Remote Sensing of Environment 298 (2023) 113835

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Extracting flowering phenology from grassland species mixtures using time-lapse cameras

Davide Andreatta [a,b,*], Christoph Bachofen [c,d], Michele Dalponte [b], Valentin H. Klaus [e,f], Nina Buchmann [e]

[a] Department of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova, Legnaro, Padova, Italy
[b] Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, Trento, Italy
[c] Plant Ecology Research Laboratory PERL, School of Architecture, Civil and Environmental Engineering, EPFL, Lausanne, Switzerland
[d] Functional Plant Ecology, Community Ecology Unit, Swiss Federal Institute for Forest, Snow and Landscape WSL, Birmensdorf, Switzerland
[e] Institute of Agricultural Sciences, Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland
[f] Forage Production and Grassland Systems, Agroscope, Zürich, Switzerland

ARTICLE INFO

Edited by Dr Marie Weiss

Keywords:
Flower phenology
Time-lapse camera
Phenocam
Biodiversity
Image classification

ABSTRACT

Understanding the impacts of climate change on plant phenology is crucial for predicting ecosystem responses. However, accurately tracking the flowering phenology of individual plant species in grassland species mixtures is challenging, hindering our ability to study the impacts of biotic and abiotic factors on plant reproduction and plant-pollinator interactions. Here, we present a workflow for extracting flowering phenology from grassland species mixtures using near-surface time-lapse cameras. We used 89 image series acquired in plots with known species composition at the Jena trait-based experiment (Germany) to develop random forest classifiers, which were used to classify images and compute time series of flower cover for each species. The high temporal resolution of time-lapse cameras allowed to select images in proper light conditions, and to extract vegetation indices and texture metrics to improve discrimination among flowering species. The random forest classifiers showed a high accuracy in predicting the cover of *Leucanthemum vulgare, Ranunculus acris,* and *Knautia arvensis* flowers, whereas graminoid flowers were harder to predict due to their green-to-brownish colours. The proposed workflow can be applied in climate change studies, ecosystem functioning, plant community ecology, and biodiversity change research, including the investigation of effects of species richness on individual species' flowering phenology. Our method could be a valuable tool for understanding the impacts of climate change on plant reproduction and ecosystem dynamics.

1. Introduction

Global change affects plant communities and their functioning in various ways. Consistent changes in the timing of phenological events are clear indicators of the impact of global change on plant life cycles (Piao et al., 2008; Schwartz, 2013). For instance, warming tends to advance the green-up and to delay the end of the vegetation growing season (Estiarte and Peñuelas, 2015; Liu et al., 2020; Menzel et al., 2006; Shen et al., 2011). However, recent studies revealed that the analysis of phenology at whole-ecosystem scale is not always suitable for describing the effect of global change on individual plant species, because phenological responses to climate change can differ among species of the same ecosystem (Thackeray et al., 2016). Furthermore, Collins et al. (2021) challenged the expectation that all phenological events will advance in unison to warming. Instead, they observed that vegetative and reproductive phenology are differentially affected by experimental warming, suggesting that different aspects of phenology should be separately investigated. Moreover, it was observed that many plant species

Abbreviations: B, Blue Digital Numbers; CV, Computer Vision; FCTS, Flower Cover Time Series; G, Green Digital Numbers; GLI, Green Leaf Index; NGRDI, Normalised Green Red Difference Index; PC, PhenoCam; R, Red Digital Numbers; RF, Random Forest; RGBVI, Red Green Blue Vegetation Index; RS, Remote Sensing; SFFS, Sequential Floating Forward Selection; TBE, Trait-Based Biodiversity Experiment; VARI, Visible Atmospherically Resistant Index.

* Corresponding author at: Department of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova, Legnaro, Padova, Italy & at: Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, Trento, Italy.

E-mail addresses: davide.andreatta@fmach.it (D. Andreatta), christoph.bachofen@epfl.ch (C. Bachofen), michele.dalponte@fmach.it (M. Dalponte), valentin.klaus@usys.ethz.ch (V.H. Klaus), nina.buchmann@usys.ethz.ch (N. Buchmann).

https://doi.org/10.1016/j.rse.2023.113835
Received 17 March 2023; Received in revised form 31 August 2023; Accepted 26 September 2023

flowered earlier in response to reductions in diversity, so that declining diversity could exacerbate phenological changes attributed to rising global temperatures (Wolf et al., 2017). To further investigate these multifaceted processes, there is an increasing need for effective methods to track single species flowering phenology in species mixtures.

Thanks to their planetary-scale analysis capabilities and short revisiting time, remote sensing (RS) and proximal sensing are opening new possibilities for phenological studies, overcoming laborious and time-consuming ground-based vegetation observations (Szigeti et al., 2016). The large-scale observation potential of RS has been applied to track vegetation reproductive phenology (Gonzales et al., 2022), but the coarse spatial resolution of satellite images restricts flowering estimation to massive homogeneous flowering events in rather homogeneous ecosystems such as of eucalypt species (Dixon et al., 2021), oil seed rape fields (d'Andrimont et al., 2020), almond (Chen et al., 2019) and pear plantations (Wouters et al., 2013). Flowers of different functional groups were mapped for the first time by Landmann et al. (2015), in African savannas using hyperspectral imagery. Images at finer spatial resolution captured by drones recently allowed Gallmann et al. (2022) to recognize flower species in permanent grasslands, a task that would not have been possible with satellite-resolution images. Frequent (sub-weekly) drones flights, however, are usually too expensive for agricultural and ecological phenological studies, and other technologies are therefore needed.

PhenoCams (PCs), i.e., digital cameras configured to capture time-lapse images, can bridge the gap between satellite monitoring and traditional ground-based vegetation observations (Brown et al., 2016; D'Odorico et al., 2015; Richardson et al., 2010). Compared to RS, PC imagery can provide a very fine temporal and spatial resolution, allowing to explore the inter- and intraspecific variability in plant phenology at a sub-daily scale to a much lower cost than repeated drone flights. However, only in 2022, PC images were used for the first time to map flowering phenology of two Arctic species, the mountain avens *Dryas octopetala* and *Dryas integrifolia* (Mann et al., 2022). For more complex ecosystems, such as multi-species and multi-layered grasslands, different automation routines and analysis processes still need to be developed.

Tracking floral phenology using time-lapse camera in grasslands is challenging due to many reasons: i) images are acquired under various light conditions, ii) sensors usually measure reflectance values only in the visible spectral region, iii) flower structures are relatively small and only cover a few pixels, iv) flowers might be occluded by vegetative plant parts, and v) grasslands are biodiversity rich compared to other ecosystems, to name a few (Andrew and Ustin, 2008; Gallmann et al., 2022; Mann et al., 2022). Even though PC imagery has been used in phenological studies at whole-ecosystem scale to track greenness, and despite floral phenology (typically determined manually) being a key trait of grasslands ecosystems, no processing workflow to track flowering phenology in grasslands has been proposed so far. Nevertheless, a workflow to extract flower cover time series (FCTS) from PC imagery is urgently needed to, for example, study the response of reproductive phenology to environmental and biotic drivers. Moreover, such a workflow could easily be applied to different questions of biodiversity and climate impact research as well as land management to assess plant-pollinator interactions, grasslands cultural services evaluation, and grassland productivity monitoring, providing important inputs to vegetation and biogeochemical models (Inouye, 2020; Richardson et al., 2012; Wolf et al., 2017).

Here, we suggest a processing workflow to extract FCTS from RGB time-lapse cameras (RGB as Red, Green, Blue digital numbers). To address the aforementioned challenges, we i) leveraged the high temporal resolution of PC imagery by selecting only images in proper light condition, ii) based our classification on vegetation indices derived from RGB reflectance, iii) included texture metrics to improve discrimination among flower species by their shapes. More specifically, we applied image filtering, calculated features of selected pixels (vegetation indices and texture metrics), and then used the subset of features with highest accuracy to train random forest classifiers. Finally, we extracted FCTS and derived phenological metrics for single or groups of plant species. We present an example application on experimental grassland plots of different diversity levels.

2. Materials and methods

2.1. Study site

Images used in this study were acquired in 2014 within the Trait-Based Biodiversity Experiment (TBE; Ebeling et al., 2014) at the field site of the Jena Experiment (Thuringia, Germany; 50°55 N, 11°35 E, 130 m a.s.l.) (Roscher et al., 2005). The TBE was established in 2010 following a design which covers gradients in plant species and functional richness, ranging from 1 to 8 species in 138 plots (3.5 m × 3.5 m). Species not belonging to the initially sown species pool were weeded every year in April, July and October to maintain the species richness gradient. Grasslands were mowed two times per year to mimic local traditional management. We included 89 plots, covering the whole species richness gradient. Selected plots were sown with a combination of 13 species, of which seven were grasses: *Anthoxanthum odoratum, Avenula pubescens, Dactylis glomerata, Festuca rubra, Holcus lanatus, Phleum pratense, Poa pratensis;* and six were forbs: *Centaurea jacea, Geranium pratense, Knautia arvensis*, *Leucanthemum vulgare*, *Plantago lanceolata,* and *Ranunculus acris.* Selected species were cultivated as monocultures (14 plots), in 2-species mixtures (32 plots), in 3-species mixtures (23 plots), in 4-species mixtures (18 plots), and in 8-species mixtures (2 plots).

2.2. Image acquisition

In spring 2014, 92 time-lapse cameras (TLC 100, Brinno) were installed on 1.5 m poles pointing north at 60° angle from horizontal, capturing an area of 3.5 m^2 in each frame. Images were recorded hourly during daylight according to the automatic mode of the cameras from April 12th through August 22th, 2014 (Fig. 1). Here, we focus on the spring growing period, i.e., between the spring weeding (April 24th) and the first mowing (May 29th). Plots 20, 27 and 33 of the TBE plots were discarded because of failure in image collection. A total of 52′678 images stored in jpg format (1280 × 1040 pixels) were considered in the study.

2.3. Workflow for data processing

The proposed workflow (Fig. 2) can be divided in four main phases: 1) image selection and pixel labelling, 2) feature computation, 3) feature selection and final classifier compilation, 4) FCTS extraction, smoothing and calculation of phenological metrics. All analyses were performed with the R version 4.3.0 (R Core Team, 2023).

2.3.1. Image selection and pixel labelling (phase 1)

In Phase 1, image selection phase aimed at increasing the spectral separability between six pixel classes. *C. jacea* and *G. pratense* did not flower during the spring and could therefore not be considered in the analyses, while the green-greyish *P. lanceolata* flowers were not big enough to be labelled separately from "Green vegetation". The graminoid species were combined for the flower identification because their flowers were not distinguishable, whereas the other three flowering species (*R. acris*, *K. arvensis*, and L. *vulgare*) were considered separately. The resulting classes used for labelling elements inside the plots were: "Green vegetation", "Soil", "Graminoids flowers", "*K. arvensis* flowers", "*L. vulgare* flowers", "*R. acris* flowers". Light conditions heavily affect pixel colours: images with high brightness were usually foggy, and images with high contrast were usually acquired in direct sunlight conditions. We calculated brightness and contrast for all images using the "extractVIs" function of the R package "Phenopix" (Filippa et al., 2016) and tested which brightness and contrast combinations allow the

Fig. 1. Examples of analysed images. Left: Plot where *G. pratense, H. lanatus, P. pratense* and *P. lanceolata* were sown; image acquired May 27th, 2014. Right: Plot where *C. jacea, K. arvensis* and *L. vulgare* were sown; image acquired May 23th, 2014.

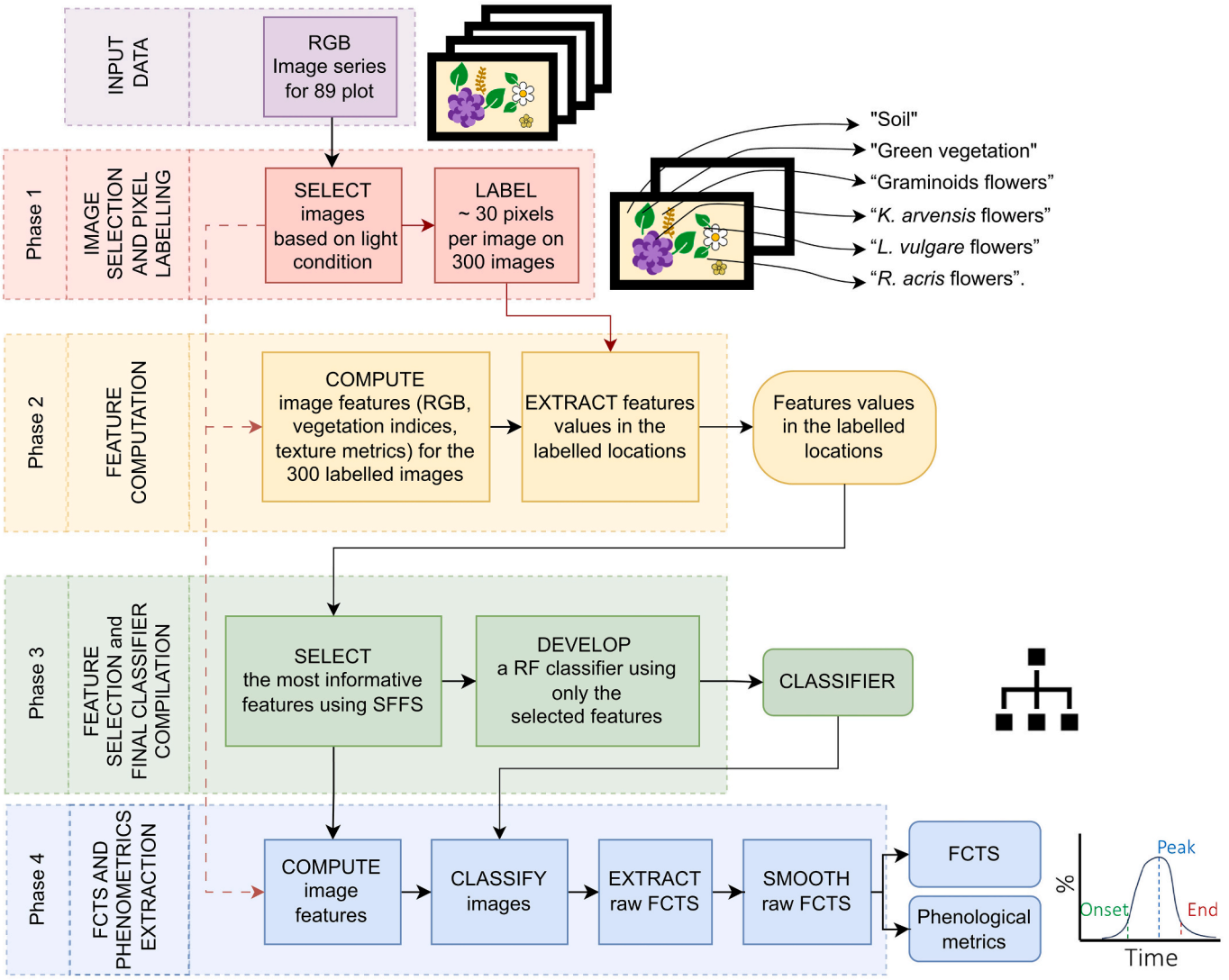

Fig. 2. Structure of the proposed workflow with four phases to extract flower phenology from PhenoCam pictures of a grassland biodiversity experiment. Abbreviations are as follows: RF = Random Forest; SFFS = Sequential Floating Forward Selection; FCTS = Flower Cover Time Series.

selection of images acquired in homogeneous light conditions. Images with uniform light conditions were retrieved by selecting brightness and contrast between the 10th and the 40th percentile within a 3-day window. The selection of the best images within this 3-day window avoided including images taken on days with sub-optimal observations (e.g., all foggy or high contrast images).

To develop a labelled dataset, 300 images were randomly selected (60 images in the period between Apr 24 and May 5; 60 images between May 6 and May 18; 180 images between May 19 and May 29, 2014). For each image, a 200 pixels × 200 pixels image patch was randomly selected and plotted in RGB colours using the "plotRGB" function of the "raster" package (Hijmans, 2022). Around 30 pixels per image were labelled by clicking on the image to retrieve the x and y coordinates using the "locator" function of the "graphics" package and assigning to each pixel the class to which it belongs (see subsection 2.1). To prevent duplicated pixels after downscaling the images (see the subsequent section for downscaling details), any labelled pixels that were within a distance of eight pixels from one another were removed from the dataset. The labelling phase resulted in a table where the class and pixel coordinates were stored.

2.3.2. Feature computation (phase 2)

To increase the spectral separability of pixels between different classes, we computed RGB-based features: vegetation indices and texture metrics, described in detail in Table 1. We selected four vegetation indices well established in colour analysis literature (Lussem et al., 2018; Zhao, 2021). For pixels with a specific shade of purple colour, the calculation of the Visible Atmospherically Resistant Index (VARI) resulted in infinite values (for definition, see Table 1). Since only finite values can be used for classifier development, infinite VARI values were replaced with the highest finite value sampled (or lowest in case of negative infinite values), which occurred in <0.1% of the labelled pixels. The image textures were derived from co-occurrence matrices for each colour band, since we expected that the flower colours differed from the background (green vegetation or soil) surfaces (Guru et al., 2010), using the "glcm" package in R Studio (Zvoleff, 2020; Haralick et al., 1973). Homogeneity, Contrast, Dissimilarity, Entropy, Second Moment, Mean, and Variance were computed in four directions (0°, 45°, 90° and 135°) and then averaged to one rotation-invariant texture as commonly used in texture analysis (e.g., Guru et al., 2010). For the computation of texture metrics, we needed to define the size of the window used for co-occurrence matrices. Moreover, downscaling the images to a lower resolution before feature extraction can give the best detection accuracy while also vastly increasing processing speed compared to higher resolution images (Mann et al., 2022). We tested the influence of window size and downscaling factor on classification accuracy and processing time, and found that a downscaling factor equal to four and a window size equal to eleven resulted in the highest accuracy (Fig. S1). Processing time of the downscaled image (4 × 4 pixels) was 16 times shorter than the processing time of the full resolution image (24 s per image vs. 395 s per image, respectively). The feature values of the labelled pixels were then extracted. The feature computation phase resulted in a table where class, and features values of the labelled pixels were stored.

Table 1
Image features tested for Phase 2. $P_{i,j}$ is the probability of values i and j occurring in adjacent pixels in the original image within the window defining the neighbourhood. i and j are the labels of the columns and rows (respectively) of the co-occurrence matrixes. Because of the construction of the co-occurrence matrixes, i refers to the value of a target pixel, and j is the value of its immediate neighbour (Rook's case).

Feature name	Equation	Reference
Red Digital Number	R	
Green Digital Number	G	
Blue Digital Number	B	
RGBVI (Red Green Blue Vegetation Index)	$\frac{((G^2)-(R^*B))}{((G^2)+(R^*B))}$	Bendig et al. (2015)
GLI (Green Leaf Index)	$\frac{(2^*G-R-B)}{(2^*G+R+B)}$	Louhaichi et al. (2001)
VARI (Visible Atmospherically Resistant Index)	$\frac{(G-R)}{(G+R-B)}$	Gitelson et al. (2002)
NGRDI (Normalised Green Red Difference Index)	$\frac{(G-R)}{(G+R)}$	Tucker (1979)
Homogeneity	$\sum_{i,j=0}^{N-1}\frac{P_{ij}}{\left(1+(i-j)^2\right)}$	Haralick et al. (1973)
Contrast	$\sum_{i,j=0}^{N-1}P_{ij}(i-j)^2$	Haralick et al. (1973)
Dissimilarity	$\sum_{i,j=0}^{N-1}P_{ij}\lvert i-j\rvert$	Haralick et al. (1973)
Entropy	$\sum_{i,j=0}^{N-1}P_{ij}\left(-lnP_{i,j}\right)$	Haralick et al. (1973)
Second Moment	$\sum_{i,j=0}^{N-1}P_{i,j}^2$	Haralick et al. (1973)
Mean	$\mu=\sum_{i,j=0}^{N-1}i\left(P_{i,j}\right)$	Haralick et al. (1973)
Variance	$\sigma^2=\sum_{i,j=0}^{N-1}P_{ij}(1-\mu)^2$	Haralick et al. (1973)

2.3.3. Feature selection and final classifier development (phase 3)

In Phase 3, we selected a set of best suitable features to optimize processing time, and to reduce redundancy of highly correlated features. Decreasing the number of features typically increases the classifier generalisation capability, because it avoids overfitting (Ho, 1995). First, we randomly assigned 70% of images for training, and 30% of images for validation. Validating a classifier on a separate part of the dataset is a common technique used to evaluate the performance of the classifier and to avoid overfitting. For the feature selection, we used the training dataset and applied the "varSelSFFS" function from the "varSel" package, which performs feature selection using the Sequential Forward Floating Selection search strategy and the Jeffries-Matusita distance (Bruzzone et al., 1995; Dalponte and Ørka, 2021; Pudil et al., 1994). The Jeffries-Matusita distance saturates at square root of two, when including a new feature does not increase class separation. Thus, the number of features to select was defined according to the saturation, as described in Richards and Jia (2006). In addition, we investigated the capability of RGB bands, vegetation indices, and texture metrics to distinguish classes. For this, we compared the accuracies of RF models trained on different subsets of features from the training dataset, including: i) features selected by SFFS, ii) RGB bands alone, iii) RGB bands combined with vegetation indices, iv) RGB bands combined with texture metrics, and v) all features. The accuracies were measured on the validation dataset.

To perform RF classifications, we used the "randomForest" function of the "randomForest" package (Liaw and Wiener, 2002). The metric to calculate the accuracy of the RF classifiers was the mean F1 score of the six classes. The F1 score is derived from precision and recall metrics as described in eq. 1. The precision is intuitively the ability of the classifier not to label a sampled pixel as positive when it is negative, whereas the recall is the ability of the classifier to find all the positive sampled pixels. Precision and recall are described in eqs. 2 and 3, where *tp* is the number of true positives, *tn* is the number of true negatives, *fp* the number of false positives, and *fn* the number of false negatives. All the described metrics have their best score at 1 and their worst score at 0 (Congalton and Green, 2009).

$$F1=\frac{2^*(precision^*recall)}{precision+recall} \quad (1)$$

$$precision = \frac{tp}{tp + fp} \quad (2)$$

$$recall = \frac{tp}{tp + fn} \quad (3)$$

We calculated the processing time for the calculation of all features and subsequent image classification for one image. For this, we used one core of an AMD Ryzen 73,700 U processor (CPU, 2300 MHz) with 16 GB RAM, and 500 GB solid-state drive storage device. The feature combination providing the best trade-off between accuracy and processing time was selected for the RF final classifier compilation.

2.3.4. Extraction of flower cover time series and phenological metrics (phase 4)

Once the final RF classifier had been trained, the percentage of pixels in each class was computed for each image. For this, images were selected (see section 2.3.1), for each image the selected features were computed, and percentages of each class within each image were calculated using the RF classifier developed in subsection 2.3.3. We identified and removed outliers from the derived flower cover time series using the "tsclean" function of the "forecast" R package which is based on Friedman's SuperSmoother for non-seasonal series (Hyndman and Khandakar, 2008). Values were aggregated at daily temporal resolution by taking the arithmetic mean. A Local Polynomial Regression function was fitted to smooth the time series using the "loess" function of the "stats" package (R Core Team, 2023). Time series calculated from nine TBE plots were displayed and analysed to show the potential applications and limitations of the proposed workflow. To obtain further insights into the reliability of FCTS, we conducted an analysis of flower cover of species that were not sown in each plot. For each image series (i.e., for each plot), we identified the predicted FCTS of unsown species, e.g., the predicted FCTS of L.

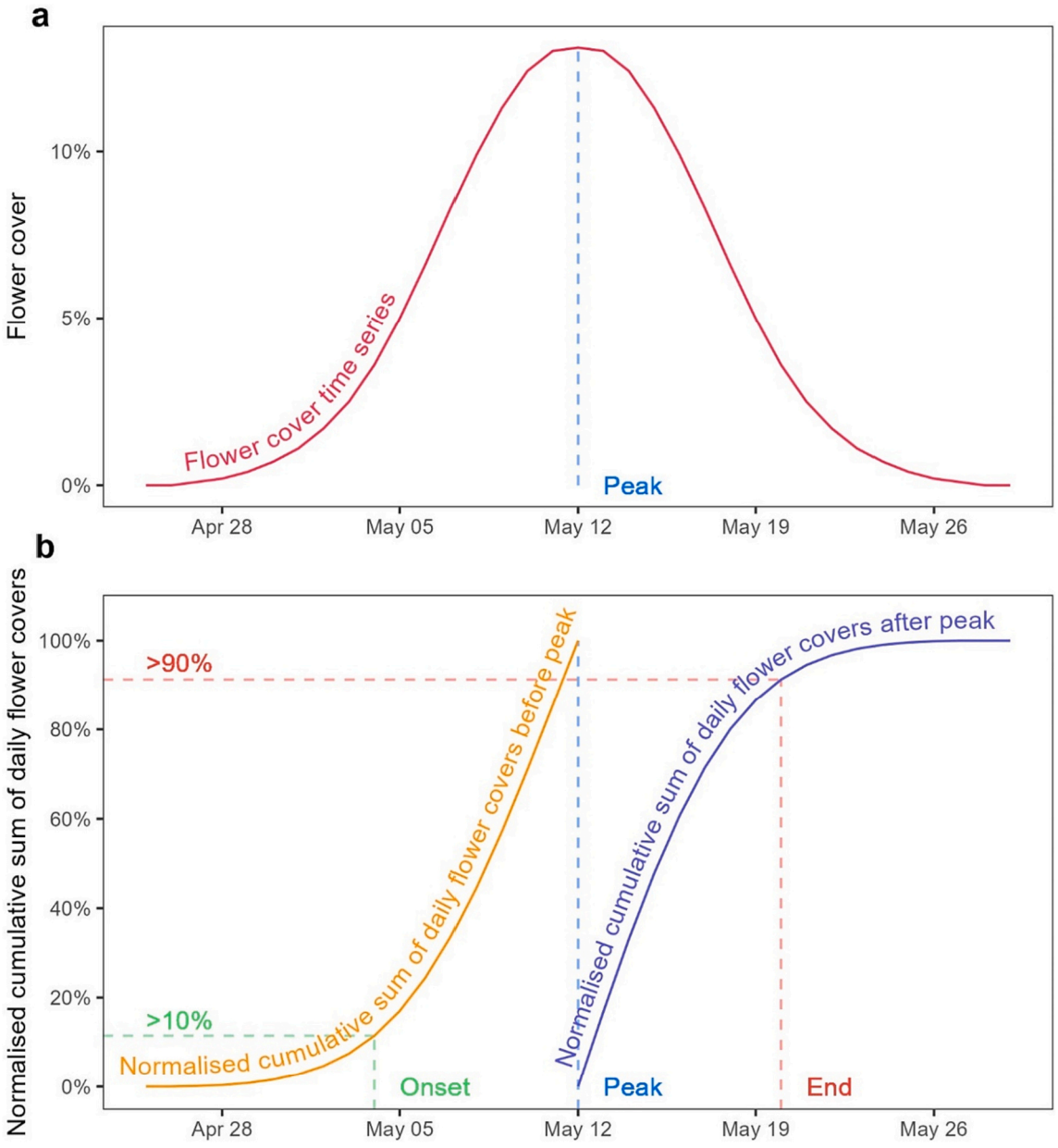

Fig. 3. Identification of flowering phenological metrics based on time-lapse cameras. Panel a: Example of flower cover time series. Panel b: Flowering phenological metrics identification for the flower cover time series in panel a; onset was defined as the first day above 10% of the normalised cumulative sum of daily flower covers before the peak; end of the season was identified as the first day above 90% of the normalised cumulative sum of daily flower covers after the peak. Both approaches allow determination of onset and end of season when logistics or management prohibited recording the full flowering season (see main text).

vulgare in plots where L. *vulgare* was not sown.

For each FCTS, onset, peak, and end of flowering were extracted. The peak was identified as the day of maximum in the FCTS, when the value was higher than the values before and after it. The onset of flowering was identified on the basis of the normalised cumulative sum of daily flower covers before the peak, whereas the end of flowering was identified on the basis of the normalised cumulative sum of daily flower covers after the peak. This allowed the identification of flowering onset in FCTS when the end of the flowering was not observable (e.g., because of mowing) as well as the identification of the end of flowering in FCTS when the onset of flowering was not observable (e.g. image acquisition started later). The cumulative sums were min-max normalised (0%–100%), and the onset was defined as the first day when the normalised cumulative sum of daily flower cover exceeded 10% (Fig. 3). Moreover, the end of the season was identified as the first day when the normalised cumulative sum of daily flower cover after the peak exceeds 90%. The 10% and 90% thresholds were chosen as a compromise between robustness against outliers and timely identification of changes. The onset of flowering was determined exclusively for FCTS exhibiting a low flower cover (< 1%) at the start of the observation period to avoid errors in plots for which the observation period started after the onset of flowering. Similarly, the end of flowering was defined exclusively for FCTS with a low flower cover (< 1%) at the end of the observation period to prevent the mischaracterization of the end of flowering in plots for which the observation period ended prior to the end of flowering. We expected a flower cover of unsown species above 0% due to wrongly classified pixels and therefore did not extract phenological metrics from time series for which the peak of the sown species was lower than 1% to avoid potential misclassification. The phenological metrics of single species that were calculated with this approach can easily be compared between treatments (i.e., multiple image time-series), and summary statistics can be derived from multiple plots, such as mean and standard deviation as well as further statistical analyses.

3. Results

With the proposed workflow we were able to successfully develop a RF model tailored to the recorded PC images, and thereby extract flower cover time-series and flowering phenology metrics of single species or groups of species from 89 image series. After the image selection based on light conditions in Phase 1 (see Fig. 2), there were on average more than three valid images per day per plot. The median number of images per plot in the period of interest was thereby reduced from 592 to 137 images per plot, leaving in total 11′472 images out of the originally 52′678. Tables and figures showing image availability before and after image selection are available in the supplementary material (Fig. S2, Table S1, Table S2).

The dataset used for the RF classifiers training and validation consisted of 9073 pixels. The "Green vegetation" class was the most represented, with 4281 pixels from 300 images. 1184 pixels were labelled as "Soil" from 139 images, 1570 pixels were labelled as "Graminoids flowers" from 115 images, 1160 as "*L. vulgare* flowers" from 65 images, 506 as "*K. arvensis* flowers" from 36 images, and 372 as "*R. acris* flowers" from 40 images. The average number of labelled pixels per image was 30. Labelling 9073 pixels in 300 images took around 300 min (labelled pixels highlighted on RGB images are available as supplementary materials in the ETH Zurich repository).

The distribution of pixel classes in the RGB space (Fig. 4) suggested a good spectral separability of some classes (e. g., "*R. acris* flowers" vs. "*K. arvensis* flowers"; "Green vegetation" vs "*R. acris* flowers"), whereas some other spectral signatures were not easily distinguishable in the RGB space (e.g., "Soil" vs. "Graminoids flowers" vs. "*K. arvensis* flowers"). The classifier developed using RGB bands resulted in an accuracy of 0.791 (Table 2). The addition of vegetation indices and texture

Table 2
Accuracy (mean F1 score of the six classes) and processing time (s image^{-1}) of random forest classifiers developed using different combinations of features (number of features in brackets). The set of eleven features selected using sequential floating forward selection (SFFS) gave a slightly lower accuracy compared to the model including all 28 features (0.888 vs. 0.905) but required less than half of the time for image processing (11 vs. 24 s). It was therefore chosen as the best feature set.

	RGB (3)	RGB + vegetation indices (7)	RGB + texture metrics (24)	All features (28)	Selected using SFFS (11)
Mean F1 score	0.791	0.800	0.883	0.905	0.888
Processing time (s image^{-1})	4	4	23	24	11

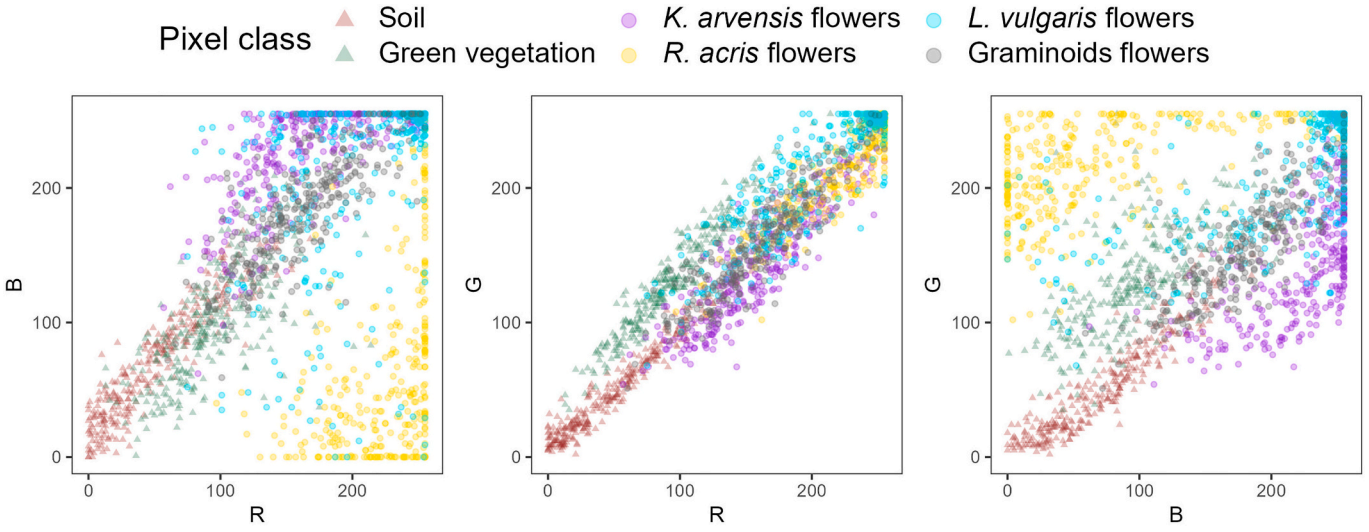

Fig. 4. Spectral separability of the labelled pixels in the RGB space. Based on two digital numbers (R and B, R and G, B and G), the overlaps of the six classes are presented. R is the red band digital number, G is the green band digital number, B is the blue band digital number. Values were extracted from 300 pixels per class without downscaling. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

metrics increased this accuracy up to 0.905. Sequential floating forward selection led to the identification of eleven features out of a total of 28 as the most informative (Phase 3). This reduction in feature number reduced the processing time from 24 s to 11 s per image, without remarkable changes in accuracy (mean F1 scores of the six classes were 0.888 vs. 0.905). The eleven selected features were red, green and blue digital number, GLI, NGRDI, RGBVI, VARI, Second Moment computed on the red band, as well as Contrast, Second Moment, and Entropy computed on the blue band.

The confusion matrix of the classification performed with the best RF classifiers gave insights into the quality of our proposed workflow (Table 3). All six classes had precisions above 0.78, indicating a low proportion of wrong pixels in the classified classes. "Graminoids flowers" were difficult to distinguish from "Soil", and "*K. arvensis* flowers" were difficult to distinguish from "Graminoids flowers" (see also Fig. 4). The recall of "Graminoids flower" and "*K. arvensis* flowers" classes was therefore the lowest (0.79 and 0.76, respectively). 3.4% of the pixels labelled as "Soil" were classified as "Graminoids flower". Even though this value appears to be low, it will result in a substantial overestimation of graminoids flowers, since a large number of the pixels in the images were classified as soil pixels at the start of the season. At the end of the season, conversely, "Graminoids flower" cover could be underestimated, since 9% of the pixels labelled as "Graminoids flower" were misclassified as soil.

In Phase 4, we extracted time series of flower cover for all the plots, e.g., a plot where *L. vulgare*, *R. acris, P. pratensis* and *G. pratense* were sown (Fig. 5). In this example plot, we could observe that L. *vulgare* and *R. acris* were the dominant flowering species (left panel) and that classified images showed a good match with RGB images. Moreover, *L. vulgare* flowered later than *R. acris*, reaching its peak five days before mowing date (May 25th). Graminoids flower cover showed positive values around 1.5% in the fitted time series, even though the flowers of *P. pratensis,* the only graminoid species sown, were not present in the RGB images (Fig. 5, right panel). This indicated that in this case the graminoids flower cover was overestimated.

We applied the developed workflow and extracted time series for all plots. Here we show the result for nine exemplary plots, dominated by L. *vulgare* (Fig. 6 A, B, and C), by graminoids (Fig. 6 D, E, F), or by *R. acris* and *K. arvensis* (Fig. 6 G, H, I). *L. vulgare* and graminoid-dominated plots showed the highest maximum flower covers, whereas *K. arvensis* and *R. acris* showed lower flower covers. The peak day of flowering of each species differed among plots: graminoids started flowering more slowly compared to the other species, before developing faster than other species after mid-May.

We further investigated the seasonal average of FCTS of species that were not sown in the plots (Figs. 7 and S3). Our findings showed that in four out of 89 plots, the flower cover of these unsown species exceeded 10%, primarily due to the presence of pixels misclassified as graminoids flowers. However, the average flower cover for unsown species across all 89 plots was relatively low with 2%. Furthermore, when the graminoids class was excluded, the error was almost negligible, being just 0.6%. In the experimental setting of the TBE, flowering started in some cases before the observation period had begun (i.e., before the spring weeding took place). These occurrences were identified (see Section 2.3.4) and the onset of flowering was not extracted for these cases. Similarly, in some cases flowering did not reach its peak before the end of the observation period (i.e., grassland mowing on May 30th) and consequently the end of flowering was not extracted. Peak day was extracted from 33 time-series, onset day from 16 time-series, and end of flowering from eight time-series.

4. Discussion

We propose a workflow to efficiently track flowering phenology of individual plant species or groups of plant species in grasslands using time-lapse cameras, which are widely applied in ecological studies. Therefore, sensor availability and installation are no limiting factors for ecologists who can use the proposed workflow for various applications (Brown et al., 2016; see subsection 4.1). Specifically, we propose an automated selection of vegetation indices and texture metric features to enhance the accuracy and processing time of a random forest classifier. The workflow can easily be replicated following FAIR principles and can be applied to new case studies. The codes have been developed in the free software R (R Core Team, 2023, GNU General Public License), and a tutorial is provided (https://github.com/andreattad/Flower_covers_phenocams).

4.1. Possible applications of flower cover extraction workflow

Multiple opportunities to apply the developed procedure in basic and applied ecological research exist. Fields of application span from climate change studies over ecosystem functioning to plant community ecology and biodiversity change research, with both experimental as well as observational settings.

Application is possible in biodiversity research such as biodiversity-ecosystem functioning experiments. The experimental site where the current study was conducted was designed to investigate species interactions and to mechanistically understand biodiversity-ecosystem functioning relationships (Ebeling et al., 2014). However, manual assessments of flowering phenology are very labour-intensive and cannot be carried out regularly. In contrast, time-lapse cameras with the proposed processing workflow can be applied to investigate if and how individual plant species change their flowering phenology. The proposed workflow opens new possibilities in the study of flowering phenology of individual species in response to a wide range of biotic and abiotic drivers, for example to assess the effects of increased carbon dioxide concentrations and higher temperatures, heat and drought stress on reproductive phenology (Collins et al., 2021; Dorji et al., 2020; Fernández-Pascual et al., 2019). Pollinator ecology is another research field that could benefit strongly from the availability of the proposed

Table 3
Confusion matrix of the final random forest classifier on the validation dataset used in Phase 3. The mean F1 Score of the six classes was 0.888, the mean recall was 0.888, and the mean precision was 0.888.

		Reference pixel class							
		K. arvensis flowers	*L. vulgare* flowers	Graminoids flowers	*R. acris* flowers	Green vegetation	Soil	Total	Precision
Predicted pixel class	*K. arvensis* flowers	117	21	12	0	0	0	150	0.78
	L. vulgare flowers	6	363	21	1	1	0	392	0.93
	Graminoids flowers	28	11	334	0	12	14	399	0.84
	R. acris flowers	0	2	1	163	2	0	168	0.97
	Green vegetation	2	4	16	5	1157	7	1191	0.97
	Soil	0	1	38	0	34	394	467	0.84
	Total	153	402	422	169	1206	415	2767	
	Recall	0.76	0.90	0.79	0.96	0.96	0.95		

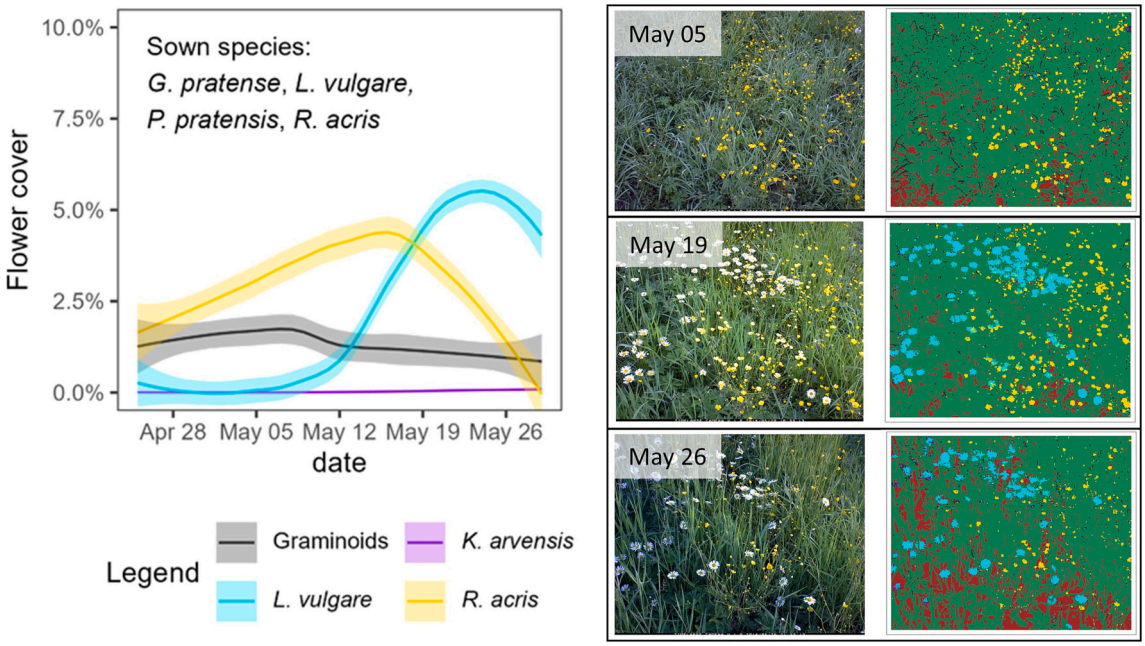

Fig. 5. Left panel: Time series of flower cover extracted from images from an example plot where L. *vulgare, R. acris, P. pratensis* and *G. pratense* were growing. Time series were fitted using Local Polynomial Regression, and 0.95 confidence intervals are displayed. Right panel: RGB and classified images acquired on the same example plot on May 5th, May 19th and May 26th, 2014 are presented. Green and brown pixels represent the "Green vegetation" and "Soil" classes, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

workflow, since plant-pollinator interactions are strongly time-sensitive (Byers, 2017; Dicks et al., 2021; Freimuth et al., 2022; Vasiliev and Greenwood, 2021; Vázquez et al., 2023).

Grasslands do not only provide animal feed, pollen and nectar, but they also provide cultural services, which are relevant for tourism, recreation, mental and physical human health, aesthetically appreciated, inspire art as well as design, and are considered a typical feature of cultural landscape in many world regions (Richter et al., 2021). Animal feed production could also benefit from the proposed workflow as grassland management requires accurate data on plant phenology in near real time as a predictor of forage quality. Flowering phenology is crucial also in plant breeding (Arzani et al., 2004; Jung and Müller, 2009).

4.2. Challenges in flower detection and limitations of the proposed workflow

Classifier development for PC image classification is challenging since light conditions vary substantially during the recording times, and some classes are likely to be strongly underrepresented, for example flowers of rare species. Here we propose an efficient labelling phase with analyses of image patches from many images, allowing the representation of also rare species in the labelled sample. We aggregate all graminoids flowers in a single pixel class, since taxonomically and phenotypically close plant species are often too similar to be distinguished reliably, especially in the case of sedge, rush and grass species. We expect that a similar aggregation of different species in one class might also be necessary in future applications in biodiversity-rich grasslands that include many closely-related species.

Not all flower species can be spectrally easily distinguished. For example, young *K. arvensis* head colours are very similar to green-greyish graminoids flowers. On the other hand, *D. glomerata* mature flowerheads may be red- to purple-tinged, very similar to *K. arvensis* flowers. Their classes were described by very similar feature characteristics, which can result in lower classification accuracy (cf. Table 3 and Fig. 4). Following these observations, we expect that with increasing complexity of the study system in terms of plant diversity, maintaining the accuracy of the method will become more challenging, even though texture metrics considerably increased the separability of different flower structures in our study. But in such very rich grasslands, plant biodiversity is then often described with plant functional types, e.g., grasses, forbs and legumes, instead of plant species, and their flower separation using time-lapse images could follow the presented, albeit further developed approach. However, low-diversity grasslands are common both in nature and as a research infrastructure, where sown swards are studied in field or pot experiments, typically to investigate the role of biodiversity and environmental factors on ecosystem functioning (Jentsch et al., 2009; Roscher et al., 2005; Wolf et al., 2017).

Phenological metrics that are automatically extracted through the proposed method can be related to metrics identified with traditional field methods. Field observation of plant flowering phenology is usually repeated at daily to weekly intervals and thus describes plant development at various degrees of detail. Simple metrics such as the first and last day with flowering individuals in the plots, or the day with the highest number of flowering individuals per plot are frequently used in ecological studies dealing with the effect of global change on plant phenology (Cleland et al., 2006; Dorji et al., 2020; Wolf et al., 2017). They are conceptually similar to the metrics proposed here. However, very detailed scales for the description of plant development such as the BBCH scale (Meier et al., 2009) with >50 distinct plant development stages, separately determined for groups of species or single species, can be more difficult to relate to the metrics proposed here. The BBCH scale is not based on flower count or flower cover, but on the description of developmental characteristics, which currently cannot be derived from images through the proposed automated workflow. Thus, future research should attempt to implement pathways to measure more traditional plant phenology metrics that are currently not assessed via automated remote sensing techniques.

When investigating plots with highest average seasonal flower cover of unsown species (i.e., the nine plots shown in Fig. S3), we found these had higher soil cover compared to the overall seasonal average (25% vs. 10%). As already reported in the confusion matrix in Table 3, some soil pixels were misclassified as flowers, and the number of pixels in the flower classes were therefore overestimated. The image dataset we

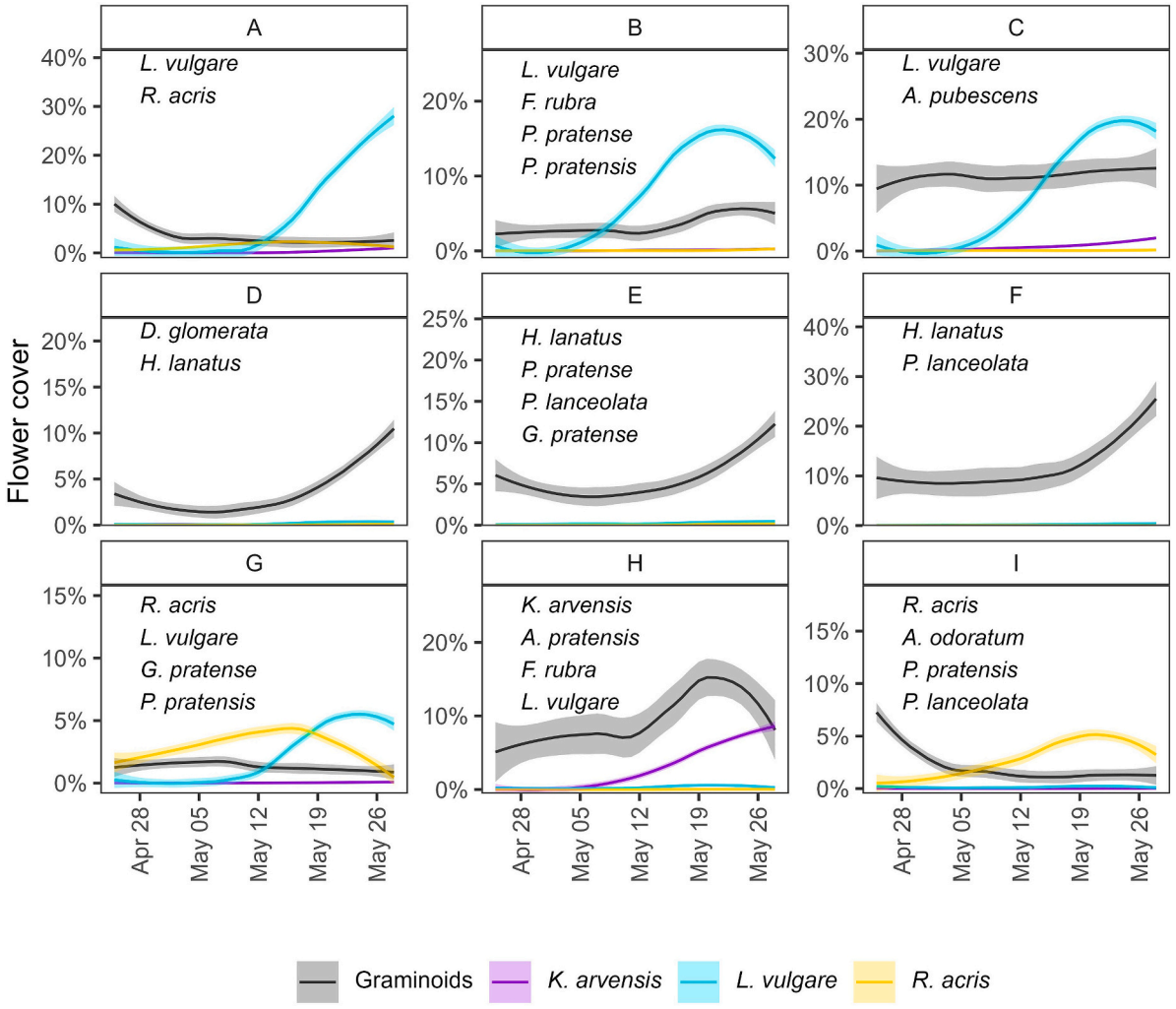

Fig. 6. Time series of flower cover extracted from images acquired in nine experimental plots in spring 2014 (before first mowing). Time series are fitted using Local Polynomial Regression. Confidence intervals (0.95) and sown species names are given in each panel.

analysed in the current study captured an area of 3.5 m^2, which resulted in a pixel size (ground sampling distance) lower than ¼ of the flower size. This ¼ is the suggested minimum ratio between pixel and object dimension for accurate classification (Hengl, 2006). PC images are usually acquired capturing larger areas to describe vegetation patches at landscape scale (Wingate et al., 2015). Thus, for the study design of future studies on the monitoring of reproductive phenology, an optimised field of view should be chosen, considering flowers size and PC resolution. For example, a camera pointing at an angle of 60° from horizontal, the field of view of 49.5° and an optical resolution of 1280 × 1040 pixels should be installed at a maximum height of 1.5 m to classify flowers of size 5 mm and at a maximum height of 3 m to classify flowers of size 10 mm. To ensure the observation of the onset of flowering, it is crucial to establish an experimental setting for phenological observation that spans the entire growing season, whenever possible. However, in this study, this was not always possible as the observation period began after spring weeding of the experiment, preventing the observation of the onset of flowering in some cases.

4.3. Paths for further investigation

New technologies and methodologies are opening new possibilities in grassland phenological studies. Active learning is a promising methodology for balanced sample collection with reduced labelling effort, and was already proposed, for example, for the reduction of sampling effort in forestry inventories (Malek et al., 2019; Persello et al., 2014). The use of active learning could facilitate the labelling phase in biodiversity rich grasslands.

Computer vision (CV) techniques are increasingly being used in ecological studies and have recently been applied to classify grassland images acquired from drones as well as PC images of arctic vegetation (Gallmann et al., 2022; Mann et al., 2022; Wäldchen and Mäder, 2017). However, to our knowledge, no study about grassland flowering phenology using CV techniques to classify PC imagery has been carried out so far. In this study, we applied a pixel approach rather than CV techniques, and quantified flower cover rather than flower count as abundance metric, because CV techniques require a much larger labelling effort and computational capacity, and might not be suited for graminoids flowers and occluded or overlapping flowers. Furthermore, we favoured pixel classification over CV techniques, because the former requires highly specialised knowledge, which is not always available to both biologists and ecologists (Wäldchen and Mäder, 2018). Flower cover is an informative metric also because it can better capture the effect of flowering on greenness, which is widely used to describe vegetation status (Shen et al., 2010). Nevertheless, the increasing availability of pre-trained models that can be fine-tuned, and the recent higher accessibility of these techniques suggest that, in some cases, they could be applied to estimate species flower cover in grassland mixtures (Kirillov et al., 2023; Mann et al., 2022; Wäldchen and Mäder, 2018). Thus, the availability of the proposed workflow opens up new

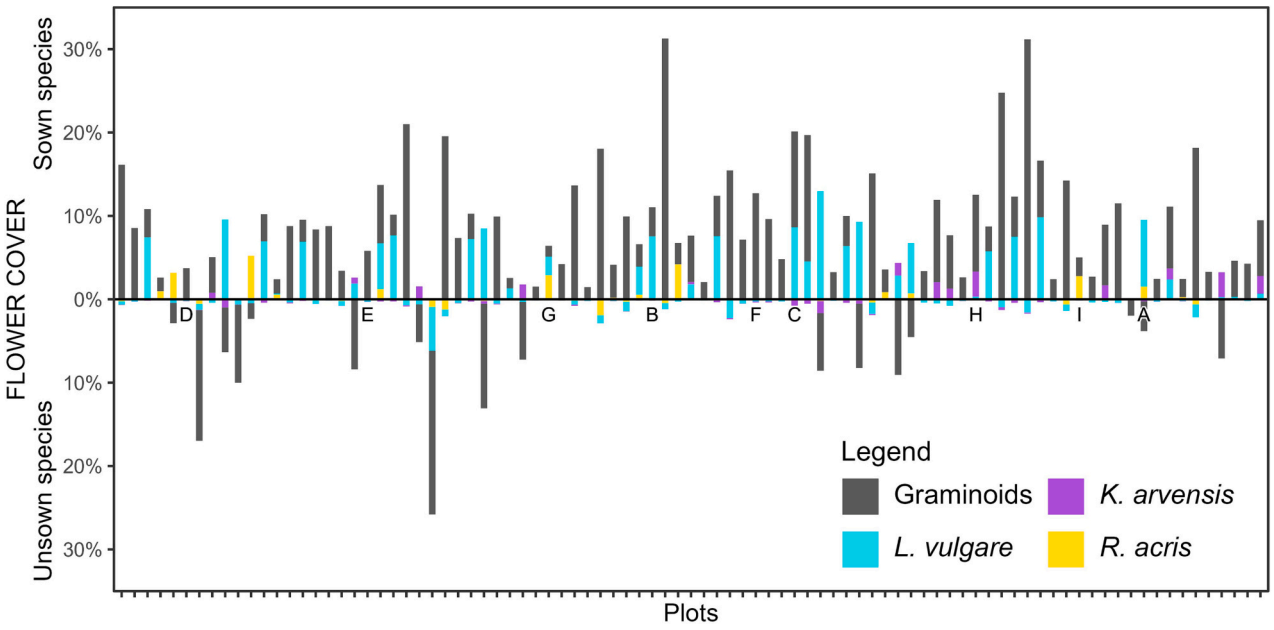

Fig. 7. Seasonal average of flower cover per class, according to the experimental design with different species mixtures. Upper panel: Sown species are displayed. Lower panel: Unsown species are displayed. The plot IDs shown in Fig. 6 (A to I) are reported below the upper panel, indicating zero flower cover. The mean flower cover of unsown species in each plot was 2% with, and 0.6% without the graminoids class.

possibilities in many ecological research fields, including the investigation of species richness effects on individual species flowering phenology.

Author contributions

All authors contributed to the intellectual development and conception of this manuscript; Davide Andreatta performed the analyses and led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All the PhenoCam images, regions of interest, labelled pixels displayed on RGB image details and labelled pixel coordinates have been made available via the ETH Zurich repository (https://doi.org/10.3929/ethz-b-000634004). We provide a tutorial to guide the reader in the application of the proposed methodology. The tutorial is available at the link (https://github.com/andreattad/Flower_covers_phenocams). Three figures were provided in the supplementary materials. Fig. S1 shows accuracy of classifiers based on features calculated at different downscaling factors and window sizes. Fig. S2, Table S1 and Table S2 showing image availability before and after image selection are available in the Supplementary information. Fig. S3 shows the time series of flower cover of the nine experimental plots with the highest seasonal average of flower cover of unsown species.

Acknowledgements

Special thanks to Thomas Lanners who acquired all the images in 2014. We acknowledge all the coordinators, the gardening team from the Jena Experiment, and many field assistants for maintaining the field. The Jena Experiment is funded by the German Research Foundation (DFG, FOR 1451) and the Swiss National Science Foundation (SNF). NB acknowledges funding for the SNF project 310030E-147487; VK acknowledges funding by the AGROSCOPE research programme "Indicate".

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2023.113835.

References

Andrew, M.E., Ustin, S.L., 2008. The role of environmental context in mapping invasive plants with hyperspectral image data. Remote Sens. Environ. 112 (12), 4301–4317. https://doi.org/10.1016/J.RSE.2008.07.016.

Arzani, H., Zohdi, M., Fish, E., Amiri, G.H.Z., Nikkhah, A., Wester, D., 2004. Phenological effects on forage quality of five grass species. J. Range Manag. 57 (6), 624–629. https://doi.org/10.2111/1551-5028(2004)057.

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M.L., Bareth, G., 2015. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. Int. J. Appl. Earth Obs. Geoinf. 39, 79–87. https://doi.org/10.1016/J.JAG.2015.02.012.

Brown, T.B., Hultine, K.R., Steltzer, H., Denny, E.G., Denslow, M.W., Granados, J., Henderson, S., Moore, D., Nagai, S., Sanclements, M., Sánchez-Azofeifa, A., Sonnentag, O., Tazik, D., Richardson, A.D., 2016. Using phenocams to monitor our changing earth: toward a global phenocam network. Front. Ecol. Environ. 14 (2), 84–93. https://doi.org/10.1002/FEE.1222.

Bruzzone, L., Roli, F., Serpico, S.B., 1995. An extension of the Jeffreys-Matusita distance to multiclass cases for feature selection. IEEE Trans. Geosci. Remote Sens. 33 (6), 1318–1321. https://doi.org/10.1109/36.477187.

Byers, D.L., 2017. Studying plant–pollinator interactions in a changing climate: a review of approaches. Appl. Plant Sci. 5 (6), 1700012. https://doi.org/10.3732/apps.1700012.

Chen, B., Jin, Y., Brown, P., 2019. An enhanced bloom index for quantifying floral phenology using multi-scale remote sensing observations. ISPRS J. Photogramm. Remote Sens. 156, 108–120. https://doi.org/10.1016/j.isprsjprs.2019.08.006.

Cleland, E.E., Chiariello, N.R., Loarie, S.R., Mooney, H.A., Field, C.B., 2006. Diverse responses of phenology to global changes in a grassland ecosystem. Proc. Natl. Acad. Sci. 103 (37), 13740–13744. https://doi.org/10.1073/pnas.0600815103.

Collins, C.G., Elmendorf, S.C., Hollister, R.D., Henry, G.H.R., Clark, K., Bjorkman, A.D., Myers-Smith, I.H., Prevéy, J.S., Ashton, I.W., Assmann, J.J., Alatalo, J.M., Carbognani, M., Chisholm, C., Cooper, E.J., Forrester, C., Jónsdóttir, I.S., Klanderud, K., Kopp, C.W., Livensperger, C., Suding, K.N., 2021. Experimental warming differentially affects vegetative and reproductive phenology of tundra plants. Nat. Commun. 12 (1), 1–12. https://doi.org/10.1038/s41467-021-23841-2.

Congalton, R.G., Green, K., 2009. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, 2nd ed. CRC Press/Taylor & Francis.
Dalponte, M., Ørka, H.O., 2021. varSel: Sequential Forward Floating Selection using Jeffries-Matusita Distance. R package version 0.2. https://CRAN.R-project.org/package=varSel.
d'Andrimont, R., Taymans, M., Lemoine, G., Ceglar, A., Yordanov, M., van der Velde, M., 2020. Detecting flowering phenology in oil seed rape parcels with Sentinel-1 and -2 time series. Remote Sens. Environ. 239, 111660. https://doi.org/10.1016/J.RSE.2020.111660.
Dicks, L.v., Breeze, T.D., Ngo, H.T., Senapathi, D., An, J., Aizen, M.A., Basu, P., Buchori, D., Galetto, L., Garibaldi, L.A., Gemmill-Herren, B., Howlett, B.G., Imperatriz-Fonseca, V.L., Johnson, S.D., Kovács-Hostyánszki, A., Kwon, Y.J., Lattorff, H.M.G., Lungharwo, T., Seymour, C.L., Potts, S.G., 2021. A global-scale expert assessment of drivers and risks associated with pollinator decline. Nat. Ecol. Evol. 5 (10), 1453–1461. https://doi.org/10.1038/s41559-021-01534-9.
Dixon, D.J., Callow, J.N., Duncan, J.M.A., Setterfield, S.A., Pauli, N., 2021. Satellite prediction of forest flowering phenology. Remote Sens. Environ. 255, 112197. https://doi.org/10.1016/J.RSE.2020.112197.
D'Odorico, P., Gonsamo, A., Gough, C.M., Bohrer, G., Morison, J., Wilkinson, M., Hanson, P.J., Gianelle, D., Fuentes, J.D., Buchmann, N., 2015. The match and mismatch between photosynthesis and land surface phenology of deciduous forests. Agric. For. Meteorol. 214–215, 25–38. https://doi.org/10.1016/J.AGRFORMET.2015.07.005.
Dorji, T., Hopping, K.A., Meng, F., Wang, S., Jiang, L., Klein, J.A., 2020. Impacts of climate change on flowering phenology and production in alpine plants: the importance of end of flowering. Agric. Ecosyst. Environ. 291, 106795. https://doi.org/10.1016/J.AGEE.2019.106795.
Ebeling, A., Pompe, S., Baade, J., Eisenhauer, N., Hillebrand, H., Proulx, R., Roscher, C., Schmid, B., Wirth, C., Weisser, W.W., 2014. A trait-based experimental approach to understand the mechanisms underlying biodiversity-ecosystem functioning relationships. Basic Appl. Ecol. 15 (3), 229–240. https://doi.org/10.1016/j.baae.2014.02.003.
Estiarte, M., Peñuelas, J., 2015. Alteration of the phenology of leaf senescence and fall in winter deciduous species by climate change: effects on nutrient proficiency. Glob. Chang. Biol. 21, 1005–1017.
Fernández-Pascual, E., Mattana, E., Pritchard, H.W., 2019. Seeds of future past: climate change and the thermal memory of plant reproductive traits. Biol. Rev. 94 (2), 439–456. https://doi.org/10.1111/BRV.12461.
Filippa, G., Cremonese, E., Migliavacca, M., Richardson, A., Galvagno, M., Forkel, M., 2016. Phenopix: Pixel Based Phenology. September.
Freimuth, J., Bossdorf, O., Scheepens, J.F., Willems, F.M., 2022. Climate warming changes synchrony of plants and pollinators. Proc. R. Soc. B 289 (1971), 20212142. https://doi.org/10.1098/RSPB.2021.2142.
Gallmann, J., Schüpbach, B., Jacot, K., Albrecht, M., Winizki, J., Kirchgessner, N., Aasen, H., 2022. Flower mapping in grasslands with drones and deep learning. Front. Plant Sci. 12, 774965. https://doi.org/10.3389/fpls.2021.774965.
Gitelson, A.A., Kaufman, Y.J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetation fraction. Remote Sens. Environ. 80 (1), 76–87. https://doi.org/10.1016/S0034-4257(01)00289-9.
Gonzales, D., Hempel de Ibarra, N., Anderson, K., 2022. Remote sensing of floral resources for pollinators – new horizons from satellites to drones. Front. Ecol. Evol. 10 (414), 869751. https://doi.org/10.3389/FEVO.2022.869751.
Guru, D.S., Sharath, Y.H., Manjunath, S., 2010. Texture features and KNN in classification of flower images. IJCA 1, 21–29. Special Issue on RTIPPR.
Haralick, R.M., Dinstein, I., Shanmugam, K., 1973. Textural features for image classification. IEEE Trans. Syst. Man Cybern. SMC-3 (6), 610–621. https://doi.org/10.1109/TSMC.1973.4309314.
Hengl, T., 2006. Finding the right pixel size. Comput. Geosci. 32 (9), 1283–1298. https://doi.org/10.1016/J.CAGEO.2005.11.008.
Hijmans, R.J., 2022. raster: Geographic Data Analysis and Modeling. https://cran.r-project.org/package=raster.
Ho, T.K., 1995. Random decision forests. In: Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, 1, pp. 278–282. https://doi.org/10.1109/ICDAR.1995.598994.
Hyndman, R.J., Khandakar, Y., 2008. Automatic time series forecasting: the forecast package for R. J. Stat. Softw. 27 (3) https://doi.org/10.18637/jss.v027.i03.
Inouye, D.W., 2020. Effects of climate change on alpine plants and their pollinators. Ann. N. Y. Acad. Sci. 1469 (1), 26–37. https://doi.org/10.1111/NYAS.14104.
Jentsch, A., Kreyling, J., Boettcher-Treschkow, J., Beierkuhlein, C., 2009. Beyond gradual warming: extreme weather events alter flower phenology of European grassland and heath species. Glob. Chang. Biol. 15 (4), 837–849. https://doi.org/10.1111/j.1365-2486.2008.01690.x.
Jung, C., Müller, A.E., 2009. Flowering time control and applications in plant breeding. Trends Plant Sci. 14 (10), 563–573. https://doi.org/10.1016/J.TPLANTS.2009.07.005.
Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., Dollár, P., Girshick, R., 2023. Segment Anything, ArXiv. http://arxiv.org/abs/2304.02643.
Landmann, T., Piiroinen, R., Makori, D.M., Abdel-Rahman, E.M., Makau, S., Pellikka, P., Raina, S.K., 2015. Application of hyperspectral remote sensing for flower mapping in African savannas. Remote Sens. Environ. 166, 50–60. https://doi.org/10.1016/J.RSE.2015.06.006.
Liaw, A., Wiener, M., 2002. Classification and regression by random forest. R News 2 (3), 1822.
Liu, Y., Wang, J., Dong, J., Wang, S., Ye, H., 2020. Variations of vegetation phenology extracted from remote sensing data over the Tibetan plateau hinterland during 2000–2014. J. Meteorol. Res. 34 (4), 786–797. https://doi.org/10.1007/s13351-020-9211-x.
Louhaichi, M., Borman, M.M., Johnson, D.E., 2001. Spatially Located Platform and Aerial Photography for Documentation of Grazing Impacts on Wheat. Geocarto International 16 (1), 65–70. https://doi.org/10.1080/10106040108542184.
Lussem, U., Bolten, A., Gnyp, M.L., Jasper, J., Bareth, G., 2018. Evaluation of RGB-Based Vegetation Indices from UAV Imagery to Estimate Forage Yield in Grassland. https://doi.org/10.5194/isprs-archives-XLII-3-1215-2018.
Malek, S., Miglietta, F., Gobakken, T., Næsset, E., Gianelle, D., Dalponte, M., 2019. Optimizing fiesld data collection for individual tree attribute predictions using active learning methods. Remote Sens. 11 (8), 949. https://doi.org/10.3390/rs11080949.
Mann, H.M.R., Iosifidis, A., Jepsen, J.U., Welker, J.M., Loonen, M.J.J.E., Høye, T.T., 2022. Automatic flower detection and phenology monitoring using time-lapse cameras and deep learning. Remote Sens. Ecol. Conserv. 8 (6), 765–777. https://doi.org/10.1002/rse2.275.
Meier, U., Bleiholder, H., Buhr, L., Feller, C., Hack, H., Heß, M., Lancashire, P., Schnock, U., Stauß, R., van den Boom, T., Weber, E., Zwerger, P., 2009. The BBCH system to coding the phenological growth stages of plants-history and publications. J. Kult. 61 (2).
Menzel, A., Sparks, T.H., Estrella, N., Koch, E., Aaasa, A., Ahas, R., Alm-Kübler, K., Bissolli, P., Braslavská, O., Briede, A., Chmielewski, F.M., Crepinsek, Z., Curnel, Y., Dahl, Å., Defila, C., Donnelly, A., Filella, Y., Jatczak, K., Måge, F., Zust, A., 2006. European phenological response to climate change matches the warming pattern. Glob. Chang. Biol. 12 (10), 1969–1976. https://doi.org/10.1111/J.1365-2486.2006.01193.X.
Persello, C., Boularias, A., Dalponte, M., Gobakken, T., Naesset, E., Scholkopf, B., 2014. Cost-sensitive active learning with lookahead: optimizing field surveys for remote sensing data classification. IEEE Trans. Geosci. Remote Sens. 52 (10), 6652–6664. https://doi.org/10.1109/TGRS.2014.2300189.
Piao, S., Ciais, P., Friedlingstein, P., Peylin, P., Reichstein, M., Luyssaert, S., Margolis, H., Fang, J., Barr, A., Chen, A., Grelle, A., Hollinger, D.Y., Laurila, T., Lindroth, A., Richardson, A.D., Vesala, T., 2008. Net carbon dioxide losses of northern ecosystems in response to autumn warming. Nature 451 (7174), 49–52. https://doi.org/10.1038/nature06444.
Pudil, P., Novovičová, J., Kittler, J., 1994. Floating search methods in feature selection. Pattern Recogn. Lett. 15 (11), 1119–1125. https://doi.org/10.1016/0167-8655(94)90127-9.
R Core Team, 2023. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/.
Richards, J.A., Jia, Xiuping, 2006. Remote Sensing Digital Image Analysis : An Introduction. Springer.
Richardson, A.D., Black, T.A., Ciais, P., Delbart, N., Friedl, M.A., Gobron, N., Hollinger, D.Y., Kutsch, W.L., Longdoz, B., Luyssaert, S., Migliavacca, M., Montagnani, L., Munger, J.W., Moors, E., Piao, S., Rebmann, C., Reichstein, M., Saigusa, N., Tomelleri, E., Varlagin, A., 2010. Influence of spring and autumn phenological transitions on forest ecosystem productivity. Philos. Trans. R. Soc. B 365 (1555), 3227–3246. https://doi.org/10.1098/RSTB.2010.0102.
Richardson, A.D., Anderson, R.S., Arain, M.A., Barr, A.G., Bohrer, G., Chen, G., Chen, J. M., Ciais, P., Davis, K.J., Desai, A.R., Dietze, M.C., Dragoni, D., Garrity, S.R., Gough, C.M., Grant, R., Hollinger, D.Y., Margolis, H.A., McCaughey, H., Migliavacca, M., Xue, Y., 2012. Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis. Glob. Chang. Biol. 18 (2), 566–584. https://doi.org/10.1111/j.1365-2486.2011.02562.x.
Richter, F., Jan, P., el Benni, N., Lüscher, A., Buchmann, N., Klaus, V.H., 2021. A guide to assess and value ecosystem services of grasslands. Ecosyst. Serv. 52, 101376. https://doi.org/10.1016/J.ECOSER.2021.101376.
Roscher, C., Temperton, V.M., Scherer-Lorenzen, M., Schmitz, M., Schumacher, J., Schmid, B., Buchmann, N., Weisser, W.W., Schulze, E.D., 2005. Overyielding in experimental grassland communities - irrespective of species pool or spatial scale. Ecol. Lett. 8 (4), 419–429. https://doi.org/10.1111/J.1461-0248.2005.00736.X.
Schwartz, M., 2013. *Phenology: An Integrative Environmental Science* (M. D. Schwartz, Ed.). Springer, Netherlands. https://doi.org/10.1007/978-94-007-6925-0.
Shen, M., Chen, J., Zhu, X., Tang, Y., Chen, X., 2010. Do flowers affect biomass estimate accuracy from NDVI and EVI? Int. J. Remote Sens. 31 (8), 2139–2149. https://doi.org/10.1080/01431160903578812.
Shen, M., Tang, Y., Chen, J., Zhu, X., Zheng, Y., 2011. Influences of temperature and precipitation before the growing season on spring phenology in grasslands of the central and eastern Qinghai-Tibetan plateau. Agric. For. Meteorol. 151 (12), 1711–1722. https://doi.org/10.1016/J.AGRFORMET.2011.07.003.
Szigeti, V., Kőrösi, Ádam, Haros, A., Nagy, J., Kis, J., 2016. Measuring floral resource availability for insect pollinators in temperate grasslands - a review. Ecol. Entomol. 41 (3), 231–240. https://doi.org/10.1111/een.12298.
Thackeray, S.J., Henrys, P.A., Hemming, D., Bell, J.R., Botham, M.S., Burthe, S., Helaouet, P., Johns, D.G., Jones, I.D., Leech, D.I., Mackay, E.B., Massimino, D., Atkinson, S., Bacon, P.J., Brereton, T.M., Carvalho, L., Clutton-Brock, T.H., Duck, C., Edwards, M., Wanless, S., 2016. Phenological sensitivity to climate across taxa and trophic levels. Nature 535 (7611), 241–245. https://doi.org/10.1038/nature18608.
Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8 (2), 127–150. https://doi.org/10.1016/0034-4257(79)90013-0.
Vasiliev, D., Greenwood, S., 2021. The role of climate change in pollinator decline across the Northern Hemisphere is underestimated. Sci. Total Environ. 775, 145788. https://doi.org/10.1016/j.scitotenv.2021.145788.

Vázquez, D.P., Vitale, N., Dorado, J., Amico, G., Stevani, E.L., Leal, R., 2023. Phenological mismatches and the demography of solitary bees. Proc. R. Soc. B 290 (1990). https://doi.org/10.1098/RSPB.2022.1847, 20221847.

Wäldchen, J., Mäder, P., 2017. Plant species identification using computer vision techniques: a systematic literature review. Arch. Comput. Methods Eng. 25 (2), 507–543. https://doi.org/10.1007/S11831-016-9206-Z.

Wäldchen, J., Mäder, P., 2018. Machine learning for image based species identification. Methods Ecol. Evol. 9 (11), 2216–2225. https://doi.org/10.1111/2041-210X.13075.

Wingate, L., Cremonese, E., Migliavacca, M., Brown, T., D'Odorico, P., Peichl, M., Gielen, B., Lukas, H., 2015. PROTOCOL Phenocamera: automated phenology monitoring. ICOS, p. 17.

Wolf, A.A., Zavaleta, E.S., Selmants, P.C., 2017. Flowering phenology shifts in response to biodiversity loss. Proc. Natl. Acad. Sci. U. S. A. 114 (13), 3463–3468. https://doi.org/10.1073/pnas.1608357114.

Wouters, N., de Ketelaere, B., de Baerdemaeker, J., Saeys, W., 2013. Hyperspectral waveband selection for automatic detection of floral pear buds. Precis. Agric. 14 (1), 86–98. https://doi.org/10.1007/S11119-012-9279-0/TABLES/3.

Zhao, Y., 2021. The segmentation of plants on RGB images with index based color analysis. In: 2021 5th International Conference on Robotics and Automation Sciences, ICRAS 2021, pp. 221–225. https://doi.org/10.1109/ICRAS52289.2021.9476331.

Zvoleff, A., 2020. Package 'glcm.' Calculate Textures from Grey-Level Co-Occurence Matrices (GLCMs). *Available Online: Https://CRAN.R-Project.Org/Package=Glcm*.