Five years of phenological monitoring in a mountain grassland: inter-annual patterns and evaluation of the sampling protocol

(Article begins on next page)

17 July 2024

Abstract

The increasingly important effect of climate change and extremes on alpine phenology highlights the need to establish accurate monitoring methods to track inter-annual variation (IAV) and long-term trends in plant phenology. We evaluated four different indices of phenological development (two for plant productivity, i.e., green biomass and leaf area index; two for plant greenness, i.e., greenness from visual inspection and from digital images) from a 5-year monitoring of ecosystem phenology, here defined as the seasonal development of the grassland canopy, in a subalpine grassland site (NW Alps). Our aim was to establish an effective observation strategy that enables the detection of shifts in grassland phenology in response to climate trends and meteorological extremes. The seasonal development of the vegetation at this site appears strongly controlled by snowmelt mostly in its first stages and to a lesser extent in the overall development trajectory. All indices were able to detect an anomalous beginning of the growing season in 2011 due to an exceptionally early snowmelt, whereas only some of them revealed a later beginning of the growing season in 2013 due to a late snowmelt. A method is developed to derive the number of samples that maximise the trade-off between sampling effort and accuracy in IAV detection in the context of long-term phenology monitoring programmes. Results show that spring phenology requires a smaller number of samples than autumn phenology to track a given target of IAV. Additionally, productivity indices (leaf area index and green biomass) have a higher sampling requirement than greenness derived from visual estimation and from the analysis of digital images. Of the latter two, the analysis of digital images stands out as the more effective, rapid and objective method to detect IAV in vegetation development.

Introduction

Global change is altering plant phenology in the northern hemisphere (Menzel et al. 2006; Linderholm 2006; Parmesan 2007) and also in the alpine and subalpine belt in temperate ecosystems (Inouye 2008; Inouye and Wielgolaski 2013). In particular, in mountain areas, the suggested increase in frequency of climate anomalies, such as early snowmelts, summer droughts or warm winter spells (Gobiet et al. 2014; Beniston 2005), highlights the need to establish accurate monitoring methods to detect inter-annual variation (IAV) in plant phenology (Rutishauser et al. 2008; Orsenigo et al. 2014). In turn, a better knowledge of factors regulating the timing and dynamics of phenological responses would provide reliable information that can be incorporated into existing land surface models. Accordingly, Richardson et al. (2013) highlighted the need for an accurate representation of phenological processes in models that couple the land surface to the climate system. Among alpine and subalpine ecosystems, grasslands are regarded as hotspots where the impact of climate change will be very likely stronger than on other mountain vegetation because of their high diversity (Körner 2005) and due to pasture abandonment and the wide range of micro-climatic conditions occurring over short distances (Zeeman MJ et al. 2010; Wohlfahrt et al. 2003).

The sensitivity of alpine ecosystems to the IAV in meteorological drivers has been well acknowledged for a long time (Walker et al. 1995). Recent studies not only indicate a strong climatic control on carbon dioxide (CO_2) exchange of alpine meadows but also reveal a strong acclimation capacity due to ecosystem response to IAV (Marcolla et al. 2011; Chen et al. 2014). An estimate of future trends in alpine plants growth suggests that shifts in timing of snowmelt will affect the onset of plant growth and biomass production in the Swiss Alps (Rammig et al. 2010).

Additionally, experimental studies indicate that subalpine meadows are sensitive to climate change manipulations, reporting either community or species-specific responses (Cornelius et al. 2013; Dunne et al. 2003).

Concerning phenology, much work has been conducted in temperate forest ecosystems (Richardson et al. 2010), while fewer studies have been devoted to grasslands and shrublands, particularly in the alpine and subalpine belt (Richardson et al. 2013; Inouye and Wielgolaski 2013). Field measurements of grassland phenology are traditionally based on the observation of individual plants or plots to detect vegetative and reproductive phases of single species (Wipf et al. 2009; Inouye 2008; Cleland et al. 2006). Additionally, observations are usually conducted focusing on the first part of the growing season, with much less attention devoted to the senescence (Richardson et al. 2013), although there is increasing evidence that green-up phenology may show different signals compared to senescence phenology (CaraDonna et al. 2014).

Recently, methods have been developed to investigate larger portions of the vegetation community (Aldridge et al. 2011; Diez et al. 2012; CaraDonna et al. 2014), therefore shifting from species phenology to the community or ecosystem phenology and from single-point analysis of a given phenological event to the description of seasonal trajectories (Eklundh et al. 2011; Gu 2009; Richardson et al. 2006). The objective of this upscaling from individuals to ecosystem is to obtain information on processes that affect and are affected by phenology at the ecosystem level, such as carbon and water fluxes, albedo, etc. Accordingly, remote sensing has been widely used to explore land surface phenology at the landscape level (Jones et al. 2013; Ma and Zhou 2012; Studer et al. 2007; De Beurs and Henebry 2010; Henebry and De Beurs 2013). Field measurements that attempt to track spatially integrated phenological signals include the estimation of greenness of a given vegetation type. Greenness has been traditionally estimated by an observer in the field, but recently digital image analysis has been proven to be a valuable tool for the computation of greenness index (e.g. Richardson et al. (2009)). Furthermore, productivity of the canopy can be measured by sampling the plant biomass that could be subsequently weighted (biomass estimation) or analysed for the determination of leaf area index (LAI). The spatial scale of these methods can be viewed as in between the observation of single-species phenology and remotely sensed land surface phenology (Studer et al. 2007). Compared to the first, these methods have the advantage of integrating the response of multiple species in a single index, whereas with respect to the latter, they have a much higher spatial resolution. Field work for data collection and/or analysis usually represents a big effort for researchers, so a key question would be which experimental protocol (number and distribution of observation units) is required to track IAV of a given phenological variable and, among different observational approaches, which are more suitable to investigate IAV in the context of a long-term monitoring programme. Reducing sampling effort is laudable especially in the monitoring of remote areas (Cornelius et al. 2011). Previous sample size analyses suggest that a sample size of 15 sampling units and a fortnightly frequency is the minimum requirement for tracking a seasonal phenological trajectory in deciduous forest ecosystems (Morellato et al. 2010). Liang and Schwartz (2009) report that a sample size of 20 plants accurately described the population phenology of a single tree species, *Populus tremuloides*. More recently, Liang et al. (2011) used a minimum number of 20 observations per species to describe population phenology of a mixed deciduous forest. Schwartz et al. (2013) reported that 30 individuals and a 4-day sampling frequency minimises data uncertainty and field work expenses in a mixed deciduous/evergreen forest. Yet, whether the suggested number of observations is sufficient to describe population phenology in a subalpine grassland is an unanswered question. Additionally, a coherent set of sampling rules and methods for phenology is strongly requested (Hudson 2010) and under recent development (Denny et al. 2014).

In this work, we used 5 years of phenological data obtained from weekly to biweekly sampling in a subalpine grassland in the northwestern Italian Alps to

(1) illustrate the seasonal trajectories of phenological indices and their IAV at the community level, and the relationship between indices and the meteorological conditions;

(2) evaluate which indices are more suitable to determine IAV in the seasonal trajectories of phenology;

(3) assess if and how the experimental design could be optimised by reducing sampling effort without losing statistical power when evaluating IAV in seasonal trajectories of phenological indices.

Materials and methods

Study site

The study was carried out in a subalpine grassland, in the northwestern Italian Alps. The site is an unmanaged grassland located in the Aosta Valley region at an elevation of 2160 m asl (45 ° 50′

40″ n, 7 ° 34′ 41″ e). Vegetation consists of the dominant *Nardus stricta* L., *Festuca nigrescens* all., *Arnica montana* L., *Carex sempervirens* vill., *Geum montanum* L., *Anthoxanthum alpinum* L., *Potentilla aurea* L. and *Trifolium alpinum* L. The terrain slopes gently (4 °), and in areas with a relatively steeper slope, where the soil is shallower, *Arnica montana*, *Carex sempervirens* and *Geum montanum* are dominant and a lower *Nardus stricta* cover occurs. The soil is classified as Cambisol (FAO/ISRIC/ISS). The site is characterised by an intra-alpine semi-continental climate, with mean annual temperature of 3.1 °C and mean annual precipitation of about 880 mm. On average, from the end of October to late May, the site is covered by a thick snow cover (90–120 cm) that limits the growing period to an average of 5 months. Long-term snow depth and air temperature averages were computed on the basis of a 30-year record (1978–2008) from the weather station of Cignana (45 ° 52′ 31″ n, 7 ° 35′ 19″

e), located nearby the Torgnon site, and at the same altitude.

Experimental design

Four phenological indices were evaluated in this study, two related to plant productivity and two to greenness: the first two are green biomass (GB) and LAI. GB was determined after sampling a 30×30 cm quadrat by clipping vegetation to 2 mm above the ground. The material was then separated to green and dry mass, dried to constant weight at 60 °C for 48 h and weighed. LAI was determined on the same material: sample leaves were run through an area meter (model LI-3100, LI-COR, Inc., Lincoln NE) and the hemi-surface area of the green material was determined. LAI was obtained by dividing the total hemi-surface area of the harvested material with the cut ground surface area (Bréda 2003).Greenness indices consist of the following:

(i) visual estimation of vegetation greenness (VG) obtained by assigning a percentage of green cover by visual observation; visual observation of greenness was performed by different observers over years, trained by one main responsible for this task; additionally, the main responsible supervised the sampling during about 80 % of the sampling dates over the 5-year study;

(ii) the greenness index (Gillespie et al. 1987; Richardson et al. 2009) was obtained from the analysis of nadiral digital images (NG). Nadiral images were acquired with a reflex camera CANON EOS 50D from an height of about 1.5 m. The collected images were then analysed in the R environment (R Core Team 2014) to compute the green chromatic coordinate that represents the proportion of green of each image (Klosterman et al. 2014).

For each index, observations were conducted on four vertexes of three rectangular plots, approximately 200 m apart from each other. Each plot was 40×15 m, and the minimum distance between vertexes was therefore about 15 m. Because of the destructive sampling procedure, GB and LAI were sampled in the surroundings of the identified vertexes, whereas VG and NG were always evaluated in the same 30×30 cm sampling square.

Samples were collected every 1–2 weeks from the snowmelt (May–Jun) to the end of the growing season (Oct–Nov) for 5 years (2009–2013). Each year sampling began no later than 10 days after snowmelt, except in 2012, when the first sample was collected 1 month after snowmelt.

Data analysis

Seasonal trajectories

Data from each vertex were used to depict seasonal trajectories of a given index. A cubic spline was then fitted to the normalised seasonal trajectories of each index and phenological phases were computed as quantiles on the curve. Accordingly, from the seasonal course of each index, 10 phases were extracted, 5 of them corresponding to the day of the year (doy) when 10th, 25th, 50th, 75th and 90th percentile of the maximum seasonal value was reached (greening phases, G10–G90) and 5 of them corresponding to the decreasing phases (yellowing phases, Y10–Y90).

Inter-annual variation

Linear mixed models (MM) were used to test for significant differences among years for a given phase and index. Year was set as fixed effect and plots and vertices as random effects. Tukey HSD was used to test for differences among years, at a significance level of $p < 0.05$.

Comparison with gross primary production

At the study site, continuous eddy covariance measurements of CO_2 exchange between the ecosystem and the atmosphere are available. Details on instrumental setup, measurements and data processing are provided in Galvagno et al. (2013). In detail, the gap-filling and partitioning methods described in Reichstein et al. (2005) and Lasslop et al. (2010) were applied separately and then averaged to obtain gross primary production (GPP) data. Seasonal trajectories of the investigated indices were compared to seasonal trajectories of GPP.

Evaluation of experimental design

We tested two methods to determine the ability of our indices in detecting IAV: (1) a traditional power analysis and (2) an analysis based on a combination of sample removal and MM.

(1) The power analysis allows the extraction of a sample size (n) required to detect a given effect size (d) at a fixed power ($1 - \beta$). The effect size was obtained from the variance components (i.e.

between and within variance), estimated by the MM. The sample size was calculated with a fixed β at 0.20, as suggested by Cohen ([1988]).

(2) For each index and each phase, we tested whether there were significant differences between paired years by means of MM. We then sequentially reduced the number of vertices and plots in all possible combinations. This resulted in one complete model (three plots and four vertices) and 10 reduced models, each with a different number of possible combinations of vertices and plots (Table [1]). For each combination, we fitted a MM, extracted the model p value and computed the mean difference (d_{mean}, expressed in days) in the doy of occurrence of a given phase between the two years in comparison. We then investigated the relationship between d_{mean}s and p values. Obviously, a higher d_{mean} (i.e. a larger difference in days of occurrence of a given threshold between 2 years) results in lower p values. The inverse relationship between d_{mean} and p values was parametrised using a cubic spline and the d_{mean} corresponding to a confidence level of $p = 0.05$ ($d_{p=0.05}$) was predicted from the fitting. Hence, a predicted $d_{p=0.05}$ was extracted for each index, phase and sample size. $d_{p=0.05}$ was used as an estimate of the minimum detectable difference at a confidence level of $p = 0.05$ for a given index, phase and sample size. Uncertainty on $d_{p=0.05}$ was estimated by a bootstrap method with 500 replications.

Table 1

Number of possible combinations for each reduced sampling size

N of vertices	3 plots	2 plots	1 plot
4	1	3	3
3	64	48	12
2	216	108	18
1	64	48	/

Results and discussion

Snowfall and air temperature

The 5 years of study showed quite distinct meteorological conditions (Fig. [1]). Winter 2009 was the most snowy, whereas winter 2012 showed the lowest cumulative snow amount. However, in 2012, a late spring snowfall prevented the complete snowmelt from occurring in April. The earliest snowmelt occurred in 2011 and the latest was in 2013.
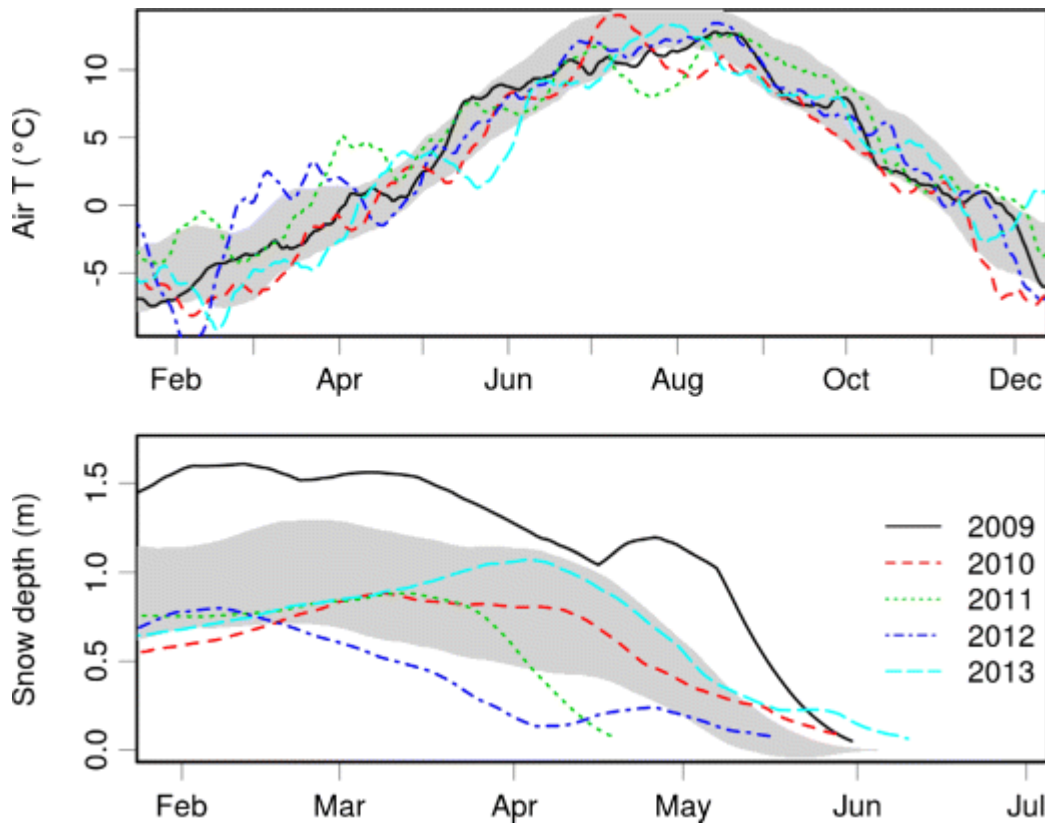
Fig. 1

Seasonal course of air temperature and snow depth for the 5-year study period (21-day window moving average). The *grey polygon* denotes the long-term (1978–2008) average (±the interquartile distance)

Air temperature across years was slightly less variable than snow amount, with some exceptions. For example, a departure from the 30-year variability envelope was recorded in 2012 before the snow melt, with air temperatures higher than the average in March. In 2013, late May was colder than the 30-year average. Summer temperatures were within the 30-year envelope for all investigated years.

Inter-annual variability of indices and extracted thresholds

The seasonal trajectories of all indices are reported in Fig. 2. All indices showed an earlier beginning of the growing season 2011 compared to other years, as a result of the earlier snowmelt in 2011 (Galvagno et al. 2013). All indices also showed a later beginning of the growing season in 2013 as a result of the later snowmelt, even if in terms of extracted thresholds the anomaly in 2013 is much less clear than in 2011 (see below). VG is the index with lower inter-annual variation, with the only significant departure from the 5-year average trajectory being represented by the earlier onset in 2011 and the later onset in 2013. The same is true for NG, except that in 2013, also the seasonal maximum occurred later than the 5-year average. GB is the index with highest inter-annual variation. Its seasonal courses suggest that the highest biomass production occurred in 2009 and the lowest in 2013. LAI is in agreement with GB with respect to the lowest plant production in 2013. The shifted trajectory in all indices for 2013 may be explained by the combination of later snowmelt and colder temperatures in May, which may have delayed plant development. The earlier onset of the growing season in 2011 did not result in higher plant production nor greenness. A previous study (Galvagno et al. 2013) clarified that the earlier onset of 2011 growing season resulted in reduced photosynthetic activity compared to average years 2009 and 2010. It also demonstrated that

this feature could be attributed to the biotic response of the ecosystem to an exceptional climate event (i.e. the early snowmelt) rather than to direct limiting weather conditions of the summer period. The seasonal development of the vegetation at this site appears therefore strongly controlled by snowmelt in its first stages and to a lesser extent in the overall development trajectory, whereas other factors (namely temperature and precipitation) concomitantly modulate the short-term changes during the growing season.
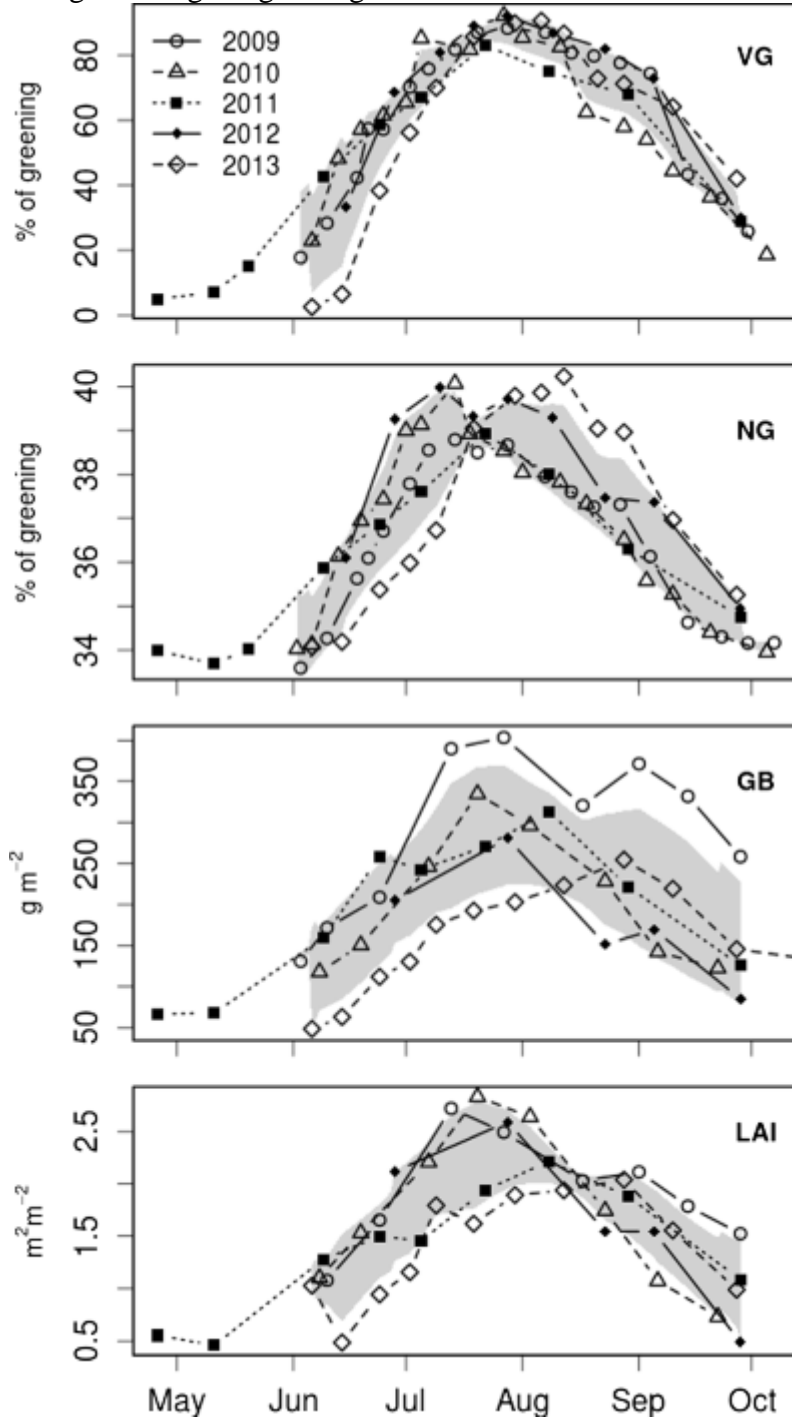

Fig. 2

Seasonal trajectories of all investigated indices in 3 years. Indices are averaged over all plots and vertexes for each sampling date. The *grey area* denotes the 5-year mean ± 2se. *GB* green biomass, *LAI* leaf area index, *NG* nadiral greening, *VG* visual greening

From the seasonal course of all indices, the cubic spline fitting and thresholding allowed us to extract greening and yellowing phases. Because the main objective of the threshold extraction is to establish whether there are changes from year to year in the beginning and end of the growing season, we will mainly refer to only one early and one late phase (G25 and Y75, respectively). These two were chosen instead of first (G10) and last (Y90) phase because the latter may be more sensitive to outliers and more influenced by the sampling schedule (Fig. 2). For example, for VG, the first spring sampling in years 2010–2012 occurred probably too late and failed to catch the baseline of the increasing trajectory. The G25 phase appears therefore more suitable in this case and, in general, in remote alpine grasslands that can be reached only after the complete snow melt, when the beginning of the growing season may already have occurred.

Patterns discussed for the seasonal trajectories are apparent also for the extracted phases (Fig. 3). In addition, the phase extraction allows to show the inter-annual differences illustrated for the seasonal trajectories in a quantitative way (Fig. 2) by means of MM and the post hoc tests. For all indices, phase G25 in 2011 occurred significantly earlier than all other years, but only G25 extracted for NG occurred significantly later in 2013 compared to other years. For the yellowing phases (Y75), it is difficult to find a pattern of inter-annual variation consistent across indices. However, productivity indices generally show higher IAV compared to greenness indices.
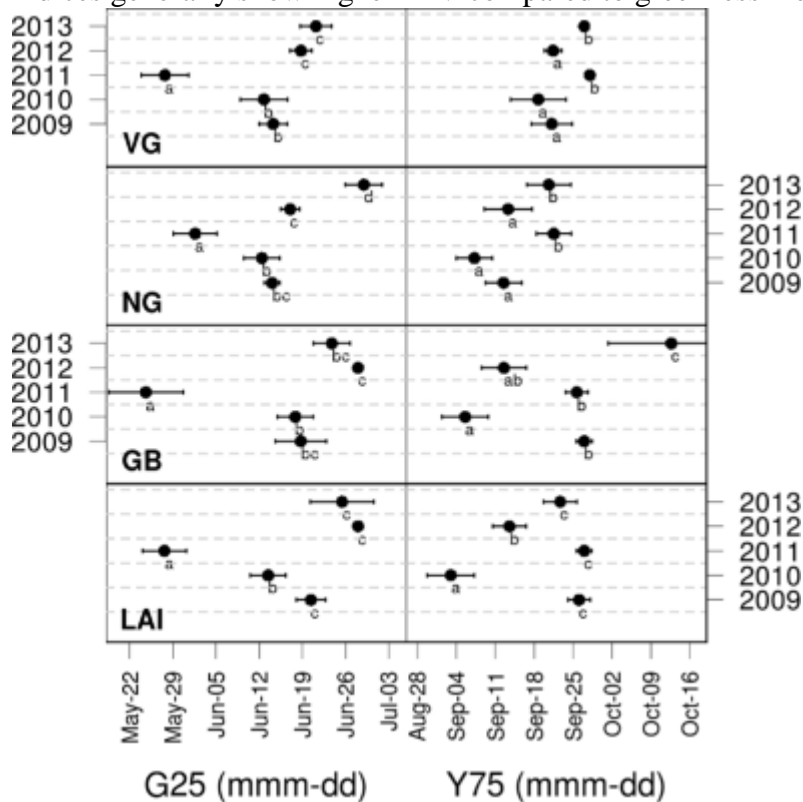

Fig. 3

Mean date of occurrence of phases G25 and Y75 extracted from percentiles on the spline fitting for all indices in 5 years. *Error bars* represent ±2se of the mean. *Different letters* indicate significant difference ($p < 0.05$) between the means (Tukey HSD). *VG* visual greening, *NG* nadiral greening, *GB* green biomass, *LAI* leaf area index

Inter-annual patterns apparent for the spring phases (e.g. the anticipated green-up in 2011) do not necessarily result in corresponding patterns in the yellowing phases (absence of a lag effect).

Relationship between indices and gross primary production

Field measurements are often associated to/substituted by automated systems to track phenology at the ecosystem level (Gonsamo et al. 2013; Noormets et al. 2009). Moreover, increasing attention is devoted to understanding the impact that phenological shifts or inter-annual variability can have on photosynthesis and carbon sequestration, i.e. the phenology of CO_2 ecosystem exchange (Richardson et al. 2010; Dragoni et al. 2011). Figure 4 reports a scatter plot between our indices and daily average GPP derived from eddy covariance measurements. All indices show some degree of correlation with GPP; however, greenness indices (VG and NG) show a stronger relationship compared to productivity indices (GB and LAI). In particular, NG appears to be the index that better correlates with GPP because it shows a slightly higher coefficient of determination and a higher inter-annual consistency compared to VG. The few previous comparisons between plant greenness and GPP report regression coefficients similar to those found in this study (e.g. Peichl et al. (2014)). As regards productivity indices, a number of studies compare GPP and biomass or LAI, some of them reporting correlations similar to ours (Hirota et al. 2010) and some reporting a stronger relationship (e.g. Flanagan et al. (2002) and Xu and Baldocchi (2004)). The relatively lower relationship between GPP and GB may reflect the higher spatial and temporal variability of the latter compared to greenness indices. In fact, GB is the index showing the highest IAV (Fig. 2).
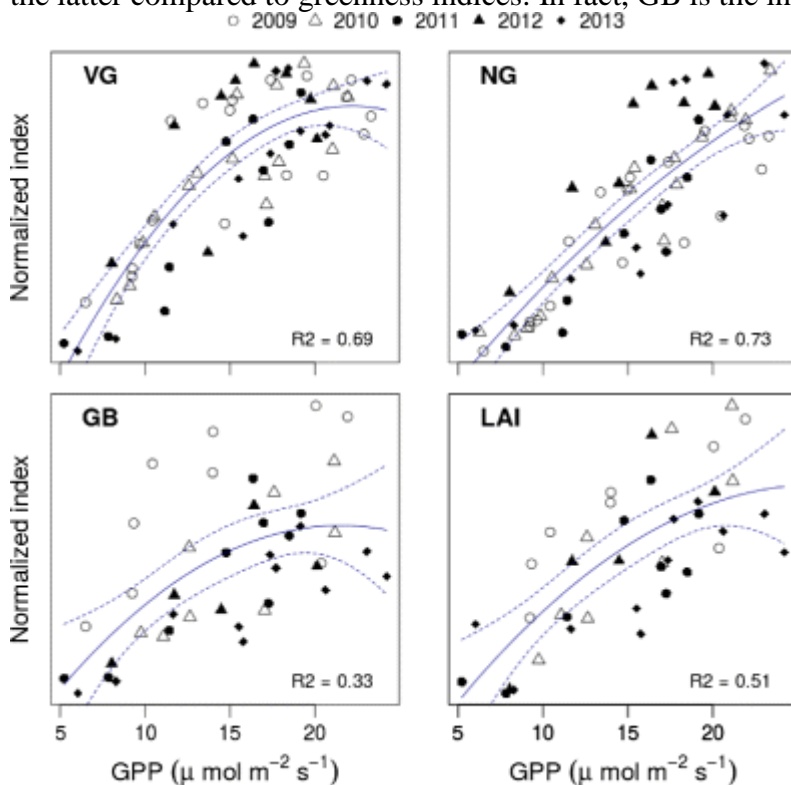

Fig. 4

Scatterplot between seasonal trajectories of the investigated indices and gross primary production (GPP). GPP data represent the mean values for 7-day windows including only fluxes measured during periods with incident photosynthetic photon flux density (PPFD) >1400 μmol m^{-2} s^{-1}. Weekly time windows are centred on the day samples for indices were collected (Flanagan et al. 2002). The coefficient of determination (r^2) of a quadratic fit and its standard error envelope are shown. *VG* visual greening, *NG* nadiral greening, *GB* green biomass, *LAI* leaf area index

Relationship between sample size and ability to detect inter-annual variation

Table 2 reports the number of sampling units required to obtain a significant inter-annual difference ($p < 0.05$) based on a power analysis. This analysis can be interpreted as the ability of the four investigated indices to track inter-annual variations in the phenology of this grassland.

Table 2

Number of sampling points at $\beta = 0.20$ (Cohen [1988]) required to obtain significant inter-year differences ($p < 0.05$), based on a power analysis

	G10	G25	G50	G75	G90	Y10	Y25	Y50	Y75	Y90
VG	4	11	27	60	91	125	34	18	12	23
NG	4	8	12	15	16	18	32	51	17	37
GB	3	9	19	25	34	33	18	21	17	28
LAI	4	7	34	40	20	26	33	23	7	8

The analysis shows that early greening phases have a very low sampling requirement compared to central and yellowing phases. It shows furthermore that all indices except LAI have a high sampling requirement for the late yellowing phases. Greenness indices (VG and NG) perform slightly worse than productivity indices (GB and LAI) in the late yellowing phases. Additionally, NG shows a relatively low sampling requirement also in the intermediate phases near the maximum seasonal development.

Although the number of sampling points obtained by the power analysis already provides an indication of sampling design in this subalpine grassland, a question still remains on whether it is preferable to arrange the sampling points into different plots and how many points per plot are required. We therefore performed an additional analysis based on resampling technique. We reduced the sample size by sequentially removing vertices and plots and fitted a MM for each possible combination of reduced dataset (Table 1). We then parametrised the inverse relationship between inter-annual differences and correspondent p values from the MM, to predict the minimum detectable difference at $p = 0.05$ ($d_{p = 0.05}$). The $d_{p = 0.05}$ was extracted for each index, phase and sample size.

Figure 5 illustrates the relationship between the number of plots and vertices and the minimum detectable difference at $p = 0.05$ for two selected phases (G25 and Y75) representative of the early and late growing season. As expected, increasing the number of vertices allows a reduction of the minimum detectable difference. However, also increasing the number of plots always leads to a reduction of the minimum detectable difference. This is noteworthy because one could expect that increasing the number of plots would capture more spatial variability and eventually result in higher representativeness but lower replicability. This occurrence will be further discussed later.
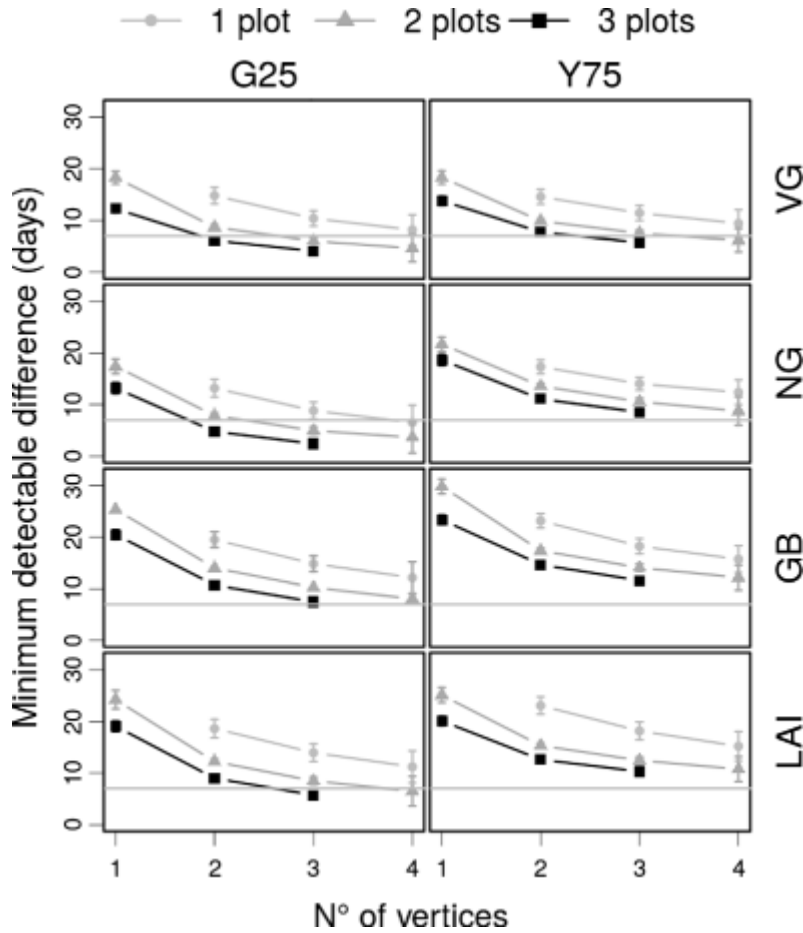
Fig. 5

Minimum detectable difference ($d_{\,p\,=\,0.05}$, expressed in days) as a function of number of plots and vertices for phase G25 (*left column*) and Y75 (*right column*) of all investigated indices from the sample removal analysis. For reference, a *horizontal grey line* at 7 days (1 week) is shown. *VG* visual greening, *NG* nadiral greening, *GB* green biomass, *LAI* leaf area index

When comparing greening and yellowing phases, the minimum detectable difference is consistently lower in G25 compared to Y75, suggesting that not only the selected indices are more suitable to track greening than yellowing dynamics but also that yellowing dynamics are more noisy due to, e.g. a higher spatial variability. Across indices, the ones related to greenness (NG and VG) always lead to a lower minimum detectable difference. In particular, NG performs slightly better than VG in the greening phases, whereas the opposite is true for the yellowing phases. The two productivity indices, i.e. GB and LAI always perform worse than the greenness indices because these indices are more sensitive to the spatial variability in plant species composition and the consequent differences in plant biomass production and LAI.

Figure 5 allows us to speculate on which would be the best trade-off between sampling effort and power in detecting inter-year differences in order to design future sampling strategies. To illustrate it, we have chosen a threshold of 7 days of inter-annual difference (the horizontal grey line in Fig. 5). This threshold was chosen because weekly sampling frequency is common in phenology sampling protocols (Richardson et al. 2006; Norby et al. 2003). Schwartz et al. (2013) report a 4-day sampling frequency as the best effort to minimise data uncertainty in field phenology of a mixed deciduous/evergreen forest. On the other hand, the temporal resolution of satellite-derived phenology at sufficient spatial resolution is likely higher than 7 days. Hence, our 7-day threshold

can be considered conservative compared to satellite products' temporal resolution and is likely the maximum affordable frequency for a manual sampling in remote mountain areas.

According to our analysis, the target of 7 days of inter-annual difference in phase G25 can be achieved by sampling at least three plots and three vertices per plot (nine samples in total) for GB and LAI. NG and VG show lower sampling requirement because for both indices, we would be able to track an inter-annual difference lower than 7 days by sampling three plots and two vertices (six samples) or two plots and three vertices (again six samples). For none of the investigated indices is one plot sufficient to track an inter-annual difference lower than 7 days. For the yellowing phases, the target of 7 days of inter-annual difference is hardly achieved with any index except for visual greening, for which either eight or nine samples arranged in two or three plots allow detecting a 7-day difference.

Consistent with the power analysis, sample removal also indicates a worse performance of the productivity indices compared to greenness indices. These appear therefore to be more robust against sample removal than productivity indices, and this may be due to the higher inter-plot and intra-plot variability of the latter.

The influence of sample size has been rarely evaluated in the past (Hudson 2010; Hemingway and Overdorff 1999), especially in mountain grasslands. Morellato et al. (2010) reported that 15 to 25 observations for a single species per sampling date are optimal to reduce uncertainty on phenological phases in a very diverse ecosystem such as tropical forest. We showed that in a rather homogeneous subalpine grassland such as in our study site, a smaller number of samples is needed to detect a 1-week IAV.

Figure 5 suggests that the total number of samples rather than their distribution in different plots is more influential in reducing the minimum detectable difference, especially for greenness indices. To clarify this concept further, we present in Fig. 6 a conceptual model drawn from our data, illustrating the relationship between minimum detectable difference, mean data variability and sample size. The mean spatial variability is expressed here as the average standard deviation of phases G25 and Y75 across all years, but standard deviations are computed for each year separately and then averaged, so to exclude the effect of inter-annual differences. Spatial variability clearly increases with sample size, whereas the opposite is true for the minimum detectable difference. Ranges of spatial variability are lower for greening than for yellowing phases and this is especially true for productivity indices. This supports the idea that spring phenology has a lower sampling requirement than autumn phenology in mountain grasslands because either autumn phenology is more variable in space or more difficult to sample with our methods.
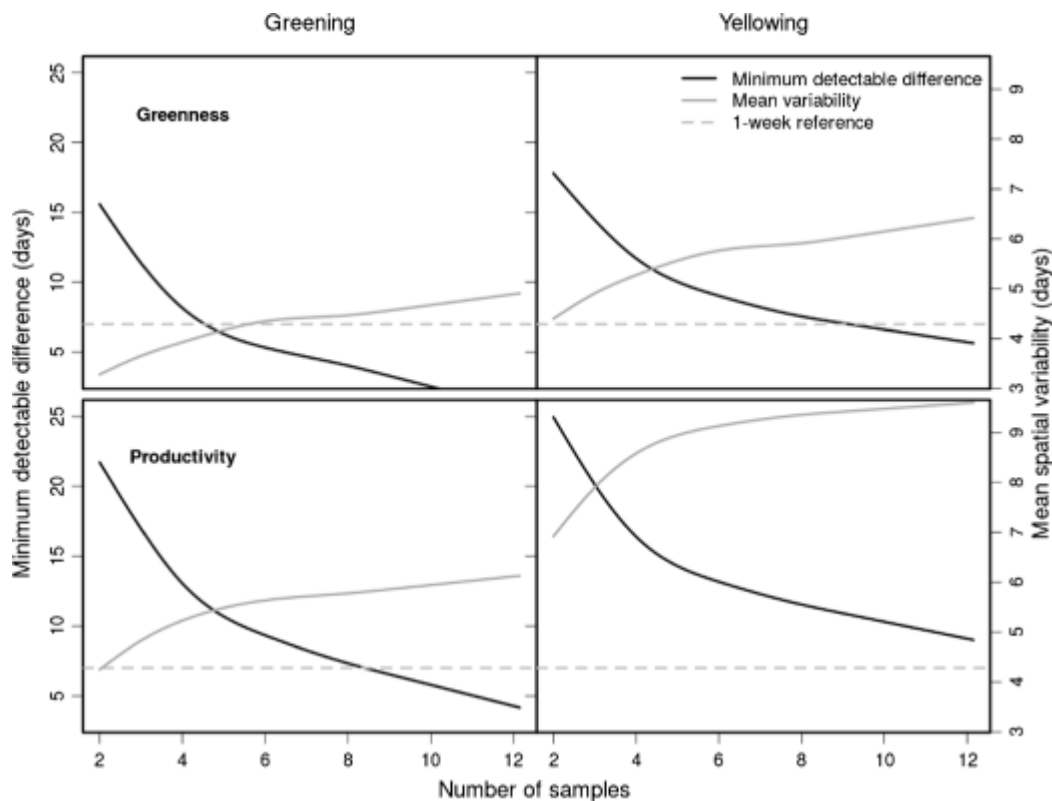
Fig. 6

Conceptual model of the relationship between sample size, spatial variability and minimum detectable difference. Model is drawn based on data of all years together. Indices are grouped in greenness indices, including NG and VG (*upper panels*) and productivity indices, including GB and LAI (*lower panels*). *Left panels* represent greening phases (G25) and *right panels* yellowing phases (Y25)

Additionally, mean variability is higher for productivity indices (GB and LAI) than for greenness indices (VG and NG). This can again indicate a higher spatial variability in GB and LAI or that these two methods are more prone to sampling error and most likely a combination of both.

Summary and recommendation for ecosystem phenology sampling protocols

At the subalpine grassland investigated in this study, an earlier snowmelt in 2011 led to an earlier but slower development of plants, whereas a later snowmelt in 2013 led to a later occurrence of early greening phases. The snowmelt appears therefore to control the beginning of the growing season. All indices were able to show these features and also agreed in showing that yellowing phases are generally more variable than greening phases.

Two of the evaluated indices (GB and LAI) require manual sampling of vegetation and subsequent analysis. One method requires visual observation of greenness (VG) and one method (NG) is based on image analysis.

GB and LAI sampling are the most time-consuming methods and show a high variability related to the species composition of the samples. As a consequence, more samples are needed to achieve a given target of accuracy in detecting IAV. The high variability in productivity indices may result from the combination of spatial variability and measurement error associated with the complexity of

the method. However, it has to be noted that GB and LAI remain important indices because they provide detailed quantitative information on the seasonal development of the structural and biophysical vegetation properties.

NG and VG perform much better than GB and LAI in detecting IAV in the phenological trajectories. Also, VG has an intrinsic deficiency, the subjectivity of the measurement, which would require that always the same observer conduct the observations. In contrast to VG, NG does not lack in objectivity and therefore probably represents the most accurate tool for monitoring grassland vegetative phenology in climate change studies (Ahrends et al. 2008; Crimmins and Crimmins 2008; Richardson et al. 2009; Migliavacca et al. 2011). Moreover, NG is the index that better correlates with gross primary production. We therefore suggest the use of nadiral pictures to depict seasonal trajectories of grassland vegetative phenology as the best trade-off between accuracy and sampling effort for long-term monitoring programmes of ecosystem productivity.

The sampling strategy discussed in this study may be used as an indication to optimise sampling protocols for grassland community phenology in the context of climate change studies.

Acknowledgments

References

1. Ahrends HE, 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(G4):G04,004
2. Aldridge G, Inouye DW, Forrest JRK, Barr WA, Miller-Rushing AJ (2011) Emergence of a mid-season period of low floral resources in a montane meadow ecosystem associated with climate change. J Ecol 99(4):905–913
3. Beniston M (2005) Warm winter spells in the Swiss Alps: strong heat waves in a cold season? A study focusing on climate observations at the Saentis high mountain site. Geophys Res Lett 32(1):L01,812
4. De Beurs K, Henebry G (2010) Spatio-temporal statistical methods for modeling land surface phenology. In: Hudson I, MR K (eds) Phenological research: methods for environmental and climate change analysis. Springer, New York, NJ, pp 157–174
5. Bréda NJJ (2003) Ground?based measurements of leaf area index: a review of methods, instruments and current controversies. J Exp Bot 54(392):2403–2417
6. CaraDonna P, Iler A, Inouye D (2014) Shifts in flowering phenology reshape a subalpine plant community. Proc Natl Acad Sci U S A Am 111(5):4916–21Chen X, Li J, Xu L, Liu L, Ding D (2014) Modeling greenup date of dominant grass species in the inner Mongolian grassland using air temperature and precipitation data. Int J Biometeorol 58(4):463–471
7. Cleland EE, Chiariello NR, Loarie SR, Mooney HA, Field CB (2006) Diverse responses of phenology to global changes in a grassland ecosystem. Proc Natl Acad Sci U S A Am 103(37):13,740–4Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd ed. Erlbaum Associates, Hillsdale, NJ
8. Cornelius C, Petermeier H, Estrella N, Menzel A (2011) A comparison of methods to estimate seasonal phenological development from bbch scale recording. Int J Biometeorol

55(6):867–877Cornelius C, Leingärtner A, Hoiss B, Krauss J, Steffan-Dewenter I, Menzel A (2013) Phenological response of grassland species to manipulative snowmelt and drought along an altitudinal gradient. J Exp Bot 64(1):241–51

9. Crimmins MA, Crimmins TM (2008) Monitoring plant phenology using digital repeat photography. Environ Manag 41(6):949–58

10. Denny E, Gerst K, Miller-Rushing A, Tierney G, Crimmins T, Enquist C, Guertin P, Rosemartin A, Schwartz M, Thomas K, Weltzin J (2014) Standardized phenology monitoring methods to track plant and animal activity for science and resource management applications. Int J Biometeorol 58(4):591–601Diez J M, Ibz I, Miller-Rushing AJ, Mazer SJ, Crimmins TM, Crimmins MA, Bertelsen CD, Inouye DW (2012) Forecasting phenology: from species variability to community patterns. Ecol Lett 15(6):545–553

11. Dragoni D, Schmid HP, Wayson CA, Potter H, Grimmond CSB, Randolph JC (2011) Evidence of increased net ecosystem productivity associated with a longer vegetated season in a deciduous forest in south-central Indiana, USA. Glob Chang Biol 17(2):886–897Dunne JA, Harte J, Taylor KJ (2003) Subalpine meadow flowering phenology responses to climate change: integrating experimental and gradient methods. Ecological Monographs 73(1)

12. Eklundh L, Jin H, Schubert P, Guzinski R, Heliasz M (2011) An optical sensor network for vegetation phenology monitoring and satellite data calibration. Sensors (Basel, Switzerland) 11(8):7678–709

13. Flanagan LB, Wever LA, Carlson PJ (2002) Seasonal and interannual variation in carbon dioxide exchange and carbon balance in a northern temperate grassland. Glob Chang Biol 8(7):599–615

14. 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(2):025,008

15. Gillespie AR, Kahle AB, Walker RE (1987) Color enhancement of highly correlated images. II. Channel ratio and chromaticity transformation techniques. Remote Sens Environ 22(3):343–365

16. Gobiet A, Kotlarski S, Beniston M, Heinrich G, Rajczak J, Stoffel M (2014) 21st century climate change in the European Alps—a review. Sci Total Environ 493:1138–51

17. Gonsamo A, Chen JM, D'Odorico P (2013) Deriving land surface phenology indicators from CO_2 eddy covariance measurements. Ecol Indic 29(0):203–207Gu L (2009) Characterizing the seasonal dynamics of plant community photosynthesis across a range of vegetation types. In: Noormets A (ed) Phenology of ecosystem processes: applications in global change research. Springer New York, NY

18. Hemingway CA, Overdorff DJ (1999) Sampling effects on food availability estimates: phenological method, sample size, and species composition. Biotropica 31(2):354–364Henebry G, De Beurs K (2013) Remote sensing of land surface phenology: a prospectus. In: Schwartz M (ed) Phenology: an integrative environmental science. Springer, New York, NJ, pp 197–210

19. Hirota M, Zhang P, Gu S, Shen H, Kuriyama T, Li Y, Tang Y (2010) Small-scale variation in ecosystem CO_2 fluxes in an alpine meadow depends on plant biomass and species richness. J Plant Res 123(4):531–541

20. Hudson IL (2010) Interdisciplinary approaches: towards new statistical methods for phenological studies. Clim Chang 100(1):143–171

21. Inouye D, Wielgolaski F (2013) Phenology at high altitudes. In: Schwartz M (ed) Phenology: an integrative environmental science. Springer, New York, NJ, pp 249–272

22. Inouye DW (2008) Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. Ecology 89(2):353–62Jones MO, Kimball JS, Small

EE, Larson KM (2013) Comparing land surface phenology derived from satellite and GPS network microwave remote sensing. International journal of biometeorology

23. 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(16):4305–4320

24. Körner C (2005) The green cover of mountains in a changing environment. In: Huber U, Bugmann H (eds) Global change and mountain regions. An overview of current knowledge. Springer, Berlin, DE

25. Lasslop G, Reichstein M, Papale D, Richardson AD, Arneth A, Barr A, Stoy P, Wohlfahrt G (2010) Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. Glob Chang Biol 16(1):187–208

26. Liang L, Schwartz MD (2009) Landscape phenology: an integrative approach to seasonal vegetation dynamics. Landsc Ecol 24(4):465–472

27. Liang L, Schwartz MD, Fei S (2011) Validating satellite phenology through intensive ground observation and landscape scaling in a mixed seasonal forest. Remote Sens Environ 115(1):143–157Linderholm HW (2006) Growing season changes in the last century. Agric For Meteorol 137(1):1–14

28. Ma T, Zhou C (2012) Climate-associated changes in spring plant phenology in China. Int J Biometeorol 56(2):269–75

29. Marcolla B, Cescatti A, Manca G, Zorer R, Cavagna M, Fiora A, Gianelle D, Rodeghiero M, Sottocornola M, Zampedri R (2011) Climatic controls and ecosystem responses drive the inter-annual variability of the net ecosystem exchange of an alpine meadow. Agric For Meteorol 151(9):1233–1243

30. Menzel A, Sparks TH, Estrella N, Koch E, Aasa A, Ahas R, Alm-Kübler K, Bissolli P, Braslavská O, Briede A, Chmielewski F M, Crepinsek Z, Curnel Y, As Dahl, Defila C, Donnelly A, Filella Y, Jatczak K, MåGe F, Mestre A, Nordli OY, Peñuelas J, Pirinen P, Remišová V, Scheifinger H, Striz M, Susnik A, Van Vliet AJH, Wielgolaski FE, Zach S, Zust A (2006) European phenological response to climate change matches the warming pattern. Glob Chang Biol 12(10):1969–1976

31. 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 AD (2011) Using digital repeat photography and eddy covariance data to model grassland phenology and photosynthetic CO_2 uptake. Agric For Meteorol 151(10):1325–1337Morellato L, Camargo M, DEca Neves F, Luize B, Mantovani A, IL H (2010) The influence of sampling method, sample size, and frequency of observations on plant phenological patterns and interpretation in tropical forest trees. In: Hudson I, MR K (eds) Phenological research: methods for environmental and climate change analysis. NE, Dordrecht

32. Noormets A, Chen J, Gu L, Desai A (2009) The phenology of gross ecosystem productivity and ecosystem respiration in temperate hardwood and conifer chronosequences. In: Noormets A (ed) Phenology of ecosystem processes. Springer, New York, pp 59–85

33. Norby RJ, Hartz-Rubin JS, Verbrugge MJ (2003) Phenological responses in maple to experimental atmospheric warming and CO_2 enrichment. Global Change Biology 9(12):1792–1801f

34. Orsenigo S, Mondoni A, Rossi G, Abeli T (2014) Some like it hot and some like it cold, but not too much: plant responses to climate extremes. Plant Ecol 215(7):677–688

35. Parmesan C (2007) Influences of species, latitudes and methodologies on estimates of phenological response to global warming. Glob Chang Biol 13(9):1860–1872

36. Peichl M, Sonnentag O, Nilsson M (2014) Bringing color into the picture: using digital repeat photography to investigate phenology controls of the carbon dioxide exchange in a boreal mire. Ecosystems pp 1–17

37. R Core Team R (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria,

38. Rammig A, Jonas T, Zimmermann NE, Rixen C (2010) Changes in alpine plant growth under future climate conditions. Biogeosciences 7(6):2013–2024Reichstein M, Falge E, Baldocchi D, Papale D, Aubinet M, Berbigier P, Bernhofer C, Buchmann N, Gilmanov T, Granier A, Grnwald T, Havrnkov K, Ilvesniemi H, Janous D, Knohl A, Laurila T, Lohila A, Loustau D, Matteucci G, Meyers T, Miglietta F, Ourcival JM, Pumpanen J, Rambal S, Rotenberg E, Sanz M, Tenhunen J, Seufert G, Vaccari F, Vesala T, Yakir D, Valentini R (2005) On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Glob Chang Biol 11(9):1424–1439

39. Richardson AD, Bailey AS, Denny EG, Martin CW, O'Keefe J (2006) Phenology of a northern hardwood forest canopy. Glob Chang Biol 12(7):1174–1188

40. Richardson AD, Braswell BH, Hollinger DY, Jenkins JP, Ollinger SV (2009) Near-surface remote sensing of spatial and temporal variation in canopy phenology. Ecol Appl 19(6):1417–28

41. Richardson AD, Black TA, Ciais P, Delbart N, Friedl MA, Gobron N, Hollinger DY, Kutsch WL, Longdoz B, Luyssaert S, Migliavacca M, Montagnani L, Munger JW, Moors E, Piao S, Rebmann C, Reichstein M, Saigusa N, Tomelleri E, Vargas R, Varlagin A (2010) Influence of spring and autumn phenological transitions on forest ecosystem productivity. Philosophical transactions of the Royal Society of London Series B. Biol Sci 365(1555):3227–46

42. Richardson AD, Keenan TF, 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 169:156–173Rutishauser T, Luterbacher J, Defila C, Frank D, Wanner H (2008) Swiss spring plant phenology 2007: extremes, a multi-century perspective, and changes in temperature sensitivity. Geophys Res Lett 35(5):L05,703

43. Schwartz MD, Hanes JM, Liang L (2013) Comparing carbon flux and high-resolution spring phenological measurements in a northern mixed forest. Agric For Meteorol 169:136–147

44. Studer S, Stöckli R, Appenzeller C, Vidale P L (2007) A comparative study of satellite and ground-based phenology. Int J Biometeorol 51(5):405–14

45. Walker MD, Ingersoll RC, Webber PJ, Ecology S, Jun N (1995) Effects of interannual climate variation on phenology and growth of two alpine forbs. Ecology 76(4):1067–1083

46. Wipf S, Stoeckli V, Bebi P (2009) Winter climate change in alpine tundra: plant responses to changes in snow depth and snowmelt timing. Clim Chang 94(1-2):105–121

47. Wohlfahrt G, Bahn M, Newesely C, Sapinsky S, Tappeiner U, Cernusca A (2003) Canopy structure versus physiology effects on net photosynthesis of mountain grasslands differing in land use. Ecol Model 170(2-3):407–426Xu L, Baldocchi DD (2004) Seasonal variation in carbon dioxide exchange over a mediterranean annual grassland in California. Agric For Meteorol 123(1-2):79–96

48. Zeeman MJ, Hiller R, Gilgen AK, Michna P, Plüss P, Buchmann N, Eugster W (2010) Management and climate impacts on net CO_2 fluxes and carbon budgets of three grasslands along an elevational gradient in Switzerland. Agric For Meteorol 150(4):519–530