Observations of satellite land surface phenology suggest that maximum leaf greenness affects global vegetation productivity more than growing season length

Xiaojie Gao[1], Ian McGregor[2], Josh Gray[2], Mark Friedl[3], and Minkyu Moon[3]

[1]Center for Geospatial Analytics, North Carolina State University
[2]Center for Geospatial Analytics
[3]Boston University

November 24, 2022

Abstract

Vegetation green leaf phenology directly impacts gross primary productivity (GPP) of terrestrial ecosystems. Satellite observations of land surface phenology (LSP) provide an important means to monitor the key timing of vegetation green leaf development. However, differences between satellite-derived LSP proxies and in-situ measurements of GPP make it difficult to quantify the impact of climate-induced changes in green leaf phenology on annual GPP. Here we used 1,110 site-years of GPP measurements from eddy-covariance towers in association with time series of satellite LSP observations from 2000-2014 to show that while satellite LSP explains a large proportion of variation in annual GPP, changes in green-leaf-based growing season length (GSL) had less impact on annual GPP by ˜30% than GSL changes in GPP-based photosynthetic duration. Further, maximum leaf greenness explained substantially more variance in annual GPP than green leaf GSL, highlighting the role of future vegetation greening trends on large-scale carbon budgets. We conclude that satellite LSP-based inferences regarding large-scale dynamics in GPP need to consider changes in both green leaf GSL and maximum greenness.

Observations of satellite land surface phenology suggest that maximum leaf greenness affects global vegetation productivity more than growing season length

Xiaojie Gao[1*], Ian R. McGregor[1], Josh M. Gray[1,2], Mark A. Friedl[3], and Minkyu Moon[3]

[1] *Center for Geospatial Analytics, North Carolina State University, Raleigh, NC, USA.*

[2] *Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, USA*

[3] *Department of Earth and Environment, Boston University, USA*

[*] *Corresponding author. Email: xgao26@ncsu.edu*

Abstract

Vegetation green leaf phenology directly impacts gross primary productivity (GPP) of terrestrial ecosystems. Satellite observations of land surface phenology (LSP) provide an important means to monitor the key timing of vegetation green leaf development. However, differences between satellite-derived LSP proxies and in-situ measurements of GPP make it difficult to quantify the impact of climate-induced changes in green leaf phenology on annual GPP. Here we used 1,110 site-years of GPP measurements from eddy-covariance towers in association with time series of satellite LSP observations from 2000-2014 to show that while satellite LSP explains a large proportion of variation in annual GPP, changes in green-leaf-based growing season length (GSL) had less impact on annual GPP by ∼30% than GSL changes in GPP-based photosynthetic duration. Further, maximum leaf greenness explained substantially more variance in annual GPP than green leaf GSL, highlighting the role of future vegetation greening trends on large-scale carbon budgets. We conclude that satellite LSP-based inferences regarding large-scale dynamics in GPP need to consider changes in both green leaf GSL and

maximum greenness.

1 Introduction

The timing and duration of vegetation growing seasons have changed across much of the Earth's terrestrial ecosystems over the last several decades (Buitenwerf et al., 2015; Hua et al., 2021; Park et al., 2016; Piao et al., 2007). Observed trends of advanced spring and delayed autumn increased the growing season length (GSL) globally. The increased GSL is generally thought to result in increased gross primary productivity (GPP) (Buermann et al., 2018; Dragoni et al., 2011; Keenan et al., 2014; Piao et al., 2019), which reflects the total amount of carbon absorbed by vegetated ecosystems during a unit time period. However, GSL changes in green leaf development might not always synchronize with changes in GPP photosynthetic duration, and geographically heterogeneous patterns of leaf greenness intensity changes, so called greening and browning, complicate this relationship, particularly at high latitudes (C. Chen et al., 2019; Huang et al., 2018; Liu et al., 2021; Zhu et al., 2016). Since global greening and browning trends alter the photosynthetic competence of vegetation in sequestering carbon, they significantly affect ecosystem GPP. Understanding the interactive roles of changes in GSL and leaf greenness is critical to infer global carbon dynamics and future climate change.

Eddy-covariance (EC) flux towers provide direct measurements of GPP, but over limited spatial extents (Baldocchi, 2020). By using EC measurements, previous studies have investigated the sensitivity of annual GPP (ΣGPP) to GPP-based photosynthetic seasonal and physiological covariates. For example, Xia et al., 2015 found that over 90% of the ΣGPP variation can be explained by photosynthetic GSL and annual maximum GPP. Following the same method, Zhou et al., 2016 and Zhou et al., 2017 supported the conclusion and found that changes in maximum GPP explain more variability in ΣGPP than the start and end of photosynthetic timing. However, whether a similar relationship holds for leaf-greenness-based phenology and

physiology remains unclear. More importantly, although GPP upscaling methods that relies on leaf greenness related measurements (Jung et al., 2019; Le Quéré et al., 2016) have been developed to estimate large spatial scale vegetation productivity, the comparison between impacts of GSL changes in green leaf development on vegetation productivity with that of GSL changes in photosynthetic duration over large spatial scales have rarely been studied.

In contrast, satellite remote sensing provides spatially continuous long-term observations of green leaf development and GSL at global scale (Friedl et al., 2019; Ganguly et al., 2010). Satellite-observed land surface phenology (LSP) data is the only source of global green leaf GSL information (Caparros-Santiago, 2021; Piao et al., 2019), and has been widely incorporated into both process-based and data-driven ecological models to upscale field-based measurements of GPP from eddy covariance towers to larger areas (Falge et al., 2002; Richardson et al., 2010; Richardson et al., 2013). However, because LSP measures the period of green leaf development rather than photosynthesis, ecosystem-scale phenology measured by LSP data sets do not always accurately capture the seasonality of GPP, and annual maximum GPP in Xia et al., 2015 is not necessarily reflected by annual maximum of leaf greenness. Moreover, although large scale ΣGPP responses to leaf greenness changes has been investigated (Huang et al., 2018; Keenan et al., 2014), few studies have explored how changes in green-leaf-based GSL and maximum greenness jointly regulate geographic and interannual variation in ΣGPP, even though both processes are occurring concurrently in many ecosystems. Therefore, investigating the nature and magnitude of how satellite LSP observations explain ΣGPP variation, as well as the joint effect of changes in green-leaf-based GSL and maximum greenness on ΣGPP, are helpful to understanding how satellite observations can be used to study climate change induced dynamics in the global carbon cycle.

In this study, by using extensive EC measurements provided by the FLUXNET2015 project (Pastorello et al., 2020) in combination with global time series of LSP observations from

NASA's Moderate Spatial Resolution Imaging Spectroradiometer (MODIS) from 2000 to 2014, we investigated the strength of the covariance between LSP and GPP seasonality and modeled how satellite-observed green leaf phenology and maximum greenness control ΣGPP across global terrestrial ecosystems. We aimed to evaluate the performance of using satellite LSP-derived green leaf GSL and maximum greenness in inferring ΣGPP, compared to GPP-based photosynthetic GSL and maximum GPP derived from EC measurements. Using Bayesian hierarchical models, we were able to quantify the variation of the relationship between ΣGPP and green leaf phenology within and across global biomes and flux sites. We found that satellite LSP-derived GSL had less ΣGPP sensitivity to GPP derived GSL by ~30% and that the GSL-ΣGPP relationship varies by biome type and flux site. Importantly, we also found that maximum greenness exerted stronger control on ΣGPP than LSP-derived GSL, suggesting that future leaf greening trends, represented by trends of increasing maximum greenness, would increase global vegetation productivity more than extending the time period of green leaf development.

2 Methods

2.1 Datasets

We obtained GPP measurements of 166 EC flux sites distributed across the globe (Fig. 1) from the FLUXNET2015 project (https://fluxnet.org/data/fluxnet2015-dataset/). Although the flux dataset provides decades of EC measurements until 2014, we only used the EC data within 2000-2014 because this time period matches with MODIS observations. Then, we filtered out site-years data that contain continuous missing values with more than 45 days. After data preprocessing, there were 1,110 site-years data left, representing 11 biome types (Fig. 1). To identify which GPP variable in the dataset was better for studying phenology, we undertook a sensitivity analysis (Fig. S1) by calculating GPP-based phenometrics from each daytime GPP variable in the FLUXNET2015 dataset. As no significant differences

were found among those GPP variables, we chose GPP_DT_VUT_REF, which represents daytime GPP using a Variable U-star (U^*) Threshold (Pastorello et al., 2020), to conduct this research. We extracted ΣGPP measurements for each site-year as well as daily GPP time series from the FLUXNET2015 data to conduct this analysis.

To analyze EC measurements with satellite LSP observations, we extracted satellite LSP data from 2000 to 2014 for flux site locations from the MCD12Q2 v6 product (https://lpdaac.usgs.gov/products/mcd12q2v006/) (Friedl et al., 2019; Ganguly et al., 2010). This data product provides global, annual LSP estimates at 500 m spatial resolution based on time series of the two-band enhanced vegetation index (EVI2; Jiang et al., 2008). The MCD12Q2 LSP product uses spline functions to smooth the EVI2 time series and percentage thresholds to extract phenological dates. The estimated phenometrics include Greenup, MidGreenup, Maturity, Peak, Senescence, MidGreendown, and Dormancy, representing 15%, 50%, 90%, 100% of annual EVI2 amplitude in spring and autumn respectively.

2.2 GPP-based seasonality estimation and evaluation

To compare with satellite LSP phenometrics, we applied the MCD12Q2 LSP algorithm (Friedl et al., 2019) to the daily GPP time series to retrieve annual GPP-based seasonality metrics as well as annual maximum GPP (GPP_{max}) and annual minimum GPP (GPP_{min}). We followed the same procedure of MCD12Q2 phenometrics estimation, whereby for any particular year, we first gathered daily GPP measurements for the full calendar year plus six months before and after, and then smoothed the time series using a spline function. The GPP-based seasonality metrics were then estimated by the same percentage thresholds as MCD12Q2 but were based on GPP amplitude. We obtained 866 site-years of GPP seasonality metrics out of the 1,110 site-years of EC measurements from the algorithm. There were 244 site-years data that had GPP values but the annual amplitudes derived from the smoothed spline function were too low ($< 4\ gCm^{-2}yr^{-1}$) to estimate reliable seasonality.

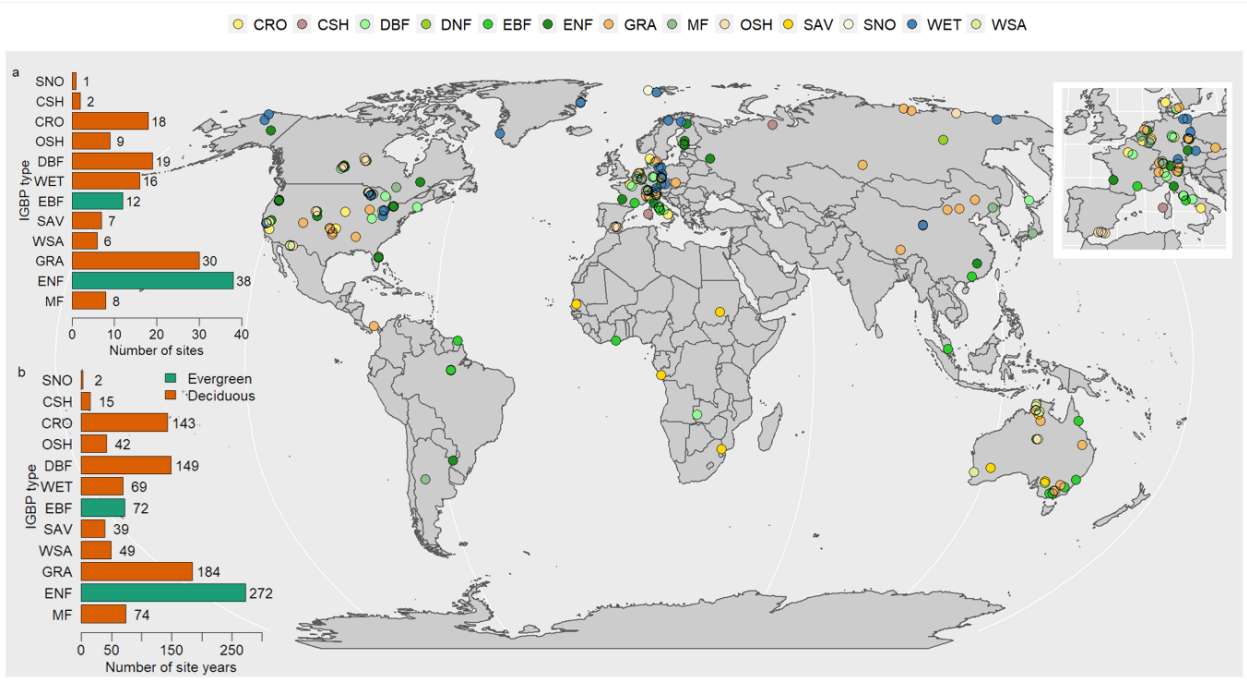

Figure 1: The distribution of eddy-covariance flux towers used in this study. **(a)** Number of sites and **(b)** number of site-years per biome type used in this study. The biome type is determined by the International Geosphere-Biosphere Programme (IGBP) classification data provided in the FLUXNET2015 dataset.

To evaluate the consistency between the GPP-based seasonality metrics and the satellite LSP data, we iteratively searched through each GPP seasonality metric for a potential match of the same type of LSP phenometric within a certain time period window ($\pm$185 days), centered on the GPP metric date. If all of the corresponding LSP phenometrics in a site-year were found, the site-year was marked as "match", otherwise "no-match". The search window is wide enough to capture potential matches, and the consistency between the GPP seasonality metrics and the LSP phenometrics was then evaluated by linear regression. Among the matched phenometrics and GPP metrics, we linearly regressed GPP seasonality metrics against LSP phenometrics. Besides reporting R^2 and root mean squared error (RMSE) from the linear regression, the mean relative deviation (MRD) and mean absolute deviation (MAD) were also calculated to quantify the differences between LSP phenometrics and GPP seasonality metrics. Specifically, MRD evaluates bias in the GPP and EVI2 metrics, while

MAD quantifies the absolute deviation between the two data sources. With N defined as the total number of matched site-years, the formulas of MRD and MAD were used as follows:

$$MRD = 1/N * \Sigma_{i=1}^{N}(GPP\ metric_i - EVI2\ phenometric_i) \quad (1)$$

$$MAD = 1/N * \Sigma_{i=1}^{N}|GPP\ metric_i - EVI2\ phenometric_i| \quad (2)$$

The "no-match" site-years were also investigated by comparing their time series data of MODIS EVI2 and EC measured GPP, respectively. Five reasons for their mismatch were investigated: EVI2 data missing, GPP data missing, EVI2 amplitude too low to retrieve LSP, GPP amplitude too low to retrieve GPP seasonality metrics, and both LSP and GPP metrics exist but do not match. Note that since the MCD12Q2 LSP product only processed pixels with annual EVI2 amplitude greater than 0.1, to distinguish between EVI2 data missing and EVI2 amplitude too low to retrieve LSP, we utilized the MCD43A4 Nadir Bidirectional Reflectance Distribution Function (BRDF)-Adjusted Reflectance (NBAR) dataset (Schaaf and Wang, 2015).

We also investigated the consistency of GPP seasonality metrics with LSP from single pixels, 3-by-3 pixel windows, and 5-by-5 pixel windows, respectively. The values in pixel windows were aggregated by mean and median. We found that single pixel LSP represents GPP seasonality metrics the best with higher R^2 values, lower RMSE values, compared with pixel window based aggregated LSP (result not shown). So, we used single-pixel LSP to conduct further analysis in this study.

2.3 Annual GPP model analysis

To understand the phenology and physiology effects on ΣGPP, inspired by Xia et al., 2015, we introduced five nested Bayesian hierarchical regression models representing different

hypotheses. The Bayesian framework allows partial pooling (Gelman and Hill, 2006) that accounts for the unbalanced number of site-years for categories of biome types and flux sites. More importantly, the Bayesian hierarchical models help us understand the site-level and biome-level effects by capturing the variability of model coefficients among flux sites and biome types. The models considered are:

$$Model\ 1: Y \sim N(\beta_0 + \beta_1 GSL + \beta_2 Z_{max} + \beta_3 Z_{min}, \sigma_y^2) \tag{3}$$

$$\begin{aligned} Model\ 2: Y &\sim N(\beta_{0j} + \beta_1 GSL + \beta_2 Z_{max} + \beta_3 Z_{min}, \sigma_y^2) \\ \beta_{0j} &\sim N(\mu_0, \sigma_0^2) \end{aligned} \tag{4}$$

$$\begin{aligned} Model\ 3: Y &\sim N(\beta_{0j} + \beta_{1j} GSL + \beta_{2j} Z_{max} + \beta_{3j} Z_{min}, \sigma_y^2) \\ \beta_{nj} &\sim N(\mu_n, \sigma_n^2), n = 0, 1, 2, 3 \end{aligned} \tag{5}$$

$$\begin{aligned} Model\ 4: Y &\sim N(\beta_{0k} + \beta_{1j} GSL + \beta_{2j} Z_{max} + \beta_{3j} Z_{min}, \sigma_y^2) \\ \beta_{0k} &\sim N(\mu_0, \sigma_0^2) \\ \beta_{nj} &\sim N(\mu_n, \sigma_n^2), n = 1, 2, 3 \end{aligned} \tag{6}$$

$$\begin{aligned} Model\ 5: Y &\sim N(\beta_{0k} + \beta_{1k} GSL + \beta_{2k} Z_{max} + \beta_3 k Z_{min}, \sigma_y^2) \\ \beta_{nk} &\sim N(\eta_{nj}, \sigma_0^2) \\ \eta_{nj} &\sim N(\mu_n, \tau_n^2), n = 0, 1, 2, 3 \end{aligned} \tag{7}$$

where Y represents ΣGPP; j represents biome type; k represents flux site; Z_{max} and Z_{min} represent the annual maximum and minimum of GPP or EVI2; $\mu_n, \eta_n, n = (0, 1, 2, 3)$ represent the mean values of the population distribution of intercepts and slopes at biome and site levels, respectively; σ_y^2, σ_n^2, and $\tau_n^2, n = (0, 1, 2, 3)$ are the corresponding variances. Model 1 is the simplest model that explores whether ΣGPP can be explained by GSL and Z_{max} and Z_{min}; Model 2 considers background GPP/EVI2 variability among biome types; Model 3 explores the variability of covariates among biome types; Model 4 adds site-level intercepts to

test the importance of considering site-level background GPP/EVI2 rates; Model 5 is the full model that considers both site-level and biome-level variability. Cross validation was used to determine the significance of model fit improvements when the model complexity increases. By using GPP-metrics-based models as benchmarks, we are able to understand the effects of phenology and physiology on annual carbon uptake and evaluate the capability of satellite observations in capturing these effects. The GSL values were calculated by the duration between MidGreenup and MidGreendown dates, representing the timing of the time series reaching 50% of amplitude in spring and autumn, estimated from GPP time series for GPP-metrics-based models or obtained from the MCD12Q2 LSP product for EVI2-metrics-based models. The MidGreenup and MidGreendown were selected because they are more robust and reliable in satellite LSP observations than other phenometrics. Note that evergreen plants also have leaf-based seasonality and that can be captured by satellite observations (R. Wang et al., 2019). To compare the sensitivity of ΣGPP to model covariates, the standardized models with centered and scaled covariates were also implemented. Note that to exclude the influence of multiple phenological cycles and cross-calendar-year phenology on ΣGPP calculation, we focused on the data with a single phenological cycle and locate at northern hemisphere only in this modeling analysis. The GPP-based phenological cycles were determined by the fitted spline function same as producing the MCD12Q2 LSP product (Friedl et al., 2019).

The Bayesian hierarchical models were implemented by JAGS software (v4.3.0) and R programming language (v3.6.3). All parameters in the models were assigned uninformative prior distributions to let data dominate the calculation of posterior distributions. We summarized the parameter posterior distributions by median values and 95% credible intervals (CIs) obtained from samples of Markov Chain Monte Carlo (MCMC) by JAGS. All the R scripts are available online (See section Code availability).

3 Results

3.1 Phenometrics comparison

To compare the GPP seasonality metrics with MODIS LSP phenometrics, we first aligned their dates individually (Fig. 2). Of the 1,110 site year GPP seasonality metrics data, 758 (68%) matched phenometrics with the MODIS LSP; 124 (11%) did not have GPP seasonality metrics due to missing GPP values; 96 (9%) were left because MODIS LSP did not provide phenometric values; 11 partially matched with the MODIS LSP phenological cycle; and only one of them did not find a match at all. Note that 120 (11%) of the site-years of data had GPP and EVI2 amplitudes that were too low to reliably estimate GPP seasonality metrics and MODIS LSP phenometrics; this could also be considered consistent between EC measurements and satellite LSP observations.

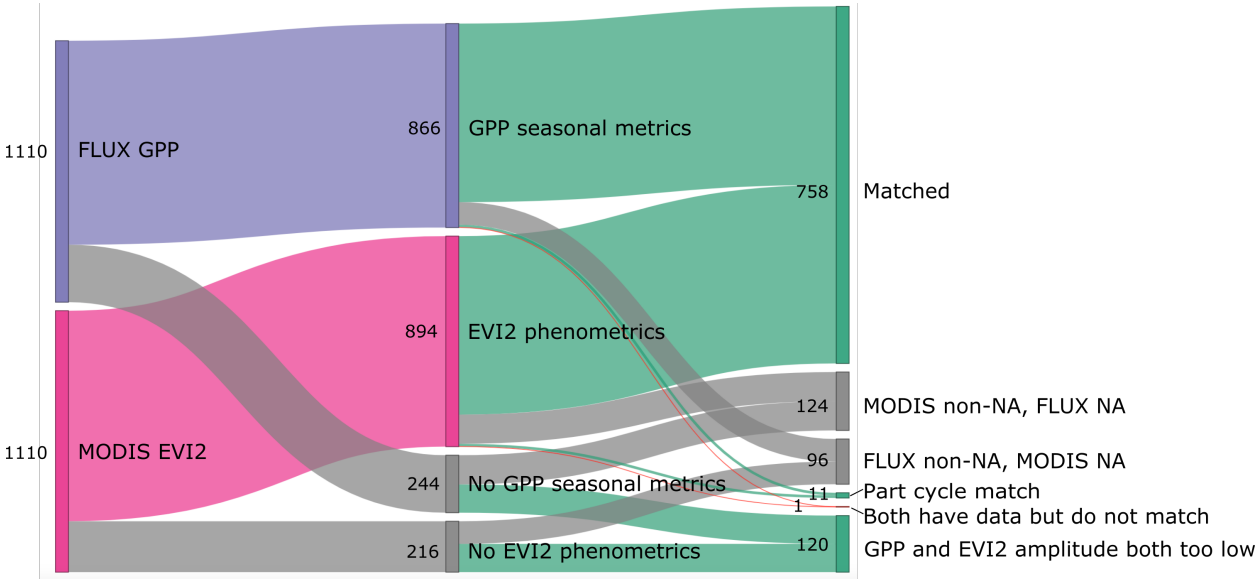

Figure 2: The match process of MODIS EVI2 phenometrics and flux towered measured GPP seasonality metrics. Numbers in the figure are number of site-years.

The regression analysis suggests that LSP measurements from MODIS had general agreement with EC measurements across sites, but exhibit systematic bias among deciduous and evergreen vegetation. Specifically, MidGreenup and MidGreendown from LSP measurements estimate later start of GSL for evergreen vegetation and later end of GSL for deciduous vegetation,

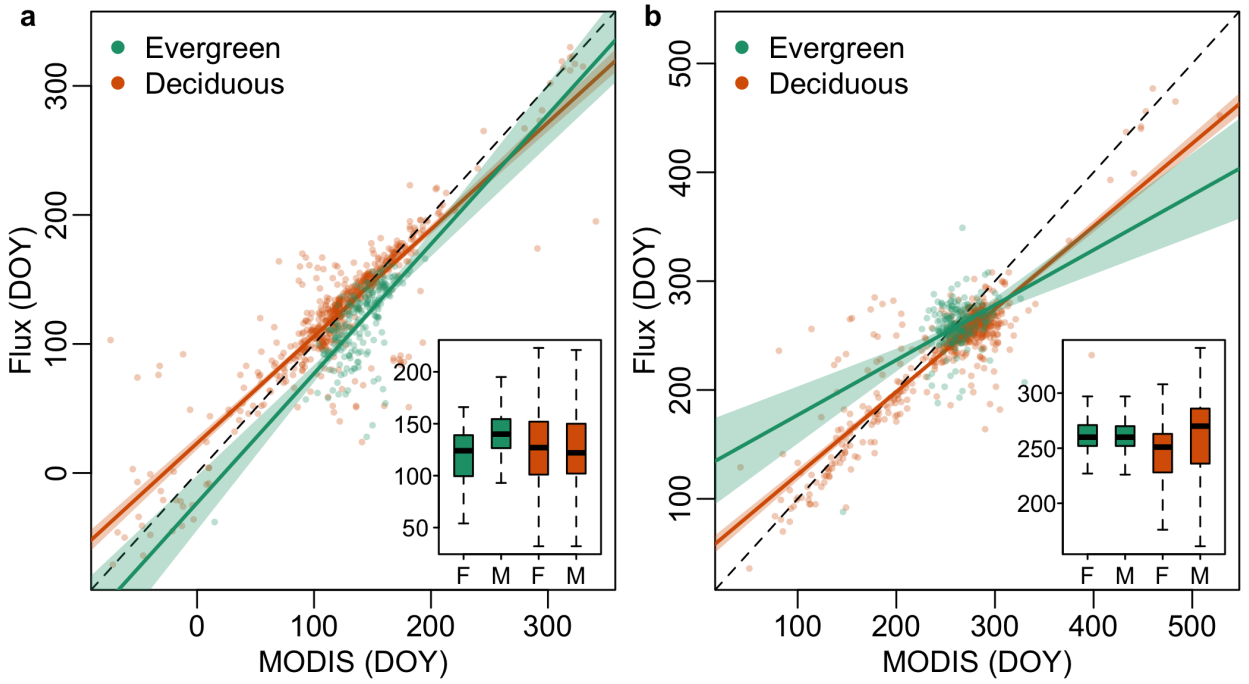

Figure 3: Phenometrics comparison. Comparison of the phenometrics derived from Flux GPP time series and MODIS EVI2 time series for MidGreenup (50% greenness in spring) (a) and MidGreendown (50% greenness in autumn) (b). Box plots show the distributions of flux phenometrics ("F") and MODIS phenometrics ("M").

relative to corresponding metrics from EC measurements (Fig. 3, Table 1). For deciduous sites, the satellite-observed MidGreenup showed strong agreement with corresponding estimates derived from EC measurements (MRD = 2 days, R^2 = 0.78, p-value $<$ 0.05). However, MidGreendown LSP measurements showed significant systematic bias relative to EC-derived estimates (MRD = -14 days, R^2 = 0.79, p-value $<$ 0.05). At evergreen sites, however, MidGreenup from LSP observation was systematically late relative to EC estimates (MRD = -23 days, R^2 = 0.47, p-value $<$ 0.05) and was only weakly correlated with MidGreendown dates from EC measurements (MRD = -3 days, R^2 = 0.16, p-value $<$ 0.05). The consistent MidGreenup and MidGreendown biases for deciduous and evergreen are representative for all phenometrics considered in spring and autumn (Fig. S5).

Table 1: Regression statistics for comparison of phenometrics in Fig. 3 from EC versus LSP measurements. MRD = mean relative deviance; MAD = mean absolute deviation.

Phenometric	Slope	Intercept	R2	MRD	MAD
Evergreen MidGreenup	1.00 (±0.15)	-23.2 (±21.07)	0.47	-22.81	23.39
Deciduous MidGreenup	0.83 (±0.03)	23.12 (±4.93)	0.78	1.82	15.36
Evergreen MidGreendown	0.51 (±0.16)	126.42 (±42.26)	0.16	-3.32	16.84
Deciduous MidGreendown	0.76 (±0.03)	46.16 (±8.45)	0.79	-14.03	23.30

3.2 Annual GPP sensitivity

Nearly all of the variance in large-scale ΣGPP is explained by three characteristics of seasonality in the EC measurements: growing season length (GSL), growing season maximum GPP (GPP_{max}), and growing season minimum GPP (GPP_{min}) (Fig. 4 a,b); the model with both biome- and site-level effects explained 98% of variance in ΣGPP, with a root mean squared error (RMSE) of 73.65 $gCm^{-2}yr^{-1}$ (Fig. 4b). This corresponds to roughly 5% of the average ΣGPP across all site-years. These results are consistent with results from previous studies (Xia et al., 2015; Zhou et al., 2016), but is based on more site-years of EC measurements and a different model structure.

The results in Fig. 4a and 4b demonstrate that GPP phenology effectively explains geographic and interannual variance in ΣGPP measured at EC tower sites. However, because EC towers provide a sparse and non-representative sample of global terrestrial ecosystems, it is difficult to use these data to make inferences regarding large-scale dynamics in GPP arising from changes in phenology. To explore how well leaf greenness based phenological and physiological metrics derived from satellite LSP observations explain ΣGPP, we estimated a Bayesian hierarchical model with the same basic form, but using the matched MODIS LSP metrics (GSL, minimum and maximum EVI2) at northern hemisphere as proxies for corresponding metrics derived from EC measurements (Fig. 4c,d). Compared to models fitted using EC-derived seasonality metrics, models estimated using MODIS LSP metrics yielded weaker agreement with in-situ measurements of ΣGPP (Fig. 4c,d). The LSP-based model with site-level effects showed strong overall correlation with ΣGPP ($R^2 = 0.88$), but the RMSE was

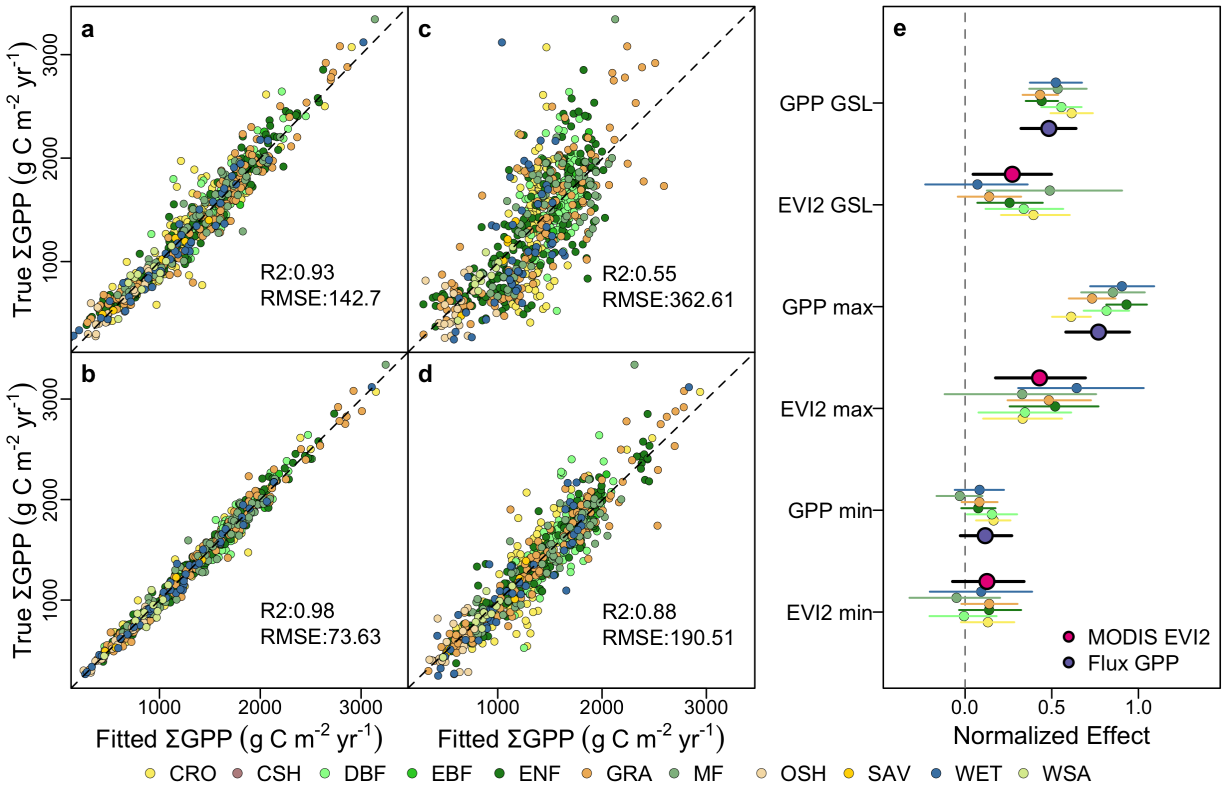

Figure 4: Bayesian hierarchical model results. (a) GPP metrics-based model with biome-level intercepts and slopes; (b) GPP metrics-based model with site-level intercepts and slopes; (c) MODIS LSP-based model with biome-level intercepts and slopes; (d) MODIS LSP-based model with site-level intercepts and slopes; (e) Comparison showing the normalized effect of GPP metrics- and LSP-based models on model results with site-level intercepts and slopes. Bars show 95% Bayesian credible intervals. The larger red and blue points in (e) show the overall effect across all biomes from the LSP- and GPP-derived metrics, respectively. ΣGPP is the annual GPP at each EC tower.

nearly double that obtained using metrics based on EC measurements (190.45 $gCm^{-2}yr^{-1}$). These results indicate that satellite-based LSP metrics are able to estimate ΣGPP, but include substantial uncertainty. Consistent with results based on EC-derived seasonality metrics, the strong positive relationship between ΣGPP and LSP-derived GSL and $EVI2_{max}$ (Fig. 4e) demonstrate that satellite-based observations of green leaf duration and maximum greenness (e.g. Keenan et al., 2014; Park et al., 2016) explain a large proportion of variability in ΣGPP across global terrestrial ecosystems.

Models estimated from MODIS LSP metrics suggests a smaller magnitude of green-leaf-based GSL effect on ΣGPP relative to the EC metrics-based models. To quantify this, we estimated models using standardized EC- and LSP-derived metrics, which allowed us to compare the magnitude of coefficients (i.e., the relative sensitivity of ΣGPP) for each predictor variable across models (Fig. 4e). After controlling for $EVI2_{max}$ and $EVI2_{min}$, the influence of satellite LSP-derived GSL was roughly half the magnitude of GSL derived from EC measurements after controlling for GPP_{max} and GPP_{min}. An increase of one standard deviation in EC-derived GSL increased the standard deviation in ΣGPP by 0.48 (0.33 to 0.63, 95% Bayesian credible interval), versus 0.27 (0.04 to 0.50) for GSL derived from MODIS LSP, controlling for GPP_{max} and GPP_{min}. Controlling for GPP_{max} and GPP_{min} and extending the photosynthetic GSL by one day in the model estimated from EC-derived metrics leads to an increase in ΣGPP of 7.2 $gCm^{-2}yr^{-1}$, but only 5.0 $gCm^{-2}yr^{-1}$ in the corresponding LSP-metrics-based model controlling for $EVI2_{max}$ and $EVI2_{min}$. Stated more directly, the result suggests that GSL changes in green leaf development had roughly 30% less effect on ΣGPP on average across biomes compared to changes in photosynthetic duration.

The magnitude of the GSL effect in models estimated using both EC metrics and MODIS LSP metrics varied across biomes. Overall, variance in the GSL effect across biomes was smaller in the EC metrics-based models (4.7 $gCm^{-2}yr^{-1}$) than in the LSP metrics-based models (7.6 $gCm^{-2}yr^{-1}$) (Fig. 4e, Fig. 5). Croplands (CRO) and deciduous broadleaf forests (DBF) had the largest GSL effects, with values of 8.4 $gCm^{-2}yr^{-1}$ and 8.2 $gCm^{-2}yr^{-1}$ respectively in the EC metrics-based model, and 7.2 $gCm^{-2}yr^{-1}$ and 6.2 $gCm^{-2}yr^{-1}$, respectively, in the LSP metrics-based model. Evergreen needleleaf forests (ENF) and grasslands (GRA) showed lower GSL effects, with a value of 6.0 $gCm^{-2}yr^{-1}$ and 6.3 $gCm^{-2}yr^{-1}$ estimated by the EC metrics-based model, compared with 3.9 $gCm^{-2}yr^{-1}$, and 1.9 $gCm^{-2}yr^{-1}$ estimated by the LSP metrics-based model. While the EC metrics-based model identified a substantial GSL effect on ΣGPP in Wetlands (WET) (7.67 $gCm^{-2}yr^{-1}$), the LSP metrics-based model found almost no effect of GSL on GPP. In general, the LSP metrics-based model had a smaller

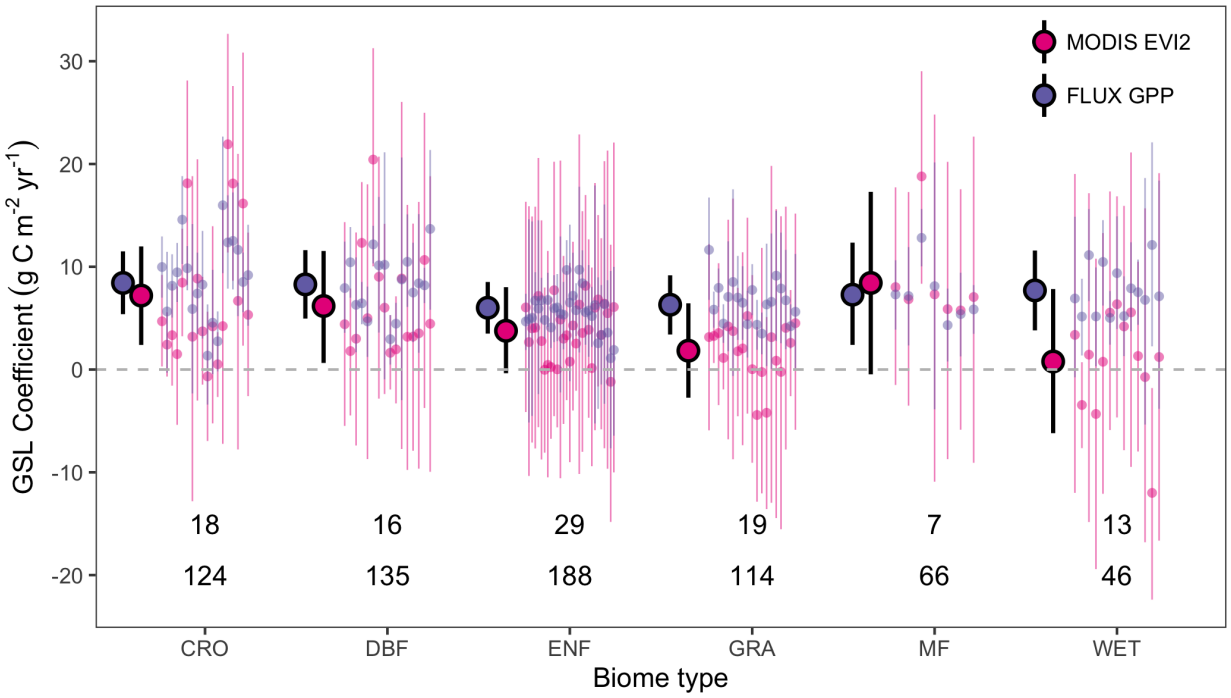

Figure 5: Growing season length coefficients for each site and biome type from the EC metrics- and LSP metrics-based models. The numbers at the bottom of the figure show the number of sites (the first row) and number of site-years (the second row) for each biome type. Biome types are cropland (CRO), deciduous broadleaf forest (DBF), evergreen needleleaf forest (ENF), grassland (GRA), mixed forest (MF), and wetland (WET). Vertical lines show Bayesian 95 percent credible intervals.

estimated GSL effect on ΣGPP and larger uncertainty ranges for most biome types compared to the EC metrics-based model. The magnitude of the differences regarding the role of GSL between EC metrics- versus LSP metrics-based models varies among different places on Earth depending on their dominant vegetation types.

Relative to the EC metrics-based models, the LSP metrics-based models showed greater sensitivity to site-level variability in ΣGPP. In the EC metrics-based models, the biome-level model explained the large majority of variance in ΣGPP ($R^2 = 93\%$), and accounting for site-level variability provided only modest improvement ($R^2 = 98\%$) (Fig. 4a, b). Indeed, cross-validation experiments indicate that accounting for site-level variability did not significantly improve the model (Fig. S6). In the LSP metrics-based models, however, inclusion of site-level variability increased the proportion of explained variance in ΣGPP from 55% to 84% (Fig. 4c,

d), with similar results achieved in cross-validation experiments (Fig. S6). Stated another way, ΣGPP modeled using LSP-derived phenology metrics is more sensitive to site-to-site variation in phenological metrics than corresponding metrics and models based on EC measurements. In fact, inclusion of site-specific intercepts explained the largest proportion of variance in ΣGPP in the LSP metrics-based models (Fig. S6). This suggests that phenological and physiological metrics derived from LSP observations do not capture differences in overall productivity across the EC sites included in this analysis; i.e., sites with the same GSL, $EVI2_{max}$, and $EVI2_{min}$ derived from LSP metrics can have significantly different ΣGPP.

The EC metrics- and LSP metrics-based models indicate that ΣGPP is more sensitive to GPP_{max} and $EVI2_{max}$, respectively, than GSL (Fig. 4e). Normalized GPP_{max} and $EVI2_{max}$ effects were both about 60% larger than normalized GSL effects. Holding GSL and GPP_{min} constant, an increase in GPP_{max} of one standard deviation increases ΣGPP by 0.77 (0.58 to 0.94) standard deviations. The corresponding sensitivity in the LSP-based model was 0.43 standard deviations (0.17 to 0.70). Similar to GSL, biome-level variability in ΣGPP associated with variability in GPP_{max} and $EVI2_{max}$ was higher in the EC metrics-based model. In fact, $EVI2_{max}$, which represents maximum leaf greenness was significantly related to the GPP_{max} at each site (Fig. S7). At the same time, variance in GPP_{max} increases with $EVI2_{max}$ (Fig. S7), which suggests that while maximum leaf greenness is a good indicator of mean maximum vegetation productivity, other factors exert substantial control on GPP_{max} at local scale.

GPP_{min} and $EVI2_{min}$ play a modest role in regulating ΣGPP uptake in most, but not all, biome types (Fig. 4e). The normalized coefficient of GPP_{min} ranged from -0.03 to 0.27 (40% less than the normalized GPP GSL effect) and normalized coefficient of $EVI2_{min}$ ranged from -0.07 to 0.33 (20% less than the normalized EVI2 GSL effect). The effects of both variables are slightly lower than 95% significance level (Amrhein et al., 2019) based on our data, but ignoring the minimum seasonal productivity or aggregating maximum and minimum

311 metrics into the seasonal amplitude may obscure important factors that are diagnostic of
312 total seasonal carbon update, especially in evergreen systems which have higher minimum
313 greenness and smaller greenness amplitude.

314 4 Discussion

315 The result that $EVI2_{max}$ had a larger effect on ΣGPP suggests that increases in maximum
316 leaf greenness alters ΣGPP more than increases in the growing season duration of green leaves.
317 Previous studies have shown that while regional decreases in leaf greenness are present in the
318 satellite record (Jong et al., 2012; Sulla-Menashe et al., 2018), so-called 'greening' of global
319 vegetated land areas has been ongoing since at least the early 1980's (C. Chen et al., 2019;
320 Huang et al., 2018). Satellite observations and ecological models suggest that this greening
321 is diagnostic of enhanced terrestrial vegetation productivity, and has potential to mitigate
322 climate warming by increasing the terrestrial carbon sink (Piao et al., 2020). However,
323 a variety of studies have also suggested that increases in early and mid-growing season
324 productivity can negatively impact end-of-season GPP, effectively offsetting early season
325 increases in GPP (Buermann et al., 2018; Piao et al., 2008; Zani et al., 2020). Therefore,
326 improved understanding of how changes in leaf greenness and GSL jointly impact ΣGPP
327 is required to forecast future change in large-scale carbon budgets. Our results (Fig. 4e)
328 indicate that maximum greenness increasing caused greening trends might have a larger
329 impact on net carbon uptake of terrestrial vegetation than changes in growing season length
330 of leaf development.

331 Our results showing that satellite LSP-derived metrics had a smaller GSL effect on ΣGPP
332 compared to EC-based metrics might have important implications for the use of remotely
333 sensed LSP metrics to infer vegetation productivity at regional, continental, and global scales
334 (e.g., Keenan et al., 2014; Richardson et al., 2010; Richardson et al., 2013). The smaller
335 magnitude of green leaf based GSL effect on ΣGPP has the potential to bias understanding

regarding if and how changes in the satellite LSP-derived growing season of terrestrial ecosystems impact the sign and magnitude of net carbon fluxes. Future warming is expected to extend both the leaf and photosynthetic GSLs in many ecosystems, thereby potentially increasing ΣGPP (Hua et al., 2021; Piao et al., 2019). However, our results suggest that leaf GSL changes had smaller effect on ΣGPP than photosynthetic GSL changes, but the extended leaf GSL in spring and autumn (Buermann et al., 2018; Piao et al., 2008; Piao et al., 2007; Wolf et al., 2016) might increase carbon loss by ecosystem respiration, and thus reduce the total net carbon uptake.

Differences in the timing of phenology from LSP and EC measurements may explain differences in model results from each data source. Our comparisons of phenophase transition dates derived from LSP and EC measurements of GPP are broadly consistent with prior work (D'Odorico et al., 2015; Lu et al., 2018; Shen et al., 2014), but are based on a much larger data set that supports additional and more nuanced interpretation. First, at deciduous sites, the timing of autumn phenology in LSP measurements was biased late compared to the timing estimated from EC measurements (Fig. 3b, 6a). The reasons for this are unclear, but this result almost certainly reflects complexity in the relationship between the timing of leaf coloration and decline in photosynthesis late in the growing season (X. Wang et al., 2020). As a consequence, LSP-derived leaf GSL was systematically longer than EC-derived photosynthetic GSL (Fig. S3, Fig. S4), which explains why the LSP-based model showed smaller GSL effect on ΣGPP for deciduous sites. Second, at evergreen sites, the timing of spring phenology from LSP measurements is biased late relative to corresponding timing from EC measurements (Fig. 3a, 6d), and the timing of autumn phenology from the two sources was only weakly correlated (Fig. 3, Table 1). This result has been previously noted (e.g., Melaas et al., 2013) and arises from the fact that photosynthesis in conifers starts well before the timing of leaf flushing and pigment changes later in the spring (Barr et al., 2009; Gao et al., 2021). These differences yielded shorter leaf GSL from LSP measurements relative to photosynthetic GSL from EC measurements and large site-level uncertainty for evergreen

363 vegetation observed by satellite LSP observations compared to EC measurements (Fig. S3,
364 Fig. S4). These results highlight the importance of developing methods to better match
365 remotely sensed phenology with vegetation photosynthetic activities such as solar-induced
366 chlorophyll fluorescence (SIF) and better vegetation indices (Gonsamo et al., 2012; Jin and
367 Eklundh, 2014; Mohammed et al., 2019).

368 The result that site-level variability contributes a substantial proportion of total variance
369 in ΣGPP modeled by LSP metrics across sites and years is consistent with Butterfield
370 et al., 2020; Richardson et al., 2010; Richardson et al., 2013, who found that the remotely
371 sensed phenology-productivity relationship was strong across flux sites but does not capture
372 interannual variability in ΣGPP at individual sites. Local environmental factors such as
373 temperature, precipitation, forest age, and soil moisture are more important regulators of
374 GPP than leaf phenology and physiology at fine spatial scales (Barr et al., 2009; Churkina
375 et al., 2005; Piao et al., 2009; Richardson et al., 2010). However, when investigating ΣGPP
376 variability across large spatial scales, we found these local environmental factors tended to be
377 averaged out, so the effects of remotely sensed leaf phenology and physiology on ΣGPP were
378 stabilized. In addition, factors complicating the relationship between GPP measurements
379 and remotely sensed LSP metrics also contribute to the site-level variability (X. Chen et al.,
380 2018; Peng et al., 2017; Zhang et al., 2017). Further, EC measurements are affected by
381 site-specific characteristics such as wind direction and measurement height (Chu et al., 2021;
382 Schmid, 2002), factors that cannot be captured by satellite LSP observations. Thus, the
383 magnitude of the estimated GSL-ΣGPP relationship at any particular site depends on
384 both the natural variability of the relationship and the interaction with local characteristics.
385 Our results support the conclusion that it is feasible to infer large-scale spatio-temporal
386 patterns in ΣGPP from satellite-observed leaf GSL, but large uncertainty at fine spatial scales.
387 Developing ways to explain this site-level variability, perhaps using ecological covariates, has
388 the potential to substantially improve our models designed to infer large scale ΣGPP using
389 satellite LSP observations.

5 Conclusion

In summary, this study suggests that satellite LSP-based green leaf phenological and physiological metrics are capable of inferring vegetation productivity over large spatial areas for most biome types, and satellite observed leaf GSL trends are meaningful for projecting carbon cycle impacts into the future. However, caution must be used as satellite observed leaf GSL changes do not synchronize photosynthetic GSL changes for evergreen vegetation in spring and deciduous vegetation in autumn. Changes in leaf GSL had a smaller effect on ΣGPP compared to changes in photosynthetic GSL. Moreover, although changes in leaf GSL have a significant impact on ΣGPP, trends of vegetation greening or browning indicated by maximum leaf greenness changes might have more carbon impacts than the extended leaf GSL caused by current climate warming. Therefore, changes in both leaf GSL and maximum greenness need to be considered in satellite LSP-based inferences regarding large-scale dynamics of vegetation productivity.

Acknowledgements

This work was funded by the NASA grant #80NSSC18K0334 (*An Operational Multisource Land Surface Phenology Product from Landsat and Sentinel 2, PI: Mark Friedl*). We gratefully acknowledge the support and excellent work of the MuSLI team at Boston University and Lund University who provide good suggestions for this work. We would also like to thank the FLUXNET2015 team for providing reliable flux tower measurements, otherwise this analysis would not be possible.

Code availability

All R scripts used to produce this study will be released on GitHub once the manuscript is accepted.

References

Amrhein, V., Greenland, S., & McShane, B. (2019). Retire statistical significance. *Nature*, *567*, 3.

Baldocchi, D. D. (2020). How eddy covariance flux measurements have contributed to our understanding of *Global Change Biology*. *Global Change Biology*, *26*(1), 242–260. https://doi.org/10.1111/gcb.14807

Barr, A., Black, T. A., & McCaughey, H. (2009). Climatic and Phenological Controls of the Carbon and Energy Balances of Three Contrasting Boreal Forest Ecosystems in Western Canada. In A. Noormets (Ed.), *Phenology of Ecosystem Processes: Applications in Global Change Research* (pp. 3–34). Springer. https://doi.org/10.1007/978-1-4419-0026-5_1

Buermann, W., Forkel, M., O'Sullivan, M., Sitch, S., Friedlingstein, P., Haverd, V., Jain, A. K., Kato, E., Kautz, M., Lienert, S., Lombardozzi, D., Nabel, J. E. M. S., Tian, H., Wiltshire, A. J., Zhu, D., Smith, W. K., & Richardson, A. D. (2018). Widespread seasonal compensation effects of spring warming on northern plant productivity. *Nature*, *562*(7725), 110–114. https://doi.org/10.1038/s41586-018-0555-7

Buitenwerf, R., Rose, L., & Higgins, S. I. (2015). Three decades of multi-dimensional change in global leaf phenology. *Nature Climate Change*, *5*(4), 364–368. https://doi.org/10.1038/nclimate2533

Butterfield, Z., Buermann, W., & Keppel-Aleks, G. (2020). Satellite observations reveal seasonal redistribution of northern ecosystem productivity in response to interannual climate variability [ZSCC: 0000005]. *Remote Sensing of Environment*, *242*, 111755. https://doi.org/10.1016/j.rse.2020.111755

Caparros-Santiago, J. A. (2021). Land surface phenology as indicator of global terrestrial ecosystem dynamics: A systematic review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 18.

Chen, C., Park, T., Wang, X., Piao, S., Xu, B., Chaturvedi, R. K., Fuchs, R., Brovkin, V., Ciais, P., Fensholt, R., Tømmervik, H., Bala, G., Zhu, Z., Nemani, R. R., & Myneni, R. B. (2019). China and India lead in greening of the world through land-use management [Bandiera_abtest: a Cg_type: Nature Research Journals Number: 2 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Agriculture;Environmental sciences;Forestry;Geography Subject_term_id: agriculture;environmental-sciences;forestry;geography]. *Nature Sustainability*, *2*(2), 122–129. https://doi.org/10.1038/s41893-019-0220-7

Chen, X., Wang, D., Chen, J., Wang, C., & Shen, M. (2018). The mixed pixel effect in land surface phenology: A simulation study. *Remote Sensing of Environment*, *211*, 338–344. https://doi.org/10.1016/j.rse.2018.04.030

Chu, H., Luo, X., Ouyang, Z., Chan, W. S., Dengel, S., Biraud, S. C., Torn, M. S., Metzger, S., Kumar, J., Arain, M. A., Arkebauer, T. J., Baldocchi, D., Bernacchi, C., Billesbach, D., Black, T. A., Blanken, P. D., Bohrer, G., Bracho, R., Brown, S., … Zona, D. (2021). Representativeness of Eddy-Covariance flux footprints for areas surrounding AmeriFlux sites. *Agricultural and Forest Meteorology*, *301-302*, 108350. https://doi.org/10.1016/j.agrformet.2021.108350

Churkina, G., Schimel, D., Braswell, B. H., & Xiao, X. (2005). Spatial analysis of growing season length control over net ecosystem exchange. *Global Change Biology*, *11*(10), 1777–1787.

D'Odorico, P., Gonsamo, A., Gough, C. M., Bohrer, G., Morison, J., Wilkinson, M., Hanson, P. J., Gianelle, D., Fuentes, J. D., & Buchmann, N. (2015). The match and mismatch between photosynthesis and land surface phenology of deciduous forests. *Agricultural and Forest Meteorology*, *214-215*, 25–38. https://doi.org/10.1016/j.agrformet.2015.07.005

Dragoni, D., Schmid, H. P., Wayson, C. A., Potter, H., Grimmond, C. S. B., & Randolph, J. C. (2011). Evidence of increased net ecosystem productivity associated

with a longer vegetated season in a deciduous forest in south-central Indiana, USA [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2486.2010.02281.x]. *Global Change Biology*, *17*(2), 886–897. https://doi.org/10.1111/j.1365-2486.2010.02281.x

Falge, E., Baldocchi, D., Tenhunen, J., Aubinet, M., Bakwin, P., Berbigier, P., Bernhofer, C., Burba, G., Clement, R., Davis, K. J., Elbers, J. A., Goldstein, A. H., Grelle, A., Granier, A., Guðmundsson, J., Hollinger, D., Kowalski, A. S., Katul, G., Law, B. E., … Wofsy, S. (2002). Seasonality of ecosystem respiration and gross primary production as derived from FLUXNET measurements. *Agricultural and Forest Meteorology*, *113*(1-4), 53–74. https://doi.org/10.1016/S0168-1923(02)00102-8

Friedl, M., J., G., & D, S.-M. (2019). MCD12Q2 MODIS/Terra+Aqua Land Cover Dynamics Yearly L3 Global 500m SIN Grid V006. *NASA EOSDIS Land Processes DAAC*. https://doi.org/https://doi.org/10.5067/MODIS/MCD12Q2.006.Accessed2020-01-23

Ganguly, S., Friedl, M. A., Tan, B., Zhang, X., & Verma, M. (2010). Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product. *Remote Sensing of Environment*, *114*(8), 1805–1816. https://doi.org/10.1016/j.rse.2010.04.005

Gao, X., Gray, J., Cohrs, C. W., Cook, R., & Albaugh, T. J. (2021). Longer greenup periods associated with greater wood volume growth in managed pine stands. *Agricultural and Forest Meteorology*, *297*, 108237. https://doi.org/10.1016/j.agrformet.2020.108237

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.

Gonsamo, A., Chen, J. M., Wu, C., & Dragoni, D. (2012). Predicting deciduous forest carbon uptake phenology by upscaling FLUXNET measurements using remote sensing data. *Agricultural and Forest Meteorology*, *165*, 127–135. https://doi.org/10.1016/j.agrformet.2012.06.006

Hua, X., Sirguey, P., & Ohlemüller, R. (2021). Recent trends in the timing of the growing season in New Zealand's natural and semi-natural grasslands [ISBN: 1548-1603 Publisher: Taylor & Francis]. *GIScience & Remote Sensing*, 1–21.

Huang, K., Xia, J., Wang, Y., Ahlström, A., Chen, J., Cook, R. B., Cui, E., Fang, Y., Fisher, J. B., Huntzinger, D. N., Li, Z., Michalak, A. M., Qiao, Y., Schaefer, K., Schwalm, C., Wang, J., Wei, Y., Xu, X., Yan, L., … Luo, Y. (2018). Enhanced peak growth of global vegetation and its key mechanisms. *Nature Ecology & Evolution*, *2*(12), 1897–1905. https://doi.org/10.1038/s41559-018-0714-0

Jiang, Z., Huete, A., Didan, K., & Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, *112*(10), 3833–3845. https://doi.org/10.1016/j.rse.2008.06.006

Jin, H., & Eklundh, L. (2014). A physically based vegetation index for improved monitoring of plant phenology. *Remote Sensing of Environment*, *152*, 512–525. https://doi.org/10.1016/j.rse.2014.07.010

Jong, R. d., Verbesselt, J., Schaepman, M. E., & Bruin, S. d. (2012). Trend changes in global greening and browning: Contribution of short-term trends to longer-term change [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2486.2011.02578.x]. *Global Change Biology*, *18*(2), 642–655. https://doi.org/10.1111/j.1365-2486.2011.02578.x

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., & Reichstein, M. (2019). The fluxcom ensemble of global land-atmosphere energy fluxes. *Scientific data*, *6*(1), 1–14.

Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., Munger, J. W., O'Keefe, J., Schmid, H. P., Wing, I. S., Yang, B., & Richardson, A. D. (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/nclimate2253

519 Le Quéré, C., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., Peters, G. P.,
520 Manning, A. C., Boden, T. A., Tans, P. P., Houghton, R. A., et al. (2016). Global
521 carbon budget 2016. *Earth System Science Data*, *8*(2), 605–649.
522 Liu, Y., Wu, C., Wang, X., Jassal, R. S., & Gonsamo, A. (2021). Impacts of global change on
523 peak vegetation growth and its timing in terrestrial ecosystems of the continental US.
524 *Global and Planetary Change*, *207*, 103657. https://doi.org/10.1016/j.gloplacha.2021.
525 103657
526 Lu, X., Liu, Z., Zhou, Y., Liu, Y., An, S., & Tang, J. (2018). Comparison of Phenology
527 Estimated from Reflectance-Based Indices and Solar-Induced Chlorophyll Fluorescence
528 (SIF) Observations in a Temperate Forest Using GPP-Based Phenology as the Standard
529 [Number: 6 Publisher: Multidisciplinary Digital Publishing Institute]. *Remote Sensing*,
530 *10*(6), 932. https://doi.org/10.3390/rs10060932
531 Melaas, E. K., Friedl, M. A., & Zhu, Z. (2013). Detecting interannual variation in decidu-
532 ous broadleaf forest phenology using Landsat TM/ETM+ data. *Remote Sensing of*
533 *Environment*, *132*, 176–185. https://doi.org/10.1016/j.rse.2013.01.011
534 Mohammed, G. H., Colombo, R., Middleton, E. M., Rascher, U., van der Tol, C., Nedbal, L.,
535 Goulas, Y., Pérez-Priego, O., Damm, A., Meroni, M., Joiner, J., Cogliati, S., Verhoef,
536 W., Malenovský, Z., Gastellu-Etchegorry, J.-P., Miller, J. R., Guanter, L., Moreno, J.,
537 Moya, I., ... Zarco-Tejada, P. J. (2019). Remote sensing of solar-induced chlorophyll
538 fluorescence (SIF) in vegetation: 50 years of progress. *Remote Sensing of Environment*,
539 *231*, 111177. https://doi.org/10.1016/j.rse.2019.04.030
540 Park, T., Ganguly, S., Tømmervik, H., Euskirchen, E. S., Høgda, K.-A., Karlsen, S. R.,
541 Brovkin, V., Nemani, R. R., & Myneni, R. B. (2016). Changes in growing season
542 duration and productivity of northern vegetation inferred from long-term remote
543 sensing data [ZSCC: NoCitationData[s0]]. *Environmental Research Letters*, *11*(8),
544 084001. https://doi.org/10.1088/1748-9326/11/8/084001

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Reichstein, M., Ribeca, A., van Ingen, C., Vuichard, N., Zhang, L., Amiro, B., Ammann, C., ... Papale, D. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific Data*, *7*(1), 225. https://doi.org/10.1038/s41597-020-0534-3

Peng, D., Zhang, X., Zhang, B., Liu, L., Liu, X., Huete, A. R., Huang, W., Wang, S., Luo, S., Zhang, X., & Zhang, H. (2017). Scaling effects on spring phenology detections from MODIS data at multiple spatial resolutions over the contiguous United States. *ISPRS Journal of Photogrammetry and Remote Sensing*, *132*, 185–198. https://doi.org/10.1016/j.isprsjprs.2017.09.002

Piao, S., Ciais, P., Friedlingstein, P., Peylin, P., Reichstein, M., Luyssaert, S., Margolis, H., Fang, J., Barr, A., Chen, A., Grelle, A., Hollinger, D. Y., Laurila, T., Lindroth, A., Richardson, A. D., & Vesala, T. (2008). Net carbon dioxide losses of northern ecosystems in response to autumn warming. *Nature*, *451*(7174), 49–52. https://doi.org/10.1038/nature06444

Piao, S., Friedlingstein, P., Ciais, P., Peylin, P., Zhu, B., & Reichstein, M. (2009). Footprint of temperature changes in the temperate and boreal forest carbon balance. *Geophysical Research Letters*, *36*(7). https://doi.org/10.1029/2009GL037381

Piao, S., Friedlingstein, P., Ciais, P., Viovy, N., & Demarty, J. (2007). Growing season extension and its impact on terrestrial carbon cycle in the Northern Hemisphere over the past 2 decades: PHENOLOGY AND CARBON CYCLE IN NH. *Global Biogeochemical Cycles*, *21*(3), n/a–n/a. https://doi.org/10.1029/2006GB002888

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., & Zhu, X. (2019). Plant phenology and global climate change: Current progresses and challenges. *Global Change Biology*, *25*(6), 1922–1940. https://doi.org/10.1111/gcb.14619

Piao, S., Wang, X., Park, T., Chen, C., Lian, X., He, Y., Bjerke, J. W., Chen, A., Ciais, P., Tømmervik, H., Nemani, R. R., & Myneni, R. B. (2020). Characteristics, drivers

and feedbacks of global greening. *Nature Reviews Earth & Environment*, *1*(1), 14–27. https://doi.org/10.1038/s43017-019-0001-x

Richardson, A. D., Andy Black, T., Ciais, P., Delbart, N., Friedl, M. A., Gobron, N., Hollinger, D. Y., Kutsch, W. L., Longdoz, B., Luyssaert, S., Migliavacca, M., Montagnani, L., William Munger, J., Moors, E., Piao, S., Rebmann, C., Reichstein, M., Saigusa, N., Tomelleri, E., ... Varlagin, A. (2010). Influence of spring and autumn phenological transitions on forest ecosystem productivity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1555), 3227–3246. https://doi.org/10.1098/rstb.2010.0102

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 [ZSCC: 0001267]. *Agricultural and Forest Meteorology*, *169*, 156–173. https://doi.org/10.1016/j.agrformet.2012.09.012

Schaaf, C., & Wang, Z. (2015). Mcd43a4 modis/terra+ aqua brdf/albedo nadir brdf adjusted ref daily l3 global-500m v006. nasa eosdis land processes daac. *USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota (https://lpdaac.usgs. gov)*.

Schmid, H. P. (2002). Footprint modeling for vegetation atmosphere exchange studies: A review and perspective. *Agricultural and Forest Meteorology*, *113*(1), 159–183. https://doi.org/10.1016/S0168-1923(02)00107-7

Shen, M., Tang, Y., Desai, A. R., Gough, C., & Chen, J. (2014). Can EVI-derived land-surface phenology be used as a surrogate for phenology of canopy photosynthesis? *International Journal of Remote Sensing*, *35*(3), 1162–1174. https://doi.org/10.1080/01431161.2013.875636

Sulla-Menashe, D., Woodcock, C. E., & Friedl, M. A. (2018). Canadian boreal forest greening and browning trends: An analysis of biogeographic patterns and the relative roles of

disturbance versus climate drivers. *Environmental Research Letters*, *13*(1), 014007. https://doi.org/10.1088/1748-9326/aa9b88

Wang, R., Chen, J. M., Luo, X., Black, A., & Arain, A. (2019). Seasonality of leaf area index and photosynthetic capacity for better estimation of carbon and water fluxes in evergreen conifer forests. *Agricultural and Forest Meteorology*, *279*, 107708.

Wang, X., Dannenberg, M. P., Yan, D., Jones, M. O., Kimball, J. S., Moore, D. J. P., Leeuwen, W. J. D., Didan, K., & Smith, W. K. (2020). Globally Consistent Patterns of Asynchrony in Vegetation Phenology Derived From Optical, Microwave, and Fluorescence Satellite Data. *Journal of Geophysical Research: Biogeosciences*, *125*(7). https://doi.org/10.1029/2020JG005732

Wolf, S., Keenan, T. F., Fisher, J. B., Baldocchi, D. D., Desai, A. R., Richardson, A. D., Scott, R. L., Law, B. E., Litvak, M. E., & Brunsell, N. A. (2016). Warm spring reduced carbon cycle impact of the 2012 US summer drought. *Proceedings of the National Academy of Sciences*, *113*(21), 5880–5885.

Xia, J., Niu, S., Ciais, P., Janssens, I. A., Chen, J., Ammann, C., Arain, A., Blanken, P. D., Cescatti, A., Bonal, D., Buchmann, N., Curtis, P. S., Chen, S., Dong, J., Flanagan, L. B., Frankenberg, C., Georgiadis, T., Gough, C. M., Hui, D., ... Luo, Y. (2015). Joint control of terrestrial gross primary productivity by plant phenology and physiology. *Proceedings of the National Academy of Sciences*, *112*(9), 2788–2793. https://doi.org/10.1073/pnas.1413090112

Zani, D., Crowther, T. W., Mo, L., Renner, S. S., & Zohner, C. M. (2020). Increased growing-season productivity drives earlier autumn leaf senescence in temperate trees [ZSCC: 0000001]. *Science*, *370*(6520), 1066. https://doi.org/10.1126/science.abd8911

Zhang, X., Wang, J., Gao, F., Liu, Y., Schaaf, C., Friedl, M., Yu, Y., Jayavelu, S., Gray, J., Liu, L., Yan, D., & Henebry, G. M. (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

Zhou, S., Zhang, Y., Caylor, K. K., Luo, Y., Xiao, X., Ciais, P., Huang, Y., & Wang, G. (2016). Explaining inter-annual variability of gross primary productivity from plant phenology and physiology. *Agricultural and Forest Meteorology*, *226-227*, 246–256. https://doi.org/10.1016/j.agrformet.2016.06.010

Zhou, S., Zhang, Y., Ciais, P., Xiao, X., Luo, Y., Caylor, K. K., Huang, Y., & Wang, G. (2017). Dominant role of plant physiology in trend and variability of gross primary productivity in north america. *Scientific Reports*, *7*(1), 1–10.

Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., Canadell, J. G., Ciais, P., Sitch, S., Friedlingstein, P., Arneth, A., Cao, C., Cheng, L., Kato, E., Koven, C., Li, Y., Lian, X., Liu, Y., Liu, R., Mao, J., ... Zeng, N. (2016). Greening of the Earth and its drivers [Bandiera_abtest: a Cg_type: Nature Research Journals Number: 8 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Climate-change ecology;Ecological modelling Subject_term_id: climate-change-ecology;ecological-modelling]. *Nature Climate Change*, *6*(8), 791–795. https://doi.org/10.1038/nclimate3004

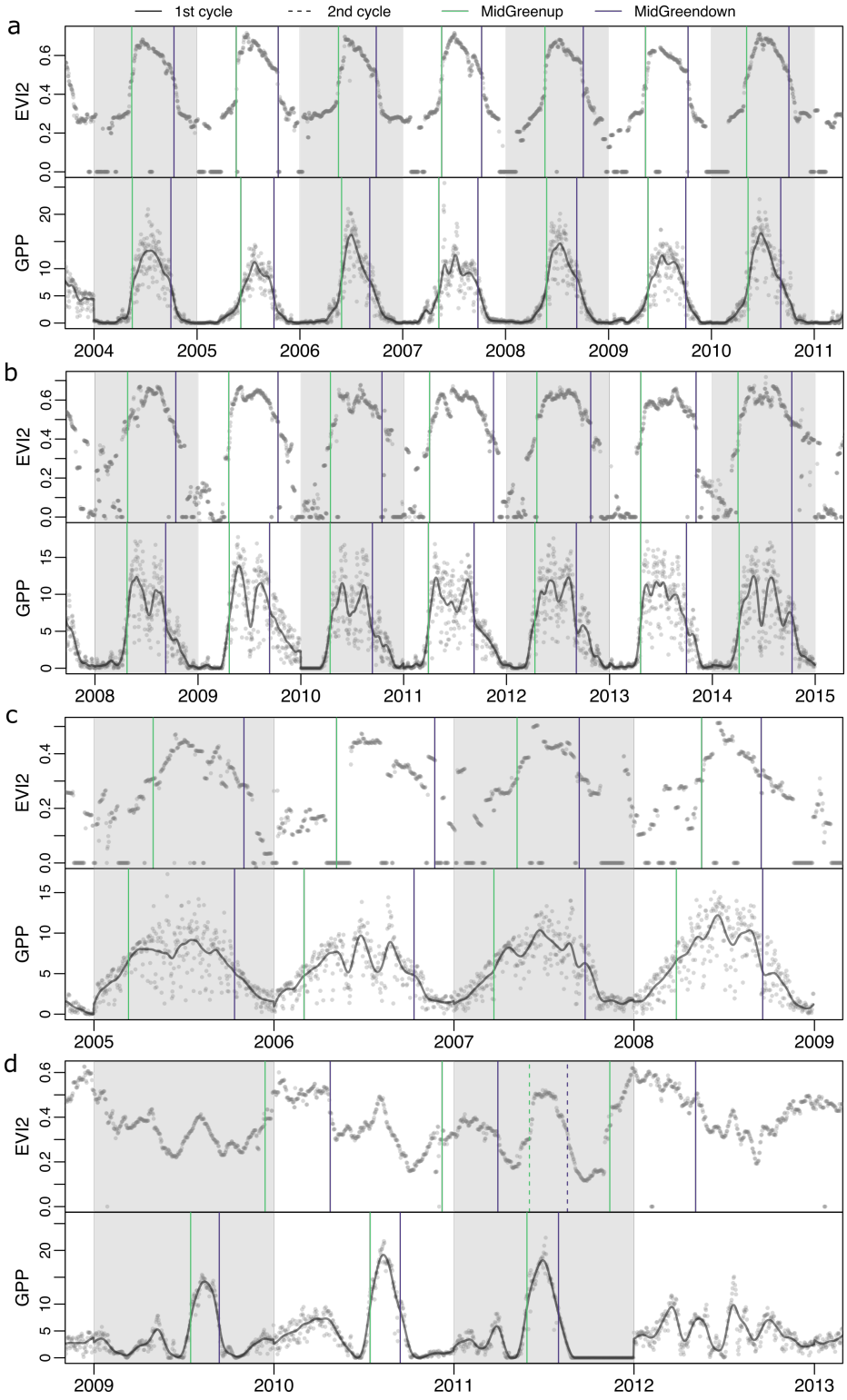

Figure 6: Representative GPP and EVI2 time series along with estimated MidGreenup and MidGreedown dates. Black lines in the GPP time series show fitted splines. (a) Deciduous broadleaf forest (DBF) site US-Ha1; (b) Grassland (GRA) site CH-Fru; (c) Evergreen needleleaf forest (ENF) site CZ-BK1; (d) Cropland (CRO) site IT-BCi. GPP values in $gCm^{-2}d^{-1}$

Observations of satellite land surface phenology suggest that maximum leaf greenness affects global vegetation productivity more than growing season length

Xiaojie Gao[1*], Ian R. McGregor[2], Josh M. Gray[3], Mark A. Friedl[4], and Minkyu Moon[5]

[1]Center for Geospatial Analytics, North Carolina State University, Raleigh, NC, 27695, USA.
[2]Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, 27695, USA
[3]Department of Earth and Environment, Boston University, USA
[*]Corresponding author. Email: xgao26@ncsu.edu

Supplementary figures and tables

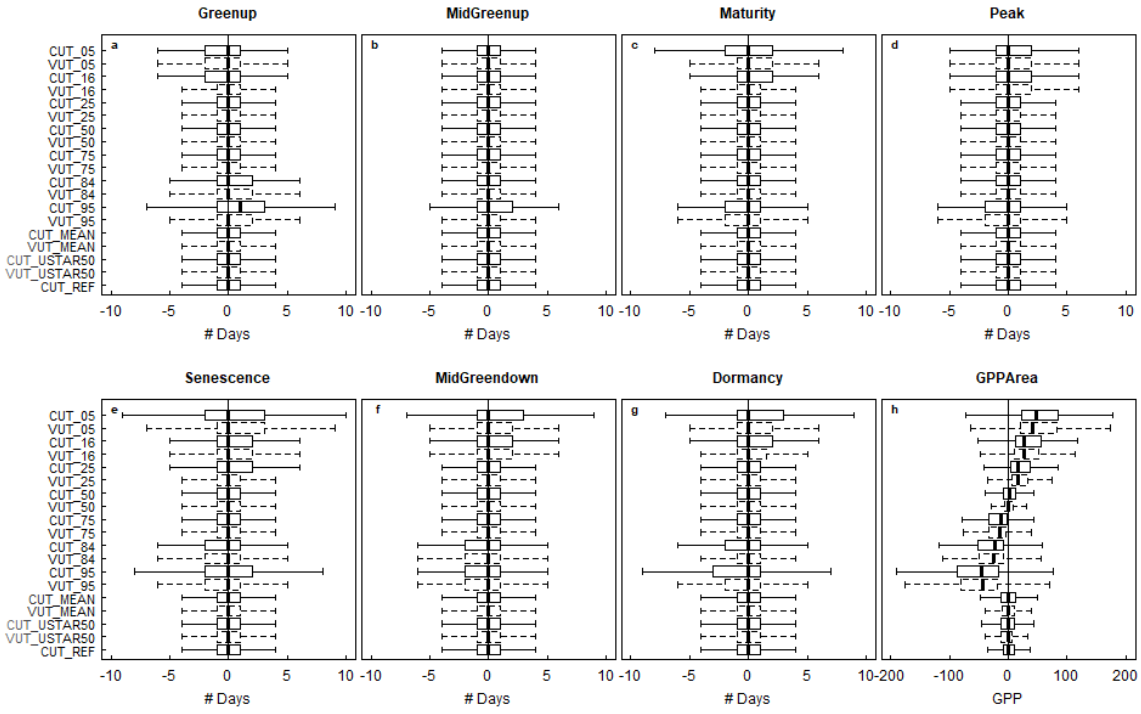

Figure 1: Sensitivity analysis of the different FLUXNET2015 GPP_DT variables. Each variable (y-axis) was compared against our focal variable, GPP_DT_VUT_REF using the difference in days between phenometrics (a-g) or GPP itself (h).

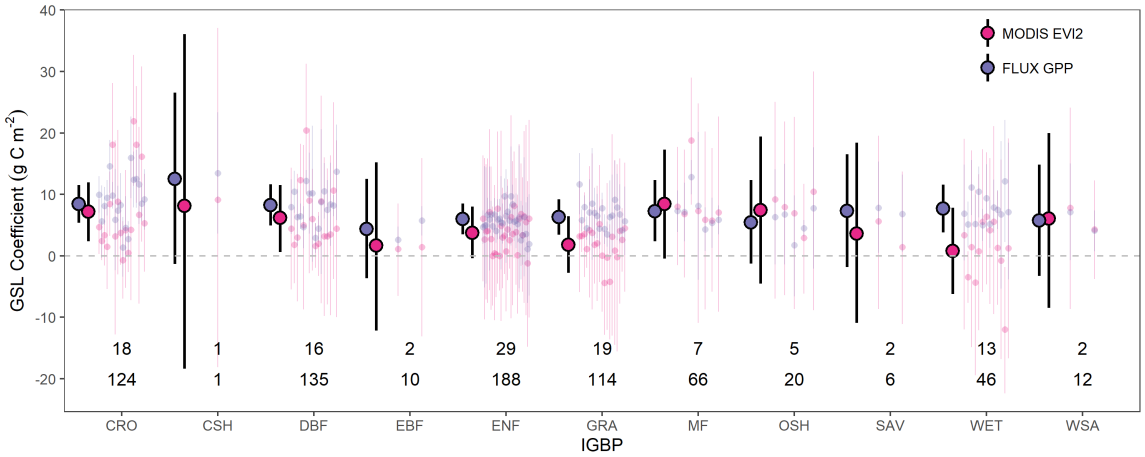

Figure 2: Growing season length coefficients for each site and biome type from the EC metrics- and LSP metrics-based models. The numbers at the bottom of the figure show the number of sites (the first row) and number of site-years (the second row) for each biome type. Biome types are cropland (CRO), deciduous broadleaf forest (DBF), evergreen needleleaf forest (ENF), grassland (GRA), mixed forest (MF), and wetland (WET). Vertical lines show Bayesian 95 percent credible intervals.

Table 1: Regression results for Figure 5

	Phenometric	Slope	Intercept	R2	MRD	MAD
1	Greenup_ever	0.54 (±0.13)	24.64 (±14.94)	0.24	-25.14	32.55
2	Greenup_deci	0.77 (±0.05)	24.28 (±4.86)	0.67	4.21	24.17
3	Mid-Greenup_ever	1 (±0.15)	-23.2 (±21.07)	0.47	-22.81	23.39
4	Mid-Greenup_deci	0.83 (±0.03)	23.12 (±4.93)	0.78	1.82	15.36
5	Maturity_ever	0.77 (±0.16)	28.03 (±27.62)	0.31	-11.14	16.6
6	Maturity_deci	0.81 (±0.03)	31.65 (±5.68)	0.78	2.34	15.39
7	Peak_ever	0.61 (±0.17)	66.29 (±32.38)	0.21	-7.61	15.86
8	Peak_deci	0.77 (±0.03)	40.68 (±6.91)	0.73	-0.16	16.97
9	Senescence_ever	0.6 (±0.19)	79.54 (±40.55)	0.17	-5.55	16.29
10	Senescence_deci	0.78 (±0.03)	41.15 (±7.92)	0.73	-3.65	18.81
11	Mid-Greendown_ever	0.51 (±0.16)	126.42 (±42.26)	0.16	-3.32	16.84
12	Mid-Greendown_deci	0.76 (±0.03)	46.16 (±8.45)	0.79	-14.03	23.3
13	Dormancy_ever	0.3 (±0.1)	210.03 (±31)	0.14	-2.6	21.23
14	Dormancy_deci	0.84 (±0.04)	28.59 (±10.83)	0.78	-16.59	24.65

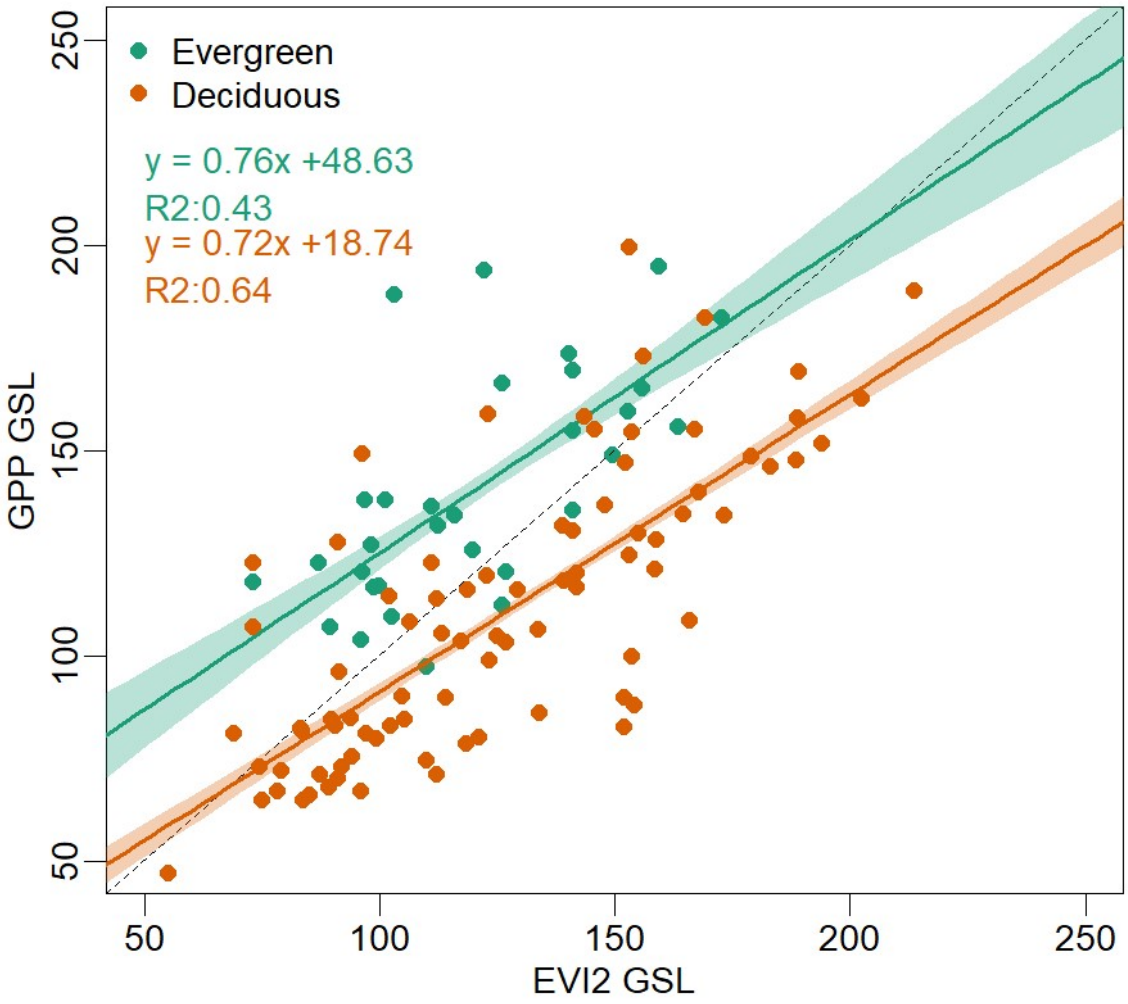

Figure 3: Comparison of multi-year mean growing season length for eddy-covariance flux site locations derived from MODIS EVI2 time series and FLUX GPP time series.

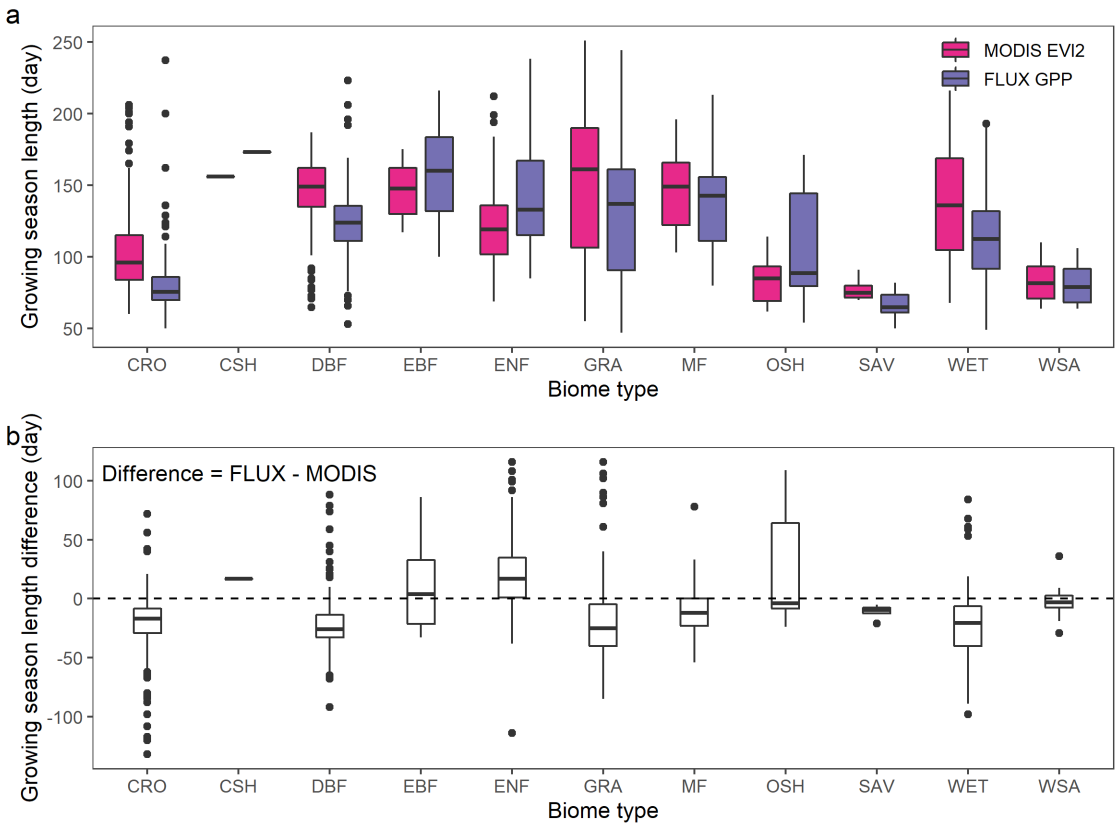

Figure 4: The growing season length derived from MODIS EVI2 time series and FLUX GPP time series. (a) Growing season length distribution for each biome type; (b) The difference between growing season length derived from the two sources of observations.

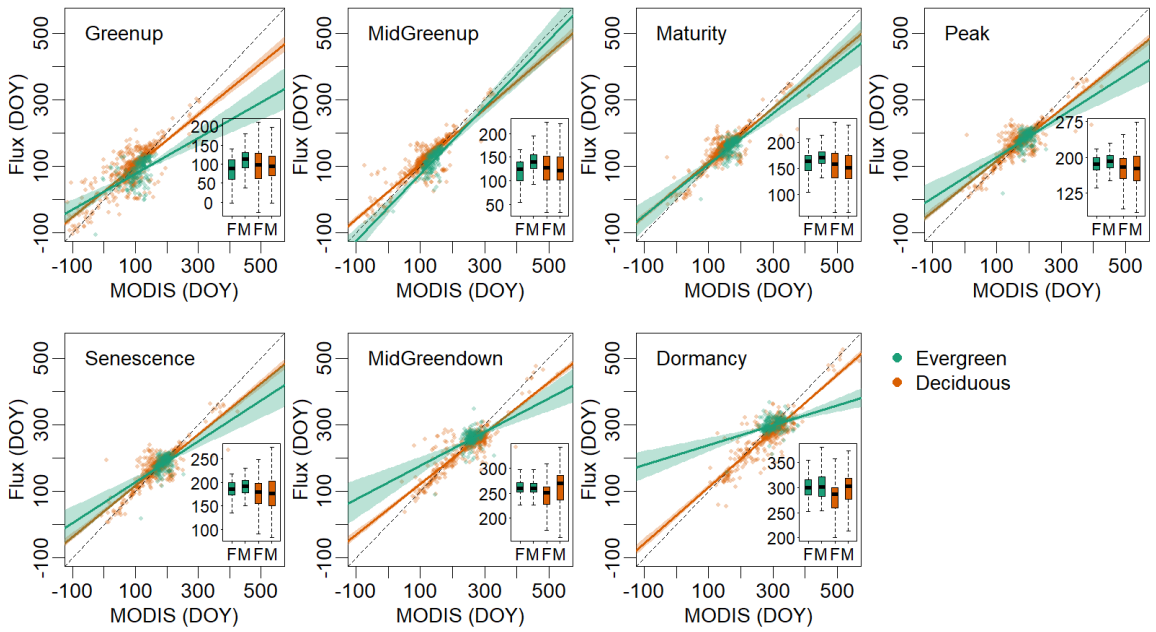

Figure 5: The comparison of phenometrics derived from MODIS EVI2 time series and Flux GPP time series.

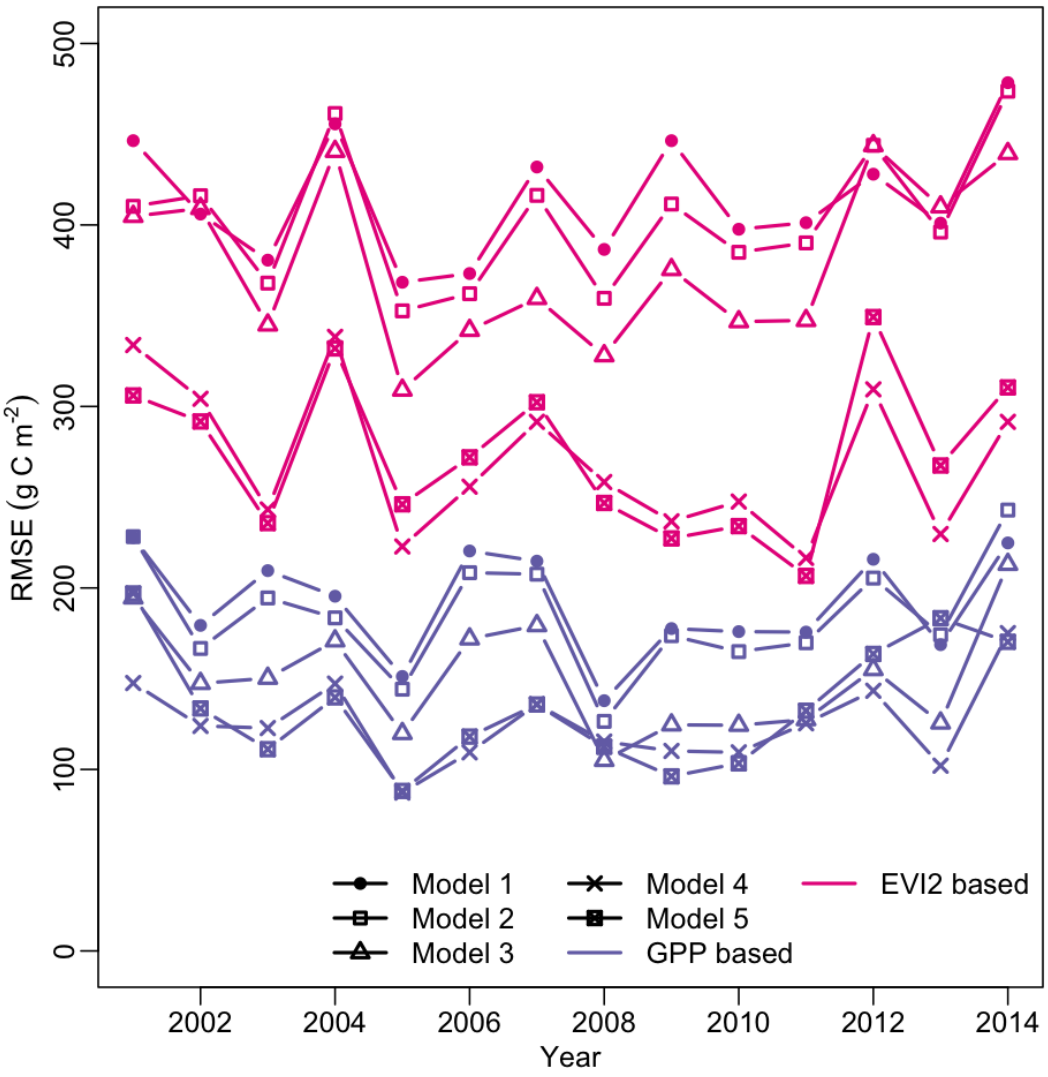

Figure 6: Leave-one-year-out cross validation results of EC metrics- and LSP metrics-based models representing different assumptions. Model 1 pools all data together; Model 2 considers biome-level intercepts; Model 3 considers biome-level intercepts and slopes; Model 4 considers site-level intercepts and biome-level slopes; Model 5 considers both site-level intercepts and site-level slopes.

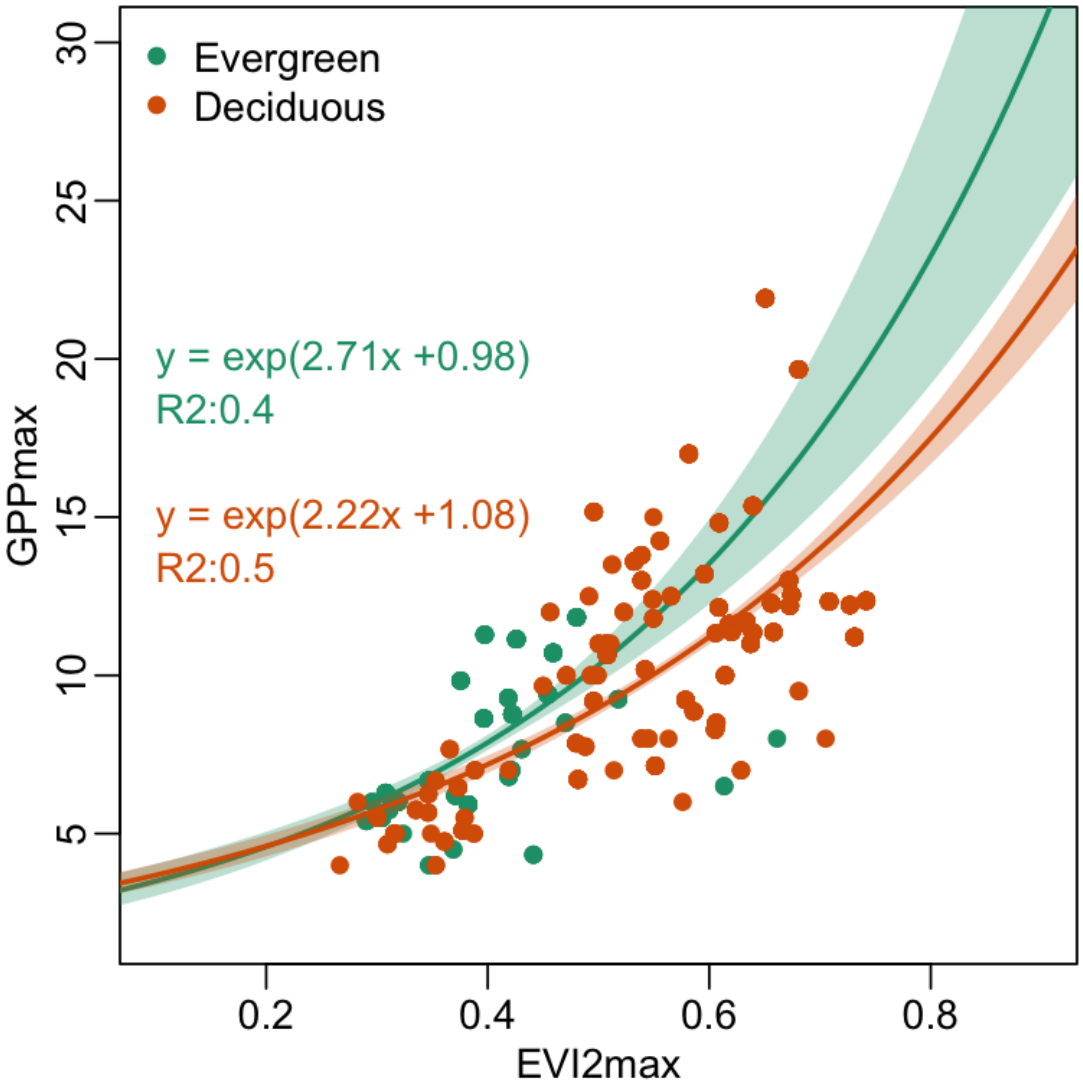

Figure 7: The relationship between flux site mean EVI2 maximum and GPP maximum across years. Lines show linear regression results and polygons show 95% confidence intervals.