Remote Sensing of Environment 186 (2016) 452–464

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Multisite analysis of land surface phenology in North American temperate and boreal deciduous forests from Landsat

Eli K. Melaas [a,*], Damien Sulla-Menashe [a], Josh M. Gray [a], T. Andrew Black [b], Timothy H. Morin [c], Andrew D. Richardson [d], Mark A. Friedl [a]

[a] Department of Earth and Environment, Boston University, Boston, MA, USA
[b] Biometeorology and Soil Physics Group, The University of British Columbia, Vancouver, British Columbia, Canada
[c] Department of Civil and Environmental Engineering and Geoditic Science, Ohio State University, Columbus, OH, USA
[d] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

ARTICLE INFO

Article history:
Received 20 May 2016
Received in revised form 16 August 2016
Accepted 14 September 2016
Available online 19 September 2016

Keywords:
Landsat
Phenology
PhenoCam
Eddy covariance
Temperate forests
Boreal forests

ABSTRACT

Forests play important roles in the Earth's climate system and global carbon cycle. Therefore, a critical need exists to improve our understanding of how the growing seasons of forests are changing, and by extension, how the composition and function of forests will respond to future climate change. Coarse spatial resolution satellite remote sensing has been widely used to monitor and map the phenology of terrestrial ecosystems at regional to global scales, and despite widespread agreement that the growing season of Northern Hemisphere forests is changing, the spatial resolution of these data sources imposes significant limitations on the character and quality of inferences that can be drawn from them. In particular, the spatial resolution afforded by instruments such as MODIS does not resolve ecologically important landscape-scale patterns in phenology. With this issue in mind, here we evaluate the ability of a newly developed Landsat phenology algorithm (LPA) to reconstruct a 32-year time series for the start and end of the growing season in North American temperate and boreal forests. We focus on 13 "sidelap" regions located between overlapping Landsat scenes that span a large geographic range of temperate and boreal forests, and evaluate the quality and character of LPA-derived start and end of growing season (SOS and EOS) dates using several independent data sources. On average, SOS and EOS dates were detected for about two-thirds of the 32 years included in our analysis, with the remaining one-third missing due to cloud cover. Moreover, there was generally better agreement between ground observations and LPA-derived estimates of SOS dates than for EOS across the 13 sites included in our study. Our results demonstrate that, despite the presence of time series gaps, LPA provides a robust basis for retrospective analysis of long-term changes in spring and autumn deciduous forest phenology over the last three decades. Finally, our results support the potential for monitoring land surface phenology at 30 m spatial resolution in near real-time by combining time series from multiple sensors such as the Landsat Operational Land Imager and the Sentinel 2 MultiSpectral Instrument.

1. Introduction

Forests play an important role in the Earth's climate system and global carbon cycle (Bonan, 2008). They influence planetary albedo and energy partitioning, and are an important sink for atmospheric carbon dioxide (Betts, 2000; Eugster et al., 2000; Pan et al., 2011). They also provide critical ecosystem services to society (Bennett et al., 2009). Over the last 30 years, most Northern Hemisphere forests have experienced warming, and evidence from both ground and satellite observations suggests that the timing and duration of growing seasons in many Northern Hemisphere forests is changing (Soja et al., 2007; Pudas et al., 2007; Beaubien & Freeland, 2000; Friedl et al., 2014; Keenan et al., 2014a; Gill et al., 2015). For all these reasons, a critical need exists to improve understanding of how the phenology of Northern Hemisphere forests is changing, and by extension, how the composition and function of these forests will respond to future climate change.

Satellite remote sensing has been widely used to monitor and map the phenology of terrestrial ecosystems at regional to global scales (Justice et al., 1985; Reed et al., 1994; Zhang et al., 2003). In particular, seasonal variation in spectral vegetation indices such as the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI) has been shown to be correlated with seasonal dynamics in photosynthetically active leaf area (Myneni & Hall, 1995; Huete et al., 2002). Exploiting this, numerous studies have used coarse spatial resolution instruments such as the Advanced Very High Resolution Radiometer (AVHRR) and the Moderate Resolution Imaging Spectroradiometer (MODIS) to characterize large-scale trends and variability in vegetation phenology (e.g., Myneni et al., 1997; Myneni et al., 1998; Zhou et al.,

* Corresponding author at: 675 Commonwealth Ave. Room 132, Boston University, Boston, MA 02215, USA.
E-mail address: emelaas@bu.edu (E.K. Melaas).

http://dx.doi.org/10.1016/j.rse.2016.09.014
0034-4257/

2003; Slayback et al., 2003; Jeong et al., 2011; Keenan et al., 2014a). These studies have identified changes in the growing season of terrestrial ecosystems across the Northern Hemisphere that are generally consistent with results from ground studies (Richardson et al., 2013). Specifically, results from both remote sensing and ground-based studies indicate that leaf emergence is happening earlier and leaf senescence is happening later in many ecosystems, leading to longer growing seasons.

Despite widespread agreement that the growing season of Northern Hemisphere forests is changing, the nature of remote sensing datasets used for these studies imposes significant limitations on the character and quality of inferences that can be drawn from them. In particular, results from AVHRR NDVI time series are limited by their coarse spatial resolution and uncertainties in sensor radiometry, geolocation, cloud screening and atmospheric correction (Goward et al., 1991; Huete et al., 2002). In regions where seasonal snow cover is common, detection and correction or removal of NDVI values contaminated by snow is challenging, even in MODIS data (Studer et al., 2007; Jönsson et al., 2010; Eklundh et al., 2011). Alternative vegetation indices (e.g., the Normalized Difference Wetness Index (Delbart et al., 2005) and Plant Phenology Index (Jin & Eklundh, 2014)) and heuristic approaches (e.g., Beck et al., 2006; Zhang et al., 2006) show promise for resolving this issue, but further testing of these methods is needed. Perhaps most importantly, disturbance regimes (fires, insects, humans), topography, and variation in edaphic, climate, and moisture controls produce landscapes in many forested ecosystems that are heterogeneous at scales well below the resolution of AVHRR or MODIS (Ahl et al., 2006; Serbin et al., 2013; Hogg et al., 2008). As a result, analyses based on coarse spatial resolution data are not able to resolve important modes and sources of variation in the growing season of terrestrial ecosystems (Badeck et al., 2004; Klosterman et al., 2014).

With these issues in mind, the goal of this paper is to evaluate the quality and utility of Landsat image time series for characterizing and quantifying long-term changes in North American temperate and boreal forest phenology over broad geographic scales. With over three decades of observations available at 30 m spatial resolution, the Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) instruments provide time series of stable and well-calibrated remote sensing observations that are well-suited for landscape-scale reconstruction and analysis of land surface phenology over the last 30 years. To perform our analysis, we exploit a newly developed Landsat phenology algorithm (LPA; Melaas et al., 2013a; Nijland et al., 2016), which uses all available images to estimate the timing of leaf emergence and fall senescence for pixels with well-defined deciduous phenology. In previous work, Melaas et al. (2013a) demonstrated that the LPA worked well at a single forested site in New England. In this paper, we use a larger and more diverse set of *in situ* observations to assess and interpret the nature and quality of Landsat-based estimates of deciduous temperate and boreal forest phenology. More specifically, we explore not only the agreement between traditional and Landsat-based measurements of phenology, but also examine a suite of complementary *in situ* data sources that provide a more comprehensive and critical assessment of the ecological meaning and utility of phenological information obtained from the LPA.

2. Methods

2.1. Landsat data and phenology algorithm

We used the Landsat phenology algorithm (LPA) described by Melaas et al. (2013a) to estimate the long-term average and annual day of year (DOY) associated with leaf emergence (start of growing season: SOS) and autumn senescence (end of growing season: EOS) at 30 m spatial resolution from time series of Landsat 4, 5, and 7 imagery. The original LPA described by Melaas et al. used all available images to model the phenology of observed Landsat EVI as a function of DOY based on separate logistic functions for spring and autumn. Long-term average SOS and EOS were estimated at each pixel based on the day of year when the spring and autumn logistic functions reached 50% of their amplitude. Observations during stable portions of each curve (e.g., the green and orange dots in Fig. 1c–d) were then used to estimate interannual anomalies in SOS and EOS, where spring and autumn anomalies were calculated as the difference between the date of each observation and the date on which the logistic curve reached the corresponding EVI value (Fig. 1e). Annual SOS and EOS dates were then estimated based on the retrieved anomalies in Spring (Autumn) and the long-term mean SOS (EOS) dates.

Because the LPA was originally developed at a relatively uniform and undisturbed temperate deciduous forest site, we implemented several revisions to the algorithm that allowed us to estimate SOS and EOS in forested landscapes that encompass more spatial heterogeneity in land cover and plant functional types. First, instead of using logistic functions fit to time series of EVI, we used cubic splines to model the mean annual phenology of EVI at each pixel. We find that cubic splines provide a more flexible basis for modeling mean annual phenology and reduce bias in estimated SOS and EOS dates that can be caused by asymmetric growth or decay in EVI, which is both common and not captured by logistic functions (Verma et al., 2016).

Second, the LPA assumes that EVI time series at each pixel follow a relatively consistent seasonal pattern from one year to the next. Hence, disturbance events that introduce large changes to the seasonal profile of EVI (e.g., wildfires, clear-cuts, or insect outbreaks) need to be identified and removed from further analysis. To do this, we used the Continuous Change Detection and Classification algorithm (Zhu & Woodcock, 2014), and excluded from our analysis pixels that experienced disturbance before 1999 or pixels that had more than two disturbance events during the entire record. For pixels with disturbance events that occurred after 1998, we estimated phenology dates using EVI observations acquired prior to the detected disturbance date. Thus, results from the LPA were only generated over extended undisturbed periods at each pixel. Third, to account for possible changes in peak-summer maximum greenness at each pixel (e.g., Goetz et al., 2005) that can introduce spurious trends in the timing of EOS and SOS estimated from the LPA, we normalized the Landsat time series using the 10th and 90th percentiles of EVI at each pixel within sequential three-year windows. After applying this normalization, both the range and the amplitude of seasonal EVI were consistent for each year over the time series at each pixel, and any biases in estimated SOS and EOS dates resulting from trends in greenness, either positive or negative, were eliminated. Finally, to account for radiometric differences between the TM and ETM+ sensors, we calibrated ETM+ EVI observations using the following linear correction derived by Sulla-Menashe et al. (2016):

$$EVI_{corr.} = (EVI_{obs.} - 0.019)/1.038$$

where $EVI_{corr.}$ and $EVI_{obs.}$ are corrected and observed ETM+ EVI values, respectively.

Our analysis used all available Landsat TM/ETM+ images at each of our study sites (see below) from 1982 to 2013 that had less than 90% cloud cover. Images were atmospherically corrected using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) atmosphere correction tool (Vermote et al., 1997; Masek et al., 2008), and clouds, cloud shadows, and snow were screened using the Fmask algorithm described by Zhu & Woodcock, (2012). Landsat images were clipped at each site to exclude pixels that were not located in sidelap regions between adjacent scenes (thereby doubling the number of images available for analysis), and were reprojected to the same Universal Transverse Mercator (UTM) Zone, as necessary.

2.2. Deciduous forest stratification

In Melaas et al. (2013a), the LPA was only applied to pixels that were identified as having strong deciduous phenology. Specifically, Melaas et

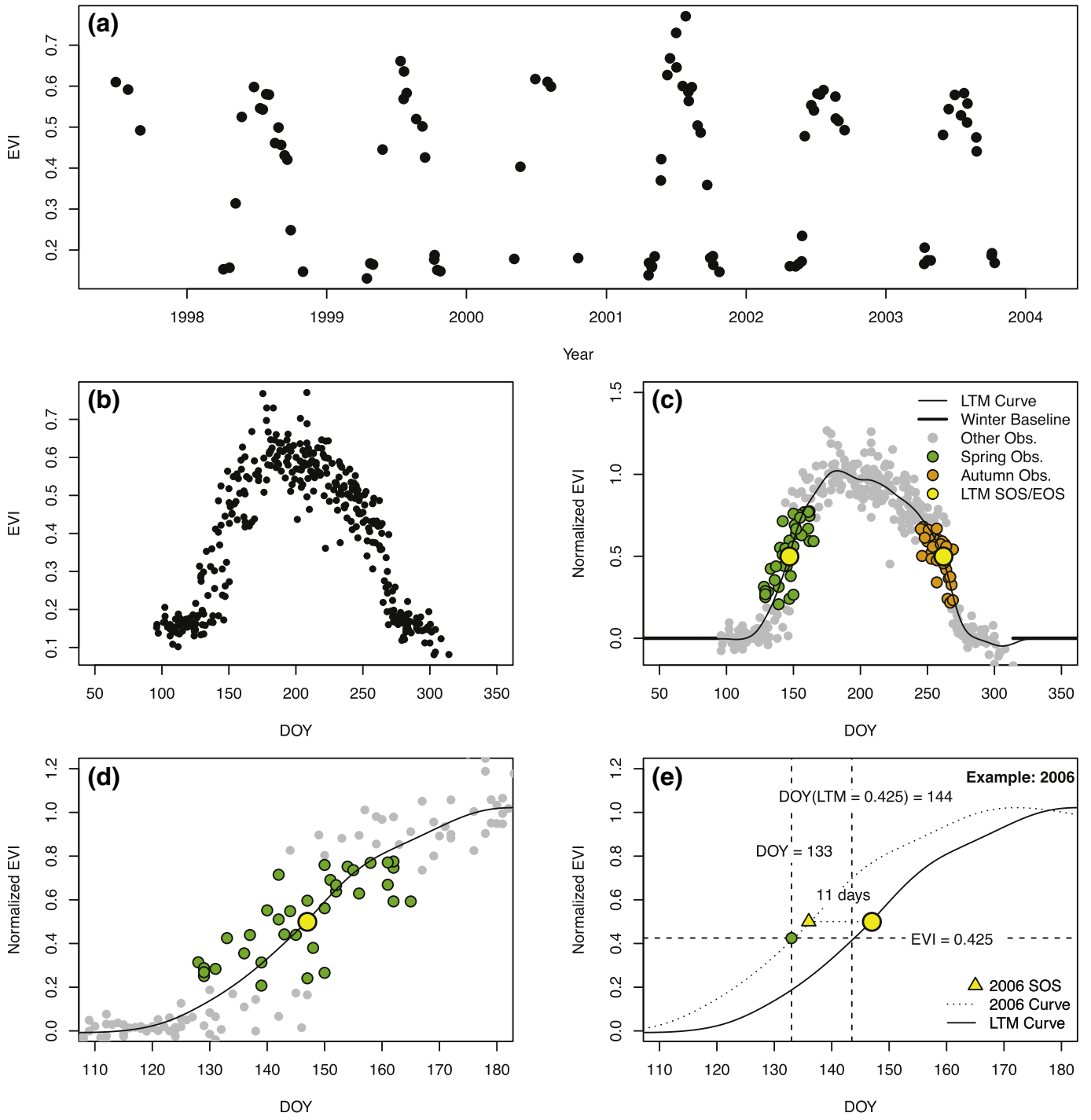

Fig. 1. Synopsis of the Landsat phenology algorithm for a sample deciduous forest pixel in the BP Old Aspen sidelap region (see Fig. 2). (a) Sample subset of full time series of EVI for a single pixel. (b) EVI plotted as a function of day of year (DOY). (c) Normalized EVI with long-term mean spline curve, long-term mean (LTM) transition dates (yellow dots) and green and orange dots identifying observations used to estimate interannual anomalies in SOS and EOS, respectively. (d) A portion of (c) showing 20+ day variation in spring phenology. (e) Sample detection of 2006 SOS. Spring and autumn anomalies are calculated as the difference between the date of each observation and the date when the spline reaches the corresponding EVI value. For more detailed information on the algorithm, please refer to Melaas et al. (2013a).

al. only included pixels where the lower asymptote of EVI for fitted logistic functions was between 0.1 and 0.25 and the upper asymptote was between 0.6 and 0.9. Because these thresholds were estimated based on local knowledge and were somewhat specific to the temperate mixed forest site where the LPA was developed, we adopted a more general approach here that makes the LPA applicable over a broader range of forest types and environments.

To do this, we used methodology based on the correlation between observed and spline-predicted EVI values that exploits the fact that pixels with large and relatively stable seasonal variation in EVI (i.e., dominated by deciduous vegetation) also show high correlations between EVI time series and cubic spline fits. Conversely, evergreen vegetation and other deciduous plant functional types (e.g., grasses, shrubs) tend to exhibit less seasonal variation in EVI and/or greater year-to-year variability in phenology. Hence, pixels dominated by these plant functional types tend to have lower correlation between observed EVI values and mean annual EVI time series. To identify pixels dominated by deciduous vegetation, we explored correlations between observed EVI and

long-term average EVI phenology for a range of land cover and plant functional types in boreal and temperate forest ecosystems. Based on this analysis, we identified a minimum correlation of 0.85 between the fitted splines and the EVI time series at any pixel to be a robust predictor of deciduous vegetation cover. In addition, we used the United States National Land Cover Dataset 2006 and Canadian Earth Observation for Sustainable Development of Forests (EOSD) land cover maps to identify and remove any agricultural pixels that were not excluded by the minimum correlation criterion (Fry et al., 2011; Wulder et al., 2003).

2.3. Assessment of LPA Results

Assessment and validation of remotely sensed estimates of SOS and EOS dates is challenging because direct observations of leaf- and canopy-level phenology are uncommon, and where they are available, time series of observations are usually quite short (generally less than 5 years). Such assessments are further complicated by mismatch in the spatial scale of ground measurements, which are point-based measurements, and remotely sensed observations, which integrate measurements over relatively large areas (in the case of Landsat, 900 m^2). We address this challenge here using multiple sources of surface, "near-surface", and remotely sensed observations of canopy and ecosystem phenology at 13 sites located in North American temperate and boreal forests (Fig. 2 and Table 1).

2.3.1. AmeriFlux data

In a previous study, Melaas et al., 2013b used eddy-covariance measurements to explore and model the timing of spring onset using 173 site-years of daily gross ecosystem productivity (GEP) measurements collected at 29 forest sites included in the FLUXNET "La Thuile" community data set. Here we used a similar strategy that focused on four deciduous broadleaf forest sites in the AmeriFlux network (ameriflux.ornl.gov) that are located in (or adjacent to) Landsat sidelap regions and that have at least five years of data: Morgan Monroe State Forest (MMSF), Harvard Forest (HF), the University of Michigan Biological Station (UMBS), and the Boreal Ecosystems Research and Monitoring Sites Old Aspen site (OAS) in Saskatchewan. To estimate the timing of EOS and SOS, we fit smoothing splines to time series of daily GEP at each site and identified ecosystem SOS and EOS as corresponding to the day of year when the smoothed GEP time series reached 50% of its growing season maximum in the spring (increasing GEP) and fall (decreasing GEP), respectively, for each site-year. Because eddy-covariance measurements integrate fluxes over footprints surrounding tower locations that can be quite large, we compared SOS and EOS dates estimated from GEP time series with mean SOS and EOS dates from Landsat using all pixels located within 1.5 km of each tower. One site (Morgan Monroe in Indiana) is located about 7.5 km east of the nearest sidelap region, and so we assumed that the phenology of forests in the nearest Landsat sidelap region is representative of those in the tower measurement footprint, and used all pixels in the sidelap region located within 10 km of the tower.

2.3.2. PhenoCam Data

PhenoCam is a continental-scale network of digital cameras that currently includes over 600 site-years of imagery that have been collected with the explicit goal of providing high-quality and high temporal resolution canopy-scale information related to vegetation phenology (http://phenocam.sr.unh.edu/webcam/). For this work, we used all deciduous forest sites in the PhenoCam network located within sidelap regions with at least five years of data (10 sites total; Table 1). To estimate the timing of SOS and EOS from PhenoCam imagery, we used the procedure described by Sonnentag et al., (2012) to extract time series of the Green Chromatic Coordinate (GCC). This procedure generates GCC time series at 3-day time steps based on a single region of interest in each PhenoCam field of view that was manually delineated to be representative of forest canopy conditions at each site. Three-day GCC data were interpolated to daily values using cubic smoothing splines, and SOS and EOS dates were estimated using the same approach that we used for the GEP data; i.e., SOS and EOS were identified as the day of

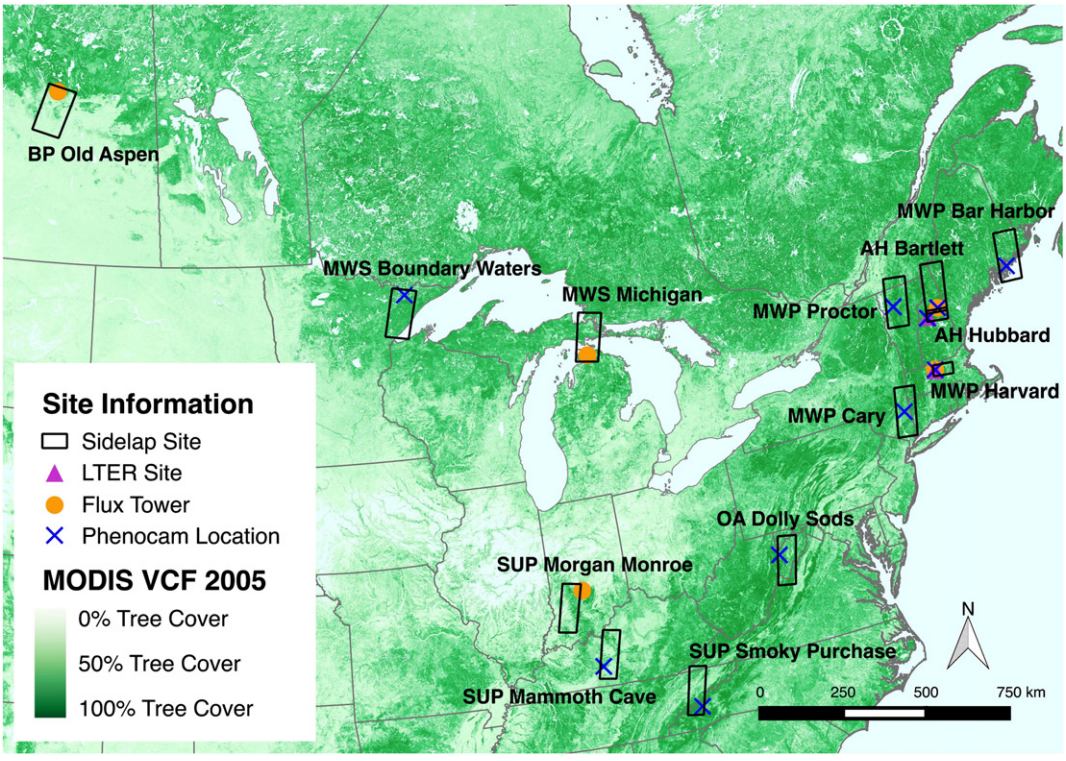

Fig. 2. Map of sidelap regions and corresponding ground observations used in the study. The background image shows percent tree cover across the study region from the 2005 Collection 5 MODIS Vegetation Continuous Fields product.

Table 1
Names, locations and general information about each sidelap region used in this analysis.

Site name	Level 1 ecoregion[1]	Latitude/longitude	Elevation (m)	Mean annual temp. (°C)	Dominant deciduous species[2]	*In situ* observations	Reference
Bar Harbor	MWS	44.38, −68.26	158	7.1	N/A	PhenoCam	Klosterman et al., 2014
Bartlett	AH	44.06, −71.29	268	6.2	Red maple, American beech	PhenoCam	Richardson et al., 2007
Boundary Waters	MWS	47.95, −91.50	519	2.8	N/A	PhenoCam	Klosterman et al., 2014
Cary	MWP	41.78, −73.73	127	9.6	Red oak	PhenoCam	klima.sr.unh.edu
Dolly Sods	OA	39.11, −79.43	1141	7.2	Sugar maple, red maple, American Beech	PhenoCam	Sonnentag et al., 2012
Harvard	MWP	42.53, −72.19	353	8.5	Red oak, red maple	Eddy covariance, ground, PhenoCam	Urbanski et al., 2007; Richardson & O'Keefe, 2009; Klosterman et al., 2014
Hubbard	AH	43.95, −71.72	500–700	5.5	Sugar maple, yellow birch, American beech	Ground, PhenoCam	Bailey, 2016; klima.sr.unh.edu
Mammoth Cave	SUP	37.19, −86.10	226	14.0	N/A	PhenoCam	Sonnentag et al., 2012
Michigan	MWS	45.56, −84.71	230	6.9	Bigtooth aspen, red maple, red oak, trembling aspen, paper birch	Eddy covariance	Gough et al. 2008
Morgan Monroe	SUP	39.32, −86.41	275	11.5	Tulip poplar, white oak, red oak, sugar maple	Eddy covariance	Schmid et al. 2000
Old Aspen	BP	53.63, −106.20	580	0.4	Aspen overstory with hazelnut understory	Eddy covariance	Krishnan et al., 2006
Proctor	MWP	45.53, −72.87	403	6.4	Sugar maple, red maple, American beech	PhenoCam	White et al., 2014
Smoky Purchase	SUP	35.59, −83.08	1550	9.0	Yellow birch, American Beech, red maple, tulip poplar	PhenoCam	Sonnentag et al. 2012

[1] AH = Atlantic Highlands; BP = Boreal Plain; MWP = Mixed Woods Plains; MWS = Mixed Woods Shield; OA = Ozark/Ouachita-Appalachian Forests; SUP = Southeastern USA Plains.
[2] For sites listed with dominant deciduous species as "N/A" we were unable to verify trees located within the PhenoCam footprint.

year when the interpolated GCC time series for each site-year passed 50% of its annual amplitude at the beginning and end of the growing season, respectively. In addition, we identified the timing of peak autumn color as the date when the Red Chromatic Coordinate index (RCC; a measure of canopy color (Sonnentag et al., 2012)) reached its annual maximum. To compare Landsat dates with PhenoCam dates, we calculated mean SOS and EOS dates using all Landsat pixels located within 500 m of each camera location. We used a 500 m radius to ensure that a sufficiently large sample of Landsat pixels were available each year-year, while also assuring that the pixels were in relatively close proximity of the camera and representative of the site.

2.3.3. Ground-based observations of phenology

Ground-based observations of phenology were derived from data collected by field observers at the Harvard Forest and Hubbard Brook Long Term Experimental Forests (Richardson et al., 2006). These data have been collected in the spring and fall for over 25 years at each site, encompass a large number of samples, and include multiple deciduous tree species located along a 2 km transect at Harvard Forest and a set of eight plots at Hubbard Brook (for details, see O'Keefe, 2000 and Bailey, 2016). Specific measurements and field protocols are different at each site, but the goal in each case is to characterize the timing of leaf emergence and development in the spring and coloration and leaf drop in the fall. At Harvard Forest we compared SOS and EOS dates from Landsat with time series of leaf length and leaf color measurements at 5% intervals from 5% to 95%. Results from this analysis indicated SOS and EOS dates were most highly correlated to dates when leaf length reached 25% of its final size and leaf coloring reached 50%, respectively, and where dates in the ground-observations were computed as weighted averages based on the relative composition of three dominant overstory species at the site: red oak (*Quercus rubra*), red maple (*Acer rubrum*) and yellow birch (*Betula alleghaniensis*). Because the observations were collected along a 2-km transect, we compared annual measurements of ground phenology with mean annual SOS and EOS dates for all Landsat pixels located in the Prospect Hill Tract of Harvard Forest (361 ha) where the observations were collected.

At Hubbard Brook, ground observations are collected for three dominant tree species (yellow birch (*Betula alleghaniensis*), sugar maple (*Acer saccharum*) and American beech (*Fagus grandifolia*)) at fixed field plots that span a range of elevation and slope aspects. To identify the DOY corresponding to the start of leaf emergence for each year at each plot location we used the date when the average canopy development index, which ranges from 0 to 4 and is used by field observers to quantify leaf development in the spring and senescence/drop in the fall, reached 3 in spring ('leaves 1/2 of final length') and decreased to 2 in autumn ('most leaves colored, few have fallen'). To compare our Landsat-based results with SOS and EOS dates derived from ground observations at Hubbard Brook, we used values for individual Landsat pixels corresponding to each plot location.

2.3.4. MODIS collection 6 land surface phenology

We compared annual SOS and EOS dates estimated from Landsat with corresponding dates estimated using the Collection 6 MODIS Land Cover Dynamics algorithm (Gray et al., *in prep*) between 2001 and 2013 across each of the Landsat sidelap regions where the ground observations, PhenoCam, and FLUXNET sites described above are located. The MODIS algorithm uses cubic splines to interpolate daily values of the two-band EVI (EVI2; Jiang et al., 2008), which is computed from the MODIS NBAR surface reflectance product (MDC43A4), and then estimates the timing of SOS and EOS as the DOY when the EVI2 reaches 50% of the seasonal amplitude of EVI each fall and spring at each pixel. For the purposes of this comparison, we used the average SOS and EOS for each year based on all Landsat pixels classified as deciduous forest in each MODIS pixel. Landsat pixels were classified as deciduous forest using the approach described in Section 2.2, and only MODIS pixels with at least 75 percent deciduous forest were included in the analysis. Note that while this does not provide a comparison of Landsat results with ground observations, this analysis allows us to assess the agreement (or lack thereof) between SOS and EOS data derived from an independent remote sensing source.

2.4. Analysis methods

To evaluate the agreement between SOS and EOS dates derived from Landsat versus those obtained from each of the independent data sets described above, we used several common measures of statistical agreement, including the root-mean-square-error (RMSE) and mean bias error (MBE) for each case. Because the ground and satellite observations both include significant measurement uncertainty, we used reduced major axis regression to estimate the coefficient of determination (R^2),

Table 2
Number of Landsat TM/ETM + images and retrieved SOS and EOS dates between 1982–2013 for each sidelap region. Values in parentheses indicate one standard deviation. Note: 500 m totals provide corresponding statistics when the data were aggregated to 500 m spatial resolution.

Site	Number of TM/ETM + Images in Stack	Average # of cloud-/snow-free EVI obs.	Average # of SOS years detected at 30 m	Average # of EOS years detected at 30 m	Average # of SOS years detected at 500 m	Average # of EOS years detected at 500 m
Bar Harbor	1127	368	20.0 (3.6)	22.0 (3.4)	21.2 (3.3)	23.5 (2.6)
Bartlett	1190	312	15.6 (4.1)	21.1 (3.6)	16.0 (3.8)	22.1 (2.8)
Boundary Waters	1203	348	20.5 (3.6)	22.2 (3.1)	21.4 (2.9)	23.2 (2.3)
Cary	1286	469	18.6 (4.3)	24.2 (2.8)	19.3 (4.0)	25.5 (2.1)
Dolly Sods	1233	440	17.1 (3.2)	25.7 (2.9)	17.8 (2.8)	27.1 (2.1)
Harvard	1257	463	19.2 (3.1)	24.0 (3.0)	19.8 (2.2)	25.8 (1.9)
Hubbard	1214	372	18.5 (3.7)	22.7 (3.2)	19.1 (3.3)	24.0 (2.6)
Mammoth Cave	1254	575	19.4 (3.5)	25.8 (2.9)	21.1 (3.0)	27.3 (1.9)
Michigan	1231	374	18.9 (4.4)	17.9 (4.1)	19.4 (4.6)	18.4 (4.1)
Morgan Monroe	1212	555	21.6 (3.5)	26.3 (2.9)	22.7 (3.1)	27.3 (2.4)
Old Aspen	951	331	18.5 (4.2)	18.4 (4.1)	19.2 (4.1)	19.2 (3.8)
Proctor	1154	320	17.6 (3.0)	21.0 (3.0)	18.3 (2.8)	22.0 (2.6)
Smoky Purchase	1103	452	17.7 (3.7)	23.5 (3.1)	19.0 (3.4)	24.9 (2.2)

slope, intercept, and associated 95% confidence intervals on linear regressions between each independent data set and the Landsat-derived EOS and SOS dates.

3. Results

3.1. Landsat phenology algorithm retrieval statistics

The number of available Landsat TM/ETM + images was relatively constant across the 13 sidelap zones we examined, with the exception of OAS (see Table 2), which had about 25% less imagery available. More importantly, because of differences in climate across the geographic range of these sites, the average number of cloud-free and snow-free EVI observations at each 30 m Landsat pixel varied between roughly 25–45% across the available 32 year time series (1982–2013). On average, the LPA detected SOS at any given Landsat pixel for roughly 50–72%, and EOS for roughly 40–85% of available years (Table 2). Missing retrievals were largely caused by clouds and cloud shadows, or by the ETM + scan-line corrector failure. Both SOS and EOS detection rates increased modestly when Landsat results at 30 m were spatially aggregated to 500 m for comparison with MODIS.

3.2. Comparison of LPA Results with AmeriFlux data

We compared the timing of LPA-detected SOS and EOS dates with corresponding dates estimated from GEP time series at four AmeriFlux towers (Section 2.4): HF, MMSF, OAS, and UMBS. Summary results for across- and within-site statistical assessments are presented in Fig. 3 and Table 3, respectively. Across these four sites, the timing of SOS derived from Landsat and GEP was highly correlated ($R^2 = 0.88$; $p < 0.01$), although the slope of this relationship was slightly less than 1 (Fig. 3a). Within each site, we also found strong correspondence between Landsat- and GEP-derived SOS ($R^2 = 0.54$–0.78), with slopes that were not significantly different from 1 (Table 3). Agreement between Landsat- and GEP-derived EOS for these sites was lower ($R^2 = 0.57$) than for SOS dates, and Landsat-derived EOS dates were biased late by roughly two weeks relative to GEP dates (Fig. 3b). Within sites, R^2 values between Landsat and GEP-derived EOS dates were variable (0.12–0.64). Moreover, the slope of this relationship was not consistent

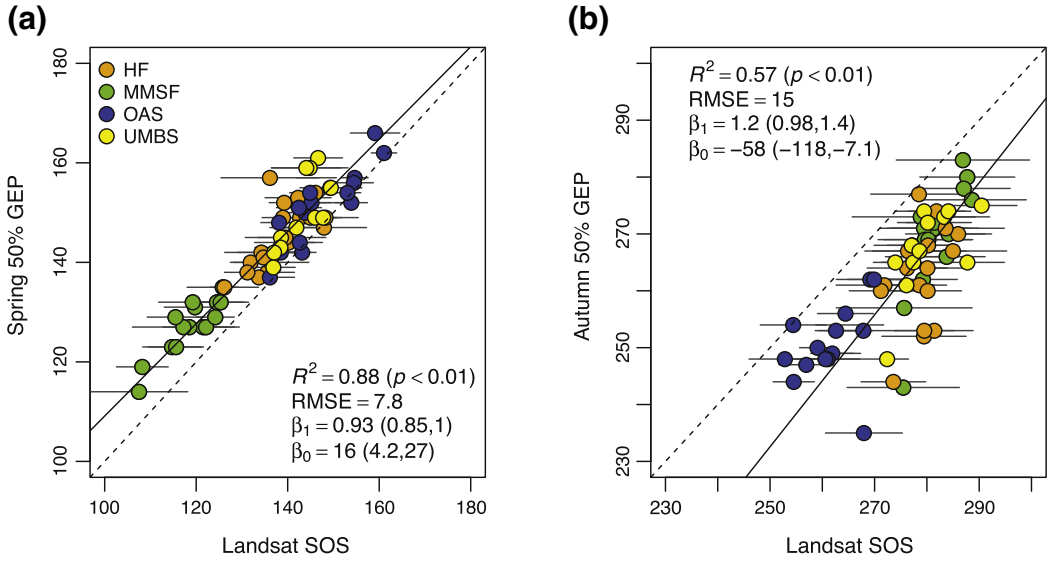

Fig. 3. Relationships between Landsat- and CO_2 flux-derived (a) SOS and (b) EOS across eddy covariance sites. Dashed lines are 1:1 and solid lines are reduced major axis (RMA) regression where between-site correlations are statistically significant ($p < 0.01$). RMA slope (β_1) and intercept (β_0) coefficients are provided with 95% confidence intervals. HF = Harvard Forest; MMSF = Morgan Monroe State Forest; OAS = Old Aspen Site; UMBS = University of Michigan Biological Station. Horizontal bars indicate one standard deviation in SOS or EOS during each site year across a 1.5 km radius centered on each flux tower (10 km at MMSF).

Table 3
Within-site reduced major axis regression statistics for Landsat and eddy covariance-derived SOS and EOS dates at each site. GEP = gross ecosystem productivity. * indicates statistical significance of correlations at $p < 0.05$.

Site	Years	R^2	RMSE	Slope (±95% CI)	Intercept (±95% CI)
Spring GEP 50%					
Morgan Monroe	15	0.78*	8.9	0.97 (0.74, 1.28)	11.9 (−25.1, 39.9)
Harvard	22	0.54*	8.1	1.05 (0.76, 1.46)	−0.4 (−57.6, 40.8)
Old Aspen	15	0.77*	5.3	0.98 (0.74, 1.30)	6.9 (−39.8, 42.2)
Michigan	12	0.58*	8.5	1.67 (1.05, 2.68)	−89.4 (−233, 0.3)
Autumn GEP 50%					
Morgan Monroe	15	0.64*	14.3	2.33 (1.64, 3.32)	−388 (−667, −193)
Harvard	22	0.14	17.4		
Old Aspen	15	0.12	13.1		
Michigan	12	0.43*	14.0	1.40 (0.84, 2.32)	−124 (−384, 32.8)

across each of the four sites and was significantly different from 1 at the MMSF and UMBS sites (Table 3).

3.3. Comparison of LPA Results with PhenoCam Data

We compared SOS and EOS dates from Landsat with corresponding dates estimated from PhenoCam imagery at 10 sites located in Landsat sidelap regions that had at least five years of camera imagery (Table 4). Across these sites, the timing of Landsat SOS was highly correlated with the timing of PhenoCam SOS ($R^2 = 0.86$; $p < 0.01$), with Landsat SOS biased late by an average of about 6 days (Fig. 4a). Within-site correlations between Landsat and PhenoCam SOS dates were also high, but varied substantially from site to site. Note that the time series of available data was quite short at some sites (<7 years), so statistical inference on within-site relationships was challenging (Table 4).

In contrast to SOS dates, agreement between Landsat EOS dates and Phenocam EOS dates estimated from GCC time series across sites was relatively weak ($R^2 = 0.25$; $p < 0.01$; Fig. 4b). Within sites, there was strong correlation between Landsat and PhenoCam EOS dates for the Boundary Waters, Proctor, and Hubbard Brook PhenoCam sites, but correlations were not significant at the seven other sites (Table 4). EOS dates based on the timing of peak RCC values, on the other hand, showed stronger agreement with EOS dates from Landsat ($R^2 = 0.44$; $p < 0.01$; Fig. 4c). However, Landsat EOS dates were biased early relative to peak RCC EOS dates, with significant site-to-site variation in the magnitude of this bias (e.g., Cary vs. Bartlett). Unlike SOS, where agreement was robust and strong both within and across sites, weaker agreement between Landsat- and PhenoCam-derived EOS dates makes it clear that even though both sources of data exhibit strong phenological patterns that seem to reflect the timing of senescence and leaf abscission, time series of vegetation indices derived from Landsat clearly measure canopy dynamics that are different from those measured by PhenoCam GCC time series.

3.4. Comparison of LPA Results with In Situ Ground Observations

Agreement between SOS and EOS dates derived from Landsat versus corresponding dates estimated using *in situ* measurements at Harvard Forest and Hubbard Brook was similar to that described above from the comparisons with AmeriFlux and PhenoCam Data. At Harvard Forest, correlation between Landsat and *in situ* dates was quite strong, but agreement was better for SOS than for EOS dates ($R^2 = 0.70$ vs. 0.50; Figs. 5a and 5b). To explore these patterns more closely at the species level, Fig. 6 presents boxplots for the root-mean-square error (RMSE), mean bias error (MBE) and R^2 between Landsat-derived and *in situ* measurements at Hubbard Brook of SOS (Fig. 6a–c) and EOS (Fig. 6d–f). In each subplot, the red, white, and blue boxes correspond

Table 4
Within-site reduced major axis regression statistics for Landsat and PhenoCam-derived SOS and EOS dates at each site. "Years" indicates number of years of camera imagery. GCC = green chromatic coordinate index. RCC = red chromatic coordinate index. * indicates statistical significance at $p < 0.05$.

Site	Years	R^2	RMSE	Slope (±95% CI)	Intercept (±95% CI)
Spring GCC					
Mammoth Cave	12	0.90*	3.3	1.17 (0.85, 1.60)	−21.1 (−70.1, 14.5)
Smoky Purchase	10	0.41	4.1		
Boundary Waters	8	0.93*	6.4	1.40 (1.03, 1.90)	−64.6 (−139.1, −9.9)
Dolly Sods	10	0.39	4.0		
Cary	6	0.83*	7.6	2.02 (1.17, 3.49)	−135.9 (−322.0, −28.1)
Proctor	5	0.20	8.3		
Bar Harbor	7	0.70*	7.6	1.13 (0.56, 2.30)	−26.4 (−192, 55.1)
Bartlett	6	0.91*	4.7	1.62 (0.95, 2.78)	−87.8 (−243.3, 2.9)
Harvard	6	0.97*	6.1	0.92 (0.66, 1.29)	4.7 (−44.8, 40.1)
Hubbard	5	0.82	7.7		
Autumn GCC					
Mammoth Cave	12	0.00	10.6		
Smoky Purchase	10	0.48*	8.2	1.39 (0.76, 2.56)	−118.8 (−449.4, 61.4)
Boundary Waters	8	0.65*	5.0	0.87 (0.46, 1.63)	35.1 (−162.8, 140.1)
Dolly Sods	10	0.07	10.7		
Cary	6	0.09	18.5		
Proctor	5	0.98*	1.1	0.93 (0.62, 1.40)	17.5 (−110.2, 102.6)
Bar Harbor	7	0.02	13.1		
Bartlett	6	0.11	14.2		
Harvard	6	0.11	5.1		
Hubbard	5	0.63	9.2		
Peak RCC					
Mammoth Cave	12	0.17	13.8		
Smoky Purchase	10	0.75*	9.1	1.23 (0.80,1.88)	−55.2 (−243, 66.7)
Boundary Waters	8	0.62*	11.5	0.79 (0.41, 1.54)	63.6 (−128, 163)
Dolly Sods	10	0.00	11.1		
Cary	6	0.18	25.4		
Proctor	5	0.83	7.4		
Bar Harbor	7	0.25	18.9		
Bartlett	6	0.05	3.7		
Harvard	6	0.34	20.6		
Hubbard	5	0.92*	18.8	3.57 (1.61, 7.90)	−672 (−1840, −146)

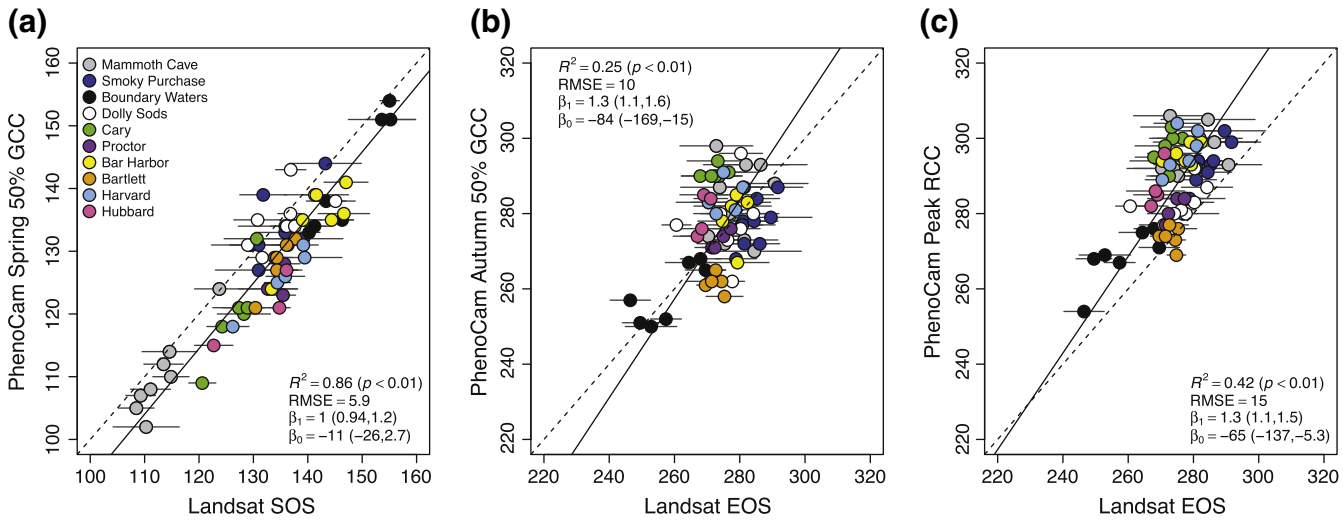

Fig. 4. Relationships between Landsat- and PhenoCam-derived SOS (a) and EOS (b and c) dates across PhenoCam sites. Dashed lines are 1:1 and solid lines are reduced major axis (RMA) regression models where between-site correlations are statistically significant ($p < 0.01$). RMA slope (β_1) and intercept (β_0) coefficients are provided with 95% confidence intervals. Horizontal bars indicate one standard deviation in SOS or EOS during each site year across a 500 m radius centered on each camera location.

to results for sugar maple (*Acer saccharum*), yellow birch (*Betula alleghaniensis*) and American beech (*Fagus grandifolia*), respectively. In spring, RMSEs between ground and Landsat SOS dates were consistently less than one week, with relatively low bias and high correlation across all species. In contrast, RMSEs for yellow birch and American beech EOS dates were twice as large as SOS RMSEs for these species, Landsat EOS was consistently biased early by more than one week relative to *in situ* measurements for all species, and R^2 for relationships between ground and Landsat dates is low (<0.20, on average, for all species).

3.5. Comparison of LPA Results with MODIS Collection 6

To summarize results from our comparison of SOS and EOS dates from Landsat versus MODIS across the 13 sidelap regions included in our analysis, we present two main sets of results. First, we report the mean and standard deviation for the RMSE, MBE, and R^2 between SOS and EOS dates derived from each sensor across all 500 m MODIS pixels with at least 75% deciduous forest cover located in each sidelap region (Table 5). Second, we compute mean annual anomalies in SOS and EOS from Landsat and MODIS for each sidelap region, and then use these data to assess correspondence in interannual variation in SOS and EOS measured by MODIS and Landsat (Fig. 7).

Table 5 shows that spring phenology dates estimated by the LPA and MODIS Land Cover Dynamics product (MCD12Q2) are in general agreement. Average RMSE was less than one week, and values of R^2 across the sidelap regions ranged between 0.5 and 0.8. MBE values in spring were generally small and negative, implying that the MODIS product estimates SOS a bit earlier than Landsat. In autumn, RMSE values were marginally higher and R^2 values are lower than in spring, while MBE values were more varied across the different sidelap regions. Fig. 7 shows that site-averaged annual anomalies in SOS and EOS from Landsat compare favorably with those from MODIS with R^2s for SOS anomalies being slightly higher than for EOS anomalies (0.83 vs. 0.68). Investigation of Landsat time series for the Old Aspen sidelap region revealed anomalous decreases in autumn greenness in two years (2002 and 2009)

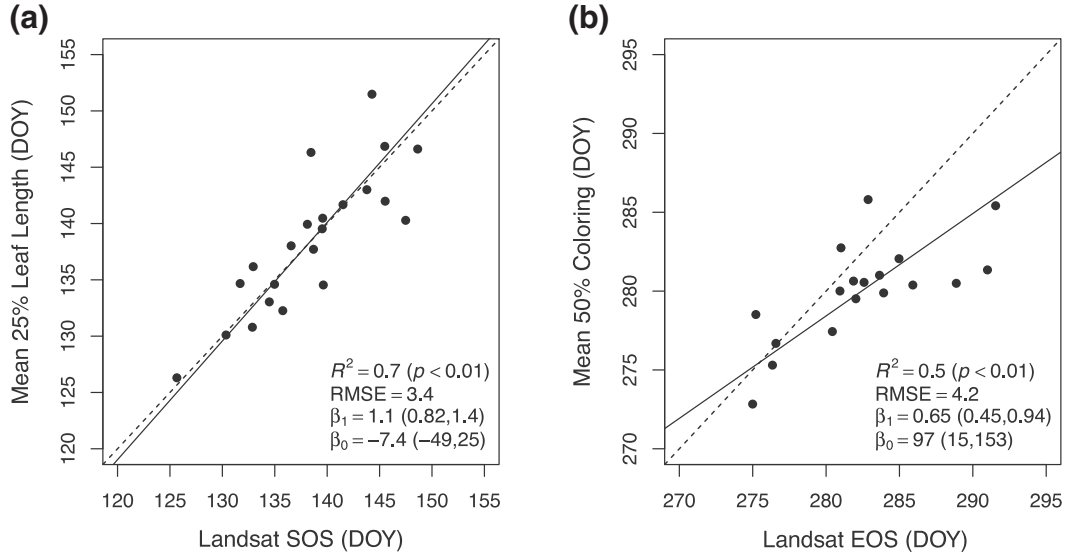

Fig. 5. (a) Landsat-derived SOS date versus the mean date when leaf length of common deciduous tree species reached 25% of maximum across the Prospect Hill Tract at Harvard Forest. (b) Landsat-derived EOS versus the mean date when leaves reached 50% coloring. Dashed lines are 1:1 and solid lines are reduced major axis (RMA) regression where between-site correlations are statistically significant ($p < 0.05$). RMA slope (β_1) and intercept (β_0) coefficients are provided with 95% confidence intervals.

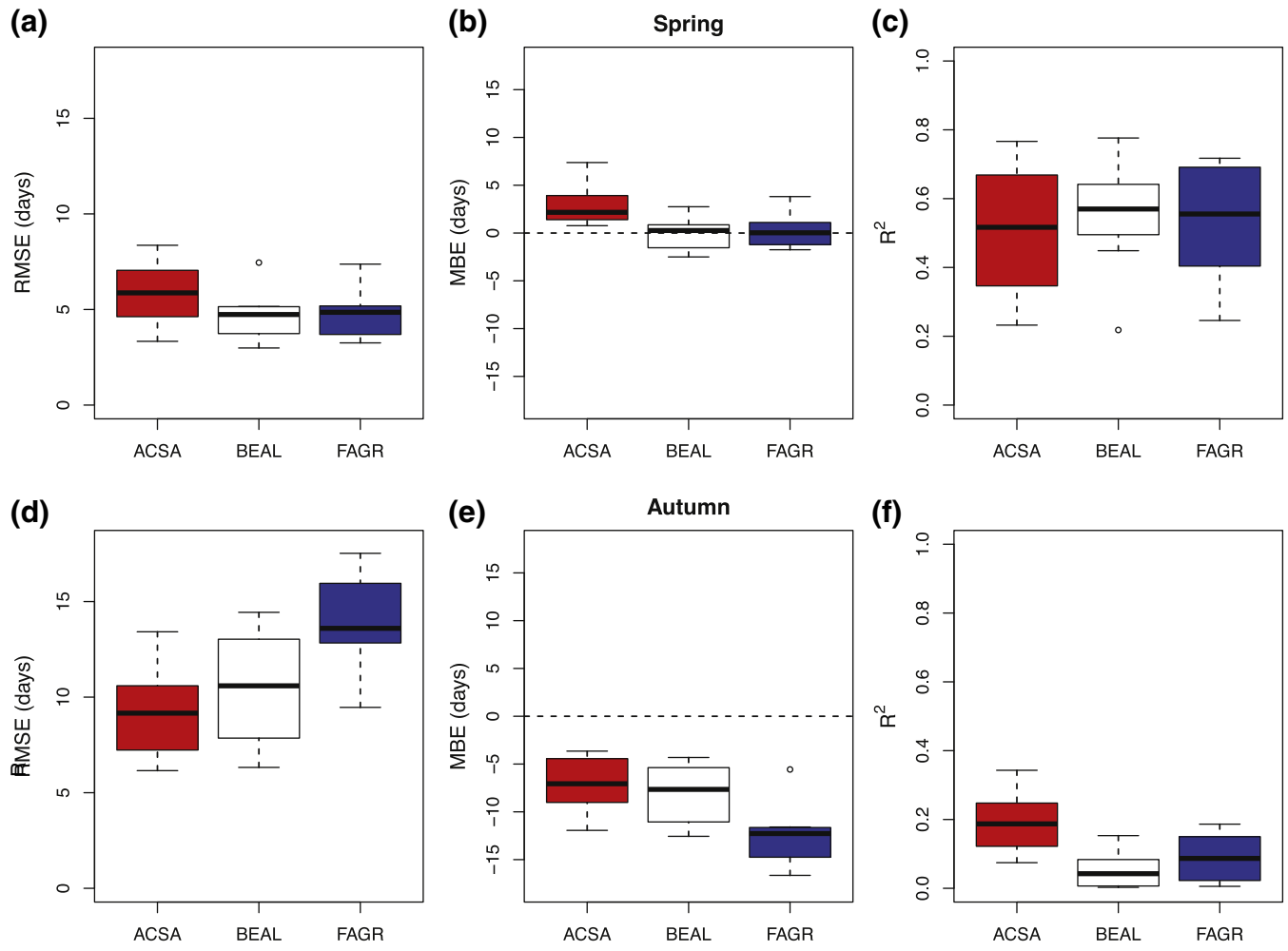

Fig. 6. Boxplots of statistical agreement between Landsat-derived spring (a–c) and autumn (d–f) phenology dates and *in situ* phenology dates across eight plots at Hubbard Brook. ACSA = *Acer saccharum* (Sugar maple); BEAL = *Betula alleghaniensis* (Yellow birch); FAGR = *Fagus grandifolia* (American beech).

followed by delayed senescence in the ensuing weeks. Because the Landsat observations associated with these years were acquired early in the fall, EOS dates estimated by the LPA were biased early relative to MODIS, which is clearly evident in Fig. 7 (light blue triangles).

4. Discussion and conclusions

The timing of spring leaf out and autumn senescence are widely used indicators of climate change impacts on ecosystem function in temperate and boreal forests (Richardson et al., 2013; Wu et al., 2012; Liu et al., 2016). Remote sensing has been widely used to explore and characterize both long-term trends and interannual variation in large-scale forest phenology (e.g., Xu et al., 2013; Friedl et al., 2014). Until recently, however, the vast majority of remote sensing-based studies of phenology have relied on coarse resolution imagery from sensors such as AVHRR, SPOT-Vegetation, and MODIS, which do not resolve important landscape-scale dynamics in phenology (e.g., Fisher & Mustard, 2007; Elmore et al., 2012). In this paper, we built upon and extended

Table 5
Summary statistics of the relationship between Landsat- and MODIS-detected spring and autumn phenology dates for each sidelap region. Numerical values represent the mean (and corresponding standard deviation) across all MODIS pixels with at least 75% deciduous forest cover in each sidelap region. Mean bias error (MBE) is calculated as MODIS – Landsat.

Site	Spring			Autumn		
	RMSE (days)	MBE (days)	R^2	RMSE (days)	MBE (days)	R^2
Bar Harbor	4.9 (1.8)	−2.4 (2.1)	0.52 (0.25)	5.6 (1.7)	−1.1 (2.7)	0.33 (0.24)
Bartlett	6.1 (3.4)	−3.9 (3.5)	0.55 (0.22)	5.5 (2.0)	2.0 (2.5)	0.35 (0.22)
Boundary Waters	3.0 (1.7)	−2.7 (1.8)	0.72 (0.18)	5.4 (2.1)	−0.3 (2.3)	0.53 (0.22)
Cary	3.0 (1.0)	−1.3 (1.2)	0.72 (0.15)	4.4 (1.3)	−1.4 (2.2)	0.55 (0.20)
Dolly Sods	4.7 (1.5)	−1.8 (2.0)	0.67 (0.19)	6.0 (2.1)	−1.7 (2.9)	0.45 (0.21)
Harvard	4.4 (1.4)	−2.5 (1.7)	0.66 (0.17)	4.7 (1.2)	−1.2 (2.0)	0.41 (0.21)
Hubbard	4.7 (3.0)	−1.9 (3.2)	0.58 (0.23)	4.8 (1.7)	1.3 (2.5)	0.41 (0.22)
Mammoth Cave	3.7 (1.2)	−0.9 (1.7)	0.67 (0.20)	7.3 (2.8)	−4.0 (2.6)	0.41 (0.21)
Michigan	4.7 (2.0)	−3.2 (2.2)	0.71 (0.15)	5.2 (1.5)	−0.1 (2.4)	0.41 (0.23)
Morgan Monroe	3.3 (1.1)	−0.4 (1.5)	0.70 (0.15)	6.9 (2.5)	−5.0 (2.2)	0.53 (0.21)
Old Aspen	3.8 (1.3)	−1.3 (1.6)	0.80 (0.13)	5.2 (1.8)	0.1 (2.1)	0.64 (0.18)
Proctor	4.4 (1.8)	−2.5 (2.3)	0.50 (0.22)	5.5 (1.9)	1.8 (2.5)	0.30 (0.22)
Smoky Purchase	4.4 (2.3)	2.2 (0.7)	0.54 (0.25)	7.1 (4.7)	−3.2 (3.3)	0.46 (0.26)

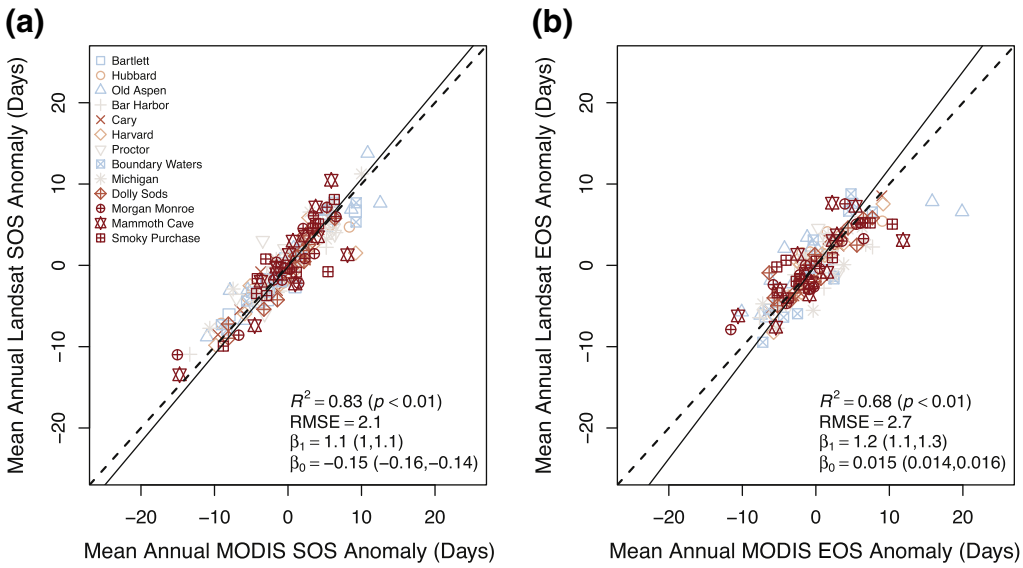

Fig. 7. Relationships between sidelap-mean annual anomalies in Landsat- and MODIS-detected (a) SOS and (b) EOS. Positive anomalies indicate phenology occurred earlier than normal, while negative anomalies indicate phenology occurred later than normal. Dashed lines are 1:1 and solid lines are reduced major axis (RMA) regression model slopes, where between-site correlations are statistically significant ($p < 0.01$). RMA slope (β_1) and intercept (β_0) coefficients are provided with 95% confidence intervals.

the methods and results reported by Melaas et al. (2013a)who developed an algorithm that uses Landsat TM/ETM + data to detect inter-annual variations in spring and autumn phenology of temperate deciduous forests at the 30 m spatial resolution afforded by Landsat. Specifically, to better understand the utility and information content of results from this algorithm, we used a variety of independent data sets to evaluate the quality and character of SOS and EOS dates retrieved from Landsat imagery in 13 Landsat sidelap regions that span a large range of North American temperate and boreal forest sub-types.

On average, the LPA successfully detected spring and autumn phenology at any given pixel for approximately two-thirds of the 32 years included in our analysis at each site. Missing retrievals were largely caused by clouds and cloud shadows, or by the ETM + scan-line corrector failure. Interestingly, we found that the frequency of missing values was only slightly reduced when 30 m results were aggregated to 500 m spatial resolution. An important practical result from this study is therefore that reliable estimation of land surface phenology at annual time steps and medium spatial resolution (i.e., 10–30 m) requires denser within-season time sampling than is currently afforded by Landsat alone. Fortunately, the successful launch of Sentinel 2A (and in the near future Sentinel 2B) will address this need. At the same time, because we were able to successfully retrieve SOS and EOS dates for roughly 20 out of the last 30 years, our results demonstrate that the LPA provides a useful basis for retrospective analysis of historical changes in spring and autumn phenology over the Landsat record, which was a primary motivation for developing the algorithm.

A key aspect of the results reported in this paper is that our analysis is restricted to Landsat pixels located within sidelap regions, where the likelihood of obtaining cloud-free observations is much higher than in other locations. For example, in our previous work (Melaas et al., 2013a) we used a single Landsat scene and were only able to estimate 9 SOS dates and 10 EOS dates (on average) over 30 years at Harvard Forest, while in the current study we detected 19 SOS and 24 EOS dates across the same geographic area (Table 2). For studies exploring long-term trends analyses or the impacts of interannual climate variability on phenology, sidelap regions clearly have significant advantages. While it is feasible to apply the LPA outside of sidelap zones, Melaas et al. (2013a)showed that uncertainty in SOS and EOS detection increases significantly when the number of cloud-free Landsat observations is below 200. As a result, care should be taken when applying the LPA to non-sidelap pixels in locations with frequent cloudy conditions (e.g., Bartlett and Proctor) or in Landsat scenes where the image archive is not dense.

A key goal of the analysis presented in this paper was to test how robust and effective the LPA is over a large range of sites in both temperate and boreal forests. Our results show that there was generally better agreement between ground observations and Landsat SOS dates than for EOS dates, both within and across sites included in our study. This result is consistent with previous studies that used moderate spatial resolution satellite remote sensing and is explained by differences in the structure and ecophysiology of forest canopy dynamics in spring versus fall (Garrity et al., 2011; Hufkens et al., 2012; Hmimina et al., 2013; Klosterman et al., 2014; D'Odorico et al., 2015; Nijland et al., 2016). Specifically, remotely sensed vegetation indices such as the EVI and GCC are strongly correlated with green leaf area (Huete et al., 2002; Keenan et al., 2014a). Therefore, the timing of leaf expansion in spring (and associated onset of photosynthesis) in deciduous canopies tends to be captured quite well by EVI from Landsat, GCC from PhenoCams, and eddy-covariance data (Barr et al., 2004; Richardson et al., 2010; Keenan et al., 2014b). Further, leaf expansion during spring tends to be relatively rapid as trees transition from dormancy to active photosynthesis, which provides a distinct signal that is easy to detect from remote sensing. During late summer and autumn, on the other hand, changes in leaf color (and eventual leaf drop) and associated declines in EVI, GCC, and GEP are more gradual, and also tend to be quite variable in space because of species-specific differences in the timing and controls on leaf aging, senescence, and abscission (e.g., see Fig. 8; Jenkins et al., 2007; Garrity et al., 2011; Yang et al., 2014). Further, during dry summers or prolonged periods of drought, remotely sensed greenness indices and productivity in deciduous forests (e.g., GEP) can become decoupled. For example, between 2001 and 2003 at Old Aspen in central Saskatchewan, annual precipitation averaged 263 mm—roughly half the annual normal total (480 mm between 1994–2015). However, greenness remained close to normal throughout the summer and autumn, despite strong drought conditions that reduced GEP by the summer of 2002 (Barr et al., 2004; Drolet et al., 2005; Krishnan et al., 2006). As a result, EOS in 2003 estimated using the LPA was 32 days late relative to the timing of EOS estimated from GEP time series (Fig. 3b).

MODIS-derived EOS and SOS dates agreed well with Landsat-derived EOS and SOS dates across all sidelap regions. This result is

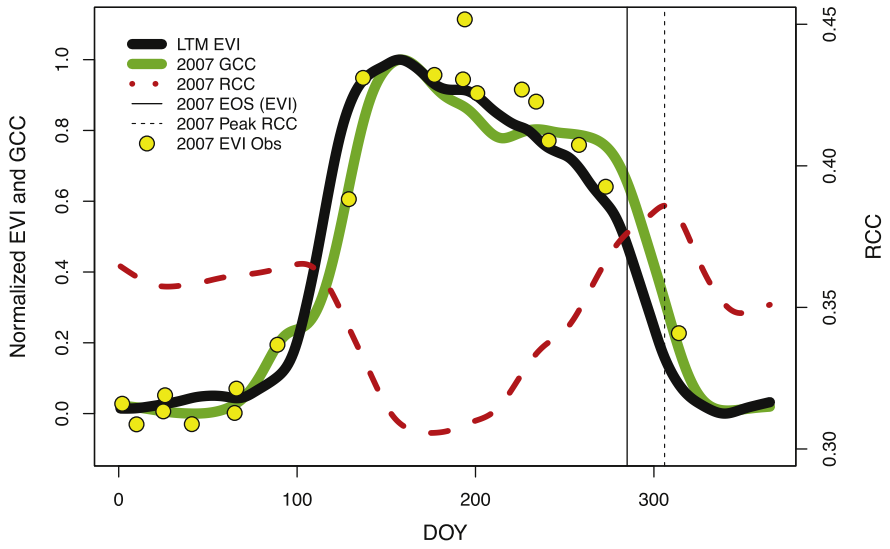

Fig. 8. 2007 smoothed daily time series of the green chromatic coordinate index (GCC; solid green line) and red chromatic coordinate index (RCC; dashed red line), smoothed long-term average enhanced vegetation index (EVI; solid black line) and 2007 EVI observations (yellow dots) for a single Landsat pixel at the Mammoth Cave PhenoCam site. The solid vertical line corresponds to the estimated timing of EOS in 2007 according to the Landsat phenology algorithm. The dashed vertical line corresponds to the timing of peak RCC.

consistent with previous findings by Fisher & Mustard (2007), who found strong correspondence between long-term average Landsat and MODIS phenology for forested locations in southeastern New England. Our results, in combination with those of Fisher & Mustard, (2007) suggest that in areas with significant deciduous forest cover (greater than 75 percent), it may be reasonable to gap-fill missing Landsat phenology dates with MODIS phenology. At the same time, it is important to note that local disagreement between MODIS and Landsat phenology can be substantial, because fine-scale controls on microclimatic arising from topography and land cover introduce variation in phenology at scales well below the spatial resolution of MODIS. Stated another way, if landscape-scale patterns are of interest, then land surface phenology information from MODIS is probably not sufficient. At sites located in more southern locations, which included substantial proportions of croplands in the Landsat sidelap regions (e.g., Morgan Monroe and Mammoth Cave), LPA Results had near-zero bias in spring but were biased late in autumn relative to MODIS. Because harvesting in croplands tends to be earlier in autumn relative to nearby natural vegetation (e.g., Zhang et al., 2006), it is possible that the presence of croplands (which we exclude from our Landsat sample but is included in the MODIS field of view) is biasing the timing of SOS and EOS dates from MODIS relative to Landsat. Data fusion algorithms such as the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM; Gao et al., 2006) that blend MODIS with Landsat data provide a potential means of leveraging the strengths of both Landsat and MODIS. However, because MODIS data are not available prior to 2000, this approach does not provide a solution for studies focused on retrospective analysis of long-term changes in the growing season of temperate and boreal forests.

Moving forward, results from this work support the utility of land surface phenology information derived from Landsat for improving information and understanding of ecosystem processes at landscape-scales. Specifically, forest fragmentation arising from land use and disturbance is prevalent across the eastern temperate and boreal forests of North America — in the United States, 62 percent of forests are located within 150 m of a forest edge (Riitters et al., 2002). Because of this, coarse spatial resolution sensors such as MODIS, which integrate surface reflectances over hundreds of meters, are insufficient to monitor fine scale patterns in land surface phenology arising from land use and land cover patterns. Moderate spatial resolution instruments such as the Landsat Operational Land Imager (OLI), which was explicitly designed for applications focused on land use and land cover, are able to capture and differentiate seasonal dynamics in land surface phenology at a spatial resolution that is an order of magnitude higher than MODIS. The results presented in this paper confirm the viability of the LPA, and lay the foundation for future algorithm development that exploits the growing availability of moderate spatial resolution data from sensors such as the Landsat OLI and Sentinel 2 MultiSpectral Instrument.

Acknowledgements

The authors acknowledge the following AmeriFlux sites for their data records: US-Ha1, US-MMS, US-UMB, CA-Oas. The majority of the funding for maintaining CA-Oas came from NSERC and CFCAS. In addition, funding for AmeriFlux data resources was provided by the U.S. Department of Energys Office of Science. The authors also thank three anonymous reviewers for their constructive comments and gratefully acknowledge phenology data provided by John O'Keefe at the Harvard Forest and Amy Bailey at the Hubbard Brook Experimental Forest. This work was partially supported by NASA grant numbers NNX15AK60G and NNX14AD57G and USGS grant number G12PC00070. ADR acknowledges support from the National Science Foundation's Macrosystems Biology (award EF-1065029) and LTER (awards DEB-1237491, DEB-1114804) programs.

References

Ahl, D.E., Gower, S.T., Burrows, S.N., Shabanov, N.V., Myneni, R.B., Knyazikhin, Y., 2006. Monitoring spring canopy phenology of a deciduous broadleaf forest using MODIS. Remote Sens. Environ. 104 (1), 88–95. http://doi.org/10.1016/j.rse.2006.05.003.

Badeck, F.-W., Bondeau, A., Böttcher, K., Doktor, D., Lucht, W., Schaber, J., Sitch, S., 2004. Responses of spring phenology to climate change. New Phytol. 162 (2), 295–309. http://doi.org/10.1111/j.1469-8137.2004.01059.x.

Bailey, A., 2016. Routine Phenology Measurements, Hubbard Brook Data Archive. Retrieved January 25, 2016 from http://hubbardbrook.org/data/dataset.php?id=51.

Barr, A.G., Black, T.A., Hogg, E.H., Kljun, N., Morgenstern, K., Nesic, Z., 2004. Inter-annual variability in the leaf area index of a boreal aspen-hazelnut forest in relation to net ecosystem production. Agric. For. Meteorol. 126 (3–4), 237–255. http://doi.org/10.1016/j.agrformet.2004.06.011.

Beaubien, E.G., Freeland, H.J., 2000. Spring phenology trends in Alberta, Canada: links to ocean temperature. Int. J. Biometeorol. 44 (2), 53–59. http://doi.org/10.1007/s004840000050.

Beck, P.S.A., Atzberger, C., Høgda, K.A., Johansen, B., Skidmore, A.K., 2006. Improved monitoring of vegetation dynamics at very high latitudes: a new method using MODIS NDVI. Remote Sens. Environ. 100 (3), 321–334. http://doi.org/10.1016/j.rse.2005.10.021.

Bennett, E.M., Peterson, G.D., Gordon, L.J., 2009. Understanding relationships among multiple ecosystem services. Ecol. Lett. 12 (12), 1394–1404. http://doi.org/10.1111/j.1461-0248.2009.01387.x.

Betts, R.A., 2000. Offset of the potential carbon sink from boreal forestation by decreases in surface albedo. Nature 408 (6809), 187–190. http://doi.org/10.1038/35041545.

Bonan, G.B., 2008. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. Science 320 (5882), 1444–1449. http://doi.org/10.1126/science.1155121.

Delbart, N., Kergoat, L., Le Toan, T., Lhermitte, J., Picard, G., 2005. Determination of phenological dates in boreal regions using normalized difference water index. Remote Sens. Environ. 97 (1), 26–38. http://doi.org/10.1016/j.rse.2005.03.011.

D'Odorico, P., Gonsamo, A., Gough, C.M., Bohrer, G., Morison, J., Wilkinson, M., ... Buchmann, N., 2015. The match and mismatch between photosynthesis and land surface phenology of deciduous forests. Agric. For. Meteorol. 214–215, 25–38. http://doi.org/10.1016/j.agrformet.2015.07.005.

Drolet, G.G., Huemmrich, K.F., Hall, F.G., Middleton, E.M., Black, T.A., Barr, A.G., Margolis, H.A., 2005. A MODIS-derived photochemical reflectance index to detect inter-annual variations in the photosynthetic light-use efficiency of a boreal deciduous forest. Remote Sens. Environ. 98 (2–3), 212–224. http://doi.org/10.1016/j.rse.2005.07.006.

Eklundh, L., Jin, H., Schubert, P., Guzinski, R., Heliasz, M., 2011. An optical sensor network for vegetation phenology monitoring and satellite data calibration. Sensors 11 (8), 7678–7709. http://doi.org/10.3390/s110807678.

Elmore, A.J., Guinn, S.M., Minsley, B.J., Richardson, A.D., 2012. Landscape controls on the timing of spring, autumn, and growing season length in mid-Atlantic forests. Glob. Chang. Biol. 18 (2), 656–674.

Eugster, W., Rouse, W.R., Pielke Sr., R.A., Mcfadden, J.P., Baldocchi, D.D., Kittel, T.G.F., ... Chambers, S., 2000. Land–atmosphere energy exchange in Arctic tundra and boreal forest: available data and feedbacks to climate. Glob. Chang. Biol. 6 (S1), 84–115. http://doi.org/10.1046/j.1365-2486.2000.06015.x.

Fisher, J.I., Mustard, J.F., 2007. Cross-scalar satellite phenology from ground, Landsat, and MODIS data. Remote Sens. Environ. 109 (3), 261–273. http://doi.org/10.1016/j.rse.2007.01.004.

Friedl, M.A., Gray, J.M., Melaas, E.K., Richardson, A.D., Hufkens, K., Keenan, T.F., ... O'Keefe, 2014. A tale of two springs: using recent climate anomalies to characterize the sensitivity of temperate forest phenology to climate change. Environ. Res. Lett. 9 (5), 054006. http://doi.org/10.1088/1748-9326/9/5/054006.

Fry, J.A., Xian, G., Jin, S., Dewitz, J.A., Homer, C.G., Limin, Y., ... Wickham, J.D., 2011. Completion of the 2006 national land cover database for the conterminous United States. Photogramm. Eng. Remote. Sens. 77 (9), 858–864. http://doi.org/10.1016/j.rse.2014.02.001.

Gao, F., Masek, J., Schwaller, M., Hall, F., 2006. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. IEEE Trans. Geosci. Remote Sens. 44 (8), 2207–2218. http://10.1109/TGRS.2006.872081.

Garrity, S.R., Bohrer, G., Maurer, K.D., Mueller, K.L., Vogel, C.S., Curtis, P.S., 2011. A comparison of multiple phenology data sources for estimating seasonal transitions in deciduous forest carbon exchange. Agric. For. Meteorol. 151 (12), 1741–1752. http://doi.org/10.1016/j.agrformet.2011.07.008.

Gill, A.L., Gallinat, A.S., Sanders-DeMott, R., Rigden, A.J., Gianotti, D.J.S., Mantooth, J.A., Templer, P.H., 2015. Changes in autumn senescence in northern hemisphere deciduous trees: a meta-analysis of autumn phenology studies. Ann. Bot. mcv055. http://doi.org/10.1093/aob/mcv055.

Goetz, S.J., Bunn, A.G., Fiske, G.J., Houghton, R.A., 2005. Satellite-observed photosynthetic trends across boreal North America associated with climate and fire disturbance. Proc. Natl. Acad. Sci. U. S. A. 102 (38), 13521–13525. http://doi.org/10.1073/pnas.0506179102.

Gough, C.M., Vogel, C.S., Schmid, H.P., Su, H.-B., Curtis, P.S., 2008. Multi-year convergence of biometric and meteorological estimates of forest carbon storage. Agric. For. Meteorol. 148 (2), 158–170. http://doi.org/10.1016/j.agrformet.2007.08.004.

Goward, S.N., Markham, B., Dye, D.G., Dulaney, W., Yang, J., 1991. Normalized difference vegetation index measurements from the advanced very high resolution radiometer. Remote Sens. Environ. 35 (2), 257–277. http://doi.org/10.1016/0034-4257(91)90017-Z.

Hmimina, G., Dufrêne, E., Pontailler, J.-Y., Delpierre, N., Aubinet, M., Caquet, B., ... Soudani, K., 2013. Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: An investigation using ground-based NDVI measurements. Remote Sens. Environ. 132, 145–158. http://doi.org/10.1016/j.rse.2013.01.010.

Hogg, E.H.(.T.)., Brandt, J.P., Michaelian, M., 2008. Impacts of a regional drought on the productivity, dieback, and biomass of western Canadian aspen forests. Can. J. For. Res. 38 (6), 1373–1384. http://doi.org/10.1139/X08-001.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83 (1–2), 195–213. http://doi.org/10.1016/S0034-4257(02)00096-2.

Hufkens, K., Friedl, M., Sonnentag, O., Braswell, B.H., Milliman, T., Richardson, A.D., 2012. Linking near-surface and satellite remote sensing measurements of deciduous broadleaf forest phenology. Remote Sens. Environ. 117, 307–321. http://doi.org/10.1016/j.rse.2011.10.006.

Jenkins, J.P., Richardson, A.D., Braswell, B.H., Ollinger, S.V., Hollinger, D.Y., Smith, M.-L., 2007. Refining light-use efficiency calculations for a deciduous forest canopy using simultaneous tower-based carbon flux and radiometric measurements. Agric. For. Meteorol. 143 (1–2), 64–79. http://doi.org/10.1016/j.agrformet.2006.11.008.

Jeong, S.-J., Ho, C.-H., Gim, H.-J., Brown, M.E., 2011. Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982–2008. Glob. Chang. Biol. 17 (7), 2385–2399. http://doi.org/10.1111/j.1365-2486.2011.02397.x.

Jiang, Z., Huete, A.R., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. Remote Sens. Environ. 112 (10), 3833–3845. http://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 Sens. Environ. 152, 512–525. http://doi.org/10.1016/j.rse.2014.07.010.

Jönsson, A.M., Eklundh, L., Hellström, M., Bärring, L., Jönsson, P., 2010. Annual changes in MODIS vegetation indices of Swedish coniferous forests in relation to snow dynamics and tree phenology. Remote Sens. Environ. 114 (11), 2719–2730. http://doi.org/10.1016/j.rse.2010.06.005.

Justice, C.O., Townshend, J.R.G., Holben, B.N., Tucker, C.J., 1985. Analysis of the phenology of global vegetation using meteorological satellite data. Int. J. Remote Sens. 6 (8), 1271–1318. http://doi.org/10.1080/01431168508948281.

Keenan, T.F., Darby, B., Felts, E., Sonnentag, O., Friedl, M.A., Hufkens, K., ... Richardson, A.D., 2014a. Tracking forest phenology and seasonal physiology using digital repeat photography: a critical assessment. Ecol. Appl. 24 (6), 1478–1489. http://doi.org/10.1890/13-0652.1.

Keenan, T.F., Gray, J., Friedl, M.A., Toomey, M., Bohrer, G., Hollinger, D.Y., ... Richardson, A.D., 2014b. Net carbon uptake has increased through warming-induced changes in temperate forest phenology. Nat. Clim. Chang. 4 (7), 598–604. http://doi.org/10.1038/nclimate2253.

Klosterman, S.T., Hufkens, K., Gray, J.M., Melaas, E., Sonnentag, O., Lavine, I., ... Richardson, A.D., 2014. Evaluating remote sensing of deciduous forest phenology at multiple spatial scales using PhenoCam imagery. Biogeosciences 11 (16), 4305–4320. http://doi.org/10.5194/bg-11-4305-2014.

Krishnan, P., Black, T.A., Grant, N.J., Barr, A.G., Hogg, E.(.T.).H., Jassal, R.S., Morgenstern, K., 2006. Impact of changing soil moisture distribution on net ecosystem productivity of a boreal aspen forest during and following drought. Agric. For. Meteorol. 139 (3-4), 208–223. http://doi.org/10.1016/j.agrformet.2006.07.002.

Liu, Y., Wu, C., Peng, D., Xu, S., Gonsamo, A., Jassal, R.S., ... Chen, J.M., 2016. Improved modeling of land surface phenology using MODIS land surface reflectance and temperature at evergreen needleleaf forests of central North America. Remote Sens. Environ. 176, 152–162. http://doi.org/10.1016/j.rse.2016.01.021.

Masek, J.G., Huang, C., Wolfe, R., Cohen, W., Hall, F., Kutler, J., Nelson, P., 2008. North American forest disturbance mapped from a decadal Landsat record. Remote Sens. Environ. 112 (6), 2914–2926. http://doi.org/10.1016/j.rse.2008.02.010.

Melaas, E.K., Friedl, M.A., Zhu, Z., 2013a. Detecting interannual variation in deciduous broadleaf forest phenology using Landsat TM/ETM + data. Remote Sens. Environ. 132, 176–185. http://doi.org/10.1016/j.rse.2013.01.011.

Melaas, E.K., Richardson, A.D., Friedl, M.A., Dragoni, D., Gough, C.M., Herbst, M., ... Moors, E., 2013b. Using FLUXNET data to improve models of springtime vegetation activity onset in forest ecosystems. Agric. For. Meteorol. 171–172, 46–56. http://doi.org/10.1016/j.agrformet.2012.11.018.

Myneni, R.B., Hall, F.G., 1995. The interpretation of spectral vegetation indexes. IEEE Trans. Geosci. Remote Sens. 33 (2), 481–486. http://doi.org/10.1109/36.377948.

Myneni, R.B., Keeling, C.D., Tucker, C.J., Asrar, G., Nemani, R.R., 1997. Increased plant growth in the northern high latitudes from 1981 to 1991. Nature 386 (6626), 698–702. http://doi.org/10.1038/386698a0.

Myneni, R.B., Tucker, C.J., Asrar, G., Keeling, C.D., 1998. Interannual variations in satellite-sensed vegetation index data from 1981 to 1991. J. Geophys. Res.-Atmos. 103 (D6), 6145–6160. http://doi.org/10.1029/97JD03603.

Nijland, W., Bolton, D.K., Coops, N.C., Stenhouse, G., 2016. Imaging phenology; scaling from camera plots to landscapes. Remote Sens. Environ. 177, 13–20. http://doi.org/10.1016/j.rse.2016.02.018.

O'Keefe, J. (n.d.). Phenology of Woody Species at Harvard Forest since 1990. Retrieved January 25, 2016, from http://harvardforest.fas.harvard.edu:8080/exist/apps/datasets/showData.html?id=hf003

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., ... Hayes, D., 2011. A large and persistent carbon sink in the world's forests. Science 333 (6045), 988–993. http://doi.org/10.1126/science.1201609.

Pudas, E., Leppälä, M., Tolvanen, A., Poikolainen, J., Venäläinen, A., Kubin, E., 2007. Trends in phenology of Betula pubescens across the boreal zone in Finland. Int. J. Biometeorol. 52 (4), 251–259. http://doi.org/10.1007/s00484-007-0126-3.

Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W., Ohlen, D.O., 1994. Measuring phenological variability from satellite imagery. J. Veg. Sci. 5 (5), 703–714. http://doi.org/10.2307/3235884.

Richardson, A.D., O'Keefe, J., 2009. Phenological differences between understory and overstory: a case study using the long-term harvard forest records. In: Noormets, A. (Ed.), Phenology of ecosystem processes. Springer, New York, NY, pp. 87–117.

Richardson, A.D., Bailey, A.S., Denny, E.G., Martin, C.W., O'keefe, J., 2006. Phenology of a northern hardwood forest canopy. Glob. Chang. Biol. 12 (7), 1174–1188. http://doi.org/10.1111/j.1365-2486.2006.01164.x.

Richardson, A.D., Black, T.A., Ciais, P., Delbart, N., Friedl, M.A., Gobron, N., ... Varlagin, A., 2010. Influence of spring and autumn phenological transitions on forest ecosystem productivity. Philosophical Transactions of the Royal Society of London B: Biological Sciences. 365, pp. 3227–3246 (1555). http://doi.org/10.1098/rstb.2010.0102.

Richardson, A.D., Jenkins, J.P., Braswell, B.H., Hollinger, D.Y., Ollinger, S.V., Smith, M.-L., 2007. Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. Oecologia 152 (2), 323–334.

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. Agric. For. Meteorol. 169, 156–173. http://doi.org/10.1016/j.agrformet.2012.09.012.

Riitters, K.H., Wickham, J.D., O'Neill, R.V., Jones, K.B., Smith, E.R., Coulston, J.W., ... Smith, J.H., 2002. Fragmentation of continental United States forests. Ecosystems 5 (8), 0815–0822. http://doi.org/10.1007/s10021-002-0209-2.

Schmid, H.P., Grimmond, C.S.B., Cropley, F., Offerle, B., Su, H.-B., 2000. Measurements of CO2 and energy fluxes over a mixed hardwood forest in the mid-western United

States. Agric. For. Meteorol. 103 (4), 357–374. http://doi.org/10.1016/S0168-1923(00)00140-4.
Serbin, S.P., Ahl, D.E., Gower, S.T., 2013. Spatial and temporal validation of the MODIS LAI and FPAR products across a boreal forest wildfire chronosequence. Remote Sens. Environ. 133, 71–84. http://doi.org/10.1016/j.rse.2013.01.022.
Slayback, D.A., Pinzon, J.E., Los, S.O., Tucker, C.J., 2003. Northern hemisphere photosynthetic trends 1982–99. Glob. Chang. Biol. 9 (1), 1–15. http://doi.org/10.1046/j.1365-2486.2003.00507.x.
Soja, A.J., Tchebakova, N.M., French, N.H.F., Flannigan, M.D., Shugart, H.H., Stocks, B.J., ... Stackhouse Jr., P.W., 2007. Climate-induced boreal forest change: predictions versus current observations. Glob. Planet. Chang. 56 (3–4), 274–296. http://doi.org/10.1016/j.gloplacha.2006.07.028.
Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A.M., Friedl, M., Braswell, B.H., ... Richardson, A.D., 2012. Digital repeat photography for phenological research in forest ecosystems. Agric. For. Meteorol. 152, 159–177. http://doi.org/10.1016/j.agrformet.2011.09.009.
Studer, S., Stöckli, R., Appenzeller, C., Vidale, P.L., 2007. A comparative study of satellite and ground-based phenology. Int. J. Biometeorol. 51 (5), 405–414. http://doi.org/10.1007/s00484-006-0080-5.
Sulla-Menashe, D., Friedl, M.A., Woodcock, C.E., 2016. Sources of bias and variability in long-term Landsat time series over Canadian boreal forests. Remote Sens. Environ. 177, 206–219. http://doi.org/10.1016/j.rse.2016.02.041.
Urbanski, S., Barford, C., Wofsy, S., Kucharik, C., Pyle, E., Budney, J., ... Munger, J.W., 2007. Factors controlling CO2 exchange on timescales from hourly to decadal at Harvard Forest. J. Geophys. Res. Biogeosci. 112 (G2), G02020. http://doi.org/10.1029/2006JG000293.
Verma, M., Friedl, M.A., Finzi, A., Phillips, N., 2016. Multi-criteria evaluation of the suitability of growth functions for modeling remotely sensed phenology. Ecol. Model. 323, 123–132. http://doi.org/10.1016/j.ecolmodel.2015.12.005.
Vermote, E.F., El Saleous, N., Justice, C.O., Kaufman, Y.J., Privette, J.L., Remer, L., ... Tanré, D., 1997. Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: Background, operational algorithm and validation. J. Geophys. Res.-Atmos. 102 (D14), 17131–17141. http://doi.org/10.1029/97JD00201.
White, K., Pontius, J., Schaberg, P., 2014. Remote sensing of spring phenology in northeastern forests: a comparison of methods, field metrics and sources of uncertainty. Remote Sens. Environ. 148, 97–107. http://doi.org/10.1016/j.rse.2014.03.017.
Wu, C., Gonsamo, A., Chen, J.M., Kurz, W.A., Price, D.T., Lafleur, P.M., ... Munger, J.W., 2012. Interannual and spatial impacts of phenological transitions, growing season length, and spring and autumn temperatures on carbon sequestration: a North America flux data synthesis. Glob. Planet. Chang. 92–93, 179–190. http://doi.org/10.1016/j.gloplacha.2012.05.021.
Wulder, M.A., Dechka, J.A., Gillis, M.A., Luther, J.E., Hall, R.J., Beaudoin, A., Franklin, S.E., 2003. Operational mapping of the land cover of the forested area of Canada with Landsat data: EOSD land cover program. For. Chron. 79 (6), 1075–1083. http://doi.org/10.5558/tfc791075-6.
Xu, L., Myneni, R., Chapin Iii, F., Callaghan, T., Pinzon, J., Tucker, C., 2013. Temperature and vegetation seasonality diminishment over northern lands. Nat. Clim. Chang. 3 (6), 581–586.
Yang, X., Tang, J., Mustard, J.F., 2014. Beyond leaf color: Comparing camera-based phenological metrics with leaf biochemical, biophysical, and spectral properties throughout the growing season of a temperate deciduous forest. J. Geophys. Res. Biogeosci. 119 (3), 2013JG002460. http://doi.org/10.1002/2013JG002460.
Zhang, X., Friedl, M.A., Schaaf, C.B., 2006. Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): evaluation of global patterns and comparison with in situ measurements. J. Geophys. Res. Biogeosci. 111 (G4), G04017. http://doi.org/10.1029/2006JG000217.
Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., ... Huete, A., 2003. Monitoring vegetation phenology using MODIS. Remote Sens. Environ. 84 (3), 471–475. http://doi.org/10.1016/S0034-4257(02)00135-9.
Zhou, L., Kaufmann, R.K., Tian, Y., Myneni, R.B., Tucker, C.J., 2003. Relation between interannual variations in satellite measures of northern forest greenness and climate between 1982 and 1999. J. Geophys. Res.-Atmos. 108 (D1), 4004. http://doi.org/10.1029/2002JD002510.
Zhu, Z., Woodcock, C.E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sens. Environ. 118, 83–94. http://doi.org/10.1016/j.rse.2011.10.028.
Zhu, Z., Woodcock, C.E., 2014. Continuous change detection and classification of land cover using all available Landsat data. Remote Sens. Environ. 144, 152–171. http://doi.org/10.1016/j.rse.2014.01.011.