Agricultural and Forest Meteorology 220 (2016) 141–150

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

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

Phenopix: A R package for image-based vegetation phenology

Gianluca Filippa[a,*], Edoardo Cremonese[a], Mirco Migliavacca[b], Marta Galvagno[a], Matthias Forkel[b,1], Lisa Wingate[c], Enrico Tomelleri[d], Umberto Morra di Cella[a], Andrew D. Richardson[e]

[a] Environmental Protection Agency of Aosta Valley, ARPA Valle dAosta, Climate Change Unit, Italy
[b] Max Planck Institute for Biogeochemistry, Department Biogeochemical Integration, Jena, Germany
[c] INRA, UMR ISPA, Bordeaux, France
[d] Institute for Applied Remote Sensing, EURAC, Bolzano, Italy
[e] Harvard University, Department of Organismic and Evolutionary Biology, Cambridge, MA, USA

ARTICLE INFO

Article history:
Received 15 October 2015
Received in revised form
29 December 2015
Accepted 7 January 2016

Keywords:
Image analysis
Community ecology
Pixel-based analysis
Phenology

ABSTRACT

In this paper we extensively describe new software available as a R package that allows for the extraction of phenological information from time-lapse digital photography of vegetation cover. The `phenopix` R package includes all steps in data processing. It enables the user to: draw a region of interest (ROI) on an image; extract red green and blue digital numbers (DN) from a seasonal series of images; depict greenness index trajectories; fit a curve to the seasonal trajectories; extract relevant phenological thresholds (phenophases); extract phenophase uncertainties.

The software capabilities are illustrated by analyzing one year of data from a selection of seven sites belonging to the PhenoCam network (http://phenocam.sr.unh.edu/), including an unmanaged subalpine grassland, a tropical grassland, a deciduous needle-leaf forest, three deciduous broad-leaf temperate forests and an evergreen needle-leaf forest. One of the novelties introduced by the package is the spatially explicit, pixel-based analysis, which potentially allows to extract within-ecosystem or within-individual variability of phenology. We examine the relationship between phenophases extracted by the traditional ROI-averaged and the novel pixel-based approaches, and further illustrate potential applications of pixel-based image analysis available in the `phenopix` R package.

1. Introduction

Traditional monitoring of plant phenology relies on direct human observations of discrete phenological events, or phenophases, such as bud-burst, flowering, autumn decolouring, and leaf-fall (e.g. Lechowicz, 1984; Richardson et al., 2006; Galvagno et al., 2013; Migliavacca et al., 2008; Filippa et al., 2015). Such observations are typically made on a limited number of individual organisms, across a limited geographic area (i.e., often at a specific research site). On the other hand, satellite remote sensing allows observing land surface phenology on regional to global scales but has a limited representativeness for phenological changes at ecosystem or species-level (White and Nemani, 2006; Delbart et al., 2005; Busetto et al., 2010; Hufkens et al., 2012; Forkel et al., 2015). At an intermediate scale, near-surface remote sensing of phenology makes use of radiometric instruments or imaging sensors. Near-surface remote sensing quantifies, at high temporal resolution, and with a flexible degree of spatial integration (i.e., the potential to look across the canopy as a whole, but at the same time focus on individual organisms), seasonal changes in the optical properties of vegetated surfaces (e.g. Jenkins et al., 2007; Richardson et al., 2007; Soudani et al., 2012). Recent studies have demonstrated the soundness of digital cameras as multi-channel imaging sensors (Richardson et al., 2009; Klosterman et al., 2014; Wingate et al., 2015; Migliavacca et al., 2011).

In the recent literature, different approaches have been used to process digital images of the vegetation canopy. Several researchers have exhaustively addressed the issues of data quality and data filtering in order to reduce noise in the seasonal trajectories of greenness (e.g. Sonnentag et al., 2012; Julitta et al., 2014; Migliavacca et al., 2011). Other authors focused on curve fitting/smoothing methods for extracting dates from phenological time-series (i.e. phenophases, see e.g. Zhang et al., 2003; Beck et al., 2006; Gu et al., 2009; Elmore et al., 2012; Klosterman et al., 2014

* Corresponding author.
E-mail address: g.filippa@arpa.vda.it (G. Filippa).
[1] Now at: Technische Universitäen Department of Geodesy and Geoinformation Research Group Remote Sensing, Vienna, Austria.

http://dx.doi.org/10.1016/j.agrformet.2016.01.006
0168-1923/

Table 1
Main characteristics of the selected PhenoCam sites.

Site ID	Coords (°)	Elev (m)	PFT	Camera type	Dominant species
Torgnon-ND	45.8 N 7.6 E	2160	GRA	Nikon D5000	*Nardus stricta*
Kamuela	20.0 N −155.7 E	850	GRA	Stardot	C3/C4 grasses
Bartlett	44.1 N −71.3 E	268	DBF	Axis 211	*Acer rubrum*; *Fagus grandifolia*
Harvard	42.5 N −72.1 E	340	DBF	Stardot	*Quercus rubra*; *Acer rubrum*
Mammothcave	37.2 N −86.1 E	226	DBF	Stardot	*Quercus sp.*; *Carya sp.*
Torgnon-LD	45.8 N 7.6 E	2091	DNF	Nikon D5000	*Larix decidua*
Harvardhemlock	42.5 N −72.2 E	345	ENF	Stardot	*Tsuga canadensis*

PFT: plant functional type, GRA: grassland, DBF: deciduous broadleaf forest, DNF: deciduous needle-leaf forest, ENF: evergreen needle-leaf forest.

and the `greenbrown` R package). The full exploitation of the existing plethora of methods and approaches would benefit from a comprehensive framework in which to compare them across ecosystem types and climate conditions.

Phenological networks based on the analysis of digital images are growing worldwide. In the US, PhenoCam network includes over 200 sites that use a standard camera and follow a standard protocol (phenocam.sr.unh.edu/webcam/). In Europe, the EUROPHEN network consists of about 60 flux sites equipped with digital cameras (Wingate et al., 2015). Similar networks are also present in Australia (www.aceas.org.au, Brown et al. in preparation) and Asia (Nasahara and Nagai, 2015). The increasing use of repeated digital

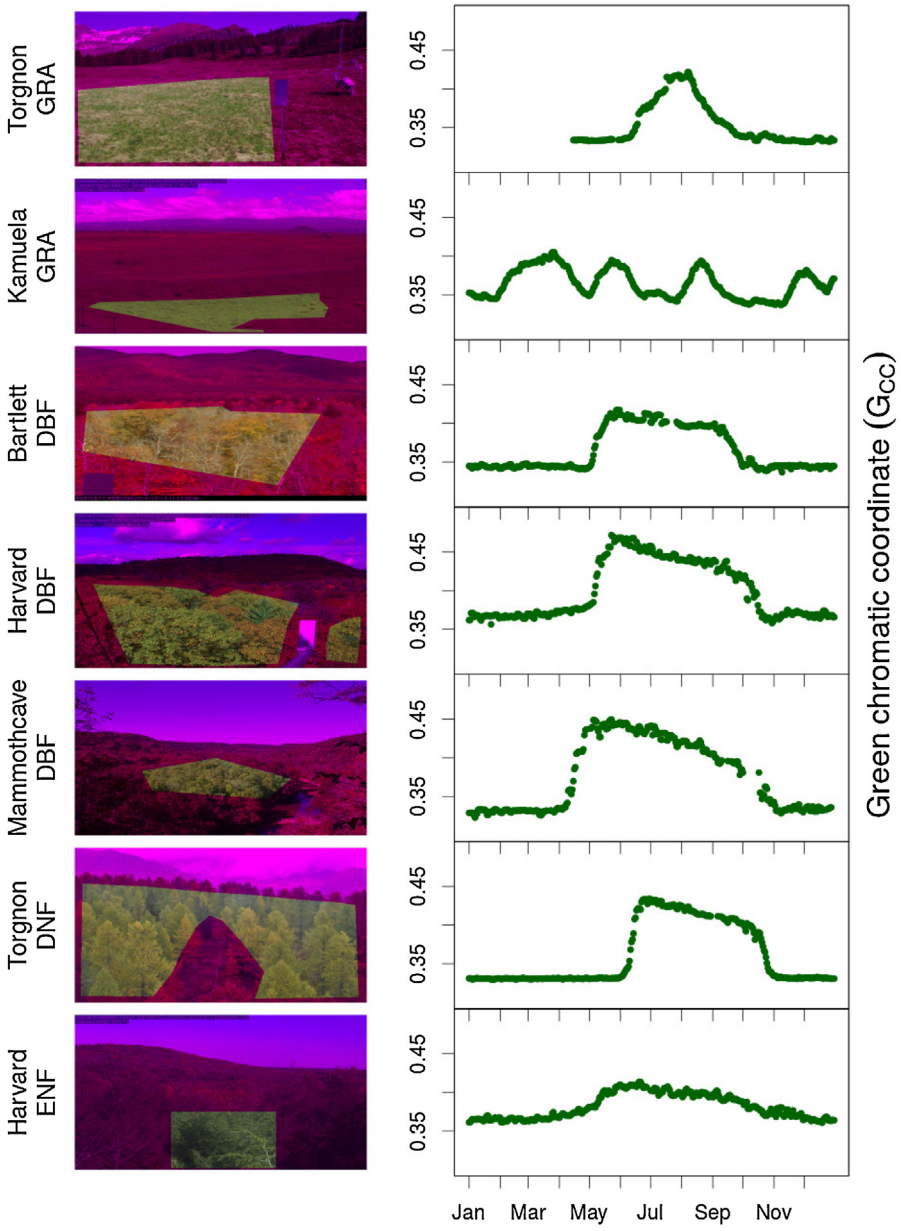

Fig. 1. A sample image from all PhenoCam sites included in this study masked on the chosen region of interest (ROI), along with the seasonal trajectories of filtered G_{CC} in year 2013. Ecosystem type abbreviations are as in Table 1.

photography for phenological research highlights the need of easy-to-use, open source and flexible software tools for the processing of the images. Additionally, the existing networks are poorly coordinated in terms of camera types, settings and sampling protocols, thus presenting a big challenge to build a flexible tool capable of facing the large diversity of ecosystems, image quality and setups.

In this paper we present a collection of functions packed in a software available as a R package (R Core Team, 2015), called `phenopix` (http://r-forge.r-project.org/projects/phenopix/). We will first show the main features of the package and then illustrate its application on a selection of sites belonging to the PhenoCam dataset. Lastly, a section focused on pixel-based analysis will: (1) examine the relationship between phenological thresholds (phenophases) extracted from the average seasonal trajectory of greenness over a region of interest (ROI-averaged approach) and from each pixel of such region (pixel-based approach); (2) illustrate potential applications of pixel-based image analysis to discriminate between subtly different phenological seasonal trajectories within the same image scene.

2. PhenoCam sites

The study sites used to illustrate the functionality of the `phenopix` package belong to the PhenoCam network (phenocam.sr.unh.edu/webcam/) and are illustrated in Table 1 and Fig. 1.

3. Main functions

The typical work-flow of the `phenopix` package is summarised in the flowchart shown in Fig. 2. First, one (or more) region(s) of interest (ROI) is (are) chosen, then digital colour numbers are extracted from the ROI of each image, and processed to obtain a seasonal time series. After filtering the time series, data are fitted with either a double logistic equation or a smoothing curve, on which phenological thresholds (phenophases) are extracted. Finally, uncertainty of the fit and of phenophases can be computed.

3.1. Regions of interest (ROIs)

The scene of the picture rarely includes only the targeted vegetation canopy, thus the user will want to choose a particular region within the scene for analysis. Even more often, more than one region may be of interest, for example in a mixed forest one might independently analyse different deciduous species and evergreen trees (e.g. Ahrends et al., 2008). The function `DrawROI()` allows the user to draw one or more regions of interest on-screen, using the mouse cursor on a chosen reference picture.

3.2. Extract vegetation indices

From the digital colour values of each image the green chromatic coordinate (G_{CC}) is computed. G_{CC} is a vegetation index derived from photographic image and quantifies the greenness relative to the total brightness. G_{CC} is computed as follows:

$$G_{CC} = \frac{G_{DN}}{R_{DN} + G_{DN} + B_{DN}} \quad (1)$$

where G_{DN}, R_{DN}, B_{DN} are the green, red and blue digital numbers, respectively (Gillespie et al., 1987). Similarly, chromatic coordinates of red and blue (R_{CC} and B_{CC}) are also computed. Several indices based on RGB colours have been developed in the last years, including for example the green excess index (GEI) (Woebbecke et al., 1995; Mizunuma et al., 2013). Some authors also used a combination of G_{CC} and R_{CC} to extract autumn phenophases (e.g. Klosterman et al., 2014). For simplicity, all subsequent analyses will be focused on G_{CC} but `phenopix` allows the analysis of the whole variety of colour indices, including the computation of new ones.

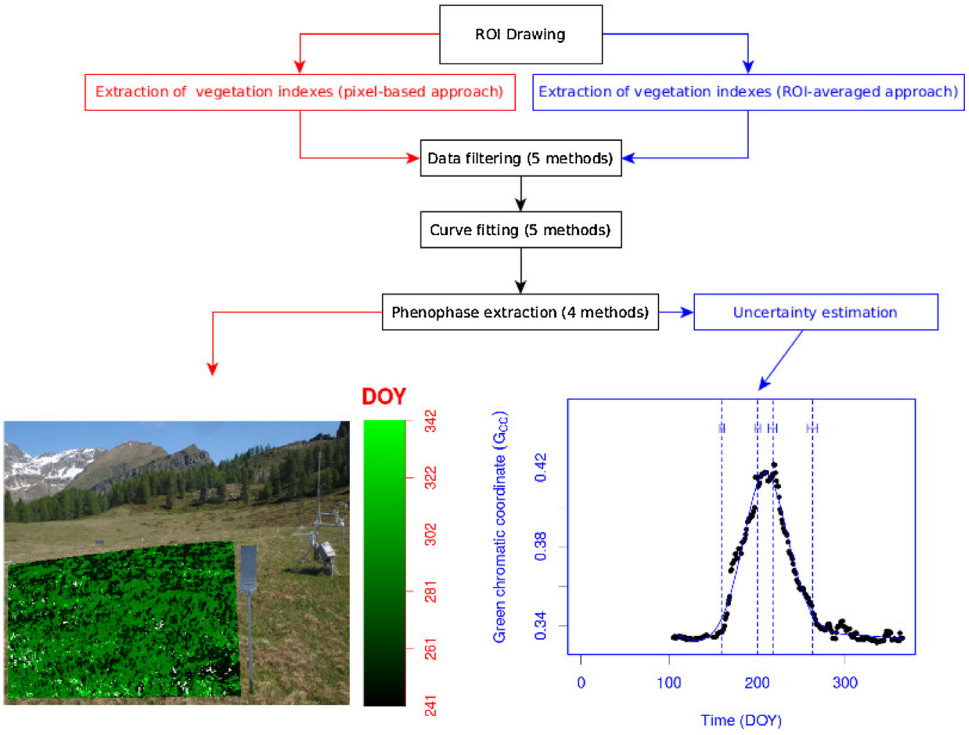

Fig. 2. The work-flow of the processing chain in the `phenopix` R package. Red and blue objects denote specific features of pixel-based and ROI-averaged approaches, respectively. Pixel-based approach produces a phenophase map (left), whereas ROI-averaged approach results in a G_{CC} seasonal trajectory and a certain number of phenophases, depending on the method chosen (right). The example is from Torgnon-ND grassland site in year 2013. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Double logistic model options available in `phenopix`, along with their option names as implemented in the package. The first two methods were first implemented in the `greenbrown` R package and described in Forkel et al. (2015).

Equation	Reference	Option name
$f(t) = mn + (mx - mn) \cdot \left(\frac{1}{1+e^{(-rsp*(t-sos))}} + \frac{1}{1+e^{(rau*(t-eos))}}\right)$	Beck et al. (2006)	'beck'
$f(t) = m_1 + (m_2 - m_7 t) \cdot \left(\frac{1}{1+e^{(m'_3-t)/m'_4}} - \frac{1}{1+e^{(m'_5-t)/m'_6}}\right)$	Elmore et al. (2012)	'elmore'
$f(t) = (a_1 t + b_1) + (a_2 t^2 + b_2 t + c) \cdot \left(\frac{1}{[1+q_1 e^{-h_1(t-n_1)}]^{v_1}} - \frac{1}{[1+q_2 e^{-h_2(t-n_2)}]^{v_2}}\right)$	Klosterman et al. (2014)	'klosterman'
$f(t) = y_0 + \frac{a_1}{[1+e^{-(t-t_{01})/b_1}]^{c_1}} - \frac{a_2}{[1+e^{-(t-t_{02})/b_2}]^{c_2}}$	Gu et al. (2009)	'gu'

Table 3
Methods for phenophase extraction available in `phenopix`.

Name	Phenophases	Description
trs	sos, eos, los, pop, mgs, peak, msp, mau	Based on a user-defined threshold of seasonal development of G_{CC}
derivatives	as TRS plus rsp and rau	Based on local extremes in the first derivative
klosterman	Greenup, Maturity, Senescence, Dormancy	Based on local extremes in the rate of change of curvature k (Kline, 1998)
gu	UD, SD, DD, RD, prr, psr	Based on a combination of local maxima in the first derivative (Gu et al., 2009)

sos: start of season, eos: end of season, los: length of season, pop: peak of season position, peak: maximum seasonal G_{CC}, mgs: mean growing season G_{CC}, msp: mean spring G_{CC}, mau: mean autumn G_{CC}, rsp: rate of spring greenup, rau: rate of autumn senescence, UD: upturn date, SD: stabilisation date, DD: downturn date, RD: recession date, prr: peak recovery rate, psr: peak senescence rate.

The function `extractVIs()` extracts raw red, green and blue digital numbers from each pixel in the ROI, and computes colour chromatic coordinates as in Eq.(1). Vegetation indices can be computed on ROIs with 2 approaches (Fig. 2): (1) the ROI-averaged approach: colour chromatic coordinates are computed for each pixel and then averaged over the whole ROI, and (2) the pixel-based approach, where each pixel belonging to the ROI is analysed separately (Section 5). The procedure is repeated for each image in the archive. A time-stamp is retrieved from the file name of the image and a time series of the computed indices is returned. Specific rules must be followed in naming the image files, and the reader is referred to the package help pages for more details.

3.3. Data filtering

Data retrieved from images often need robust methods for filtering the time series. Bad weather conditions, low illumination and dirty lenses are among the most common issues that determine noise in the time series of vegetation indices. Previous studies (e.g. Sonnentag et al., 2012; Julitta et al., 2014; Migliavacca et al., 2011; Papale et al., 2006) provide the background upon which the filtering techniques implemented in `phenopix` were chosen. We designed a function `autoFilter()` based on 5 different approaches.

(1) A filter called *night* which removes G_{CC} records lower than a certain threshold, by default 0.2. This filters removes night data or images with very scarce illumination (e.g. early morning or sunset images, cloudy days, etc.).

(2) A filter called *blue*, based on blue chromatic coordinate (B_{CC}) (Julitta et al., 2014). First, hourly B_{CC} data is aggregated at a daily resolution and daily means and standard deviations are computed. A quantile (by default the 5th percentile) is then calculated on standard deviations. This threshold is then added and subtracted from the daily mean B_{CC} to generate a seasonal envelope. Raw data falling outside this envelope are discarded. The *blue* filter was designed to remove images with dominant clouds and/or snow on the canopy, because the blue chromatic coordinate was found to be very sensitive to such conditions.

(3) A filter called *mad*, following the method of Papale et al. (2006), an outlier detection based on the double-differenced time series, using the median of absolute deviation about the median (MAD) that is a robust outlier estimator.

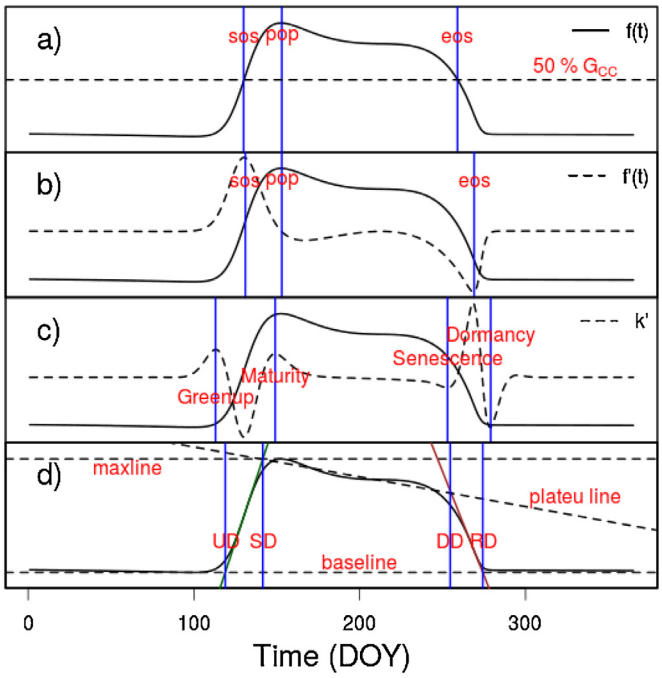

Fig. 3. Illustration of the methods used to extract phenological thresholds (phenophases) in a seasonal G_{CC} trajectory (f(t)). (a) *trs method*: Phenophases are defined as the day of year (DOY) when the 50% G_{CC} is reached either during greenup (sos) and autumn (eos), the DOY of maximum G_{CC} is called pop; (b) *derivatives method*: Phenophases are defined as the DOY when f′(t) shows the absolute maximum and minimum; (c) *Klosterman method*: Phenophases are defined as the two local maxima (greenup) and two local minima (autumn) in the rate of change in curvature k′ (Kline, 1998; Klosterman et al., 2014) and are named after Zhang et al. (2003); (d) *Gu method*: Maximum and minimum of f′(t) are used to define slopes of recovery and senescence lines tangent to the curves (green and brown lines). The intersection between these lines and baseline and maxline define the four phenophases in the original formulation (Gu et al., 2009). To account for the mid season decrease in G_{CC} we have further defined a plateau line as a linear fit to G_{CC} values between SD and DD, in order to adjust the definition of phenophase DD. For phenophase abbreviations and details see Table 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

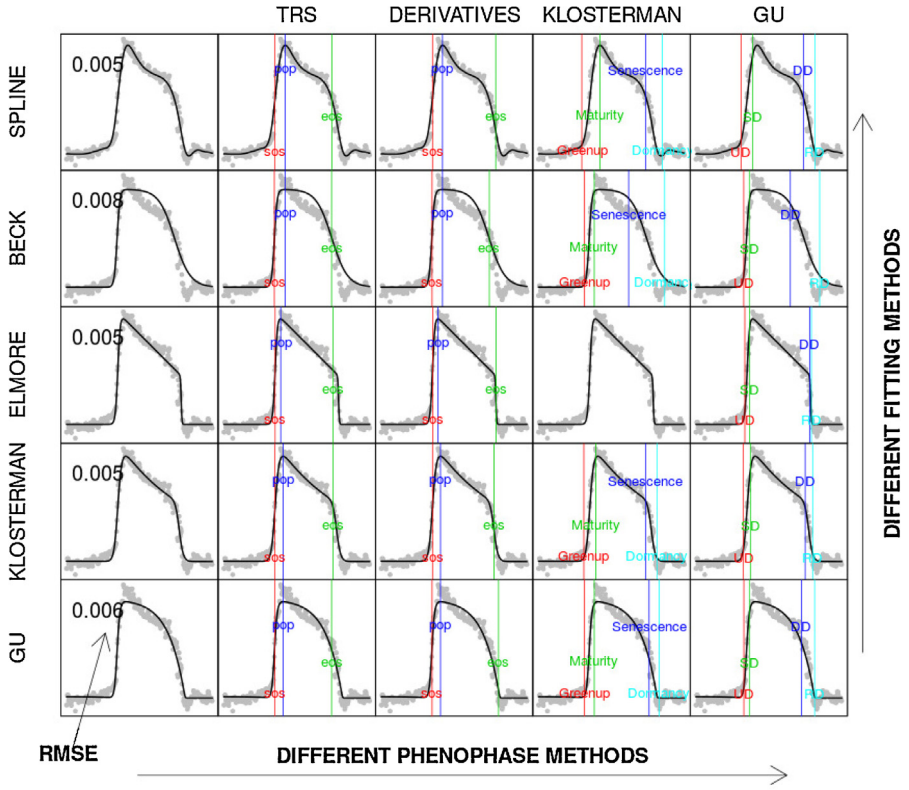

Fig. 4. All fitting and phenophase methods applied to Harvard deciduous forest G_{CC} seasonal trajectory in year 2013. This plot is the output from the functions `greenExplore()` and `plotExplore()`. Plots in the same row share the same fitting method, whereas those in the same column share the same phenophase extraction method. The RMSE for each fitting is also annotated in the first column.

(4) A filter called *spline*, following the method of Migliavacca et al. (2011), based on recursive spline smoothing and residual computation followed by removal of outliers falling outside a given residual envelope.

(5) A filter called *max*, following the method of Sonnentag et al. (2012), based on the identification of the 90th percentile values in three-days moving windows.

Filters can be applied alone or in sequence so that the user can choose the best combination of filters that suits the image archive to be processed. Additionally, each filter has its own discarding criteria that may be tuned by the user according to the quality of the input G_{CC} time series. The default behaviour of the filtering function is to use a sequence of *night*, *spline* and *max* filters. In general, this sequence will be effective enough to properly filter the G_{CC} time series. The user is advised to apply the *max* filter to effectively minimise the impact of changes in scene illumination (Sonnentag et al., 2012).

3.4. Fit a curve to the G_{CC} seasonal course and extract phenophases

The extraction of phenophases is done in two steps (following the approach of Forkel et al., 2015; Klosterman et al., 2014). First, a curve is fitted to the G_{CC} seasonal curve to reduce the influence of single observation and to better capture the seasonal behaviour. In a second step different extraction methods can be used to retrieve phenophase metrics from the fitted seasonal curve. We selected four different double logistic equations to be included in the package (Table 2). The equations differ in the number of parameters to be optimised and hence in the flexibility of the fitting curves. For a thorough examination of the fittings and explanation of curve parameters, see the correspondent publications. In addition to the double logistic equations, an approach based on a smoothed cubic spline is also available. There are cases, e.g. with very noisy time series or weak signal (low seasonal amplitude in greenness) where the double logistic fits fail. In such cases, the spline method is the only possibility to extract a seasonal trajectory.

After curve fitting, a suite of different approaches is available to extract phenophases, as summarised in Table 3, and illustrated in Fig. 3. The phenophase extraction methods lead to the definition of dates that may have markedly different ecological meanings and be useful for several applications in the field of environmental

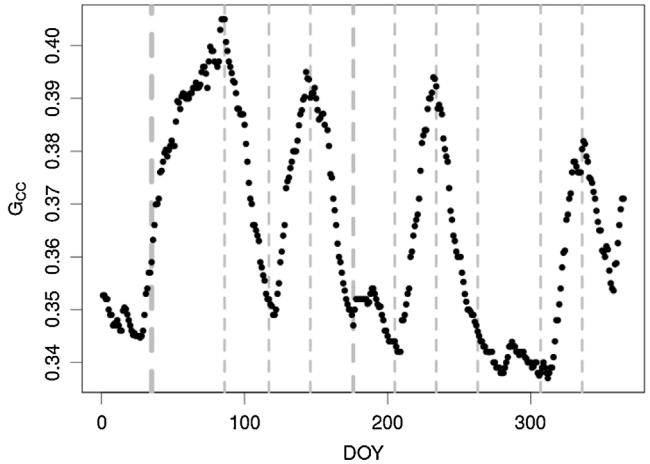

Fig. 5. Phenophases as determined by the break point analysis (`PhenoBP()` function) applied to Kamuela grassland G_{CC} seasonal trajectory in year 2013. The thickness of the vertical dashed lines is proportional to the uncertainty of the extracted break points (same scale as the *x*-axis). Uncertainty is estimated using the function `confint()` as the 95% confidence interval.

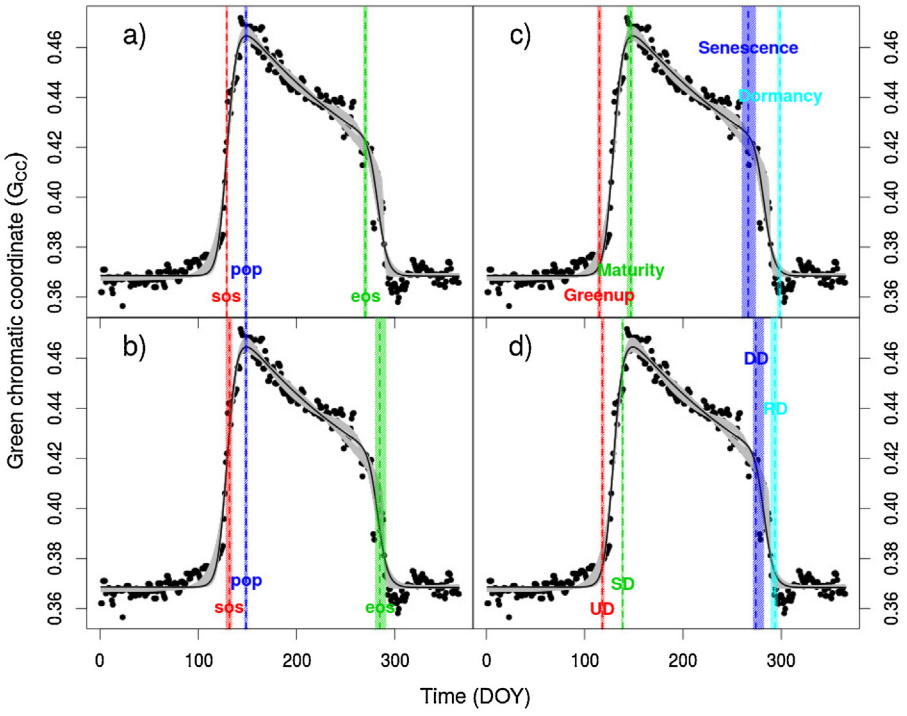

Fig. 6. The uncertainty analysis applied to Harvard deciduous forest G_{CC} seasonal trajectory in year 2013. The grey lines represent the 500 simulated curves on which phenophases are extracted. Shaded areas on phenophases represent 10th and 90th percentiles of the 500 replications. Data are fitted using the Klosterman method. Phenophases are extracted with (a) the trs method; (b) the derivatives method, (c) Klosterman method, (d) Gu method.

sciences. However the discussion of such meanings is beyond the objectives of this presentation paper, and the reader is referred to the publications listed in Table 2 for further information. Here it is important to point out that the availability of very simple methods (e.g. *trs*) and more sophisticated ones (e.g. *Klosterman* method) allow the user to choose the extraction method better suited to the phenological trajectory of the investigated ecosystem and/or based on the quality of the input data.

The combination of four fitting methods plus the cubic spline smoothing with the four phenophase extraction methods produces 20 different available combinations in output from a single yearly G_{CC} time series. A specific function (`greenExplore()`) is designed to give an overview of all fittings and phases for a given dataset (Fig. 4). The RMSE for each fit is also shown, so that the user can use it as a criterion for fit selection. Once the fit is chosen, by examining the plot along a row the user can evaluate which phenophase method is more suitable to the G_{CC} time series. The possibility to combine different fitting and phenophase methods to a G_{CC} time series provides a framework in which to compare in detail the different methods currently in use. This need was highlighted by Keenan et al. (2014) in a recent attempt to link forest phenology as described by time-lapse photography and physiology as described by traditional plant trait measurements.

In addition to the above described methods, `phenopix` also implements a phenophase extraction method (function `PhenoBP()`) based on linear piecewise regression and correspondent break points in the time series (Wingate et al., 2015). This method was designed to accommodate multiple greening peaks during the same season, which is typical, for example, of water limited ecosystems such as Kamuela (Fig 5) or managed grasslands and croplands.

4. Estimation of the uncertainty

Traditionally, the analysis of digital images for phenology has rarely included the estimation of uncertainty on phenophase extraction, despite its paramount importance (e.g. for the evaluation of method robustness and as input data for the optimisation of process-based or land surface models, where phenology is currently poorly represented (Richardson et al., 2013)).

The `phenopix` package provides two approaches for the estimation of uncertainty, one for the fitted double logistic equations and one for the smoothing spline method. For the fitted equations the residuals between the model fit and the observed data are used to generate random noise and fitting is applied recursively to randomly-noised original data. This procedure results in an ensemble of curves and curve parameters on which phenophase extraction is performed, thus providing an uncertainty estimate of phenophases. For the spline smoothing method, the uncertainty is estimated by combining a random noise as for the fitted equations with multiple changes in the spline's degrees of freedom, to account for the arbitrariness in their choice. In particular, spline degrees of freedom are set at 5% of the time series length and then let vary recursively between 1% and 5% of the time series length to generate the ensemble of smoothing curves on which the uncertainty is estimated. In addition to the residuals method, for the fitted equations only, we are also developing a method based on the Hessian matrix of the curve parameters.

Table 4
Estimated computation time (hours) required to complete different steps of the pixel-based analysis processing 10,000 pixels from a seasonal time series of 5000 images using one single processor on three different computers.

Step	3.07 GHz	2.60 GHz	2.13 GHz
Filtering	28	49	52
Spline	0.02	0.02	0.04
Beck	36	63	69
Elmore	0.7	0.8	1.5
Klosterman	10	16	19
Gu	22	37	42

Fig. 6 shows the uncertainty analysis for Harvard deciduous 2013 G_{CC} seasonal trajectory, with 500 replicates. The fitting method applied in this case was Klosterman. By default, the confidence interval is represented by the 10th and 90th percentile of the distribution of all extracted phenophases, but other options are available. In any case, the user can access the uncertainty data tables of (1) curve equation parameters and (2) extracted phenophases to customise the computation of the uncertainty envelope.

5. Phenopix application: spatial analysis

Based on an analysis run on all sites included in this paper except for Kamuela (i.e. six year-sites), we present in this paragraph the application of the pixel based analysis to investigate the spatial variability of phenology within a ROI (pixel-based approach).

To date, few studies have explored the possibility of analyzing a set of images pixel by pixel rather than averaging colour values on the overall region of interest (Julitta et al., 2014; Ide and Oguma, 2013). Therefore, an emergent question is how the ROI-averaged analysis is representative of all ROI pixels in terms of extracted phenophases? And, is this relationship consistent across fitting methods and ecosystem types? The `phenopix` package allows the user to perform pixel based image analysis and provides a number of functions to display and analyse the results. It must be noted that pixel-by-pixel analysis is computationally intense. We report in Table 4 the computation times required to analyse a ROI containing 10,000 pixels on a seasonal time series of 5000 images with all fits available in phenopix for 3 different computers, using one single processor. Computation time may be as high as 70 h, with very low values for spline smoothing and higher for data filtering and fitted equations. However, parallel computation available in the spatial functions of `phenopix` enable considerably reduced computation times compared to those reported in Table 4.

We evaluate the relationship between approaches (i.e. ROI-average vs pixel-based) and fitting/phenophase methods across sites in Fig. 7. Two different fitting methods (namely Klosterman and spline) applied to each pixel of the ROIs and then averaged are in good agreement with each other (Fig. 7, first above diagonal panel), with a slightly higher correlation for spring than for autumn phases. The error bars for this relationship suggest a higher variability for autumn than for spring phenophases (first below diagonal panel). When comparing the ROI-averaged and the pixel-based approach, the best relationship is found with the Klosterman fitting method applied to both approaches, with correlation coefficients of 0.97 and 0.75 for spring and autumn phenology, respectively. However, the relationship between spline fit applied pixel-by-pixel and Klosterman fit applied to the average ROI also shows very good agreement. In contrast, the spline method applied to the average ROI (last column in Fig. 7) has a consistently weaker relationship with all other methods, resulting in earlier autumn phases and a much higher variability.

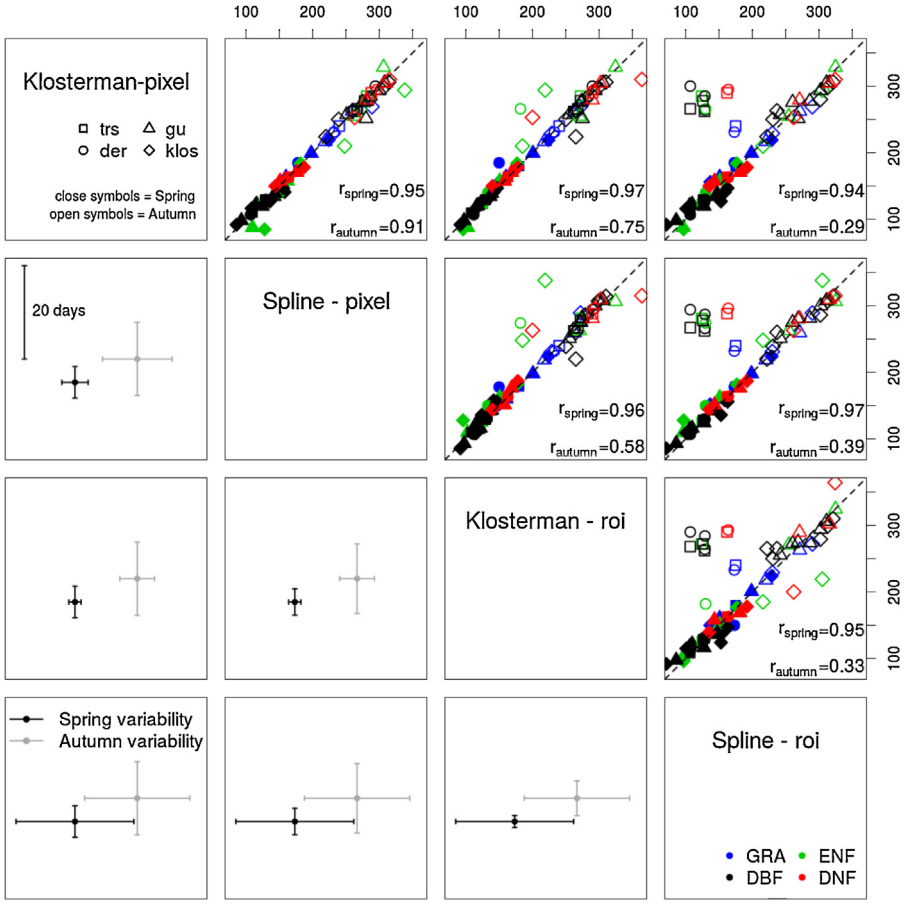

Fig. 7. Scatterplot matrix showing: (a) above the diagonal, the relationship between phenophases extracted with two different approaches (pixel-based and ROI-averaged) and two different fitting methods (spline smoothing and Klosterman fit), and (b) below the diagonal, errors associated. Error bars for pixel approach represent the mean absolute deviation of all pixels, whereas error bars for ROI-averaged approach represent the uncertainty as described in Section 4. Note that the scale for error bars is provided in the first below-diagonal panel. Different symbols denote different phenophase extraction methods. Close symbols are for spring phenology and open symbols for autumn phenology. Different colours represent different PFTs (abbreviations for PFTs are as in Table 1). Data from all sites except for Kamuela were used in this analysis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Across phenophase methods (different symbols in Fig. 7), there is no evidence of a systematic bias between the two approaches for any particular phase extraction method. Across ecosystem types, needle-leaf forest (ENF, i.e. harvardhemlock site) shows a consistent departure from the 1:1 line even in the best relationships (third panel from left in the first row of Fig. 7), probably due to the lower signal to noise ratio in the seasonal trajectory of evergreen trees. However, the most striking difference is the consistently lower correlation for autumn than for spring phases, suggesting that in the same ecosystem and possibly also within a single species autumn phases are more variable than spring phases.

In summary, the faster spline smoothing applied pixel-by-pixel leads to the identification of phenophases in substantial agreement with the more computationally-intense Klosterman fit. Additionally, being more flexible, the spline fitting leads to a much lower number of unfitted pixels compared to the Klosterman fit. Hence, from this analysis, spline smoothing is preferred over curve fitting for pixel-by-pixel analysis. For the ROI-averaged approach, it is strongly preferable to perform curve fitting over spline smoothing, provided that the G_{CC} seasonal trajectory does not show multiple peaks, related to e.g. water stress or management practices.

One interesting application of the pixel-based analysis stems from the examination of the frequency distribution of a given phenophase across all pixels. When a phenophase shows a bi- (or multi-) modality and a consistent spatial distribution, pixels may be grouped according to the values of one (or more) phenophase(s) and separate seasonal trajectories can be obtained (Fig. 8). By applying this procedure to the Torgnon-ND grassland site we were able to identify a bimodal distribution in maturity phase (i.e. the end of spring growth), as extracted after Klosterman curve fitting. We then selected pixels with maturity onset dates falling around the two modes (i.e. DOY 180±3 and DOY 205±3) and computed an average trajectory for these two groups. This was possible because in the pixel-based approach each pixel has associated curve parameters to reconstruct the seasonal trajectory. The resulting average trajectories are markedly distinct around the seasonal peak, and reflect the different ecology of forbs and grasses occurring over very short distances (<20 cm) at this grassland site (Julitta et al., 2014). Further interpretation of this behaviour is beyond the objective of this

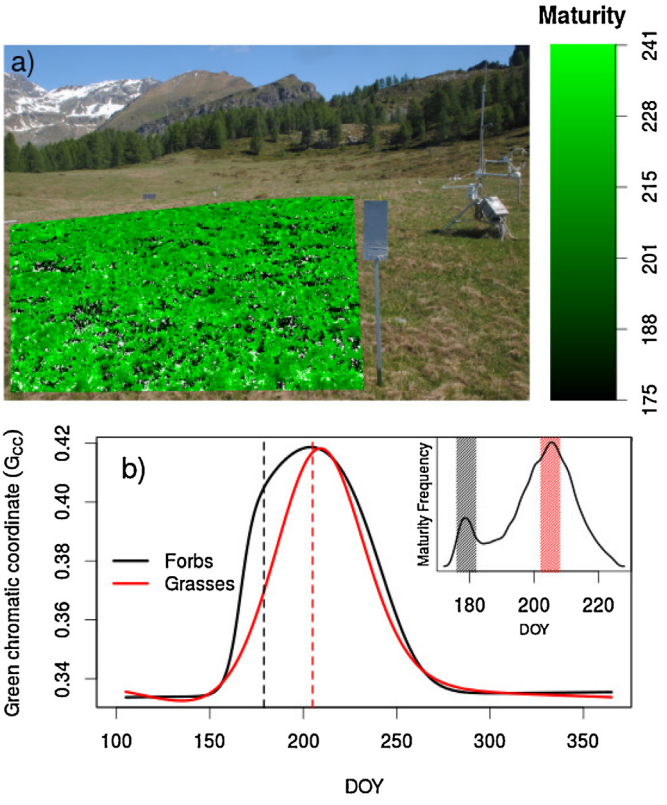

Fig. 8. Results of the spatial analysis conducted on Torgnon-ND grassland site. (a) Map of the spatial distribution of maturity (Klosterman fitting method, Klosterman phenophases method), with darker colours indicating earlier occurrence of the phase (in DOYs). (b) Average seasonal trajectory of pixels clustered according to the bimodal distribution of maturity. Patches with forbs dominating show an earlier maturity compared to grasses (dashed lines). In the inset, density distribution of Maturity phase across all pixels. Shaded areas denote the maturity intervals used to subset ROI pixels. The resulting subsets were in turn used to extract the seasonal trajectories of forbs- and grass-dominated portions of the ROI. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

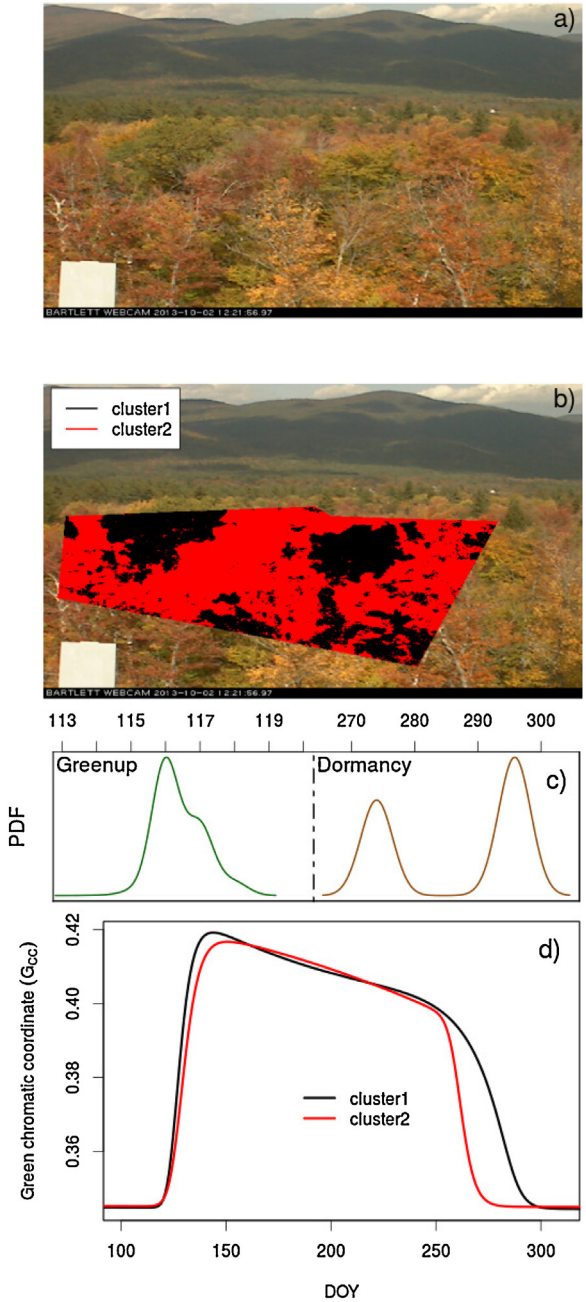

Fig. 9. Results of the cluster analysis run on Bartlett pixel-based phenophases. (a) Sample image from DOY 275, (b) distribution of clusters in the ROI, (c) density distribution of Greenup and Dormancy phases across all pixels, and (d) averaged seasonal trajectories for the two clusters. All extracted phases (from the four methods presented in Section 3) were used in input to clustering.

paper, but this preliminary analysis highlights the potential use of spatially explicit image processing to identify different phenology related to different plant functional types.

More sophisticated clustering methods can be applied to pixel values in order to distinguish different phenological patterns. For example, at the Bartlett site (a mixed deciduous forest) we might expect a different phenology in early or late flushing species (Fig. 9a). With a cluster analysis (Hartigan and Wong, 1979) using phenophases extracted with all available methods as an input matrix, we were able to clearly distinguish two portions of the canopy (Fig. 9b). In this example the cluster analysis was used to define two sub-ROIs, that reentered the processing chain (ROI-averaged approach, Fig. 2) and led to the identification of two markedly different seasonal trajectories (Fig. 9c and d). Beech-dominated portions of the canopy (cluster 2) show a later yellowing/browning compared to maple, prevailing in cluster 1 (Richardson et al., 2009).

6. Summary and outlook

We presented a new R package called `phenopix` that collects the most up-to-date processing techniques for repeated digital photography of vegetation canopy in the context of phenological and ecological research. Together with the basic and well established processing steps, the package implements several new features, including the possibility to combine different fitting and phenophase methods, to compute uncertainty on extracted phenophases, and to conduct pixel-by-pixel analysis. Phenophase maps obtained by the pixel-based approach may represent a promising tool for a variety of ecological analyses, including for example: (1) the study of species-specific phenology; (2) the examination of spatial variability of apparently homogeneous ecosystems such as grasslands; (3) the potential to discriminate between overstory and understory or age-related differences in phenological patterns of open canopy forests.

One major future step in the evolution of `phenopix` package includes the computation of the so called camera-NDVI, in the formulation proposed by Petach et al. (2014).

The `phenopix` R package offers a suite of standardised and reproducible processing code that may be adopted, deployed and further developed by the existing phenocameras networks worldwide.

Acknowledgments

ADR acknowledges support from the National Science Foundation, through the Macrosystems Biology program (grant EF-1065029).

The package was developed in the framework of the ALCOTRA Interreg Project n 227 e-PHENO Reti fenologiche nelle Alpi.

We thank Morgan Furze, Harvard University, for assistance with editing the manuscript and two anonymous reviewers for valuable comments on a previous version of the manuscript.

References

Ahrends, H.E., Brügger, R., Stöckli, R., Schenk, J., Michna, P., Jeanneret, F., Wanner, H., Eugster, W., 2008. Quantitative phenological observations of a mixed beech forest in northern Switzerland with digital photography. J. Geophys. Res. 113, G04004, http://dx.doi.org/10.1029/2007JG000650.

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, 321–334, http://dx.doi.org/10.1016/j.rse.2005.10.021, http://www.sciencedirect.com/science/article/pii/S0034425705003640.

Busetto, L., Colombo, R., Migliavacca, M., Cremonese, E., Meroni, M., Galvagno, M., Rossini, M., Siniscalco, C., Morra Di Cella, U., Pari, E., 2010. Remote sensing of larch phenological cycle and analysis of relationships with climate in the Alpine region. Glob. Change Biol., 10.1111/J.1365-2486.2010.02189.X.

Delbart, N., Kergoat, L., Toan, T.L., Lhermitte, J., Picard, G., 2005. Determination of phenological dates in boreal regions using normalized difference water index. Remote Sens. Environ. 97, 26–38, http://dx.doi.org/10.1016/j.rse.2005.03.011, http://www.sciencedirect.com/science/article/pii/S0034425705001288.

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. Change Biol. 18, 656–674, http://dx.doi.org/10.1111/j.1365-2486.2011.02521.x.

Filippa, G., Cremonese, E., Galvagno, M., Migliavacca, M., Morra di Cella, U., Petey, M., Siniscalco, C., 2015. Five years of phenological monitoring in a mountain grassland: inter-annual patterns and evaluation of the sampling protocol. Int. J. Biometeorol., http://dx.doi.org/10.1007/s00484-015-0999-5.

Forkel, M., Migliavacca, M., Thonicke, K., Reichstein, M., Schaphoff, S., Weber, U., Carvalhais, N., 2015. Codominant water control on global interannual variability and trends in land surface phenology and greenness. Glob. Change Biol. 21, 3414–3435, http://dx.doi.org/10.1111/gcb.12950.

Galvagno, M., Wohlfahrt, G., Cremonese, E., Rossini, M., Colombo, R., Filippa, G., Julitta, T., Manca, G., Siniscalco, C., Morra di Cella, U., Migliavacca, M., 2013. Phenology and carbon dioxide source/sink strength of a subalpine grassland in response to an exceptionally short snow season. Environ. Res. Lett. 8, 025008.

Gillespie, A.R., Kahle, A.B., Walker, R.E., 1987. Color enhancement of highly correlated images. II. Channel ratio and chromaticity transformation techniques. Remote Sens. Environ. 22, 343–365.

Gu, L., Post, W., Baldocchi, D., Black, T., Suyker, A., Verma, S., Vesala, T., Wofsy, S., 2009. Characterizing the seasonal dynamics of plant community photosynthesis across a range of vegetation types. In: Noormets, A. (Ed.), Phenology of Ecosystem Processes. Springer, New York, pp. 35–58, 10.1007/978-1-4419-0026-5_2.

Hartigan, J.A., Wong, M.A., 1979. Algorithm as 136: a k-means clustering algorithm. J. R. Stat. Soc. Ser. C (Appl. Stat.) 28, 100–108, http://www.jstor.org/stable/2346830.

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://dx.doi.org/10.1016/j.rse.2011.10.006, http://linkinghub.elsevier.com/retrieve/pii/S0034425711003543.

Ide, R., Oguma, H., 2013. A cost-effective monitoring method using digital time-lapse cameras for detecting temporal and spatial variations of snowmelt and vegetation phenology in alpine ecosystems. Ecol. Inform. 16, 25–34, http://dx.doi.org/10.1016/j.ecoinf.2013.04.003, http://www.sciencedirect.com/science/article/pii/S1574954113000290.

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. Forest Meteorol. 143, 64–79, http://dx.doi.org/10.1016/j.agrformet.2006.11.008, http://www.sciencedirect.com/science/article/pii/S0168192306003613.

Julitta, T., Cremonese, E., Migliavacca, M., Colombo, R., Galvagno, M., Siniscalco, C., Rossini, M., Fava, F., Cogliati, S., di Cella, U.M., Menzel, A., 2014. Using digital camera images to analyse snowmelt and phenology of a subalpine grassland. Agric. Forest Meteorol. 198–199, 116–125, http://dx.doi.org/10.1016/j.agrformet.2014.08.007, http://www.sciencedirect.com/science/article/pii/S0168192314001981.

Keenan, T.F., Darby, B., Felts, E., Sonnentag, O., Friedl, M., Hufkens, K., O' Keefe, J.F., Klosterman, S., Munger, J.W., Toomey, M., Richardson, A.D., 2014. Tracking forest phenology and seasonal physiology using digital repeat photography: a critical assessment. Ecol. Appl., http://dx.doi.org/10.1890/13-0652.1, 140206175103002. http://www.esajournals.org/doi/abs/10.1890/13-0652.1.

Kline, M., 1998. Calculus: An Intuitive and Physical Approach (Second Edition) (Google eBook). Dover Publications, http://books.google.com/books?id=tZy8AQAAQBAJ&pgis=1.

Klosterman, S.T., Hufkens, K., Gray, J.M., Melaas, E., Sonnentag, O., Lavine, I., Mitchell, L., Norman, R., Friedl, M.A., Richardson, A.D., 2014. Evaluating remote sensing of deciduous forest phenology at multiple spatial scales using PhenoCam imagery. Biogeosciences 11, 4305–4320.

Lechowicz, M.J., 1984. Why do temperate deciduous trees leaf out at different times? Adaptation and ecology of forest communities. Am. Nat. 124, 821–842, http://www.jstor.org/stable/2461303.

Migliavacca, M., Cremonese, E., Colombo, R., Busetto, L., Galvagno, M., Ganis, L., Meroni, M., Pari, E., Rossini, M., Siniscalco, C., Morra di Cella, U., 2008. European larch phenology in the Alps: can we grasp the role of ecological factors by combining field observations and inverse modelling? Int. J. Biometeorol. 52, 587–605, http://dx.doi.org/10.1007/s00484-008-0152-9, http://www.ncbi.nlm.nih.gov/pubmed/18437430.

Migliavacca, M., Galvagno, M., Cremonese, E., Rossini, M., Meroni, M., Sonnentag, O., Cogliati, S., Manca, G., Diotri, F., Busetto, L., Cescatti, A., Colombo, R., Fava, F., Morra di Cella, U., Pari, E., Siniscalco, C., Richardson, A.D., 2011. Using digital repeat photography and eddy covariance data to model grassland phenology and photosynthetic CO_2 uptake. Agric. Forest Meteorol. 151, 1325–1337.

Mizunuma, T., Wilkinson, M., Eaton, L., Mencuccini, E.M., Morison, I.L., Grace, J.J., 2013. The relationship between carbon dioxide uptake and canopy colour from two camera systems in a deciduous forest in southern England. Funct. Ecol. 27, 196–207, http://dx.doi.org/10.1111/1365-2435.12026, http://doi.wiley.com/10.1111/1365-2435.12026.

Nasahara, K., Nagai, S., 2015. Review: Development of an in-situ observation network for terrestrial ecological remote sensing – the phenological eyes

network (pen). Ecol. Res. 30, 211–223, http://dx.doi.org/10.1007/s11284-014-1239-x.
Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., Yakir, D., 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences 3, 571–583.
Petach, A.R., Toomey, M., Aubrecht, D.M., Richardson, A.D., 2014. Monitoring vegetation phenology using an infrared-enabled security camera. Agric. Forest Meteorol. 195–196, 143–151, http://dx.doi.org/10.1016/j.agrformet.2014.05.008, http://linkinghub.elsevier.com/retrieve/pii/S0168192314001257.
R Core Team R., 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Austria, Vienna, http://www.R-project.org/.
Richardson, A., Jenkins, J., Braswell, B., Hollinger, D., Ollinger, S., Smith, M.L., 2007. Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. Oecologia 152, 323–334, http://dx.doi.org/10.1007/s00442-006-0657-z.
Richardson, A.D., Bailey, A.S., Denny, E.G., Martin, C.W., O'Keefe, J., 2006. Phenology of a northern hardwood forest canopy. Glob. Change Biol. 12, 1174–1188.
Richardson, A.D., Braswell, B.H., Hollinger, D.Y., Jenkins, J.P., Ollinger, S.V., 2009. Near-surface remote sensing of spatial and temporal variation in canopy phenology. Ecol. Appl. 19, 1417–1428.
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. Forest Meteorol. 169, 156–173.
Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A.M., Friedl, M., Braswell, B.H., Milliman, T., O'Keefe, J., Richardson, A.D., 2012. Digital repeat photography for phenological research in forest ecosystems. Agric. Forest Meteorol. 152, 159–177.
Soudani, K., Hmimina, G., Delpierre, N., Pontailler, J.Y., Aubinet, M., Bonal, D., Caquet, B., de Grandcourt, A., Burban, B., Flechard, C., Guyon, D., Granier, A., Gross, P., Heinesh, B., Longdoz, B., Loustau, D., Moureaux, C., Ourcival, J.M., Rambal, S., Saint André, L., Dufrêne, E., 2012. Ground-based Network of NDVI measurements for tracking temporal dynamics of canopy structure and vegetation phenology in different biomes. Remote Sens. Environ. 123, 234–245, http://dx.doi.org/10.1016/j.rse.2012.03.012.
White, M.A., Nemani, R.R., 2006. Real-time monitoring and short-term forecasting of land surface phenology. Remote Sens. Environ. 104, 43–49, http://dx.doi.org/10.1016/j.rse.2006.04.014, http://www.sciencedirect.com/science/article/pii/S0034425706001660.
Wingate, L., Ogée, J., Cremonese, E., Filippa, G., Mizunuma, T., Migliavacca, M., Moisy, C., Wilkinson, M., Moureaux, C., Wohlfahrt, G., Hammerle, A., Hörtnagl, L., Gimeno, C., Porcar-Castell, A., Galvagno, M., Nakaji, T., Morison, J., Kolle, O., Knohl, A., Kutsch, W., Kolari, P., Nikinmaa, E., Ibrom, A., Gielen, B., Eugster, W., Balzarolo, M., Papale, D., Klumpp, K., Köstner, B., Grünwald, T., Joffre, R., 2015. Interpreting canopy development and physiology using the EUROPhen camera network at flux sites. Biogeosci. Discuss. 12, 7979–8034, http://dx.doi.org/10.5194/bgd-12-7979-2015, http://www.biogeosciences-discuss.net/12/7979/2015/.
Woebbecke, D., Meyer, G., Von Bargen, K., Mortensen, D., 1995. Color indices for weed identification under various soil, residue, and lighting conditions. Trans. ASAE 38, 259–269, http://cat.inist.fr/?aModele=afficheN&cpsidt=3503524.
Zhang, X., Friedl, M., Schaaf, a., Strahler, C.B., Hodges, A.H., Gao, J.C.F., Reed, F., Huete, B.C.A., 2003. Monitoring vegetation phenology using MODIS. Remote Sens. Environ. 84, 471–475, doi: 10.1016/S0034-4257(02)00135-9.