Modeling growing season phenology in North American forests using seasonal mean vegetation indices from MODIS

Chaoyang Wu, Alemu Gonsamo, Christopher M. Gough, Jing M. Chen, Shiguang Xu

Abstract

The phenology of vegetation exerts an important control over the terrestrial ecosystem carbon (C) cycle. Remote sensing of key phenological phases in forests (e.g., the spring onset and autumn end of growing season) remains challenging due to noise in time series and the limited seasonal variation of canopy greenness in evergreen forests. Using 94 site-years of C flux data from four deciduous broadleaf forests (DBF) and six evergreen needleleaf forests (ENF) in North America, we examine whether growing season phenology can be remotely sensed from mean vegetation indices (VIs) derived from spring (Apr.–May) and autumn (Sep.–Nov) observations. Five VIs were used based on Moderate Resolution Imaging Spectroradiometer (MODIS) data, including the normalized difference vegetation index (NDVI), the land surface water index (LSWI), the enhanced vegetation index (EVI), the wide dynamic range vegetation index (WDRVI) and the optimized soil-adjusted vegetation index (OSAVI). Our results show that growing season transitions can be inferred from mean seasonal VIs, though the different VIs varied in their predictive strength across sites and plant functional types. Widely used NDVI and EVI exhibited limited potential in tracking growing season phenology of ENF ecosystems, while indices sensitive to water (i.e., LSWI) or less influenced by soil (i.e., OSAVI) may have unrevealed powers in indicating phenological transitions. OSAVI was shown to be a strong predictor of the end of the growing season in ENF ecosystems, suggesting that this VI may offer a new strategy for modeling the phenology of ENF sites. We conclude that combinations of multiple indices may improve the remote sensing of land surface phenology, as evidenced by the good agreement between modeled and observed growing season transitions and its length in our evaluation.

1. Introduction

Phenology describes the timing of plant processes and is an important control on carbon (C) sequestration in terrestrial ecosystems (e.g., Richardson et al., 2013; Wu et al., 2013). Remote sensing data from satellites provide broad coverage of useful information on vegetation phenology for diverse ecosystems at various scales (Brown, de Beurs, & Marshall, 2012; Hmimina et al., 2013; Karlsen et al., 2008; Kross, Fernandes, Seaquist, & Beaubien, 2011; White et al., 2009). Growing season phenology from remote sensing is determined by detecting the seasonal dynamics of green vegetation using spectral signals from sensors on satellite images or flux towers (Garrity, Maurer, Mueller, Vogel, & Curtis, 2011; Melaas et al., 2013; Sonnentag et al., 2012).

Time series of satellite remote sensing observations, e.g., the Moderate Resolution Imaging Spectroradiometer (MODIS), are commonly used to derive phenological metrics, with several vegetation indices (VIs) having potential to indicate phenological transitions (Hmimina et al., 2013). Commonly used VIs include the normalized difference vegetation index (NDVI), the enhanced vegetation index (EVI), the wide dynamic range vegetation index (WDRVI) and the optimized soil-adjusted vegetation index (OSAVI).

Keywords: Remote sensing, Phenology, Growing season transitions, MODIS, Forest

Article info

Article history:
Received 20 August 2013
Received in revised form 3 November 2013
Accepted 7 March 2014
Available online 22 March 2014

Keywords:
Remote sensing, Phenology, Growing season transitions, MODIS, Forest
Though substantial advances have been made in predicting growing season phenology from remote sensing data, challenges remain that limit the predictive strength of these tools. First, most contemporary methods require remote sensing time series data at fine temporal scales (from daily to 16 day), which increases the burden of data collection and noise removal (Atkinson et al., 2012). Second, most analyses were conducted in deciduous forests or crop ecosystems, but the same methods are sometimes limited when applied to evergreen species that show little seasonal variation in canopy greenness (Guyon et al., 2011; Hufkens et al., 2012; Melaas, Friedl, & Zhu, 2013). For example, Hnimina et al. (2013) showed that MODIS is unable to accurately infer phenological patterns for evergreen forests. A third limitation is modeling the end of the growing season, particularly for evergreen forests, since the growing transition in autumn is difficult to detect (Richardson et al., 2013). Lastly, most previous studies examined the utility of a single index (e.g., either EVI or NDVI), but the potential of various VIs that focus on different canopy characteristics (e.g., greenness, water) remains unknown.

To address these issues, here we use 94 site-years of C flux data from 10 forest sites in North America to evaluate whether the growing season phenologies (i.e., onset of spring and end of autumn) of deciduous and evergreen forests can be predicted by several different spectral indices derived from MODIS data. The underlying assumption for an expected relationship between the phenological transition dates of growing season and seasonal mean VI is the reported impacts of spring and autumn seasonality on annual vegetation productivity. For example, interannual variations in spring (Black et al., 2000; Chen et al., 2003) and autumn (Piao et al., 2008; Wu et al., 2012) climate are important determinants of annual carbon sequestration. Our specific objectives are to: (1) determine whether longer, less data intensive mean seasonal signatures (i.e., mean VIs for spring and autumn) derived from MODIS data capture growing season phenology in deciduous broadleaf and evergreen needleleaf forests, (2) compare this method with the traditional approach of deriving phenological indicators from more temporally detailed seasonal patterns of EVI, and (3) compare the capabilities of different VIs in tracking phenology of canopy photosynthesis.

2. Materials and methods

2.1. Study sites

We identified ten forest ecosystems in North America flux networks (AmeriFlux and Fluxnet-Canada) that report at least 8 years of complete data with less than 20% gap-filled in each year. Data from these sites covered a variety of forest ecosystems in North America that are classified broadly into two plant functional types (PFT): deciduous broadleaf forests (DBF, n = 4 sites) and evergreen needleleaf forests (ENF, n = 6 sites) (Fig. 1). Detailed descriptions of these sites are given in Table 1.

2.2. Flux measurements and data processing

Half-hourly CO₂ fluxes were continuously measured at each site using the eddy-covariance technique (Baldocchi et al., 2001). Standard procedures were applied to partition net ecosystem exchange (NEE) into gross primary productivity (GPP) and total ecosystem respiration (Re) (Barr et al., 2004; Reichstein et al., 2005). Because these sites belong to two different regional flux networks within North America, gap-filling and NEE partitioning approaches differ. For Fluxnet-Canada sites, GPP and Re estimation, and gap-filling were conducted using the network’s standard approaches described in Barr et al. (2004). For the AmeriFlux sites, level-4 products were used which contain gap-filled and u*-filtered records of C fluxes at varying time intervals using the Artificial Neural Network (ANN) method (Papale & Valentini, 2003) and/or the Marginal Distribution Sampling (MDS) method (Reichstein et al., 2005).

Desai et al. (2008) suggested that though different decomposition techniques are applied to flux data from various flux networks, they generally have a moderate impact on modeled GPP (i.e., most methods tended to cluster on similar results to within 10%). Therefore, it is reasonable to use sites from multiple networks in this study, which also agrees with other flux analyses reported on sites from global Fluxnet dataset (Lasslop et al., 2010; Wu et al., 2013).

Fig. 1. Spatial distribution of ten flux tower sites of this study. DBF (■) and ENF (▲) represent deciduous broadleaf forest and evergreen needleleaf forest, respectively.
2.4. MODIS surface reflectance product and vegetation indices

We used the 8-day composite MODIS Surface Reflectance Atmospheric Correction Algorithm Product (MOD09A1, 500 m, collection 5) to derive five commonly used VIs. Surface reflectance is computed from seven MODIS Level 1B bands (centered at 648 nm, 858 nm, 470 nm, 555 nm, 1240 nm, 1640 nm, and 2130 nm, respectively) after the removal of cloud contaminated pixels (Vermote et al., 1997). Based on the geo-location information (latitude and longitude) of the flux tower sites, reflectance was extracted from 3 × 3 MODIS pixels (~1.5 km × 1.5 km) centered on the flux tower footprint (Wu, Niu, & Kuang, 2010). The five VIs adopted in this study to evaluate the performance of MODIS data in tracking growing season phenological transitions included NDVI, the land surface water index, (LSWI), EVI, the wide dynamic range vegetation index (WDRVI) and the optimized soil-adjusted vegetation index (OSAVI). Detailed descriptions of these indices are provided in Table 2.

Unlike previous studies that derived phenological metrics based on the seasonal pattern of a VI, we focused on mean seasonal values of these Vls for modeling growing season phenology. Since the growing season onset and end occur in spring and in autumn, respectively, we selected spring months (April–May, Wang et al., 2011) and autumn months (September–November, Wu, Gonsamo, et al., 2012) to calculate the mean values of each VI (i.e., mean values of Vls during the spring and autumn months were used to model growing season phenology).

2.5. Deriving phenological metrics from seasonal patterns of EVI

We also derived conventional phenological measures at each site using seasonal patterns of EVI for comparison with our seasonally averaged metrics. Logistic models are commonly used to determine the transition dates from local maxima and minima in the fitted function (Melas, Friedli, & Zhu, 2013; Zhang, Friedli, & Schaaf, 2006, Zhang et al., 2003). Recent analysis from Beck, Atzberger, Hegda, Johansen, and Skidmore (2006) suggested that a double logistic function can more accurately describe duration of the growing season by effectively removing outliers. Hird and MacDermid (2009) similarly showed that a double logistic function optimally reduces noise in remote sensing time series data (Gonsamo, Chen, Price, Kurz, & Wu, 2012). In this analysis, we used a double logistic function to fit the time series of EVI data for each year:

\[ f(t, x_1, x_2, x_3, x_4) = \frac{1}{1 + \exp \left( \frac{x_1 - t}{x_2} \right)} - \frac{1}{1 + \exp \left( \frac{x_3 - t}{x_4} \right)} \]  

where \( x_1 \) and \( x_3 \) represent the positions of the left (i.e., growing season onset date) and right (i.e., growing season end date) inflection points and \( x_2 \) and \( x_4 \) determine the rates of changes at these inflection points (Fig. 3).

2.6. Statistical analysis and modeling procedure

To compare the potential of various VIs to estimate the phenology of GPP, we correlated VIs with GPP transitions (both GPP onset and end)
using least-squares linear regression. Coefficients of determination (R²) and p-value were used to assess their performance. For each site, VIs that showed the highest R² with GPP onset or end were considered optimal candidates for predicting GPP transitions using the leave-one-out cross-validation approach (Gonsamo, Chen, Wu & Dragoni, 2012; Shao, 1993). After GPP transitions were modeled, GSL was calculated as the difference between modeled GPP end and GPP onset, and root mean square error (RMSE) and R² were adopted as indicators of model accuracy for each plant functional type.

3. Results

3.1. Growing season transitions of DBF sites

We found substantial variation in the ability of seasonally averaged VIs to track the growing season phenology of individually DBF sites (Table 3). Spring NDVI was negatively correlated with GPP onset at CA-OAS (R² = 0.69, p = 0.003) and US-UMB (R² = 0.67, p = 0.001) sites. Spring LSWI was significantly (negatively) correlated with GPP onset at CA-OAS (R² = 0.52, p = 0.018) and US-MMS (R² = 0.60, p = 0.015) sites. Spring EVI demonstrated relatively high potential as an indicator of GPP onset, with significant negative relationships found at three of four sites examined (CA-OAS, US-UMB and US-WCR). For WDRVI and OSAVI, we found significant negative correlations at CA-OAS (R² of 0.83 and 0.72, respectively) and US-UMB (R² of 0.83 and 0.87, respectively) sites. An R² of 0.80 (p < 0.001) was observed between spring WDRVI and the GPP onset for the overall dataset of these four DBF sites. Fig. 4 shows VIs that correlated significantly (all negative) with GPP onset during the spring at individual sites. When considering phenological metrics derived from seasonal patterns of EVI (EVI_SP), we found significant positive correlations at three individual sites (CA-OAS, US-UMB, US-MMS), with an R² of 0.67 (p < 0.001) for the overall dataset.

We also related GPP end during autumn at DBF sites with each of the five VIs, with none of the indices performing well across all DBF sites. None of the indices correlated with GPP end at US-WCR. Autumn NDVI was only significantly (positively) correlated with GPP end at US-MMS site (R² = 0.85, p < 0.001). Autumn LSWI and EVI were significantly correlated with GPP end at CA-OAS, US-UMB, US-MMS, and WDRVI and OSAVI with GPP end at US-UMB and US-MMS (all correlations were positive). Autumn LSWI showed the highest positive correlation with GPP end at CA-OAS site, but autumn OSAVI and WDRVI were the best measures (positive) of GPP end for US-UMB and US-MMS, respectively (Fig. 4). Conventional phenological indicators of GPP end developed from seasonal patterns of EVI did not exhibit better results than the seasonally averaged VIs, with significant positive correlations only at two sites (CA-OAS and US-MMS).

3.2. Growing season transitions of ENF sites

Compared with the mixed significant correlations found at DBF sites, the five seasonally averaged VIs had very limited potential to track GPP onset at ENF sites (Table 4). Spring NDVI (negatively) and LSWI (positively) were only significantly correlated with GPP onset at US-NR1 (R² = 0.67, p = 0.012) and CA-TP4 (R² = 0.42, p < 0.001), respectively. For EVI, no significant correlation was found with GPP onset at any site. WDRVI both negatively correlated with GPP onset at CA-OJP (R² = 0.58, p = 0.017) and US-NR1 sites (R² = 0.64, p = 0.017). Spring OSAVI showed greater potential for GPP onset modeling and it showed significant correlations with GPP onset at three sites (CA-OBS (positive), CA-MAN (positive), and US-NR1 (negative), Fig. 5). Analyzing seasonal patterns of EVI only gave a significant positive measure of GPP onset at US-NR1 site (R² = 0.65, p = 0.015).

Better results were observed for modeling GPP end than spring onset in ENF sites. Autumn NDVI was significantly (positively) correlated with GPP end at US-NR1. LSWI significantly predicted GPP end at CA-OJP (R² = 0.60, p = 0.014, negative), US-HO1 (R² = 0.51, p = 0.031, positive) and US-NR1 (R² = 0.56, p = 0.033, negative). Seasonally averaged EVI in autumn, however, was significantly correlated with GPP end only at CA-OJP site (R² = 0.50, p = 0.034, negative). For autumn WDRVI, we found significant positive correlations at both CA-MAN and US-NR1 (R² = 0.58, p = 0.018 and p = 0.029, respectively). Surprisingly, the best results were observed for OSAVI, which was positively correlated with GPP end at five of six sites (R² ranged from 0.44 (p = 0.048) at CA-MAN to 0.67 (p = 0.007) at US-HO1). Phenological indicators from seasonal patterns of EVI showed somewhat moderate results that GPP end of two sites can be explained (CA-OJP and US-HO1, both correlations were positive).

---

Table 2

<table>
<thead>
<tr>
<th>Vegetation indices</th>
<th>Abbreviations</th>
<th>Wavebands</th>
<th>Equations</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized difference vegetation index</td>
<td>NDVI</td>
<td>1, 2</td>
<td>( \text{NDVI} = (R_{NIR} - R_{RED})/(R_{NIR} + R_{RED}) )</td>
<td>Rouze, Haas, Schell, Deering, and Harlan (1974)</td>
</tr>
<tr>
<td>Land surface water index</td>
<td>LSWI</td>
<td>2, 5</td>
<td>( \text{LSWI} = (R_{NIR} - R_{SWIR})/(R_{NIR} + R_{SWIR}) )</td>
<td>Xiao et al. (2002)</td>
</tr>
<tr>
<td>Enhanced vegetation index</td>
<td>EVI</td>
<td>1, 2, 3</td>
<td>( \text{EVI} = 2.5 \times (R_{NIR} - R_{RED}) - 1 \times R_{RED} - 0.5 \times R_{BLUE} )</td>
<td>Huete et al. (2002)</td>
</tr>
<tr>
<td>Wide dynamic range vegetation index</td>
<td>WDRVI</td>
<td>1, 2</td>
<td>( \text{WDRVI} = (0.2R_{NIR} - R_{RED})/(0.2R_{NIR} + R_{RED}) )</td>
<td>Gitelson (2004)</td>
</tr>
<tr>
<td>Optimized soil-adjusted vegetation index</td>
<td>OSAVI</td>
<td>1, 2</td>
<td>( \text{OSAVI} = (1 + 0.16) \times R_{NIR} - R_{RED} + 0.16 )</td>
<td>Rondeaux et al. (1996)</td>
</tr>
</tbody>
</table>

---

Fig. 3. An example of determining growing season onset and end from seasonal patterns of EVI with data at CA-OAS for 2005. A double logistic function was fitted to the 8-day EVI data and the inflection positions in the left and right sides of the curve were taken as the onset and end of a growing season. Time duration between these two positions was calculated as the growing season length.

Fig. 4. Growing season transitions of DBF sites demonstrated by VIs and the fitted curves. Fig. 5. Growing season transitions of ENF sites demonstrated by VIs and the fitted curves.

---

NIR and SWIR represent near infrared and shortwave infrared, respectively.

a Wavebands refer to the MODIS standard settings and see text for details.

b Also called normalized difference water index (NDWI) as stated in Gao (1996).
3.3. Modeling growing season transitions and GSL using optimal VIs

In this section, three DBF (CA-OAS, US-UMB and US-MMS) and five ENF sites (CA-OAS, CA-TP4, CA-OJP, CA-MAN and US-NR1) were chosen to test the potential of these VIs in modeling both growing season transitions (GPP onset and end) and GSL using the cross validation approach. The reason for selecting these sites is that both GPP onset and end are significantly correlated with at least one of the five VIs so that we can continue to model the GSL after these transitions are being estimated.

The growing season transitions for each site were modeled with a combination of two VIs displaying the strongest correlations with growing season onset and end. For example, GSL at CA-OAS was inferred as the difference between modeled GPP onset from WDRVI and modeled GPP end from LSWI. Growing season onset of DBF sites was well predicted and $R^2$ of 0.91 ($p < 0.001$) was observed for the overall dataset with RMSE of 3.6 days (Fig. 6a). However, results for ENF sites showed lower accuracy with an $R^2$ of 0.59 ($p < 0.001$) and RMSE of 7.2 days (Fig. 6b). Growing season end of all considered sites was successfully modeled with $R^2$ higher than 0.90 ($p < 0.001$) for both the overall dataset and for each plant functional type (RMSE within 4 days). GSL was well predicted for both plant functional types, with $R^2$ of 0.96 ($p < 0.001$) and 0.89 ($p < 0.001$) observed between modeled and observed GSL for DBF and ENF, respectively (Fig. 6c). Furthermore, the correlations were both close to the 1:1 line, with an RMSE of 3.4 and 7.3 days for DBF and ENF sites, respectively. RMSE of 6.4 days was found for the overall dataset that included the two plant functional types.

Results for individual sites were also promising, with the RMSE of GPP onset for three deciduous sites ranging from 3.4 days at US-MMS to 3.8 days at CA-OAS site. Similar results were found for the GPP end of DBF sites (RMSE within 3.3 days). Phenological transitions for each ENF site were predicted with slightly less certainty, with RMSE of GPP onset ranging from 3.4 days for CA-TP4 to 9.5 days at CA-MAN. GPP end in ENF sites was better constrained as the largest RMSE was 4.6 days for US-NR1. The RMSE of GSL for individual DBF sites was between 3.5 (US-UMB) and 5.4 (CA-OAS) days while the value for each ENF site was from 4.3 days at CA-TP4 to 8.2 days at CA-MAN.

4. Discussion

4.1. Seasonal vegetation indices for phenology analysis

Instead of using relatively fine temporal scale metrics derived from daily to 16-day time series of remote sensing data, we used in our analysis seasonal (spring and autumn) indices that have much longer temporal resolutions. As a comparison with our approach, we also used conventional phenological metrics derived from seasonal patterns of EVI. Our findings show that seasonal signatures may be as robust as standard methods in tracking the growing season phenology of forests, though the potential varies among VIs and plant functional types. Phenological metrics derived from seasonal patterns of EVI correlated with observations moderately well at DBF sites though it did not explain the GPP onset at US-WCR or the GPP end at US-UMB and US-WCR. Similar to other studies, this standard approach showed limited potential to track ENF growing season transitions, only explaining GPP onset at one site (out of six, US-NR1) and GPP end at two of six sites. Furthermore, for all sites seasonally averaged VIs correlated more strongly with GPP onset and end than did seasonal patterns of EVI.

Our results indicate that seasonally averaged VIs may perform as well or better than standard approaches to remote sensing phenology, while substantially reducing data frequency requirements. This is especially important for regions where cloud cover has significant effect on data frequency. Furthermore, averaging seasonal data eliminates artifacts caused by complicated curve fitting or data smoothing, which may result in substantial differences in derived phenology metrics (Atkinson et al., 2012).

4.2. The potential of various VIs for phenology analysis

Vegetation indices differed substantially in their ability to track growing season phenology. NDVI and EVI are among the most commonly used VIs to indicate vegetation phenology (Melaas, Friedl, & Zhu, 2013; Zhang & Goldberg, 2011; Zhang et al., 2003). However, our evaluation suggests that neither NDVI nor EVI is superior over the other three VIs examined in tracking growing season transitions. For these two widely used VIs, only in one case did each exhibit the highest correlation with growing season transitions (NDVI vs. GPP onset at US-NK1 and EVI vs. GPP onset at US-WCR). Limited potential in NDVI was particularly acute for ENF sites for which this metric significantly explained GPP onset and end at one of six sites. EVI was similarly limited, exhibiting a correlation with GPP end only at CA-OJP site. The better performance of EVI than NDVI for growing season phenology modeling of DBF sites was found, possibly due to the improved sensitivity of EVI in reducing residual atmospheric effects and influence of soil and background as explained by Huete et al. (2002). The “water” index, LSWI, was significantly correlated with GPP onset at two DBF sites and one ENF site, and GPP end at six (three DBF and three ENF) sites. These observations are consistent with previous evaluations that show LSWI is closely related to phenology (Xiao et al., 2004; Jin et al., 2013). WDRVI also had better potential than NDVI and EVI to track growing season phenology. The best results were observed for OSAVI, EVI and LSWI, enhancing vegetation index, wide dynamic range vegetation index and optimized soil-adjusted vegetation index, respectively. EVI SP represents results of phe-

Note: + and − represent positive and negative correlations, respectively. NS represents no significant correlation was found. See text for details of site_ID. LSWI, EVI, WDRVI and OSAVI are land surface water index, enhanced vegetation index, wide dynamic range vegetation index and optimized soil-adjusted vegetation index, respectively. EVI SP represents results of phenological metrics derived from seasonal patterns of EVI using the double logistic function. Bold formats indicate the highest correlations among indices when they are related to GPP onset or end for each site.

<table>
<thead>
<tr>
<th>Phenology transitions</th>
<th>Site_ID</th>
<th>Phenological metrics</th>
<th>NDVI</th>
<th>LSWI</th>
<th>EVI</th>
<th>WDRVI</th>
<th>OSAVI</th>
<th>EVI SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPP onset (Day of year)</td>
<td>CA-OAS (n = 10)</td>
<td>0.69 (−), p = 0.003</td>
<td>0.52 (−), p = 0.018</td>
<td>0.81 (−), p &lt; 0.001</td>
<td>0.83 (−), p &lt; 0.001</td>
<td>0.72 (−), p = 0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>US-UMB (n = 12)</td>
<td>0.67 (−), p = 0.001</td>
<td>NS</td>
<td>0.83 (−), p &lt; 0.001</td>
<td>0.83 (−), p &lt; 0.001</td>
<td>0.87 (−), p &lt; 0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>US-MMS (n = 9)</td>
<td>NS</td>
<td>0.60 (−), p = 0.015</td>
<td>NS</td>
<td>NS</td>
<td>0.60 (+), p = 0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>US-WCR (n = 9)</td>
<td>NS</td>
<td>0.52 (−), p = 0.027</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>All data (n = 40)</td>
<td>0.76 (−), p &lt; 0.001</td>
<td>0.78 (−), p &lt; 0.001</td>
<td>0.80 (−), p &lt; 0.001</td>
<td>0.79 (−), p &lt; 0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPP end (Day of year)</td>
<td>CA-OAS (n = 10)</td>
<td>NS</td>
<td>0.58 (+), p = 0.010</td>
<td>0.81 (+), p = 0.000</td>
<td>0.91 (+), p &lt; 0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>US-UMB (n = 12)</td>
<td>NS</td>
<td>0.44 (+), p = 0.018</td>
<td>0.62 (+), p = 0.002</td>
<td>0.89 (+), p &lt; 0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>US-MMS (n = 9)</td>
<td>0.85 (+), p &lt; 0.001</td>
<td>0.72 (+), p = 0.004</td>
<td>0.84 (+), p &lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>US-WCR (n = 9)</td>
<td>NS</td>
<td>NS</td>
<td>0.91 (+), p &lt; 0.001</td>
<td>0.89 (+), p &lt; 0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>All data (n = 40)</td>
<td>0.75 (+), p &lt; 0.001</td>
<td>0.31 (+), p &lt; 0.001</td>
<td>0.72 (+), p &lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: + and − represent positive and negative correlations, respectively. NS represents no significant correlation was found. See text for details of site_ID. LSWI, EVI, WDRVI and OSAVI are land surface water index, enhanced vegetation index, wide dynamic range vegetation index and optimized soil-adjusted vegetation index, respectively. EVI SP represents results of phenological metrics derived from seasonal patterns of EVI using the double logistic function. Bold formats indicate the highest correlations among indices when they are related to GPP onset or end for each site.
Fig. 4. Relationships between growing season phenological variations (gross primary productivity onset, GPP onset, day of year; gross primary productivity end, GPP end, day of year) and spring (April–May) and autumn (September–November) vegetation indices for deciduous broadleaf forests. WDRVI, OSAVI, LSWI, and EVI are wide dynamic range vegetation index, optimized soil-adjusted vegetation index, land surface water index and enhanced vegetation index, respectively. Solid and dash lines represent regression and 95% confidence level for mean prediction, respectively.

which significantly explained GPP onset at five (two DBF and three ENF) sites and GPP end at eight (two DBF and all six ENF) sites.

There are several potential reasons for the strong performance of OSAVI. First, most remote sensing based spectral indices are, to some degree, contaminated by soil impacts but OSAVI was developed to maximize the capability in resistance to soil impacts from the SAVI family (Rondeaux, Steven, & Barret, 1996). Sensitivity analyses (e.g., Vincini & Frazzi, 2011; Liu, Pattey, & Jégo, 2012) consistently show that OSAVI is less influenced (only in 4–5% absolute difference) by solar and viewing zenith angles, soil background, and atmospheric visibility than other VIs (e.g., EVI) as reported in Sims, Rahman, Vermote, and Jiang (2011). Second, land surface phenology from remote sensing tracks seasonal dynamics of green vegetation and is therefore related to canopy chlorophyll content. The OSAVI consistently demonstrates the greatest accuracy among VIs to detect chlorophyll content because of the low interference from soils and atmosphere (Haboudane, Tremblay, Miller, & Vigneault, 2008; Stuckens et al., 2011; Wu, Niu, Tang, & Huang, 2008; Zarco-Tejada et al., 2005). These results have important implications for phenology modeling using remote sensing data. Present approaches may underestimate the potential of MODIS (or remote sensing) data in phenology modeling because most previous analyses only used one index (Hmimina et al., 2013; Zhang et al., 2003). Instead, our results suggest that a combination of indices serving as proxies of different indicators of phenology (e.g., water, greenness) may improve the accuracy of phenology modeling. However, because a single pair of VIs did not emerge as superior for all sites, additional investigation is required to determine when and why combined indices most accurately predict phenology.

### 4.3. Differences between plant functional types

A common observation is that the phenology of DBF ecosystems is more readily detected than that of ENF sites. This is likely because DBF ecosystems require leaf production and expansion to activate photosynthesis and a coinciding spectral shift has a large evident effect on canopy greenness. Similarly, leaf senescence and drop significantly change greenness (Wu et al., 2013). For ENF ecosystems, however, leaf expansion does not coincide with the start of GPP. In evergreen plants, only the oldest and least-efficient foliage is discarded in autumn, and the remaining leaves overwinter (Drenkhan, Kurkela, & Hanso, 2006; Xiao, 2003). Consequently, spectral signatures in evergreens change less as season shift and are less coupled with seasonal variation in photosynthesis (Coursolle et al., 2006; Doi & Takahashi, 2008). Our results generally support prior observations that GPP onset is more successfully detected in DBF sites than in ENF sites. An unexpected result, however, is that GPP end of ENF sites was better modeled than GPP onset, in particular when using OSAVI. This suggests that while canopy greenness is an adequate indicator of phenology in many remote sensing phenology applications, OSAVI’s removal of irrelevant spectral signatures (e.g., background, soil) may improve the tracking of phenological transitions.

### 4.4. Limitations and challenges

Our results inform the application of sensors that have a similar design of spectral bands. For example, the requirement of red and near-infrared bands to construct OSAVI can be met for most satellite sensors, and historical variations of phenology may be reconstructed given the long records accumulated from AVHRR and Landsat data over the past three decades. The largest obstacle in applying these methods is that the general relationship and pair of VIs that best explain patterns of phenology are highly site dependent, varying both within and across plant functional types. For example, OSAVI exhibits opposite relationships with GPP onset within the ENF ecosystems (CA-OBS vs. US-NR1). As another example, EVI was negatively correlated with GPP end at CA-OJP and conversely positively correlated with three DBF sites. Furthermore, none of the five indices were significantly correlated with the GPP end of US-WCR and the GPP onset at US-HO1. These results may also explain the non-significant correlations when VIs are used to model phenology for the combined data of all ENF sites. Such observations on one hand indicate the limitations of our analysis, but on the other hand they also demonstrate the difficulty of upscaling phenology using remote sensing. The underlying mechanisms explaining such high divergence, particularly within plant functional types, are poorly characterized at present, making it difficult to precisely determine the diverse phenological variation across large scales (Wang et al., 2011) as well as contrasting responses within a site (Cleland, Chiarrello, Loarie, Mooney, & Field, 2006). For an operational application, our method using the mean seasonal VI may be limited by the latitude of a region since remote sensing data can be potentially contaminated by the presence of snow that VIs do not change significantly until late May. In such case, the distribution of forests might be also strictly limited (Pan, Birdsey, Phillips, & Jackson, 2013; Zhu & Toutin, 2013). Consequently, identifying the mechanisms responsible for observed variability in the performance of VIs will greatly improve their application and help identify which spectral information is broadly useful in detecting phenology.
Fig. 5. Relationships between growing season phenological variations (gross primary productivity onset, GPP onset, day of year; gross primary productivity end, GPP end, day of year) and spring (April–May) and autumn (September–November) vegetation indices for evergreen needleleaf forests. WDRVI, OSAVI, LSWI, and EVI are wide dynamic range vegetation index, optimized soil-adjusted vegetation index, land surface water index, and enhanced vegetation index, respectively. Solid and dash lines represent regression and 95% confidence level for mean prediction, respectively.

No significant correlation was found at US-HO1 site.
mainly focused on a single VI at a finer temporal scale, we examined five seasonally averaged VIs that detect different properties of the forest canopy (e.g., water, greenness). We have shown that commonly used greenness indices (NDVI or EVI) had limited potential to track growing season phenology of ENF ecosystems while less commonly employed indices that remove background spectral noise may better detect phenological variations. One of the most consistently high performing VIs was OSAVI, which correlated strongly with GPP end at ENF sites, suggesting a new way to predict the end of growing season for ENF sites. Though none of the indices evaluated provided a consistent measure of phenology (both onset and end) across plant functional types, our results demonstrate that a combination of multiple indices may substantially improve the estimation of site-level phenology.

5. Conclusions

Using 94 site-years of flux data from ten forest sites in North America, we explored the potential of MODIS-derived seasonal vegetation indices to track season phenology. Unlike previous analyses that

- **Fig. 6.** Relationships between (a) modeled gross primary productivity onset (GPP onset) and observed GPP onset, (b) modeled gross primary productivity end (GPP end) and observed GPP end, and (c) modeled growing season length (GSL) and observed GSL for three deciduous broadleaf forests (DBF, CA-OAS, US-UMB and US-MMS) and five evergreen needleleaf forests (ENF, CA-OBS, CA-TP4, CA-MAN, CA-OJP and US-NR1). Solid and dash lines represent regression and 95% confidence level for mean prediction, respectively.

**Acknowledgments**

We are grateful to the principal investigators of these sites for providing the data and explanations. Constructive comments from three reviewers are appreciated. This work was funded by the Major State Basic Research Development Program of China (2013CB733405), the Knowledge Innovation Program of Chinese Academy of Sciences (KZCX2-EWQN302) and the National Natural Science Foundation of China (Grant No. 41001210, 41371013, 41271412). The flux data were acquired by the FLUXNET community and in particular by the following networks: AmeriFlux and Fluxnet-Canada, US-UMB site were supported by U.S. Department of Energy grant (DE-SC0006708) and NSF grant (DEB-0911461).

**References**

melaas, e. k., friedl, m. a., & zhu, z. (2013). detecting interannual variation in deciduous forest carbon exchange. agricultural and forest meteorology, 151, 1741–1752.
gonsamo, a., chen, j. m., price, d. t., kurz, w. a., & wu, c. (2012). predicting deciduous forest carbon uptake by upscaling fluxnet measurements using remote sensing data. agricultural and forest meteorology, 155, 127–135.
goyon, d., guillot, m., vittase, y., cardot, h., haguie, o., delzon, s., et al. (2011). monitoring elevation variations in leaf phenology of deciduous broadleaf forests from spot/vegetation time-series. remote sensing of environment, 115, 615–627.
holben, b., aber, d., dail, r., davidson, e. a., goltz, s. m., hughes, h., et al. (2004). spatial and temporal variability in forest-atmosphere CO2 exchange. global change biology, 10, 1689–1706.