A Semiprognostic Phenology Model for Simulating Multidecadal Dynamics of Global Vegetation Leaf Area Index

Qinchuan Xin1,2, Xuewen Zhou3, Nan Wei1, Hua Yuan3, Zurui Ao1, and Yongjiu Dai3

1School of Planning and Geography, Sun Yat-sen University, Guangzhou, China, 2State Key Laboratory of Desert and Oasis Ecology, Research Center for Ecology and Environment of Central Asia, Chinese Academy of Sciences, Urumqi, China, 3School of Atmospheric Sciences, Sun Yat-sen University, Guangzhou, China

Abstract  Vegetation leaf phenology, often reflected by the dynamics in leaf area index (LAI), influences a variety of land surface processes. Robust models of vegetation phenology are pivot components in both land surface models and dynamic global vegetation models but remain challenging in terms of the model accuracy. This study develops a semiprognostic phenology model that is suitable for simulating time series of vegetation LAI. This method establishes a linear relationship between the steady-state LAI (i.e., the LAI when the environment conditions remain unchanging) and gross primary productivity, meaning that the LAI an unchanged environment can carry is proportional to the photosynthetic products produced by plant leaves and implements with a simple light use efficiency algorithm of MOD17 to form a closed set of equations. We derive an analytical solution based on the Lambert W function to the closed equations and then apply a simple restricted growth process model to simulate the time series of actual LAI. The results modeled using global climate data demonstrate that the model is able to capture both the spatial pattern and intra-annual and interannual variation of LAI derived from the satellite-based product on a global scale. The results modeled using the flux tower data suggest that the developed model is able to explain over 70% variation in daily LAI for each plant functional type except evergreen broadleaf forest. The developed semiprognostic approach provides a simple solution to simulating the spatiotemporal variation in vegetation LAI across plant functional types on the global scale.

Plain Language Summary  Modeling vegetation leaf phenology remains challenging in terms of the model accuracy. We develop a semiprognostic vegetation phenology method that is suitable for simulating the variation in the time series of leaf area index. This approach establishes a linear relationship between LAI and vegetation productivity and implements with a simple light use efficiency algorithm of MOD17. The results modeled using global climate data demonstrate that the model is able to capture both the spatial pattern and intra-annual and interannual variation of LAI derived from the remote sensing data on the global scale.

1. Introduction

Vegetation phenology denotes the cyclic and seasonal natural phenomena related to plant life. The change in plant leaves is one of the most apparent activities in vegetation phenology. The amount of plant leaves is often quantified by leaf area index (LAI), which is defined as total plant leaf area per unit land area (Richardson et al., 2009). Accompanying the periodic leaf cycles of terrestrial plants, seasonal variation in LAI is significant vegetation dynamics on global land surface (Jiang et al., 2017; Lee et al., 2017; Verger et al., 2016). Vegetation phenology is often characterized by key phenophases of leaves, such as the start of the season (SOS) and the end of the season (EOS) (Qiao et al., 2019). Vegetation leaf phenology controls biogeochemical processes such as photosynthesis and evapotranspiration and largely influences global carbon and water cycles (M. Chen, Melaas, et al., 2016; Keenan et al., 2014; Wolf et al., 2016). Observing and modeling vegetation leaf phenology over large areas provide valuable information to our understanding of the land surface processes.

As vegetation leaf phenology is sensitive to climate variation, understanding long-term variation of global vegetation phenology including dynamics in LAI has become a research frontier in the field of climate change (Broich et al., 2014; Piao et al., 2019). Tremendous efforts have been made to quantify and analyze
vegetation leaf phenology changes in the past decades using both in situ and satellite observational data, leading to a number of valuable scientific discoveries (Liu et al., 2016; Parmesan, 2007; Richardson et al., 2006). For example, studies on satellite remote sensing data have shown that global green leaf area has continuously increased since the new century, whereas increases in green leaf area in China and India account for nearly one third of global increases (Chen et al., 2019). Plant phenology responses to climate variation are nonlinear and differ across species (Menzel et al., 2006). In a few key regions, the results derived from observational phenology data remain controversial. A study based on the remote sensing data have found delayed spring vegetation leaf phenology in the Qinghai-Tibet Plateau due to winter and spring warming (Shen et al., 2014; Yu et al., 2010), and another argued that spring phenology in the Qinghai-Tibet Plateau continuously advanced in the past few decades (Zhang et al., 2013). Both the data sources and the processing methods of satellite remote sensing data could affect the derived vegetation leaf phenology and subsequently its changing trend. Despite of the issues such as data contamination and spatial scaling in satellite observations (Zhang et al., 2017), LAI time series derived from remote sensing data have become one critical data source for studying the terrestrial ecosystems and the land surface processes.

Comparing to numerous studies that monitor and quantify vegetation phenological changes using long-term observational data, our abilities on modeling vegetation phenology are largely limited to date (Piao et al., 2019). There are large uncertainties associated with vegetation phenology modeling in the current terrestrial biosphere models (Chen, Rao, Shen, et al., 2016). Even for the extensively studied deciduous forests, it has found that almost all 14 models participated in the North American Carbon Program Site Synthesis have misrepresented phenology schemes and make predictions with considerable errors on the timing of key phenophases (Richardson et al., 2012).

The current terrestrial biosphere models often adopt the methods of the prescribed phenology (i.e., using static LAI derived from satellite data) and the prognostic phenology (i.e., predicting LAI based on climatic variables) or the hybrid method of the semiprognostic and semiprescribed phenology (Hu et al., 2018; Lawrence et al., 2011). Although applying satellite-derived LAI in the land surface model is suitable for simulating vegetation-related processes over the past few decades, the prescribed phenology method is unable to make prediction on future responses of phenology to climate change (Zhu et al., 2018). Among the prognostic phenology models, the growing-degree day (GDD) model and its derivative models are the most common phenomenological simulation schemes, of which the basic principle is to predict the timing of key phenophases by accumulating daily air temperature starting from a certain day until reaching a specific heat forcing threshold that is required for vegetation growth (Chuine et al., 1999). Based on the early GDD model that only considers heating temperature accumulation, studies have developed phenology models that account for the other climate variables, such as photoperiod, vapor pressure deficit, and chilling temperature (Hufkens et al., 2018; Melaas et al., 2013). White et al. (1997) developed vegetation phenology models for deciduous forests and grasslands using Advanced Very High Resolution Radiometer (AVHRR) satellite remote sensing data. The developed phenology simulation scheme that is now embedded in both Biome-BGC and the Community Land Model consists of several limiting equations to account for the effects such as accumulated soil temperature and soil moisture on vegetation growth (Oleson et al., 2013; White et al., 2000) and, therefore, is able to simulate the phenology processes of both nonstressed and water-stressed vegetation. As the Community Land Model accounts for as much as 17 plant functional types in modeling the land surface processes but only three vegetation types (i.e., evergreen, deciduous, and stressed-deciduous vegetation) in modeling phenology (Oleson et al., 2013), the phenology module requires substantial improvement. Recent studies attempted to use both remote sensing and large-scale meteorological data to calibrate and evaluate multiple GDD-based phenology models on modeling the timing of key phenophases across vegetation types (Liu et al., 2018; Xin et al., 2015).

In addition to numerical models that simulate the timing of key phenophases via climate variables, studies also attempted to simulate time series of vegetation LAI throughout the growing season in a more direct way. The growing season index (GSI) model proposed by Jolly et al. (2005) considered the limitation of daily temperature, vapor pressure deficit, and photoperiod on vegetation growth and was applied to simulate LAI time series (Savoy & Mackay, 2015). LPJ-Ml, a dynamic global vegetation model (Schaphoff et al., 2018), adopts the modified GSI approach for phenology modeling and accounts for the limiting effects of cold temperature, light, water availability, and heat stress on the daily phenology status (Forkel et al., 2014). Xin et al. (2018) proposed a steady-state approximation method for simulating seasonal dynamics of LAI. The method...
iteratively solved the developed closed equations between steady-state LAI and vegetation productivity and subsequently modeled time series of vegetation LAI at the daily scale by applying the simple moving average method to the obtained time series of steady-state LAI. A crop model, the De-Nitrification De-Composition model, defines an optimal curve of plant LAI during the growing season and then calculates daily LAI time series under environmental stress conditions by accounting for the scalar factors of water, carbon, and nitrogen stresses (Chen, Yu, Li, et al., 2016).

While robust models of vegetation phenology are indispensable components in both the land surface models and dynamic global vegetation models, the current phenological model is largely empirical and relies on the relationship or thresholding established by statistical regression methods (Hu et al., 2017; Polgar & Primack, 2011). There is still a lack of deep understanding on the physiological mechanisms of plants in response to climate variation at the current stage. The phenology models often make predictions with large errors and biases across biomes when compared with ground observations. While it requires full-scale and simultaneous measurements of both meteorological and phenological variables to support the development of the phenology models, ground phenological observation networks have only limited numbers of both observational sites and records. Although the remote sensing technique provides continuous observation of LAI dynamics on the land surface, it remains a challenging task to match the modeled phenophases on the continental and global scales with satellite data, because satellite-derived phenophases are dependent on the data quality and the inversion algorithms. In essence, developing prognostic or semiprognostic phenology models, if the modeling results could well match with ground and remote sensing observations, could make a valuable contribution to our understanding of terrestrial vegetation dynamics and the vegetation-climate feedback.

The objectives of this study are to (1) develop a semiprognostic phenology model for simulating multidecadal variation in the LAI time series, (2) conduct parameterization and assessment of the developed phenology model across plant functional types using both field and satellite data, and (3) analyze the modeled results across time and space on a global scale.

2. Methods

Vegetation phenology is closely related to the ability of vegetation photosynthesis as reflected by gross primary productivity (GPP). On the one hand, vegetation GPP is nonlinearly dependent on LAI under given meteorological conditions because plant leaves are the basic elements for photosynthesis. On the other hand, vegetation GPP that reflects how the meteorological conditions are suitable for plant photosynthetic activities determines the process that plant allocates biomasses to leaves, because the strategy of plant leaf phenology is to gain competitive advantages in photosynthesis. Under a continuously changing environment, leaf photosynthesis responds to micrometeorological variation within minutes, and the response of leaf phenology to climate variation takes up from days to months (Sellers et al., 1996). Xin et al. (2018) proposed the concept of steady-state LAI (i.e., the canopy LAI that reaches an equilibrium state under an unchanging environment) and established closed equations (i.e., the number of unknown variables is the same as the number of independent equations) for both steady-state LAI and GPP. An approximate solution to the closed equations can be obtained through iterations, and the details regarding to the model and its solution can be found in Xin et al. (2018). Our previous approach has proven successful in modeling the phenology of deciduous broadleaf forest (DBF), but it implemented with a sophisticated canopy photosynthesis model, making it difficult in parameterization and computationally expensive in global simulations.

Here we present a straightforward solution that implements with a simple vegetation GPP model (i.e., the light use efficiency model of MOD17, which is used to produce global MODIS GPP products routinely) to model LAI time series (Zhao & Running, 2010). Given daily meteorological conditions, we modeled the steady-state LAI at the corresponding day by solving the closed equations as follows:

\[ \text{LAI}_s = m \times \text{GPP}_s, \]

\[ \text{GPP}_s = \varepsilon_{\text{max}} \times \text{PAR} \times f(TMIN) \times f(VPD) \times (1 - \exp(-k\text{LAI}_s)), \]

where \( \text{LAI}_s \) denotes steady-state LAI; \( m \) denotes the ratio of LAI to GPP under the canopy closure condition; \( \text{GPP}_s \) denote GPP that is corresponding to steady-state LAI; \( \varepsilon_{\text{max}} \) denotes maximum light use efficiency; \( \text{PAR} \) denotes photosynthetically active radiation; \( f(TMIN) \) and \( f(VPD) \) denote the scalar functions.
that account for the limiting impacts of daily minimum temperature and vapor pressure deficit, respectively, on canopy photosynthesis; and $k = 0.5$ denotes canopy light extinction coefficient.

Equation 1 represents that the LAI an unchanging environment can carry is proportional to the photosynthetic products produced by plant leaves, and Equation 2 follows the light use efficiency model of MOD17 (Zhao & Running, 2010). According to MOD17, the scalar functions of $f(TMIN)$ and $f(VPD)$ are calculated by the following equations, respectively,

$$f(TMIN) = \begin{cases} 0 & TMIN \leq TMIN_{min} \\ \frac{TMIN - TMIN_{min}}{TMIN_{max} - TMIN_{min}} & TMIN_{min} < TMIN < TMIN_{max} \\ 1 & TMIN \geq TMIN_{max} \end{cases}$$

(3)

$$f(VPD) = \begin{cases} 0 & VPD \geq VPD_{max} \\ \frac{VPD_{max} - VPD}{VPD_{max} - VPD_{min}} & VPD_{min} < VPD < VPD_{max} \\ 1 & VPD \leq VPD_{min} \end{cases}$$

(4)

where $TMIN$ denotes daily minimum temperature; $VPD$ denotes daily vapor pressure deficit; $TMIN_{max}$ and $TMIN_{min}$ denote the upper and lower boundaries of daily minimum temperature that limit canopy photosynthesis, respectively; and $VPD_{max}$ and $VPD_{min}$ denote the upper and lower boundaries of daily vapor pressure deficit that limit canopy photosynthesis, respectively.

Given that Equations 1 and 2 are closed equations, we can adopt the Newton-Raphson method to obtain an approximate solution. Alternatively, because the MOD17 model as represented by Equations 2, 3, and 4 has a simple form, it is possible to obtain an analytical solution using the following approach.

Let $\mu$ be an intermediate variable as follows:

$$\mu = m \times \varepsilon_{max} \times PAR \times f(TMIN) \times f(VPD).$$

(5)

Substitute $\mu$ into Equations 1 and 2, and we can derive

$$LAI_s = \mu(1 - \exp(-kLAI_s)).$$

(6)

Let $y$ be another intermediate variable as follows:

$$y = k\mu - kLAI_s.$$  

(7)

Substituting $y$ into Equation 6 gives the equation

$$y\exp(-y) = k \mu \exp(-k\mu).$$

(8)

Equation 8 can be solved using the Lambert W function directly as follows:

$$y = W_0(-k \mu \exp(-k\mu)).$$

(9)

where $W_0$ denotes the Lambert W function operator.

The Lambert W function is defined as the inverse relation of the function $f(z) = ze^z$, where $e^z$ is an exponential function and $z$ is a complex number (Corless et al., 1996). The value of $y$ can be derived from the Lambert W function once the values of both $k$ and $\mu$ are known. Note that $-k\mu \exp(-k\mu)$ has the range of $(-e^{-1}, 0]$ because $k\mu \geq 0$. Solutions to the Lambert W function have already been embedded in some commercial software such as Matlab.

Substituting Equation 9 into Equation 7 yields the analytical solution as follows:

$$LAI_s = \mu + \frac{1}{k}W_0(-k \mu \exp(-k\mu)).$$

(10)

where the solved steady-state LAI must satisfy the condition $LAI_s \leq LAI_{max}$, and $LAI_{max}$ denotes the maximum canopy LAI during a growing season. $LAI_{max}$ could be obtained from the remote sensing data on a pixel-by-pixel basis.
Due to dramatic fluctuations in meteorological conditions at the daily scale, the steady-state LAI obtained by Equation 10 undergoes large variation from day to day. To account for the lagging effect of vegetation phenology processes on changing meteorological conditions, we can use a simple restricted growth process model to simulate time series of actual LAI as follows:

\[ \frac{dLAI}{dt} = k_{\text{leaf}}(LAI_s - LAI), \]  

where \( LAI \) denotes leaf area index; \( dt \) denotes the unit time, and \( k_{\text{leaf}} \) denotes a time constant that represents the lagged responses of plant in allocating biomasses to leaves to climate variation.

3. Data and Model Calibration

3.1. Flux Tower Data (Model Input Data)

The eddy covariance flux tower is a core data source for observing and monitoring of the surface energy balance and the land-atmosphere interaction. A global flux tower data set, the FLUXNET 2015 Tier One data set as obtained online (http://fluxnet.fluxdata.org/), was used to develop and evaluate the phenology model at the site scale. The used data were acquired at 133 flux towers. We used the site-year data spanning the time period from 2000 to 2014 (1,171 site-year data in total), including 38 evergreen needleleaf forest (ENF) sites, 12 evergreen broadleaf forest (EBF) site, 19 DBF sites, 8 mixed forest (MIF) sites, 2 closed shrubland (CSH) site, 10 open shrubland (OSH) sites, 6 woody savanna (WSA) sites, 8 savanna (SAV) sites, and 30 grassland (GRA) sites (Table S1). The raw eddy covariance data have been processed in terms of gap filling, data screening, and quality checking, whereas the processing details can be found in the user manual at the official website. The site-scale simulation studies require environmental variables such as incident solar radiation, minimum temperature, and vapor pressure deficit at the daily scale as input data to the models. As the provided daily data do not contain daily minimum temperature, the hourly and half-hourly temperature data were processed to obtain daily minimum temperature.

3.2. Global Climate Data (Model Input Data)

Data from the Global Land Data Assimilation System (GLDAS) were downloaded from the Goddard Earth Sciences Data and Information Services Center (ftp://hydro1.sci.gsfc.nasa.gov/) and used to model dynamic LAI time series on a global scale. The used data spanned from 1982 to 2016 with a spatial resolution of 0.25° and a temporal resolution of 3 hr (Rodell et al., 2003). Modeling global vegetation LAI requires land surface meteorological variables such as temperature, vapor pressure deficit, and incident short-wave radiation on a daily basis. The extracted climate data were processed from 0.25° to (1/12)° spatial resolution using the bilinear resampling method via Geospatial Data Abstraction Library (GDAL) and were processed from 3 hr to daily time scale.

3.3. Satellite LAI Data (Model Evaluation Data)

As the flux tower data do not include observational LAI data, we used LAI time series derived from Version 6 Moderate Resolution Imaging Spectroradiometer (MODIS) LAI and fraction of photosynthetically active radiation product (MOD15A2H). MOD15A2H is publically available from the website of the Land Processes Distributed Active Archive Center (https://lpdaac.usgs.gov/) and provides 8-day LAI data at 500-m resolution (Myneni et al., 2002). The 8-day LAI time series were extracted from MOD15A2H for each flux tower site. The extracted LAI time series were screened using the quality control data in MOD15A2H for cloud contamination, and the cloudy data were replaced using a 3-point median filtering method. A Hamel filter (Pearson et al., 2016) was used to eliminate the outliers, and the missing values were gap-filled based on an autoregressive modeling method. The obtained 8-day LAI time series were then smoothed using the Savitzky-Golay filter and were further linearly interpolated to generate time series at the daily scale.

The global-scale study used the LAI fourth-generation product (Zhu et al., 2013) produced by Boston University (hereinafter referred to as BULAI4g). The BULAI4g product spanned from 1982 to 2016 has a spatial resolution of (1/12)° and a temporal resolution of half a month. BULAI4g is a remote sensing data set generated using a feed-forward neural network algorithm based on the AVHRR data and ground observational LAI data. We linearly interpolated the half-month BULAI4g data to generate daily global LAI time series by accounting for different days in each individual month.
3.4. Simulated Historical LAI Data in Climate Model Intercomparison Project Phase 6 (Model Evaluation Data)

We downloaded the LAI data set modeled by EC-Earth3-Veg in the AMIP experiment from World Climate Research Programme (WCRP) Climate Model Intercomparison Project Phase 6 (CMIP6) data archive (https://esgf-node.llnl.gov/search/cmip6/) for comparison. The data set (hereinafter referred to as EC-Earth3-Veg) covering the time period from 1982 to 2016 has a 100-km spatial resolution and a monthly temporal resolution. The raw data were resampled to match the spatial resolution of BULAI4g using a bilinear resampling approach and were processed to daily time scale using linearly interpolation.

3.5. Model Calibration

We calibrated the model parameters for different plant functional types using 10% flux tower data. We used the boundaries of the parameters in the MOD17 model to constrain the parameter calibration. We used an automated approach, namely, the Shuffled Complex Evolution, to speed up the model calibration process without exploring the entire parameter spaces (Duan et al., 1992). The approach is based on a synthesis of four concepts that have proved successful for global optimization, including combination of probabilistic and deterministic approaches, clustering, systematic evolution of a complex of points spanning the space, and competitive evolution (Duan et al., 1993). We determined the optimal parameters by minimizing the root mean squared error (RMSE) between modeled and observed data. The specific parameter settings after model calibration are shown in Table 1.

The required LAI\textsubscript{max} for one individual year are extracted for the flux tower sites and for each individual pixel from time series of the remote sensing data such as MOD15A2H and BULAI4g, respectively. Based on the above modeling approach, we are able to simulate time series of vegetation LAI for different plant functional types using climate variables. The LAI time series derived from the remote sensing data were used for model assessment. We use three common evaluation metrics, namely, the Pearson correlation coefficient (r), RMSE, and mean bias error (Bias) to assess the model performance. A more accurate model generally achieves higher correlation coefficient and lower RMSE and Bias errors when compared with observational data.

### Table 1

| PFT   | \( \varepsilon_{\text{max}} \) (gC/MJ) | TMIN\textsubscript{min} (°C) | TMIN\textsubscript{max} (°C) | VPD\textsubscript{min} (kPa) | VPD\textsubscript{max} (kPa) | m     | \( k_{\text{leaf}} \) (s\textsuperscript{-1}) |
|-------|----------------------------------|-----------------------------|-----------------------------|----------------------------|----------------------------|-------|-----------------------------|
| ENF   | 1.291                           | -10.00                      | 2.77                        | 0.50                       | 2.98                       | 0.645  | 0.0360                      |
| EBF   | 1.232                           | -10.00                      | 7.25                        | 0.50                       | 2.80                       | 0.973  | 0.0060                      |
| DNF   | 1.692                           | -10.00                      | 6.32                        | 0.50                       | 2.75                       | 0.399  | 0.0466                      |
| DBF   | 1.391                           | -6.93                       | 13.68                       | 0.50                       | 3.06                       | 0.622  | 0.0692                      |
| MIF   | 1.163                           | -8.39                       | 5.17                        | 0.50                       | 6.00                       | 0.559  | 0.0559                      |
| CSH   | 0.855                           | -9.03                       | 0.00                        | 0.50                       | 2.06                       | 0.998  | 0.0352                      |
| OSH   | 1.088                           | -10.00                      | 0.91                        | 0.50                       | 2.80                       | 0.431  | 0.0448                      |
| WSA   | 1.644                           | -10.00                      | 28.66                       | 0.50                       | 3.28                       | 0.371  | 0.0695                      |
| SAV   | 1.662                           | 3.90                        | 25.40                       | 0.68                       | 3.93                       | 0.414  | 0.0264                      |
| GRA   | 1.363                           | -9.67                       | 0.12                        | 0.50                       | 2.52                       | 0.609  | 0.0403                      |

Note. PFT = plant functional type; ENF = evergreen needleleaf forest; EBF = evergreen broadleaf forest; DNF = deciduous needleleaf forest; DBF = deciduous broadleaf forest; MF = mixed forest; CSH = closed shrubland; OSH = open shrubland; WSA = woody savanna; SAV = savanna; GRA = grassland.

### 4. Results

#### 4.1. Site-Scale Assessment

On the site scale, the simulated LAI and GPP time series have seasonal fluctuations that are consistent with the observed ones for different plant functional types (Figure 1). The LAI time series for the ENF, DBF, and MIF sites all have distinctive seasonal variation that is well captured by the model. Although Figure 1 for the GRA site shows a regular seasonal cycle, variation in the LAI time series is often irregular with multiple cycles affected by local meteorological conditions such as precipitation for many studied grassland flux towers. The LAI time series in the EBF site have no distinct seasonal cycle and fluctuate with high values...
throughout the entire growing season. The WSA, CSH, and OSH sites tend to have low LAI values during the growing season. For each plant functional type, the growing cycle in GPP matches with LAI, but GPP exhibits large daily variation. Vegetation GPP modeled using our approach well agrees with the GPP measured at flux towers. It is worth noting that the sites shown in Figures 1g and 1h are both located in the Southern Hemisphere and have the growing cycles that are opposite to the sites located in the Northern Hemisphere. Variation in both LAI and GPP for the sites in the Southern Hemisphere is also well modeled.

Modeled LAI agree with satellite-derived LAI on a daily basis for each individual plant functional type (Figure 2). For ENF, DBF, MIF, CSH, OSH, WSA, SAV, and GRA, the model could explain 78.1%, 87.8%, 79.9%, 93.3%, 81.5%, 70.3%, 79.6%, and 70.7% variance of satellite-derived LAI, respectively. Compared with the satellite data, the modeled results have RMSE values less than 1 m²/m² and the absolute values of Bias.
less than 0.3 m²/m² for all plant functional types except EBF. For EBF, a plant functional type without noticeable growing cycles, the model still accounts for 44.3% of the variance in satellite-derived LAI with RMSE of 1.571 m²/m² and negative Bias of −0.613 m²/m². The model performs reasonably well for plant functional types with distinctive growing cycles such as DBF, ENF, and MIF.

In addition to modeling at the daily time scale, the model is able to capture the variation in yearly average LAI and yearly sum GPP across plant functional types (Figure 3). When using all available data, the model could explain 86.6% variance of satellite-derived yearly average LAI (RMSE = 0.470 m²/m² and Bias = 0.217 m²/m²) and 62.3% variance of field-observed yearly sum GPP (RMSE = 473.81 gC/m²/year and Bias = 119.63 gC/m²/year). Table 2 demonstrates that the model can well simulate LAI and GPP on the yearly basis for individual plant functional types. The correlation coefficient between satellite-derived and modeled yearly average LAI is larger than 0.85 for all plant functional types except DBF (r = 0.83) and MIF (r = 0.77), and RMSE is less than 0.60 m²/m² except EBF (RMSE = 1.04 m²/m²). The correlation coefficient between satellite-derived and modeled yearly sum GPP is low for DBF, MIF, OSH, and SAV.

Figure 4 shows the spatial distribution of RMSE between satellite-derived and modeled LAI for each individual station. In the FLUXNET 2015 Tier One data set, most of the available flux tower sites are located in the developed countries in North America, Europe, and Australia. Compared with remote sensing data, the modeled results generally have RMSE less than 1 m²/m² for all sites except the EBF sites. The four stations marked as red dots in Figure 4 (RMSE > 2.0 m²/m²) are CN-Din, GF-Guy, BR-Sa3, and Au-Rob, of which the plant functional types are all EBF. As the flux tower sites for modeling are not evenly distributed and the sites in areas such as China and Russia are sparse, the accuracy of the modeled results is relatively low in those areas, indicating there is a need to increase flux towers in the corresponding areas. For South
4.2. Global-Scale Modeling

Regarding to the spatial distribution of multiyear average LAI (1982–2016), our modeled LAIs are consistent with BULAI4g on a global scale, but EC-Earth3-Veg LAIs are considerably higher than BULAI4g for many places in the world (Figure 5). Both satellite-derived and modeled multiyear average LAIs generally decrease as latitude increases. Our modeled results have consistent latitudinal fluctuations with the satellite-derived product for multiyear average LAI, given that the correlation coefficient is 0.994 and the differences are less than 0.5 m²/m² across the latitude. By comparison, the correlation coefficient of latitudinal average LAI between EC-Earth3-Veg and BULAI4g is 0.878, and EC-Earth3-Veg has positive biases in both northern and southern temperate latitudes. When comparing our modeled with satellite-based multiyear average LAI, the correlation coefficient is as high as 0.975 for all data in the Northern Hemisphere and 0.984 for those in the Southern Hemisphere.

Table 2

| PFT  | Yearly average LAI | Yearly sum GPP |
|------|--------------------|----------------|
|      |                    |                |
|      | r                  | RMSE (m²/m²)   | Bias (m²/m²) | SD_MODIS (m²/m²) | SD_model (m²/m²) | r     | RMSE (gC/m²/year) | Bias (gC/m²/year) | SD_tower (gC/m²/year) | SD_model (gC/m²/year) |
| ENF  | 0.96               | 0.34           | 0.23         | 0.82            | 0.99            | 0.76              | 423.72          | 101.70               | 614.46               | 570.69               |
| EBF  | 0.87               | 1.04           | 0.62         | 1.24            | 1.65            | 0.91              | 429.10          | 86.72                | 929.61               | 681.79               |
| DBF  | 0.83               | 0.29           | 0.09         | 0.50            | 0.55            | 0.08              | 463.23          | 120.19               | 365.58               | 292.88               |
| MIF  | 0.77               | 0.29           | 0.05         | 0.45            | 0.32            | −0.02             | 635.45          | −36.83               | 600.94               | 189.98               |
| CSH  | 0.99               | 0.32           | 0.30         | 1.62            | 1.65            | 0.89              | 185.29          | 60.62                | 335.73               | 383.75               |
| OSH  | 0.89               | 0.11           | 0.01         | 0.21            | 0.24            | −0.02             | 296.09          | −29.96               | 225.35               | 243.89               |
| WSA  | 0.97               | 0.25           | 0.12         | 0.46            | 0.64            | 0.91              | 371.14          | 131.23               | 630.81               | 811.75               |
| SAV  | 0.98               | 0.20           | 0.15         | 0.38            | 0.49            | 0.47              | 608.76          | 134.46               | 542.45               | 628.00               |
| GRA  | 0.87               | 0.59           | 0.29         | 1.05            | 0.94            | 0.80              | 528.96          | 268.75               | 741.61               | 686.51               |

Note. Reported metrics include correlation coefficient (r), root mean square error (RMSE), and standard deviation (SD), where SD_MODIS denotes standard deviation for the MODIS-derived data, SD_model denotes standard deviation for the modeled data, and SD_tower denotes standard deviation for the flux tower data.
In term of the LAI trend in 1982–2016, BULAI4g and our modeled results have similar spatial pattern, but EC-Earth3-Veg overestimates the increasing trends in LAI (Figure 6). In BULAI4g, significant greening ($p < 0.10$ under the Mann-Kendall test) only occurred in areas such as Central Africa, Europe, southern China, and eastern United State, while significant browning ($p < 0.10$ under the Mann-Kendall test) occurred in northern high latitudes such as northern Canada and northern Russia. Our modeled results are basically consistent with satellite observations except for the differences in central Africa. Unlike BULAI4g, EC-Earth3-Veg exhibits a large degree of greening almost globally. In the latitudinal direction, our simulation results have the same trend as BULAI4g with the correlation coefficient as high as 0.763, while the correlation coefficient between EC-Earth3-Veg and BULAI4g is only 0.256. When comparing all pixels with significant trends, the correlation coefficient between our modeled results and BULAI4g is 0.552 in the Northern Hemisphere and 0.843 in the Southern Hemisphere, while the correlation coefficient between EC-Earth3-Veg and BULAI4g is 0.278 in the Northern Hemisphere and $-0.008$ in the Southern Hemisphere.

Taking the year of 2005 as an example, Figure 7 illustrates the model performance on capturing intra-annual LAI variation by comparing the modeled results with satellite data. In the tropical region, the correlation between our modeled and satellite-derived daily LAI time series is low, and the corresponding RMSEs are greater than 1 m²/m² with positive Bias greater than 1 m²/m². In the EBF-dominated areas, the modeled LAI time series appear to saturate with values that are close to maximum LAI. At temperate and high latitudes; our modeled results are close to BULAI4g with high correlation coefficient ($R > 0.8$) and low errors (RMSE < 1 m²/m² and absolute Bias value < 0.5 m²/m²). Comparing with our modeled results, the correlation coefficient between EC-Earth3-Veg and BULAI4g is lower globally with much higher RMSE and larger Bias. EC-Earth3-Veg generally makes much larger LAI prediction than the satellite-based product of BULAI4g.

On the interannual scale, our modeled results have significant correlation with remote sensing data in most of global land surface areas that are covered by natural vegetation (Figure 8). The correlation between modeled and observed results is lower in areas with lower annual average LAI, such as high latitudes in the Northern Hemisphere, indicating that it is still difficult to model interannual variation in LAI in sparsely vegetated areas. Areas with large errors, such as those with RMSE greater than 1.0 m²/m², are mainly located in the tropics and the Nordic region. Similar to the results on the intra-annual scale, both RMSE and Bias generally decrease as latitude increases on the interannual scale. By comparison, EC-Earth3-Veg largely deviates from BULAI4g as indicated by the insignificant correlation globally between EC-Earth3-Veg and BULAI4g on the interannual scale. Compared to BULAI4g, annual average LAI in EC-Earth3-Veg has RMSE that is greater than 1.5 m²/m² and positive Bias that is larger than 1.5 m²/m² for many places in the tropics and the Nordic region.
the world. Overall, these results demonstrate the ability of the developed model on capturing both intra-annual and interannual variation in global LAI.

5. Discussion

Our modeling approach is a semiprognostic phenology model because it requires obtaining the annual maximum LAI for each site or each pixel to constrain the simulated LAI. We have attempted to use prescribed maximum LAI for each plant functional type, but the preliminary modeling results only perform well in a few regions on a global scale, although they seem acceptable at the site scale. Annual maximum LAI,
even for the same individual plant functional type, exhibits strong spatial heterogeneity. Kim and Wang (2005) also used remote sensing data to define annual maximum LAI at the 0.5° × 0.5° grids on the global scale and simulate the LAI time series as key inputs to the Integrated Biosphere Simulator model. The developed semiprognostic phenology model only requires obtaining annual maximum LAI and thus simplifies the phenology modeling problem, but in the long run, it is necessary to develop methods to simulate annual maximum LAI via climate variables.
Our modeling approach allows for obtaining an analytical solution directly based on the Lambert W function to the closed equations of both steady-state LAI and GPP, because it implements with a simple light use efficiency model, MOD17 as described in Equation 2. The MOD17 algorithm only accounts for the impacts of daily minimum temperature and vapor pressure deficit on vegetation productivity using the scalar functions. If using light use efficiency models that have similar forms to MOD17, the analytical solution of steady-state LAI can still be derived based on the Lambert W function simply by replacing the expression of $\mu$ in Equation 5. Because there are models that account for the influence of different environmental factors on light use efficiency, for example, EC-LUE considers the influence of moisture on vegetation productivity (Yuan et al., 2014) and considers the effects of cloud cover on radiation partitioning and canopy light use efficiency (Xin et al., 2016), it implies that applying our method can potentially account for the impacts of different environmental factors such as moisture and cloud cover on the simulated vegetation phenology and LAI time series. These considerations remain to be tested in further researches. If implementing with a sophisticated vegetation productivity model, it is not able to obtain an analytical solution of the steady-state LAI, but we can use the Newton-Raphson method or similar approaches to obtain an approximate solution and then model the actual LAI time series using Equation 11.

The main contribution of this study is to develop a semiprognostic phenology model that allows for simulating time series of vegetation LAI. The success of this method confirms that there is a strong interdependence between vegetation LAI time series and its photosynthesis capacity; that is, on the one hand, LAI and meteorological conditions together determine GPP on the corresponding day, and, on the other hand,
vegetation photosynthesis capacity affects LAI at a certain time lag. Currently, the semipredicted vegetation phenology model cannot be directly applied to LSM for vegetation phenology simulation. When using prescribed maximum LAI for each pixel according to satellite data, the developed approach is able to provide independent LAI model product that could be used as important reference for phenology studies and for benchmarking and diagnosing LSM phenology schemes. In addition, if both Equations 1 and 11 are synthesized into LSM, LSM is able to simulate the time series of LAI, but there is a need to further study and validate the modeling accuracies. In the long run, when a reliable method that allows for simulating maximum LAI based on climate variables becomes available, the method together with our developed semiprosthetic method could form a relatively complete prognostic solution for simulating vegetation LAI time series.

The largest differences between our modeled and satellite-derived LAI time series occurred in the tropic areas, where they are dominated by EBF with unapparent phenology cycles. The effects such as cloud contamination and signal saturation are also pronounced in low-latitude regions and contribute to the quality issues of the satellite data. As field-measured LAI data were not available, time series of LAI extracted from the MODIS product were used as surrogates, but they contributed errors to model assessment. Due to limited field sites and ground observations, it is difficult but obviously necessary to improve vegetation GPP and LAI models for EBF. In addition, we did not attempt to use the developed approach to simulate the LAI time series for croplands because crop phenology is largely dependent on human activities such as planting and harvesting. In the land surface models, crop phenology and crop-related processes are normally simulated independently of natural vegetation.

Figure 8. The spatial distribution of correlation coefficient (a and d), root mean square error (b and e), and mean bias error (c and f) between our modeled and satellite-derived annual average LAI time series (left column) and between EC-Earth3-veg and satellite-derived annual average LAI time series (right column) in 1982–2016 on a global scale. N = 35 and r = 0.2826 correspond to the 0.100 significance level. Cropland, snow/ice, and nonvegetated areas are shown in gray.
6. Conclusions

This study developed a new phenology modeling approach that is able to simulate time series of vegetation LAI. This approach establishes a linear relationship between LAI and vegetation productivity and implements with a simple light use efficiency algorithm of MOD17. The analytical solution to the closed equations can be obtained by the Lambert W equation, and the obtained steady-state LAI is used to derive time series of actual LAI using a simple restricted growth process model. The results modeled using flux tower data show that the developed phenology model can well simulate daily and yearly LAI for each plant functional type on the site scale. The results modeled using global climate data demonstrate that the model is able to capture both the spatial pattern and intra-annual and interannual variation of LAI derived from the remote sensing data on the global scale. The semiprospective phenology model developed in this study provides a straightforward solution to simulating and understanding the spatiotemporal variation in vegetation LAI across plant functional types.

Data Availability Statement

This work used eddy covariance data acquired and shared by the FLUXNET community, including the following networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCONSiberia, and USCC. The ERA-Interim reanalysis data are provided by ECMWF and processed by LSCE. The FLUXNET eddy covariance data processing and harmonization were carried out by the European Fluxes Database Cluster, AmeriFlux Management Project, and Fluxdata project of FLUXNET, with the support of CDiac and ICOS Ecosystem Thematic Center, and the OzFlux, ChinaFlux, and AsiaFlux offices. The model we proposed and data from the simulations are available online (via https://zenodo.org/record/3779917#.Xqv3BpCRXCI). For more details please contact the authors.

Acknowledgments

This research is supported by National Key R&D Program of China (Grants 2017YFA0604308 and 2017YFA0604400), National Natural Science Foundation of China (Grant 41875122), Western Talents (Grant 2018XYJRC004), Guangdong Top Young Talents (Grant 2017TQ04Z359), Innovation and Development Projects for Excellent Postgraduate at Sun Yat-sen University (Grant 19gyjsh10), and One Hundred Talents Program of the Chinese Academy of Science (Grant Y764141001). We also thank the anonymous reviewers for their constructive comments.

References

Broich, M., Huete, A., Tulbure, M. G., Ma, X., Xin, Q., Paget, M., et al. (2014). A generalized, bioclimatic index to predict foliar phenology in response to climate. *Biogeosciences*, 11(8), 5181–5198. https://doi.org/10.5194/bg-11-5181-2014

Chen, C., Park, T., Wang, X., Piao, S., Xu, B., Chaturvedi, R. K., et al. (2019). China and India lead in greening of the world through land-use management. *Nature Sustainability*, 2(2), 122–129. https://doi.org/10.1038/s41893-019-0220-7

Chen, H., Yu, C., Li, C., Xin, Q., Huang, X., Zhang, J., et al. (2016). Modeling the impacts of water and fertilizer management on the ecosystem service of rice rotated cropping systems in China. *Agriculture, Ecosystems & Environment*, 219, 49–57. https://doi.org/10.1016/j.agee.2015.11.023

Chen, J., Yao, Y., Shen, M., Wang, C., & Yang, X. (2016). A simple method for detecting Phenological change from time series of vegetation index. *IEEE Transactions on Geoscience and Remote Sensing*, 54(6), 3436–3449. https://doi.org/10.1109/tgrs.2016.2518167

Chen, M., Melas, E. K., Gray, J. M., Friedli, M. A., & Richardsdon, 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. https://doi.org/10.1111/gcb.13326

Chuine, I., Cour, P., & Rouseau, D. D. (1999). Selecting models to predict the timing of flowering of temperate trees: Implications for tree phenology modelling. *Plant, Cell and Environment*, 22(1), 1–13. https://doi.org/10.1046/j.1365-3040.1999.00395.x

Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J., & Knuth, D. E. (1996). On the LambertW function. *Advances in Computational Mathematics*, 5(1), 329–359. https://doi.org/10.1007/BF02124750

Duan, Q., Gupta, V. K., & Sorooshian, S. (1993). Shuffled complex evolution approach for effective and efficient global minimization. *Journal of Optimization Theory and Applications*, 76(3), 501–521. https://doi.org/10.1007/104902190301

Duan, Q., Sorooshian, S., & Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, 28(4), 1015–1031. https://doi.org/10.1029/91WR02985

Forkel, M., Carvalhais, N., Schaphoff, S., v.Bloh, W., Migliavacca, M., Thurner, M., & Thonicke, K. (2014). Identifying environmental controls on vegetation greenness phenology through model-data integration. *Biogeosciences*, 11(23), 7025–7050. https://doi.org/10.5194/bg-11-7025-2014

Hu, Z., Shi, H., Cheng, K., Ying-Ping, W., Piao, S., Li, Y., et al. (2018). Joint structural and physiological control on the inter-annual variation in productivity in a temperate grassland: A data-model comparison. *Global Change Biology*, 24(3), 2965–2979. https://doi.org/10.1111/gcb.14274

Hu, Z., Wu, G., Zhang, L., Li, S., & Yu, G. (2017). Modeling and partitioning of regional evapotranspiration using a satellite-driven water-carbon coupling model. *Remote Sensing*, 9(1), 54–75. https://doi.org/10.3390/rs9010054

Hufkens, K., Basler, D., Milliman, T., Melas, E. K., & Richardsdon, A. D. (2018). An integrated phenology modelling framework in R. *Methods in Ecology and Evolution*, 9(5), 1276–1285. https://doi.org/10.1111/2041-210X.12970

Jiang, C., Ryu, Y., Fang, H., Myeneni, R., Clevier, M., & Zhu, Z. (2017). Inconsistencies of interannual variability and trends in long-term satellite leaf area index products. *Global Change Biology*, 23(10), 4133–4146. https://doi.org/10.1111/gcb.13787

Jolly, W. M., Nemani, R., & Running, S. W. (2005). A generalized, bioclimatic index to predict foliar phenology in response to climate. *Global Change Biology*, 11(4), 619–632. https://doi.org/10.1111/j.1365-2486.2005.00930.x

Keenan, T. F., Gray, J., Friedli, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., et al. (2014). Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nature Climate Change*, 4(7), 598–604. https://doi.org/10.1038/nclimate2553
Yuan, W., Cai, W., Xia, J., Chen, J., Liu, S., Dong, W., et al. (2014). Global comparison of light use efficiency models for simulating terrestrial vegetation gross primary production based on the LaThuile database. Agricultural and Forest Meteorology, 192–193(0), 108–120. https://doi.org/10.1016/j.agrformet.2014.03.007

Zhang, G., Zhang, Y., Dong, J., & Xiao, X. (2013). Green-up dates in the Tibetan plateau have continuously advanced from 1982 to 2011. Proceedings of the National Academy of Sciences, 110(11), 4309–4314. https://doi.org/10.1073/pnas.1210423110

Zhang, X., Wang, J., Gao, F., Liu, Y., Schaaf, C., Friedl, M., et al. (2017). Exploration of scaling effects on coarse resolution land surface phenology. Remote Sensing of Environment, 190, 318–330. https://doi.org/10.1016/j.rse.2017.01.001

Zhao, M., & Running, S. W. (2010). Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. Science, 329(5994), 940–943. https://doi.org/10.1126/science.1192666

Zhu, W., Zheng, Z., Nan, J., & Zhang, D. (2018). A comparative analysis of the spatio-temporal variation in the phenologies of two herbaceous species and associated climatic driving factors on the Tibetan plateau. Agricultural and Forest Meteorology, 248, 177–184. https://doi.org/10.1016/j.agrformet.2017.09.021

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., et al. (2013). Global data sets of vegetation leaf area index (LAI) 3g and fraction of Photosynthetically active radiation (FPAR) 3g derived from global inventory modeling and mapping studies (GIMMS) normalized difference vegetation index (NDVI3g) for the period 1981 to 2011. Remote Sensing, 5(2), 927–948. https://doi.org/10.3390/rs5020927