Evaluation of leaf phenology of different vegetation types from local to hemispheric scale in CLM

Xiaolu Li[1], Carlos M. Carrillo[1], Toby Ault[1], Andrew Richardson[2], Mark A. Friedl[3], and Steve Frolking[4]

[1]Cornell University
[2]Northern Arizona University
[3]Boston University
[4]University of New Hampshire

May 2, 2023

Abstract

Accurate simulation of plant phenology is important in Earth system models as phenology modulates land-atmosphere coupling and the carbon cycle. Evaluations based on grid-cell average leaf area index (LAI) can be misleading because multiple plant functional types (PFT) may be present in one model grid cell and PFTs with different phenology schemes have different LAI seasonal cycles. Here we examined PFT-specific LAI amplitudes and seasonal cycles in the Community Land Model versions 5.0 and 4.5 (CLM5.0 and CLM4.5) and their relationship with the onset of growing season triggers in the Northern Hemisphere. LAI seasonal cycle and spring onset in CLM show the best agreement with MODIS for temperature-dominated deciduous PFTs. Although the agreement in LAI amplitude between CLM5.0 and MODIS is better than CLM4.5, the agreement in seasonal cycles is worse in CLM5.0. CLM5.0 also simulates higher soil moisture and shows lower influences of soil moisture on LAI amplitudes and seasonal cycles. While productivity depends on the environmental factors to which the plant is exposed during any given growing season, differences in phenology sensitivity to its environment necessitate a decoupling between the seasonality of LAI and GPP, which in turn could lead to biases in the carbon cycle as well as surface energy balance and hence land-atmosphere interactions. Because the discrepancy not only depends on parameterizing phenology but phenology-environment relationship, future improvements to other model components (e.g., soil moisture) could better align the seasonal cycle of LAI and GPP.

Hosted file

960767_0_art_file_10917589_rtg9cp.docx available at https://authorea.com/users/542218/articles/639719-evaluation-of-leaf-phenology-of-different-vegetation-types-from-local-to-hemispheric-scale-in-clm

Hosted file

960767_0_supp_10917543_rtg7t3.docx available at https://authorea.com/users/542218/articles/639719-evaluation-of-leaf-phenology-of-different-vegetation-types-from-local-to-hemispheric-scale-in-clm

Evaluation of leaf phenology of different vegetation types from local to hemispheric scale in CLM

Authors: Xiaolu Li[1,*], Carlos M. Carrillo[1], Toby Ault[1], Andrew D Richardson[2,3], Mark A. Friedl[4], and Steve Frolking[5]

[1] Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, NY 14850, USA

[2] School of Informatics, Computing and Cyber Systems, Northern Arizona University, Flagstaff AZ 86011 USA and

[3] Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff AZ 86011 USA

[4] Department of Earth and Environment, Boston University, Boston, MA 02215, USA

[5] Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham NH 03824 USA

*Corresponding author: Xiaolu Li (xl552@cornell.edu)

Key Points:

- CLM LAI exhibits the best agreement with MODIS in seasonal deciduous PFTs and deciduous broadleaf trees;
- LAI amplitudes are sensitive to environmental factors while LAI seasonal cycle is mostly determined by the phenology scheme;
- Discrepancies in LAI result in biases in GPP, but improvements in one variable may not lead to better results in the other.

Abstract

Accurate simulation of plant phenology is important in Earth system models as phenology modulates land-atmosphere coupling and the carbon cycle. Evaluations based on grid-cell average leaf area index (LAI) can be misleading because multiple plant functional types (PFT) may be present in one model grid cell and PFTs with different phenology schemes have different LAI seasonal cycles. Here we examined PFT-specific LAI amplitudes and seasonal cycles in the Community Land Model versions 5.0 and 4.5 (CLM5.0 and CLM4.5) and their relationship with the onset of growing season triggers in the Northern Hemisphere. LAI seasonal cycle and spring onset in CLM show the best agreement with MODIS for temperature-dominated deciduous PFTs. Although the agreement in LAI amplitude between CLM5.0 and MODIS is better than CLM4.5, the agreement in seasonal cycles is worse in CLM5.0. CLM5.0 also simulates higher soil moisture and shows lower influences of soil moisture on LAI amplitudes and seasonal cycles. While productivity depends on the environmental factors to which the plant is exposed during any given growing season, differences in phenology sensitivity to its environment necessitate a decoupling between the seasonality of LAI and GPP, which in turn could lead to biases in the carbon cycle as well as surface energy balance and hence land-atmosphere interactions. Because the discrepancy not only depends on parameterizing phenology but phenology-environment relationship, future improvements to other model components (e.g., soil moisture) could better align the seasonal cycle of LAI and GPP.

Plain Language Summary

Leaf phenology modifies rates of water, carbon, and energy exchange between the land and the atmosphere. However, large discrepancies have been identified between leaf area simulations from land surface models and remote sensing estimates. Here we evaluated how plant types (e.g., evergreen or deciduous) differed in leaf area magnitude and seasonality in the Community Land Model versions 5.0 and 4.5 (CLM5.0 and CLM4.5) and their relationships with environmental factors in the Northern Hemisphere. Timing of spring leaf growth and growing season length show the best agreement with satellite data for temperature-dominated deciduous vegetation. CLM5.0 had better agreement than CLM4.5 in winter-summer range in the amount of leaf area, but worse agreement on seasonal timing. Because simulated leaf phenology in CLM is primarily determined by how the phenology scheme is parameterized while carbon production is governed by environmental factors, this leads to a decoupling between the seasonality of leaf area and that of plant productivity. This could lead to model biases in both the carbon cycle and land-atmosphere coupling. Since the mismatch depends on both the phenology scheme and environmental factors, future improvements to other model components (e.g., soil moisture) could better align the seasonality of leaf area and plant productivity.

1 Introduction

Plant phenophase modulates ecosystem function, the flow of carbon through the terrestrial biosphere, and land-atmosphere coupling (e.g. Fitzjarrald et al., 2001; Morisette et al. 2009; Lawrence & Chase, 2010; Richardson et al. 2013; Renner and Zohner 2018). In temperate and boreal regions, plant phenology modifies the terrestrial carbon cycle by governing the onset and duration of the growing season (Morisette et al. 2009; Richardson et al. 2009, 2010). In addition, plant phenophase changes also regulate the energy and momentum exchanges between the land surface and the atmosphere (Schwartz 1992; Richardson et al. 2013), and therefore influence land-atmosphere coupling strength, as demonstrated by both observations (Berg et al. 2016; Findell et al. 2015; Green et al. 2017) and model experiments (Guillevic et al. 2002; Levis and Bonan 2004; Lorenz et al. 2013; Puma et al., 2013; Xu et al. 2020; Li et al., 2023). In particular, leaf phenology links the biogeophysical and biogeochemical processes in land surface models and therefore influences both the land surface and the atmosphere above it. However, large discrepancies are present in both the amplitude of leaf area index (LAI) and the start and end of the growing season between model simulations and remote sensing estimates (e.g. Richardson et al., 2012; Mahowald et al. 2016; Peano et al. 2019; Park and Jeong, 2021; Song et al. 2021; Li et al., 2022). Because both the amplitude and seasonal cycle of leaf phenology modulate land-atmosphere interactions as well as carbon and other biogeochemical cycles in LSMs, evaluating both LAI values and seasonality is critical for accurately simulating land-atmosphere coupling and the carbon fluxes.

Although leaf phenology is simulated at a daily or higher temporal resolution and separately for different plant functional types (PFT; e.g. Sitch et al. 2003; Krinner et al. 2005; Lawrence et al., 2019), LAI is often recorded as grid cell averages and at a coarser temporal resolution in the history files and therefore most of the large-scale evaluations are based on monthly averages (Richardson et al., 2012; Mahowald et al. 2016; Peano et al. 2019; Park and Jeong 2021; Song et al. 2021). In land surface models that explicitly simulate plant phenology, natural terrestrial ecosystems are classified into different phenological plant functional types (phenoPFTs; e.g. seasonal/cold deciduous and stress/drought deciduous) and phenophases of the PFTs within the same phenoPFT are controlled by the same set of environmental control factors such as temperature, soil moisture, and precipitation. As a result, LAIs of different PFTs within the same grid cell can differ greatly though they are forced by the same meteorological conditions. Because water and carbon fluxes are calculated at the PFT level, evaluating plant phenology at high temporal resolution and PFT level is critical for simulating water and energy exchanges between the land and the atmosphere as well as the carbon cycle, but the skill of land surface models to represent plant phenology at PFT level at fine resolution across large spatial scales are yet to be evaluated.

Here we used the Community Terrestrial System Model (CTSM, previously the Community Land Model, or CLM), which is the land surface model of the Community Earth

System Model (CESM; Danabasoglu et al., 2020). CLM simulates complex biogeophysical and biogeochemical processes over the land surface (Lawrence et al., 2019) and has been widely adopted into other earth system models as well as regional climate models and weather models (e.g. CMCC-ESM2, Lovato et al., 2022; NorESM2, Seland et al., 2020; RegCM, Steiner et al., 2009). In CLM, coupling between vegetated land surface and the atmosphere happens at the plant functional type level, although both underlying soil layers and the atmosphere above receive output from the land surface as aggregated over all PFTs within the grid cell. Each year, the start and end of growth for deciduous vegetation types are triggered by a subroutine that determines leaf phenophase (Oleson et al., 2013; Lawrence et al., 2019). These subroutines are informed by studies that used both in-situ measurements and remote sensing at individual sites or for specific PFTs to fit parameters, characterize ecological processes, and evaluate model performance. For instance, Scholze et al. (2017) used a data assimilation approach at the site level and tuned parameterizations in land surface models to reproduce observed seasonal cycles. Dahlin et al. (2015, 2017) examined stress deciduous phenology in CLM in tropical drylands and found that a precipitation criterion is necessary to prevent rapid onset of growing seasons due to soil moisture fluctuations. Chen et al. (2016) implemented different spring onset triggers for seasonal deciduous trees in CLM and improved the model's simulation of plant productivity. Birch et al. (2021) developed alternate phenology and photosynthesis schemes and adjusted carbon allocation parameters in arctic-boreal regions and improved gross primary production (GPP) simulation. While these studies are crucial for developing model parameters and evaluating ecological processes, it is unclear how well CLM can simulate key aspects of spring "green up" across different climate zones and over varying vegetation types at the hemispheric and continental scales, which will be most important for future climate feedback and carbon sinks or sources.

Previous studies have found large land surface phenology disagreements between CLM and MODIS and also between different versions of CLM (e.g. Richardson et al., 2012; Mahowald et al., 2016; Scholze et al., 2017; Albergel et al. 2018; Li et al., 2022). These discrepancies emerge from different sources including definitions of spring onset, derivation procedure of simulated and observed variables, as well as biases in observations and land surface models. Satellite-derived LAI values, for instance, are based on reflectance and therefore subject to influences of snow cover and atmospheric conditions like cloud cover and diffuse radiation (Myneni et al., 2015). In addition, large biases in model simulations also emerge from how phenology is parameterized and/or how it responds to environmental factors and other model components. Comparisons based on gridcell averages show that even though CLM5.0 exhibits better representations of LAI values and seasonal amplitude than CLM4.5, it fails to represent LAI annual cycle and interannual variability accurately (Li et al., 2022). Because phenology is simulated at the PFT level, reducing these biases and characterizing distinct sources of uncertainty requires comparisons at the vegetation type level.

In this study, we compare leaf phenology at the plant functional type (CLM) - land cover type (MODIS) level at both grid cell level and hemispheric scale in the Northern Hemisphere and seek to answer: (1) across PFTs, how well can CLM simulate LAI variability estimated from MODIS and how does the CLM-MODIS agreement change with PFT and location? (2) how may these disagreements influence how CLM simulates the carbon cycle? and (3) how sensitive is simulated plant phenology to environmental factors such as soil temperature and soil moisture?

2 Data and Methods

2.1 CLM land surface data structure and plant functional types

To represent spatial heterogeneity in the land surface, CLM adopts a "nested subgrid hierarchy" (Lawrence et al., 2019). Under this scheme, grid cells are divided into multiple land units and if a grid cell has natural vegetation (e.g., trees, shrubs, and grasses) in it, then one of its land units is a vegetated land unit. Each grid cell can have multiple land units and each land unit can have multiple columns in it. The vegetated land unit only consists of one single soil column that is further divided into different plant functional types. When the crop model is disabled in CLM, each grid cell can have a maximum of 14 natural plant functional types (PFTs), two generic crop PFTs, and bare ground. The area weight of different PFTs in CLM5.0 for a case run with a standard "year 2000" land surface configuration (e.g., land use/land cover and atmospheric boundary conditions are taken to be representative of the early 2000s) is shown in Figure S1. Vegetation carbon and energy fluxes as well as state variables (e.g., leaf area index and vegetation temperature) are defined and simulated at the PFT level.

Coupling between CLM and the atmosphere component (either a "data" atmosphere or dynamical atmospheric model), happens at the grid cell level. That is, individual PFTs and columns are not assigned a specific location within a grid cell and all the PFTs, together with other land units within the same grid cell, are forced with the same atmospheric conditions at each time step and their fluxes are combined to supply information back to the atmosphere based on the area weight of each PFT. Interactions between the land surface and below-ground soil layers take place at the column level. That is, because all natural PFTs belong to the same soil column, they are faced with the same soil conditions at each time step. Crop PFTs each have their separate soil column when the crop model runs, but because our focus here is on natural ecosystems and their phenology, we excluded crop PFTs and focused only on natural PFTs in our analysis.

The phenology subroutines in CLM govern carbon and nitrogen fluxes for leaf development and litter fall for natural vegetation types. These routines also partially regulate biogeophysical processes, like photosynthesis and canopy hydrology (among others), by modifying LAI over the course of the year. There are three distinct phenology parameterizations in CLM--*seasonal deciduous, stress deciduous,* and *evergreen*--and each parameterization affects LAI in at least one of the model's 14 natural PFTs (Table S1). For example, a growing degree day (GDD) threshold triggers leaf emergence and growth in PFTs that use the *seasonal deciduous* phenology routine (White et al., 1997; Oleson et al., 2013; Lawrence et al., 2019). PFTs governed by *stress deciduous* phenology start growing only after their chilling requirements are met (except for tropical stress deciduous PFTs, e.g., broadleaf deciduous tropical trees, that do not have a chilling requirement) and in response to GDD thresholds and soil moisture (White et al., 1997; Oleson et al., 2013), as well as an antecedent precipitation requirement introduced in CLM5.0 (Dahlin et al., 2015; Lawrence et al., 2019). Evergreen phenology, in contrast, has a fixed background litter fall rate and no leaf onset/offset trigger (Oleson et al., 2013; Lawrence et al., 2019), so new leaf production depends on the current rate of photosynthesis and respiration (Table S1). During each simulation year, PFTs using the deciduous phenology routine allocate some portion of their carbon and nitrogen for leaf development during the next growing season (Thornton and Zimmermann, 2007; Oleson et al., 2013; Lawrence et al., 2019). Onset of leaf development occurs when the environmental thresholds listed above are reached, and CLM begins to allocate stored carbon (and nitrogen) from the previous growing season to increase LAI over a fixed 30-day period for non-evergreen PFTs. However, stress deciduous PFTs can have multiple growing seasons within one year, or a long growing season and no dormancy when conditions are favorable. LAI for each PFT and grid cell average at an illustrative multi-PFT grid cell is shown in Figure S2.

In our experiments, we ran CLM (4.5 and 5.0) with the GSWP3 historical forcing dataset (Muller Schmied et al., 2016). Briefly, GSWP3 data is dynamically downscaled from the 20th Century Reanalysis (20CR, Compo et al., 2011) and corrected using observations derived from satellite remote sensing or station data. Due to its temporal coverage, GSWP3 was used as boundary conditions over the period from January 1, 1970 through December 31, 2014. Although other boundary conditions have been used to examine the skill of CLM (e.g., CRUNCEP in Wang et al., 2014), our findings did not depend strongly on the historical forcing data. We discarded the first 33 years as spin-up and used 2003-2014 to compare CLM and MODIS.

2.2 MODIS land cover types and LAI for different land cover types

Our remote sensing LAI originated from the MODIS Terra MOD15A2H.v006 (Myneni et al., 2015) 500-m product. To compare against different PFTs in CLM, we separated MODIS LAI into different land cover types. The land cover type was based on the MODIS land cover type data product (MCD12Q1.v006, Friedl et al., 2010). Within each CLM finite-volume grid cell, raw 500-m LAIs with a "good quality" flag (i.e., MODIS QA flag = 0) from each land cover type were averaged together to produce a 1° product with eight-day temporal resolution. A cubic spline was then fit to each time series of the 1° grid cell average for each year and used to interpolate the data to daily time steps. Figure S3 shows the spatial distribution and area weight of each MODIS land cover type. Table S2 reports the match between CLM PFTs and MODIS land cover types.

2.3 LAI ratios

To investigate the agreement between CLM and MODIS leaf phenology at a hemispheric scale, we employed two sets of indicators. *LAI ratio* denotes how well LAI amplitude matches while *seasonal ratio* shows how well seasonality matches between CLM and MODIS leaf phenology. At grid points where both a CLM PFT and its corresponding MODIS land cover (Table S2) are present, we defined the annual mean LAI ratio as the annual mean CLM LAI divided by the annual mean MODIS LAI (Equation 1).

$$LAI\ ratio = \frac{\overline{LAI_{CLM}}}{\overline{LAI_{MODIS}}} \qquad \text{(Equation 1)}$$

Here bars over the variable denote the average value. LAI ratio ranges from 0 to ∞. An LAI ratio of 0 means CLM simulates zero LAI. When LAI ratio equals or is close to 1, annual mean CLM LAI values are close to mean MODIS LAI, although no information is provided on how well their seasonal cycles agree. If the LAI ratio is above 1, then MODIS LAI is smaller than CLM LAI, suggesting CLM may have overestimated LAI at those locations.

2.4 Seasonal ratios

We also calculated the root mean square error (RMSE) between normalized CLM LAIs and MODIS LAIs and named it the seasonal ratio ($RMSE_{normLAI}$). For each year between 2003 and 2014, we first normalized daily LAIs from MODIS and CLM, respectively, to remove the impact of differences in LAI amplitudes. Then we calculated the RMSE of normalized LAIs to determine how well the seasonal cycle of LAI agrees between CLM and MODIS (Equation 2).

$$RMSE_{normLAI} = \sqrt[2]{mean[(\frac{LAI_{CLM}-\overline{LAI_{CLM}}}{\sigma(LAI_{CLM})} - \frac{LAI_{MODIS}-\overline{LAI_{MODIS}}}{\sigma(LAI_{MODIS})})^2]} \qquad \text{(Equation 2)}$$

Here bars over the variable denote the average and σ denotes the standard deviation. If $RMSE_{normLAI}$ is large, then normalized CLM LAIs differ from normalized MODIS LAIs substantially over the year, suggesting the CLM seasonal cycle differs from that in MODIS. When $RMSE_{normLAI}$ is close to 0, CLM and MODIS LAI have good agreement on their seasonal variation (Figure S4). Here we focused on results from GSWP3 forced CLM5.0 and CLM4.5, but $RMSE_{normLAI}$ remains little-changed when using a different model-forcing combination.

2.5 LAI threshold-based spring onset timing

We developed a suite of indicators to investigate the start of spring timing in CLM and MODIS. The indices of primary interest here are all based on LAI, which is either derived from

MODIS or calculated internally by CLM. We define the annual dynamical range of LAI as the difference between minimum (winter) and maximum (summer) LAI each year. We then focus on the 50% thresholds of the annual dynamical range of LAI (Figure S5). Using threshold-based indicators reduces the influence of land use change in remote sensing records as well as differences in peak LAI from one year to the next.

2.6 Difference in terrestrial production

To investigate how the discrepancies in LAI seasonal cycles influence plant production, we also computed the gross primary production (GPP) simulated by different CLM PFTs during the differences in the duration of the peak growing season between CLM and MODIS (ΔGPP, see also Li et al., 2022 for grid cell averaged influences on net primary production). We defined peak growing season as days within the year when LAI is above 75% of its annual dynamical range for each PFT. We then calculated the difference between CLM and MODIS peak growing season and estimated the GPP simulated by each PFT during that difference window. CLM GPP is counted as positive when CLM LAI is within its peak growing season but MODIS is not, and negative vice versa. After computing the annual GPP difference induced by different peak growing windows in CLM and MODIS, we averaged across all years to characterize the potential influence of errors in modeled phenology on terrestrial carbon cycle simulations. We only examined ΔGPP induced by differences in peak growing season to diminish the influence of LAI differences and focus on differences due to plant phenology. We also calculated how large ΔGPP is compared to the total annual GPP in CLM and named the index ΔGPP_{pheno}. To verify the inferred GPP and LAI seasonal cycles, we adopted site-level LAI and GPP estimates from two AmeriFlux flux tower sites (Pastorello et al., 2020) at US-Me2 (44.4523°N, 121.5574°W, mostly evergreen forest; Law, 2022) and US-Ho2 (45.2091°N, 68.7470°W, ~90% ENF and 10% DBF in 1km around tower; across region forest type is mixed evergreen and deciduous broadleaf forest; Hollinger, 2021).

2.7 Soil temperature and soil moisture

To investigate how sensitive the indicators are to meteorological variables and the relative importance of forcing versus PFT, we investigated the relationship between the indicators and the temperature and soil moisture from the simulations. We used soil temperature and soil moisture in the top 10 centimeters because CLM uses soil properties from the third soil layer (0.08m) to determine the start of growing season (Table S1; Lawrence et al., 2019).

3 Results

We examined the agreements between CLM and MODIS LAI amplitudes, LAI seasonal cycles, spring onset, their influences on GPP, and how the agreements vary with soil moisture and temperature. We found that (1) CLM phenology agrees best with MODIS in seasonal deciduous PFTs and deciduous broadleaf trees; (2) Phenology-induced biases in GPP are smaller than in LAI, but the discrepancies still result in large biases in GPP; and (3) LAI amplitudes are sensitive to environmental factors while LAI seasonal cycle and spring onset are mostly determined by the phenology scheme. We discuss these results in more detail below.

3.1 Amplitude and seasonal cycle of annual LAI

Simulated LAIs in CLM show the best agreement with MODIS LAI seasonal cycle in seasonal deciduous PFTs, but LAI amplitudes are overestimated at high latitudes (>60°N, Table 1, attached at the end of the manuscript, Figure S6). Over the Northern Hemisphere, CLM5.0 and MODIS display similar LAI seasonal amplitude and variation in deciduous broadleaf dominated boreal and temperate regions (Figures 1gh, 2gh). LAI ratio is close to or lower than 1 and $RMSE_{normLAI}$ is around 0.5 in Eastern US, Europe, East Asia, and along 55°N in Central Asia, suggesting that CLM LAIs are close to or lower than MODIS LAIs in these regions, but the seasonal cycles are similar. $RMSE_{normLAI}$ also decreases when moving to lower latitudes in these regions (Figure 2gh). For broadleaf deciduous boreal shrub, like other high-latitude PFTs, CLM overestimates LAI values but agrees with MODIS on seasonality (Figure 1k, 2k). Across high-latitude regions in North America and Eurasia, CLM5.0 exhibits an LAI ratio of either zero or above five, suggesting that CLM5.0 boreal shrubland LAI is either zero or too high in these regions (Figure 1k). $RMSE_{normLAI}$ is around 0.5 except in north Russia, indicating that CLM and MODIS LAI agree on seasonality over boreal shrublands (Figure 2k). C3 arctic grass generally displays overestimated LAI values but good agreement on LAI seasonal cycle. For the C3 arctic grass PFT in CLM5.0, CLM exhibits larger LAIs with the LAI ratio higher than 6 in some regions in Northern Canada and Northern Russia. $RMSE_{normLAI}$ is around 0.5 or lower, suggesting CLM5 leaf seasonal cycles in seasonal deciduous grasslands agree with MODIS (Figure 2i).

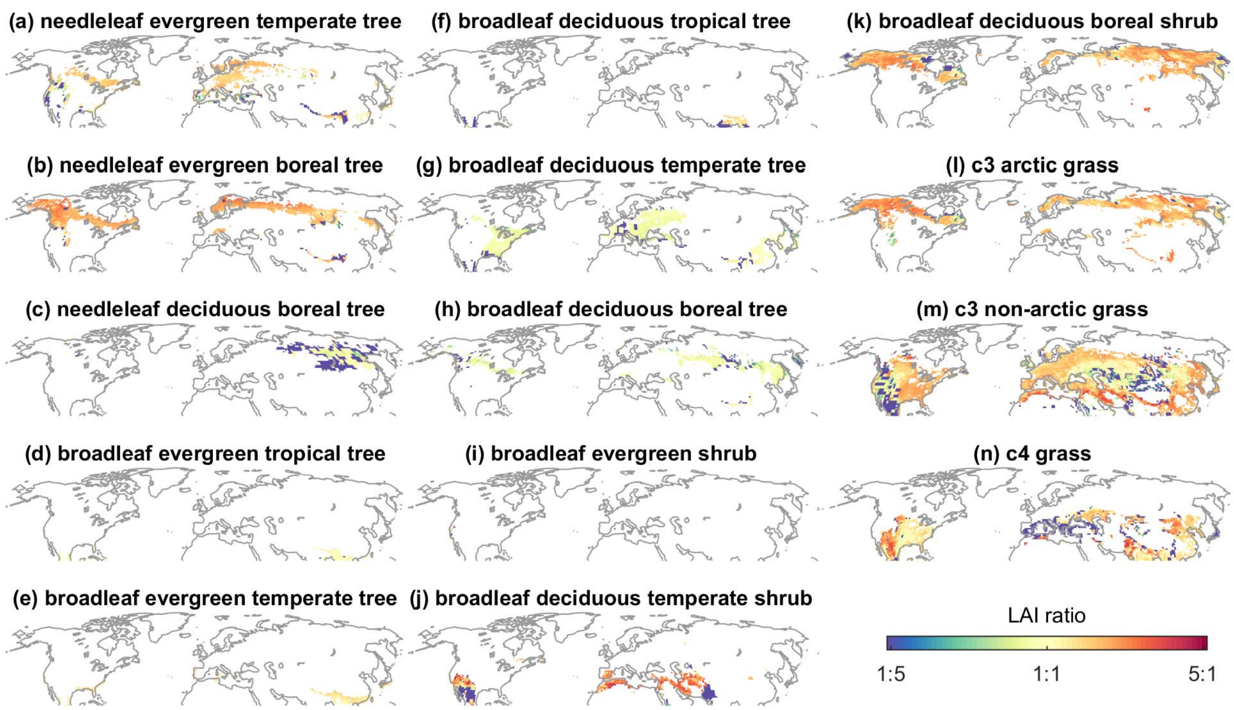

Figure 1. Maps showing agreement of LAI amplitude between GSWP3 forced CLM5.0 and MODIS for different PFTs. Agreement of LAI amplitude, or LAI ratio, is defined as mean annual CLM LAI divided by mean annual MODIS LAI for the corresponding PFT. This ratio reflects how well CLM LAI values match with MODIS. An LAI ratio close to 1 means a good match. If the LAI ratio is less than 1, then MODIS LAI is larger than CLM LAI over the course of a year, and the opposite if LAI ratio is larger than one.

Agreements between MODIS LAI estimates and CLM simulations depend on plant functional type, model version, and atmospheric forcing in stress deciduous PFTs (Table 1, Figure S7). Stress deciduous tree and shrub PFTs display lower agreement between CLM and MODIS in both LAI amplitude and seasonality than the seasonal deciduous trees and shrubs in colder environments. Over the Northern Hemisphere, the LAI ratio of temperate shrub is slightly lower than that of boreal shrub but is still over 3. This suggests that if a temperate shrub has a non-zero LAI (i.e. the PFT survives) in CLM5.0, the LAI values are much higher than MODIS. $RMSE_{normLAI}$ is around 1.5 in the Rocky Mountain regions in the US, the Mediterranean, and Central Asia, indicating a large mismatch between CLM and MODIS LAI seasonal cycles. For locations occupied by tropical deciduous broadleaf trees in Southeast Asia, India, and Mexico, LAI ratio is either zero or higher than one and the RMSEs are generally above 1, indicating that the PFT is either dead or over-productive in CLM and exhibits different seasonal cycles from MODIS (Figures 1f, 2f). Agreement between grassland LAI in CLM5.0 and MODIS varies dramatically across space (Table 1; Figures 1imn, 2imn). For the C3 non-arctic grass PFT in

CLM5.0, CLM exhibits larger LAIs with the LAI ratio higher than 1.5 in Eastern US, South Canada, Central America, the majority of Europe, as well as Central and Eastern Asia. LAI ratio is lower than 0.5 (but still non-zero) at a few locations in the Western US, Northern UK, Eastern Europe, and Central Asia, and decreases to zero in the surrounding regions (Figure 1mn). $RMSE_{normLAI}$ is the largest and more than 1.5 in the Southeast US and Western Europe and decreases from lower to higher latitudes (Figure 2mn). Overall, CLM best simulates MODIS grassland LAIs at locations around 50°N in Eastern Europe, Central Asia, and part of North America. The C4 grass PFT in CLM5.0 exhibits a better agreement on LAI values and seasonal cycles than C3 non-arctic grass. Overall, CLM overestimates growing season length and may simulate multiple growing seasons at locations where only one growing season is presented in MODIS. Matches between CLM and MODIS LAI amplitudes are better over tropical stress deciduous PFTs (e.g., broadleaf deciduous tropical tree and C4 grass; Figure S7ad) whereas CLM overestimates LAI in temperate regions (e.g. broadleaf deciduous temperate shrub; Figure S7b).

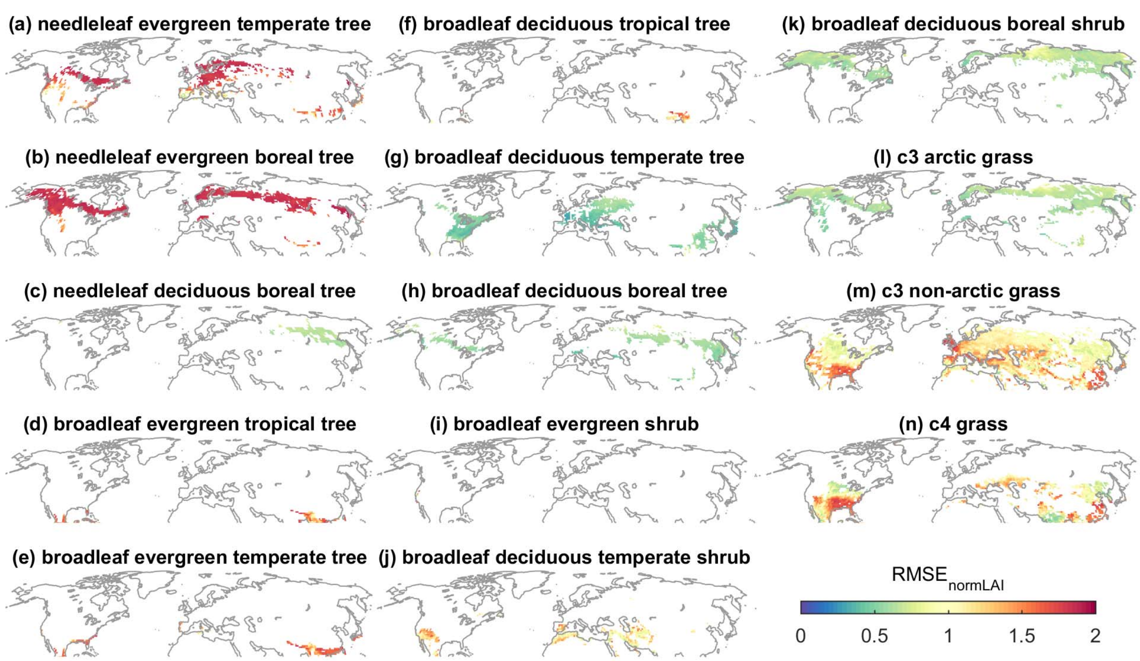

Figure 2. Maps showing agreement of LAI seasonal cycle between GSWP3 forced CLM5.0 and MODIS for different PFTs. Agreement of LAI seasonal cycle, or $RMSE_{normLAI}$, is defined as RMSE between annually normalized LAIs in CLM and MODIS, averaged over 2003-2014. This ratio indicates agreement of LAI seasonal cycles between CLM and MODIS. Smaller $RMSE_{normLAI}$ means better agreement.

Evergreen PFTs exhibit the least influence of atmospheric forcing and exhibit little seasonal variation in CLM simulations (Table 1; Figure S8). While MODIS needleleaf evergreen LAIs can drop to close to zero in winter in temperate and boreal regions, CLM LAI shows much less intra and interannual variability. Peak growing season is also delayed in CLM relative to MODIS over mid-to-high latitude regions (Figure S8ab). CLM disagrees with MODIS LAI in both LAI values and seasonal variation in evergreen needleleaf tree PFTs (Figures 1ab, 2ab), though part of this discrepancy may be due to uncertainties in satellite-derived LAI estimates when evergreen trees can be snow-covered. For evergreen broadleaf forests, CLM5.0 LAI estimates are close to those in MODIS, especially for tropical trees (Figures 1de, S8cd). CLM5.0 overestimates LAIs (LAI ratio>4) for evergreen broadleaf temperate trees in subtropical regions in East and Southeast Asia, Southeast US, and Mexico, while CLM LAIs are close to, or a little lower than, MODIS values (LAI ratio=<1) for evergreen broadleaf tropical trees in Southeast Asia, Central America, and Sahel (Figure 1abde). $RMSE_{normLAI}$ is generally greater than 1 across these regions, indicating that the LAI seasonal cycle in CLM differs from the spline-fitted seasonal variations in MODIS LAI (Figure 2abde).

Agreement between MODIS and CLM LAI values and seasonal cycles depend largely on the plant functional type in CLM. Across high-latitude regions (around 60°N) in North America and Eurasia, LAI ratio is 2 or higher while $RMSE_{normLAI}$ is 1 or lower, suggesting that CLM LAIs are much larger than MODIS values, although they display similar seasonal cycles. Both LAI ratio and $RMSE_{normLAI}$ decrease when moving to lower latitudes, indicating that LAI seasonal amplitude and variation of these two communities are more similar at lower latitudes. In general, CLM LAIs are higher than corresponding MODIS land cover types for needleleaf evergreen trees, shrubs, and grass PFTs (Figures 1, S9). Overall, CLM shows most agreement with MODIS LAI in broadleaf deciduous temperate tree PFT and least agreement in shrub PFTs (Table 1, Figure S9).

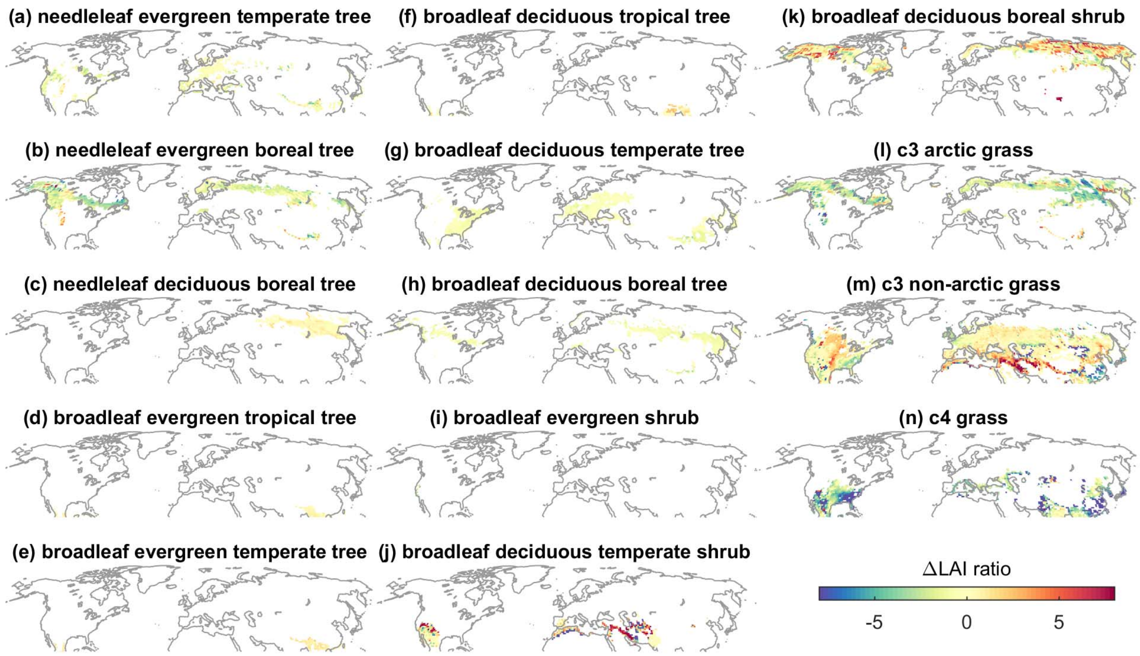

Figure 3. Difference between LAI ratio of GSWP3 forced CLM5.0-MODIS and CLM4.5-MODIS for different PFTs. Agreement of LAI amplitude, or LAI ratio, is defined as mean annual CLM LAI divided by mean annual MODIS LAI for the corresponding PFT.

Across PFTs, CLM5.0 shows improvements from CLM4.5 in the match between PFT and MODIS land cover, reducing the excessive number of growth cycles in stress deciduous PFTs with more than one growing seasons each year, and reduction in LAI overestimation and zero LAIs, but exhibits worse disagreement with MODIS LAI seasonal cycle in stress deciduous and evergreen PFTs. CLM5.0 reduces areas with zero LAI estimates in needleleaf deciduous boreal trees (Figures 1c, 3c), broadleaf deciduous boreal shrub (Figures 1k, 3k), and C3 non-arctic grass (Figures 1m, 3m), and decreases LAI ratio in broadleaf deciduous trees (from >2 in CLM4.5 to ~1; Figures 1gh, 3gh), broadleaf deciduous boreal shrub (Figures 1k, 3k), and C3 arctic grass (from >8 in CLM4.5 to ~4; Figures 1l, 3l). However, estimation of LAI seasonal cycles is largely unchanged for deciduous PFTs in temperate and boreal regions and $RMSE_{normLAI}$ increases by around 1 in needleleaf evergreen trees from CLM4.5 to CLM5.0, suggesting that CLM5.0 experiences larger disagreement with MODIS LAI seasonal cycle than CLM4.5 in evergreen and some stress deciduous PFTs (Figures 4, S9, and Table 1).

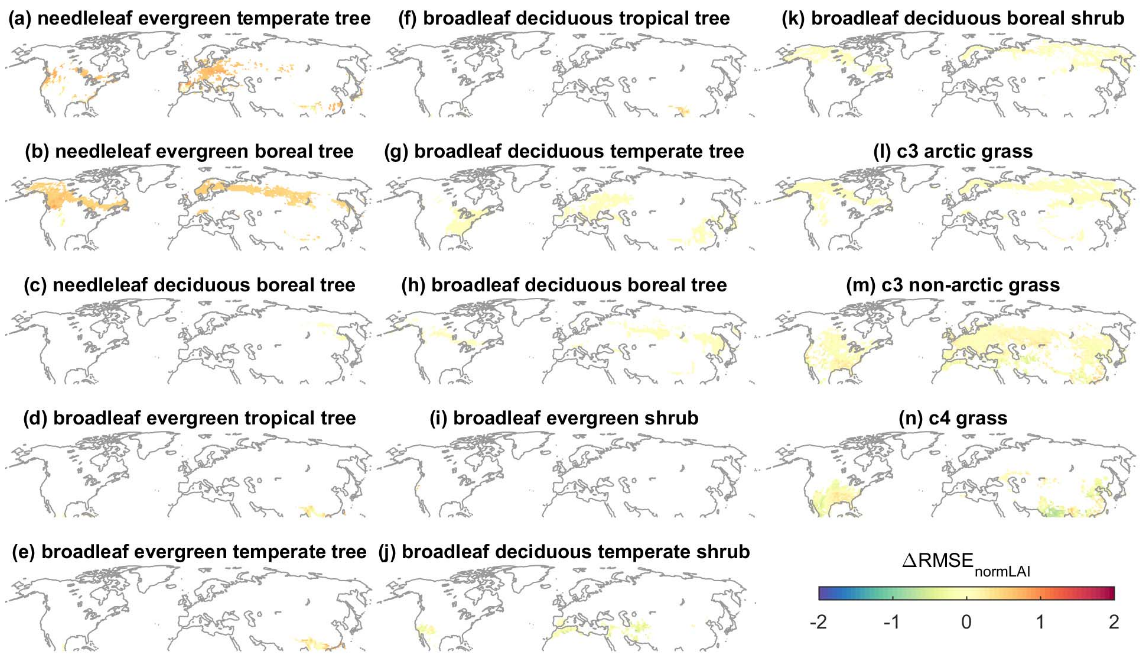

Figure 4. Difference between LAI seasonal cycle agreement of GSWP3 forced CLM5.0-MODIS and CLM4.5-MODIS for different PFTs. Agreement of LAI seasonal cycle, or $RMSE_{normLAI}$, is defined as RMSE between annually normalized LAIs in CLM and MODIS, averaged over 2003-2014. This ratio indicates agreement of LAI seasonal cycle between CLM and MODIS and smaller $RMSE_{normLAI}$ means better agreement.

3.2 Start of growing season and GPP

Across different PFTs in the Northern Hemisphere, CLM5.0 simulates later spring onset than MODIS (Figure 5), except for a few grass-dominated mid-latitude locations, but the difference between CLM5.0 and CLM4.5 phenology varies across vegetation types (Figure S10). CLM5.0 spring onset exhibits the largest difference from MODIS in high-latitude regions that are dominated by needleleaf evergreen trees, broadleaf evergreen trees, broadleaf deciduous shrubs, and non-arctic grasses (Figure 5abejmn), where DOYs of LAI 50% threshold are more than 50 days later in CLM5.0 than MODIS. This difference remains more than 20 days in boreal regions in other PFTs, such as broadleaf deciduous shrub and C3 arctic grass but decreases in mid-latitude regions across the Northern Hemisphere. Across the Northern Hemisphere, CLM5.0 spring onset timing is less than 5 days from CLM4.5 in seasonal deciduous PFTs, while CLM5.0 exhibits later spring onset than CLM4.5 over needleleaf evergreen trees (>50 days), some non-arctic grass PFTs (30 days and above), and some temperate vegetation types in subtropical regions (Figure S10).

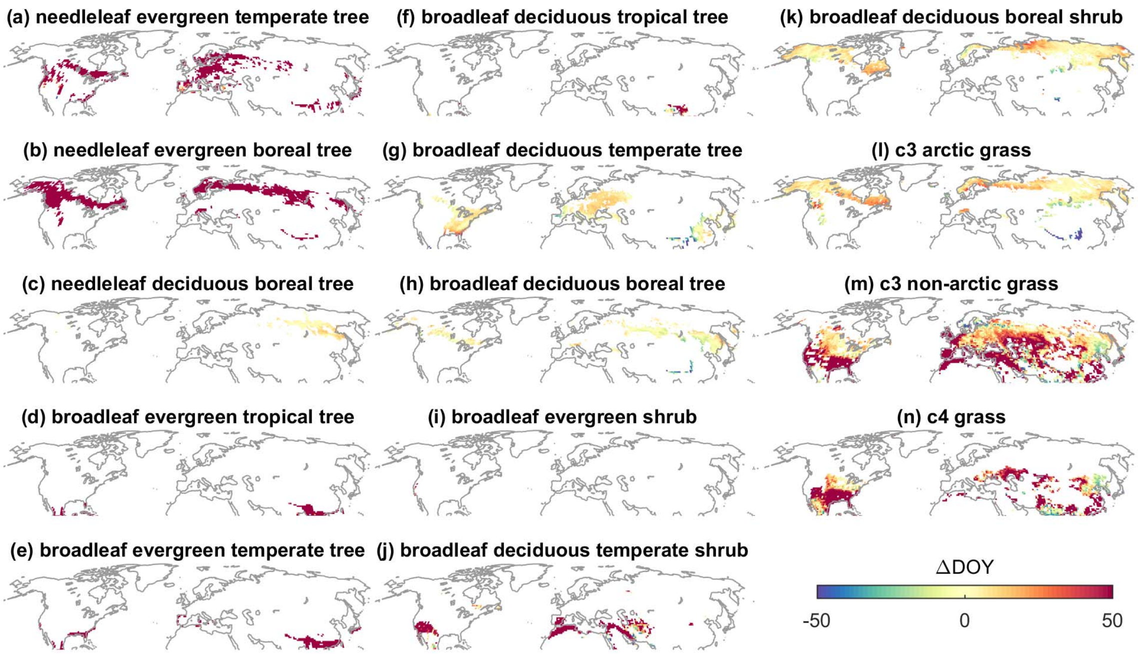

Figure 5. Difference between mean day of year when LAI reaches 50% threshold of LAI annual dynamical range in CLM5.0 and MODIS for each PFT. LAI annual dynamical range is defined as annual maximum LAI minus annual minimum LAI.

Correlations between the DOY indicators in CLM5.0 and MODIS are larger at higher latitudes, suggesting a strong temperature influence (Figure 6). In deciduous PFTs, the highest correlations (>0.8) are present in temperature and boreal deciduous trees, boreal shrubs, and arctic grasses (Figure 6cghki). Evergreen needleleaf trees in temperate and boreal regions display small or even negative correlations with MODIS DOYs (Figure 6ab). From CLM4.5 to CLM5.0, correlations exhibit small and non-uniform changes (Figure S11). Overall, correlations depend largely on the phenology scheme adopted in CLM and are high in deciduous PFTs at high latitudes (Figure 6).

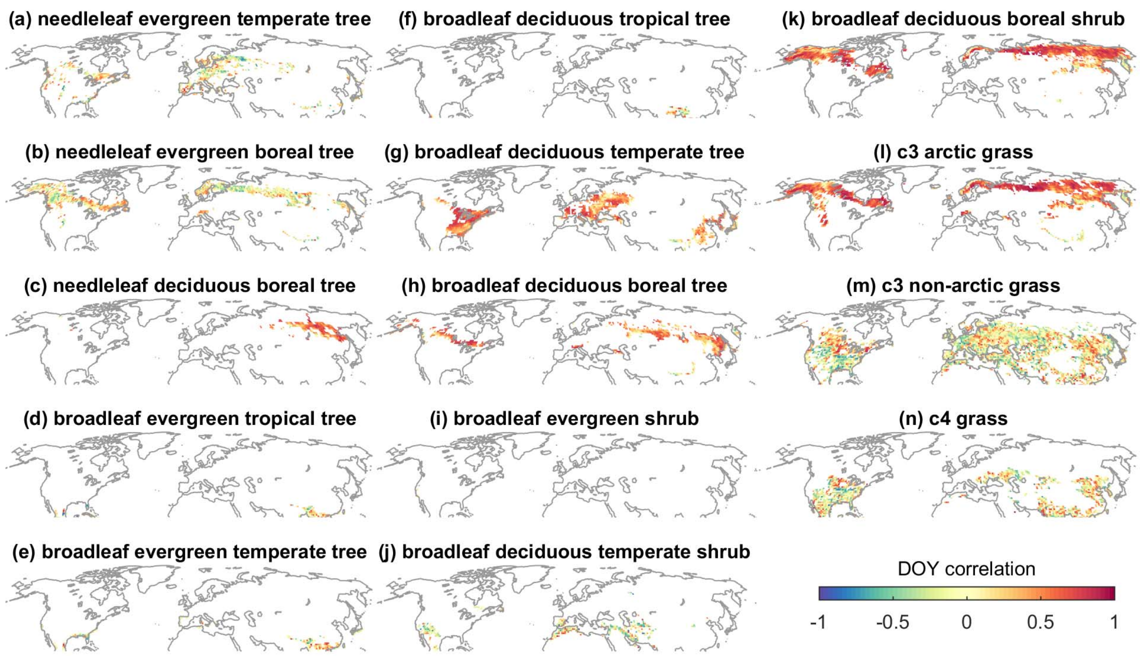

Figure 6. Correlation between day of year when LAI reaches 50% threshold of LAI annual dynamical range in for each PFT in CLM5.0 and corresponding land cover type in MODIS over 2003-2014.

Disagreements between CLM and MODIS leaf phenology lead to underestimation of GPP in CLM in evergreen PFTs and overestimation in deciduous PFTs during the peak growing season (Table 1, Figure 7). On average, GPP bias caused by phenology mismatch is larger in boreal PFTs than temperate and tropical PFTs with the same phenology scheme (Table 1). For evergreen PFTs, peak LAI occurs much later than peak LAI in MODIS, but GPP is more consistent with favorable environmental conditions, causing leaf phenology to be biased later than the seasonality of primary production (Figures 5, 7, S8). For deciduous PFTs, because the start of growing season is later in CLM than MODIS and growing season length is longer (Figures 5, S6-S7), the overall longer peak growing season has a positive influence on GPP (Figure 7). Notably, although phenology induces consistent over- or underestimation of GPP in most PFTs, C3 and C4 grass PFTs in Eastern US, Southern Europe, and Eastern China exhibit underestimated GPP, while in Central and Eastern NA, Central Eurasia and Northern China, GPP is overestimated. Differences between CLM5.0 and CLM4.5 ΔGPP_{pheno} are small, but GPP biases are generally larger in CLM5.0 as CLM5.0 has less agreement with MODIS LAI in the seasonal cycle (Table 1, Figure S12). However, site-level comparisons show that the mismatch between GPP and LAI seasonal cycle in evergreen PFTs may be due to biases in both CLM and MODIS LAI estimates (Figure S13) as MODIS may overestimate LAI seasonal variations in evergreen PFTs (Myneni et al., 2015). However, the potential mismatch between CLM LAI and GPP seasonal cycles could still cause underestimation of GPP in evergreen PFTs.

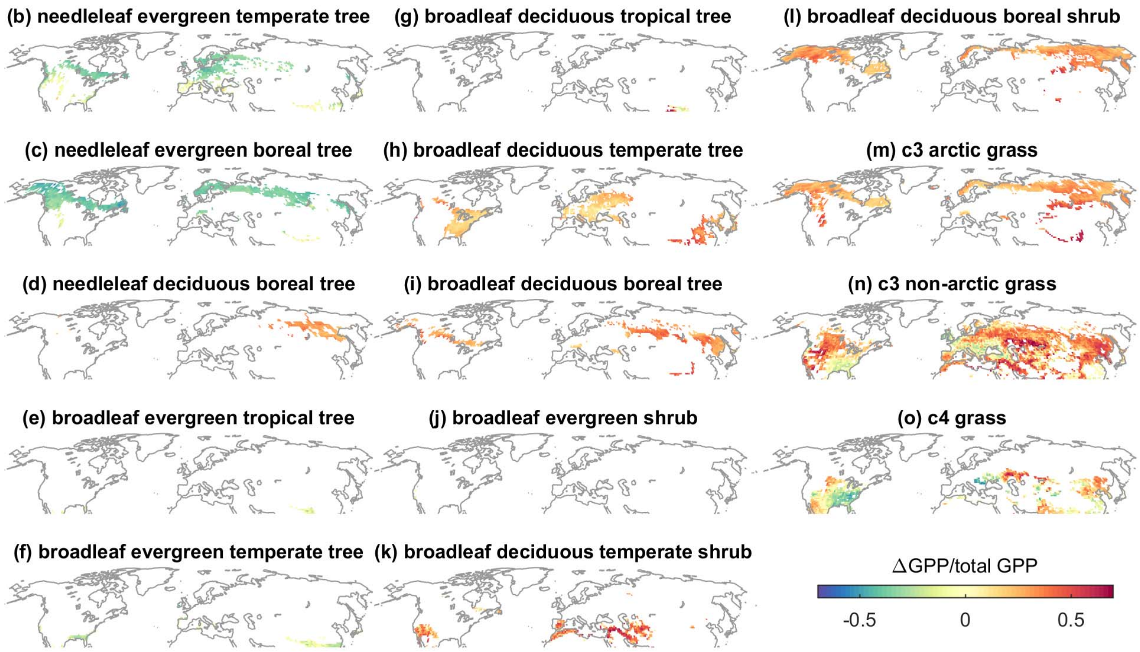

Figure 7 Proportion of GPP produced by CLM during the difference in CLM and MODIS peak growing season devided by total annual GPP in CLM. Peak growing season is defined as days when LAI is over the 75% threshold of its annual dynamical range.

3.3 Environmental factors

To investigate the roles of environmental forcing, model version, and PFT type in modulating leaf phenology, we also examined the relationship between the four sets of indicators (LAI ratio, $RMSE_{normLAI}$, DOY correlation, and ΔGPP_{pheno}) and onset triggers (soil temperature and soil moisture). Soil temperature and PFT play an important role in determining the agreement between CLM and MODIS LAI, whereas soil moisture influences LAI estimation more in CLM4.5 than in CLM5.0. PFT and phenology schemes modulate the influences of soil states on LAI amplitudes and seasonal cycles. Generally, LAI ratio and DOY correlation decrease with increasing mean spring temperature, while $RMSE_{normLAI}$ increases with increasing spring temperature (Figure 8a-f), suggesting LAI seasonal cycles agree better between CLM and MODIS with lower spring temperature, though PFTs play an important role. Influences of soil moisture on $RMSE_{normLAI}$ are largely modified by PFTs while LAI ratio and DOY correlation generally increase with higher spring soil moisture (Figure 9 a-f). For evergreen PFTs, higher spring temperature and lower soil moisture leads to smaller phenology-induced GPP bias, while for deciduous PFTs, lower spring temperature and higher soil moisture generally leads to smaller GPP difference (Figure 8gh, 9gh).

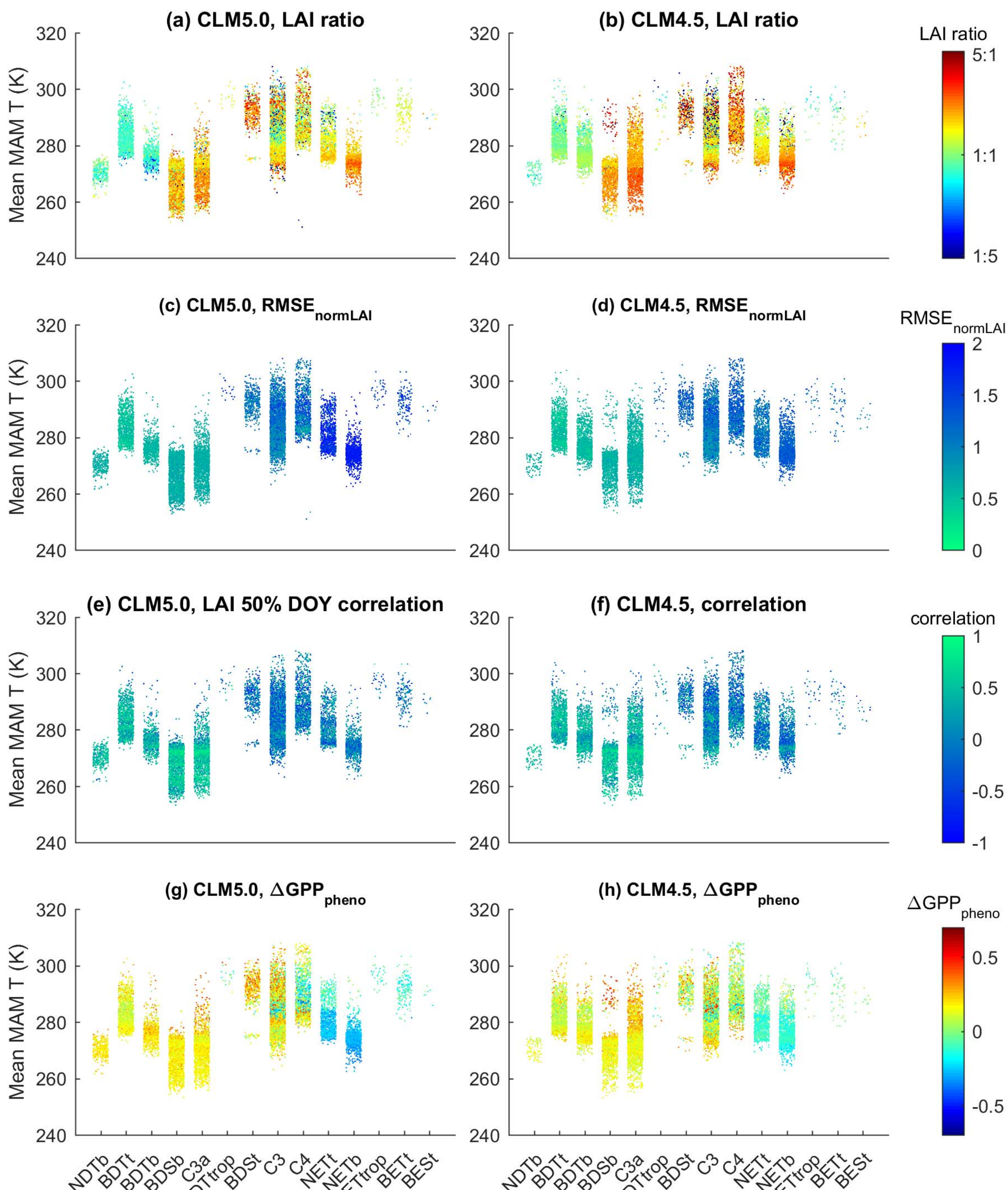

Figure 8. Relationships between each plant functional type (PFT; abbreviations on the X axis), mean spring soil temperature (Y axis), and different indices (different panels). Y axis shows mean March, April, and May 0-10cm soil temperature (mean MAM T) and X axis denotes CLM PFTs: NETt: needleleaf evergreen temperate tree; NETb: needleleaf evergreen boreal tree; NDTb: needleleaf deciduous boreal tree; BETtrop: broadleaf evergreen tropical tree; BETt: broadleaf evergreen temperate tree; BDTtrop: broadleaf deciduous tropical tree; BDTt: broadleaf deciduous temperate tree; BDTb: broadleaf deciduous boreal tree; BES: broadleaf evergreen

temperate shrub; BDSt: broadleaf deciduous temperate shrub; BDSb: broadleaf deciduous boreal shrub; C3a: C3 arctic grass; C3: C3 non-arctic grass; C4: C4 grass.

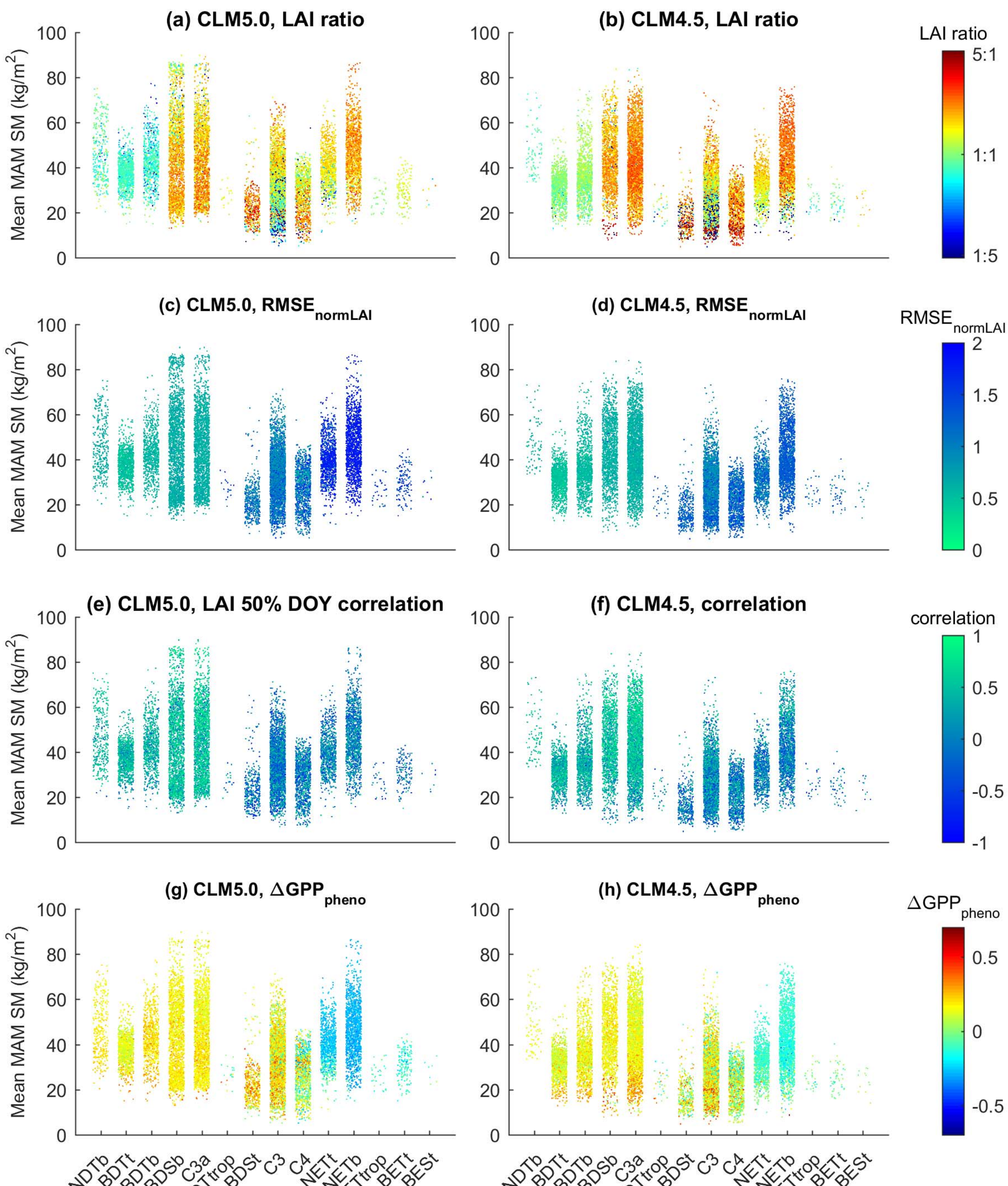

Figure 9. Relationships between each plant functional type (abbreviations on the X axis), mean spring soil moisture (Y axis), and different indices (different panels). Y axis shows mean March, April, and May 0-10cm soil moisture (mean MAM SM) and x axis denotes CLM PFTs. PFT names are the same as in Figure 8.

Within each PFT, LAI ratios and ΔGPP_{pheno} are most sensitive to environmental factors, while LAI seasonal cycles experience relatively small impacts of both soil temperature and moisture except for some stress deciduous PFTs (Figure 8-9). Across PFTs, LAI ratios mostly decrease with increasing spring temperature (Figure 9ab). $RMSE_{normLAI}$ and spring onset timing exhibit little changes when mean spring temperature changes except for stress deciduous where $RMSE_{normLAI}$ increases with higher mean spring temperature. Higher mean spring soil moisture also leads to higher LAI ratios and lower $RMSE_{normLAI}$, especially in stress deciduous PFTs (Figure 9). Overall, soil moisture has greater influences on the indicators in CLM4.5 than CLM5.0. Although lower soil temperature and higher soil moisture generally lead to better estimation of LAI seasonal cycle and start of spring timing, within each PFT, environmental factors have relatively small influences on the agreement of seasonality between CLM and MODIS (Figure 8-9).

4 Discussion

Here we investigated the relative importance of model version and forcing, phenology scheme, and soil moisture and temperature to four sets of indicators of leaf phenology and their influences on vegetation productivity. PFTs that employ the seasonal deciduous phenology scheme exhibit better agreement with MODIS LAI than PFTs using either of the other two routines. Because temperature is the dominant control on seasonal deciduous ecosystems, they respond similarly in both the model and the observational data, though the amplitudes of the LAI seasonal cycle are often too high for shrubs and grasslands at high latitudes. Evergreen PFTs show the lowest agreement between LAI seasonal cycles in CLM and MODIS and display a later start of growing season in CLM than MODIS, though comparisons at AmeriFlux sites show the discrepancy is partially due to biases in MODIS LAI estimates and how growing season is measured across different datasets. For stress deciduous PFTs, both LAI amplitudes and seasonal cycles agree better between CLM and MODIS at higher latitudes. Over the NH, CLM agrees best with MODIS in both LAI amplitudes and seasonal cycles over broadleaf deciduous trees in temperate and boreal regions. PFT and phenology schemes play a critical role in regulating agreements between CLM and MODIS leaf phenology, and their disagreements can cause large mismatches between leaf phenology and carbon production in CLM (Figure 10).

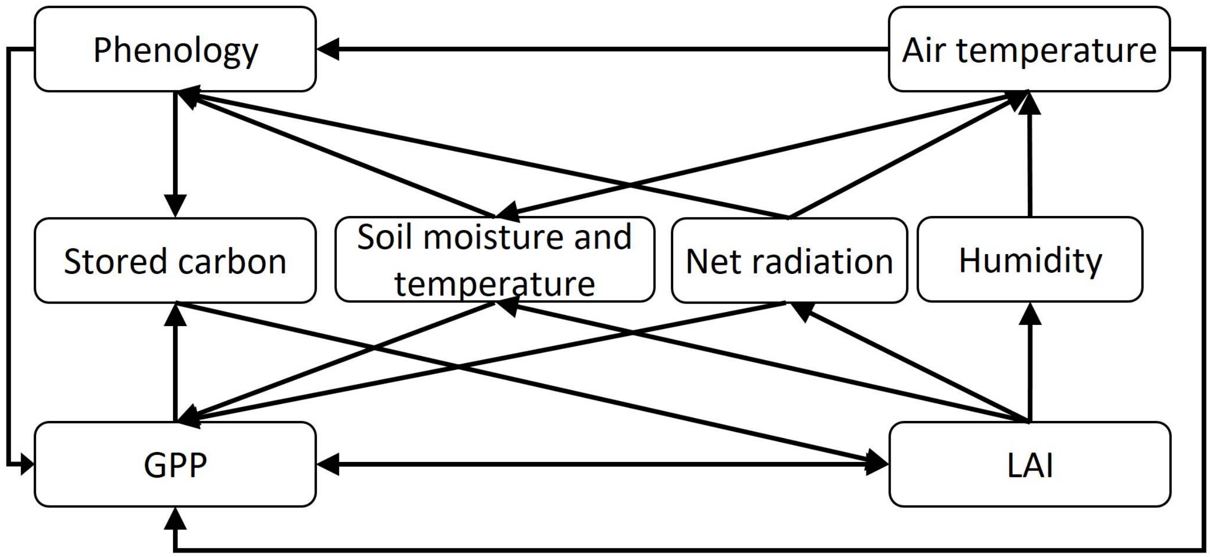

Figure 10. Conceptual diagram showing the relationships between plant phenology, GPP, and environmental factors. Phenology controls the onset and offset of growing season and productivity for deciduous PFTs and carbon and nitrogen fluxes from the stored carbon and nitrogen pools. PFTs and their phenology triggers are listed in Table S1.

Despite uncertainties in both satellite observations and CLM simulations, grid cell and Northern Hemispheric scale PFT-level comparison reveal fundamental differences in leaf development process in these two state-of-the-art large-scale products. LAI of evergreen PFTs in boreal and temperate regions peaks too late and may cause underestimations of GPP. Stress deciduous PFTs in moisture-limited temperate and tropical regions can have several growing seasons within one year, while usually only one or two growing periods are displayed in MODIS LAI. Agreement with MODIS also changes from CLM4.5 to CLM5.0. Over most PFTs, CLM5.0 improves simulations of LAI amplitudes, but the alignment of the LAI seasonal cycle between MODIS and CLM decreases from CLM4.5 to CLM5.0. Although site-level analysis and comparison can provide critical information in model development and validation (Dahlin et al., 2015; Kim et al., 2015; Stöckli et al., 2008), because vegetation within the same PFT in CLM shares the same set of parameters (Table S1), it is essential to evaluate model simulations against observations at both large spatial scales and plant functional type level.

Vegetation types determine the agreement between CLM and MODIS LAI, especially agreements on seasonal cycles. CLM LAIs are more likely to be larger than MODIS LAIs with lower soil temperature and higher soil moisture, however, once the phenology scheme is determined, agreements between the seasonal cycles exhibit only small influences of soil states, even for the more moisture-limited stress deciduous PFTs. This insensitivity may be due to the uniform soil water potential threshold across stress deciduous PFTs and relatively moist soil in CLM5.0 (Li et al., 2022). In addition, although adding in the 10-day total precipitation criteria helped limit the number of growing seasons each year in stress deciduous PFTs, depending on the forcing, there can still be more than one growing season present each year. Because large differences are present in soil moisture in different model-forcing combinations (Koster et al., 2009), one fixed soil water potential threshold may not fit all stress deciduous PFTs and forcing datasets. Therefore, the soil moisture threshold and/or the water stress period need to be adaptive to the meteorological variables as the temperature thresholds are.

Although the phenology scheme is little changed from CLM4.5 to CLM5.0, plant phenology displays large differences between the two model versions, suggesting phenology is greatly modified by other biogeophysical and biogeochemical processes in the model. Soil temperature plays an important role in regulating spring onset in CLM5.0 while both soil temperature and soil moisture influence phenology in CLM4.5. CLM5.0 simulates higher soil moisture with the same GSWP3 forcing and less influence of soil moisture on leaf phenology. Soil moisture simulations in LSMs are highly model-dependent and have critical influences on other processes (Koster et al., 2009), so changes in soil moisture states require new parameterization for moisture-sensitive PFTs. Meanwhile, although PFTs and phenology schemes are pre-determined by their environment in the models, leaf phenology in turn modulates momentum and gas exchange between the land and the atmosphere as well as temperature and humidity (Figure 10). Differences in LAI amplitudes and seasonal cycles also result in large biases in primary production and the carbon cycle (e.g. Birch et al., 2021; Li et al., 2022).

Decoupling between LAI and GPP seasonal cycles may cause biases in GPP estimation, and the biases in GPP may further feedback to LAI simulations and influence energy and water exchanges between the land and the atmosphere. Notably, as the influence of LAI disagreement on GPP estimates differs at the same grid cell for different PFTs, the biases can cancel out and simulate an overall smaller bias at the grid cell level (e.g. as for deciduous broadleaf trees and grasses in Eastern US), but the PFT-level biases may still result in discrepancies in simulations of the carbon cycle and land-atmosphere coupling. In addition, improving phenology in CLM may not simply be a matter of adopting new phenology parameterizations—even if these new parameterizations are more accurate in their representation of observed phenology. It is therefore important to develop new indicators and assess model performance at large scale and at PFT level. Furthermore, in a changing climate, the relative importance of temperature versus soil moisture in regulating plant growth in deciduous PFTs may change (e.g. Green et al., 2017; Denissen et al., 2022). Diagnosing the ecological dynamics of those changes in sensitivity will require researchers to archive a larger number of required LAI and land surface variables to compute the indices used here, which we argue are critical for identifying sources of model bias that are not typically seen using other standard diagnostic metrics.

5 Conclusions

Phenology modulates the flow of energy, moisture, and gasses through the biosphere and atmosphere. Here we compared PFT-specific LAI between CLM and MODIS and found that mismatches between CLM phenology and observations do not arise solely from how phenology is parameterized. Instead, those discrepancies emerge from a combination of biases in the land surface model and the details of phenology schemes employed by individual PFTs. Among all the phenology schemes, seasonal deciduous shows the best agreement. Among the PFTs, deciduous broadleaf trees have the best agreements in both LAI amplitudes and seasonal cycles. While PFT has a strong influence on leaf phenology, environmental factors influence the agreement both by determining the PFT that is present and by influencing LAI amplitudes. Spring onset timing and LAI seasonal cycle are more consistent across vegetation types at higher latitudes and earlier in the growing season, indicating that there may be a temperature-dominated signal in spring phenological changes in both CLM and observations. However, CLM displays large cross-PFT variation in LAI values, seasonal amplitude, and seasonal cycle, and they are influenced by both coexisting PFTs within the grid cell and the location. This information may be lost when aggregated to grid cell level averages, resulting in possibly good LAI and GPP simulations for the wrong reason. Therefore, it is critical to examine LAI variability and related fluxes at PFT-level over large spatial scales.

Despite large uncertainties in both MODIS and CLM LAIs, there are fundamental differences between leaf phenology in different versions of CLM and MODIS. While LAI seasonal cycles agree well over seasonal deciduous PFTs, evergreen and stress deciduous PFTs exhibit large differences between CLM and MODIS LAI amplitudes and seasonal variation, especially over boreal and temperate regions. Compared to CLM4.5, CLM5.0 improves over LAI values but exhibits less agreement with LAI seasonal cycles in MODIS, potentially causing a decoupling between LAI and GPP seasonal variations. Future improvements should bring GPP and phenology into better agreement because these two features of plant growth are intimately related. However, as currently implemented, improvements in one component will not necessarily translate into improvements in the other. Moreover, due to the wetter soil in CLM5.0, it may be critical to adapt soil moisture criteria in the phenology scheme to soil moisture estimates and increase phenology sensitivity to environmental factors. However, as currently implemented, improvements in one component will not necessarily translate into improvements in the other. Accordingly, future improvements to CLM (and other LSMs) stand to benefit most from efforts that simultaneously optimize for phenology alongside GPP, soil moisture, and other variables. Such efforts would not only improve the fidelity of the annual cycle of vegetation growth, but would also make model simulations of land-atmosphere interactions more accurate overall.

Acknowledgments

This work is supported by the National Science Foundation Macrosystems Biology award [grant numbers DEB-1702551, DEB-1702697, DEB-1702727, DEB-2017815] and the National Science Foundation Career Award [grant number AGS-1751535]. The authors would like to thank Dr. Eli Melaas and Dr. Adam M. Young for their helpful comments and feedback during the development of this project. We acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory. The authors declare no conflicts of interest with this publication.

Open Research

The MODIS LAI data is publicly available online through USGS website: https://lpdaac.usgs.gov/products/mod15a2hv006/. LAI and GPP simulations from the CESM experiments are available upon reasonable request from the corresponding author.

References

Albergel, C., Dutra, E., Munier, S., Calvet, J. C., Munoz-Sabater, J., Rosnay, P. D., & Balsamo, G. (2018). ERA-5 and ERA-Interim driven ISBA land surface model simulations: which one performs better?. *Hydrology and Earth System Sciences*, 22(6), 3515-3532.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., … Evans, R. (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, *82*(11), 2415–2434.

Berg, A., Findell, K., Lintner, B., Giannini, A., Seneviratne, S. I., Van Den Hurk, B., ... & Milly, P. C. (2016). Land–atmosphere feedbacks amplify aridity increase over land under global warming. *Nature Climate Change*, *6*(9), 869-874.

Birch, L., Schwalm, C. R., Natali, S., Lombardozzi, D., Keppel-Aleks, G., Watts, J., ... & Rogers, B. M. (2021). Addressing biases in Arctic–boreal carbon cycling in the Community Land Model Version 5. *Geoscientific Model Development*, *14*(6), 3361-3382.

Chen, M., Melaas, E. K., Gray, J. M., Friedl, M. A., & Richardson, A. D. (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), 3675–3688.

Dahlin, K. M., Fisher, R. A., & Lawrence, P. J. (2015). Environmental drivers of drought deciduous phenology in the Community Land Model. *Biogeosciences*, *12*(16), 5061–5074.

Dahlin, K. M., Ponte, D. Del, Setlock, E., & Nagelkirk, R. (2017). Global patterns of drought deciduous phenology in semi-arid and savanna-type ecosystems. *Ecography*, *40*(2), 314–323. https://doi.org/10.1111/ecog.02443

Dai, Y., Zeng, X., Dickinson, R. E., Baker, I., Bonan, G. B., Bosilovich, M. G., ... & Yang, Z. L. (2003). The common land model. *Bulletin of the American Meteorological Society*, 84(8), 1013-1024.

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., ... & Strand, W. G. (2020). The community earth system model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, 12(2).

Denissen, J., Teuling, A. J., Pitman, A. J., Koirala, S., Migliavacca, M., Li, W., ... & Orth, R. (2022). Widespread shift from ecosystem energy to water limitation with climate change. *Nature Climate Change*, *12*(7), 677-684.

Findell, K. L., Gentine, P., Lintner, B. R., & Guillod, B. P. (2015). Data length requirements for observational estimates of land–atmosphere coupling strength. *Journal of Hydrometeorology*, *16*(4), 1615-1635.

Fitzjarrald, D. R., Acevedo, O. C., & Moore, K. E. (2001). Climatic consequences of leaf presence in the eastern United States. *Journal of Climate*, *14*(4), 598–614.

Friedl, M. A., 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*(1), 168–182.

Green, J. K., Konings, A. G., Alemohammad, S. H., Berry, J., Entekhabi, D., Kolassa, J., ... & Gentine, P. (2017). Regionally strong feedbacks between the atmosphere and terrestrial biosphere. *Nature geoscience*, *10*(6), 410-414.

Guillevic, P., Koster, R. D., Suarez, M. J., Bounoua, L., Collatz, G. J., Los, S. O., & Mahanama, S. P. P. (2002). Influence of the interannual variability of vegetation on the surface energy balance—A global sensitivity study. *Journal of Hydrometeorology*, *3*(6), 617-629.

Hollinger, D., (2021), AmeriFlux BASE US-Ho2 Howland Forest (west tower), Ver. 4-5, AmeriFlux AMP, (Dataset). https://doi.org/10.17190/AMF/1246062. [Accessed 2022/11/10]

Kim, Y., Moorcroft, P. R., Aleinov, I., Puma, M. J., & Kiang, N. Y. (2015). Variability of phenology and fluxes of water and carbon with observed and simulated soil moisture in the Ent Terrestrial Biosphere Model (Ent TBM version 1.0.1.0.0). *Geoscientific Model Development*, *8*(12), 3837–3865. https://doi.org/10.5194/gmd-8-3837-2015

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., ... & Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, *19*(1).

Koster, R. D., Guo, Z., Yang, R., Dirmeyer, P. A., Mitchell, K., & Puma, M. J. (2009). On the nature of soil moisture in land surface models. *Journal of Climate*, *22*(16), 4322-4335.

Law, B., (2022). AmeriFlux BASE US-Me2 Metolius mature ponderosa pine, Ver. 18-5, AmeriFlux AMP, (Dataset). https://doi.org/10.17190/AMF/1246076. [Accessed 2022/11/10]

Lawrence, P. J., & Chase, T. N. (2010). Investigating the climate impacts of global land cover change in the community climate system model. *International Journal of Climatology*, *30*(13), 2066–2087.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., and Kluzek, E. (2019). The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems*.

Levis, S., & Bonan, G. B. (2004). Simulating springtime temperature patterns in the community atmosphere model coupled to the community land model using prognostic leaf area. *Journal of Climate*, *17*(23), 4531-4540.

Li, X., Ault, T., Richardson, A. D., Carrillo, C. M., Lawrence, D. M., Lombardozzi, D., ... & Moon, M. (2023). Impacts of shifting phenology on boundary layer dynamics in North America in the CESM. *Agricultural and Forest Meteorology*, *330*, 109286.

Li, X., Melaas, E., Carrillo, C. M., Ault, T., Richardson, A. D., Lawrence, P., ... & Young, A. M. (2022). A comparison of land surface phenology in the Northern Hemisphere derived from satellite remote sensing and the Community Land Model. *Journal of Hydrometeorology*, *23*(6), 859-873.

Lorenz, R., Davin, E. L., Lawrence, D. M., Stöckli, R., & Seneviratne, S. I. (2013). How important is vegetation phenology for European climate and heat waves?. *Journal of Climate*, *26*(24), 10077-10100.

Lovato, T., Peano, D., Butenschön, M., Materia, S., Iovino, D., Scoccimarro, E., ... & Navarra, A. (2022). CMIP6 Simulations With the CMCC Earth System Model (CMCC-ESM2). *Journal of Advances in Modeling Earth Systems*, *14*(3), e2021MS002814.

Mahowald, N., Lo, F., Zheng, Y., Harrison, L., Funk, C., Lombardozzi, D., & Goodale, C. (2016). Projections of leaf area index in earth system models. *Earth System Dynamics*, 7(1), 211–229.

Morisette, J. T., Richardson, A. D., Knapp, A. K., Fisher, J. I., Graham, E. A., Abatzoglou, J., ... & Liang, L. (2009). Tracking the rhythm of the seasons in the face of global change: phenological research in the 21st century. *Frontiers in Ecology and the Environment*, *7*(5), 253-260.

Myneni, R., Knyazikhin, Y., & Park, T. (2015). MOD15A2H MODIS/terra leaf area index/FPAR 8-day L4 global 500 m SIN grid V006. *NASA EOSDIS Land Processes DAAC*.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Charles, D., … Sacks, W. (2013). CLM 4.5 NCAR Technical Note, (July). https://doi.org/10.1007/s11538-011-9690-0

Park, H., & Jeong, S. (2021). Leaf area index in Earth system models: how the key variable of vegetation seasonality works in climate projections. *Environmental Research Letters*, *16*(3), 034027.

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y. W., ... & Law, B. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific data*, *7*(1), 1-27.

Peano, D., Materia, S., Collalti, A., Alessandri, A., Anav, A., Bombelli, A., & Gualdi, S. (2019). Global variability of simulated and observed vegetation growing season. *Journal of Geophysical Research: Biogeosciences*, *124*(11), 3569-3587.

Puma, M. J., Koster, R. D., & Cook, B. I. (2013). Phenological versus meteorological controls on land-atmosphere water and carbon fluxes. *Journal of Geophysical Research: Biogeosciences*, *118*(1), 14-29.

Renner, S. S., & Zohner, C. M. (2018). Climate change and phenological mismatch in trophic interactions among plants, insects, and vertebrates. *Annual review of ecology, evolution, and systematics*, *49*(1), 165-182.

Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G., … Xue, Y. (2012). Terrestrial biosphere models need better representation of vegetation phenology: Results from the North American Carbon Program Site Synthesis. *Global Change Biology*, 18(2), 566–584.

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., & Toomey, M. (2013). Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agricultural and Forest Meteorology*, *169*, 156–173. https://doi.org/10.1016/j.agrformet.2012.09.012

Richardson, A. D., Hufkens, K., Milliman, T., Aubrecht, D. M., Chen, M., Gray, J. M., … Kosmala, M. (2018). Tracking vegetation phenology across diverse North American biomes using PhenoCam imagery. *Scientific Data*, *5*, 180028.

Scholze, M., Buchwitz, M., Dorigo, W., Guanter, L., & Quegan, S. (2017). Reviews and syntheses: Systematic Earth observations for use in terrestrial carbon cycle data assimilation systems. *Biogeosciences*, *14*(14), 3401–3429. https://doi.org/10.5194/bg-14-3401-2017

Schwartz, M. D. (1992). Phenology and springtime surface-layer change. *Monthly weather review*, *120*(11), 2570-2578.

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., ... & Schulz, M. (2020). Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geoscientific Model Development*, *13*(12), 6165-6200.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., ... & Venevsky, S. (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global change biology*, *9*(2), 161-185.

Song, X., Wang, D. Y., Li, F., & Zeng, X. D. (2021). Evaluating the performance of CMIP6 Earth system models in simulating global vegetation structure and distribution. *Advances in Climate Change Research*, *12*(4), 584-595.

Steiner, A. L., Pal, J. S., Rauscher, S. A., Bell, J. L., Diffenbaugh, N. S., Boone, A., ... & Giorgi, F. (2009). Land surface coupling in regional climate simulations of the West African monsoon. *Climate Dynamics*, *33*, 869-892.

Stöckli, R., Lawrence, D. M., Niu, G., Oleson, K. W., Thornton, P. E., Yang, Z., … Running, S. W. (2008). Use of FLUXNET in the Community Land Model development. *Journal of Geophysical Research: Biogeosciences*, *113*(G1).

White, M. A., De Beurs, K. M., Didan, K., Inouye, D. W., Richardson, A. D., Jensen, O. P., … Lauenroth, W. K. (2009). Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. *Global Change Biology*, *15*(10), 2335–2359. https://doi.org/10.1111/j.1365-2486.2009.01910.x

Xu, X., Riley, W. J., Koven, C. D., Jia, G., & Zhang, X. (2020). Earlier leaf-out warms air in the north. *Nature Climate Change*, *10*(4), 370-375.

780 **Table 1.** Mean and one standard deviation of the three indicators of each PFT

PFT		LAI ratio (close to 1 means better estimation)		$RMSE_{normLAI}$ (lower means better seasonality)		Correlation (higher means better estimation)		ΔGPP_{pheno} (closer to 0 means smaller GPP bias)	
		clm5	clm4.5	clm5	clm4.5	clm5	clm4.5	clm5	clm4.5
seasonal deciduous	needleleaf deciduous boreal tree	0.84±0.41	0.75±0.17	0.65±0.04	0.64±0.05	0.51±0.24	0.57±0.17	0.30±0.03	0.22±0.02
	broadleaf deciduous temperate tree	0.70±0.23	1.19±0.46	0.50±0.10	0.50±0.10	0.46±0.24	0.46±0.26	0.24±0.09	0.22±0.10
	broadleaf deciduous boreal tree	0.62±0.33	1.29±0.43	0.60±0.07	0.55±0.10	0.43±0.28	0.42±0.28	0.32±0.06	0.25±0.08
	broadleaf deciduous boreal shrub	3.93±1.97	4.93±3.07	0.66±0.09	0.64±0.11	0.53±0.29	0.51±0.30	0.30±0.05	0.30±0.09
	c3 arctic grass	4.17±2.15	5.83±2.27	0.64±0.08	0.62±0.10	0.54±0.30	0.49±0.30	0.30±0.08	0.30±0.10
stress deciduous	broadleaf deciduous tropical tree	2.27±0.49	1.67±1.43	1.54±0.18	1.26±0.20	0.10±0.43	0.20±0.36	0.16±0.27	0.02±0.25
	broadleaf deciduous temperate shrub	6.98±5.43	8.61±11.93	1.12±0.21	1.25±0.19	-0.07±0.39	-0.02±0.37	0.35±0.15	0.18±0.17
	c3 non-arctic grass	3.17±2.98	3.43±4.26	1.13±0.26	1.11±0.25	0.01±0.37	0.11±0.38	0.23±0.17	0.17±0.17
	c4 grass	3.33±3.23	8.67±8.48	1.17±0.34	1.27±0.22	0.02±0.38	-0.00±0.36	0.06±0.19	0.05±0.22
evergreen	needleleaf evergreen temperate tree	2.46±1.12	2.69±1.26	1.66±0.27	1.09±0.23	0.06±0.34	0.06±0.34	-0.25±0.08	-0.07±0.08
	needleleaf evergreen boreal tree	4.54±1.76	5.29±2.71	1.81±0.08	1.28±0.17	0.07±0.32	0.09±0.35	-0.36±0.05	-0.13±0.06
	broadleaf evergreen tropical tree	1.18±0.19	0.79±0.30	1.53±0.24	1.23±0.25	-0.17±0.41	-0.15±0.40	-0.11±0.10	-0.06±0.11
	broadleaf evergreen temperate tree	1.99±0.48	0.93±0.43	1.51±0.22	1.20±0.27	-0.07±0.35	0.01±0.40	-0.14±0.09	-0.06±0.09
	broadleaf evergreen shrub	2.96±2.11	3.50±0.96	1.42±0.47	0.89±0.22	-0.11±0.30	-0.17±0.42	-0.08±0.10	0.06±0.05

781
782