Int J Appl Earth Obs Geoinformation 84 (2020) 101974

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

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

Land surface phenology from VEGETATION and PROBA-V data. Assessment over deciduous forests

Kevin Bórnez[a,b,*], Adrià Descals[a,b], Aleixandre Verger[a,b], Josep Peñuelas[a,b]

[a] CREAF, Cerdanyola del, Vallès 08193, Catalonia, Spain
[b] CSIC, Global Ecology Unit, Cerdanyola del Vallès 08193, Catalonia, Spain

ARTICLE INFO

Keywords:
Land surface phenology
SPOT-VEGETATION
PROBA-V
Leaf area index
Ground observations

ABSTRACT

Land surface phenology has been widely retrieved although no consensus exists on the optimal satellite dataset and the method to extract phenology metrics. This study is the first comprehensive comparison of vegetation variables and methods to retrieve land surface phenology for 1999–2017 time series of Copernicus Global Land products derived from SPOT-VEGETATION and PROBA-V data. We investigated the sensitivity of phenology to (I) the input vegetation variable: normalized difference vegetation index (NDVI), leaf area index (LAI), fraction of absorbed photosynthetically active radiation (FAPAR), and fraction of vegetation cover (FCOVER); (II) the smoothing and gap filling method for deriving seasonal trajectories; and (III) the method to extract phenological metrics: thresholds based on a percentile of the annual amplitude of the vegetation variable, autoregressive moving averages, logistic function fitting, and first derivative methods. We validated the derived satellite phenological metrics (start of the season (SoS) and end of the season (EoS)) using available ground observations of *Betula pendula*, *B. alleghaniensis, Acer rubrum*, *Fagus grandifolia*, and *Quercus rubra* in Europe (Pan-European PEP725 network) and the USA (National Phenology Network, USA-NPN). The threshold-based method applied to the smoothed and gap-filled LAI V2 time series agreed best with the ground phenology, with root mean square errors of ˜10 d and ˜25 d for the timing of SoS and EoS respectively. This research is expected to contribute for the operational retrieval of land surface phenology within the Copernicus Global Land Service.

1. Introduction

Phenology is the study of the timing of recurrent biological and seasonal events and their biotic and abiotic factors (Beaumont et al., 2015). Studies of plant phenology focus on how these events and factors are influenced by seasonal and interannual variations in climate and how they modulate abundance and diversity (Beaumont et al., 2015). Phenology is, moreover, key to control physicochemical and biological processes, especially albedo, surface roughness, canopy conductance and fluxes of carbon, water and energy (Peñuelas et al., 2009; Richardson et al., 2013). Phenological metrics are thus relevant parameters for modeling land surface processes and the global carbon cycle (Wu et al., 2014).

Phenological metrics are estimated based on ground observations and data derived from satellites. Ground observations provide accurate timing of vegetation phenophases but cannot cover continuously large-scale areas (Garrity et al., 2011; Yu et al., 2017). Satellite sensors with moderate spatial resolutions, including AVHRR, MODIS, MERIS, SPOT-VEGETATION and PROBA-V, provide long-term time series of daily observations that allow improving the characterization of land surface phenology on a global scale (Verger et al., 2016; Atkinson et al., 2012; Zhang et al., 2004). However, the noise in the data and missing observations mainly due to cloud contamination may induce significant uncertainties in the estimation of phenological metrics (Kandasamy et al., 2013; Verger et al., 2013). The literature shows a broad variety of time-series processing methods designed to reconstruct gap-filled vegetation seasonal trajectories from noisy satellites signals. This includes the best index slope method (Viovy et al., 1992), mean filters (Reed et al., 1994), moving-window filters (Sweets et al. 1999), asymmetric Gaussian functions (Jönsson and Eklundh, 2002), Savitzky–Golay filters (Chen et al., 2004) or the Whittaker smoother (Eilers, 2003). However, no single method always performs better than others for smoothing vegetation time series (Cai et al., 2017) and their performance vary spatially and temporally with land surface conditions and cloud influence (Atkinson et al., 2012; Kandasamy and Fernandes, 2015).

A broad variety of statistical methods have been designed to extract phenological metrics from satellite time series. Metrics typically include the start of the season (SoS), the end of the season (EoS), the timing of

* Corresponding author at: CREAF, Cerdanyola del Vallès, 08193, Catalonia, Spain.
E-mail address: k.bornez@creaf.uab.cat (K. Bórnez).

https://doi.org/10.1016/j.jag.2019.101974
Received 9 February 2019; Received in revised form 22 August 2019; Accepted 9 September 2019
Available online 03 October 2019

maximum growth and the length of the growing season (LoS) (Reed et al., 1994; Zhang et al., 2004). de Beurs and Henebry (2010) provided a comprehensive review of the exiting phenology retrieval approaches that can be classified in four main categories: thresholds and percentile based methods (Atzberger and Eilers, 2011; Verger et al., 2016), moving averages (Reed et al., 1994), first derivatives (White et al., 2009) and fitted models (de Beurs and Henebry, 2005). White et al. (2009) compared ten different phenology retrieval methods applied to AVHRR NDVI in North America and found large discrepancies of up to two months in the detection of the SoS.

In addition to the sensitivity to the smoothing and phenological extraction algorithm, the derived phenological metrics are also dependent on the sensor, spatial and temporal resolution, processing chain, and satellite data set. The satellite-derived spectral vegetation indices (e.g. the Normalized Difference Vegetation Index (NDVI)) vary in their strength of phenological prediction across sites and plant functional types (Wu et al., 2014). Unlike previous studies based on vegetation indices, the present study aimed to characterize the phenology not only with NDVI but also with biophysical variables: the leaf area index (LAI), the fraction of absorbed photosynthetically active radiation (FAPAR), the fraction of vegetation cover (FCOVER). We used NDVI version V2.1 (Toté et al., 2017), LAI, FAPAR and FCOVER V1 (Baret et al., 2013) and V2 (Verger et al., 2014) time series derived within the Copernicus Global Land Service (CGLS) from SPOT-VEGETATION and PROBA-V data. Verger et al. (2017) showed that the phenology derived from the interannual climatology of LAI V1 improved other existing products including MODIS-EVI when compared to ground observations for the average date of the SoS and EoS. However, their study was limited to the baseline LAI phenology as derived from a single extraction method. This paper is a continuation of the previous paper by Verger et al. (2017) and we address now the interannual variation of the yearly phenology, the impact of the input vegetation variable and the phenological extraction method. Further, we incorporate LAI, FAPAR and FCOVER V2 that improved continuity (no missing data in V2) and smoothness as compared to V1.

Our study had two main objectives: to select the best biophysical variable or vegetation index for estimating phenological metrics on a global scale within the portfolio of the CGLS vegetation products (NDVI, LAI, FCOVER or FAPAR) and to define the method that best matched the ground data.

2. Materials and methods

2.1. Phenological ground observations

Ground-based phenological data from PEP725 and USA-NPN were examined, focusing on the dates of leaf out and leaf senescence for *Betula* (birch) in Europe (Fig. 1a) and the USA and for *Quercus* (oak), *Fagus* (beech) and *Acer* (maple) in the USA (Fig. 1b). These genera were chosen because they are present in both Europe and the USA and have large numbers of records in the combined data set.

The PEP725 Pan-European Phenology database (Templ et al., 2018) (www.pep725.eu) has complete records from 1990 to the present. The phenophases defined in PEP725 are based on BBCH (Biologische Bundesanstalt, Bundessortenamt and Chemical industry) code (Meier et al., 2009). We used the phenophases corresponding to the first visible leaves (BBCH 11) as the reference for the timing of SoS and the date corresponding to 50% of leaves with autumn coloration (BBCH 94) for the timing of EoS.

The USA National Phenology Network was established in 2007 to collect, store and share historical and contemporary phenological data on a North American scale (Schwartz et al., 2012). The data are freely available at https://www.usanpn.org/. This network provides measurements of several phenophases. We used the phenophases “leaves” which corresponds to first visible leaves and “increasing leaf size” for SoS and “colored leaves” for EoS.

We discarded ground sites with less than four yearly measurements to obtain consistent data records over the time series. We used ground-site located pixels. The spatial heterogeneity hamper comparing ground-based phenology for individual plants with satellite phenology at a resolution of 1 km. We filtered the ground sites located in agricultural or urban areas using high-resolution images from Google Earth (https://earth.google.com/) and the ESA Land Cover Map (CCI-LC) (http://maps.elie.ucl.ac.be/CCI/viewer/index.php). In Europe, we used only the points with forest coverages > 5% using a tree cover map for European forests (Brus et al., 2012).

2.2. Satellite time series

The time series of satellite imagery used for estimating the phenological metrics were from the SPOT-VEGETATION (1999–2013) and PROBA-V satellites (2014–2017) with spatial resolutions of 1 km and temporal frequencies of 10 d. In particular, we used LAI, FAPAR, FCOVER V1 (Baret et al., 2013) and V2 (Verger et al., 2014) and NDVI V2.1 (Toté et al., 2017) products generated within the CGLS (https://land.copernicus.eu/global/themes/vegetation) (Table1).

Time series of LAI, FAPAR and FCOVER V1 and NDVI V2.1 contained frequent disturbances caused by residual cloud contamination, atmospheric variability, snow and bi-directional effects (Fig. 2). These time series required the application of smoothing and gap filling techniques to generate consistent and gap filled seasonal trajectories (section 2.3) before the extraction of phenological metrics. The LAI and FCOVER V2 products improved over the V1 products in terms of temporal consistency and continuity (Fig. 2). The V2 algorithm included multi-step data filtering, smoothing and gap-filling techniques that rendered the products suitable for phenological estimation without additional pre-processing. Filtering of outliers is based on an upper envelope approach, the pixel climatology (interannual mean) is used to fill missing data and a Savitzky-Golay filter is used for the smoothing (Verger et al., 2014).

2.3. Smoothing methods

We tested several smoothing methods for reducing noise and reconstructing gap filled seasonal trajectories from CGLS time series (Eerens and Haesen, 2015):

- *WHITTAKER smoother* (Atzberger and Eilers, 2011): It minimizes a cost function describing the balance between fidelity (quadratic difference between estimates and actual observations) and roughness (quadratic difference between successive estimates).
- *BISE (Best Index Slope Extraction) (*Viovy et al., 1992*):* It retains the good observations in a local window and replaces missing or eliminated suspect values by linear interpolation.
- *MEAN*: A linear interpolation is first applied to fill missing data. A running mean filter with a sliding window of 50 d length is then applied.
- *SWETS method (*Swets et al., 1999*)*: A linear interpolation is first applied to fill missing data. A weighted linear regression over a local window is then applied.

2.4. Methods for extracting phenological metrics

We tested four state of the art methods to extract phenological metrics from CGLS time series (Fig. 3):

- *Thresholds based on a pixel percentile value:* SoS is defined as the day of the year (DoY) when a vegetation variable exceeds a particular threshold. EoS is defined as the DoY when an index remains below a particular threshold. We established dynamic thresholds per pixel based on a percentile of the annual amplitude of the vegetation variable (Verger et al., 2016). The selected percentiles were

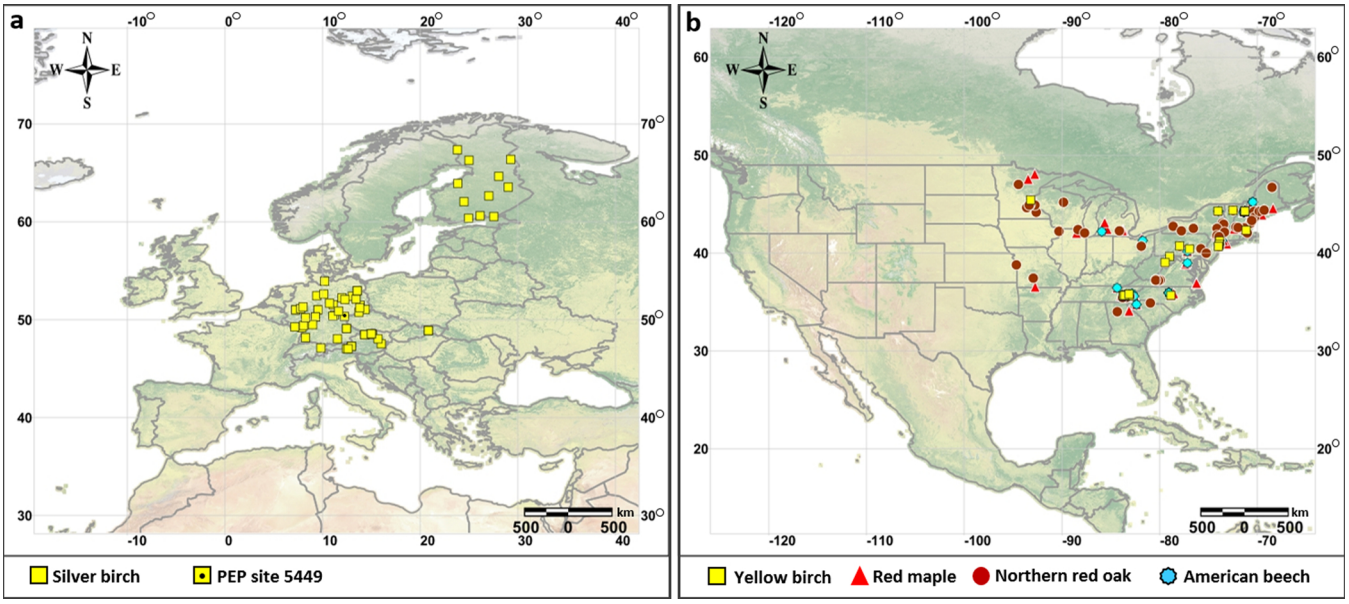

Fig. 1. Location of selected phenological ground observations at (a) PEP725 sites for Silver birch (*Betula pendula*) and (b) USA-NPN sites for Red maple (*Acer rubrum*), Yellow birch (*Betula alleghaniensis*), American beech (*Fagus grandifolia*) and Northern red oak (*Quercus rubra*).

Table 1
Algorithm principles of NDVI V2.1, and LAI, FAPAR, FCOVER V1 and V2.

	NDVI Version 2.1	LAI, FAPAR, FCOVER Version 1	LAI, FAPAR, FCOVER Version 2
Inputs	Top of the canopy (TOC) reflectances in the red and near infrared (NIR) spectral bands	Nadir normalized TOC reflectances in the red, NIR and short-wave infrared (SWIR) spectral bands, and cosine of the sun zenith angle at 10:00 local time	TOC reflectances in the red, NIR and SWIR spectral bands, and cosine of the 3 angles of sun and view directions
Temporal composition	10 d compositing period. The maximum NDVI value in the composition window is retained. Starting date of composition:1st, 11th and 21th day of the month	30 d compositing period with Gaussian weighting (minimum of two valid observations) Nominal dates: 3rd, 13th and 21-24th day of each month	Adaptive compositing within 15 and 60d semi-periods defined by the availability of 6 valid observations at each side of the date being processed Nominal dates: 10th, 20th and last day of the month
Temporal smoothing and gap filling	Not applied	Not applied	Multi-step filtering, temporal smoothing and gap-filling

determined based on the comparison with available ground measurements. We tested the 20th, 30th, 40th and 50th percentiles of the annual amplitude for SoS, and the 30th, 40th, 50th and 60th percentiles for EoS. This method is also the basis of SPIRITS phenological approach.

- *Autoregressive moving average:* A moving average is first computed at a randomly chosen time lag (Ivits et al., 2009). We tested time lags from 50 to 150 d and selected a time lag of 100 d based on the comparison with ground measurements. SoS and EoS are then defined as the DOY when the moving average curves cross the original curve of the vegetation variable.
- *First derivative:* SoS is defined as the DoY of the maximum increase (maximum first derivative) in the curve (Tateishi and Ebata, 2004). EoS is defined as the DoY of the maximum decrease in the curve.

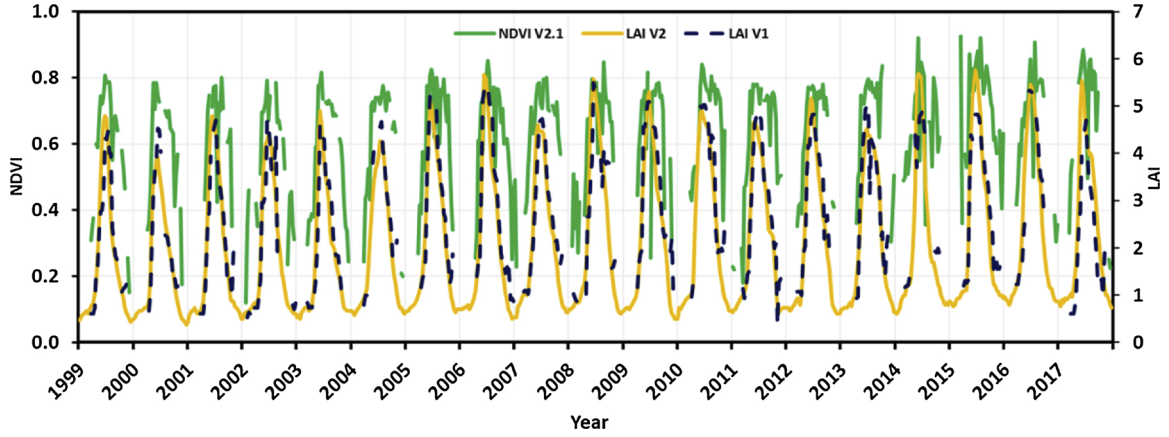

Fig. 2. LAI (V1 and V2) and NDVI (V2.1) time series for the PEP725 site 5449 (50°42′20.49″N, 13°46′59.55″E) representative of birch forest in Europe.

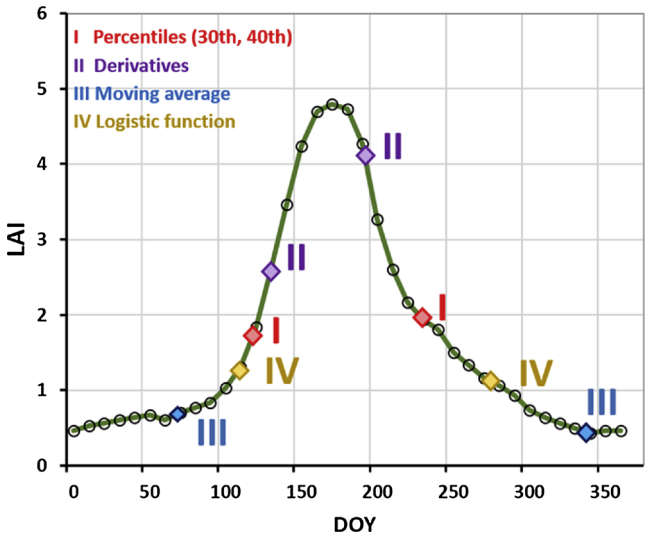

Fig. 3. Schematic representation of SoS (on the left of the peak) and EoS (on the right of the peak) retrieved with the four methods for the PEP725 site 4959 (50°42′20.49″N, 13°46′59.55″E) for 2011. The black circles correspond to the original LAI data at a 10-d frequency, and the green line corresponds to the data interpolated at daily steps, which is used for phenological estimation.

Table 2
Statistics of comparisons between the derived phenology metrics from reconstructed LAI V1 time series using the different smoothing methods and PEP725 ground measurements (*Betula pendula*) for the SoS and EoS. The SPIRITS percentile method was used to extract the phenology using 30% of LAI amplitude for the SoS and 40% for the EoS. * Significant correlations with $p < 0.05$ (**$p < 0.001$).

Metric	Method	RMSE (d)	Bias (d)	Abs. Bias (d)	*R*	Slope
SoS	WHITTAKER	23.80	−8.16	16.76	0.41	0.76
"visible	BISE	27.22	−7.52	19.84	0.43*	0.89
leaves"	MEAN	23.98	−5.84	18.43	0.39	0.74
(n = 359)	**SWETS**	**16.44**	**−1.38**	**17.39**	**0.49****	**0.67**
EoS	WHITTAKER	44.56	13.51	31.65	−0.19	−0.12
"colored	BISE	49.45	19.50	37.25	−0.03	−0.10
leaves"	MEAN	44.10	13.18	30.54	−0.08	−0.01
(n = 359)	**SWETS**	**34.28**	**12.01**	**28.96**	**−0.15**	**−0.19**

- *Logistic function:* SoS is defined as the DoY of the first local maximum rate of change in the curvature of a logistic function fitted to the time series (Zhang et al., 2003). EoS is defined as the DoY of the first local minimum rate of change in the curvature.

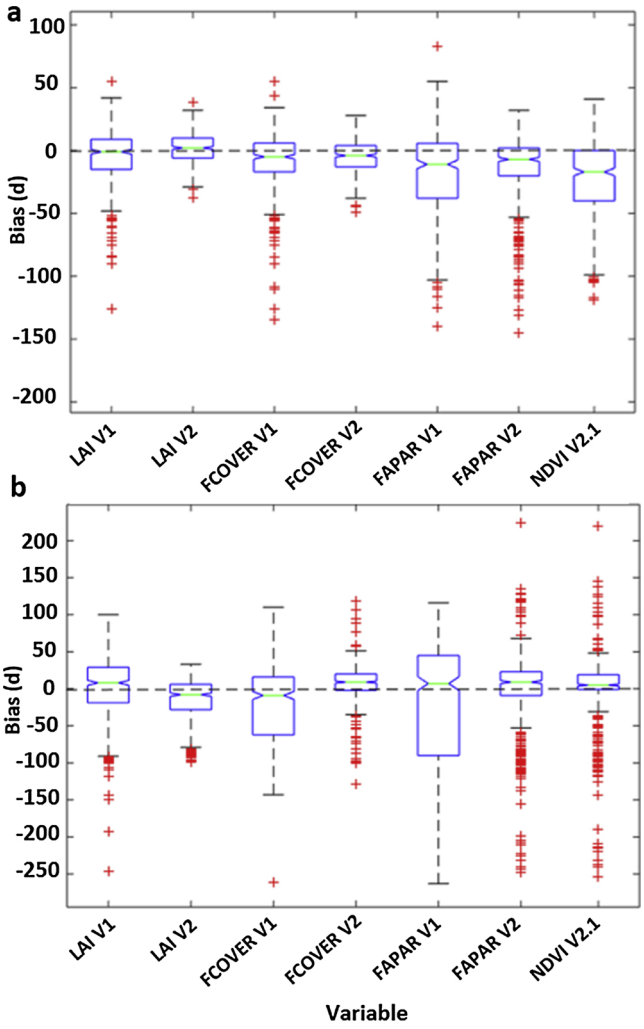

Fig. 4. Boxplots of the bias errors for (a) SoS and (b) EoS estimated from the LAI, FCOVER, FAPAR V1 and V2 and NDVI V2.1 time series minus the PEP725 ground measurements. An elongated boxplot indicates a greater dispersion of the average bias.

2.5. Methodological approach

The several satellite-derived vegetation variables, smoothing methods and phenological extraction approaches lead to a large

Table 3
Statistics of comparisons between the derived phenologies from LAI, FCOVER, FAPAR V1 and V2 and NDVI V2.1 time series and PEP725 ground measurements (*Betula pendula*) for the start of the season (SoS) and end of the season (EoS). The SPIRITS percentile method was used to extract the phenology with specific thresholds per vegetation variable and phenological metric. *mark indicates significant correlations with $p < 0.05$ (** indicates $p < 0.001$).

Metric	Index	Version	Definition	RMSE (d)	BIAS (d)	Abs. BIAS (d)	*R*	Slope
SoS	LAI	V1	30%	16.44	−1.38	17.39	0.49**	0.67
(n = 359)		**V2**	**30%**	**12.49**	**1.65**	**10.22**	**0.62****	**0.78**
	FCOVER	V1	40%	17.25	−4.44	23.54	0.52**	0.74
		V2	40%	13.95	−4.46	11.21	0.54**	0.95
	FAPAR	V1	50%	31.40	−18.35	31.23	0.48**	0.85
		V2	50%	23.01	−14.26	13.94	0.57**	0.88
	NDVI	V2.1	50%	20.70	−13.18	15.38	0.25	0.82
EoS	LAI	V1	40%	34.28	12.01	28.96	−0.15	−0.19
(n = 359)		V2	40%	32.72	−6.18	21.12	0.05	0.11
	FCOVER	V1	40%	30.69	−14.33	42.25	−0.13	−0.20
		V2	**40%**	**25.35**	**6.42**	**20.34**	**0.26**	**0.38**
	FAPAR	V1	50%	44.06	18.75	45.87	−0.23	−0.65
		V2	50%	39.99	18.61	37.97	−0.04	−0.51
	NDVI	V2.1	50%	48.35	14.75	34.58	0.02	−0.25

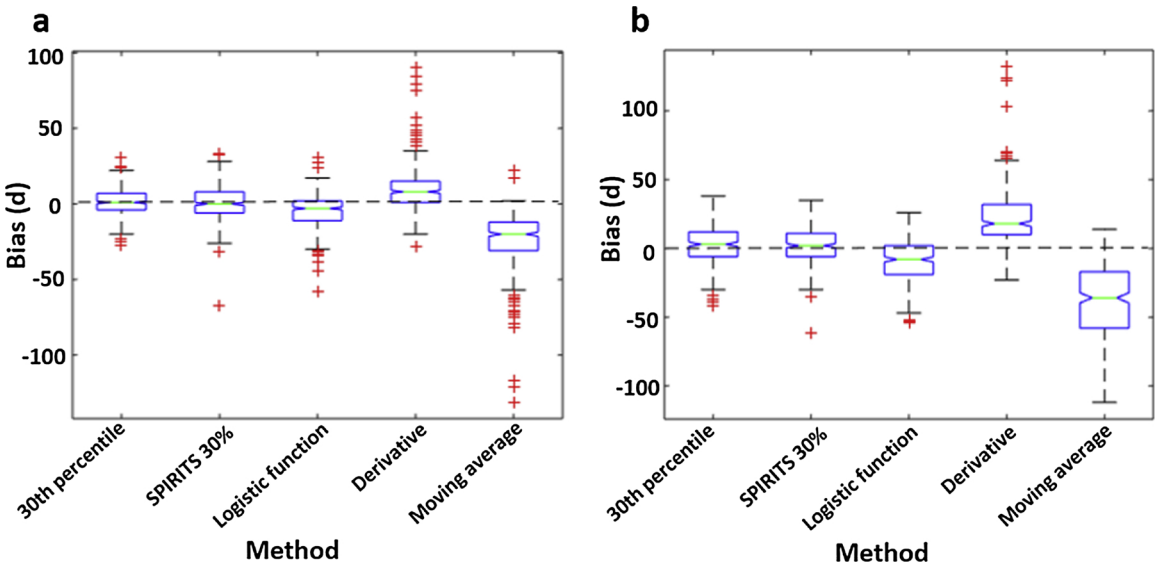

Fig. 5. Boxplots of the bias error for the SoS estimated from LAI V2 minus the ground measurements at the USA-NPN (a) and PEP725 (b) sites. An elongated boxplot indicates a greater dispersion of the average bias in each method.

Table 4
Statistics of comparisons between LAI V2 derived phenology for the start of the season (SoS) and end of the season (EoS), and the ground measurements for the various methods (percentiles, logistic function, derivative and moving average). *mark indicates significant correlations with $p < 0.05$ (** indicates $p < 0.001$).

Validation source	Metric	Method	RMSE (d)	Bias (d)	Abs. Bias (d)	*R*	Slope
USA-NPN (n = 462)	SoS "Leaves"	**30th percentile**	**9.19**	**1.59**	**6.77**	**0.81****	**0.73**
		SPIRITS 30%	13.75	0.58	7.05	0.60**	0.58
		Logistic	14.14	−5.25	8.66	0.63**	0.60
		Derivative	16.82	9.63	12.01	0.68**	0.82
		Moving average	30.06	−24.33	23.95	0.59*	0.82
USA-NPN (n = 158)	SoS "Increaising leaf size"	**30th percentile**	**8.87**	**−2.20**	**6.11**	**0.83****	**0.67**
		SPIRITS 30%	10.40	−3.46	6.18	0.78**	0.67
		Logistic	12.81	−7.24	6.86	0.74**	0.63
		Derivative	10.04	4.22	7.73	0.83**	0.84
		Moving average	34.49	−28.07	27.86	0.49*	0.70
PEP725 (n = 359)	SoS "visible leaves"	**30th percentile**	**11.50**	**1.69**	**9.84**	**0.60****	**0.90**
		SPIRITS 30%	12.49	1.65	10.22	0.62**	0.78
		Logistic	17.96	−9.22	14.21	0.53*	0.71
		Derivative	28.31	19.19	20.42	0.49*	0.88
		Moving average	56.24	−51.35	54.27	0.50*	1.14
USA-NPN (n = 241)	EoS "Colored leaves"	**40th percentile**	**25.61**	**6.39**	**17.60**	**0.14**	**0.10**
		SPIRITS 40%	30.79	−3.69	20.17	0.06	0.07
		Logistic	30.70	21.80	23.66	0.10	0.15
		Derivative	40.40	−10.11	26.70	0.02	0.01
		Moving average	58.75	50.35	44.26	0.13	0.01
PEP725 (n = 359)	EoS "Colored leaves"	**40th percentile**	**27.69**	**−5.15**	**18.89**	**0.11**	**0.15**
		SPIRITS 40%	32.72	−6.18	21.12	0.05	0.11
		Logistic	30.15	11.91	24.29	0.03	0.28
		Derivative	64.93	−43.96	48.49	0.00	−0.11
		Moving average	54.95	37.81	52.55	0.01	0.21

number of combinations. We sequentially investigated the impact of smoothing, variable and extraction method based on an initial set of modalities. The initial modalities were defined *a posteriori* based on the analysis of all the combinations:

1 Sensitivity analysis of the smoothing method: We used LAI V1 as input dataset and the percentile phenology method.
2 Sensitivity analysis of the vegetation variable: We used the SWETS smoothing method and the percentile phenology method.
3 Sensitivity analysis of the method to extract phenological metrics: We used the LAI V2. Note that in this case the application of a smoothing method is not required because LAI V2 is already smoothed and gap-filled.

The analysis 1 and 2 were carried out in Europe and for the validation we used ground measurements of *Betula Pendula*, which showed a greater latitudinal distribution. The analysis 3 was performed at the global scale. For the sensitivity analysis 1 and 2, we used the Software for the Processing and Interpretation of Remotely Sensed Image Time Series (SPIRITS) (Eerens et al., 2014; Eerens and Haesen, 2015). For the sensitivity analysis 3, we used Google Earth Engine (GEE) (https://earthengine.google.org) which allowed implementing dedicated algorithms while only the threshold method is available in SPIRITS. The input 10 d time series were linearly interpolated at daily steps before phenological retrieval. For SPIRITS, the precision of phenological estimates is limited by the frequency of the input time series (10 d in our case) (non interpolation). For pixels with multiple growing seasons, we computed the phenological metrics for the growing season having the highest LAI amplitude.

The agreement between metric estimates from satellite imagery and

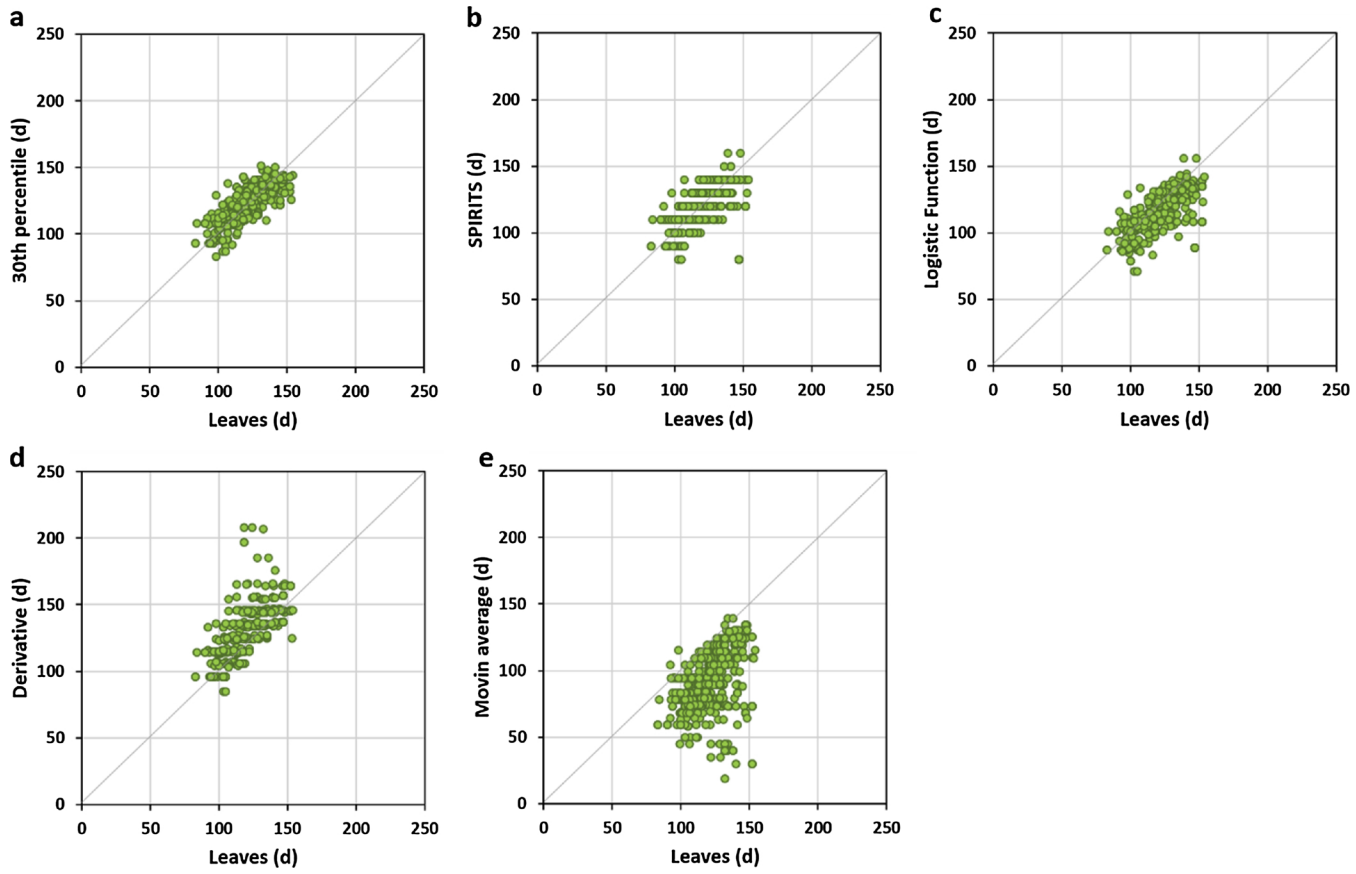

Fig. 6. Scatterplots between the SoS predicted from LAI V2 by the percentile method (a), SPIRITS (b), logistic function (c), derivative (d) and moving average (e) compared with the ground phenology (USA-NPN "leaves"). Values are given in DoY. Statistics of the comparison are indicated in Table 4.

ground-based measurements was quantified using the slope of the linear regression, the Pearson correlation coefficient (R), bias, i.e. the average difference between the satellite-derived phenology and the observed date (a positive bias indicated that SoS and EoS occurred later than the observed leaf out and autumnal coloring, respectively), the absolute bias and root mean square error (RMSE) calculated (e.g. for SoS) as:

$$\mathrm{RMSE}_{\mathrm{SoS}} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(\mathrm{SoS}_{\mathrm{ground}} - \mathrm{SoS}_{\mathrm{est.}})^2} \quad (1)$$

where n is the number of samples

3. Results

3.1. Sensitivity analysis of the smoothing method

The SWETS method performed the best for the reconstruction of seasonal trajectories and the estimation of phenological metrics:16 d in terms of RMSE and 1 d in terms of bias for the timing of SoS, and 34 d (RMSE) and 12 d (bias) for the timing of EoS (Table 2).

3.2. Sensitivity analysis of the vegetation variable

The best agreement with ground measurements for the timing of SoS was found for the 30% threshold of LAI amplitude, 40% of FCOVER, 50% of FAPAR and 50% of NDVI. The best metric definitions for the EOS were based on 40% of LAI and FCOVER amplitudes and 50% for FAPAR and NDVI. Phenological metrics derived from V2 time series improved over V1 for all variables (compare V1 and V2 statistics in Table 3).

The best performances for SoS were obtained using the LAI and FCOVER V2 time series (Fig. 4a), with RMSEs of ˜12 and 14 d, respectively (Table 3). In contrast, RMSEs were ˜21 and 23 d for NDVI V2.1 and FAPAR V2, respectively (Table 3), and contained many outliers (Fig. 4a). SoS was slightly underestimated for all cases, except when using LAI (Fig. 4a). EoS had a higher RMSE (25–48 d) and a lower R (< 0.3) (Table 3). The estimates of EoS from the LAI and FCOVER V2 time series also agreed best with ground data although no significant correlations were found (Table 3, Fig. 4b).

3.3. Sensitivity analysis of the method to extract phenological metrics

The 30th percentile and SPIRITS applied to LAI V2 provided the best performances among the different analyzed methods when compared both with USA-NPN and Europe-PEP725 measurements of the SoS (Fig. 5, Table 4). The results for the 30th percentile and SPIRITS were similar because both methods use the same definition of SoS based on the 30% threshold of annual amplitude but the 30th percentile method slightly improved SPIRITS in terms of precision (c.f. Fig. 6a–b, Fig. 7a–b) and accuracy (RMSE of 9 vs 14 d for USA-NPN "leaves", 9 vs 10 d for USA-NPN "increasing leaf size" and 11 vs 12 d for PEP725 "first visible leaves" (Table 4)) because of the daily interpolation applied to the input 10 d data. The logistic function also performed well (RMSE from 13 to 18 d) but provided slightly advanced SoS as compared to ground measurements (bias from -5 to -9 d). The derivative method provided showed higher RMSE (up to 28 d) and positive bias (up to 19 d). The moving-average method performance the worst (RMSE from 30 to 59 d) and systematically advanced the timing of SoS as compared to ground measurements (bias from -24 to -51d).

The different methods provided poorer performances for the EoS (Table 4, Fig. 8). The best agreement with ground measurements was

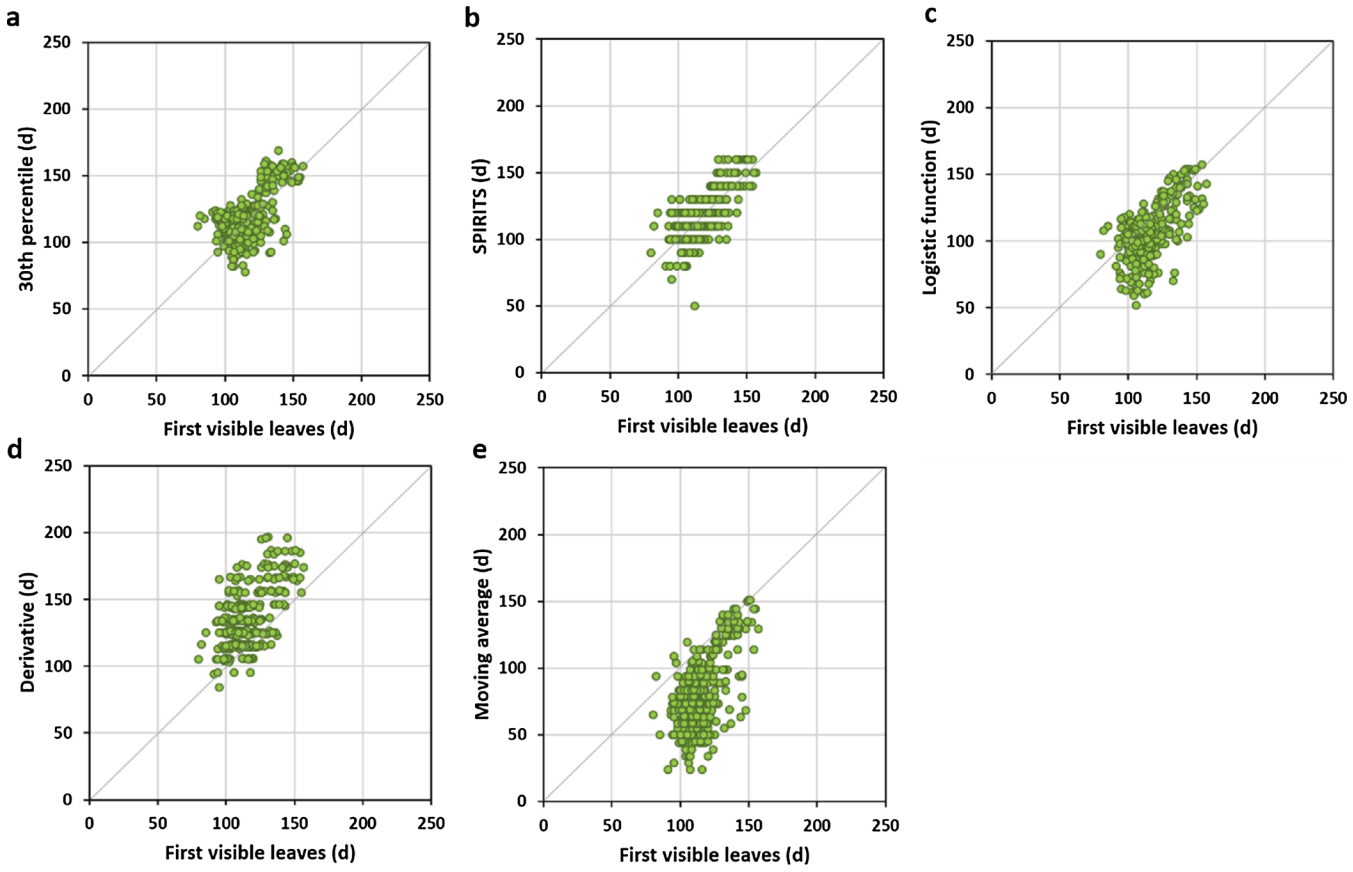

Fig. 7. Scatterplots between SoS predicted from LAI V2 by the percentile (a), SPIRITS (b), logistic function (c), derivative (d) and moving average (e) methods and ground phenology (PEP725 "first visible leaves"). Values are given in DoY. Statistics of the comparison are indicated in Table 4.

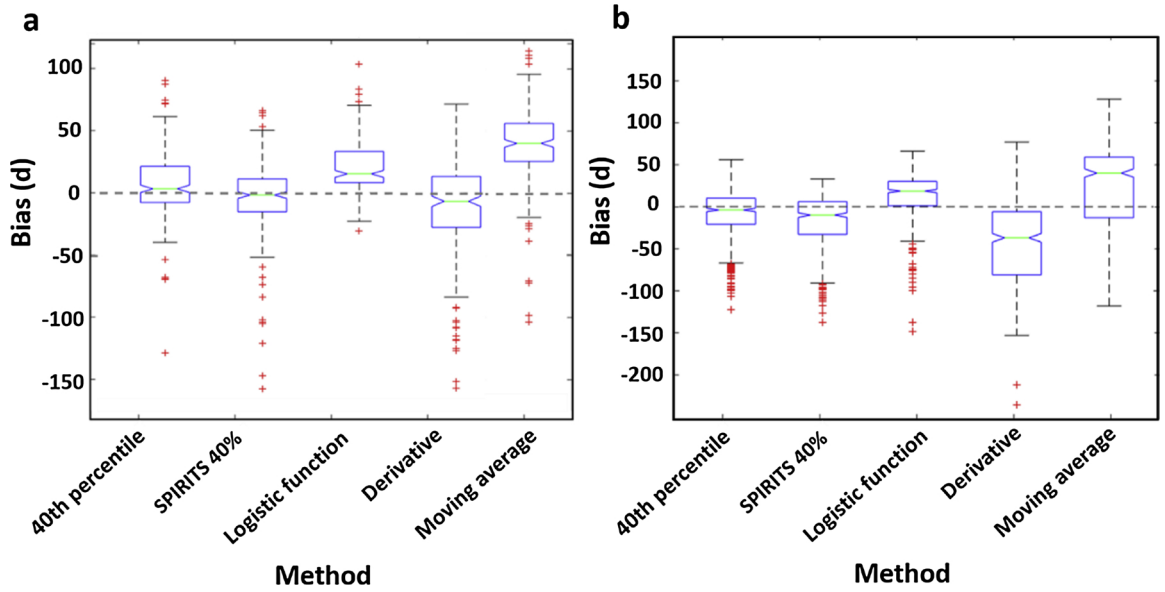

Fig. 8. Boxplots of the bias error for the EoS estimated from LAI V2 minus the ground measurements at the USA-NPN (a) and PEP725 (b) sites. An elongated boxplot indicates a greater dispersion of the average bias in each method.

found for the 40th percentile method (RMSE of 25 d for USA-NPN and 28 d for PEP725 "colored leaves" measurements). The logistic function provided similar performances in terms of RMSE but slightly over-estimated ground measurements (bias from 12 to 22 d). The derivative method showed higher scattering (RMSE from 40 to 65 d) and lower correlation than other methods (Table 4, Figs. 9 and 10). The moving average retrievals showed a positive delay as compared to ground data (bias of 38 d for PEP725 and 50 d for USA-NPN).

3.4. Spatial patterns of land surface phenology

Fig. 11 shows the average timing of SoS, EoS and LoS phenophases

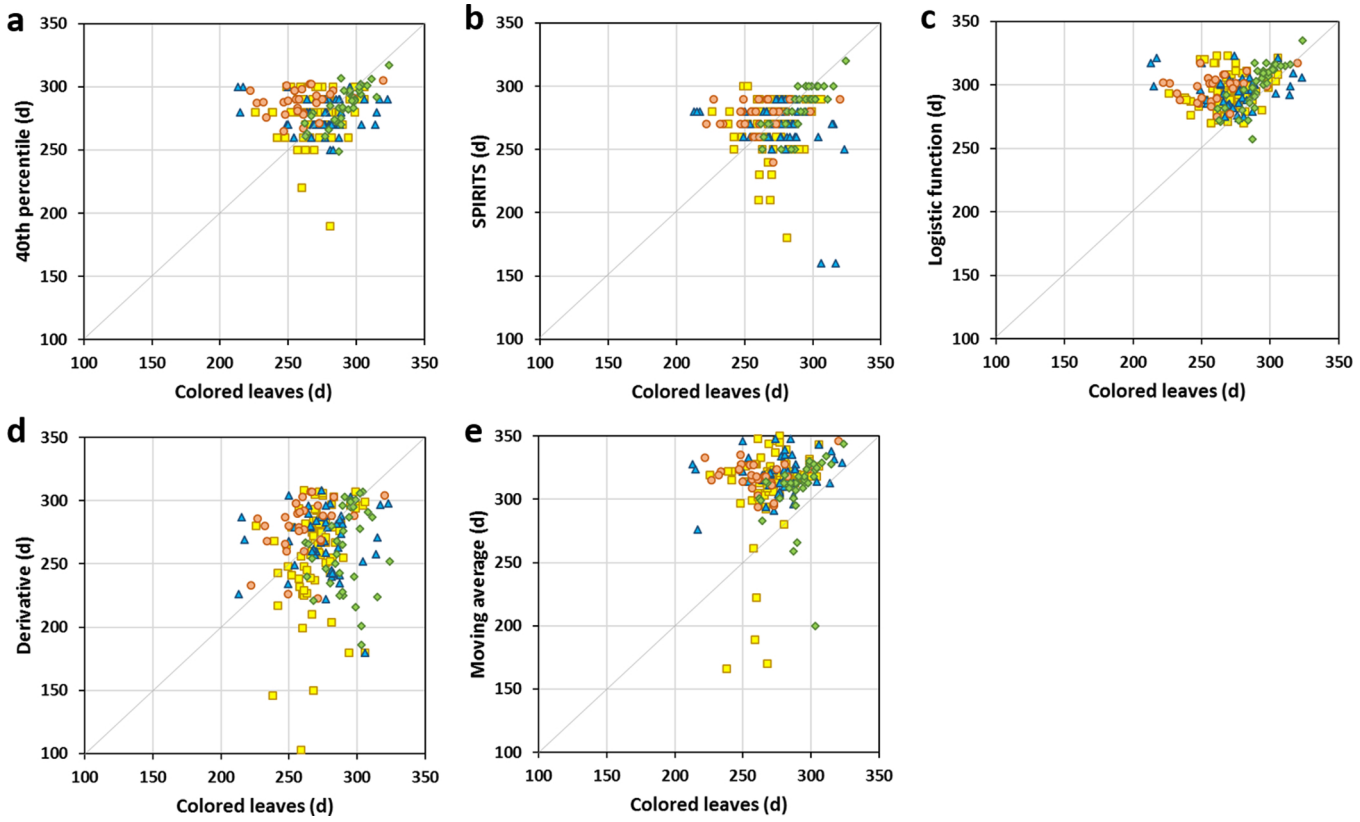

Fig. 9. Scatterplots between the EoS predicted from LAI V2 by the percentile (a), SPIRITS (b), logistic function (c), derivative (d) and moving average (e) methods and ground phenology (USA-NPN "colored leaves") for *Acer rubrum* (diamond), *Betula alleghaniensis* (circle), *Fagus grandifolia* (triangle) and *Quercus rubra* (square). Values are given in DoY. Statistics of the comparison are indicated in Table 4.

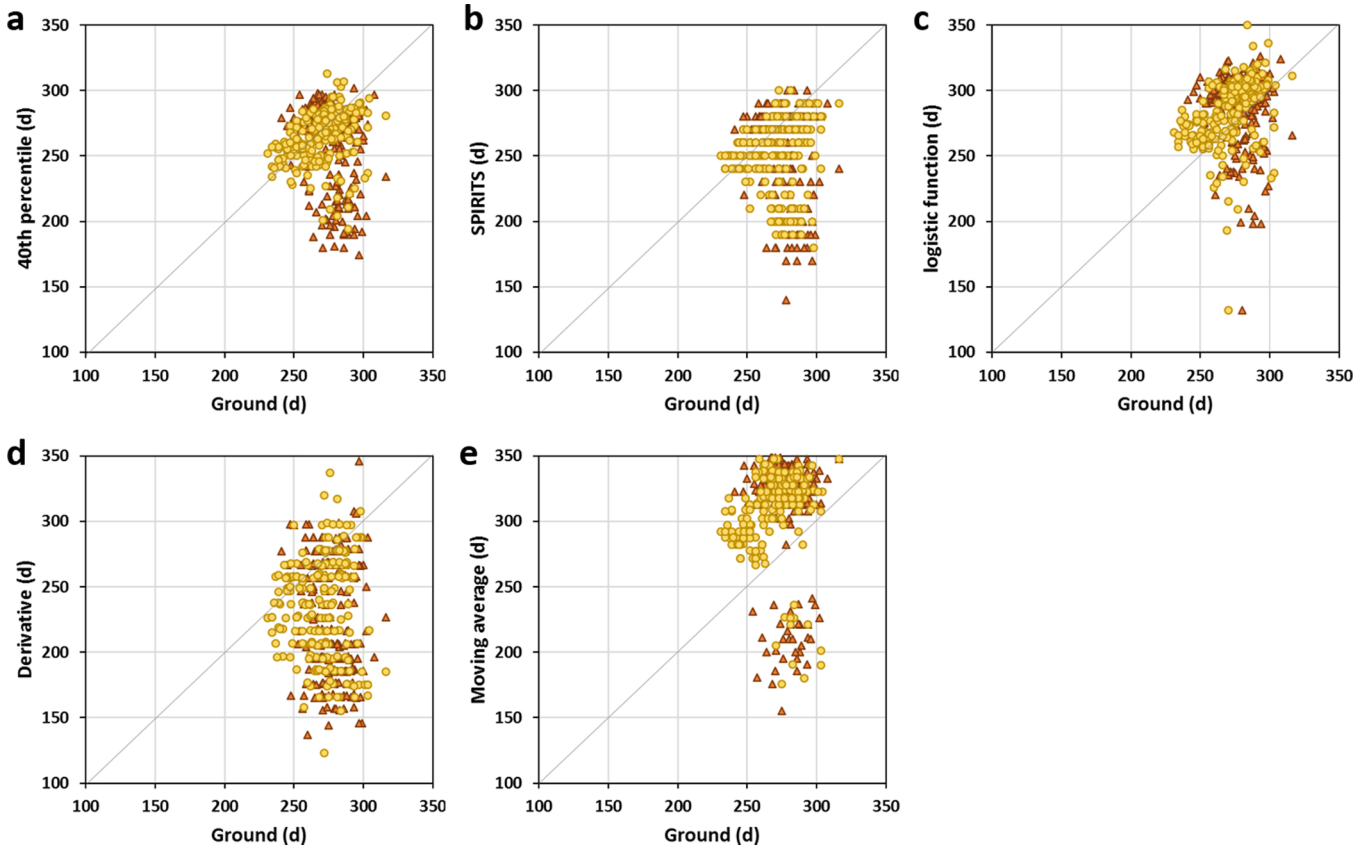

Fig. 10. Scatterplots between the EoS predicted from LAI V2 by the percentile (a), SPIRITS (b), logistic function (c), derivative (d) and moving average (e) methods and ground phenology (PEP725 "colored leaves"). Triangles for heterogeneous sites (forest cover < 50%) and circles for homogeneous sites (forest cover > 50%). Values are given in DoY. Statistics of the comparison are indicated in Table 4.

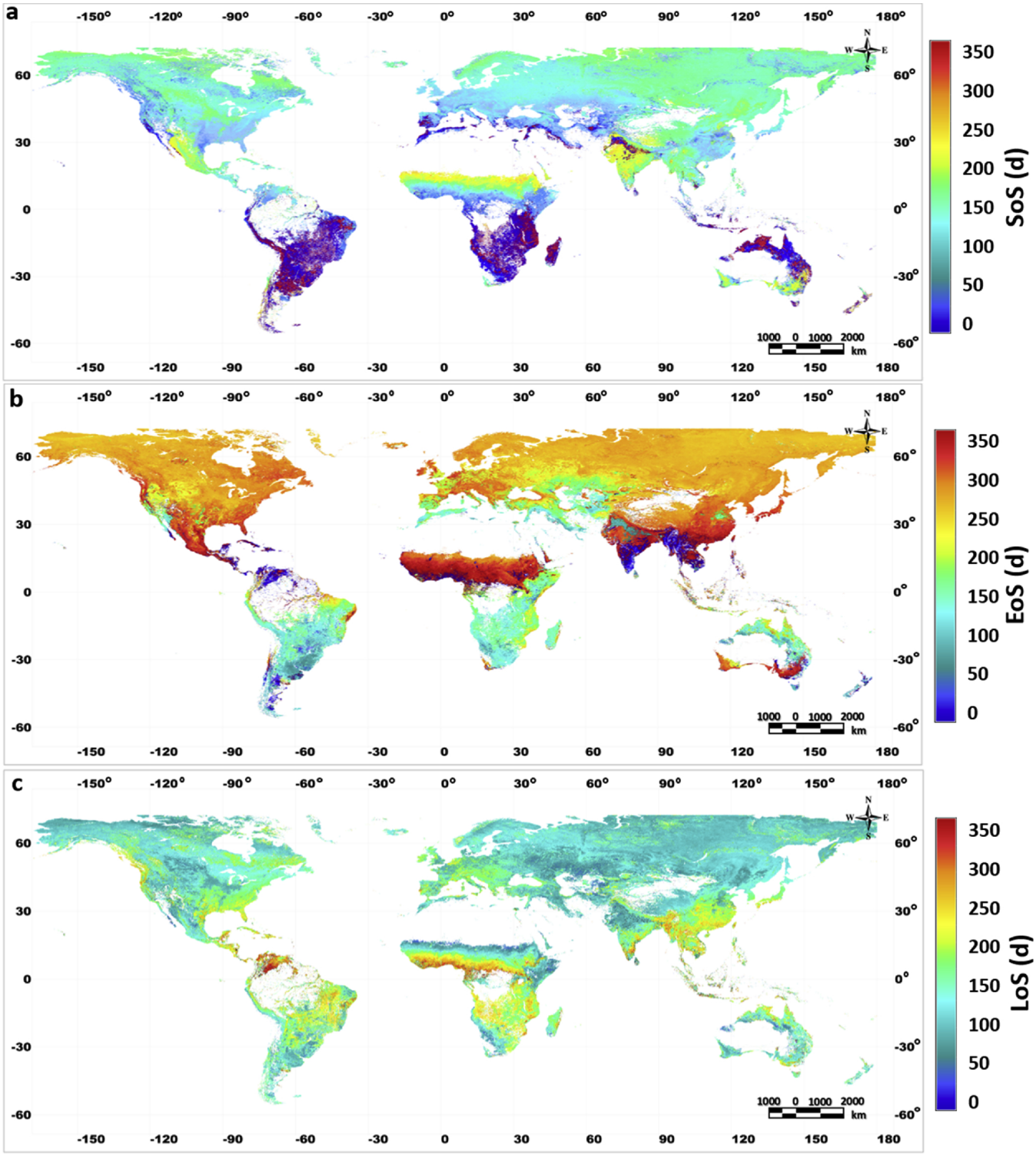

Fig. 11. Global Map of average SoS (a), EoS (b) and LoS (c) derived from the LAI V2 time series (1999–2017) and the threshold-based method. The continental areas in white are deserts and evergreen forests with limited seasonality where phenology was not computed.

at the global scale by using the percentile method (30th and 40th percentiles for SoS and EoS, respectively) and time series of V2 LAI (1999–2017). The LoS is estimated as the length of time between the SoS and the EoS. The derived maps show consistent spatial patterns of the seasonality of vegetation at the global scale which is driven by the distribution of latitudinal climatic patterns, type of vegetation and topographic elements, among other factors (Verger et al., 2015; Zhang et al., 2004). The timing of SoS (Fig. 11a) and EoS (Fig. 11b) reflected a broad variation in the range of values and their spatial pattern at the middle and high latitudes of the Northern Hemisphere strongly depend on the thermal and photoperiod latitudinal gradient (Verger et al., 2016).

The LoS (Fig. 11c) in some ecoclimatic and biogeographic regions of transition such as Sahel shows a broad range of variation from 10 to 200 d following the positive north-south gradient of rainfall (Verger et al., 2016). On the contrary, in northern latitudes > 50° the LoS

showed a limited range of variation from 10 to 60 d with shorter days at higher latitudes following the latitudinal gradient of temperature and radiation (Verger et al., 2016).

4. Discussion and conclusion

Land surface phenology (LSP) provides a synoptic view of vegetation dynamics and it can substantially improve our macroecological knowledge and the representation of phenology in earth-system models. Unlike previous studies limited to NDVI, we used three additional biophysical variables: LAI, FAPAR and FCOVER generated within CGLS from SPOT-VEGETATION (1999–2013) and PROBA-V (2014–2017) satellite imagery. We found that the phenology derived from LAI (or FCOVER) was more closely related to actual ground observation than the NDVI-derived phenology (Table 3). LAI-phenology is based on leaf development rather than on proxies provided by vegetation indices which are not driven solely by the amount of leaves but also by the canopy structure and the leaf biochemical properties (Richardson et al., 2009). LAI is more sensitive than vegetation indices such as NDVI to larger amounts of vegetation (Myneni and Williams, 1994). In addition, NDVI V2.1 is affected by variations in solar zenith and viewing angles and surface reflectance bidirectional effects (Tote et al., 2017).

In addition to the vegetation variable, the derived phenology was found to be highly sensitive to the retrieval algorithm and processing chain. We found that the retrieved phenology performed the best using LAI V2 (Table 3) due to the improved continuity (no missing data in V2) and smoothness as compared to V1. In V1 products and in NDVI time-series, the noise primarily due to cloud contamination and atmospheric effects is an important shortcoming in the study of land surface phenology. To overcome this limitation, smoothing methods were applied in SPIRITS approach as a pre-processing step to phenological retrieval. The choice of the smoothing method introduced differences of up to 50% in the performance of the phenology derived from LAI V1 (Table 2). These conclusions agree with previous literature studies highlighting the importance of the temporal reconstruction methods (Kandasamy et al., 2013; Verger et al., 2013). However, the phenology derived from original LAI V2 (smoothing and gap filling was already included in the retrieval algorithm) outperforms the phenology derived from the seasonal trajectories derived after smoothing LAI V1 (Table 3).

We tested four state of the art methods to extract phenological metrics: thresholds, logistic function, derivative and moving average. Each method has its own strengths and limitations (de Beurs and Henebry, 2010). The threshold approach based on a percentage of the annual amplitude is simple and robust but it is sensitive to the minimum and maximum values that may be affected by noise in the signal. The logistic function approach has been widely used (e.g. Zhang et al., 2003) but it is limited to the performance of the model fitting (Beck et al., 2006) and it may fail when the curvature function is too flat to determine the phenophases (de Beurs and Henebry, 2010). The derivative approach based on the maximum increase and decrease of the vegetation variable is very sensitive to the noise in the signal and the temporal smoothing and composition approach and it cannot represent short growing seasons well, especially when the increase and decrease in the annual time series occur rapidly and abruptly (Beck et al., 2006). The moving average approach is based on the assumption that vegetation growth follows a well-defined temporal profile and it may fail in cases of disturbances and abrupt changes. Further, the selection of the time lag is arbitrary.

We found that the choice of the extraction method introduced differences > 150% and > 85% in the performance of the SoS and EoS, respectively, derived from LAI V2 when compared to ground observations (PEP725 and USA-NPN) (Table 4). The percentile method agreed the best with ground measurements. The validation over ground observations indicated that the 30% threshold of the LAI amplitude was optimal for detecting SoS but that a 40% threshold was more suitable for detecting EoS in agreement with Verger et al. (2016). The accuracy for SoS using ground observations produced overall RMSEs of 9 and 11 d for the date of increasing leaf size and leaf unfolding, respectively, for the forests in Europe and the USA. The biases were < 2 d (˜10 d of standard deviation). We found poorer performances for EoS, than SoS, with higher RMSE of 25 d (28 d) and a bias of 8 d (-6 d) in USA (Europe). The lower performances for the EoS as compared to the SoS is associated to higher uncertainties of both satellite (atmospheric effects, snow and poor illumination conditions) (Delbart et al., 2005) and ground (the timing of leaf colouring is more subjective and difficult to identify than spring phenophases like leaf unfolding) (Estrella and Menzel, 2006) phenology for autumn. Richardson et al. (2009) also reported higher variabilities across the canopy of the timing and rate of foliar development in autumn than spring.

The validation of land surface phenology with ground observations presented some difficulties, such as the spatial distribution and the spatial representability of the data. The ground measurements represent the phenology for a limited number of individual plants that are not necessarily the most representative species of the 1-km satellite pixels. Conversely, satellite phenology at 1-km resolution represented an integrated response across landscapes with diverse species and phenological behaviors. The phenology of each species and their characteristics (sizes, ages, homogeneity), though, influenced the satellite signal, depending on its abundance within the pixel sampling area and on the timing of their phenophases (Delbart et al., 2015). Statistics of the comparison between LAI V2 derived EoS using the percentile method and ground measurements improved significantly when the analysis is restricted to *Acer rubrum* in USA-NPN (RMSE of 10 d, bias of -6 d and significant correlation ($p < 0.001$) of 0.8 (c.f. Table 4)) and to homogeneous sites (forest cover > 50% based on GEE high resolution imagery) in PEP725 (RMSE of 22 d, bias of 6 d and significant correlation ($p < 0.05$) of 0.4 (c.f. Table 4)).

The differences in the definition of the ground phenophases and satellite metrics hamper the comparison. The logistic function and, specially, the moving-average approaches systematically advanced the SoS as compared to ground measurements since these methods determine the SoS as the timing when the vegetation variable starts to increase (Table 4). On the contrary, the derivative approach based on the most rapid increase of the signal introduces positive delays in the SoS. The opposite trend is observed for the EoS (Table 4): the logistic function and moving-average show a delay in the detection of the EoS while the derivatives advances the timing of EoS. On the other hand, ground measurements are subjective and some ambiguity exists in the definition of phenophases. In this sense, we found positive bias of 2 d for the SoS retrieved with the 30th percentile of LAI amplitude when comparing with USA-NPN “leaves” phenophase but negative bias of -2 d when comparing to “increasing leaf size”. Further studies will focus on the comparison of the retrieved land surface phenology with continuous ground observations from PhenoCam (Zhang et al., 2018). This should ultimately lead to propose a standardization in the definition of phenological metrics.

This research is expected to contribute for the development of a dedicated algorithm for the operational retrieval of land surface phenology within CGLS. Validation using ground observations was limited to deciduous broadleaf forests. Further studies should extend the analysis to other vegetation types. The methods may need to be adapted to handle multiple and irregularly occurring vegetation growing cycles. Finally, forecasting approaches need to be developed for near-real time land surface phenology retrieval.

Acknowledgements

This research was supported by an FPU grant (Formación del Profesorado Universitario) from the Spanish Ministry of Education and Professional Training to the first author (FPU15-04798), the Copernicus Global Land Service (CGLOPS-1, 199494-JRC), the Spanish

Government grant CGL2016-79835-P, the Catalan Government grant SGR 2017-1005, and the European Research Council Synergy grant ERC-2013-SyG-610028 IMBALANCE-P.

References

Atkinson, P.M., Jeganathan, C., Dash, J., Atzberger, C., 2012. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. Remote Sens. Environ. 123, 400–417.

Atzberger, C., Eilers, P.H., 2011. Evaluating the effectiveness of smoothing algorithms in the absence of ground reference measurements. Int. J. Remote Sens. 32, 3689–3709.

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholcyzk, P., Smets, B., 2013. GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 1: principles of development and production. Remote Sens. Environ. 137, 299–309.

Beaumont, L., Hartenthaler, T., Keatley, M., Chambers, L., 2015. Shifting time: recent changes to the phenology of Australian species. Clim. Chang. Res. Lett. 63, 203–214.

Beck, P.S.A., Atzberger, C., Hogda, K.A., Johansen, B., Skidmore, A.K., 2006. Improved monitoring of vegetation dynamics at very high latitudes: a new method using MODIS NDVI. Remote Sens. Environ. 100, 321–334.

Brus, D., Hengeveld, G., Walvoort, D., Goedhart, P., Heidema, A., Nabuurs, G., Gunia, K., 2012. Statistical mapping of tree species over Europe. Eur. J. For. Res. Vol. 131, 145–157.

Cai, Z., Jönsson, P., Jin, H., Eklundh, L., 2017. Performance of smoothing methods for reconstructing NDVI time-series and estimating vegetation phenology from MODIS data. Remote Sens. 9, 1271.

Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., Eklundh, L., 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. Remote Sens. Environ. 91, 332–344.

de Beurs, K., Henebry, G., 2005. Land surface phenology and temperature variation in the International Geosphere–Biosphere Program high-latitude transects. Glob. Change Biol. 11, 779–790.

de Beurs, K.M., Henebry, G.M., 2010. Spatio-temporal statistical methods for modelling land surface phenology. In: Hudson, Keatley (Eds.), Phenological Research: Methods for Environmental and Climate Change Analysis, pp. 177–208.

Delbart, N., Kergoat, L., Le Toan, T., Lhermitte, J., Picard, G., 2005. Determination of phenological dates in boreal regions using normalized difference water index. Remote Sens. Environ. 97, 26–38.

Delbart, N., Beaubien, E., Kergoat, L., Le Toan, T., 2015. Comparing land surface phenology with leafing and flowering observations from the PlantWatch citizen network. Remote Sens. Environ. 160, 273–280.

Eerens, H., Haesen, D., 2015. User's manual of Software for the Processing and Interpretation of Remotely Sensed Image Time Series. Version 1.4.0. Available at. http://spirits.jrc.ec.europa.eu/files/SpiritsManual_140.pdf.

Eerens, H., Haesen, D., Rembold, F., Urbano, F., Tote, C., Bydekerke, L., 2014. Image time series processing for agriculture monitoring. Environ. Model. Softw. 53, 154–162.

Eilers, P.H.C., 2003. A perfect smoother. Anal. Chem. 75 (14), 3631–3636.

Estrella, N., Menzel, A., 2006. Responses of leaf colouring in four deciduous tree species to climate and weather in Germany. Clim. Res. 32, 253–267.

Garrity, S., Bohrer, G., Maurer, K., Mueller, K., Vogel, C., Curtis, P., 2011. A comparison of multiple phenology data sources for estimating seasonal transitions in deciduous forest carbon exchange. Agric. For. Meteorol. 151, 1741–1752.

Ivits, E., Cherlet, M., Mehl, W., Sommer, S., 2009. Estimating the ecological status and change of riparian zones in Andalusia assessed by multi-temporal AVHHR datasets. Ecol. Indic. 9, 422–431.

Jönsson, P., Eklundh, L., 2002. Seasonality extraction by function fitting to time-series of satellite sensor data. IEEE Trans. Geosci. Remote. Sens. 40, 1824–1832.

Kandasamy, S., Baret, F., Verger, A., Neveux, P., Weiss, M., 2013. A comparison of methods for smoothing and gap filling time series of remote sensing observations application to MODIS LAI products. Biogeosciences 10, 4055–4071.

Kandasamy, S., Fernandes, R., 2015. An approach for evaluating the impact of gaps and measurement errors on satellite land surface phenology algorithms: application to 20year NOAA AVHRR data over Canada. Remote Sens. Environ. 164, 114–129.

Meier, U., Bleiholder, H., Buhr, L., Feller, C., Hack, H., Heß, M., Lancashire, P.D., 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. Journal für Kulturpflanzen 61, 41–52.

Myneni, R.B., Williams, D.L., 1994. On the relationship between FAPAR and NDVI. Remote Sens. Environ. 49, 200–211.

Peñuelas, J., Rutishauser, T., Filella, I., 2009. Phenology feedbacks on climate change. Science 324 (5929), 887–888.

Reed, B.C., Brown, J.F., Vander, D., 1994. Measuring phenological variability from satellite imagery. J. Veg. Sci. 5, 703–714.

Richardson, A., Braswell, B., Hollinger, D., Jenkins, J., Ollinger, S., 2009. Near-surface remote sensing of spatial and temporal variation in canopy phenology. Ecol. Appl. 19, 1417–1428.

Richardson, A., Keenan, T., Migliavacca, M., Ryu, Y., Sonnentag, O., Toomey, M., 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. Agric. For. Meteorol. 169, 156–173.

Schwartz, M., Betancourt, J., Weltzin, J., 2012. From caprio's lilacs to the USA national phenology network front. Ecol. Environ. 10, 324–327.

Swets, D.L., Reed, B.C., Rowland, J.R., Marko, S.E., 1999. A weighted least-squares approach to temporal smoothing of NDVI. ASPRS Annual Conference, From Image to Information, Portland, Oregon, May 17-21, Proceedings: Bethesda, Maryland. American Society for Photogrammetry and Remote Sens.

Tateishi, R., Ebata, M., 2004. Analysis of phenological change patterns using 1982-2000 Advanced very High-Resolution Radiometer (AVHRR) data. Remote Sens. 25, 2287–2300.

Templ, B., Koch, E., Bolmgren, K., Ungersböck, M., Paul, A., Scheifinger, H., 2018. Pan European Phenological database (PEP725): a single point of access for European data. Int. J. Biometeoroly 62, 1109–1113.

Toté, C., Swinnen, E., Sterckx, S., Clarijs, D., Quang, C., Maes, R., 2017. Evaluation of the SPOT/VEGETATION collection 3 reprocessed dataset: surface reflectances and NDVI. Remote Sens. Environ. 201, 219–233.

Verger, A., Baret, F., Weiss, M., Kandasamy, S., Vermote, E., 2013. The CACAO method for smoothing, gap filling and characterizing seasonal anomalies in satellite time series. IEEE Trans. Geosci. Remote. Sens. 51, 1963–1972.

Verger, A., Baret, F., Weiss, M., 2014. Near real time vegetation monitoring at global scale. IEEE J. Stars 7, 3473–3481.

Verger, A., Bater, M., Weiss, M., Filella, Peñuelas, J., 2015. GEOCLIM A global climatology of LAI, FAPAR, and FCOVER from VEGETATION observations for 1999-2010. Remote Sens. Environ. 166, 126–137.

Verger, A., Filella, I., Baret, F., Peñuelas, J., 2016. Vegetation baseline phenology from kilometric global LAI satellite products. Remote Sens. Environ. 178, 1–14.

Verger, A., Baret, F., Weiss, M., 2017. Algorithm Theoretical Basis Document: LAI, FAPAR, FCOVER Collection 1km, Version 2, Issue I1.40. Available at: https://land.copernicus.eu/global/sites/cgls.vito.be/files/products/GIOGL1_ATBD_FAPAR1kmV2_I1.40.pdf. .

Viovy, N., Arino, O., Belward, A.S., 1992. The best index slope extraction (BISE): a method for reducing noise in NDVI time-series. Int. J. Remote Sens. 13, 1585–1590.

White, M.A., de Beurs, K., Didan, K., Inouye, D., Richardson, A., Jensen, O., Lauenroth, W., 2009. Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. Glob. Change Biol. 15, 2335–2359.

Wu, C., Gonsamo, I., Gough, A., Chen, J., Xu, S., 2014. Modeling growing season phenology in North American forests using seasonal mean vegetation indices from MODIS. Remote Sens. Environ. 147, 79–88.

Yu, L., Liu, T., Bu, K., Yan, F., Yang, J., Chang, L., Zhang, S., 2017. Monitoring the long-term vegetation phenology change in northeast China from 1982 to 2015. Sci. Rep. 7, 14770.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., Huete, A., 2003. Monitoring vegetation phenology using MODIS. Remote Sens. Environ. 84 471-475.

Zhang, X., Friedl, M.A., Schaaf, C.B., 2004. Climate controls on vegetation phenological patterns in northern mid- and high latitudes inferred from MODIS data. Glob. Change Biol. 10, 1133–1145.

Zhang, X., Jayavelu, S., Liu, L., Friedl, M.A., Henebry, G.M., Liu, Y., Schaaf, C.B., Richardson, A.D., Gray, J., 2018. Evaluation of land surface phenology from VIIRS data using time series of PhenoCam imagery. Agric. For. Meteorol. 256, 137–149.