement over the standard seasonal-deciduous spring phenology submodel in

312 CLM. By reducing errors and biases associated with SOS prediction under current climate  
 313 regimes, and by doing a better job at reproducing both recent (decadal) and longer-term (multi-  
 314 decadal) phenological trends, the PhenoCam spring phenology submodel should, therefore, give  
 315 us greater confidence in model predictions under future climate scenarios.



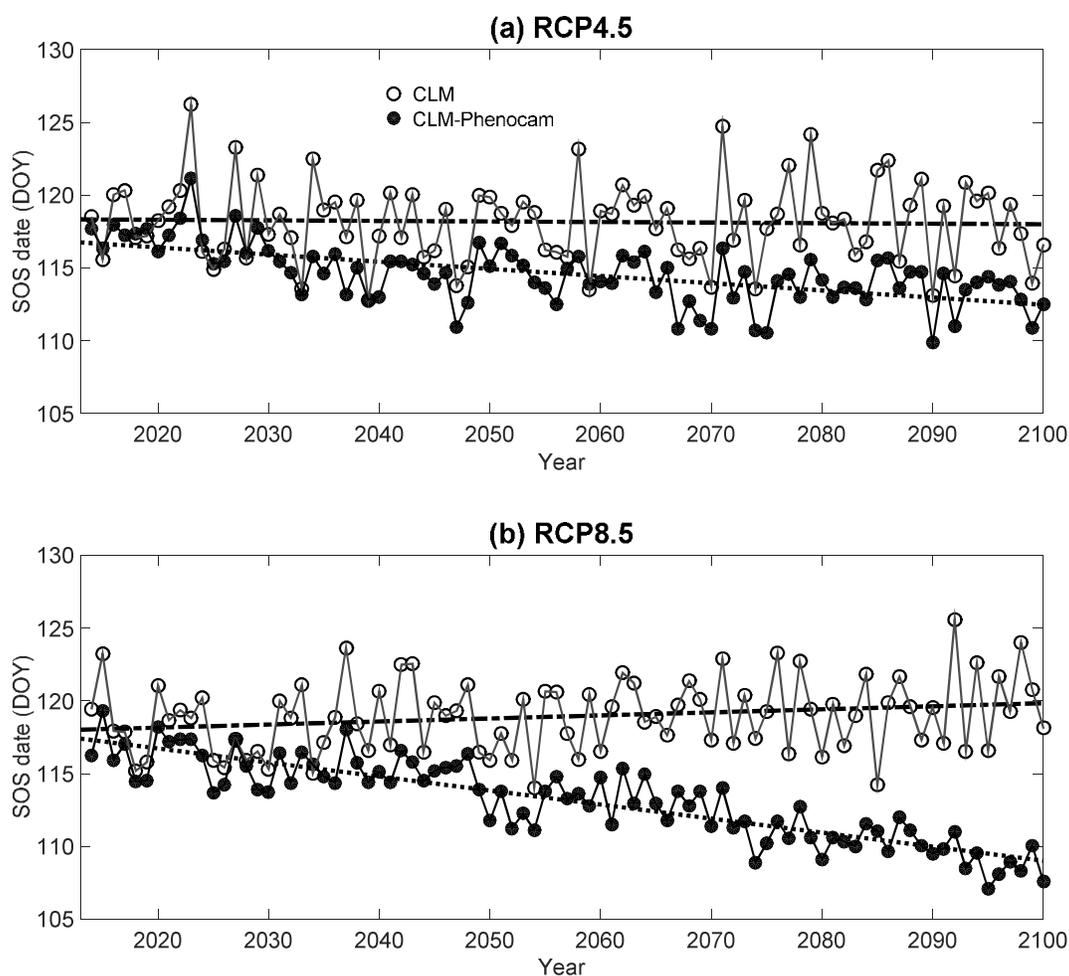
317 **Figure 4.** Differences of SOS dates predicted by CLM and CLM-PhenoCam of forward model runs under  
 318 the RCP 4.5 (the first row) and 8.5 (the second row) scenarios. (a) Changes of SOS dates predicted by  
 319 CLM between 2014-2023 and 2091-2100 under RCP 4.5 (calculated by using 2091-2100 results minus  
 320 2014-2023 results); (b) Same as (a) but from CLM-PhenoCam results; (c) Differences between (b) and (a)  
 321 [(a)-(b)]; (d)-(f): Same as (a)-(c), respectively, but under RCP 8.5 scenario.

322

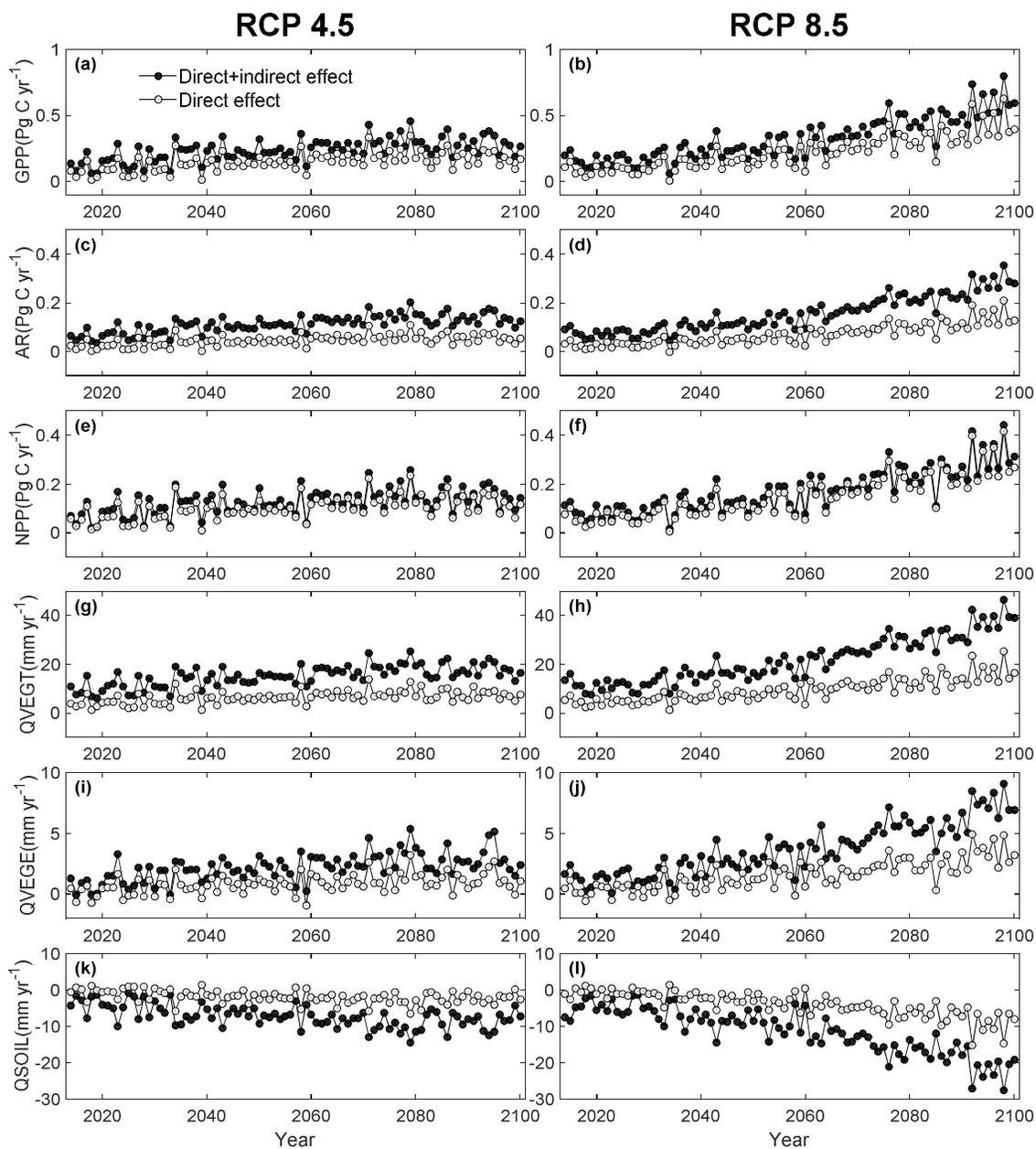
### 323 *Forward runs*

324 Comparing the end of the 21<sup>st</sup> century to the beginning of the 21<sup>st</sup> century, CLM predicts  
 325 earlier SOS in colder, higher-latitude regions such as boreal North America, northeastern China,  
 326 Siberia and Northern Europe, but later spring onset dates in warmer, lower-latitude regions (Fig.

327 2b) including the southeastern United States, Mediterranean Europe, southeastern China and  
 328 northern India (Fig. 4a,d). These patterns are most readily apparent in the RCP 8.5 model runs.  
 329 In contrast, CLM-PhenoCam generally predicts earlier SOS by the end of the 21<sup>st</sup> century across  
 330 the entire Northern Hemisphere deciduous broadleaf forest (Fig. 2c, 4b,e). At the end of the 21<sup>st</sup>  
 331 century, differences between predictions from the two phenology submodels are greatest in low  
 332 latitudes (Fig. 2b-c), including the southeastern United States, southeastern China and northern  
 333 India (Fig. 4c,f).



334  
 335 **Figure 5.** Global area-weighted mean start of spring (SOS) date predicted from 2014 to 2100 under (a)  
 336 Representative Concentration Pathway (RCP) 4.5 and (b) RCP 8.5 scenarios, for CLM's standard  
 337 seasonal-deciduous phenology submodel and a revised phenology submodel calibrated to PhenoCam data.

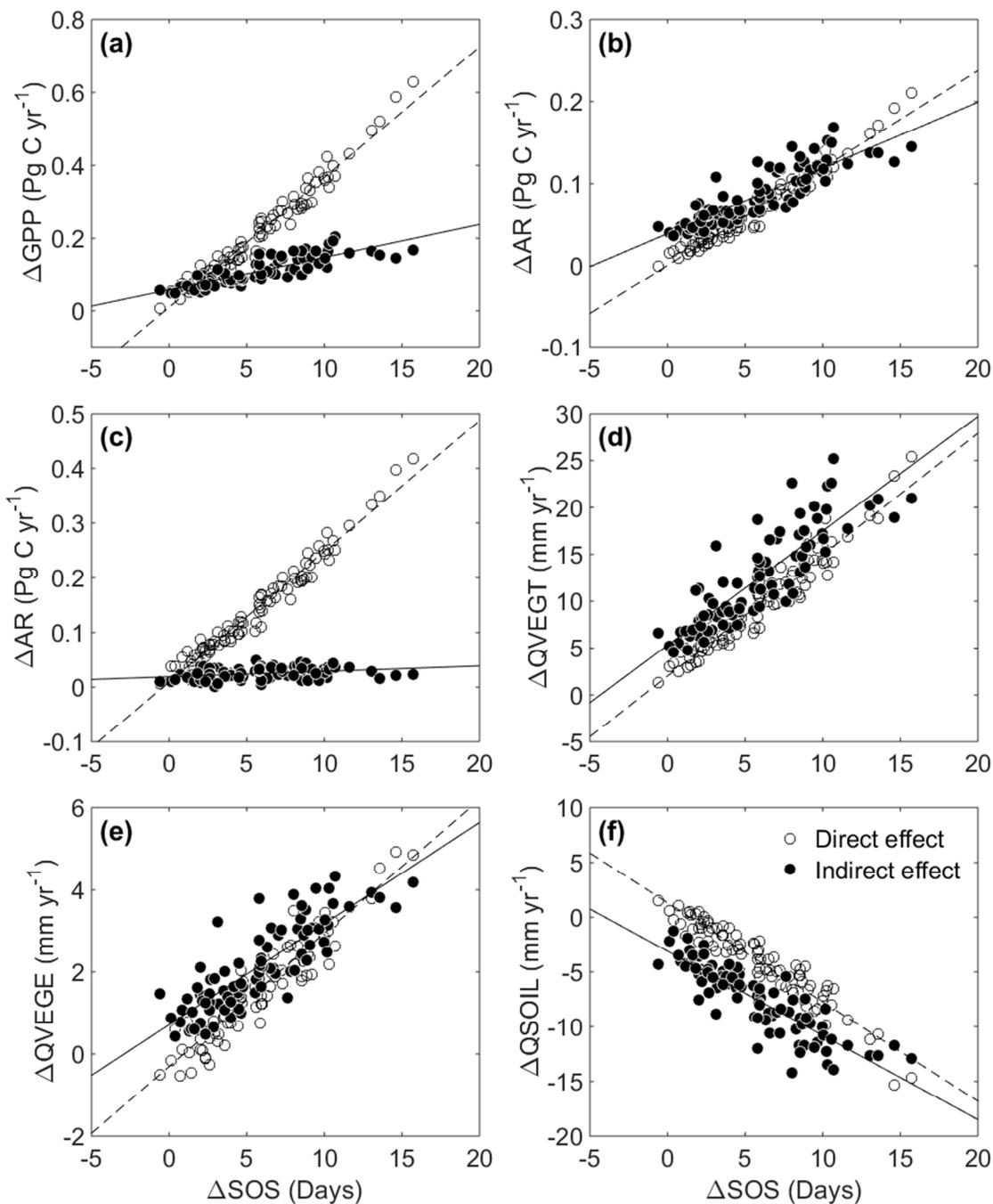


338

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

346 For the current decade (2014-2023), the global mean (area-weighted) SOS predicted by  
347 CLM, roughly day 118 under both climate scenarios, is very similar to that predicted by CLM-  
348 PhenoCam (Fig. 5). However, model predictions clearly begin to diverge around 2050. Under  
349 RCP 4.5, CLM-PhenoCam predicts SOS just 4 days earlier than CLM by 2100, but under RCP  
350 8.5 the difference is 11 days by 2100 (Fig. 5). Thus, the two models predict very different trends  
351 in SOS over the 85 years of our forward runs. CLM predicts little change in SOS under either  
352 RCP 4.5 (slightly earlier by  $0.004 \text{ d yr}^{-1}$ ,  $r=-0.04$ , trend not significant) or RCP 8.5 (slightly later  
353 by  $0.02 \text{ d yr}^{-1}$ ,  $r=-0.21$ , trend not significant). In contrast, CLM-PhenoCam predicts a statistically  
354 significant trend toward earlier SOS under both RCP 4.5 (earlier by  $0.05 \text{ d yr}^{-1}$ ,  $r = -0.60$ ,  $p <$   
355  $0.001$ ) and RCP 8.5 ( $0.10 \text{ d yr}^{-1}$ ,  $r = -0.87$ ,  $p < 0.001$ ).

356 Differencing model predictions of CLM and CLM-PhenoCam shows that the earlier SOS  
357 predicted by CLM-PhenoCam results in increased carbon assimilation (GPP) and forest  
358 productivity (NPP), but also increases in autotrophic (plant) respiration (AR) across the Northern  
359 Hemisphere deciduous broadleaf forest biome (Fig. 6). With future climate warming and the  
360 associated advancing SOS predicted by CLM-PhenoCam, the phenologically-driven  
361 enhancement of GPP reaches  $0.28 \pm 0.07$  and  $0.60 \pm 0.12 \text{ Pg C yr}^{-1}$  by the end of the 21<sup>st</sup> century  
362 for RCP 4.5 and 8.5, respectively. This is partially offset by enhanced AR, which reaches  $0.14 \pm$   
363  $0.03 \text{ Pg C yr}^{-1}$  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-  
364 PhenoCam predicts about  $0.14 \pm 0.04 \text{ Pg C yr}^{-1}$  more NPP under RCP 4.5, and  $0.32 \pm 0.08 \text{ Pg C}$   
365  $\text{yr}^{-1}$  more NPP under RCP 8.5, compared to CLM run with the standard seasonal-deciduous  
366 submodel.



367

368 **Figure 7.** Scatterplots of the direct and indirect differences of annual carbon and water fluxes from CLM  
 369 and CLM-PhenoCam vs. the SOS differences from the forward model runs under both RCP 8.5 scenario.  
 370 The relationships are essentially identical for RCP 4.5.

371

372 Phenology also affects model predictions for ecosystem water fluxes. The earlier SOS  
373 predicted by CLM-PhenoCam results in more evapotranspiration compared to the CLM run with  
374 the standard seasonal-deciduous submodel. At the end of the 21<sup>st</sup> century, both vegetation  
375 evaporation (QVEGE) and transpiration (QVEGT) are predicted by CLM-PhenoCam to be  
376 higher (by  $3 \pm 1 \text{ mm yr}^{-1}$  and  $18 \pm 3 \text{ mm yr}^{-1}$ , respectively, under RCP 4.5; and by  $7 \pm 1 \text{ mm yr}^{-1}$   
377 and  $38 \pm 5 \text{ mm yr}^{-1}$ , respectively, under RCP 8.5) compared to CLM. At the same time, soil  
378 evaporation (QSOIL) is predicted by CLM-PhenoCam to be lower (by  $9 \pm 3 \text{ mm yr}^{-1}$  and  $22 \pm 4$   
379  $\text{mm yr}^{-1}$  under RCP 4.5 and 8.5, respectively) compared to CLM, because the longer duration  
380 canopy results in lower radiant energy fluxes incident on the soil surface. Together, the net effect  
381 of these changes in transpiration and evaporation leads to an overall increase ( $12 \pm 2 \text{ mm yr}^{-1}$  and  
382  $24 \pm 3 \text{ mm yr}^{-1}$  under RCP 4.5 and 8.5, respectively) in ecosystem ET by the end of the 21<sup>st</sup>  
383 century for CLM-PhenoCam compared to CLM, resulting in drier soils and reduced runoff in  
384 CLM-PhenoCam model runs.

385 Overall, compared to the standard CLM predictions for Northern Hemisphere deciduous  
386 broadleaf forests under RCP 8.5, CLM-PhenoCam predicts  $9 \pm 2 \%$  more GPP,  $8 \pm 2\%$  more  
387 NPP,  $10 \pm 2\%$  more AR,  $8 \pm 1\%$  less soil evaporation,  $6 \pm 1\%$  more vegetation evaporation, and  
388  $12 \pm 2\%$  more transpiration, over the years 2090-2100.

389 Further, our analysis reveals evidence for both direct and indirect effects of earlier spring  
390 onset on processes related to ecosystem carbon and water cycling. We quantified direct effects  
391 by aggregating model differences strictly between the SOS day of year predicted by CLM-  
392 PhenoCam and the SOS day of year predicted by CLM. We quantified “direct+indirect” effects  
393 by aggregating model differences over the entire year, and indirect effects by the difference  
394 between direct and direct+indirect effects. The magnitudes of direct and the indirect effects of

395 earlier spring onset were linearly correlated with the difference in predicted onset date, and these  
396 relationships were essentially invariant over the course of the 85 years of our forward runs. Thus,  
397 these “phenological sensitivities” can be used for back-of-the-envelope estimates of how global  
398 deciduous forest carbon and water cycling would be altered under different assumptions (i.e.  
399 larger or smaller advance in phenology) about future phenological change. As shown in Table 1  
400 and Fig. 7, a one-day advancement of spring in CLM-PhenoCam (relative to CLM) was overall  
401 associated with a direct  $0.03 \text{ Pg C yr}^{-1} \text{ d}^{-1}$  increase in GPP, a  $0.02 \text{ Pg C yr}^{-1} \text{ d}^{-1}$  increase in NPP,  
402 and a  $0.01 \text{ Pg C yr}^{-1} \text{ d}^{-1}$  increase in AR. Indirect effects of earlier spring were slightly smaller for  
403 GPP and larger for AR (both were about 0.01, but because these tended to cancel out, there was  
404 relatively little indirect effect ( $< 0.01 \text{ Pg C yr}^{-1} \text{ d}^{-1}$ ) of earlier spring on NPP. Overall, a one-day  
405 advancement of spring in CLM-PhenoCam was associated with a direct  $1.25 \text{ mm yr}^{-1} \text{ d}^{-1}$  increase  
406 in vegetation transpiration, a  $0.32 \text{ mm yr}^{-1} \text{ d}^{-1}$  increase in vegetation evaporation, and a 0.85  
407 decrease in soil evaporation. Indirect effects of earlier spring were similar: a one-day  
408 advancement of spring in CLM-PhenoCam was associated with an indirect  $1.14 \text{ yr}^{-1} \text{ d}^{-1}$  increase  
409 in vegetation transpiration, a  $0.23 \text{ mm yr}^{-1} \text{ d}^{-1}$  increase in vegetation evaporation and a  $0.75 \text{ mm}$   
410  $\text{yr}^{-1} \text{ d}^{-1}$  decrease in soil evaporation. Together, a one-day advancement of spring in CLM-  
411 PhenoCam would result in a  $1.34 \pm 0.05 \text{ mm yr}^{-1} \text{ d}^{-1}$  ( $r=0.90, p < 0.001$ ) increase in the  
412 ecosystem ET.

413

414

415

416

417

418 **Table 1.** Sensitivities of carbon and water fluxes in the Northern Hemisphere broadleaf  
 419 deciduous forest to a one-day advancement of SOS, calculated by the difference in globally-  
 420 integrated carbon and water fluxes between CLM-PhenoCam and CLM using the standard winter  
 421 deciduous spring phenology submodel. The sensitivities were calculated based on the linear  
 422 slopes of differences in fluxes against differences in SOS date, under both RCP 4.5 and 8.5  
 423 scenarios. The statistical significance of these relationships is all  $p < 0.001$ . Units are Pg C yr<sup>-1</sup>  
 424 d<sup>-1</sup> for GPP, AR and NPP; and are mm yr<sup>-1</sup> d<sup>-1</sup> for QVEGT, QVEGE and QSOIL.

	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

425

#### 426 4. Discussion

427 Stimulated by concerns related to climate-change impacts on terrestrial ecosystems, there  
 428 has been substantial effort devoted to improving the accuracy of many widely-used land surface  
 429 models (Williams *et al.*, 2009), including CLM (Bonan *et al.*, 2011; Levis *et al.*, 2012). However,  
 430 it has also been demonstrated that phenology is one area where the performance of existing land  
 431 surface models is particularly poor (Keenan *et al.*, 2012), and it has been argued that there is no

432 reason to expect that performance will be improved under future climate scenarios (Richardson  
433 *et al.*, 2012). While there is broad consensus that future warming is likely to speed up plant  
434 developmental processes and advance spring phenology in temperate forests (Saxe *et al.*, 2001),  
435 accurately forecasting the impacts of these changes on carbon and water fluxes requires better  
436 phenology submodels to be integrated into large-scale land surface models. Our results—  
437 obtained using a phenology model tuned to an extensive dataset derived from near-surface  
438 remote sensing and validated globally using MODIS observations—show that future shifts in  
439 phenology are likely to be smallest in warmer, lower-latitude temperate forests and largest in  
440 colder, higher-latitude boreal forests. By coupling vegetation phenology to carbon and water  
441 cycling processes, our analysis also shows that use of the standard spring phenology submodel in  
442 CLM is likely to substantially under-predict C uptake and evapotranspiration across deciduous  
443 broadleaf forests in the Northern Hemisphere. This confirms that accurate prediction of spring  
444 phenological transitions is essential to reduce uncertainties in quantifying land-atmosphere  
445 exchanges of carbon and water under future climate scenarios. We note that our model runs were  
446 conducted “offline”, in that the forcing was prescribed and the biosphere does not feedback to  
447 the climate system. We expect that with fully coupled runs in the CESM, the increased carbon  
448 uptake and evapotranspiration predicted by CLM-PhenoCam could have a substantial influence  
449 on the evolution of the climate system over the next 85 years.

450         We showed that when evaluated against MODIS-derived SOS for the entire Northern  
451 Hemisphere boreal and temperate deciduous broadleaf forest, the PhenoCam spring phenology  
452 submodel made considerably more accurate SOS predictions than the standard spring deciduous  
453 phenology submodel used in CLM. Against several different types of validation data, at different  
454 levels of spatial aggregation, the PhenoCam submodel had both lower RMSE and lower bias

455 than the standard submodel (Fig. 3). This suggests that, for the most part, the PhenoCam sites  
456 (spanning 15° latitude, from 35 to 50 °N, and almost two months in spring onset date, from day  
457 of year 80 to 140) we used for model development effectively capture the dominant patterns of  
458 spatial variation of spring phenology in Northern Hemisphere deciduous broadleaf forests.  
459 However, careful examination of the results clearly reveals that model performance was poorest  
460 in warm regions such as southeastern China and the southeastern United States. In these regions,  
461 PhenoCam model predictions showed significant bias towards later SOS compared to MODIS-  
462 derived SOS. One possible source of this bias is that there are large number of missing MODIS  
463 observations in these regions, which may lead to erroneous detection of seasonality metrics of  
464 OGI and MAT and therefore the MODIS SOS dates. Meanwhile, none of the PhenoCam sites  
465 that were used to calibrate the model were located in warm, low-latitude (22-35 °N) temperate  
466 forests, probably leading the model to be over-fit to cooler, northern sites. As more sites are  
467 added to the PhenoCam network in coming years (or as complementary data become available  
468 from other networks around the globe), it will be important to re-estimate this model using a  
469 more geographically representative sample. The increasing availability of other long-term,  
470 spatially extensive phenological datasets, e.g. from the USA-National Phenology Network  
471 (Jeong & Medvigy, 2014; Melaas *et al.*, 2016), should prove invaluable for the development and  
472 testing of new phenology models with better accuracy and improved generalizability  
473 (Richardson *et al.*, 2013).

474         A few previous studies have reported adaptation (implying genetic change from natural  
475 selection) or acclimation (implying reversible physiological adjustment) of plant spring  
476 phenology to warming temperatures (e.g., van Asch *et al.*, 2007; Bennie *et al.*, 2010; Keller *et al.*,  
477 2011), suggesting an evolutionary tradeoff between advancing spring onset and avoiding

478 catastrophic disturbance (e.g., spring frost and insect outbreaks). However, these responses may  
479 vary among species and also geographically. Thus there is a lack of knowledge regarding the  
480 capacity for forest tree phenology to adjust to rising temperatures via either mechanism. We are  
481 not aware of any existing large-scale phenology model that explicitly considers adaptation to  
482 future climate change. The CLM model implicitly assumes that acclimation will occur, as the  
483 GDD threshold to trigger bud burst depends on mean air temperature in the previous year  
484 (Equation 2), but this is a short-term and reversible response. Notably, a surprising result from  
485 this model—which occurs precisely because of the acclimation effect—is the prediction that  
486 spring bud burst will actually be delayed for some warmer regions of the world by 2100,  
487 compared to present-day conditions (Fig. 4a,d). While there is some evidence that failure to meet  
488 chilling requirements may delay spring bud burst in a small selection of species (e.g. (Heide,  
489 1993; Orlandi *et al.*, 2004; Schwartz & Hanes, 2010)), we are not aware of any observational  
490 studies which have yet demonstrated this kind of phenological delay, at a regional scale, in  
491 response to recent warming trends.

492         Similar to our results, many previous studies have found that land-atmosphere exchanges  
493 of carbon and water are sensitive to vegetation phenology (Richardson *et al.*, 2013). For example,  
494 Richardson *et al.*, (2009) used ground observations of spring phenology, together with eddy  
495 covariance measurements of CO<sub>2</sub> exchange to estimate that earlier spring leaf-out increased  
496 annual GPP by about 10 g C m<sup>-1</sup> d<sup>-1</sup> in a temperate deciduous forest. By comparison, across  
497 Northern Hemisphere forests, Piao *et al.*, (2007) used a model-based analysis to estimate that a  
498 1-day extension in growing season length was associated with a 5.8 g C m<sup>-2</sup> d<sup>-1</sup> increase in GPP  
499 and a 2.8 g C m<sup>-2</sup> d<sup>-1</sup> increase in NPP. Yue *et al.*, (2015) reported GPP in the north of 30 °N has  
500 increased 0.01 Pg C yr<sup>-1</sup> d<sup>-1</sup> due to the phenological change during 1982-2011 based on global

501 model simulations. Similarly, Zha *et al.*, (2010) reported that warmer spring temperatures  
502 advanced spring leaf-out and enhanced both springtime and annual evapotranspiration in  
503 Western Canadian ecosystems. By analyzing long-term flux data and ground observations of  
504 phenology during the past two decades, Keenan *et al.* (2014) suggested that a one-day change in  
505 SOS would result in both more GPP ( $7.5 \text{ g C m}^{-2} \text{ d}^{-1}$ ) and more net C uptake ( $4.5 \text{ g C m}^{-2} \text{ d}^{-1}$ ).  
506 However, with observational studies, estimates of relationships between shifts in phenology and  
507 shifts in productivity may be confounded by processes that are simultaneously affected by the  
508 same factors driving variation in phenology, but which are not directly linked to phenology. A  
509 strength of our analysis is that it provides a defensible “model experiment” framework, with both  
510 a “control” (or null) model (in this case, the standard CLM spring phenology submodel) and a  
511 “treatment” model (CLM-PhenoCam). By differencing the predictions of the two models, we can  
512 isolate effects of phenology ecosystem processes. This approach allows us to eliminate the  
513 impact of climate change on other processes which are not phenologically-mediated but that  
514 could be mistakenly attributed to phenology if only a single model was used. (e.g., (Piao *et al.*,  
515 2007)), because of their underlying covariation with temperature and hence SOS.

516         Changes in spring leaf out dates have been hypothesized to have both direct (i.e.,  
517 immediate) and indirect (i.e., lagged) effects on ecosystem processes (Richardson *et al.*, 2009).  
518 While the empirical evidence for the indirect effects of phenological anomalies is mixed  
519 (Richardson *et al.*, 2010), our model-based analysis allowed us to distinguish between the direct  
520 and the indirect effects. Our results show that the direct effects of changes in SOS account for the  
521 majority (about 66% for GPP, 44% for AR and 86% for NPP) of the total (annual) differences in  
522 modeled carbon fluxes by the end of the 21<sup>st</sup> century (2090-2100). By comparison, for water  
523 fluxes, the direct effects of changes in SOS account for a smaller fraction (33% for soil

524 evaporation, 43% for vegetation evaporation, and 43% for vegetation transpiration) of the  
525 differences, with the indirect effects being much larger. This may have to do with available soil  
526 water serving as both a constraint and a buffer on ET (Migliavacca *et al.*, 2012). However,  
527 empirical validation of these results using FLUXNET data is advised before these conclusions  
528 can be made with confidence. In summary, the results of this study show improved estimation of  
529 spring phenology for Northern Hemisphere boreal and deciduous broadleaf forests by  
530 incorporating a new and optimized submodel into the CLM version 4.5. The PhenoCam spring  
531 phenology submodel outperforms the standard CLM seasonal-deciduous spring phenology  
532 submodel in terms of both better accuracy and precision. Our analysis shows that with the  
533 standard seasonal-deciduous spring phenology submodel, errors in modeled SOS will propagate  
534 into modeled carbon and water fluxes (Richardson *et al.*, 2012), and that these errors are  
535 exacerbated under future climate change. Our results suggest that with the standard seasonal-  
536 deciduous spring submodel, CLM may under-estimate GPP by  $0.6 \text{ Pg C yr}^{-1}$  and NPP by  $0.3 \text{ Pg}$   
537  $\text{C yr}^{-1}$  by the end of the 21<sup>st</sup> century. While the under-estimation of GPP is small relative to  
538 global terrestrial GPP (estimated at  $123 \pm 8 \text{ Pg C yr}^{-1}$  by Beer *et al.*, (2010)), the under-  
539 estimation of NPP is considerable relative to the global terrestrial C sink, which for 2013 is  
540 estimated to be  $2.5 \pm 0.9 \text{ Pg C yr}^{-1}$  (Le Quéré *et al.*, 2015). Therefore, our results argue for a  
541 reconsideration of the standard seasonal-deciduous spring phenology submodel in CLM, as  
542 substantial errors in predictions of key land-atmosphere fluxes, as well as interactions and  
543 feedbacks between the biosphere and the climate system, may otherwise result (Richardson *et al.*,  
544 2013).

**545 Acknowledgement**

546 This research was supported by the National Science Foundation's Macrosystem Biology (EF-  
547 1065029) and LTER (DEB-1237491, DEB-1114804) programs, DOE Regional and Global  
548 Climate Modeling DE-SC0016011, and by NASA Grant number NNX14AJ35G. The CESM  
549 project is supported by the National Science Foundation and the Office of Science (BER) of the  
550 U.S. Department of Energy. Computing resources were provided by the Climate Simulation  
551 Laboratory at NCAR's Computational and Information Systems Laboratory (CISL), sponsored  
552 by the National Science Foundation and other agencies. PEP725 data were provided by the  
553 members of the PEP725 project. We acknowledge the E-OBS dataset from the EU-FP6 project  
554 ENSEMBLES (<http://ensembles-eu.metoffice.com>) and the data providers in the ECA&D  
555 project (<http://www.ecad.eu>).

556 **References**

- 557 van Asch M, van Tiendren PH, Holleman LJM, Visser ME (2007) Predicting adaptation of  
558 phenology in response to climate change, an insect herbivore example. *Global Change*  
559 *Biology*, **13**, 1596–1604.
- 560 Beer C, Reichstein M, Tomelleri E et al. (2010) Terrestrial Gross Carbon Dioxide Uptake:  
561 Global Distribution and Covariation with Climate. *Science*, **329**, 834–838.
- 562 Bennie J, Kubin E, Wiltshire A, Huntley B, Baxter R (2010) Predicting spatial and temporal  
563 patterns of bud-burst and spring frost risk in north-west Europe: the implications of local  
564 adaptation to climate. *Global Change Biology*, **16**, 1503–1514.
- 565 Blanken PD, Black TA (2004) The canopy conductance of a boreal aspen forest, Prince Albert  
566 National Park, Canada. *Hydrological Processes*, **18**, 1561–1578.
- 567 Bonan GB, Lawrence PJ, Oleson KW et al. (2011) Improving canopy processes in the  
568 Community Land Model version 4 (CLM4) using global flux fields empirically inferred  
569 from FLUXNET data. *Journal of Geophysical Research: Biogeosciences*, **116**, n/a–n/a.
- 570 Burnham KP, Anderson DR (2002) *Model selection and multimodel inference: a practical*  
571 *information-theoretic approach*. Springer Science & Business Media.
- 572 Cannell MGR, Smith RI (1983) Thermal Time, Chill Days and Prediction of Budburst in *Picea*  
573 *sitchensis*. *Journal of Applied Ecology*, **20**, 951–963.
- 574 Dahlin KM, Fisher RA, Lawrence PJ (2015) Environmental drivers of drought deciduous  
575 phenology in the Community Land Model. *Biogeosciences Discuss.*, **12**, 5803–5839.
- 576 Defries RS, Hansen MC, Townshend JRG, Janetos AC, Loveland TR (2000) A new global 1-km  
577 dataset of percentage tree cover derived from remote sensing. *Global Change Biology*, **6**,  
578 247–254.
- 579 Friedl MA, Sulla-Menashe D, Tan B, Schneider A, Ramankutty N, Sibley A, Huang X (2010)  
580 MODIS Collection 5 global land cover: Algorithm refinements and characterization of new  
581 datasets. *Remote Sensing of Environment*, **114**, 168–182.
- 582 Ganguly S, Friedl MA, Tan B, Zhang X, Verma M (2010) Land surface phenology from MODIS:  
583 Characterization of the Collection 5 global land cover dynamics product. *Remote Sensing of*  
584 *Environment*, **114**, 1805–1816.
- 585 Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M (2008) A European  
586 daily high-resolution gridded data set of surface temperature and precipitation for 1950–  
587 2006. *Journal of Geophysical Research*, **113**, D20119.
- 588 Heide OM (1993) Dormancy release in beech buds (*Fagus sylvatica*) requires both chilling and  
589 long days. *Physiologia Plantarum*, **89**, 187–191.
- 590 Hollinger DY, Ollinger S V, Richardson AD et al. (2010) Albedo estimates for land surface  
591 models and support for a new paradigm based on foliage nitrogen concentration. *Global*  
592 *Change Biology*, **16**, 696–710.
- 593 Houghton RA (2007) Balancing the Global Carbon Budget. *Annual Review of Earth and*

- 594 *Planetary Sciences*, **35**, 313–347.
- 595 Hufkens K, Friedl M, Sonnentag O, Braswell BH, Milliman T, Richardson AD (2012) Linking  
596 near-surface and satellite remote sensing measurements of deciduous broadleaf forest  
597 phenology. *Remote Sensing of Environment*, **117**, 307–321.
- 598 Jeong S-J, Medvigy D (2014) Macroscale prediction of autumn leaf coloration throughout the  
599 continental United States. *Global Ecology and Biogeography*, **23**, 1245–1254.
- 600 Keenan TF, Davidson E, Moffat AM, Munger W, Richardson AD (2012) Using model-data  
601 fusion to interpret past trends, and quantify uncertainties in future projections, of terrestrial  
602 ecosystem carbon cycling. *Global Change Biology*, **18**, 2555–2569.
- 603 Keenan TF, Darby B, Felts E et al. (2014a) Tracking forest phenology and seasonal physiology  
604 using digital repeat photography: a critical assessment. *Ecological Applications*, **24**, 1478–  
605 1489.
- 606 Keenan TF, Gray J, Friedl MA et al. (2014b) Net carbon uptake has increased through warming-  
607 induced changes in temperate forest phenology. *Nature Clim. Change*, **4**, 598–604.
- 608 Keller SR, Soolanayakanahally RY, Guy RD, Silim SN, Olson MS, Tiffin P (2011) Climate-  
609 driven local adaptation of ecophysiology and phenology in balsam poplar, *Populus*  
610 *balsamifera* L. (Salicaceae). *American journal of botany*, **98**, 99–108.
- 611 Klosterman ST, Hufkens K, Gray JM et al. (2014) Evaluating remote sensing of deciduous forest  
612 phenology at multiple spatial scales using PhenoCam imagery. *Biogeosciences*, **11**, 4305–  
613 4320.
- 614 Lawrence DM, Oleson KW, Flanner MG et al. (2011) Parameterization improvements and  
615 functional and structural advances in Version 4 of the Community Land Model. *Journal of*  
616 *Advances in Modeling Earth Systems*, **3**, M03001.
- 617 Levis S, Bonan GB, Kluzek E, Thornton PE, Jones A, Sacks WJ, Kucharik CJ (2012) Interactive  
618 Crop Management in the Community Earth System Model (CESM1): Seasonal Influences  
619 on Land–Atmosphere Fluxes. *Journal of Climate*, **25**, 4839–4859.
- 620 Luysaert S, Inglima I, Jung M et al. (2007) CO<sub>2</sub> balance of boreal, temperate, and tropical  
621 forests derived from a global database. *Global Change Biology*, **13**, 2509–2537.
- 622 Meehl GA, Washington WM, Arblaster JM et al. (2013) Climate Change Projections in  
623 CESM1(CAM5) Compared to CCSM4. *Journal of Climate*, **26**, 6287–6308.
- 624 Melaas EK, Friedl MA, Richardson AD (2015) Multi-scale modeling of spring phenology across  
625 Deciduous Forests in the Eastern United States. *Global Change Biology*, in Press.
- 626 Migliavacca M, Sonnentag O, Keenan TF, Cescatti A, O’Keefe J, Richardson AD (2012) On the  
627 uncertainty of phenological responses to climate change, and implications for a terrestrial  
628 biosphere model. *Biogeosciences*, **9**, 2063–2083.
- 629 Oleson KW, Lawrence DM, Bonan GB et al. (2013) *Technical Description of version 4.5 of the*  
630 *Community Land Model (CLM)*. Boulder, CO, 422 pp.
- 631 Orlandi F, Garcia-Mozo H, Ezquerro LV, Romano B, Dominguez E, Galan C, Fornaciari M  
632 (2004) Phenological olive chilling requirements in Umbria (Italy) and Andalusia (Spain).

- 633 *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology*, **138**,  
634 111–116.
- 635 Pan Y, Birdsey RA, Fang J et al. (2011) A Large and Persistent Carbon Sink in the World's  
636 Forests. *Science*, **333**, 988–993.
- 637 Peñuelas J, Rutishauser T, Filella I (2009) Phenology Feedbacks on Climate Change. *Science*,  
638 **324**, 887–888.
- 639 Piao S, Friedlingstein P, Ciais P, Viovy N, Demarty J (2007) Growing season extension and its  
640 impact on terrestrial carbon cycle in the Northern Hemisphere over the past 2 decades.  
641 *Global Biogeochemical Cycles*, **21**, n/a–n/a.
- 642 Le Quéré C, Moriarty R, Andrew RM et al. (2015) Global carbon budget 2014. *Earth Syst. Sci.*  
643 *Data*, **7**, 47–85.
- 644 Richardson A, O'Keefe J (2009) Phenological Differences Between Understory and Overstory.  
645 In: *Phenology of Ecosystem Processes SE - 4* (ed Noormets A), pp. 87–117. Springer New  
646 York.
- 647 Richardson AD, Hollinger DY, Dail DB, Lee JT, Munger JW, O'keefe J (2009) Influence of  
648 spring phenology on seasonal and annual carbon balance in two contrasting New England  
649 forests. *Tree physiology*, **29**, 321–31.
- 650 Richardson AD, Andy Black T, Ciais P et al. (2010) Influence of spring and autumn  
651 phenological transitions on forest ecosystem productivity. *Philosophical Transactions of the*  
652 *Royal Society of London B: Biological Sciences*, **365**, 3227–3246.
- 653 Richardson AD, Anderson RS, Arain MA et al. (2012) Terrestrial biosphere models need better  
654 representation of vegetation phenology: results from the North American Carbon Program  
655 Site Synthesis. *Global Change Biology*, **18**, 566–584.
- 656 Richardson AD, Keenan TF, Migliavacca M, Ryu Y, Sonnentag O, Toomey M (2013) Climate  
657 change, phenology, and phenological control of vegetation feedbacks to the climate system.  
658 *Agricultural and Forest Meteorology*, **169**, 156–173.
- 659 Saxe H, Cannell MGR, Johnsen Ø, Ryan MG, Vourlitis G (2001) Tree and forest functioning in  
660 response to global warming. *New Phytologist*, **149**, 369–399.
- 661 Schwartz MD, Hanes JM (2010) Continental-scale phenology: warming and chilling.  
662 *International Journal of Climatology*, **30**, 1595–1598.
- 663 Sonnentag O, Hufkens K, Teshera-Sterne C et al. (2012) Digital repeat photography for  
664 phenological research in forest ecosystems. *Agricultural and Forest Meteorology*, **152**,  
665 159–177.
- 666 Thornton P., Law B., Gholz HL et al. (2002) Modeling and measuring the effects of disturbance  
667 history and climate on carbon and water budgets in evergreen needleleaf forests.  
668 *Agricultural and Forest Meteorology*, **113**, 185–222.
- 669 Thornton PE, Thornton MM, Mayer BW, Wilhelmi N, Wei Y, Devarakonda R, Cook RB (2014)  
670 Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 2.
- 671 Williams M, Richardson AD, Reichstein M et al. (2009) Improving land surface models with

- 672 FLUXNET data. *Biogeosciences*, **6**, 1341–1359.
- 673 Yue X, Unger N, Zheng Y (2015) Distinguishing the drivers of trends in land carbon fluxes and  
674 plant volatile emissions over the past 3 decades. *Atmos. Chem. Phys.*, **15**, 11931–11948.
- 675 Zha T, Barr AG, van der Kamp G, Black TA, McCaughey JH, Flanagan LB (2010) Interannual  
676 variation of evapotranspiration from forest and grassland ecosystems in western Canada in  
677 relation to drought. *Agricultural and Forest Meteorology*, **150**, 1476–1484.
- 678 Zhang X, Friedl MA, Schaaf CB et al. (2003) Monitoring vegetation phenology using MODIS.  
679 *Remote Sensing of Environment*, **84**, 471–475.
- 680
- 681

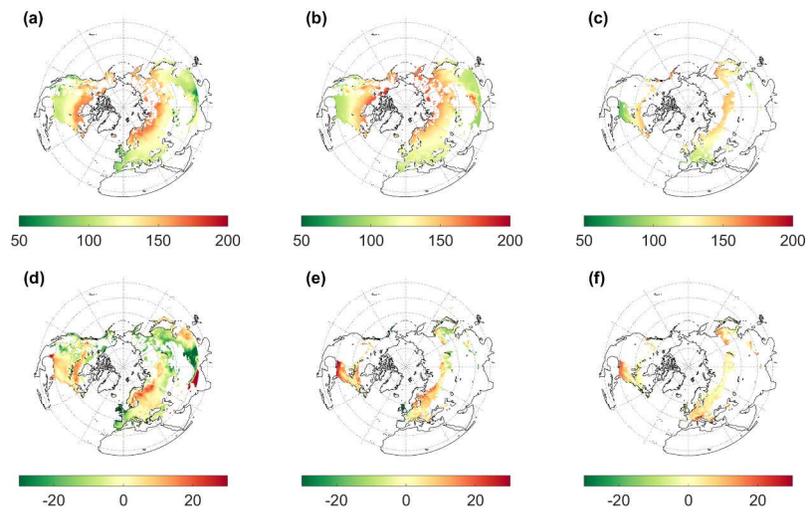


Figure 1. Average SOS (start of spring) dates predicted by CLM and CLM-PhenoCam, compared with MODIS-derived SOS, over the period 2003-2013: (a) CLM predicted SOS; (b) CLM-PhenoCam predicted SOS; (c) MODIS-derived SOS; (d) Differences between CLM and CLM-PhenoCam SOS [(a)-(b)]; (e) Differences between CLM and MODIS SOS [(a)-(c)]; (f) Differences between CLM-PhenoCam and MODIS SOS [(b)-(c)]. 677x350mm (96 x 96 DPI)

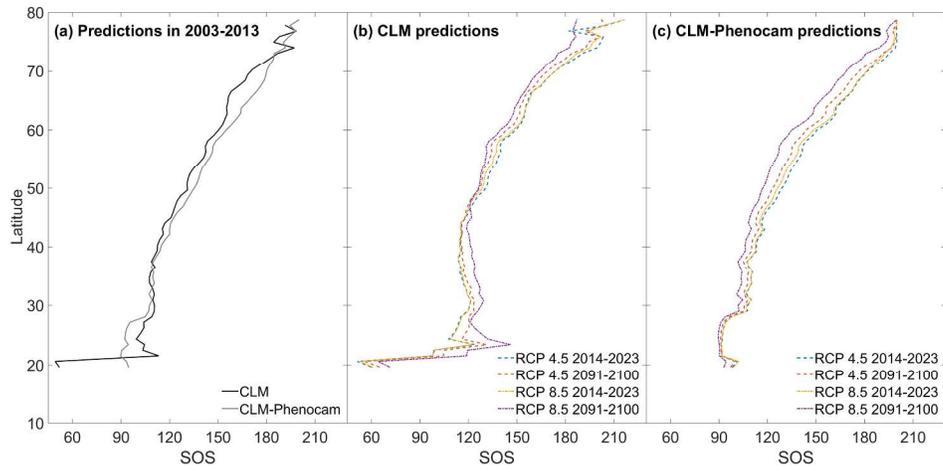


Figure 2. Latitudinal mean of the SOS dates predicted by CLM and CLM-PhenoCam: (a) hindcast predictions, 2001-2013; (b) CLM predictions at the beginning (2014-2023) and end (2091-2100) of the forward runs under RCP 4.5 and 8.5; and (c) CLM-PhenoCam predictions at the beginning (2014-2023) and end (2091-2100) of the forward runs under RCP 4.5 and 8.5.  
677x350mm (96 x 96 DPI)

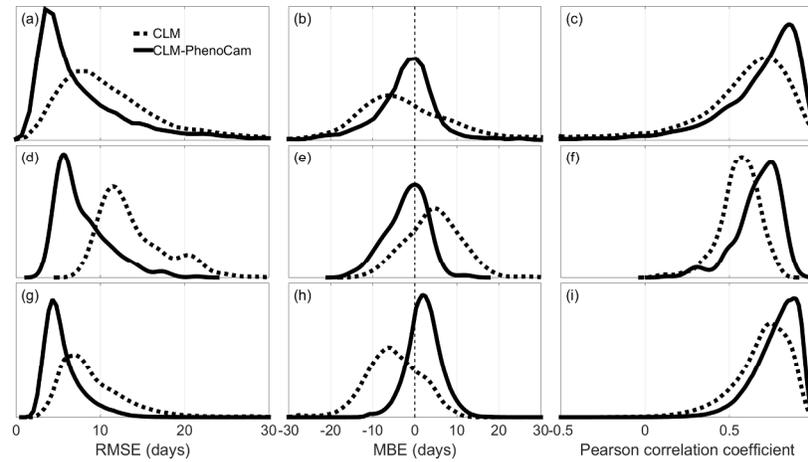


Figure 3. Probability density estimates of the root mean square error (RMSE), mean bias error (MBE), and Pearson correlation coefficient ( $r$ ), for start of spring (SOS) predicted by two models: CLM and CLM-PhenoCam. (a)-(c): comparison against Moderate Resolution Imaging Spectroradiometer (MODIS)-derived SOS across the Northern Hemisphere deciduous broadleaf forest ( $0.9 \times 1.25$  degree grid cells); (d)-(f): comparison against PEP725 data ( $0.25$  degree grid cells); (g)-(i): comparison against MODIS-derived SOS across the eastern US deciduous forest ( $1\text{km}$  grid cells).  
677x350mm (96 x 96 DPI)

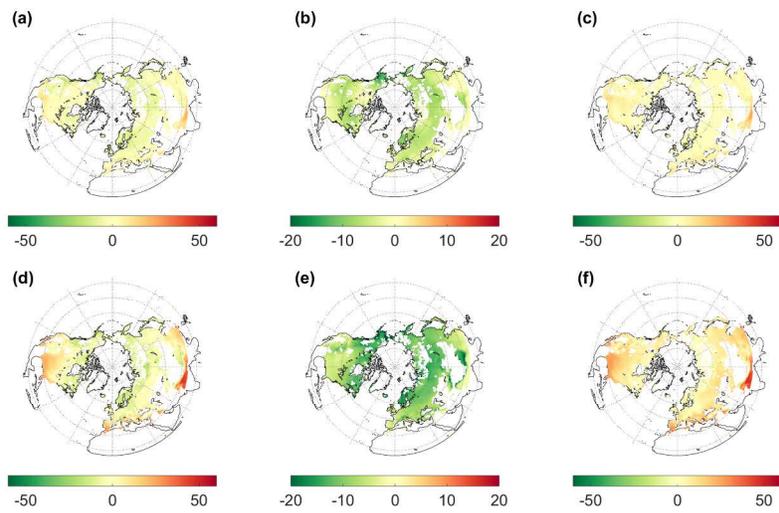


Figure 4. Differences of SOS dates predicted by CLM and CLM-PhenoCam of forward model runs under the RCP 4.5 (the first row) and 8.5 (the second row) scenarios. (a) Changes of SOS dates predicted by CLM between 2014-2023 and 2091-2100 under RCP 4.5 (calculated by using 2091-2100 results minus 2014-2023 results); (b) Same as (a) but from CLM-PhenoCam results; (c) Differences between (b) and (a) [(a)-(b)]; (d)-(f): Same as (a)-(c), respectively, but under RCP 8.5 scenario.  
677x350mm (96 x 96 DPI)

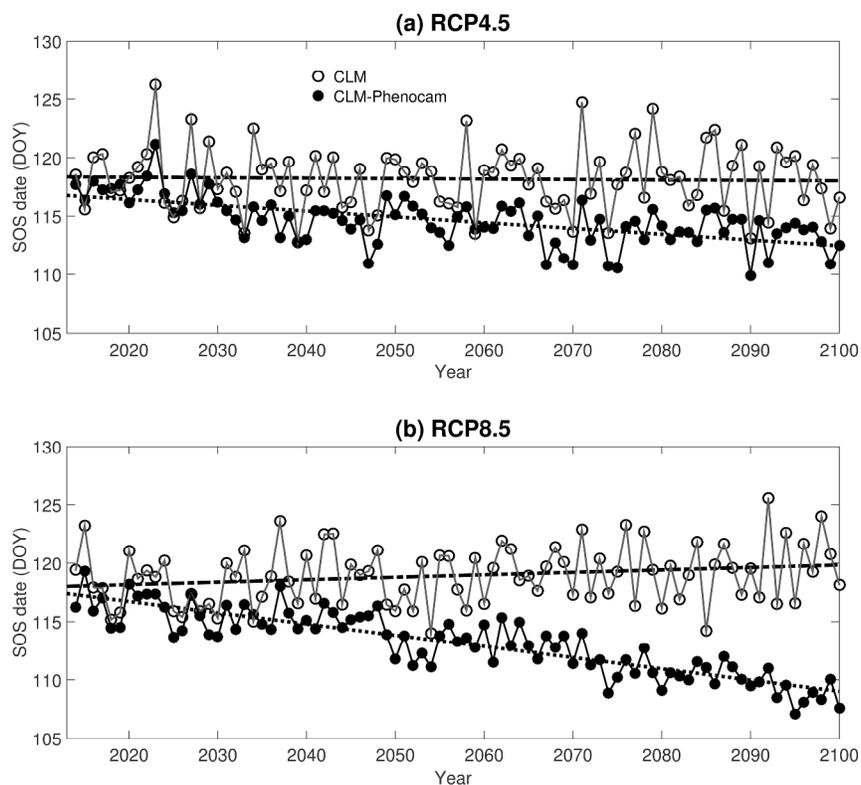


Figure 5. Global area-weighted mean start of spring (SOS) date predicted from 2014 to 2100 under (a) Representative Concentration Pathway (RCP) 4.5 and (b) RCP 8.5 scenarios, for CLM's standard seasonal-deciduous phenology submodel and a revised phenology submodel calibrated to PhenoCam data.  
347x299mm (300 x 300 DPI)

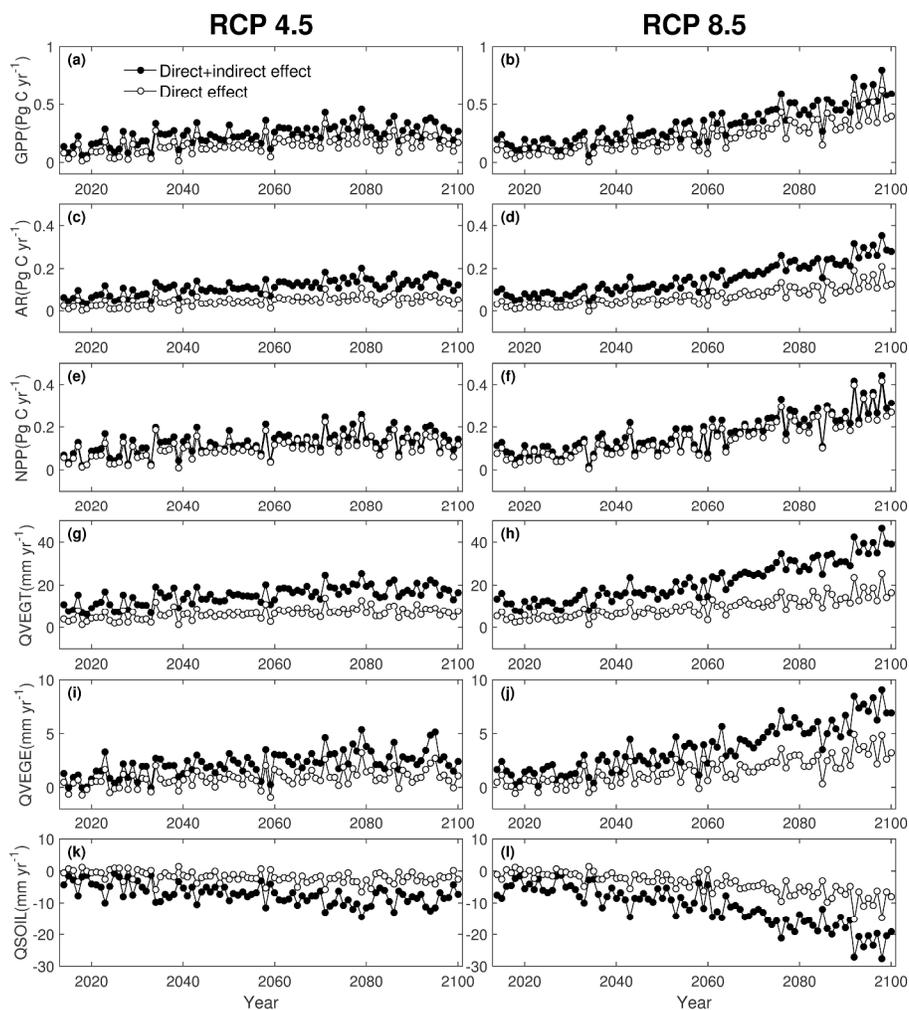


Figure 6. Impacts of start of spring (SOS) submodel on globally-integrated carbon (GPP: gross primary production; AR: autotrophic respiration; NPP: net primary production) and water (QVEGT: vegetation transpiration; QVEGE: vegetation evaporation; QSOIL: soil evaporation) fluxes for 2014 to 2100 under 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.

342x352mm (300 x 300 DPI)

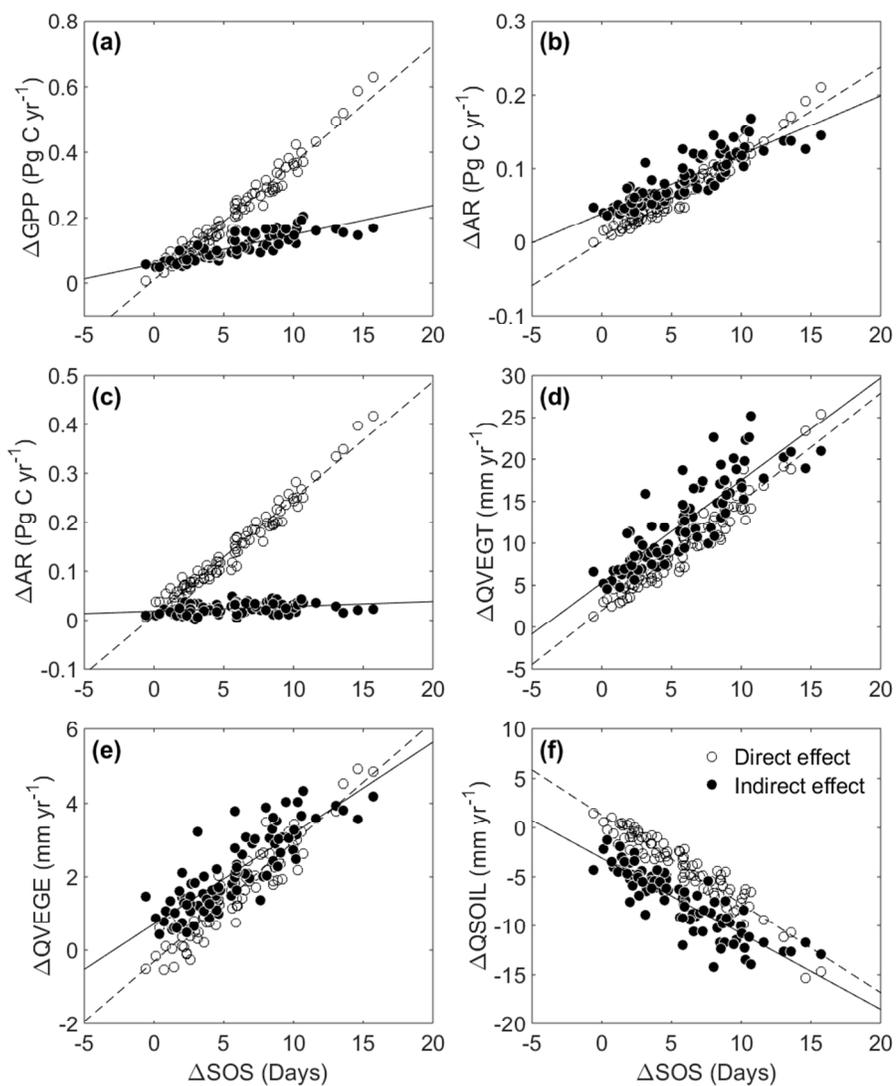


Figure 7. Scatterplots of the direct and indirect differences of annual carbon and water fluxes from CLM and CLM-PhenoCam vs. the SOS differences from the forward model runs under both RCP 8.5 scenario. The relationships are essentially identical for RCP 4.5.  
289x347mm (96 x 96 DPI)

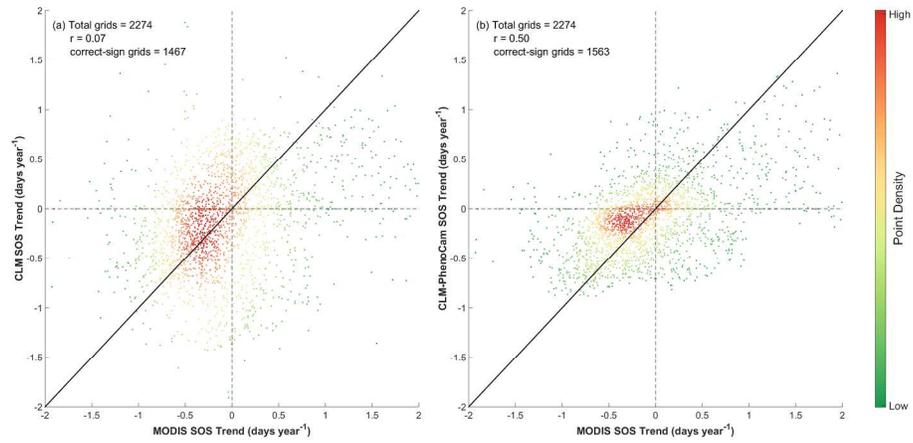


Figure S1. Scatterplots of the trend in MODIS vs. the trend in the CLM (a) and CLM-PhenoCam (b) predicted SOS time series across all Northern Hemisphere boreal and temperate deciduous broadleaf forest grid cells. 39 points fall outside the axis ranges. 677x337mm (96 x 96 DPI)

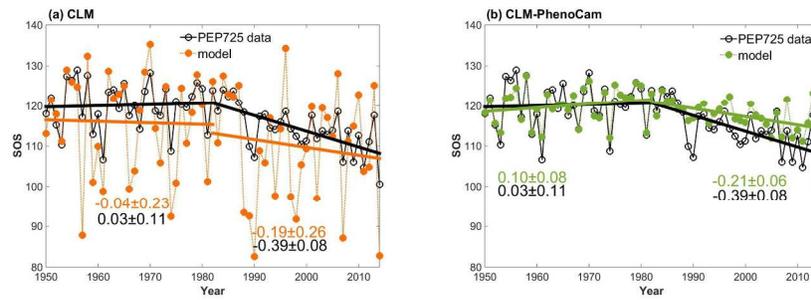


Figure S2. Breakpoint trends of aggregated PEP725 data and model predicted SOS in Europe. 901x268mm (72 x 72 DPI)