NDVI derived from near-infrared-enabled digital cameras: Applicability across different plant functional types

Gianluca Filippa, Edoardo Cremonese, Mirco Migliavacca, Marta Galvagno, Oliver Sonnentag, Elyn Humphreys, Koen Hufkens, Youngryel Ryu, Joseph Verfaillie, Umberto Morra di Cella, Andrew D. Richardson

A B S T R A C T

Time series of vegetation indices (e.g. normalized difference vegetation index [NDVI]) and color indices (e.g. green chromatic coordinate [GCC]) based on radiometric measurements are now available at different spatial and temporal scales ranging from weekly satellite observations to sub-hourly in situ measurements by means of near-surface remote sensing (e.g. spectral sensors or digital cameras). In situ measurements are essential for providing validation data for satellite-derived vegetation indices. In this study we used a recently developed method to calculate NDVI from near-infrared (NIR) enabled digital cameras (NDVIL) at 17 sites (for a total of 74 year-sites) encompassing six plant functional types (PFT) from the PhenoCam network.

The seasonality of NDVL was comparable to both NDVI measured by ground spectral sensors and by the moderate resolution imaging spectroradiometer (MODIS). We calculated site- and PFT-specific scaling factors to correct NDVL values and recommend the use of site-specific NDVI from MODIS in order to scale NDVL. We also compared GCC extracted from red-green-blue images to NDVI and found PFT-dependent systematic differences in their seasonalities. During senescence, NDVL lags behind GCC in deciduous broad-leaf forests and grasslands, suggesting that GCC is more sensitive to changes in leaf color and NDVL is more sensitive to changes in leaf area. In evergreen forests, NDVL peaks later than GCC in spring, probably tracking the processes of shoot elongation and new needle formation. Both GCC and NDVL can be used as validation tools for the MODIS Land Cover Dynamics Product (MCD12Q2) for deciduous broad-leaf spring phenology, whereas NDVL is more comparable than GCC with autumn phenology derived from MODIS. For evergreen forests, we found a poor relationship between MCD12Q2 and camera-derived phenology, highlighting the need for more work to better characterize the seasonality of both canopy structure and leaf biochemistry in those ecosystems.

Our results demonstrate that NDVL is in excellent agreement with NDVI obtained from spectral measurements, and that NDVL and GCC can complement each other in describing ecosystem phenology. Additionally, NDVL allows the detection of structural changes in the canopy that cannot be detected by visible-wavelength imagery.

1. Introduction

Vegetation phenology (the study of the timing of recurrent biological events) is highly sensitive to climate variability and change (Rosenzweig et al., 2007; Migliavacca et al., 2012; Richardson et al., 2013). Phenological time-series based on radiometric measurements are now available, covering different spatial and temporal scales ranging from weekly satellite observations to sub-hourly in situ measurements by means of, e.g. spectral sensors or digital repeat photography. Regarding satellite-based data, the trade-off between spatial and temporal resolution represents a critical
The Normalized Di\textsubscript{fference Vegetation Index \textnormal{(NDVI, Tucker, 1979)}} has been widely used to monitor the timing and magnitude of the seasonal development of the vegetation and link them to environmental factors such as temperature, precipitation and photoperiod (Jolly et al., 2005). The majority of studies focused on satellite data to retrieve seasonal development of the vegetation and link them to environmental in situ images covering visible-only and of combined visible and NIR images to understand using narrow-band spectral instruments, or retrieved from the MODIS, demonstrating the potential of NIR-enabled cameras. The PhenoCam Network (http://phenocam.sr.unh.edu/webcam/) consists now of 340 sites equipped with networked digital cameras, of which two hundred are NIR-enabled cameras such as the ones used in Petach et al. (2014), thereby allowing to extend the analysis across multiple years and different plant functional types.

In the present paper, we analyzed a large dataset of visible plus NIR images across 17 North American sites encompassing six plant functional types (PFT) for a total of 74 year-sites of data from the PhenoCam image archive. Our objectives are:

(a) to compare NDVI\textsubscript{C} and GCC\textsubscript{C} seasonal trajectories across different PFT and identify potential differences in their phenology;
(b) to compare NDVI\textsubscript{C} and spectral measurements at different scales, including NDVI from MODIS and measured by ground light-emitting diodes (LED) sensors (Ryu et al., 2010);
(c) to examine the consistency between camera-derived phenological transition dates and the MODIS Land Cover Dynamics Product (MCD12Q2) with different methods for deciduous broad-leaf and evergreen needle-leaf forests.

2. Materials and Methods

The seventeen sites included in this study belong to the PhenoCam network (http://phenocam.sr.unh.edu/webcam/) and are located in mid-latitude US and Canada (Table 1). The majority of them also belong to other observational networks such as Fluxnet (https://fluxnet.ornl.gov/). Each site is equipped with a NetCam SC IR security camera (StarDot Technologies, Buena Park, CA), featuring a Micron 1/2\textsuperscript{nd} CMOS active-pixel digital imaging sensor and configured for 1.3 megapixel (1296 × 976) output. Camera channels are centered at 600, 530, 450 nm for red, green and blue, respectively (unpublished data). The camera was set at manual (fixed) white balance and automatic exposure. Five sites were also equipped with LED sensors. Peak sensitivities (and full width half maximum, FWHM) of LED sensors were at 646 ± 2 nm, for red and NIR, respectively. These self-manufactured spectral sensors were first tested by Ryu et al. (2010) and have proven to be comparable to traditional radiometers.

\textit{Extraction of color and vegetation indices.} Digital images were processed using the R package phenopix (Filippa et al., 2016). For each site one or more region-of-interest (ROI) was chosen, which restricts all subsequent analyses on that subset of pixels. ROIs can be viewed on the PhenoCam web page for each site (http://phenocam.sr.unh.edu/webcam/). RGB images were processed to obtain seasonal trajectories of green chromatic coordinates (GCC, Gillespie et al., 1987). Camera
NDVI (NDVIC) was computed based on simultaneous processing of sequentially-captured RGB and NIR + RGB images according to Petach et al. (2014). Briefly, NDVIC is calculated by extracting the red digital numbers (DN) from a RGB-image and the NIR DN from a NIR + RGB-image acquired few seconds apart. Both values are corrected for the exposure of the respective image prior to NDVI calculation.

\[ G_{CC} \text{ time series were filtered with a combination of three filtering algorithms: (a) a threshold filter based on sun position excluding from the processing images acquired with a sun angle lower than 5° above the horizon; (b) a spline filter (Migliavacca et al., 2011) based on recursive spline smoothing and residual computation followed by removal of outliers falling outside an envelope of 5 times the standard deviation of the residuals; (c) a max filter following the method of Sonnentag et al. (2012), based on the identification of the 90th percentile values in a three-day moving window. NDVIE time series were subjected to similar filtering procedures, except that (b) was used with an envelope of 1.8 times the standard deviation of the residuals, an empirically-defined threshold based on residual distribution. We compared the spline filter to a more sophisticated method that finds the optimal span of a smoothing function (local polynomial regression, Cleveland, 1979) fitted through the data, based on Bayesian Information Criterion minimization (LOESS-BIC). In section SI1 we show that our empirical spline method results in phenological information comparable to the LOESS-BIC method. The empirical model is less computationally intense and therefore preferable for dealing with large databases. }

Additionally, we took advantage of an existing snow classification for the PhenoCam image archive flagging each image with snow on canopy or ground (Richardson et al., in review). NDVIE appears particularly sensitive to the presence of snow cover, therefore data points with snow in the image were removed and replaced with a winter baseline value (Si2).

\[ \text{Scaling factors. Petach et al. (2014) showed that NDVIE raw values must be scaled in order to compare them to NDVI measurements obtained from spectral sensors (either in situ or remotely sensed). This step is needed because the spectral response of the CMOS sensor in the R, G, B and NIR of the Stardot camera is different compared to spectral measurements. We scaled NDVIE to the MODIS MCD43A2 product (NDVI}_{MODIS}). We extracted seasonal NDVI}_{MODIS trajectories on 3 x 3 pixel windows centered on each study site. The 9 pixels trajectories were first filtered with the MCD43A2 Bidirectional Reflectance Distribution Function (BRDF) Albedo Quality product. Scaling factors were computed by linear regression between the seasonal course of NDVIE and NDVI}_{MODIS following the formula:} \]

\[ \text{NDVIE} = \text{NDVI}_{MODIS} \times b + c \]  

where \(b\) and \(c\) are the scaling factors. Because MCD43A2 is a 8-day product composing 16 days of images, we used a data-driven approach (SI3) to establish the best matching between NDVIE and NDVI}_{MODIS data, i.e. a 10-day left-aligned moving average of NDVIE.

\[ \text{Comparison of NDVIE with ground NDVI measurements. NDVIE was compared to NDVI measured by LED sensors (NDVIE}_{LED}, Ryu et al., 2010, 2014) at 5 sites: bartlett, vaira, tonzi, harvardburn2 and merblue. LED sensors are pointed nadir and have a field-of-view of 180°. NDVIE}_{LED is typically recorded every half hour. For this analysis we used daily-aggregated data, obtained by averaging values recorded in a two-hour interval around solar noon. For the five sites equipped with LED sensors, we also used NDVIE}_{LED to compute scaling factors.} \]

\[ \text{Extraction of phenological dates from G}_{CC}, \text{NDVIE and comparison with MODIS Land Cover Dynamics Product (MCD12Q2). A double logistic function (Klosterman et al., 2014) was fitted to the seasonal trajectories of NDVIE and G}_{CC}. Four different methods were then used to extract phenological transition dates on the fitted curves (namely \(rs\), \(derivatives\), \(klosterman\) and \(gu\) methods), described in detail in Filippa et al. (2016) and illustrated in Fig. S14. All methods provide an estimation of uncertainty of the transition dates based on the resampling of the time series and extraction of the dates at each iteration (\(n = 500\) as described in Filippa et al. (2016). From the uncertainty ensemble we calculated the median transition date and its mean absolute deviation (MAD).} \]

The MCD12Q2 Land Cover Dynamics Product (spatial resolution 250 m) was obtained on 3 x 3 pixel windows centered on the study sites. For each scene and year four transition dates corresponding to Greenup, Maturity, Senescence and Dormancy were obtained, computed with a method similar to Klosterman et al. (2014). From the nine pixel values we calculated the median and the MAD. The relationship between MCD12Q2-derived and NDVIE-derived dates was evaluated by means of linear regressions in a statistical framework that accounts also for uncertainty. Briefly, each regression analysis was replicated 500 times with observation points sampled (i) from the transition dates uncertainty ensemble, for NDVIE-derived dates, and (ii) from a normally distributed population generated with the 9-pixel median and MAD values for NDVIE-derived dates. This procedure generally results in poorer relationship scores compared to traditional linear regression, but allows for a more robust assessment of the uncertainty of the estimated transition dates. The coefficient of determination \(r^2\), root mean square error (RMSE) and model bias (BIAS) were computed for each of the 500 regressions and averaged.

3. Results

\[ \text{Seasonal trajectories of NDVIE and G}_{CC} across different PFTs. Representative time series of G}_{CC} and NDVIE for each PFT are show in Fig. 1 for 2013 and 2014. Deciduous broad-leaf forests are characterized by synchronous NDVIE and G}_{CC} trajectories until the seasonal maximum of canopy development is reached. During the greendown period (Elmore et al., 2012), a small but consistent departure between G}_{CC} and NDVIE occurs and gets amplified in autumn, when the decrease in NDVIE lags behind G}_{CC} by several days.} \]

\[ \text{Evergreen needle-leaf forests are characterized by an overall smaller seasonal amplitude in both G}_{CC} and NDVIE compared to DBFs. The G}_{CC} seasonal trajectory is roughly sinusoidal with a short baseline in winter. In contrast, NDVIE shows an abrupt increase around the beginning of July for both years. This lag in NDVIE with respect to G}_{CC} is consistent across ENFs. We observed an earlier drop in NDVIE in autumn (beginning of October) as compared to G}_{CC} which also shows a decrease at that time, but smaller.} \]

\[ \text{The seasonal trajectories of NDVIE and G}_{CC} for burnnassagebrush, a dry shrubland, show comparable seasonal dynamics, but differences in the year-to-year variability. In particular, G}_{CC} peak is similar for 2013 and 2014, whereas NDVIE peak is remarkably lower in 2013 than in 2014. More generally, shrubland sites included in this study range from peatlands (merblue) and dry (burnnassagebrush) sites, to water limited sites (jernort and luckyhills), therefore patterns are not all consistent. However, the seasonality of NDVIE in SHBs is in general very similar to that of G}_{CC}. The same is also true for the only cropland site included, i.e. southernreatpains, even though in this PFT, inter-year variability in the behavior of G}_{CC} and NDVIE is probably associated to different crops. Similar to DBFs, Lethbridge grassland site shows a lag in autumn phenology, with NDVIE reaching the autumn baseline about two weeks later than G}_{CC}.} \]

\[ \text{To evaluate the relationship between NDVIE and G}_{CC} across all PFTs in terms of seasonal transition dates, we used the \(gu\) method on both indices. For this analysis, sites showing multiple seasonal peaks (i.e. croplands or water limited ecosystems, for which the double-logic fitting offered poor results) were excluded. Results are shown in Fig. 2 for DBFs and GRAs, and Fig. 3 for ENFs. For DBFs and GRAs, upturn date (UD) occurs between day of year (DOY) 100 and 150 except for ibp, a dry grassland with a short seasonal cycle beginning slightly before DOY 200. Stabilization date (SD) and downturn date (DD) occur over a longer time frame (DOY 110–220 and 200–300, respectively).} \]

\[ \text{Recession date (RD) across DBFs and GRAs spans roughly 50 days (DOY} \]
250–300) for $G_{CC}$ and 100 days (DOY 250–350) for NDVI$_C$. The greening dates (UD and SD) extracted from $G_{CC}$ and NDVI$_C$ are highly correlated and close to the 1:1 line, with an RMSE of about 4–6 days and a bias of about one week. DD is the transition date showing the largest error in the single dates and lower $r^2$ in the relationship. The regression statistics of RD demonstrate a poorer relationship than for spring, but interestingly show a 19-day positive bias (i.e. NDVI$_C$ transition dates occurring later than $G_{CC}$), thereby allowing to generalize the autumn lag between $G_{CC}$ and NDVI$_C$ decrease across GRA and DBF functional types.

Unlike DBFs and GRAs, the methods based on derivatives (i.e. all except trs method) did not give satisfactory results for ENFs. The much lower signal-to-noise ratio in the time series of both NDVI$_C$ and $G_{CC}$ resulted in poor performance of the double-logistic fittings and higher uncertainty in the estimated dates. Instead, the method chosen for ENFs was based on reaching 40% of the amplitude of the seasonal trajectory. Compared to DBFs and GRAs, transition dates for ENFs (Fig. 3) are less consistent between NDVI$_C$ and $G_{CC}$. A positive bias (29 days) in the start of season (sos) was found, indicating that for all ENF sites NDVI$_C$ starts increasing later by about 1 month compared to $G_{CC}$. The DOY of peak production (pop) occurs on average 13 days later for NDVI$_C$ than for $G_{CC}$, whereas the opposite is true in autumn, with a negative bias (22 days) in end of season (eos).

Comparison of NDVI$_C$ with MODIS seasonal trajectories and scaling

**Fig. 1.** Two-year of daily NDVI$_C$ and $G_{CC}$ from a selection of sites included in this study. DBF: deciduous broad-leaf forest; ENF: evergreen needle-leaf forest; SHB: shrubland; CRO: cropland; GRA: grassland.

**Fig. 2.** Relationship between transition dates extracted from $G_{CC}$ (x-axis) and NDVI$_C$ (y-axis) time series for deciduous broad-leaf forests (DBF) and grasslands (GRA). Each data point is a site-year median date. Error bars represent the MAD. The red dashed line is the 1:1, whereas black lines represent the linear fit (solid) and its uncertainty (dashed). Site abbreviations: bar: bartlett; har: harvardbarn2; let: lethbridge; ibp: ibp; wil: willowcreek; alt: alligatorriver. Abbreviations: UD: upturn date; SD: stabilization date; DD: downturn date; RD: recession date. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
year median transition date. Error bars represent the MAD. The red dashed line is the 1:1. Site abbreviations: har: harvardbarn2; tur: turkeypointenf39; can: canadaOBS; ore: oregonMP.

First, for each site NDVI (Table 2). Therefore, for the estimation of PFT-specific scaling factors, we included only sites where the site-specific relationship was higher than 0.3, i.e. 14 out of 17 sites. On this subset we calculated the PFT-specific scaling factors. These PFT-specific scaling factors are then applied to the raw NDVI data (Fig. 4). For DBFs, there is an excellent agreement between scaled NDVI and NDVI$_{M}$, independently on whether NDVI$_{C}$ is scaled according to PFT-specific (grey triangles) or site-specific (black triangles) factors. For ENFs, the relationship between NDVI$_{M}$ and NDVI$_{C}$ is in general poor with tighter relationships for NDVI$_{C}$ scaled according to the site-specific scaling factor compared to the PFT-specific ones. In burnstagebrush, PFT-specific and site-specific scaled NDVI$_{C}$ are very close to each other. In contrast NDVI$_{M}$ shows a lower seasonal amplitude and temporal mismatch with NDVI$_{C}$, with earlier seasonal peaks in all three years included in the comparison. In lebhiride the site-scaled NDVI$_{C}$ is in good agreement with NDVI$_{M}$, whereas the PFT-scaling reduces the seasonal amplitude of NDVI$_{C}$. Spatial heterogeneity may be responsible for the temporal mismatch between NDVI$_{C}$ and NDVI$_{M}$ in southerngreatplains seasonal trajectories, however the PFT-specific scaling factor (extended from GRA) seems to better reproduce the amplitude of NDVI$_{C}$ signal compared to the site-specific factors. Scaling factors computed against data from LED sensors are different compared to those obtained from NDVI$_{M}$ at the same site, except for harvardbarn2 and merbleue (Table 2).

Comparison of in situ measurements: NDVI$_{LED}$ and NDVI$_{C}$. Fig. 5 shows a three-year time series comparing NDVI$_{LED}$ and NDVI$_{C}$. Overall, the seasonalities of NDVI$_{C}$ and NDVI$_{LED}$ are in excellent agreement. In DBF, the two sensors provide consistent seasonal NDVI trajectories both in terms of timing and magnitude. At bartlett, NDVI$_{LED}$ appears to be more affected by the presence of snow compared to NDVI$_{C}$. At vaira, the magnitude of NDVI$_{C}$ is lower than that of NDVI$_{LED}$, but the timing is similar across the three years. At tonzi and merbleue, both the magnitude and seasonality of NDVI$_{LED}$ and NDVI$_{C}$ are consistent.

<table>
<thead>
<tr>
<th>Site-specific</th>
<th>PFT</th>
<th>Slope (b)</th>
<th>Intercept (c)</th>
<th>$r^2$</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>southerngreatplains</td>
<td>CRO</td>
<td>0.3 ± 0.33</td>
<td>0.42 ± 0.01</td>
<td>0.12</td>
<td>102</td>
</tr>
<tr>
<td>alligatorriver</td>
<td>DBF</td>
<td>0.47 ± 0.070</td>
<td>0.7 ± 0.00</td>
<td>0.88</td>
<td>109</td>
</tr>
<tr>
<td>bartlett</td>
<td>DBF</td>
<td>0.55 ± 0.06</td>
<td>0.71 ± 0.00</td>
<td>0.86</td>
<td>191</td>
</tr>
<tr>
<td>bartlett (LED)</td>
<td>DBF</td>
<td>0.83 ± 0.02</td>
<td>0.60 ± 0.00</td>
<td>0.79</td>
<td>662</td>
</tr>
<tr>
<td>freemanwood</td>
<td>DBF</td>
<td>0.49 ± 0.44</td>
<td>0.54 ± 0.07</td>
<td>0.40</td>
<td>35</td>
</tr>
<tr>
<td>harvardbarn2</td>
<td>DBF</td>
<td>0.61 ± 0.09</td>
<td>0.76 ± 0.01</td>
<td>0.84</td>
<td>141</td>
</tr>
<tr>
<td>harvardbarn2 (LED)</td>
<td>DBF</td>
<td>0.56 ± 0.01</td>
<td>0.75 ± 0.00</td>
<td>0.82</td>
<td>1208</td>
</tr>
<tr>
<td>willowcreek</td>
<td>DBF</td>
<td>0.78 ± 0.1</td>
<td>0.66 ± 0.00</td>
<td>0.90</td>
<td>100</td>
</tr>
<tr>
<td>oregonMP</td>
<td>ENF</td>
<td>45.29 ± 38.07</td>
<td>-9.38 ± 8.46</td>
<td>0.00</td>
<td>152</td>
</tr>
<tr>
<td>merbleue</td>
<td>GRA</td>
<td>1.25 ± 0.38</td>
<td>0.54 ± 0.1</td>
<td>0.72</td>
<td>69</td>
</tr>
<tr>
<td>lebhiride</td>
<td>GRA</td>
<td>0.80 ± 0.14</td>
<td>0.44 ± 0.02</td>
<td>0.79</td>
<td>127</td>
</tr>
<tr>
<td>tonzi</td>
<td>GRA</td>
<td>0.60 ± 0.19</td>
<td>0.63 ± 0.06</td>
<td>0.52</td>
<td>148</td>
</tr>
<tr>
<td>tonzi (LED)</td>
<td>ENF</td>
<td>5.94 ± 3.12</td>
<td>-0.47 ± 0.6</td>
<td>0.41</td>
<td>96</td>
</tr>
<tr>
<td>vaira</td>
<td>GRA</td>
<td>0.66 ± 0.17</td>
<td>0.62 ± 0.05</td>
<td>0.63</td>
<td>137</td>
</tr>
<tr>
<td>vaira (LED)</td>
<td>ENF</td>
<td>1.06 ± 0.01</td>
<td>0.63 ± 0.00</td>
<td>0.88</td>
<td>1274</td>
</tr>
<tr>
<td>canadaOBS</td>
<td>MF</td>
<td>2.48 ± 1.07</td>
<td>0.56 ± 0.00</td>
<td>0.44</td>
<td>116</td>
</tr>
<tr>
<td>burnstagebrush</td>
<td>SHB</td>
<td>1.66 ± 0.90</td>
<td>0.64 ± 0.25</td>
<td>0.35</td>
<td>108</td>
</tr>
<tr>
<td>jernort</td>
<td>SHB</td>
<td>0.19 ± 0.09</td>
<td>0.22 ± 0.02</td>
<td>0.57</td>
<td>63</td>
</tr>
<tr>
<td>luckyhills</td>
<td>SHB</td>
<td>0.44 ± 0.64</td>
<td>0.36 ± 0.22</td>
<td>0.22</td>
<td>33</td>
</tr>
<tr>
<td>merbleue</td>
<td>SHB</td>
<td>1.88 ± 0.54</td>
<td>0.67 ± 0.00</td>
<td>0.72</td>
<td>80</td>
</tr>
<tr>
<td>merbleue (LED)</td>
<td>DBF</td>
<td>1.38 ± 0.01</td>
<td>0.65 ± 0.00</td>
<td>0.92</td>
<td>965</td>
</tr>
<tr>
<td>PFT-specific</td>
<td>ENF</td>
<td>0.57 ± 0.06</td>
<td>0.70 ± 0.00</td>
<td>0.72</td>
<td>576</td>
</tr>
<tr>
<td>enfield</td>
<td>SHB</td>
<td>1.70 ± 0.56</td>
<td>0.47 ± 0.05</td>
<td>0.41</td>
<td>212</td>
</tr>
<tr>
<td>GRA</td>
<td>0.59 ± 0.21</td>
<td>0.52 ± 0.05</td>
<td>0.20</td>
<td>481</td>
<td></td>
</tr>
<tr>
<td>shaw</td>
<td>1.83 ± 0.17</td>
<td>0.67 ± 0.03</td>
<td>0.88</td>
<td>251</td>
<td></td>
</tr>
<tr>
<td>Petach et al. (2014)</td>
<td>DBF</td>
<td>0.53 ± 0.02</td>
<td>0.84 ± 0.01</td>
<td>0.89</td>
<td>18</td>
</tr>
</tbody>
</table>

Fig. 3. Relationship between transition dates extracted from GCC (x-axis) and NDVI$_{C}$ (y-axis) for ENFs. The extraction method used here is a fixed 40% threshold (ts). Each point is a site-year median transition date. Error bars represent the MAD. The red dashed line is the 1:1. Site abbreviations: har: harvardbarn2; tur: turkeypointenf39; can: canadaOBS; ore: oregonMP. Abbreviations: trs.sos: threshold-based start of season; trs.pop: day of peak of production; trs.eos: threshold-based end of season. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
transition dates. Consistently negative BIAS scores suggest earlier MODIS dates compared to either GCC and NDVI$_C$ but with considerably lower values when dates are extracted with $g_u$ and Klosterman methods. The $r^2$ of the relationships indicate an explained variance higher than 65% for greenup, being lower for the derivatives method compared to others. For autumn dormancy of DBFs, the best relationship is provided by the $g_u$ method applied to NDVI time series, with RMSE = 10 days, bias = 4 days and $r^2 > 0.7$.

Compared to DBFs, the relationship between MODIS- and pheno-camera-derived transition dates at ENF sites was generally worse. For spring, the best result is achieved by the Klosterman method applied to NDVI ($r^2 = 0.23$, RMSE = 13, bias = $-26$). Noteworthy is the fact that we found opposite signs for BIAS for NDVI$_C$ and GCC, with NDVI$_C$ resulting in later greenup compared to MODIS and the opposite for GCC. For dormancy, the derivatives approach applied to GCC time series provides the best scores ($r^2 = 0.26$, RMSE = 28, BIAS = $-6$).

4. Discussion

4.1. Robustness of NDVI$_C$ measurements

This study provides the first analysis of a large dataset (74 site-years) of visible and visible + NIR images for the computation of paired GCC and NDVI$_C$ using an image archive acquired by NIR-enabled digital cameras. After the pioneering work of Petach et al. (2014), several questions remained open. Does the relationship between NDVI$_C$ and spectral measurements verified on a single year above a deciduous canopy hold across multiple sites and PFTs? Can the proposed scaling factors – calculated for a deciduous canopy – be generalized across sites and, potentially, PFTs?

As for the first question, Fig. 5 demonstrates excellent agreement between NDVI$_{LED}$ and NDVI$_C$ across five different sites and four PFTs. These results indicate that the seasonal trajectory of NDVI$_C$ mirrors that of NDVI measurements obtained from spectral sensors across a range of ecosystems, even those characterized by a different canopy cover fraction (as for example at tonzi, a savanna with mixed grassland and deciduous trees). At Bartlett, the NDVI$_{LED}$ signal seems to deteriorate in winter likely because of snow on the ground, a less relevant problem for NDVI$_C$ because the analyzed ROI is specifically tailored on the tree crowns, therefore avoiding the integration of bare soil, whereas the nadir view of the LED sensors is more affected by background spectroscopic properties. This advantage of digital images over spectral measurements could be fully exploited in evergreen forests, where the presence of snow on trees often causes very noisy NDVI signals. Unfortunately the lack of radiometric NDVI measurements with a similar view angle compared to digital image acquisition prevents us from further exploring this hypothesis and future work should be oriented in this direction.

As for the second question, our data demonstrate that scaling factors may have great variation across sites (Table 2). This may be due to several reasons: (a) the different view angle of pheno-cameras and satellite data, (b) the different bands used for NDVI calculation in NDVI$_M$ and NDVI$_C$ (Hufkens et al., 2012), (c) a spatial mismatch between MODIS and pheno-cameras especially in fragmented landscapes, because the ROI of the cameras is specifically tailored on the canopies. In fact at two sites with low fractional cover (tonzi and vaira) scaling...
factors computed against ground LED sensors show better relationship than with MODIS, whereas at merbleue, where the vegetation cover is homogeneous and LED, cameras and satellite likely observe similar targets, scaling factors are more similar. This supports the hypothesis that spatial mismatch plays a role in the variability of scaling factors. Overall, scaling factors for DBFs are the most consistent, and they approach the values reported by Petach et al. (2014) (see Table 2).

We recommend the use of site-specific NDVI<sub>M</sub> data to scale NDVI<sub>C</sub> values. In the case that the MODIS pixel is not representative of the vegetation targeted by the cameras (which can be checked by visually comparing the seasonal trajectories), NDVI<sub>C</sub> should be scaled with ground spectral NDVI data, if available. Otherwise, PFT-specific scaling factors reported in Table 2 can be used for DBFs. For other PFTs, scaling factors computed in this study should be used with caution.

4.2. The seasonality of NDVI<sub>C</sub> and GCC

The seasonal trajectory of NDVI<sub>C</sub> and GCC shows remarkable differences (Fig. 1) which result in different transition dates (Figs. 2 and 3), and these differences are PFT-specific. Autumn decrease occurs

Fig. 5. Time series of NDVI<sub>LED</sub>, NDVI<sub>M</sub> and scaled NDVI<sub>C</sub>. The coefficients of determination of the relationship between NDVI<sub>C</sub> and NDVI<sub>LED</sub> are 0.79, 0.82, 0.88, 0.82 and 0.92 for bartlett, harvardbarn2, vaira, tonzi and merbleue, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 6. Scores (i.e. root mean square error (RMSE), bias and r<sup>2</sup>) of the linear relationships between near surface phenology (NDVI<sub>C</sub> and GCC) and MODIS Land Cover Dynamics Product (MCD12Q2). Scores are shown for four different methods for the extraction of transition date (trs, derivatives, klosterman and gu, the four bars in the same subplot), separately for DBFs and ENFs (light and dark green backgrounds, respectively), and for NDVI<sub>C</sub>-based and GCC-based phenology (blue and red bars, respectively). Spring transition dates are featured to the left, whereas autumn dates to the right. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
somewhat slower for NDVI_{C} than for G_{CC} in DBFs and GRAs, resulting in a later start of dormancy of NDVI_{C} compared to G_{CC}. This suggests that G_{CC} is more sensitive to the amount of senescent material in the canopy (ratio of dry/green biomass) and leaf area index compared to NDVI_{C}, which in turn is more indicative of the leaf abscission and shedding processes during the senescence period. These findings agree with Keenan et al. (2014), who reported that declines in autumn leaf area index lagged declines in autumn G_{CC}, reflecting changing leaf color before actual leaf abscission in a temperate DBF.

The autumn lag is apparent also for temperate grasslands. A study combining phenocams and radiative transfer modelling in a Mediterranean grassland suggests that the lag between NDVI_{C} and G_{CC} in the senescence period (dry-down in summer) is the result of different sensitivity of G_{CC} to the amount of senescent material in the canopy compared to NDVI_{C}. G_{CC} is shown to be more sensitive to the dry/green biomass than NDVI_{C}, while NDVI_{C} responds more to variation of LAI in the range between 1 and 3 m^{2}/m^{2} (Liao et al., in review). In terms of transition dates, different temporal dynamics of NDVI_{C} and G_{CC} result in a 20-day lag in the end of season (recession date, RD) for DBFs and GRAs. This might have important implications when these data streams are used to constrain or to develop phenology models.

ENFs are characterized by a lower seasonal amplitude in both NDVI_{C} and G_{CC}, and a lower signal-to-noise ratio, preventing us from using derivative-based extraction methods. Our results suggest that in such ecosystems, the most robust method for transition date computation is threshold-based (Fig. 3). With the trs method we were able to identify a generalized later spring increase in NDVI_{C} compared to G_{CC} quantified at 29 days across all ENFs (Fig. 3).

An hypothesis for the asynchronous spring ramp in G_{CC} and NDVI_{C} is the increase in foliar biomass associated with shoot elongation (Wingate et al., 2015) and the formation of new needles (Ryu et al., 2014), with no substantial change in the canopy greenness. In Fig. 7 we show the smoothed time course of G_{CC} and NDVI_{C}, along with representative photographs at turkeypointenf39 in 2013. The evergreen canopy (eastern white pine, *Pinus strobus*) turns from brownish in early April to green in May. This transition is characterized by a sharp increase in G_{CC}, while NDVI_{C} remains low and rather constant. The following period is characterized by only slight increase in G_{CC} and a sharp NDVI_{C} ramp. This period corresponds to the new needle formation; in *Pinus strobus* new needles emerge in the third decade of May and young needles tend to be completely formed at the end of June (National Phenology Network, 2017). Our data suggests therefore that while the level of green is already close to the maximum in spring, NDVI_{C} tracks foliar biomass increase associated to new needle formation. The decrease in NDVI_{C} in September corresponds to the ripening period of seed cones which turn from green to red, and are fully developed in early October (National Phenology Network, 2017). The significant presence of red cones may cause the observed decline in NDVI_{C}. Afterwards, cones begin dropping, resulting in green canopy dominating again the region-of-interest and producing a slight increase in NDVI_{C} in November, before the onset of winter. This comparison with ground phenology suggests that in boreal and temperate ENFs, NDVI_{C} allows the detection of structural changes in the canopy that cannot be detected by G_{CC}. 

Fig. 7. Seasonal course of 20-days-moving-average G_{CC} and NDVI_{C} values at TurkeyPointenf39 in 2013, along with selected photographs. Vertical lines in the time series correspond to the dates when the above pictures were taken. Phenological information (in blue) is obtained from aggregated phenological data collected in 2013 for about 300 Eastern White Pine individuals located in the Eastern US areas (data not shown, National Phenology Network, 2017). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

![Seasonal course of 20-days-moving-average G_{CC} and NDVI_{C} values at TurkeyPointenf39 in 2013, along with selected photographs.](image-url)
4.3. Comparison of camera-derived metrics and the MODIS Land Cover Dynamics Product (MCD12Q2)

Phenocamera networks provide invaluable ground data for the validation of phenology satellite products. Our analysis suggests that both $G_{CC}$ and NDVI$_C$ can be used as validation tools for MCD12Q2 for DBF spring phenology, provided that transition dates are extracted with the $gu$ or klosterman methods. Spring greenup is detected earlier by MODIS by about one week compared to both NDVI$_C$ and $GCC$. This is likely because the oblique view of the camera solely tracks tree phenology whereas the nadir satellite view is more affected by the earlier greenup of the understory (Ryu et al., 2014). For autumn phenology, NDVI$_C$-dates extracted with the $gu$ method perform better than other methods, and better than $GCC$. A previous study comparing $GCC$ transition dates and MCD12Q2 across North American DBF showed a consistent bias for dormancy, with dates predicted from $GCC$ occurring 12 days earlier than those predicted from MCD12Q2 (Klosterman et al., 2014). In our study, the positive BIAS between $GCC$-derived and MODIS-derived dates is even higher (17 days), whereas by using NDVI$_C$-derived dates not only we obtain a better relationship than for $GCC$ but also an opposite and reduced BIAS (-4 days, Fig. 8).

Our study confirms better coherence in spring rather than in autumn DBF phenology. Hufkens et al. (2012) argued that more uncertain senescence dates are due to the fact that the rate of change in any vegetation index is more gradual in autumn. Furthermore, while greenup tends to be more homogeneous in space, higher heterogeneity in the canopy exists during senescence (Klosterman et al., 2014). This is likely the cause of higher uncertainty in MODIS-derived than in camera-derived dates (Fig. 8, right column). The discrepancy between camera- and satellite-derived senescence dates can also be associated to different degrees of deciduous or mixed canopy cover, with lower bias with increasing degree of deciduous species over evergreens (Klosterman et al., 2014).

For ENFs, the poor relationship between MCD12Q2 and camera-derived transition dates suggest the we are not yet able to properly describe the phenology of such ecosystems by means of currently available vegetation and color indices (Jönsson et al., 2010). One of the main problems in detecting evergreen phenology at high latitude from satellite is the small amplitude in the greenness signal combined with the presence of snow cover (Delbart et al., 2005). This was the case for all evergreen sites included in this study. Encouraging is the fact that the phenocamera approach can overcome the problem of background by properly choosing the region-of-interest of the analysis. Additionally, a growing body of research is currently addressing new vegetation indices including the chlorophyll/carotenoid index (CCI, Gamon et al., 2016), the green/red vegetation index (GRVI, Nasahara and Nagai, 2015), particularly suitable for describing evergreen phenology, representing promising candidates for comparison with
vegetation indices such as NDVI.

5. Conclusion

In this work we demonstrate that NDVI measurements derived from NIR-enabled digital cameras are robust and comparable to NDVI measured by ground spectral sensors (NDVIg) or satellite (MODIS) NDVI measurements. We show that scaling factors have to be applied to NDVI in order to scale it to spectral NDVI measurements. Moreover, we demonstrate that simultaneous NDVIg and GCC time series provide complementary phenological information by tracking different canopy-level processes, with potentially important implications for phenology models. We furthermore illustrate to what extent and how NDVIg and GCC can be used as ground validation for the MODIS Land Cover Dynamics Product.

In the future, satellite imagery (eg. with the recently launched Sentinel suite) will provide 5-days, 10 m resolution remotely sensed data less prone to spatial and temporal mismatch, likely improving the relationship between near- and far-remote sensing, especially in ecosystems, such as evergreen forests, where this relationship is currently poor.

Acknowledgements

The development of PhenoCam has been supported by the Northeastern States Research Cooperative, NSF's Macrosystems Biology program (award EF-1065029 and EF-1702697), DOE’s Regional and Northeastern States Research Cooperative, NSF's Macrosystems Biology program (award DE-SC0016011), and the US National Park Service Inventory and Monitoring Program and the USA National Phenology Network (grant number G10AP00129 from the United States Geological Survey). We acknowledge additional support, through the National Science Foundation’s LTER program, for research at Harvard Forest (DEB-1237491), and Bartlett Experimental Forest (DEB-1114804). The National Research Foundation of Korea (NRF) supported making LED sensors (NRF-2011-0030485).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.agrforest.2017.11.003.

References


