Universitat Autònoma de Barcelona

Spatio-temporal variability of vegetation phenology and its drivers at global scale

Adrià Descals Ferrando

Spatio-temporal variability of vegetation phenology and its drivers at global scale

PhD Thesis

Adrià Descals Ferrando

to be eligible for the Doctor degree

Supervised by:

Prof. Josep Peñuelas Reixach

Dr. Aleixandre Verger Ten

PhD in Terrestrial Ecology

Global Ecology Unit (CREAF-CSIC)

Center for Ecological Research and Forestry Applications

Universitat Autònoma de Barcelona, September 2022

Image generated from scratch by an artificial intelligence algorithm — *multimodalart* — given only the instruction “Trees orbiting the Earth”.

Table of contents

Abstract

Phenology is the study of the timing of periodic life-cycle events in plants and animals, as well as how seasonal changes in climatic factors affect these events. In recent decades, temperature has increased, leading to changes in the timing of leaf phenophases and the lengthening of the photosynthetically growing season. These changes in vegetation phenology have an impact on the water and carbon cycle; the lengthening of the growing season has increased the carbon uptake of terrestrial ecosystems offsetting part of the atmospheric carbon from human emissions. However, future responses of vegetation phenology to warming are uncertain, raising concerns about the ability of vegetation to offset atmospheric carbon emissions. The main objective of the thesis was to characterize the spatial and temporal variability of land surface phenology and link this variability to climate drivers in a context of global warming. The research focused on understanding how rising temperatures might change the climate factors affecting vegetation phenology. The first two chapters deal with the methodology for estimating land surface phenology. In Chapter 1, we propose a new method for estimating land surface phenology in cloud-based platforms that can be applied to raw time series without the need for time series preprocessing. In Chapter 2, we present 10-meter resolution maps for the continental scale, emphasizing the importance of spatial resolution. Chapters 3 to 5 cover the impact of three climate factors, temperature, light, and water availability, on the vegetation phenology. In Chapter 3, results suggest that frozen soils constraint vegetation activity, and vegetation resume photosynthesis closely after soil thawing. Chapter 4 points at the climate constraints on carbon uptake phenology, particularly the limitation of radiation in temperate and cold regions in the Northern Hemisphere. These findings suggest that the start of the growing season may still advance, although at a slower pace, while radiation restrains photosynthesis in autumn, preventing the further delay of the end of season with future warming. Chapter 5 studies another factor, the occurrence of heatwaves and droughts, that advances the end of the growing season. In this chapter, we show evidence of early leaf shedding using 10-meter resolution satellite data, and link this phenomenon to high temperature and aridity conditions. These findings reveal that early leaf shedding is more recurrent and widespread than previously reported. The last chapter, Chapter 6, is a compendium of techniques and knowledge gained in previous chapters. We model the land surface phenology estimated with the technique describe in Chapter 1 and using a model based on the climate factors studied in Chapters 3-5. We analyze the climate constraints on vegetation phenology at

the global scale using a novel technique that explains the relationship between sun-induced fluorescence and climate factors in a machine learning model. The findings of the thesis demonstrate different spatial constraints of temperature, light, and water availability at the beginning and end of the growing season, suggesting that vegetation phenology will respond differently to future climatic warming depending on the location and type of vegetation. In some regions, rising temperatures may not translate into a lengthening of the growing season because of radiation and water constraints.

General introduction

Phenology is defined as the timing of periodic life-cycle events in plants and animals and how these events are modulated by seasonal variations in climate (Schwartz, 2003). Life-cycle events in plant phenology involve changes in foliar biomass such as leaf unfolding and leaf shedding and the timing when vegetation becomes photosynthetically active, but also the timing for reproductive stages, such as flowering. Temperature, water availability, and day length have been proposed as the main environmental drivers that constrain vegetation activity and regulate plant phenology, acting eventually as confounded restraints for vegetation growth (Chuine & Régnière, 2017; Jolly et al., 2005; Kramer & Hänninen, 2009). These climate restraints induce vegetation into a stage of dormancy when the climate is adverse and a period of growth when conditions are favorable for vegetation activity.

Phenology have gained attention, particularly for the study of leaf phenophases, because such events have been an indicator of global warming and more broadly for its link with global environmental change (Badeck et al., 2004; Cleland et al., 2007; Peñuelas et al., 2009; Peñuelas & Filella, 2001). With recent climatic warming, the timing of leaf unfolding has advanced, and the leaf shedding has delayed. These changes have lengthened the growing season, as observed from in situ measurements of leaf phenophases as well as long-term time series from satellite observations. Vegetation phenology influences the carbon and water cycle (Keenan et al., 2014; Myneni et al., 1997; Richardson et al., 2010), successional transition (Chuine, 2010), and plays a role in the feedbacks between vegetation and climate (Peñuelas & Filella, 2009; Richardson et al., 2013). The lengthening of the growing season due to rising temperatures has increased the carbon uptake of terrestrial ecosystems (Le Quéré et al., 2009), offsetting part of the atmospheric carbon resulting from human emissions. However, future responses of vegetation to warming and its capacity of offsetting atmospheric carbon remains unclear (Piao et al., 2019), which makes future atmospheric carbon concentration and climate projections uncertain.

Remotely sensed data is becoming increasingly important for studying plant phenology and determining the effects of climate change on ecosystem functioning. Vegetation phenology explains the seasonal pattern of variation in vegetated land surfaces observed from satellite remote sensing. Such seasonal variations observed in remotely sensed data are known as land surface phenology. Structural vegetation indices, such as the normalized difference vegetation index (NDVI), indicate changes in foliar biomass, while physiological vegetation indices, such as

the photochemical reflectance index (PRI), are proxies of light use efficiency and changes in photosynthetic dynamics. Moreover, satellite products are derived from satellite spectral bands and provide time series of key biophysical variables such as leaf area index (LAI) and gross primary productivity (GPP). Such remotely sensed observations of land surface phenology have been validated with in situ observations of leaf phenophases, such as the PEP725 and the NPN networks, the PhenoCam network, and measurements of carbon fluxes in FLUXNET towers (Bórnez et al., 2020).

Moderate to broad-resolution satellite data, such as the Advanced Very-High-Resolution Radiometer (AVHRR) (1.1 km) and the Moderate Resolution Imaging Spectroradiometer (MODIS) (500 m), have been employed often for the estimation of land surface phenology metrics (Park et al., 2016; Zeng et al., 2011). However, coarse resolution LSP metrics tend to overestimate the length of the growth season. Recent developments in remote sensing technology have made it possible to estimate LSP with higher spatial resolution. The Sentinel-2 mission provides decametric images with frequent revisit times (5 d) and, jointly with Landsat-7 and Landsat-8, the revisit time is reduced to 2-3 days at the equator. The dense time series of satellite data at finer scale allows for the extraction of phenology metrics and the study of vegetation dynamics at the canopy scale (Bolton et al., 2020). Moreover, the development of cloud-based platforms, such as Google Earth Engine (GEE) (Gorelick et al., 2017), allows for the processing of large volumes of satellite data for planetary-scale analysis (Hansen et al., 2013; Pekel et al., 2016).

The main goal of the research presented in the thesis was to characterize the spatial and temporal variability of land surface phenology and link this variability to climate drivers within the context of global warming. The focus of the research was on gaining an understanding of how climate factors that are limiting photosynthesis at the start and end of the growing season might change as a result of rising temperatures. These climate factors included temperature, light, and water availability, and the climate variables used were mainly soil and air temperature, incoming shortwave radiation, and the aridity index. The thesis is divided in six chapters. Chapters 1 and 2 focus on methodological aspects of land surface phenology metric estimation, taking into account aspects such as the impact of the spatial resolution and time series processing in cloud-based platforms. Chapters 3 to 5 analyse the three climate factors aforementioned: temperature, and light and water availability. Chapter 6 explores a novel

method for explaining the constraints of climate factors on the phenology of carbon uptake at the global scale.

More specifically, in Chapter 1, we propose a new method for land surface phenology. The method has the capability to estimate the start and end of the growing season observed from near-surface and remotely-sensed time series, as well as binary time series (dormant versus growing stages). It improves upon state-of-the-art methods because it processes raw time series, which allow the fast and large-scale estimation of phenology metrics in cloud-based platforms. The method is also useful for the estimation of phenology metrics in the following chapters.

In Chapter 2, we present new data for the start and end of the growing season in a large area, the Arctic region, using a cloud-based platform. The phenology metrics were extracted from Sentinel-2 time series at 10-meter resolution and validated with near-surface reflectance measurements and cross-validated with MODIS data. The chapter evidences the importance of high-resolution land surface phenology estimation in heterogeneous landscapes. Such phenology metrics at 10-meter resolution are key for accurately monitoring forest disturbances as observed in Chapter 5.

In Chapter 3, we investigate the role of soil temperature at the start of the growing season, a key variable that remains understudied in phenology studies. We hypothesized that vegetation starts the growing season when frozen water melts and becomes available to plants. Results showed a correlation between the thawing date and the start of the growing season at high latitudes, but this relationship weakened in southerly latitudes.

In Chapter 4, we study the divergent spatial and temporal impacts of climate constraints on the vegetation activity at the start and end of the growing season. In particular, we examined the role of radiation on the vegetation productivity and how rising temperatures are increasing the radiation constraint at the end of the growing season. At the start of the growing season, however, vegetation is sensitive to temperature in most of the extratropical regions of the Northern Hemisphere and, thus, future global warming may still advance the start of the growing season especially for non-deciduous vegetation.

In Chapter 5, the role of water and droughts is analysed for studying the premature leaf shedding in European deciduous forests, a phenomenon that has implications on forest productivity and

tree mortality. First, we showed evidence of anomalous early leaf shedding (before September 1st) in several locations of Europe. Second, phenology metrics at the end of season were extracted from 10-meter resolution Sentinel-2 data, which that early leaf shedding is more recurrent and widespread than previously reported. Finally, we show that early leaf shedding is related to anomalous arid conditions that occurred during the study period.

In Chapter 6, we extended the analysis of climate constraints that were examined in Chapter 3 and 4, and considered the climate constraints of water as in Chapter 5. The global climate constraints on sun-induced fluorescence (SIF) were evaluated at the start and end of the growing season using a local interpretation of machine learning models, an approach that is novel for remote sensing studies. This approach estimates the contribution of the input variables –climate factors– on individual predictions of the model –SIF. Thus, this approach tackles lack of interpretability—the black box problem— in machine learning models used in remote sensing research. Results confirmed limitations of temperature, radiation, and vapor-pressure-deficit on vegetation productivity but at a finer spatial detail.

References

Badeck, F., Bondeau, A., Böttcher, K., Doktor, D., Lucht, W., Schaber, J., & Sitch, S. (2004). Responses of spring phenology to climate change. *New Phytologist*, *162*(2), 295–309.

Bolton, D. K., Gray, J. M., Melaas, E. K., Moon, M., Eklundh, L., & Friedl, M. A. (2020). Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. *Remote Sensing of Environment*, *240*, 111685.

Bórnez, K., Richardson, A. D., Verger, A., Descals, A., & Peñuelas, J. (2020). Evaluation of vegetation and proba-v phenology using phenocam and eddy covariance data. *Remote Sensing*, *12*(18), 3077.

Chuine, I. (2010). Why does phenology drive species distribution? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1555), 3149–3160.

Chuine, I., & Régnière, J. (2017). Process-Based Models of Phenology for Plants and Animals. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 159–182. https://doi.org/10.1146/annurev-ecolsys-110316-022706

Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A., & Schwartz, M. D. (2007). Shifting plant phenology in response to global change. *Trends in Ecology & Evolution*, *22*(7), 357–365. https://doi.org/10.1016/j.tree.2007.04.003

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S., Goetz, S. J., Loveland, T. R., & others. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, *342*(6160), 850–853.

Jolly, W. M., Nemani, R., & Running, S. W. (2005). A generalized, bioclimatic index to predict foliar phenology in response to climate. *Global Change Biology*, *11*(4), 619–632. https://doi.org/10.1111/j.1365-2486.2005.00930.x

Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., Munger, J. W., O'Keefe, J., Schmid, H. P., Wing, I. S., & others. (2014). Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nature Climate Change*, *4*(7), 598–604.

Kramer, K., & Hänninen, H. (2009). The Annual Cycle of Development of Trees and Process-Based Modelling of Growth to Scale Up From the Tree To the Stand. In A. Noormets (Ed.), *Phenology of Ecosystem Processes: Applications in Global Change Research* (pp. 201–227). Springer New York. https://doi.org/10.1007/978-1-4419-0026-5_9

Le Quéré, C., Raupach, M. R., Canadell, J. G., Marland, G., Bopp, L., Ciais, P., Conway, T. J., Doney, S. C., Feely, R. A., Foster, P., & others. (2009). Trends in the sources and sinks of carbon dioxide. *Nature Geoscience*, *2*(12), 831–836.

Myneni, R. B., Keeling, C., Tucker, C. J., Asrar, G., & Nemani, R. R. (1997). Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature*, *386*(6626), 698–702.

Park, T., Ganguly, S., Tømmervik, H., Euskirchen, E. S., Høgda, K.-A., Karlsen, S. R., Brovkin, V., Nemani, R. R., & Myneni, R. B. (2016). Changes in growing season duration and

productivity of northern vegetation inferred from long-term remote sensing data. *Environmental Research Letters*, *11*(8), 084001.

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422.

Peñuelas, J., & Filella, I. (2001). Responses to a warming world. *Science*, *294*(5543), 793–795.

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

Peñuelas, J., Rutishauser, T., & Filella, I. (2009). Phenology Feedbacks on Climate Change. *Science*, *324*(5929), 887–888. https://doi.org/10.1126/science.1173004

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., & Zhu, X. (2019). Plant phenology and global climate change: Current progresses and challenges. *Global Change Biology*, *25*(6), 1922–1940. https://doi.org/10.1111/gcb.14619

Richardson, A. D., Andy Black, T., Ciais, P., Delbart, N., Friedl, M. A., Gobron, N., Hollinger, D. Y., Kutsch, W. L., Longdoz, B., Luyssaert, S., & others. (2010). Influence of spring and autumn phenological transitions on forest ecosystem productivity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1555), 3227–3246.

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., & Toomey, M. (2013). Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agricultural and Forest Meteorology*, *169*, 156–173. https://doi.org/10.1016/j.agrformet.2012.09.012

Schwartz, M. D. (2003). *Phenology: An integrative environmental science*.

Zeng, H., Jia, G., & Epstein, H. (2011). Recent changes in phenology over the northern high latitudes detected from multi-satellite data. *Environmental Research Letters*, *6*(4), 045508.

Chapter 1

A threshold method for robust and fast estimation of land-surface phenology using Google Earth Engine

Adrià Descals, Aleixandre Verger, Gaofei Yin, and Josep Peñuelas

Published in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14 (2020): 601-606.* https://doi.org/10.1016/j.jag.2019.101974

Abstract

Cloud computing platforms are changing the way of analyzing remotely sensed data by providing high computational power and rapid access to massive volumes of data. Several types of studies use cloud-based platforms for global-scale analyses, but the number of land-surface phenology (LSP) studies that use cloud-based platforms is low. We analyzed the performance of state-of-the-art LSP algorithms and propose a new threshold-based method that we implemented in Google Earth Engine (GEE). This new LSP method, called Maximum Separation (MS) method, applies a moving window that estimates the ratio of observations that exceed a given threshold before and after the central day. The start and end of the growing season are the days of the year when the difference between the ratios before and after the central day are minimal and maximal. The MODIS phenology metrics estimated with the MS method showed similar performances as traditional threshold methods when compared with ground estimations derived from the PhenoCam dataset, a network of digital cameras that provides near-surface remotely-sensed observations of vegetation phenology. The main advantage of the MS method is that it can be directly applied to daily non-smoothed time series without any additional preprocessing steps. The implementation of the proposed method in GEE allowed the processing of global phenological maps derived from MODIS. The distribution of code in GEE allows the reproducibility of results and the rapid processing of LSP metrics by the scientific community.

1.1. Introduction

The study of land-surface phenology (LSP) entails the estimation of metrics from remotely sensed vegetation seasonality (Zeng et al., 2020). Its study is important for reliably estimating vegetation dynamics, commonly the start and end of the growing season (SoS and EoS), which correspond to the timing of the year when the vegetation is released from dormancy and when the growing season ends. Phenology has recently gained importance because of its linkage with global warming. Several studies have found that the length of the growing season has increased in recent decades due to the global warming (Cleland et al., 2007; Peñuelas & Filella, 2001). The influence of temperature and other climatic factors on the vegetation dynamics are still under discussion (Körner & Basler, 2010; Peñuelas & Filella, 2009), and reliable estimates of phenological metrics at a global scale are required to understand the links between the climatic factors and vegetation phenology.

Many cloud-processing platforms have recently become available to the scientific community, which has allowed the processing of large amounts of remotely sensed data. Google Earth Engine (GEE) is one of the most popular platforms (Gorelick et al., 2017). Some studies have addressed temporal aspects such as data smoothing and, particularly, phenology estimation using a curve fitting method (Li et al., 2019), but the number of LSP studies that make use of cloud-computing platforms is still low. One reason that explains the low number of LSP studies is that temporal processing in cloud platforms requires methods that can be vectorized and easily implemented. Code vectorization is the process of transforming an algorithm so that instead of processing elements on an array separately, generally with a for loop, the operations are performed on all components of the array simultaneously (Walt et al., 2011). Such elements of the array are commonly pixel values or satellite time series. Code vectorization enhances the computational power of GEE and allows the processing of satellite time series in an optimized manner. A freely available implementation of an LSP method in GEE would allow the fast processing of satellite archives in a variety of products and would ease the accessibility of data for computing large-scale LSP maps at different spatial resolutions.

The results of the state-of-the-art LSP methods, however, are affected to some extent by noise, gaps, and outliers presented in the time series, so a set of preprocessing steps are required to prepare and reconstruct the time series captured from remotely sensed data (Zeng et al., 2020). These steps commonly include outlier rejection filters, temporal compositing, smoothing, gap-

filling and interpolation of the time series to regular or daily time steps. A commonly used approach in the literature is the maximum value composite (Jonsson & Eklundh, 2002; Zhang et al., 2006). This temporal compositing method computes the maximum value of the vegetation index for a temporal window and assumes that negatively biased values are contaminated by clouds or cloud shadows. In addition to the temporal compositing, smoothing techniques are also used to reduce the noise of the time series. The most common smoothing techniques can be classified into moving-window (Bórnez et al., 2020; Ma & Veroustraete, 2006; Viovy et al., 1992) and curve-fitting methods (Beck et al., 2007; Fisher et al., 2006; Jonsson & Eklundh, 2002; Zhang et al., 2006). The first category runs a moving window in the time series, such as the moving average filter, while the second category fits a function to the time series, such as logistic and Gaussian models. These preprocessing steps have various shortcomings: a) they are highly parametric (window size, smoothing type), and results are not robust in all regions; preprocessing may work well in some places but not others, b) smoothing may affect the temporal pattern of vegetation seasonality and may thus add a bias to the LSP estimations, and c) defining the smoothing functions and their parameters requires expert knowledge and makes their implementation time consuming, although some program packages such as SPIRITS (Eerens & Haesen, 2013) and TIMESAT (Tan et al., 2010) include well-implemented methods.

The choice of the preprocessing steps is critical, because the robustness of the results of LSP methods greatly depends on their quality (Zeng et al., 2020). The complexity of these preprocessing steps also impedes the implementation of LSP methods in GEE. The aim of this study was to develop a variant of the threshold method that (i) does not require preprocessing steps and thus does not need excessive user-defined parameters, (ii) has fast computing time and can be easily implemented in cloud-processing platforms, (iii) can be directly applied to daily remotely sensed observations and still compares well with LSP metrics estimated with ground data.

1.2. Data and Methods

1.2.1. LSP extraction methods

1.2.1.1. Maximum separation method

The method we propose, which we named Maximum Separation method (MS), is a variant of the threshold method. Given a time series of paired values *{t, b}*, where *t* is the time in days and

b is the vegetation activity represented by a biophysical variable, a vegetation index, a reflectance band, or any measure of greenness (Figure 1.1a), we defined a threshold value *u* dynamically (White et al., 1997) with Eq. 1. The threshold is dynamic because it is determined by a percentage (*p*) of the amplitude (difference between the minimum (b_{min}) and maximum (b_{max}) values of the season)

$$u = (b_{max} - b_{min})\, p + b_{min} \quad (1)$$

The threshold *u* is used to classify binarily the time series into 0, when $b < u$, and 1, when $b > u$. The binary time series *b'* (Figure 1.1b) has values equal to zero during the dormant period and values equal to one during the growing season. The method then runs a moving window operation for each day of the binary time series. The operation calculates the difference between the proportion of observations with vegetation activity (*b´= 1*) before and after the central day of the moving window (Eq. 2):

$$d = \sum b'_{before} \Big/ n_{before} - \sum b'_{after} \Big/ n_{after} \quad (2)$$

where n_{before} and n_{after} are the total number of observations before and after the central day of the moving window, and b'_{before} and b'_{after} are the number of observations above the threshold before and after the central day.

The moving window generates a new time series *{t, d}*, with minimal values close to -1 when the time series exceeds the threshold and maximal values close to 1 at the end of the season (Figure 1.1c). SoS (t_{sos}) corresponds to the day when d is lowest, and EoS (t_{eos}) corresponds to the day when *d* is highest. We named the method Maximum Separation since the SoS and EoS represent the days when the difference between observations above the threshold before and after these dates is maximal.

The running window in Eq. 2 can be applied to the entire time series and, then, the minimum and maximum ratios can be searched within a natural year (e.g. from 1 January to 31 December) or, in case of double seasonality, within a specific range of months. This property simplifies the LSP estimation in the Southern Hemisphere, where the main growing season usually occurs from November to February.

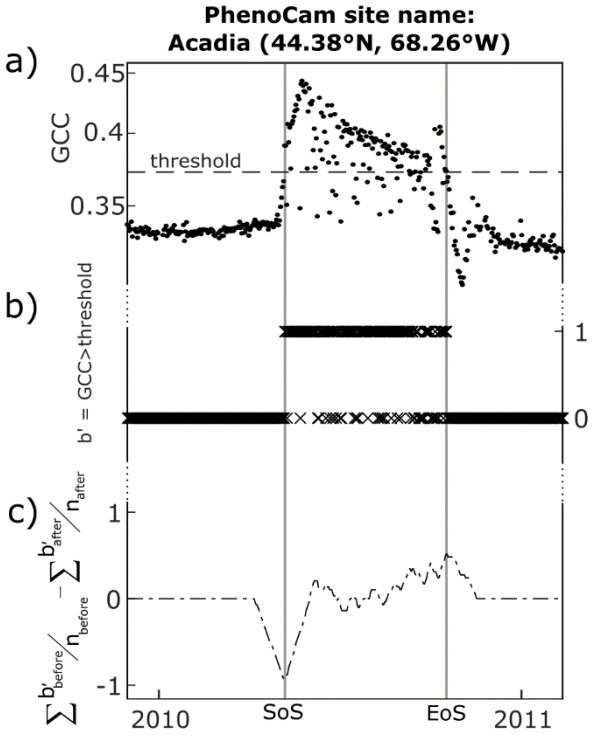


Figure 1.1. Estimation of the start of season (SoS) and end of season (EoS) using maximum separation at the Acadia PhenoCam deciduous forest site in 2010. a) The time series of Green Chromatic Coordinate Index (GCC), b) binary classification of the time series (0 = dormant period, 1 = growing season), and c) time series obtained by applying a moving window that estimates the difference in the proportion of growing season observations before and after a given day. SoS and EoS correspond to the minimum and maximum values in c), respectively. We used a dynamic threshold of 50% of the amplitude. The semi-period of the moving window when calculating the proportions is 30 days.

1.2.1.2. Threshold method

The threshold method assigns the SoS (t_{sos}) to the first day when the seasonal time series b exceeds a threshold *u* after the dormant period, while EoS (t_{eos}) is the last day of the season when *b* is greater than the threshold *u*. The threshold is computed dynamically with Eq. 1. The threshold method is inherently affected by noise in the time series and, thus, smoothing techniques are required before its application (Zeng et al., 2020). Here, we used three preprocessing methods:

1) The first preprocessing method (TH1) used a variant of the so-called maximum composite value (Holben, 1986). The maximum composite value estimates the maximum value of a moving window. We used the 80th percentile instead, which was more robust against outliers. We estimated the composite every 8 days. The snow observations were replaced with the 5th percentile value obtained from the entire snow-

free time series. This reclassification aims to reduce the sharp increase in the vegetation indices, particularly in NDVI, during the transition from snow to snow-free period, which may lead to unreliable estimate of the Start of Season. This approach has been adopted for LSP estimation in high-latitude biomes (Beck et al., 2007) and at the global scale (Gray et al., 2019). Finally, we applied a linear interpolation to convert the composite time series to daily estimates.

2) The second preprocessing method (TH2) involved an outlier rejection algorithm that excluded low values, a smoothing step with the Savitzky–Golay filter (Chen et al., 2004), and a linear interpolation. The outlier rejection excluded observations that exceeded a level of tolerance to adjacent observations. The tolerance was 0.2 times the interpolated value for the given observation. The snow-contaminated values were replaced with the 5th value of the snow-free time series and the cloud-contaminated values were filled by a linear interpolation. The Savitzky–Golay filter was then applied with a second polynomial order and a frame length of 21 days. The parameters of the Savitzky–Golay filter were set by trial and error (See an example of the compositing in Supplementary figure 1.1). These parameters have a strong influence in the reconstruction of the time series.
3) The last preprocessing method (TH3) was a curve-fitting method. We fitted a logistic function (Eq. 3) to the first half of the season and another to the second half:

$$f(t;a,b,c) = \frac{b}{1+e^{-a(t-c)}} \tag{3}$$

where t is the Day of the Year; a, b and c are the parameters to be fitted: a determines the steepness of the curve, b is the maximum value of the time series, and c represents the maximum increase or decrease in the time series and is commonly associated with the SoS and EoS. We used the curve fitting only as a smoothing step, and the phenological metrics were extracted using the threshold method over the fitted function.

1.2.2. Data

1.2.2.1. MODIS data

We extracted the phenology metrics SoS and EoS to the following MODIS products: 1) the MCD15A3H leaf area index (LAI) product (Myneni et al., n.d.), which has a temporal resolution

of four days at 500 meters of spatial resolution, and 2) the daily MOD09GA product (Vermote & Wolfe, 2020), also at 500 meters, from which we estimated the NDVI, the EVI, and the Green Chromatic Coordinate Index (GCC). The MOD09GA data is available from the year 2000, while MCD15A3H from 2002. The NDVI, EVI, and LAI are among the most commonly used vegetation indices in phenology studies (Zeng et al., 2020), while GCC is a vegetation index available in the PhenoCam dataset that has recently taken relevance in phenology studies because it is invariant to cloud shadows (Sonnentag et al., 2012) . The NDVI, EVI, and GCC are estimated directly from the spectral bands (Supplementary Table 1.1) and reflect the amount of live green vegetation in a pixel. These indices correlate with LAI, which is a biophysical variable that indicates the one-side green leaf area per pixel (m^2m^{-2}). The LAI product in MCD15A3H product is estimated from modeled radiances (Knyazikhin, 1999).

1.2.2.2. PhenoCam dataset

The LSP metrics estimated with the MODIS products were validated with the PhenoCam Dataset v2.0 (Seyednasrollah et al., 2019). PhenoCam is a network of digital cameras that records images of the vegetation at high temporal resolution (commonly 30 minutes), at close range, and for a diverse range of ecosystems. The time series of images recorded by the digital cameras provide information on the seasonal changes in vegetation greenness that can be used to validate the phenology metrics extracted from remotely-sensed satellite data. PhenoCam has been previously used for validation LSP estimated from MODIS data (Hufkens et al., 2012), VEGETATION and PROBA-V (Bórnez et al., 2020), and Landsat-8 and Sentinel-2 (Bolton et al., 2020).

We considered all the sites of the PhenoCam dataset, but rejected the sites that presented continuous gaps in the time series (with gaps >30 days) or years with missing data. A total of 212 sites were considered in the study; 9 sites in evergreen needleleaf forests (ENF), 23 sites in deciduous broadleaf forests (DBF), 77 sites in mixed forests (MX), 7 sites in open shrublands (OSH), 6 sites in woody savannah (OSH), 3 sites in savannah (SAV), 23 sites in grasslands (GRA), 48 sites in temporary crops (CRO1), and 16 sites in cropland/natural vegetation mosaics (CRO2) (See Supplementary Table 1.2 for more metadata information of the PhenoCam sites and location map in Supplementary figure 1.2). Sites that presented time series with consistent gaps were rejected from the analysis.

PhenoCam Dataset v2.0 provides not only near-surface images but also the GCC-generated time series for certain areas of the image. The different regions observed by the digital camera may cover different types of vegetation that are presented in the site. We used the GCC time series of the primary vegetation type when more than one vegetation type was present in the PhenoCam site. The time coverage of PhenoCam data depends on the date when the digital camera was installed and, therefore, differs from one site to another; the PhenoCam that provided the earliest data was for year 2001, and the latest images were taken in 2018. The temporal coverage of the PhenoCam sites overlaps with the MODIS time series; 2000-2020 for MOD09GA and 2002-2020 for MCD15A3H).

1.2.3. Experimental setup

The phenology metrics, SoS and EoS, estimated with MODIS GCC, NDVI, EVI, and LAI were compared with the same metrics estimated with the GCC time series in the PhenoCam network. We reported the comparison between the satellite and the near-surface estimates in terms of mean error (ME) and root mean squared error (RMSE) (See formulas in Supplementary Table 1.3). The ME represents a measure of the bias (mean average of the LSP differences PhenoCam – MODIS: positive ME indicates higher PhenoCam values than MODIS) and the RMSE gives an indication of the accuracy of the MODIS in relation to the PhenoCam LSP estimations (the standard deviation of an average MODIS LST estimation from the PhenoCam LSP estimation). The comparison between MODIS and PhenoCam LSP metrics aimed to test the performance of the MS in comparison to the three variants of the threshold method based on two aspects: 1) Threshold percentage (variable *p* in Eq. 1). We estimated the SoS and EoS with *p* values ranging from 10 to 90% in steps of 10%. 2) Primary vegetation type of the Phenocam site. For this case, we reported the ME and RMSE only for a threshold percentage of 50%.

We also tested the influence of the window size but, since we found that this parameter has a marginal effect when it takes large values, we set the window radius to 30 days in all experiments of the study. A window size of 30 days represents a good compromise between robustness of the method against noise and computing time.

Finally, we used Google Earth Engine to generate maps of the SoS and EoS fom the MODIS variables: GCC, NDVI, EVI, and LAI. The phenological maps in GEE were generated with the MS method with a threshold percentage of 50% and a window radius of 30 days. We used the RESOLVE Ecoregions 2017 map (Dinerstein et al., 2017) in order to mask the biomes that show

a negligible seasonality in the selected MODIS variables. The masked biomes were the tropical and subtropical dry broadleaf forests, tropical and subtropical coniferous forests, deserts and xeric shrublands, and mangroves

1.3. Results

The size of the moving window and the threshold percentage are the only parameters used in MS. Supplementary figure 1.3 shows the smoothing effect of the window size (semi-period of 10, 30, and 90 days) over a time series of daily MODIS GCC (Supplementary figure 1.3a) for the Acadia site in the PhenoCam network. The smoothing in the time series derived from Eq. 2 is more remarkable in high window sizes (Supplementary figure 1.3c and 1.3d), and the absolute maximum and minimum, corresponding to the phenology metrics, are well-defined in the time series. A window size of 10 days, however, does not capture the maximum difference between dormant and growth observation (Supplementary figure 1.3b), which in turn results in unrealistic LSP estimations.

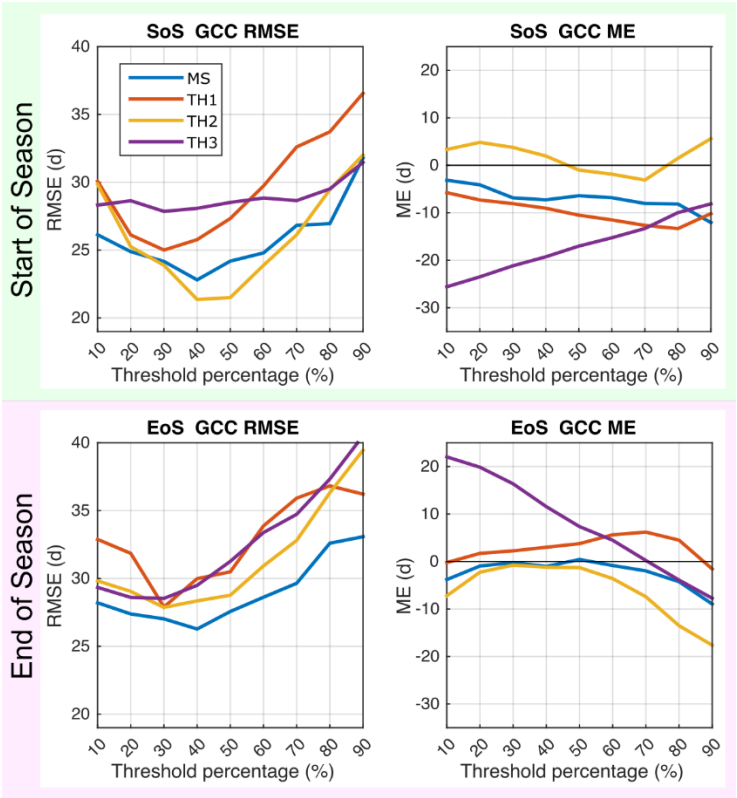


Figure 1.2. Root mean squared error (RMSE) and mean error (ME) obtained from the comparison of the LSP metrics generated with MODIS Green Chromatic Coordinate (GCC) and the PhenoCam dataset for a range of threshold values. The LSP metrics are the Start of Season (SoS) and the End of Season (EoS).

The MS showed similar results as the best threshold method, which was TH2, for all MODIS variables (figure 1.2 and Supplementary figure 1.4). TH2 showed the best results for the SoS, while the MS showed better results for the EoS, although overall differences between LSP methods were marginal and none of the MODIS variables excelled in the comparison with PhenoCam. The best result in terms of RMSE were obtained with a threshold percentage of 40% (SoS RMSE = 19.8 d in TH2 using NDVI; EoS RMSE = 23.0 d in MS using LAI). Threshold percentages that ranged from 30 to 60% showed the best results in terms of RMSE for GCC, NDVI, EVI, and LAI, and low accuracy in extreme percentages, particularly >80%. The SoS showed better results than the EoS; the minimum RMSE ranged from 20 to 25 days in the SoS, while the same statistic was between 25 to 30 days in the EoS for the GCC, NDVI, and EVI. The impact of the threshold was less significant in terms of ME for the SoS, but more apparent in the EoS. TH2 outperformed in SoS with a negligible bias for all variables. MS, TH1, and TH2 showed a similar pattern in the ME regarding the threshold percentage, while the TH3 (logistic fitting) presented a different trend in the ME, particularly for the EoS.

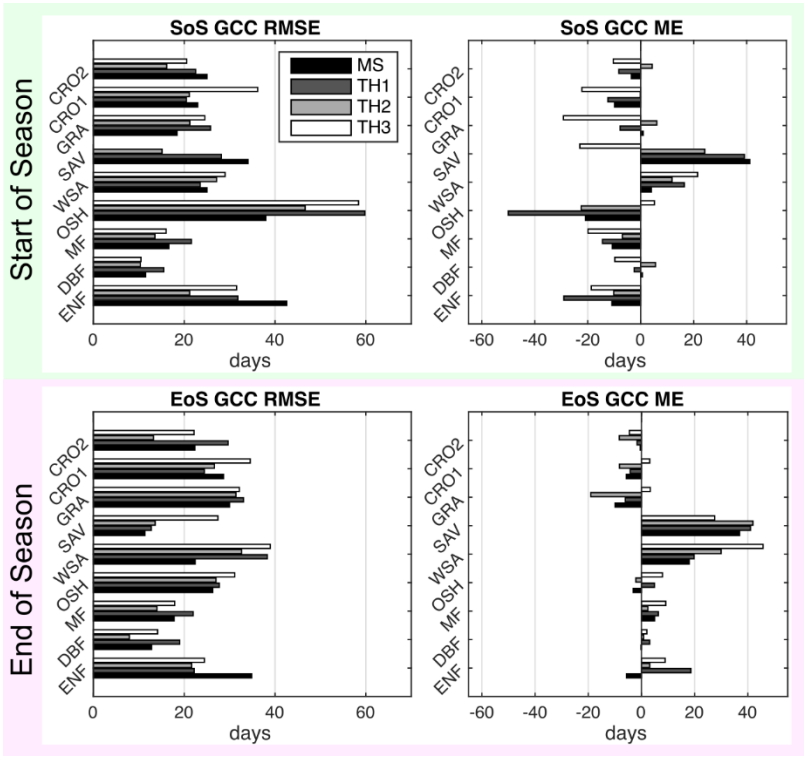


Figure 1.3. Root mean squared error (RMSE) and mean error (ME) obtained from the comparison of the LSP metrics generated with MODIS Green Chromatic Coordinate (GCC) and the PhenoCam dataset. The comparison is based on the land cover of the Phenocam sites: evergreen needleleaf forests (ENF), deciduous broadleaf forests (DBF), mixed forests (MX), open shrublands (OSH), woody savannah (WSA), savannah (SAV), grasslands (GRA), temporary crops (CRO1), and cropland/natural vegetation mosaics (CRO2). The LSP metrics are the Start (SoS) and the End of Season (EoS).

The RMSE between MODIS and PhenoCam LSP metrics depended to a great extent on the type of vegetation. The ME and RMSE differed between land covers in MODIS GCC (figure 1.3) and the rest of MODIS variables (Supplementary Figure 5). The best results were found in DBF and MX in the four methods (e.g., ME = -7.2 d and RMSE = 0.6 d for EoS and ME = -0.2 d and RMSE = 12.8 d for EoS in DBF estimated with MS). Contrarily, on average, land covers that represented evergreen showed a high RMSE

The maps of mean SoS and EoS estimated with MS using the MODIS GCC product for the period 2001-2019 are presented in Figure 1.4, and the maps generated with the NDVI, EVI, and LAI are presented in Supplementary figure 1.6. The four vegetation indices showed similar spatial patterns at the global scale.

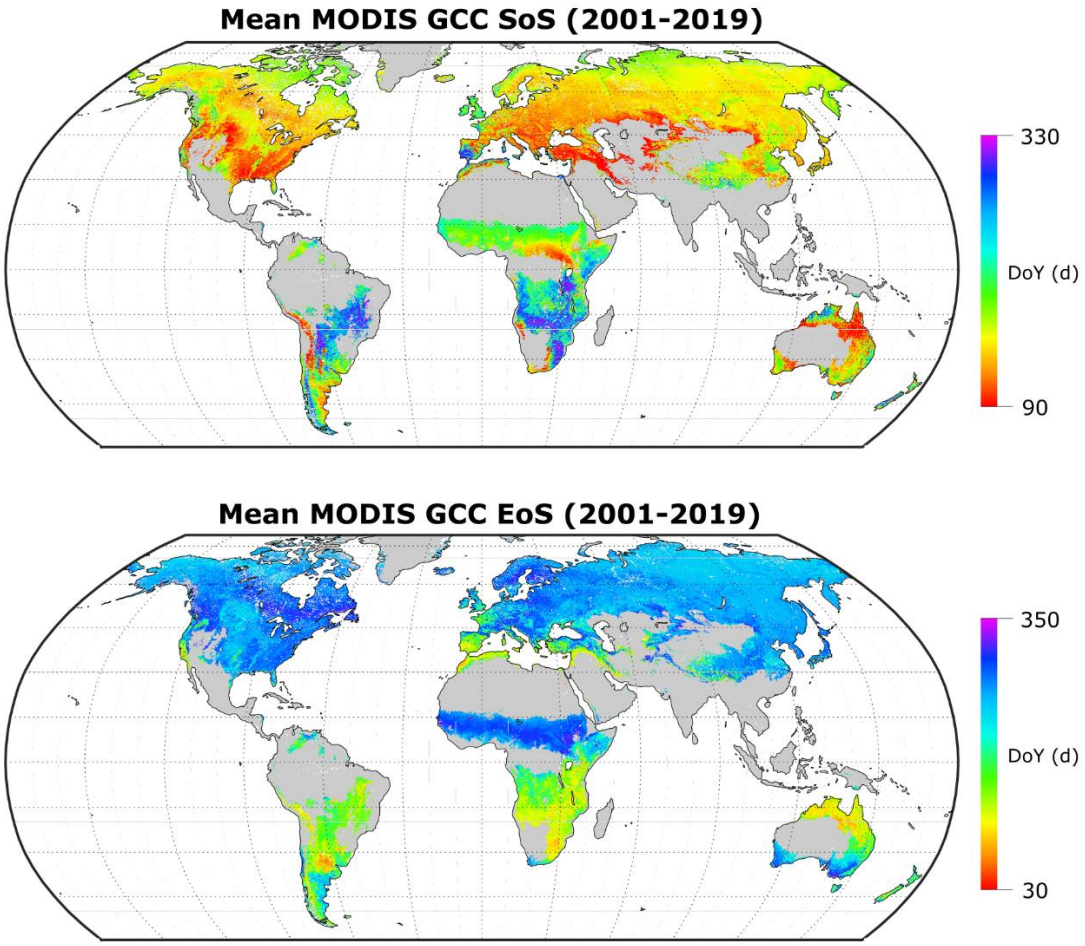


Figure 1.4. Maps of the mean start and end of season (SoS and EoS) estimated using the Maximum Separation method and a dynamic threshold of 50% of the amplitude. The SoS and EoS are shown as the Day of Year (DoY). The Maximum Separation method was applied to the MODIS GCC (MOD09GA).

1.4. Discussion

The proposed MS method showed accuracy similar to the best threshold method. Differences in accuracy were higher among land cover types and threshold values rather than LSP methods and MODIS variables. Although MS did not excel in the comparison with PhenoCam, the proposed method presents an advantage over the threshold method; MS was applied directly to daily non-smoothed time series and with fewer parameters compared to the conventional threshold method with preprocessing steps. The simplicity of the MS method allowed a straightforward implementation in GEE and the production of phenological maps at the global scale. In this way, GEE allowed the rapid processing of custom algorithms and overcame the computing time that would be necessary for processing LSP maps with a local computer.

One of the two parameters of the MS method is the size of the moving window, which can be set with prior knowledge of the length of the growing season. A large window entailed a higher computational cost, but it positively affected the estimation of the phenological metric (Supplementary figure 1.3), because the window covered more observations and the impact of noise was minimized. The main consideration in the selection of the window size is that it should not exceed the length of the growing season, particularly in vegetation types that present a short growing season such as high-latitude shrublands, alpine grasslands, and precipitation-sensitive vegetation.

The second parameter of the MS method is the threshold percentage, which had a strong influence on the accuracy of the LSP estimates (figure 1.2). The MS, and the threshold methods in general, outputs unreliable LSP estimates when the time series present significant noise and outliers. Noise and outliers may lead to unreliable threshold values in Eq. 1, which in turn results in inaccurate LSP estimates. This also explains the good results of the MS and the threshold methods in medium-range threshold percentages (figure 1.2). For the SoS for instance, a low threshold percentage is more likely to detect a noisy high value as the first observation that overpasses the threshold. These results further proves the robustness of medium threshold percentages against clouds, cloud shadows, and other contamination effects (Bolton et al., 2020; Bórnez et al., 2020).

The MS and the variants of the threshold method performed better in deciduous broadleaf forests and mixed forests. Deciduous forests generally show a well-defined seasonality that is reflected in sharp changes in the vegetation indices, particularly during the leaf unfolding period.

On the contrary, in evergreen, changes in greenness indices are marginal because leaves are perennial, which makes LSP estimation challenging, and phenology changes are observed at the photosynthetic level (Yin et al., 2020). Moreover, vegetation indices may not reflect a logistic-type of growth, which explain the poor results of TH3 and a ME that is linearly correlated with the threshold percentage (figure 1.2). (Gray et al., 2019) reported that EVI did not reflect a logistic growth in specific ecosystems and, thus, justified the use of the threshold method for the MODIS Land Cover Dynamics version 6 over the logistic modelling used in the previous version of the product.

1.5. Conclusions

This study proposed a new method (Maximum Separation, MS) ideal for fast and large-scale estimates of phenological metrics in cloud computing platforms such as GEE. The MS presents similar results as the threshold method with time series preprocessing, but the simplicity of its implementation when applied to different satellite products makes MS more pragmatic than the threshold method, particularly when the data require preprocessing for time series smoothing. We provide the GEE code and Python implementation, which can be used in other remotely sensed data. Future studies may thus benefit from this method and customize it for their regional studies.

1.6. References

Beck, P., Jönsson, P., Høgda, K.-A., Karlsen, S., Eklundh, L., & Skidmore, A. (2007). A ground-validated NDVI dataset for monitoring vegetation dynamics and mapping phenology in Fennoscandia and the Kola peninsula. *International Journal of Remote Sensing*, *28*(19), 4311–4330.

Bolton, D. K., Gray, J. M., Melaas, E. K., Moon, M., Eklundh, L., & Friedl, M. A. (2020). Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. *Remote Sensing of Environment*, *240*, 111685.

Bórnez, K., Descals, A., Verger, A., & Peñuelas, J. (2020). Land surface phenology from VEGETATION and PROBA-V data. Assessment over deciduous forests. *International Journal of Applied Earth Observation and Geoinformation*, *84*, 101974.

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 Sensing of Environment*, *91*(3–4), 332–344.

Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A., & Schwartz, M. D. (2007). Shifting plant phenology in response to global change. *Trends in Ecology & Evolution*, *22*(7), 357–365. https://doi.org/10.1016/j.tree.2007.04.003

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., & others. (2017). An ecoregion-based approach to protecting half the terrestrial realm. *BioScience*, *67*(6), 534–545.

Eerens, H., & Haesen, D. (2013). Software for the Processing and interpretation of remotely sensed image time series. *User's Manual, Version*, *1*(1).

Fisher, J. I., Mustard, J. F., & Vadeboncoeur, M. A. (2006). Green leaf phenology at Landsat resolution: Scaling from the field to the satellite. *Remote Sensing of Environment*, *100*(2), 265–279.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27.

Gray, J., Sulla-Menashe, D., & Friedl, M. A. (2019). *User Guide to Collection 6 MODIS Land Cover Dynamics (MCD12Q2) Product*. https://lpdaac.usgs.gov/documents/218/mcd12q2_v6_user_guide.pdf

Holben, B. N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, *7*(11), 1417–1434.

Hufkens, K., Friedl, M., Sonnentag, O., Braswell, B. H., Milliman, T., & Richardson, A. D. (2012). Linking near-surface and satellite remote sensing measurements of deciduous broadleaf forest phenology. *Remote Sensing of Environment*, *117*, 307–321.

Jonsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(8), 1824–1832.

Knyazikhin, Y. (1999). MODIS leaf area index (LAI) and fraction of photosynthetically active radiation absorbed by vegetation (FPAR) product (MOD 15) algorithm theoretical basis document. *Http://Eospso. Gsfc. Nasa. Gov/Atbd/Modistabls. Heml*.

Körner, C., & Basler, D. (2010). Phenology Under Global Warming. *Science*, *327*(5972), 1461–1462. https://doi.org/10.1126/science.1186473

Li, X., Zhou, Y., Meng, L., Asrar, G. R., Lu, C., & Wu, Q. (2019). A dataset of 30 m annual vegetation phenology indicators (1985–2015) in urban areas of the conterminous United States. *Earth System Science Data*, *11*(PNNL-SA-144732).

Ma, M., & Veroustraete, F. (2006). Reconstructing pathfinder AVHRR land NDVI time-series data for the Northwest of China. *Advances in Space Research*, *37*(4), 835–840.

Myneni, R., Knyazikhin, Y., & Park, T. (n.d.). *MCD15A3H MODIS/Terra+ Aqua Leaf Area Index/FPAR 4-day L4 Global 500m SIN Grid V006. 2015*.

Peñuelas, J., & Filella, I. (2001). Responses to a warming world. *Science*, *294*(5543), 793–795.

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

Seyednasrollah, B., Young, A. M., Hufkens, K., Milliman, T., Friedl, M. A., Frolking, S., & Richardson, A. D. (2019). Tracking vegetation phenology across diverse biomes using Version 2.0 of the PhenoCam Dataset. *Scientific Data*, *6*(1), 1–11.

Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A. M., Friedl, M., Braswell, B. H., Milliman, T., O'Keefe, J., & Richardson, A. D. (2012). Digital repeat photography for phenological research in forest ecosystems. *Agricultural and Forest Meteorology*, *152*, 159–177. https://doi.org/10.1016/j.agrformet.2011.09.009

Tan, B., Morisette, J. T., Wolfe, R. E., Gao, F., Ederer, G. A., Nightingale, J., & Pedelty, J. A. (2010). An enhanced TIMESAT algorithm for estimating vegetation phenology metrics from MODIS data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *4*(2), 361–371.

Vermote, E., & Wolfe, R. (2020). *MOD09GA MODIS/Terra Surface Reflectance Daily L2G Global 1kmand 500m SIN Grid V006. 2015, distributed by NASA EOSDIS Land Processes DAAC*.

Viovy, N., Arino, O., & Belward, A. (1992). The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series. *International Journal of Remote Sensing*, *13*(8), 1585–1590.

Walt, S. van der, Colbert, S. C., & Varoquaux, G. (2011). The NumPy array: A structure for efficient numerical computation. *Computing in Science & Engineering*, *13*(2), 22–30.

White, M. A., Thornton, P. E., & Running, S. W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles*, *11*(2), 217–234.

Yin, G., Verger, A., Filella, I., Descals, A., & Peñuelas, J. (2020). Divergent Estimates of Forest Photosynthetic Phenology Using Structural and Physiological Vegetation Indices. *Geophysical Research Letters*, e2020GL089167.

Zeng, L., Wardlow, B. D., Xiang, D., Hu, S., & Li, D. (2020). A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sensing of Environment*, *237*, 111511.

Zhang, X., Friedl, M. A., & Schaaf, C. B. (2006). Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements. *Journal of Geophysical Research: Biogeosciences*, *111*(G4).

Chapter 2

Improved estimates of Arctic Land Surface Phenology using Sentinel-2 time series

Adrià Descals, Aleixandre Verger, Gaofei Yin, and Josep Peñuelas

Published in *Remote Sensing* 12.22 (2020): 3738. https://doi.org/10.3390/rs12223738

Abstract

The high spatial resolution and revisit time of Sentinel-2A/B tandem satellites allow a potentially improved retrieval of land surface phenology (LSP). The biome and regional characteristics, however, greatly constrain the design of the LSP algorithms. In the Arctic, such biome-specific characteristics include prolonged periods of snow cover, persistent cloud cover, and shortness of the growing season. Here, we evaluate the feasibility of Sentinel-2 for deriving high-resolution LSP maps of the Arctic. We extracted the timing of the start and end of season (SoS and EoS, respectively) for the years 2019 and 2020 with a simple implementation of the threshold method in Google Earth Engine (GEE). We found a high level of similarity between Sentinel-2 and PhenoCam metrics; the best results were observed with Sentinel-2 enhanced vegetation index (EVI) (root mean squared error (RMSE) and mean error (ME) of 3.0 d and –0.3 d for the SoS, and 6.5 d and –3.8 d for the EoS, respectively), although other vegetation indices presented similar performances. The phenological maps of Sentinel-2 EVI compared well with the same maps extracted from the Moderate Resolution Imaging Spectroradiometer (MODIS) in homogeneous landscapes (RMSE and ME of 9.2 d and 2.9 d for the SoS, and 6.4 and –0.9 d for the EoS, respectively). Unreliable LSP estimates were filtered and a quality flag indicator was activated when the Sentinel-2 time series presented a long period (>40 d) of missing data; discontinuities were lower in spring and early summer (9.2%) than in late summer and autumn (39.4%). The Sentinel-2 high-resolution LSP maps and the GEE phenological extraction method will support vegetation monitoring and contribute to improving the representation of Artic vegetation phenology in land surface models.

2.1. Introduction

Studies of the Arctic are becoming increasingly important, particularly in the context of the onset of an expected tipping point in the function of its ecosystems as a result of ongoing climate warming (Post et al., 2019) that is lengthening the growing season and increasing vegetation productivity (Park et al., 2016; H. Zeng et al., 2011). Rising levels of vegetation productivity in the Arctic, characterized as greening, result in reduced albedo that further drives the warming trend (Chapin et al., 2005; Peñuelas & Filella, 2009) and has consequences for the regional carbon cycle (Schuur et al., 2009). For example, permafrost is thawing at an accelerating pace in a process that releases greenhouse gases and, consequently, exacerbates the warming trend in the region (Schuur et al., 2015).

Land surface phenology (LSP), the study of remotely-sensed seasonal patterns in vegetation growth (Helman, 2018), complements the sparse field observations at high latitudes and is essential in the assessment and monitoring of responses of arctic vegetation to climate warming. Reliable LSP maps of the Arctic could be used to support models that show positive feedbacks between trends in climate warming and a lengthening growing season (H. Zeng et al., 2011), while studies of Arctic LSP may reveal a link between the advancement of spring onset, drought severity, and an increase in the incidence of artic fires over recent decades (Witze, n.d.), and support reports of links between carbon uptake and phenology metrics related to vegetation greenness, such as the amplitude of the vegetation index (Myers-Smith et al., 2020).

Studies of Arctic LSP have tended to use moderate-resolution satellite data, such as the Advanced Very-High-Resolution Radiometer (AVHRR) (1.1 km) and the Moderate Resolution Imaging Spectroradiometer (MODIS) (500 m) (Park et al., 2016; H. Zeng et al., 2011). At these high latitudes, there are two key challenges for LSP analysis (Myers-Smith et al., 2020). Firstly, LSP methods tend to erroneously detect the end of a snow period as the start of season (SoS), particularly in methods that extract phenology metrics from greenness indices, such as the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI), where distinct changes in reflectance during the snowmelt period are incorrectly detected as the onset of vegetation growth. To account for this problem, vegetation indices that are insensitive to snow have been developed to improve phenology estimation (Jin & Eklundh, 2014; Wang et al., 2017). Secondly, a lack of valid satellite observations due to persistent cloud cover hampers LSP

estimation, so biome-specific algorithms based on the combination of multi-satellite data and spatio-temporal gap filling methods have been developed (H. Zeng et al., 2011).

Recent advances in remote sensing technologies present opportunities for the estimation of LSP at greater spatial resolution. Sentinel-2 mission provides decametric images with frequent revisit times (<5 d), allowing the extraction of phenology metrics (Bolton et al., 2020) and study of vegetation dynamics at the canopy scale, while the development of cloud-based platforms, such as Google Earth Engine (GEE) (Gorelick et al., 2017), allows processing of large volumes of satellite data for planetary-scale analysis (Hansen et al., 2013; Pekel et al., 2016) and increases the accessibility of high-resolution satellite archive data required for time series analyses.

The objectives of this study were (1) evaluating the feasibility of Sentinel-2 for LSP retrieval in the Arctic at a spatial resolution of 10 m, (2) proposing a novel and fast cloud computing implementation in GEE of the widely used threshold phenological extraction method, and (3) assessing the performances of the Arctic Sentinel-2 LSP maps for the SoS and end of season (EoS) for the years 2019 and 2020 based on the comparison with MODIS LSP and PhenoCam ground data.

2.2. Materials and Methods

2.2.1. Study area

We generated SoS and EoS maps over regions classified as tundra in the RESOLVE Ecoregions dataset 2017 (Dinerstein et al., 2017). High latitudes of the Arctic, where winters are cold and summers are short, are mostly uninhabited by human populations; here, tundra soils contain a layer of permafrost that prevents the growth of trees, but supports the growth of grass and shrub vegetation during the short summer period between June and August, as revealed by vegetation index time series (e.g., NDVI and EVI).

2.2.2. Data

We generated LSP metrics for 2019 and 2020 using Sentinel-2 level-2A data, which provide daily top-of-canopy reflectance at 10, 20, and 60 m of spatial resolution; from these data, we used the 10-meter resolution bands 2, 3, 4, and 8, and the 20-meter band 12. Sentinel-2A and -2B multispectral satellites were launched in 2015 and 2017, respectively, and have a revisit time of

5 d at the equator that decreases with increasing latitude. The maximum revisit time between 1 May and 30 September 2019 for latitudes between 70 and 75° was 1.7 d (average: 0.9 d) (Figure 2.1a). The LSP metrics estimated with Sentinel-2 were compared with the same metrics estimated from the 500 m MOD09GAv6 product (Vermote & Wolfe, 2015).

Maximum discontinuity in the time series from spring and early summer (1 May to 15 July) and later summer and autumn (15 July to 30 September) was plotted after non-valid observations had been filtered using values of the quality band scene classification layer (SCL) provided in the Sentinel-2 level-2A, where 1 = saturated or defective; 2 = dark area pixels; 3 = cloud shadows; 6 = water; and, 7−10 = clouds and cirrus (Figure 2.1b,c).

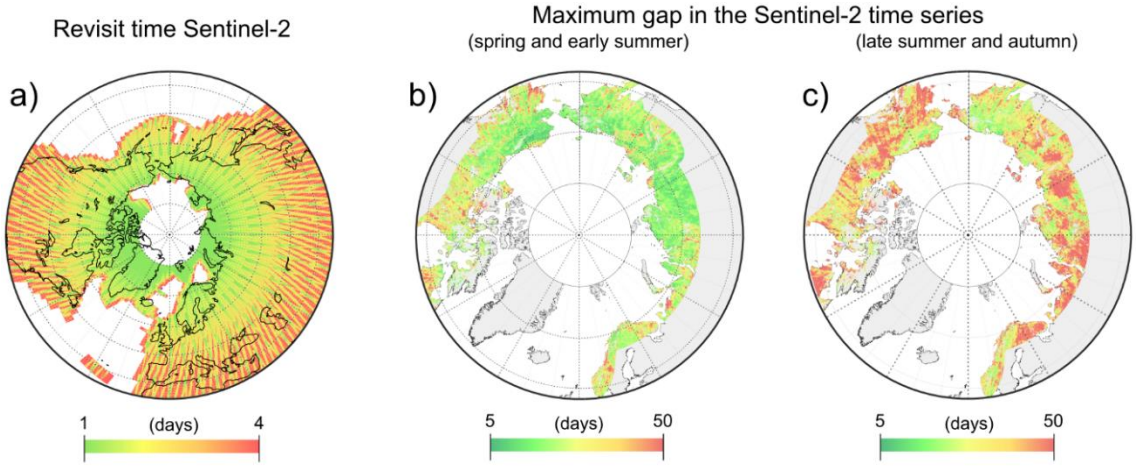


Figure 2.1. Revisit time of Sentinel-2 (a) and maximum discontinuity in the Sentinel-2 Level 2A time series after cloud masking for spring and early summer (1 May to 15 July) (b) and late summer and autumn (15 July to 30 September) 2019 (c).

2.2.3. Phenology extraction

We adopted a widely used threshold method (Bolton et al., 2020; Bórnez et al., 2020; Gray et al., 2019) that assigns the SoS and EoS as the first and last days of the season, respectively, on which a threshold *u* is exceeded; *u* may be a constant or defined dynamically for each pixel (L. Zeng et al., 2020). In this study, we estimated *u* as a dynamic value that depends on the annual amplitude of the time series (Equation (1)):

$$u = (V_{min} - V_{max}) \times p + V_{min}, \quad (1)$$

where V_{min} and V_{max} are the minimum and maximum annual values in the time series, respectively, and p is a given proportion (%) of the amplitude. In this study, we used p = 0.5 as the mid-greenup and mid-greendown of the growing season. The threshold metrics estimated with 50% of the amplitude are less affected by biases due to discontinuities in the time series (Bolton et al., 2020). We applied two variants of the threshold method as follows:

1. Threshold method without smoothing. The threshold method was applied directly to the daily time series. For the SoS, we searched for the earliest date when the daily vegetation index exceeded u; then, we applied a linear interpolation between this first observation when the vegetation index was >u and the preceding observation from which to estimate SoS as the value >u (Figure 2.2c). For the EoS, the linear interpolation was applied between the latest date when the vegetation index was >u, and the subsequent observation, where EoS corresponded to the linearly interpolated value >u.
2. Threshold method after smoothing. The time series data were smoothed prior to the extraction of LSP metrics (Figure 2.2d), as is common practice in LSP estimation to reduce noise and discontinuities of time series data (L. Zeng et al., 2020). The criteria for selection of the processing steps were based on the feasibility of their implementation in GEE, without comprising the recreation of the phenology curve (see GEE code in Supplementary Materials). Excessive smoothing of time series may lead to unrealistic recreations of the growing season. We first applied a moving average window, with an average radius of 10 d, every 20 d (Figure 2.2d); if a pixel in the 20 d composite window was empty due to a lack of valid observations, the window size was increased to 40 d. Next, we applied a cubic interpolation to convert the 20 d composites to a daily time series. The threshold was estimated from the amplitude of the interpolated time series, rather than with daily observations, and then the SoS and EoS were estimated as the first and last days, respectively, that exceeded the dynamic threshold in the interpolated time series.

2.2.4. Sentinel-2 vegetation indices

We extracted the phenology metrics from four vegetation indices: the green chromatic coordinate (GCC) (Equation (2)) (Sonnentag et al., 2012), NDVI (Equation (3)), EVI (Equation (4)) (Huete et al., 2002), and normalized difference phenology index (NDPI) (Equation (5)) (Wang et al., 2017). These spectral indexes reflect the greenness of vegetation and are calculated from

different spectral bands in Senteinl-2: Blue (Band 2), Green (Band 3), Red (Band 4), near-infrared (NIR) (Band 8), and shortwave-infrared (SWIR2) (Band 12). The formulas of these vegetation indices are:

$$\mathrm{GCC} = \frac{\mathrm{Green}}{\mathrm{Blue} + \mathrm{Green} + \mathrm{Red}}, \quad (2)$$

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{Red}}{\mathrm{NIR} + \mathrm{Red}}, \quad (3)$$

$$\mathrm{EVI} = 2.5 \frac{(\mathrm{NIR} - \mathrm{Red})}{\mathrm{NIR} + 6 \times \mathrm{Red} - 7.5\, x\, \mathrm{Blue} + 1}, \text{and} \quad (4)$$

$$\mathrm{NDPI} = \frac{\mathrm{NIR} - (\mathrm{alpha}\, x\, \mathrm{Red} + (1 - \mathrm{alpha})\, x\, SWIR2)}{\mathrm{NIR} + (\mathrm{alpha}\, x\, \mathrm{Red} + (1 - \mathrm{alpha})\, x\, SWIR2)} \quad (5)$$

where alpha was originally set to 0.74 for MODIS (Wang et al., 2017), but re-estimated to 0.51 for Sentinel-2 in the current study (see section Estimation of the optimal alpha in NDPI in Supplementary Materials, Supplementary Figure 2.1). The NDPI was specifically designed to cope with the snow observations in LSP studies.

2.2.4.1. Reclassification of snow observations in green chromatic coordinate (GCC), normalized difference vegetation index (NDVI), and enhanced vegetation index (EVI)

NDVI and EVI show a sharp increase after snowmelt in tundra vegetation (Figure 2.2a,b), because the previously snow-covered vegetation canopy is predominantly evergreen. The break in the time series attributed to the transition snow-to-vegetation and vegetation-to-snow is problematic in the estimation of LSP metrics. The SoS and EoS can be erroneously assigned to the snow transition dates instead to the actual vegetation dynamics. The reclassification of snow and post-thaw values is a common practice in Arctic LSP studies (Beck et al., 2007; Bolton et al., 2020) to ensure a consistent threshold value (Equation (1)). We reclassified the snow observations to a fixed minimum value, which is specific for each vegetation index: GCC_{min} = 0.31, $NDVI_{min}$ = 0.39, EVI_{min} = 0.2 (See histograms in Supplementary Figure 2.2), and $NDPI_{min}$ = 0.24 (Supplementary Figure 2.1). This parameter was estimated as the mean value of the first Sentinel-2 snow-free observation of the year, extracted from 400 points randomly distributed. For the NDPI, the $NDPI_{min}$ corresponded to the mean NDPI value of snow observations for the optimal *alpha* (Supplementary Figure 2.1). The pixels that presented a maximum vegetation

index in the time series below the minimum snow value were masked as non-vegetated pixels and the phenology metrics were not computed.

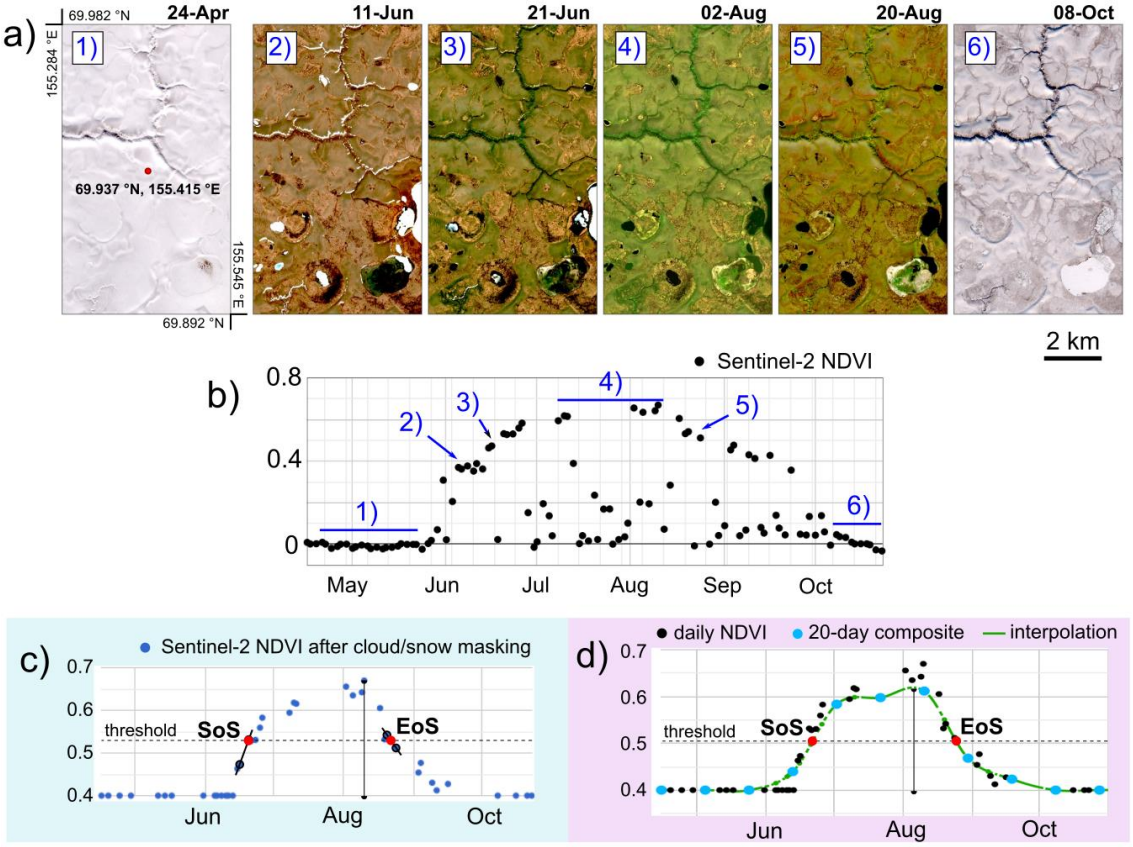


Figure 2.2. Illustration of vegetation dynamics for a randomly selected pixel in the Arctic. (a) True color compositions of Sentinel-2 time series data during 2019; (b) Sentinel-2 normalized difference vegetation index (NDVI) time series during 2019, with reference to true color compositions shown in (a); (c) Sentinel-2 NDVI time series after cloud masking and snow reclassification, where start of season (SoS) and end of season (EoS) correspond to the earliest and latest linearly interpolated NDVI values, respectively, above a threshold defined dynamically as 50% of the annual amplitude; and, panel (d) application of the threshold method over a smoothed and interpolated time series.

2.2.5. Implementation of land surface phenology algorithms in Google Earth Engine

The two variants of the threshold method were vectorized for their correct implementation in GEE. The vectorization is the process of transforming a code so that all components of an array are processed simultaneously (Walt et al., 2011). This concept is contrary to the commonly used practice in LSP estimation, in which time series are processed separately, pixel by pixel, in a *for* loop. The use of *for* loops are, however, highly discouraged in GEE in preference for the recommended *map* functions. For instance, the moving average for the 20-day composition was

implemented as a function that maps a list of dates (see code in Supplementary Materials). The function takes the dates as the input arguments to filter the Sentinel-2 collection with a window size of 20 days and, then, makes the average out of the selected images. The purpose of the map function is that each element of the array, in this case the dates, is processed separately and, consequently, each 20-day composition is generated at the same time.

2.2.6. Validation with PhenoCam

The SoS and EoS metrics extracted from Sentinel-2 were compared with the same metrics estimated from the near-surface reflectances of the PhenoCam network (Richardson et al., 2018). PhenoCam provides half-hourly images captured from digital cameras for 393 sites across North America and Europe, from which 15 cameras cover the tundra biome. At the moment of the analysis, GEE provided Sentinel-2 data from 2019 onwards, and only three sites presented available data in the Arctic for the years 2019 and 2020. The PhenoCam data used in the study was presented as provisional near-real time and subject to changes. The coordinates of these three PhenoCam cameras, used in the study, are shown in Supplementary Table 2.1. PhenoCam also provides daily time series of GCC derived from different regions of interest observed by the digital camera (Richardson et al., 2018). Such regions of interest cover the vegetation types in the site. For the selected sites, however, the camera only observed one vegetation type consisting of tundra grasses.

The GCC has been proved a good index for LSP estimation (Sonnentag et al., 2012) since the time series does not present the sharp increase during the snowmelt. However, we also reclassified the snow values in PhenoCam GCC to the minimum value of the snow-free time series in order to be consistent with the forcing applied in the other Sentinel-2 vegetation indices. The snow observations in GCC were identified and selected by their similar values before and after the growing season (see Supplementary Figures 3–5).

We extracted the LSP metrics with a 50% threshold method from the daily GCC time series for the years 2019 and 2020 and compared them with the same LSP metrics estimated from the four Sentinel-2 vegetation indices (GCC, NDVI, EVI, and NDPI). The statistics that we reported were the mean error (ME) as the bias metric and the root mean squared error (RMSE) as the accuracy metric. Although the selected PhenoCam sites cover a homogeneous area, the location of the sites were manually relocated to ensure consistency with the area observed by the digital

camera. The adjustment of the site coordinates was based on the orientation of the camera; the coordinates were relocated to 50 m from the original coordinates in the direction of the observation of the camera.

2.2.7. Comparison with the Moderate Resolution Imaging Spectroradiometer (MODIS) Land Surface Phenology

Our estimates of SoS and EoS using Sentinel-2 time series were compared with estimates based on MODIS time series data, using the same metrics and LSP extraction method; we also applied the reclassification of snow values to the MODIS time series. To compare both satellite datasets, we resized the 10 m Sentinel-2 LSP estimates to the 500 m MODIS projection. The Sentinel-2 was aggregated by averaging the pixels that lie within a 500 m MODIS pixel. The resampled SoS and EoS estimated with Sentinel-2 and the original MODIS LSP metrics, both at 500 m, were compared pixel-wise and the ME and RMSE were reported. To ensure reliable estimates in MODIS, we only considered the MODIS LSP estimates obtained from time series that presented a gap lower than 10 days. This condition did not result in an excessive filtering of pixels since the MODIS time series had a maximum gap that generally did not exceed 20 days (Supplementary Figure 2.6).

To test the reliability of our MODIS LSP estimates, we compared the SoS and EoS with the 'Mid_greenup' and 'Mid_greendown' layers, respectively, of the MODIS Land Cover Dynamics product (MCD12Q2v6) (Friedl et al., 2019). The MCD12Q2v6 product was also estimated using the threshold method, with a dynamic threshold of 50% in the 'Mid_greenup' and 'Mid_greendown' layers (Gray et al., 2019). MCD12Q2v6 data for 2019 were unavailable at the time of this analysis, so we compared our MODIS LSP estimates for 2018 with the MCD12Q2v6 layers for 2018.

2.3. Results

The Sentinel-2 vegetation index that showed the best results in the comparison with PhenoCam was the EVI (SoS and EoS ME: –0.3 and –3.8 d, and SoS and EoS RMSE: 3.0 and 6.5 d) for the threshold method without time series smoothing (Figure 2.3). Results with the threshold method with smoothing were less conclusive (Supplementary Figure 2.7); RMSE and ME results were uneven between SoS and EoS for the same LSP method and none of the vegetation indices

excelled in both ME and RMSE. Overall, the four Sentinel-2 vegetation indices compared well with PhenoCam, and the RMSE and ME were generally below 10 d for the two threshold methods (see time series of PhenoCam and Sentinel-2 in Supplementary Figures 3–5).

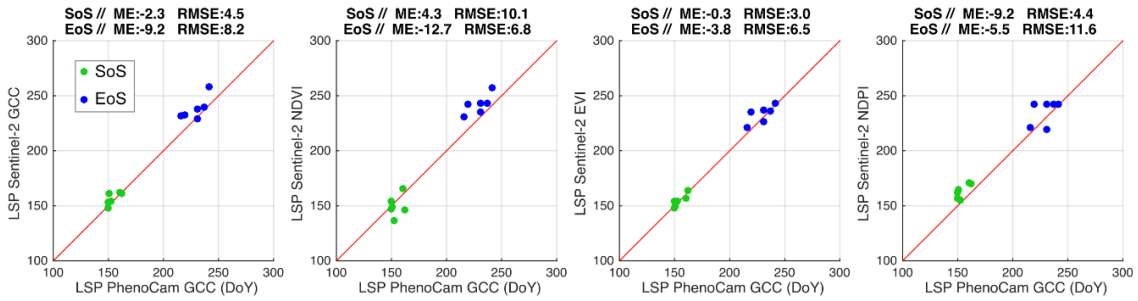


Figure 2.3. Comparison of start of season (SoS) and end of season (EoS) dates between PhenoCam and four vegetation indices estimated with Sentinel-2. The comparison was performed for three PhenoCam sites in the tundra biome for the years 2019 and 2020. The vegetation indices were the green chromatic coordinate (GCC) in PhenoCam and the GCC, normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), and normalized difference phenology index (NDPI) for Sentinel-2. The phenology metrics were extracted with a 50% threshold method without time series smoothing. The bias between PhenoCam and Sentinel-2 is reported with the mean error (ME) and the accuracy with the root mean squared error (RMSE).

LSP maps generated using Sentinel-2 EVI (Figure 2.4a,b) showed a high level of similarity at the continental scale compared with the same phenology metrics estimated with MODIS time series (Supplementary Figure 2.8). However, at the local scale, differences were apparent particularly in the EoS; similarly, the map of length of season (LoS) (Figure 2.5a), which represented the difference between EoS and SoS, showed spatial patterns and latitudinal gradients that are expected across the region, with shorter LoS in the surrounding areas of the Kara Sea and Arctic Archipelago (Figure 2.5b), and in elevated areas, such as the Central Siberian Plateau (Figure 2.5c). At the local level, the SoS and EoS maps derived from Sentinel-2 showed a high degree of detail that moderate scales of resolution were unable to capture. For instance, Figure 2.5d shows the positive relationship between LoS and elevation in the Scandinavian mountains. The altitudinal gradient is observed at the canopy scale, which allows the analysis of LSP metrics by vegetation type. Figure 2.5e shows another example of the effect of topography on the LoS in the Ural Mountains. In this second example, the LoS is different in both sites of the mountain range depending on the presence of glaciers. Supplementary Figure 2.9 shows the different spring growth onset depending on the land cover type, tundra shrublands and herbaceous cover in the delta of the Lena River.

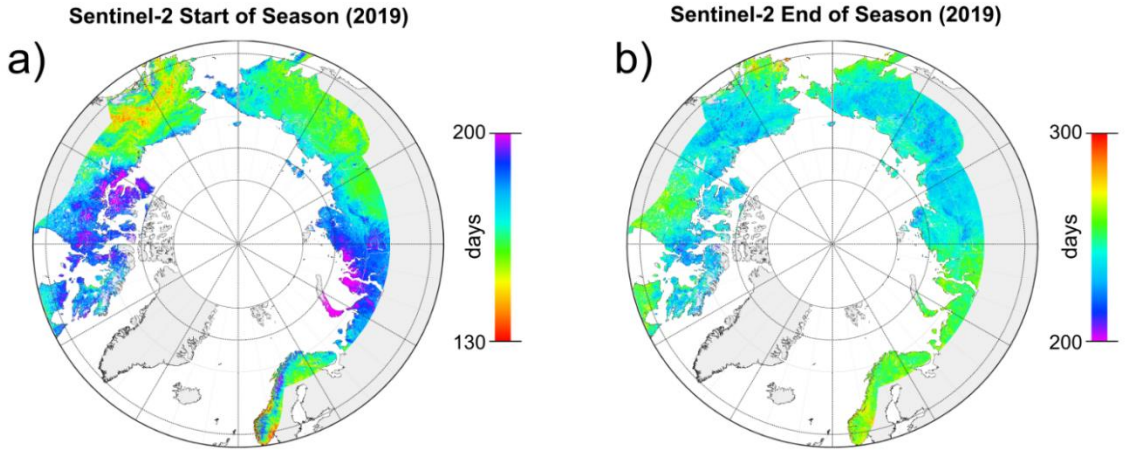


Figure 2.4. Maps of (a) start of season (SoS) and (b) end of season (EoS) in the Arctic estimated using Sentinel-2 for 2019. The phenology metrics were extracted with the 50% threshold method applied to the non-smoothed Sentinel-2 enhanced vegetation index (EVI) time series.

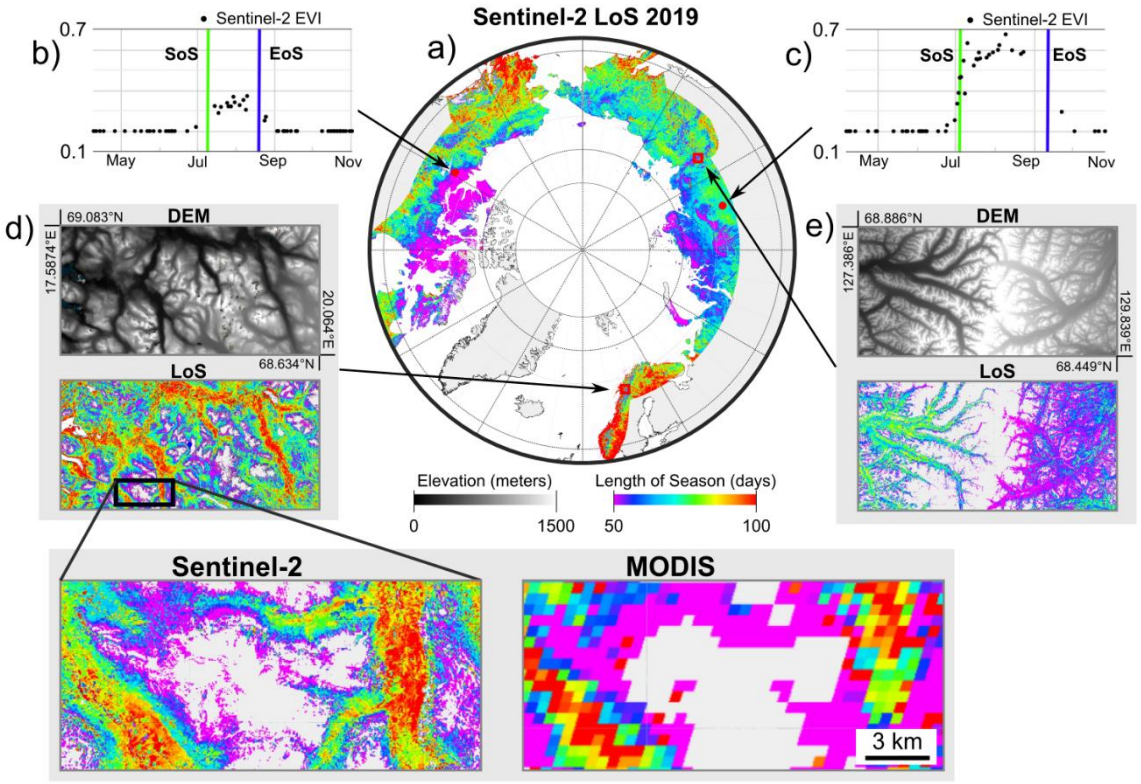


Figure 2.5. Map of length of season (LoS) in the Arctic estimated using Sentinel-2 enhanced vegetation index (EVI) time series for 2019 (a). EVI time series and start of season (SoS) and end of season (EoS) for two randomly selected pixels (b,c); note, the time series in (b) show a shorter LoS than in (c) due to the latitudinal gradient. Arctic digital elevation model map (ArcticDEM) and LoS exemplify the altitudinal gradient of phenology in the Scandinavian mountains and the Ural mountains (d,e) and comparison of LoS estimated with Sentinel-2 and Moderate Resolution Imaging Spectroradiometer (MODIS) (d). The phenology metrics were extracted with the 50% threshold method applied to the non-smoothed Sentinel-2 EVI time series.

The comparison between the LSP metrics estimated using MODIS and Sentinel-2 resized to 500 m, in which MODIS pixels were filtered when variance of the Sentinel-2 phenology estimates within the 500 m pixel were >5 d^2, showed a high level of similarity in homogeneous landscapes (Figure 2.6), regardless of data smoothing (Figure 2.6a,b), and greater similarity for SoS than EoS (e.g., threshold method with smoothing SoS and EoS RMSE: 8.1 and 10.4 d, respectively). The comparison of our MODIS estimates and the MCD12Q2v6 product (Supplementary Figure 2.10) showed a slight bias towards EoS (threshold method with smoothing ME: 4.9 and 6.4 d, respectively) and similarity with Sentinel-2 for SoS and EoS (threshold method without smoothing RMSE = 10.2 and 9.8 d for SoS and EoS, respectively).

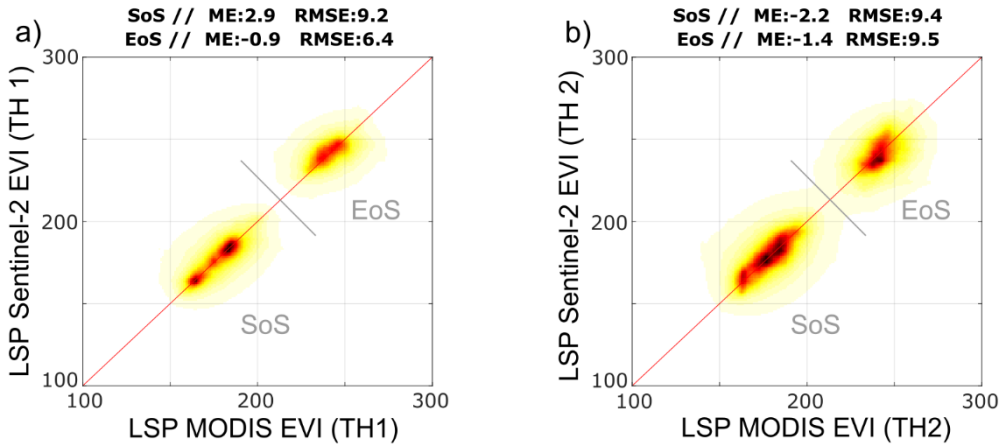


Figure 2.6. Comparison of the phenology metrics, start of season (SoS) and end of season (EoS), estimated with Sentinel-2 and MODIS (MOD09GAv6) time series using the threshold method without (a) and with (b) data smoothing. The bias between PhenoCam and Sentinel-2 is reported with the mean error (ME) and the accuracy with the root mean squared error (RMSE).

The phenology maps of the Arctic presented some unreliable LSP estimates, particularly when there were continuous gaps in the time series, due to the presence of clouds. The proportion of vegetated land pixels in the study area with a gap length that exceeded 40 days was 9.2% for the spring and early summer period and 39.4% for the late summer and autumn period. These percentages refer only to the pixels that presented an EVI value higher than 0.2 and, thus, were considered as vegetated land in the study. The percentage of pixels flagged as non-vegetated areas (EVI values lower than 0.2 during the entire time series) was 15.4%. Figure 2.7 illustrates the flags associated with the SoS estimates. The regions with low revisit times (Figure 2.1a) were prone to discontinuities in the time series.

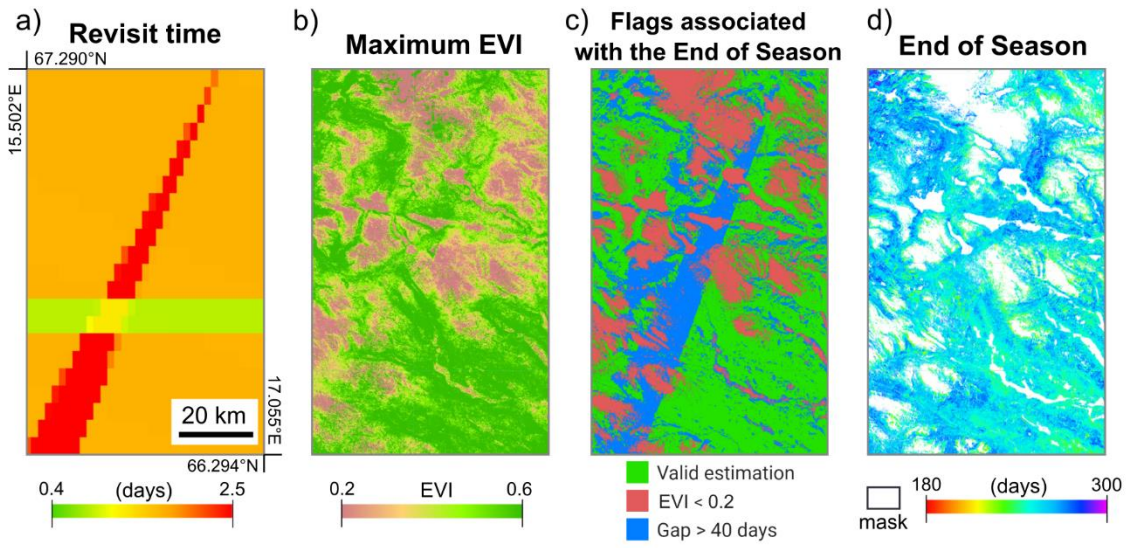


Figure 2.7. Example of flags associated with gaps in the Sentinel-2 time series in the Scandinavian mountains. (a) Revisit time; (b) maximum enhanced vegetation index (EVI) for 2019; (c) flags raised during the estimation of end-of-season (EoS); (d) EoS estimation. Red: non-vegetated pixels (EVI<0.2); blue: pixels with >40-d gap in time series; EoS was extracted with the 50% threshold method applied to the non-smoothed Sentinel-2 EVI time series.

2.4. Discussion

The high agreement between the phenology metrics estimated with Sentinel-2 and MODIS corroborates the feasibility of LSP estimation owing to the low revisit time of the Sentinel-2 time series (Bolton et al., 2020) in a region prone to cloud coverage. We found that only 9.2% of pixels showed a discontinuity of >40 d in the time series during the green-up period. In contrast, the cloud coverage was more persistent in late summer and autumn (39.4% of pixels with >40-d gap), and this high level of cloud occurrence may explain the lower similarity between EoS estimated using Sentinel-2 and MODIS. Such discontinuities in time series make the use of gap-filling techniques and robust smoothing techniques necessary for the accurate estimation of LSP metrics (Bolton et al., 2020). Moreover, when time series present continuous gaps, the combined use of Sentinel-2 and Landsat-8 is recommended as, overall, the combination of sensors provides improved LSP estimates (Kowalski et al., 2020).

The comparison with PhenoCam further proves the feasibility of LSP estimation with Sentinel-2 with a variety of vegetation indices. Results showed that the EVI performed the best. It was shown by (Kowalski et al., 2020) that EVI was more suited than NDVI in temperate deciduous forests and (Bolton et al., 2020) used EVI2 (two-band EVI) for continental-scale analysis, both

using a combination of Landsat-8 and Sentinel-2. Despite this, an extensive analysis including more in situ observations is required to further justify the use of EVI in Arctic phenology studies. Moreover, the correct reclassification of snow values seems more determinant than the variable selection in high-latitudes.

The procedure for flagging pixels with discontinuities in the time series may be used to identify good-quality pixels for further analysis in phenology studies; however, this flag only works when clouds are successfully filtered. Non-valid observations that were not filtered, mainly cloud-contaminated pixels and cloud shadows that were not reflected in the quality band, are included in the time series and may underestimate the long gaps in the Sentinel-2 time series. Furthermore, such contaminated values may change the growing season curve in the vegetation index time series and lead to erroneous estimates of the phenology.

We found high levels of similarity between the phenology metrics estimated using Sentinel-2 and MODIS and PhenoCam, even though the comparison reflects changes in vegetation greenness that do not necessarily correspond to vegetation phenophases or vegetation productivity (Helman, 2018). In situ databases, such as the National Phenology Network (NPN) [29], the Pan European Phenology database (PEP725) [30], and the FLUXNET network [31] are essential to ground-truth remotely sensed phenological changes in vegetation. Near-surface LSP, such as the PhenoCam network, supports the validation of satellite LSP estimates, as observed in the current study. Phenophase observations provide information on the timing of relevant stages in the annual life cycle of vegetation, such as the leaf-out, which may differ from the LSP metrics estimated from a satellite (Bórnez et al., 2020). Similarly, the onset and the end of the photosynthetic activity, estimated from time series of carbon fluxes in FLUXNET, do not match with the LSP metrics depending on the forest ecosystem (D'Odorico et al., 2015). Phenophase, carbon fluxes, and near-surface LSP measurements are scarce in the Arctic, with the distribution of digital cameras generally restricted to northern areas of Alaska, so data are lacking for adequate validation of high-latitude LSP maps.

The use of cloud platforms, such as GEE, has made possible the estimation of LSP using Sentinel-2 time series for extensive regions, such as the Arctic. In this study, the generation of the phenology maps for the Arctic used approximately 133,000 Sentinel-2 images, solely for the year 2019; this type of platform is a useful tool that allows the scientific community to inspect such dense, high-spatial resolution time series. Furthermore, GEE allows sharing the code and data,

and researchers can easily reproduce the algorithm and inspect the time series locally (see GEE code in Supplementary Materials).

The resulting phenology maps should be taken with caution when data availability is scarce, which occurs particularly during the EoS (39.4% of pixels with >40 d gap). The combination of Sentinel-2 and Landsat-8 is recommended when continuous gaps are present in the time series, and more elaborated phenological retrieval methods including gap-filling and robust smoothing techniques should be used instead. Despite this, sensor harmonization and temporal smoothing may improve the LSP retrievals but at the expenses of adding complexity to the preprocessing, which impedes the implementation and fast computation in GEE. Our SoS and EoS maps may benefit future Arctic phenology studies that aim to analyze spatial variability in vegetation dynamics and require a high degree of detail at the canopy level; test effects of spatial scaling on temporal changes in phenology; and identify homogeneous landscapes in which phenology dynamics are similar.

2.5. Conclusions

Here, we present the first high-spatial resolution maps (10 m) of SoS and EoS that fully cover the Arctic for the most recent years (2019–2020). We prove the feasibility of phenology metric extraction in the region solely with Sentinel-2 and using basic implementations of the threshold method in GEE; the high revisit time at high latitudes allows dense time series to be obtained under cloud-free conditions in 1–3 days and only 9.2% of pixels showed a discontinuity of >40 d in the time series during the green-up period. We propose a set of forcing values for the reclassification of snow observations in Sentinel-2 for three common vegetation indices used in LSP estimation: GCC, NDVI, and EVI; and also provide a re-adjusted parameterization of the NDPI specifically for Sentinel-2. Future work may use the adjusted parameters and forcing values along with the implementation of the threshold method in GEE for large-scale estimation of phenology metrics. The retrieved Sentinel-2 high-resolution LSP maps and the proposed GEE phenological extraction method will support monitoring vegetation changes at high-spatial resolution and are expected to contribute to the representation of Artic vegetation phenology in land surface models.

2.6. References

Beck, P., Jönsson, P., Høgda, K.-A., Karlsen, S., Eklundh, L., & Skidmore, A. (2007). A ground-validated NDVI dataset for monitoring vegetation dynamics and mapping phenology in Fennoscandia and the Kola peninsula. *International Journal of Remote Sensing*, *28*(19), 4311–4330.

Bolton, D. K., Gray, J. M., Melaas, E. K., Moon, M., Eklundh, L., & Friedl, M. A. (2020). Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. *Remote Sensing of Environment*, *240*, 111685.

Bórnez, K., Descals, A., Verger, A., & Peñuelas, J. (2020). Land surface phenology from VEGETATION and PROBA-V data. Assessment over deciduous forests. *International Journal of Applied Earth Observation and Geoinformation*, *84*, 101974.

Chapin, F. S., Sturm, M., Serreze, M. C., McFadden, J. P., Key, J., Lloyd, A. H., McGuire, A., Rupp, T. S., Lynch, A. H., Schimel, J. P., & others. (2005). Role of land-surface changes in Arctic summer warming. *Science*, *310*(5748), 657–660.

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., & others. (2017). An ecoregion-based approach to protecting half the terrestrial realm. *BioScience*, *67*(6), 534–545.

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. *Agricultural and Forest Meteorology*, *214*, 25–38.

Friedl, M., Gray, J., & Sulla-Menashe, D. (2019). MCD12Q2 MODIS/Terra+ Aqua Land Cover Dynamics Yearly L3 Global 500m SIN Grid V006. *NASA EOSDIS Land Processes DAAC; NASA: Washington, DC, USA*.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27.

Gray, J., Sulla-Menashe, D., & Friedl, M. A. (2019). User Guide to Collection 6 MODIS Land Cover Dynamics (MCD12Q2) Product. *NASA EOSDIS Land Processes DAAC: Missoula, MT, USA*.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S., Goetz, S. J., Loveland, T. R., & others. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, *342*(6160), 850–853.

Helman, D. (2018). Land surface phenology: What do we really 'see'from space? *Science of the Total Environment*, *618*, 665–673.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1–2), 195–213.

Jin, H., & Eklundh, L. (2014). A physically based vegetation index for improved monitoring of plant phenology. *Remote Sensing of Environment*, *152*, 512–525.

Kowalski, K., Senf, C., Hostert, P., & Pflugmacher, D. (2020). Characterizing spring phenology of temperate broadleaf forests using Landsat and Sentinel-2 time series. *International Journal of Applied Earth Observation and Geoinformation*, *92*, 102172.

Myers-Smith, I. H., Kerby, J. T., Phoenix, G. K., Bjerke, J. W., Epstein, H. E., Assmann, J. J., John, C., Andreu-Hayles, L., Angers-Blondin, S., Beck, P. S., & others. (2020). Complexity revealed in the greening of the Arctic. *Nature Climate Change*, *10*(2), 106–117.

Park, T., Ganguly, S., Tømmervik, H., Euskirchen, E. S., Høgda, K.-A., Karlsen, S. R., Brovkin, V., Nemani, R. R., & Myneni, R. B. (2016). Changes in growing season duration and productivity of northern vegetation inferred from long-term remote sensing data. *Environmental Research Letters*, *11*(8), 084001.

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422.

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

Post, E., Alley, R. B., Christensen, T. R., Macias-Fauria, M., Forbes, B. C., Gooseff, M. N., Iler, A., Kerby, J. T., Laidre, K. L., Mann, M. E., & others. (2019). The polar regions in a 2° C warmer world. *Science Advances*, *5*(12), eaaw9883.

Richardson, A. D., Hufkens, K., Milliman, T., Aubrecht, D. M., Chen, M., Gray, J. M., Johnston, M. R., Keenan, T. F., Klosterman, S. T., Kosmala, M., & others. (2018). Tracking vegetation phenology across diverse North American biomes using PhenoCam imagery. *Scientific Data*, *5*, 180028.

Schuur, E. A., McGuire, A. D., Schädel, C., Grosse, G., Harden, J., Hayes, D. J., Hugelius, G., Koven, C. D., Kuhry, P., Lawrence, D. M., & others. (2015). Climate change and the permafrost carbon feedback. *Nature*, *520*(7546), 171–179.

Schuur, E. A., Vogel, J. G., Crummer, K. G., Lee, H., Sickman, J. O., & Osterkamp, T. (2009). The effect of permafrost thaw on old carbon release and net carbon exchange from tundra. *Nature*, *459*(7246), 556–559.

Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A. M., Friedl, M., Braswell, B. H., Milliman, T., O'Keefe, J., & Richardson, A. D. (2012). Digital repeat photography for phenological research in forest ecosystems. *Agricultural and Forest Meteorology*, *152*, 159–177.

Vermote, E., & Wolfe, R. (2015). MOD09GA MODIS/Terra Surface Reflectance Daily L2G Global 1 km and 500 m SIN Grid V006. *NASA EOSDIS LP DAAC*.

Walt, S. van der, Colbert, S. C., & Varoquaux, G. (2011). The NumPy array: A structure for efficient numerical computation. *Computing in Science & Engineering*, *13*(2), 22–30.

Wang, C., Chen, J., Wu, J., Tang, Y., Shi, P., Black, T. A., & Zhu, K. (2017). A snow-free vegetation index for improved monitoring of vegetation spring green-up date in deciduous ecosystems. *Remote Sensing of Environment*, *196*, 1–12.

Witze, A. (n.d.). The Arctic is burning like never before-and that's bad news for climate change. *Nature*.

Zeng, H., Jia, G., & Epstein, H. (2011). Recent changes in phenology over the northern high latitudes detected from multi-satellite data. *Environmental Research Letters*, *6*(4), 045508.

Zeng, L., Wardlow, B. D., Xiang, D., Hu, S., & Li, D. (2020). A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sensing of Environment*, *237*, 111511.

Chapter 3

Soil thawing regulates the spring growth onset in tundra and alpine biomes

Adrià Descals, Aleixandre Verger, Iolanda Filella, Dennis Baldocchi, Ivan A. Janssens, Yongshuo H. Fu, Shilong Piao, Marc Peaucelle, Philippe Ciais, and Josep Peñuelas

Published in *Science of the Total Environment, 742, 140637* (2020)
https://doi.org/10.1016/j.scitotenv.2020.140637

Abstract

Soil temperature remains isothermal at 0 °C and water shifts to a liquid phase during soil thawing. Vegetation may receive this process as a signal and a key to restore physiological activity. We aimed to show the relationship between the timing of soil thawing and the spring growth onset. We estimated the delay between the soil thawing and the spring growth onset in 78 sites of the FLUXNET network. We built a soil thawing map derived from modelling for the northern hemisphere and related it to the greenness onset estimated with satellite imagery. Spring onset estimated with GPP time series occurred shortly after soil surface thawing in tundra (1.1 ± 3.5 days) and alpine grasslands (16.6 ± 5.8 days). The association was weaker for deciduous forests (40.3 ± 4.2 days), especially where soils freeze infrequently. Needleaved forests tended to start the growing season before the end of thawing (-17.4 ± 3.6 days), although observations from remote sensing (MODIS Land Cover Dynamics) indicated that the onset of greenness started after the thawing period (26.8 ± 3.2 days). This study highlights the role of soil temperature at the spring growth onset at high latitudes. Soil thawing becomes less relevant in temperate forests, where soil is occasionally frozen and other climate factors become more important.

3.1. Introduction

Phenology is defined as the timing of periodic life-cycle events (Schwartz, 2003). Phenology has recently gained attention as an indicator of global warming and more broadly for its link with global environmental change (Badeck et al., 2004; Cleland et al., 2007; Peñuelas et al., 2009; Peñuelas & Filella, 2001). Plant phenology is primarily modulated by the seasonal variation in climate, which involves a stage of dormancy when the climate is adverse and a period of growth when conditions are favorable for vegetation activity. Temperature, water availability, and day length have been proposed as the main environmental drivers that constrain vegetation activity and regulate plant phenology, acting eventually as confounded restraints for vegetation growth (Chuine & Régnière, 2017; Jolly et al., 2005; Kramer & Hänninen, 2009). Temperature is the main climatic factor regulating plant phenology for the onset of vegetation in the northern high latitudes (Schwartz, 2003). Evidence for the interaction of temperature with other limiting factors, e.g. photoperiod and water availability, though, is under discussion (Fu et al., 2019; Körner & Basler, 2010; Peñuelas et al., 2004) and, the choice of the climatic variables for modeling phenology is an open science question.

Air temperature has been widely used to simulate the transition from dormancy to growth in temperature-limited biomes, and jointly with day length, an indicator of photoperiod, for modeling plant phenology in the Northern Hemisphere (Badeck et al., 2004; Peñuelas & Filella, 2001). Some biome-specific studies, however, have suggested that soil temperature more accurately indicates the start of the growing season (Baldocchi et al., 2005; Jiang et al., 2018; Lieth, 2013; Semenchuk et al., 2016; Starr et al., 2008), although a study showed a stronger connection with air temperature than soil temperature for boreal forests (Tanja et al., 2003). Some studies even suggested that the lengthening of the growing season in tundra during recent decades was primarily caused by a change in the seasonal thaw cycle (Barichivich et al., 2013; Kimball et al., 2006). A soil warming experiment in subarctic grasslands, which reduced the occurrence of soil frost, revealed that the growing season lengthened across the entire range of IPCC warming projections for 2100 and reported no indications of photoperiod constraints within that warming range (Leblans et al., 2017).

A frozen soil strictly limits vegetation growth and overrules any impact of eventual favorable air temperatures during spring. Thawing occurs when soil remains isothermal at 0 °C, triggering some processes that may be used as signals for the vegetation to restore physiological activity.

Water only returns to its liquid form and becomes available to roots (Noormets, 2009) during and after thawing. In addition to a direct signal of soil thawing, soil temperature has been found to exert a stronger influence on plants than air temperature. For example, soil temperature has also been associated with physiological processes such as root elongation (Wang et al., 2018) and root respiration (Lloyd & Taylor, 1994). Soil-warming experiments have found that an overall increase in soil temperature affected specific traits; leaf area, leaf expansion rates, and diameter growth in deciduous trees (Farnsworth et al., 1995; Wheeler et al., 2016), needle starch content and photochemical efficiency in boreal forests (Repo et al., 2004), and flowering frequency in tundra (Khorsand et al., 2015).

These results suggest that soil temperature may play a more deterministic limiting role than air temperature at the onset of spring at high latitudes. We tested the hypothesis that soil thawing regulates spring growth onset at high latitudes but becomes less relevant where soil is only sporadically frozen.

3.2. Materials and Methods

3.2.1. Phenology derived from in situ and satellite data

We used the same methodology as Zhang *et al,* 2003 to estimate the start of season at the FLUXNET sites. We took the daily mean gross primary productivity (GPP) measurements (daytime partitioning method) and adjusted a logistic function for each year for the period January 1st to July 1st. The start of season at the FLUXNET sites (GPP SoS) was the day when the second derivative of the logistic function was maximal.

We used the MCD12Q2 V5 Land Cover Dynamics product (MODIS SoS) for the period 2001-2014. MCD12Q2 phenological metrics are estimated yearly, with a logistic function fitted over the MODIS EVI time series (Zhang et al., 2003). We used the layer 'Onset Greenness Increase', which represents the day of the year on which the positive curvature of the fitted logistic function is maximal.

3.2.2. Estimation of the start and end of thawing

We defined the end of thawing (EoT) as the last day before solstice with daily mean soil temperature <0.5 °C and defined the start of thawing (SoT) as the closest day before EoT with a

temperature below -0.5 °C. We estimated the EoT and SoT using the *in situ* records of soil temperature at the FLUXNET sites in the Profile 1 (TS_1_1_1), which corresponds to soil temperature measured at the shallowest layer, with a different depth depending on the FLUXNET site but commonly 2-5 centimeters. The EoT was also estimated using the layer of soil temperature at 10-centimeter depth available in the Global Land Data Assimilation System (GLDAS) dataset (See section 2.3.). This resulted in global maps of EoT at 0.25° resolution. We used a threshold of 0.5 °C instead of 0 °C to avoid false-positive estimates of EoT due to noise and slight variations in the soil temperature time series during thawing.

3.2.3. Global climatic data set

We used the GLDAS version 2.1 as a global climatic data set for the daily representation of soil temperature at a global scale. GLDAS is a land-surface modeling system developed by the Goddard Space Flight Center of the National Aeronautics and Space Administration and by the National Centers for Environmental Prediction of the National Oceanic and Atmospheric Administration. The version 2.1 provides data for the period 2000 up until the present time. Ground measurements (meteorological data), such as soil temperature, and satellite observations are ingested in order to constrain the land-surface model and restrain unrealistic model states. GLDAS offers 3-hourly estimates of earth-surface variables such as soil temperature at different depths. We used the soil temperature product at 10-centimeter depth with a spatial resolution of 0.25°.

3.2.4. Resizing and filtering MODIS SoS

The spatial resolutions of GLDAS EoT and MODIS SoS differed, so we resized the resolution of MODIS SoS (500 m) to the resolution of GLDAS EoT (0.25°) to make the comparison between the two datasets consistent. We first masked the yearly MODIS SoS images with the MODIS IGBP Land Cover (MCD12Q1) product. MODIS SoS was masked separately with the IGBP classes 1) deciduous broadleaved forests (deciduous forests), 2) evergreen needleleaved forests (needleleaved forests), 3) shrubland, 4) open shrubland, and 5) grassland. For the tundra biome, we only considered the pixels classified as shrubland, open shrubland, and grassland at >50° latitude. Finally, the masked MODIS SoS images were resized to 0.25° using the mean as the resampling operation. This resulted in six sets of MODIS SoS for each land cover for the period 2001-2014.

3.2.5. Experimental setup and statistical analyses

We analyzed 78 sites of the FLUXNET network and compared the timing of spring onset of ecosystem GPP with thawing date. We used all sites north of 30°N latitude that are located in tundra (13 sites) and alpine (4 sites) climates, and the sites categorized in the IGBP classification as deciduous broadleaved forests (DBF) (18 sites in temperate and 1 site boreal climate zones) or evergreen needleaved forests (ENF) (25 sites in temperate and 17 sites boreal climate zones). Supplementary Table 3.1 shows the coordinates and biome of the selected sites. We excluded the sites without in situ soil temperature records.

We investigated at which soil temperature the GPP was significantly positive in each biome. We grouped the half-hourly GPP observations in intervals of 0.5°C of soil temperature. We applied a two-tailed Student's t-test at each interval of soil temperature with a significance of 95%. We considered that GPP was positive at a certain temperature when the lower GPP confidence interval was greater than 0.5 μmol CO_2 m^{-2} s^{-1}.

The SoT and EoT estimated with the FLUXNET records were compared with two independent observations of the start of the growing season (SoS): 1) SoS estimated with daily *in situ* observations of the FLUXNET GPP time series (GPP SoS) and 2) SoS extracted from remote-sensing sources (MODIS SoS). We applied the two-tailed Student's t-test with a significance of 95% to find whether the day of the EoT and the spring onset (GPP SoS and MODIS SoS) were significantly different. We reported the mean difference, confidence interval, and coefficient of determination (R^2) between the end of thawing and the spring onset.

We also compared the spring onset observed with satellite derived data (MODIS SoS) with the thawing period estimated with the GLDAS. The MODIS SoS was filtered and resized to the resolution of the GLDAS to make the datasets comparable. Similarly to the comparison at the FLUXNET sites, we reported the R^2 and the mean difference with a confidence interval of 95% between the EoT and the MODIS SoS. The analysis covers a time period from 2000 to 2015 for the comparison between FLUXNET and MODIS, and from 2001 to 2017 for the comparison between the GLDAS and MODIS.

3.3. Results

The growing seasons were primarily driven by temperature in the 78 FLUXNET sites; GPP and air temperature had similar intra-seasonal patterns. Gross primary productivity at these sites was highest in summer, when the air and soil temperatures were also highest. The Italian site IT-Ro1 was the only exception, where GPP decreased during summer due to water stress. A representation of air and soil temperature, GPP and spring onset is shown in Supplementary Figure 1 for five FLUXNET sites. GPP was insignificant with frozen soil in all biomes (Figure 3.1). The lowest soil temperature at which GPP was significantly higher than 0.5 µmol CO_2 m^{-2} s^{-1} was 1°C in tundra, -0.5°C in alpine grasslands, -1 °C in needleaved forests, and 4.5 °C in deciduous forests.

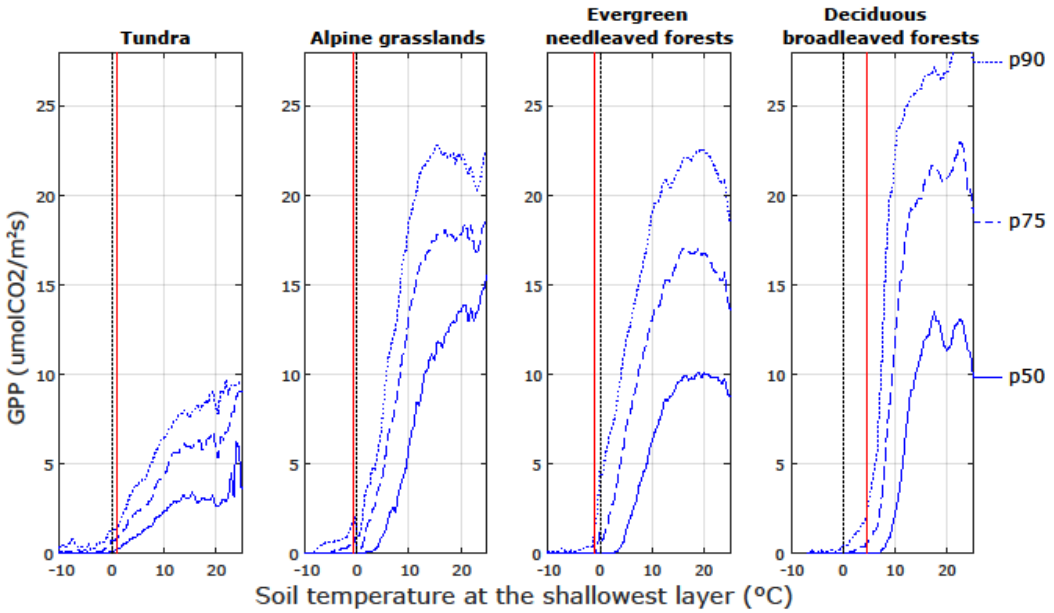


Figure 3.1. Gross primary productivity (GPP) in relation to soil temperature in the FLUXNET sites. The x-axis shows soil temperature measured at the shallowest layer. The distribution of GPP is presented with the 50th percentile (continuous line), 75th percentile (dashed line), and 90th percentile (dotted line). The red line shows the lowest temperature in which GPP is significantly higher than 0 °C. The temporal resolution of the GPP and temperature data is half-hourly.

Spring onset occurred during thawing (41% of observations for GPP SoS and 6% of the observations for MODIS SoS) or after EoT (51% of observations for GPP SoS and 59% of the observations for MODIS SoS) (Figure 3.2). Only 8% of the observations occurred when soil was frozen in GPP SoS, primarily in needleaved forests. Spring onset for tundra and alpine grasslands

tended to start shortly after the EoT (1.1 ± 3.5 d for GPP SoS and 18.5 ± 3.0 d for MODIS SoS for tundra, 16.6 ± 5.8 d for GPP SoS and 32.9 ± 8.3 d for MODIS SoS for alpine grasslands). Spring onset tended to start with a longer delay after EoT in deciduous forests with seasonally frozen soil (40.3 ± 4.2 d for GPP SoS and 46.4 ± 3.8 d for MODIS SoS), while needleleaved forests showed discrepant results between the two sources of data (-17.3 ± 3.6 d for GPP SoS and 26.8 ± 3.2 d for MODIS SoS).

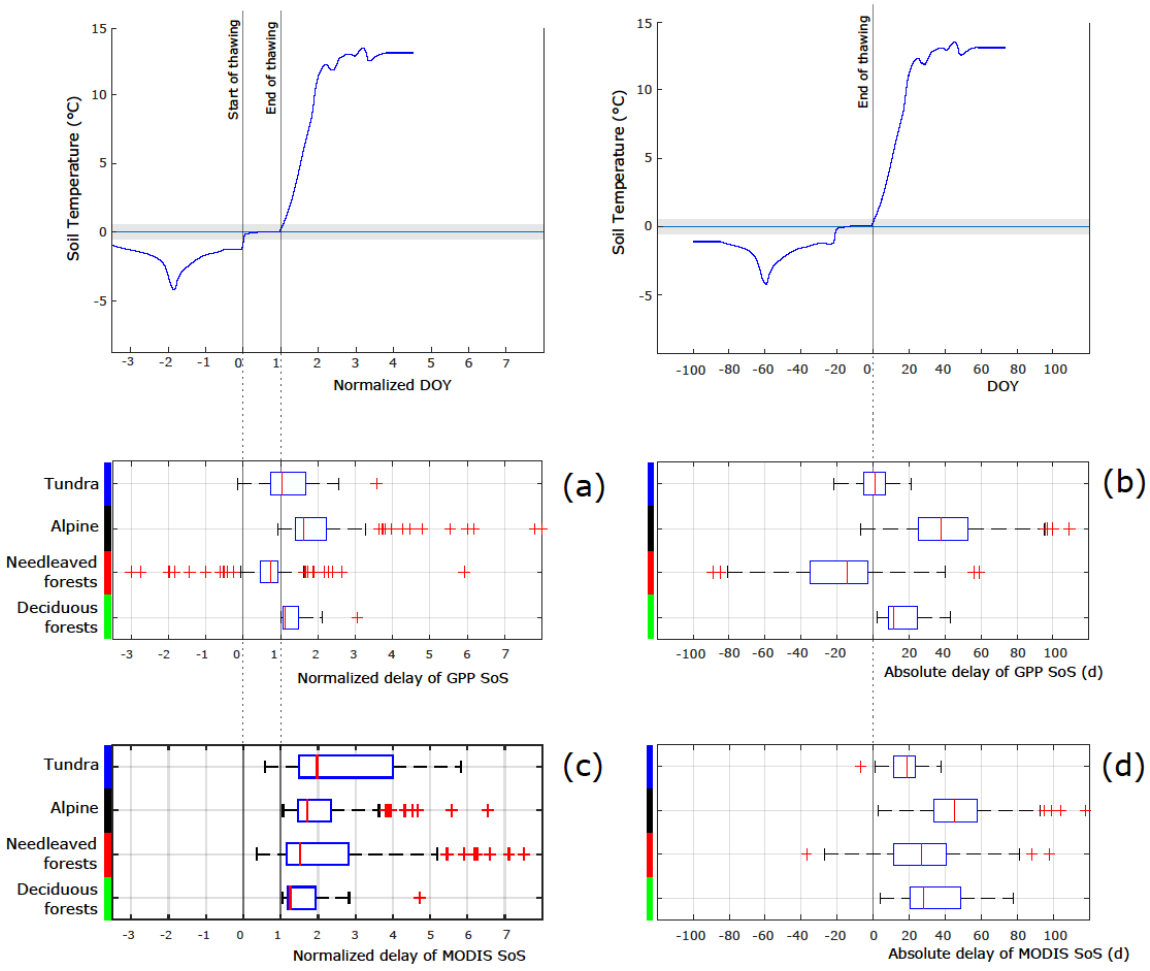


Figure 3.2. Boxplots of the normalized delay ((a) and (c)) and absolute delay ((b) and (d)) of spring onset after the end of thawing (EoT) for tundra, alpine grasslands, needleaved forests, and deciduous forests. The absolute delay is the number of days between spring onset and EoT, and the normalized delay represents spring onset relative to thawing, with start of thawing (SoT) = 0 and EoT = 1. Spring onset is presented as FLUXNET gross primary productivity (GPP) start of season (SoS) in (a) and (b) and as MODIS SoS ('MidGreenup1' band in MCD12Q2v6 product) in (c) and (d). The central mark of the boxplot indicates the median, the bottom and top edges indicate the 25th and 75th percentiles. The whiskers cover until the most extreme data points not considered outliers, and outliers are plotted with the '+' symbol.

The association between the GPP SoS and the EoT SoS is different in deciduous than in needleleaved forests (Figure 3.3a). Spring onset occurs with positive temperatures in deciduous

forests, while needleleaved forest starts their growing season before the soil has completely thawed. Tundra and alpine grasslands show the highest agreement of spring onset with the EoT (*RMSE* = 9.03 in GPP SoS and *RMSE* = 20.45 in MODIS SoS). FLUXNET sites in the boreal biome also showed a better agreement than the sites in the temperate biome in both GPP SoS and MODIS SoS (Supplementary Figure 3.2). In addition, soil did not freeze in some sites of the temperate biome, which explains the EoT values equal to 0 in Figure 3.3 and Supplementary Figure 3.2.

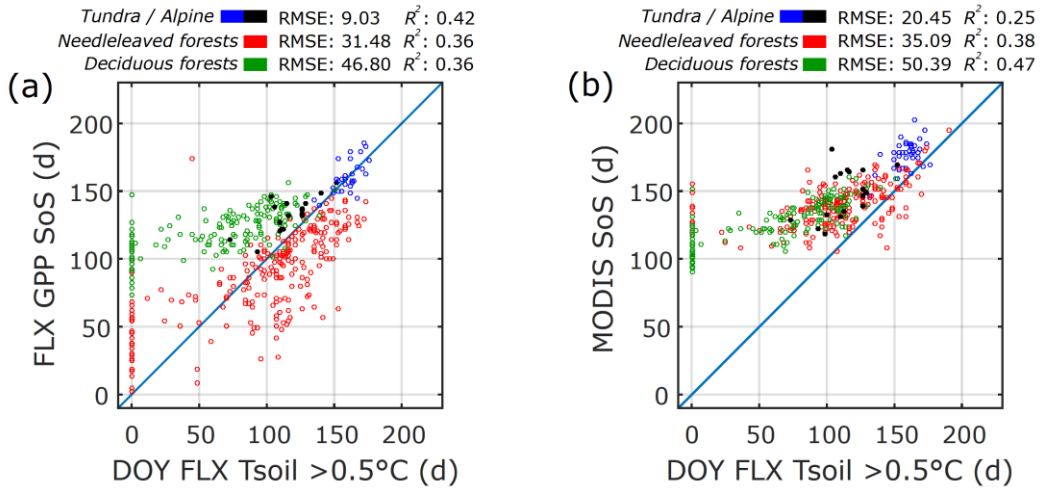


Figure 3.3. Scatter plots of spring onset (gross primary productivity (GPP) start of season (SoS) (a) and MODIS SoS ('MidGreenup1' band in MCD12Q2v6 product) (b)) and the last day before the growing season using soil temperature >0.5 °C. Point colors represent tundra (blue), alpine grasslands (black), needleaved forests (red), and deciduous forests (green). Notice that the best fit between the spring onset and the end of thawing corresponds to tundra and alpine grasslands.

The comparison of MODIS SoS with the global estimates of EoT (Figure 3.4) led to similar results as the observations at the FLUXNET sites. The onset of greenness estimated with MODIS occurred near EoT in the land cover types that represent tundra (8.63 ± 7.96 d in shrubland IGBP class, 16.97 ± 7.72 d in open shrubland IGBP class, and 18.80 ± 14.55 d in grassland IGBP class), while deciduous forests showed the highest delay (34.24 ± 18.59 d). The high delay in deciduous forests can be observed in the map showing the difference mean MODIS SoS - mean EoT (Figure 3.5) at the most southerly latitudes of the deciduous forest distribution (eastern North America, central Europe, and northern China). We observed a high interannual variability in the MODIS SoS and EoT also at low latitudes (Supplementary Figure 3.3). Contrarily to the results in the

FLUXNET sites, the MODIS SoS in needleleaved forests show a similar pattern as deciduous forests and do not present a negative delay compared to the EoT (20.03 ± 14.15 d).

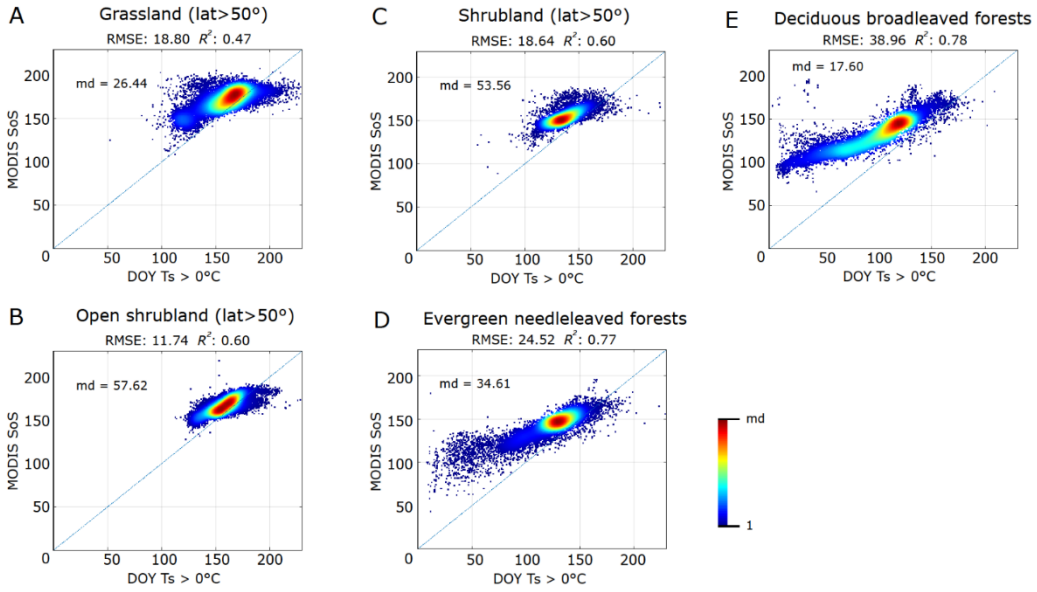


Figure 3.4. Relationships between MODIS start of season (SoS) and end of thawing (EoT) estimated using the GLDAS dataset. Each panel represents MODIS SoS ('MidGreenup1' band in MCD12Q2v6 product) masked with the MODIS IGBP Land Cover product for cover types (a) grassland, (b) open shrubland, (c) shrubland, (b) evergreen needleleaved forests, and (e) deciduous broadleaved forests. The land cover types grassland, open shrubland, and shrubland for latitudes higher than 50° represent predominantly the tundra biome. The maximum density (md) varies depending on the land cover.

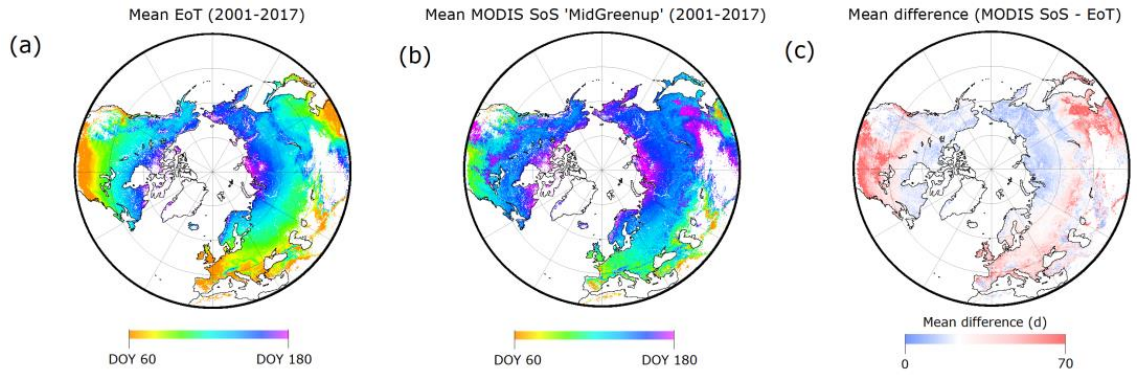


Figure 3.5. Maps of (a) mean end of thawing (EoT), (b) mean MODIS start of season (SoS), and (c) mean difference (MODIS SoS - EoT) for the period 2001-2017. EoT was estimated using the GLDAS data set as the last day of the dormant period with soil temperatures <0.5 °C. MODIS SoS is the 'MidGreenup1' band in MCD12Q2v6 product. We resized MODIS SoS to the spatial resolution of the GLDAS dataset. The map projection is the Lambert Azimuthal Equal-Area.

3.4. Discussion

Spring onset of vegetation activity and soil thawing were closely associated in tundra and alpine grasslands, and weakly associated where soil froze intermittently. Water remains inaccessible to plants in frozen soils, inhibiting vegetation activity. This process is particularly relevant in regions with deep frost, where the vegetation has maximal root depth <0.5 m in tundra or <2 m in boreal forests (Canadell et al., 1996). Likewise, most root biomass in boreal and temperate forests is in shallow soil layers, where most mineralization processes release the greatest amount of nutrients. Frozen soil, even if only the first few centimeters, may therefore imply inhibition of vegetation activity, as suggested by the results from the FLUXNET towers, where the GPP is insignificant when soil temperatures are negative (Figure 3.1). In addition, most of the observations of spring onset, both MODIS SoS and GPP SoS, coincided with positive or near freezing soil temperatures (Figure 3.2).

Water gradually shifts from the solid to the liquid form during thawing, which plants may use as a signal to start their activity. Our results, however, indicated that the time between EoT and spring onset depended on the biome. These differences can be attributed to different forcing requirements for which vegetation has different adaptive histories. Deciduous forests seem to require greater heat accumulation than tundra and needleleaved forests to initiate spring growth. Many studies have suggested that the high thermal forcing required by deciduous trees, which delays the vegetation onset date after the EoT, is an adaptation to prevent frost damage in late spring (Fu et al., 2015; Laube et al., 2014). In addition, deciduous forests have adapted their high thermal forcing depending on the background climate (Peaucelle et al., 2019), with lower heat requirements in colder climates. This would explain the high delay between the spring onset and EoT in deciduous forest in the lowest latitudes of the study, while the only FLUXNET site with deciduous forests located in the boreal zone shows a close association between the spring onset and the thawing period. Similarly to the site with deciduous forest in the boreal biome, tundra and alpine grasslands have a low heat requirements, very likely because their SoS also occurs very late in the season, when light availability is already high, and these plants may therefore have reduced their heat requirement to maximize profit from the high light conditions in late spring.

Our results show that needleaved forests can release from dormancy and start their gross primary production during the thawing period with near freezing soil temperatures, which

corroborates the photosynthetic acclimation to cold in boreal needleaved forests (Oquist & Huner, 2003). However, the results obtained from remote sensing show that the onset in greenness in needleaved forests occurs much closer to the end of the thawing period. This suggests that evergreen needleaved trees start the growing season when temperatures become favorable using the needles preserved during the dormant period, but needle emergence occurs much later than the start of the photosynthetic activity. The delay between GPP SoS and onset of greenness was also observed with remotely-sensed data in previous studies (Melaas et al., 2013; Walther et al., 2016). An alternative explanation for the early GPP SoS in boreal forests is that understory might have started the photosynthetic activity before the trees in the overstory and, thus, vegetation in the understory layer accounted for the carbon fluxes before the thawing period. This would be supported by the findings of (Ikawa et al., 2015), which showed the contributions of the understory in boreal ecosystems to the carbon and energy balances.

The hypothesis that tundra has adapted to the long period with frozen soils by reducing its heat requirement implies that spring onset is highly responsive to soil thawing. This hypothesis is supported by studies with long time series of snow melt and by soil-warming experiments reporting that spring onset is advanced by an increasingly earlier soil thawing (Barichivich et al., 2013; Kimball et al., 2006; Leblans et al., 2017). Spring onset in arctic and subarctic vegetation may thus continue to advance as soil thawing occurs earlier with climate warming. The possible restriction in vegetation growth because of insufficient levels of incoming radiation or hours of sunshine is, however, still undetermined.

Our results suggest that the role of soil temperature is secondary at southerly latitudes for deciduous and needleaved forests. Sites with favorable soil and air temperatures throughout the year may be constrained by photoperiod (Zohner et al., 2016), by chilling requirements (Chuine et al., 2010), or by both (Fu et al., 2019). These regions without soil temperature regulation had the largest differences between MODIS SoS and EoT, as observed in the difference map in Figure 3.5. The differences observed between the EoT and the GPP SoS at the FLUXNET towers with needleleaved forests cannot be depicted with the MODIS SoS and, thus, are not reflected in the map in Figure 3.5.

In summary, we demonstrate that spring onset rarely occurs with negative soil temperatures, indicating that frozen soil acted as a major constraint on vegetation activity. We also demonstrate that vegetation activity began shortly after thawing in tundra, after soil

temperatures at 2-centimeter depth surpassed the 0 °C threshold. This correlation became weaker for lower latitudes. Needleleaved forests can start the gross primary production during the thawing period, while deciduous forests delayed the spring onset especially at more southerly latitudes. These results thus provide insight in the onset of spring that can be used for improving phenological models, especially in tundra and boreal forests where soil thawing plays a major role in controlling the start of vegetation activity. Further studies may take advantage of recent advances in mapping soil thawing using microwave remote sensing to estimate the start of the growing season at northern latitudes.

3.5. References

Badeck, F., Bondeau, A., Böttcher, K., Doktor, D., Lucht, W., Schaber, J., & Sitch, S. (2004). Responses of spring phenology to climate change. *New Phytologist*, *162*(2), 295–309.

Baldocchi, D. D., Black, T. A., Curtis, P. S., Falge, E., Fuentes, J. D., Granier, A., Gu, L., Knohl, A., Pilegaard, K., Schmid, H. P., Valentini, R., Wilson, K., Wofsy, S., Xu, L., & Yamamoto, S. (2005). Predicting the onset of net carbon uptake by deciduous forests with soil temperature and climate data: A synthesis of FLUXNET data. *International Journal of Biometeorology*, *49*(6), 377–387. https://doi.org/10.1007/s00484-005-0256-4

Barichivich, J., Briffa, K. R., Myneni, R. B., Osborn, T. J., Melvin, T. M., Ciais, P., Piao, S., & Tucker, C. (2013). Large-scale variations in the vegetation growing season and annual cycle of atmospheric CO2 at high northern latitudes from 1950 to 2011. *Global Change Biology*, *19*(10), 3167–3183. https://doi.org/10.1111/gcb.12283

Canadell, J., Jackson, R., Ehleringer, J., Mooney, H., Sala, O., & Schulze, E.-D. (1996). Maximum rooting depth of vegetation types at the global scale. *Oecologia*, *108*(4), 583–595.

Chuine, I., Morin, X., & Bugmann, H. (2010). Warming, Photoperiods, and Tree Phenology. *Science*, *329*(5989), 277–278. https://doi.org/10.1126/science.329.5989.277-e

Chuine, I., & Régnière, J. (2017). Process-Based Models of Phenology for Plants and Animals. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 159–182. https://doi.org/10.1146/annurev-ecolsys-110316-022706

Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A., & Schwartz, M. D. (2007). Shifting plant phenology in response to global change. *Trends in Ecology & Evolution*, *22*(7), 357–365. https://doi.org/10.1016/j.tree.2007.04.003

Farnsworth, E. J., Nunez-Farfan, J., Careaga, S. A., & Bazzaz, F. A. (1995). Phenology and Growth of Three Temperate Forest Life Forms in Response to Artificial Soil Warming. *Journal of Ecology*, *83*(6), 967–977. https://doi.org/10.2307/2261178

Fu, Y. H., Zhang, X., Piao, S., Hao, F., Geng, X., Vitasse, Y., Zohner, C., Peñuelas, J., & Janssens, I. A. (2019). Daylength helps temperate deciduous trees to leaf-out at the optimal time. *Global Change Biology*, *0*. https://doi.org/10.1111/gcb.14633

Fu, Y. H., Zhao, H., Piao, S., Peaucelle, M., Peng, S., Zhou, G., Ciais, P., Huang, M., Menzel, A., Peñuelas, J., Song, Y., Vitasse, Y., Zeng, Z., & Janssens, I. A. (2015). Declining global warming effects on the phenology of spring leaf unfolding. *Nature*, *526*(7571), 104–107. https://doi.org/10.1038/nature15402

Ikawa, H., Nakai, T., Busey, R. C., Kim, Y., Kobayashi, H., Nagai, S., Ueyama, M., Saito, K., Nagano, H., Suzuki, R., & others. (2015). Understory CO2, sensible heat, and latent heat fluxes in a black spruce forest in interior Alaska. *Agricultural and Forest Meteorology*, *214*, 80–90.

Jiang, H., Zhang, W., Yi, Y., Yang, K., Li, G., & Wang, G. (2018). The impacts of soil freeze/thaw dynamics on soil water transfer and spring phenology in the Tibetan Plateau. *Arctic, Antarctic, and Alpine Research*, *50*(1), e1439155. https://doi.org/10.1080/15230430.2018.1439155

Jolly, W. M., Nemani, R., & Running, S. W. (2005). A generalized, bioclimatic index to predict foliar phenology in response to climate. *Global Change Biology*, *11*(4), 619–632. https://doi.org/10.1111/j.1365-2486.2005.00930.x

Khorsand, R., Oberbauer, S., Starr, G., La Puma, I., Pop, E., Ahlquist, L., & Baldwin, T. (2015). Plant phenological responses to a long-term experimental extension of growing season and soil warming in the tussock tundra of Alaska. *Global Change Biology*, *21*. https://doi.org/10.1111/gcb.13040

Kimball, J. S., McDonald, K. C., & Zhao, M. (2006). Spring Thaw and Its Effect on Terrestrial Vegetation Productivity in the Western Arctic Observed from Satellite Microwave and Optical Remote Sensing. *Earth Interactions*, *10*(21), 1–22. https://doi.org/10.1175/EI187.1

Körner, C., & Basler, D. (2010). Phenology Under Global Warming. *Science*, *327*(5972), 1461–1462. https://doi.org/10.1126/science.1186473

Kramer, K., & Hänninen, H. (2009). The Annual Cycle of Development of Trees and Process-Based Modelling of Growth to Scale Up From the Tree To the Stand. In A. Noormets (Ed.), *Phenology of Ecosystem Processes: Applications in Global Change Research* (pp. 201–227). Springer New York. https://doi.org/10.1007/978-1-4419-0026-5_9

Laube, J., Sparks, T. H., Estrella, N., Höfler, J., Ankerst, D. P., & Menzel, A. (2014). Chilling outweighs photoperiod in preventing precocious spring development. *Global Change Biology*, *20*(1), 170–182.

Leblans, N. I. W., Sigurdsson, B. D., Vicca, S., Fu, Y., Penuelas, J., & Janssens, I. A. (2017). Phenological responses of Icelandic subarctic grasslands to short-term and long-term natural soil warming. *Global Change Biology*, *23*(11), 4932–4945. https://doi.org/10.1111/gcb.13749

Lieth, H. (2013). *Phenology and seasonality modeling* (Vol. 8). Springer Science & Business Media.

Lloyd, J., & Taylor, J. A. (1994). On the Temperature Dependence of Soil Respiration. *Functional Ecology*, *8*(3), 315–323. https://doi.org/10.2307/2389824

Melaas, E. K., Richardson, A. D., Friedl, M. A., Dragoni, D., Gough, C. M., Herbst, M., Montagnani, L., & Moors, E. (2013). Using FLUXNET data to improve models of springtime vegetation activity onset in forest ecosystems. *Agricultural and Forest Meteorology*, *171*, 46–56.

Noormets, A. (2009). *Phenology of ecosystem processes: Applications in global change research*. Springer.

Oquist, G., & Huner, N. P. A. (2003). Photosynthesis of overwintering evergreen plants. *Annual Review of Plant Biology*, *54*, 329–355. https://doi.org/10.1146/annurev.arplant.54.072402.115741

Peaucelle, M., Janssens, I. A., Stocker, B. D., Descals Ferrando, A., Fu, Y. H., Molowny-Horas, R., Ciais, P., & Peñuelas, J. (2019). Spatial variance of spring phenology in temperate deciduous forests is constrained by background climatic conditions. *Nature Communications*, *10*(1), 5388. https://doi.org/10.1038/s41467-019-13365-1

Peñuelas, J., & Filella, I. (2001). Responses to a warming world. *Science*, *294*(5543), 793–795.

Peñuelas, J., Filella, I., Zhang, X., Llorens, L., Ogaya, R., Lloret, F., Comas, P., Estiarte, M., & Terradas, J. (2004). Complex spatiotemporal phenological shifts as a response to rainfall changes. *New Phytologist*, *161*(3), 837–846. https://doi.org/10.1111/j.1469-8137.2004.01003.x

Peñuelas, J., Rutishauser, T., & Filella, I. (2009). Phenology Feedbacks on Climate Change. *Science*, *324*(5929), 887–888. https://doi.org/10.1126/science.1173004

Repo, T., Leinonen, I., Ryyppö, A., & Finér, L. (2004). The effect of soil temperature on the bud phenology, chlorophyll fluorescence, carbohydrate content and cold hardiness of Norway spruce seedlings. *Physiologia Plantarum*, *121*(1), 93–100. https://doi.org/10.1111/j.0031-9317.2004.00307.x

Schwartz, M. D. (2003). *Phenology: An integrative environmental science*.

Semenchuk, P. R., Gillespie, M. A. K., Rumpf, S. B., Baggesen, N., Elberling, B., & Cooper, E. J. (2016). High Arctic plant phenology is determined by snowmelt patterns but duration of phenological periods is fixed: An example of periodicity. *Environmental Research Letters*, *11*(12), 125006. https://doi.org/10.1088/1748-9326/11/12/125006

Starr, G., Oberbauer, S. F., & Ahlquist, L. E. (2008). The Photosynthetic Response of Alaskan Tundra Plants to Increased Season Length and Soil Warming. *Arctic, Antarctic, and Alpine Research*, *40*(1), 181–191. https://doi.org/10.1657/1523-0430(06-015)[STARR]2.0.CO;2

Tanja, S., Berninger, F., Vesala, T., Markkanen, T., Hari, P., Mäkelä, A., Ilvesniemi, H., Hänninen, H., Nikinmaa, E., Huttula, T., Laurila, T., Aurela, M., Grelle, A., Lindroth, A., Arneth, A., Shibistova, O., & Lloyd, J. (2003). Air temperature triggers the recovery of evergreen boreal forest photosynthesis in spring. *Global Change Biology*, *9*(10), 1410–1426. https://doi.org/10.1046/j.1365-2486.2003.00597.x

Walther, S., Voigt, M., Thum, T., Gonsamo, A., Zhang, Y., Köhler, P., Jung, M., Varlagin, A., & Guanter, L. (2016). Satellite chlorophyll fluorescence measurements reveal large-scale decoupling of photosynthesis and greenness dynamics in boreal evergreen forests. *Global Change Biology*, *22*(9), 2979–2996.

Wang, Y., Mao, Z., Bakker, M. R., Kim, J. H., Brancheriau, L., Buatois, B., Leclerc, R., Selli, L., Rey, H., Jourdan, C., & Stokes, A. (2018). Linking conifer root growth and production to soil temperature and carbon supply in temperate forests. *Plant and Soil*, *426*(1–2), 33–50. https://doi.org/10.1007/s11104-018-3596-7

Wheeler, J. a., Gonzalez, N. m., & Stinson, K. a. (2016). Red hot maples: Acer rubrum first-year phenology and growth responses to soil warming. *Canadian Journal of Forest Research*, *47*(2), 159–165. https://doi.org/10.1139/cjfr-2016-0288

Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H., Hodges, J. C. F., Gao, F., Reed, B. C., & Huete, A. (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, *84*(3), 471–475. https://doi.org/10.1016/S0034-4257(02)00135-9

Zohner, C. M., Benito, B. M., Svenning, J.-C., & Renner, S. S. (2016). Day length unlikely to constrain climate-driven shifts in leaf-out times of northern woody plants. *Nature Climate Change*, *6*(12), 1120–1123. https://doi.org/10.1038/nclimate3138

Chapter 4

Climate-constrained boundaries cause nonuniform responses of vegetation phenology to climatic warming

Adrià Descals, Aleixandre Verger, Gaofei Yin, Iolanda Filella, Yongshuo H. Fu, Shilong Piao, Ivan A. Janssens, and Josep Peñuelas

Under review in *Global Change Biology* (2022)

Abstract

Climatic warming has lengthened the growing season in recent decades, thus affecting the functioning and biogeochemistry of ecosystems, the global carbon cycle and climate. Temperature response of phenology varies spatially and temporally, even within species, and daily total intensity of radiation and water availability may play a role. Whether radiation and water will constrain the trend towards longer growing seasons with future warming is uncertain. We empirically modelled the thresholds of temperature, radiation, and water under which daily carbon uptake is constrained and thereafter reduced the two-dimensionality of the temperature-radiation constraint to one single variable, θ, which represents the angle in a polar coordinate system for the temperature-radiation observations during the start and end of the growing season. In dryland ecosystems, water availability fully constrains the lengthening of the growing season, which indicates that future warming might reduce the length of the growing season by increasing water demand. In temperate and cold regions, we revealed that radiation is a major factor limiting photosynthetic activity that constrains the phenology response to temperature during the end-of-season. In contrast, the start of the growing season is overall highly sensitive to temperature but not constrained by radiation at the hemispheric scale. This study thus revealed that while at the end-of-season the phenology response to warming is constrained at the hemispheric scale, at the start-of-season the advance of spring onset may continue, even if it is at a slower pace.

4.1. Introduction

Seasonal changes in weather conditions drive the timing of the start and end of vegetation growth. With recent climatic warming, the growing season has lengthened, with the start of the growing season (SoS) advancing more than the end of the growing season (EoS) delaying(Menzel et al., 2006; Peñuelas & Filella, 2001). Vegetation phenology influences the carbon and water cycle(Keenan et al., 2014; Myneni et al., 1997; Richardson et al., 2010), successional transition(Chuine, 2010), and plays a role in the feedbacks between vegetation and climate(Peñuelas & Filella, 2009; Richardson et al., 2013). The warming-induced lengthening of the growing season has increased the carbon uptake(Le Quéré et al., 2009), offsetting atmospheric carbon from human emissions. However, how phenology may respond to future warming and whether vegetation will increase carbon sequestration remains unclear(Piao et al., 2019), which adds uncertainty to future atmospheric carbon concentration and, thus, climate projections.

Previous research suggests that chilling requirements will constrain the advance of leaf unfolding in deciduous forests by reducing temperature sensitivity(Fu et al., 2015), whereas other factors such as precipitation(Peaucelle et al., 2019) and photoperiod(Körner & Basler, 2010; Meng et al., 2021; Zohner et al., 2016) may increase the heat requirements during the ecodormancy stage and, as a result, slow down the warming-induced advance of the leaf unfolding. With regards to leaf senescence, a recent study found that increased productivity during the growing season counteracts the warming-induced delay in leaf senescence(Zani et al., 2020). These findings are based on the study of leaf phenophases in deciduous forests obtained from in situ observations or remotely-sensed vegetation greenness indices, such as the normalized vegetation difference index (NDVI).

The key conceptual framework in these studies is that pre-SoS and pre-EoS conditions affect the timing of tree phenophases(Chuine et al., 2013). For instance, prolonged high temperatures in spring advance the leaf unfolding date. However, a line of research has shown that the start and end of the photosynthetically active season are directly influenced by current meteorological conditions, not by pre-SoS or pre-EoS conditions. This is evidenced by a decoupling between remotely sensed vegetation greenness and proxies of photosynthetic activity(Jeong et al., 2017; Yin et al., 2020; Zhang et al., 2020), which indicates that vegetation might present leaves but these are not photosynthetically active because meteorological conditions are restricting

photosynthetic activity at the moment. For instance, vegetation productivity declines in agreement with the decrease in radiation intensity during autumn(Zhang et al., 2020), and a lack of available light may thus prompt the end of the growing season. Similarly, water availability determines the start and end of the photosynthetically active season in tropical dryland ecosystems(Eamus & Prior, 2001).

Given the link between photosynthesis and current meteorological conditions, the seasonality of carbon uptake can be modelled with meteorological variables, and the start and end of the growing season can be determined by constraint functions(Jolly et al., 2005). The constraint functions define the photosynthesis-inhibiting thresholds that restrict the carbon uptake during the growing season. The start and end of the season occur when a limiting factor, such as air temperature, radiation intensity, or water availability, exceeds the threshold under which photosynthesis is constrained. Using this phenology modelling framework, carbon uptake phenology can be thus determined by the law of limiting factors, i.e., by the most limiting factor.

Here, we aimed to clarify the role of climate constraints during the start and end of the photosynthetically active season and how these constraints change among ecosystems and over space. To achieve this, we first determined whether the law of limiting factors holds true for carbon uptake phenology. Then, we evaluated the current constraints on SoS and EoS by radiation and temperature in temperate and cold regions of the Northern Hemisphere and by water in tropical dryland ecosystems. Lastly, we estimated the potential lengthening of the growing season in a warming scenario in which radiation might take over the role as a limiting factor of photosynthetic activity phenology. To evaluate the limitations of temperature, radiation, and water, we empirically modelled the phenology of carbon uptake from FLUXNET towers using three climatic variables: daily mean air temperature, daily total incoming shortwave radiation, and mean soil water content.

4.2. Methods

4.2.1. Data

We used daily in situ records from 85 sites of the FLUXNET2015 Tier 1 data set (see Supplementary Figure 4.1 for the locations of the sites), daily mean GPP obtained using the day-time partitioning method observed at the FLUXNET towers. These 85 sites represent all sites >30°N in the tundra, boreal, and temperate biomes, excluding cropland sites. We used air

temperature and incoming shortwave radiation at 0.1 arc degrees from the ERA5-Land hourly data for the hemispheric scale analysis. Four-day clear-sky daily contiguous solar-induced chlorophyll fluorescence (CSIF) estimates at 0.05° from the OCO-2 (Orbiting Carbon Observatory-2) SIF data set were used to represent the spatial variation of vegetation phenology in temperate and cold regions of the Northern Hemisphere and tropical dryland regions for 2001-2020. SIF observations have been demonstrated to track the seasonality of GPP globally(Sun et al., 2017). The biome types at the FLUXNET sites were characterised using the MODIS MCD12Q1 V6 product for the mode of land cover during 2001-2019 and the RESOLVE Ecoregions 2017 map. The temperate and cold regions include tundra, boreal forests, temperate broadleaf and mixed forests, and temperate coniferous forests in the RESOLVE Ecoregions 2017 map (Supplementary Figure 4.1). We also masked the areas where the mean radiation during the winter solstice exceeded the 27 W m^{-2} threshold (Supplementary Figure 4.1). The mean radiation during the winter solstice was estimated using the average of shortwave radiation from the ERA5-Land observed from DoY 345 and 365 and aggregated over the 2001-2020 period. For the dryland ecosystems, we included the regions classified as tropical and subtropical grasslands, savannas and shrublands in the RESOLVE Ecoregions 2017 map.

4.2.2. Phenological estimation and modelling

Phenology was modelled using two approaches that produced similar estimates of the start of the growing season (SoS) and the end of the growing season (EoS). The first approach used a neural network that fitted the ERA5-Land daily mean air temperature, incoming shortwave radiation, and soil water content to the daily mean GPP. The fitting method was a neural network with two hidden layers with 8 and 2 neurons and trained using a gradient-descent optimisation algorithm. Each FLUXNET site obtains records with different temporal coverages, so we randomly selected two-year observations from each site to avoid overfitting the sites with the longest time series. We estimated the phenological dates using the modelled GPP time series generated from the climatic records from the FLUXNET sites. Then, we used the maximum-separation method(Descals et al., 2020) for estimating the phenological metrics (SoS and EoS). This method is robust against noise and can determine phenological metrics over raw time series, avoiding biases in phenological estimates due to the excessive smoothing of time series. The size of the window in the maximum-separation method was set to 60 days; different window sizes produced similar estimates of SoS and EoS.

The second approach for phenological modelling used a binary supervised classification, which labelled the FLUXNET daily records as class 'inhibited' when GPP was <2 and class 'active' when GPP was >2 g C m^{-2} d^{-1}. We used a predefined function (Eq. 1) as a decision boundary between the two classes, because the GPP modelled by the neural network had a decision boundary that followed this pattern (Figure 4.1a). The parameters of the decision boundary (T, S, and C) were estimated using an exhaustive grid search; the overall accuracy of all combinations of parameter C (representing the curvature of the function, from 50 to 300 in steps of 25), temperature (T, from -5 to 10 °C in steps of 0.1 °C), and radiation (S, from 0 to 150 W m^{-2} in steps of 1 W m^{-2}) were evaluated using the FLUXNET daily records, and we selected the parameters that fit best for overall accuracy. The overall accuracy was estimated as the ratio of correctly classified observations to the total number of observations. The confidence intervals of the parameters were calculated using bootstrapping; the parameters were estimated 100 times with random sampling with replacement, and the confidence interval represented the 5th and 95th percentiles of the distribution of the estimated parameters. We estimated the phenological dates from the binary time series (inhibited vs active) generated from the classification of the climatic variables at the FLUXNET towers. We used the maximum-separation method for extracting SoS and EoS, because this method can also estimate phenological metrics in a binary time series.

$$GPP = \frac{C}{Temperature - T} + S - Radiation \quad \text{(Eq. 1)}$$

The parameters T, S, and C were estimated using four fixed GPP thresholds (1, 2, 2.5, and 3). The phenological metrics reported in Figure 4.2, Supplementary Figure 4.2, and Supplementary Figure 4.3 represent the average of all SoS and EoS estimated using the classification approach.

We also used the maximum-separation method for the phenological estimates in the OCO2SIF time series. We used the averages of SoS and EoS estimated using four fixed SIF thresholds (10, 20, 25, and 30). For the potential dates on which SoS and EoS were restricted only by radiation (SoS_{pot} and EoS_{pot}), we changed the decision boundary in Eq. 1 to a constant radiation threshold of 27 W m^{-2}; the observation was considered inhibited when shortwave radiation was <27 W m^{-2} and active when it was >27 W m^{-2}. The 27 W m^{-2} represents the value S in Eq. 1 for the GPP threshold equal to 2. The potential SoS and EoS were estimated using the binary time series generated from the shortwave radiation records in the ERA5-Land data set.

The GPP modelling in tropical dryland ecosystem was done with OCO2-SIF time series, as a proxy of GPP, and with air temperature, shortwave radiation, and soil moisture at 7-28 cm depth extracted from the ERA5-Land dataset. The time series were extracted from 56 BELMANIP2 sites that are classified as tropical and subtropical grasslands, savannas and shrublands in the RESOLVE Ecoregions 2017 map. BELMANIP2(Weiss et al., 2014) is a network of sites that present a high landscape homogeneity and is used for the calibration and validation of land satellite products. Given the heterogeneity of soil moisture contents in different sites, we scaled the soil moisture time series such that the minimum soil moisture was 0, and the maximum was 1. Thus, we estimated the phenological dates as percentages of the soil moisture amplitude. We found that the start and end of the season corresponded to the 20% threshold of the soil moisture time series.

4.2.3. Estimation of constraints on phenological dates by temperature and radiation

The restriction of the onset and end of carbon uptake by temperature and radiation was estimated geometrically (Eq. 2) based on the decision boundary (Eq. 1) obtained in the classification approach. Eq. 2 calculates the angle θ given by the vertices A, B, and C, where A and B have fixed coordinates in (T, S) and (T, 25), respectively. The coordinates of C are given by the temperature (TA) and shortwave radiation (SW) at the time of the phenological event. A graphical representation of the estimates of θ is shown in Supplementary Figure 4.7; a high temperature and low amount of radiation at the time of the phenological event lead to low values of θ proximal to 0°, and a low temperature and high amount of radiation at the time of the phenological event lead to high values of θ proximal to 90°. Values of θ <0 or >90° were clamped to 0 and 90°, respectively.

$$\theta = \tan^{-1}\left(\frac{25\ x\ (SW-S)}{300\ x\ (TA-T)}\right) \qquad \text{(Eq. 2)}$$

4.2.4. Estimation of temperature and aridity dependency

The dependency of SoS and EoS to temperature was estimated for 2001-2020 using ERA5-Land and the phenological dates extracted from the OCO2SIF time series. We used two different time periods for estimating the mean air temperature: 15 and 30 days before the mean phenological date. We used Pearson correlations, which is termed temperature dependency, instead of the conventional metric of temperature sensitivity(Fu et al., 2015), because the latter overestimates

the sensitivity when the temperature time series is highly variable(Keenan et al., 2020), which could potentially weaken the spatial analysis in our study. The temperature sensitivity estimated with the OCO2SIF dataset gave the same results as the temperature dependency; temperature sensitivity was higher in the SoS than the EoS, and it was substantially lower in regions where radiation is constraining vegetation activity. In tropical dryland ecosystems, we investigated the dependency of phenological dates to the aridity index (the ratio of precipitation (P) to potential evapotranspiration (PET)) calculated for the 15 days before the mean phenological date. The precipitation was extracted from the ERA5-Land dataset and the potential evapotranspiration was calculated using a model adapted for the ERA5-Land dataset(Singer et al., 2021).

4.3. Results

4.3.1. Limiting factors in the phenology of carbon uptake

We first modelled gross primary production (GPP) in each FLUXNET site as a function of temperature and radiation. If the law of the limiting factor holds true, the seasonality of GPP would be fully explained with meteorological variables, and the difference between simulated and observed GPP during pre-SoS (15 days before the start of the season) and post-EoS (15 days after the end of the season) would be close to 0. Simulated GPP during pre-SoS presented a low bias compared to observed GPP in evergreen needleleaved forests for pre-SoS and post-EoS (+0.50 g C m^{-2} d^{-1} and +0.16 g C m^{-2} d^{-1}, respectively) (Supplementary Figure 4.2). This indicates that photosynthesis ceased as soon as temperature and radiation conditions were unfavourable. In contrast, deciduous forests presented the largest bias during the pre-SoS (+1.2 g C m^{-2} d^{-1}), indicating that conditions were favourable for photosynthesis but the ecosystem remained dormant. The bias was also positive, but lower, for simulated carbon uptake during post-EoS (+0.49 g C m^{-2} d^{-1}), suggesting that deciduous trees prepared for leaf senescence before conditions became unfavourable for vegetation growth.

Second, we used the FLUXNET observations to train a single model that explains carbon uptake as a function of temperature and radiation in the temperate and cold regions of the Northern Hemisphere. FLUXNET data records show that photosynthetic activity occurs within specific ranges of temperature and radiation (Figure 4.1a). The constraint on photosynthetic activity, when GPP is <2 g C m^{-2} d^{-1}, can be formulated as a rational function (hereafter constraint boundary) with asymptotes for temperature at 2.5 (2.0, 3.4) °C (95% confidence interval) and for shortwave radiation at 27 (11, 54) W m^{-2}. We used the constraint boundary as a model that

determines in a binary fashion whether conditions are favourable for vegetation growth or, contrarily, temperature and radiation restrict daily mean GPP. The constraint boundary was able to discriminate between dormant and growing season with an overall classification accuracy of 84.7 %, which remained similar for different vegetation types (Supplementary Table 4.1). Overall accuracy was highest in evergreen needleleaved forests (87.9 %).

Thirdly, we determined the start and end of the carbon uptake using the constraint boundary. We estimated the SoS as the timing of the year when climatic conditions become favourable for vegetation activity, and the EoS as the time of the year when these conditions cease. The SoS and EoS were simulated in the FLUXNET sites using climatic time series and then compared to the observed SoS and EoS from the FLUXNET GPP. A high agreement between simulated and observed phenological dates would indicate that the law of the minimum holds true in the regulation of carbon uptake phenology. The estimated and observed phenological metrics agreed well in evergreen needleleaved forests for a wide range of latitudes, regardless of biome (Root mean squared error (RMSE) was 9.3 d and 10.4 d, and Mean Error (ME) was -9.7 d and 13.0 d for SoS and EoS, respectively) (Supplementary Figure 4.3); constraints of radiation and temperature explained the SoS and EoS in the FLUXNET site San Rossore (IT-SRo), with a Mediterranean climate, as well as the FLUXNET site Saskatchewan (Ca-Obs), covering a boreal forest. The largest bias was found for SoS in deciduous forests (RMSE was 11.9 d and 12.6 d, and ME was 27.2 d and -13.2 d for SoS and EoS, respectively).

Lastly, we used remotely-sensed OCO2-SIF (Orbiting Carbon Observatory-2 solar-induced chlorophyll fluorescence), a proxy of GPP, to support the findings in the FLUXNET sites and extend the analysis to water-restricted regions in the tropics. For these dryland ecosystems, we included temperature, radiation, and soil moisture records as predictive variables in the GPP model. Changes in soil moisture influenced GPP seasonality (Figure 4.1d), while temperature and radiation had negligible effects on the carbon uptake. The growing season started when soil moisture increased rapidly, when the rainy season began, and ended when soil moisture returned to low levels (Figure 4.1e). The spatial patterns of observed and simulated SoS and EoS show the same spatial patterns in temperate and cold regions of the Northern Hemisphere and tropical dryland regions (Supplementary Figure 4.4 and 4.5), further corroborating that carbon uptake phenology can be determined with the most limiting meteorological factor.

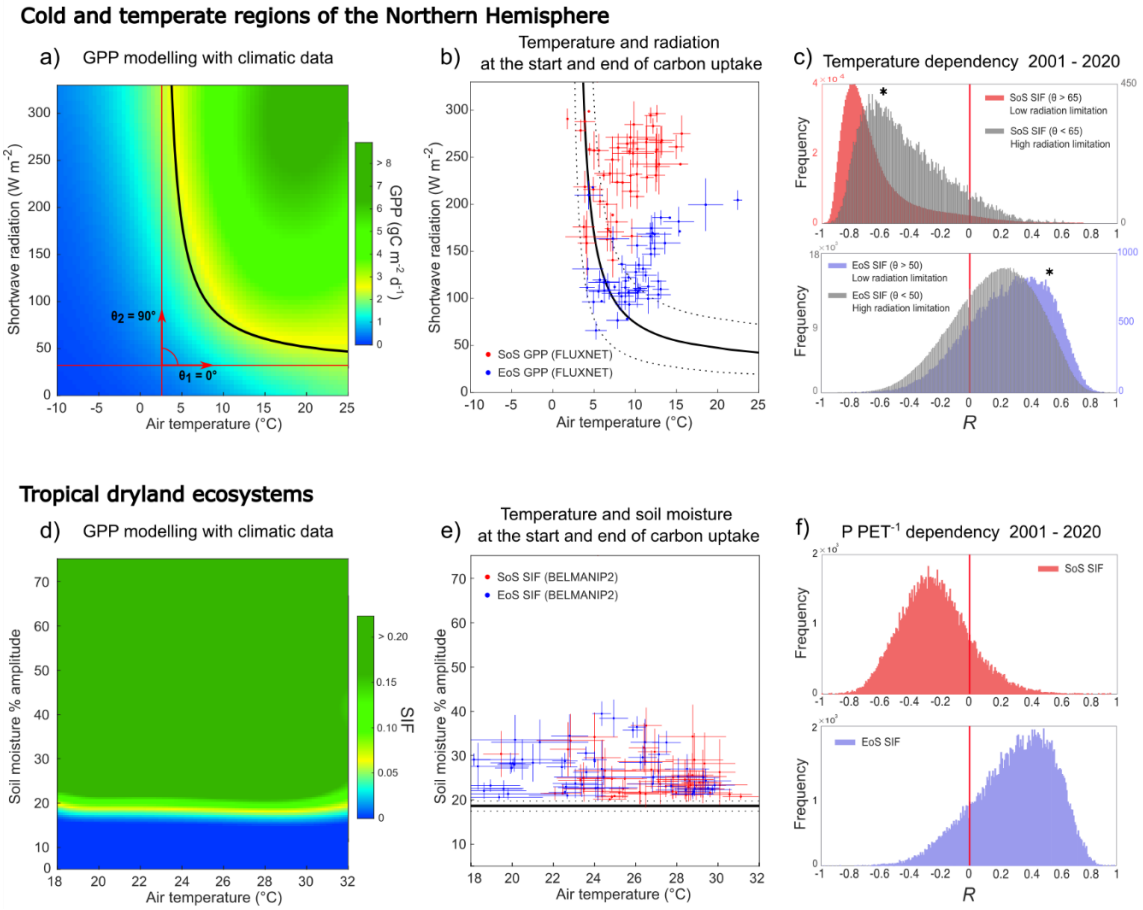


Figure 4.1. Gross Primary Production (GPP) modelling and the relationships of the start and end of the growing season (SoS and EoS, respectively) with climate factors. a) Estimated GPP for different air temperatures and amounts of shortwave radiation for the temperate and cold regions of the Northern Hemisphere, and d) for different air temperature and soil moisture levels in tropical dryland ecosystems. The restriction of the phenological dates by temperature and radiation was estimated using the angle θ, representing the gradients of temperature and radiation in constraining GPP (θ_1 and θ_2 represent the highest constraints on GPP by radiation and temperature, respectively). b) Climatic factors during SoS and EoS obtained from the GPP time series in the FLUXNET sites, and e) from SIF time series in the BELMANIP2 sites. Error bars show two standard deviations. The black line depicts the constraint boundary between GPP below and above 2 g C m^{-2} d^{-1} in b), and SIF below and above 0.05 in e). Dotted lines show the 95% confidence interval for the constraint boundary. c) Histograms of the coefficients (*R*) for the correlation between temperature (ERA5-Land) and phenological dates (estimated from the OCO-2 SIF time series) during 2001-2020. The asterisk represents that mean *R* was significantly different (p-value < 0.05) in the regions with the highest radiation constrain during the SoS (θ < 65) and in the regions with the highest temperature constrain during the EoS (θ > 50). The threshold of θ for the SoS represents the 5th percentile of all mean θ values, while the threshold of θ for the EoS is the 95th percentile of the mean θ values. f) Histograms of *R* between the aridity index (the ratio of precipitation (P) to potential evapotranspiration (PET)) and phenological dates in tropical dryland ecosystems. The simulated SIF values in d) were generated with a constant radiation of 250 W m^{-2}, which represents the approximate mean radiation in tropical dryland ecosystems.

4.3.2. Current climate constraints on carbon uptake phenology

Temperature and radiation during SoS and EoS at the FLUXNET sites were within the ranges defined by the constraint boundary (Figure 4.1b and Supplementary Figure 4.6), although on average the daily mean radiation differed between SoS (air temperature, 7.2 °C; shortwave radiation, 213 W m^{-2}) and EoS (air temperature, 8.0 °C; shortwave radiation, 107 W m^{-2}), which suggests a contrasting limitation of temperature and radiation during the start and the end of the carbon uptake season in temperate and cold regions of the Northern Hemisphere. To quantify the limitations of radiation and temperature, we reduced the two-dimensionality of the temperature-radiation constraint to one single variable; θ, which represents the angle in polar coordinate system for the temperature-radiation observations on a range from 0 to 90 (0 = fully restricted by radiation and 90 = fully restricted by temperature; See Methods and Supplementary Figure 4.7). Values of θ in temperate and cold regions were on average high during the spring onset (median and 5 - 95th percentile range was 87 [65, 98]) and low during the end of season (32 [14, 50]) (Figure 4.2b and Supplementary Figure 4.8), which indicates a high temperature limitation on the carbon uptake during the SoS and a moderate radiation limitation during the EoS. The regions where radiation limitation was highest during the SoS and EoS correspond to temperate Europe (Supplementary Figure 4.8), while temperature restricted the SoS and EoS in high-latitude and high-altitude regions. In dryland ecosystems, SIF phenology was determined by the 20% threshold in soil moisture, and temperature and radiation did not exhibit a constraint boundary.

The limitation of radiation to photosynthesis was linked to a low sensitivity of phenology to temperature in temperate and cold regions of the Northern Hemisphere. Temperature dependency was estimated using the coefficient of correlation between temperature and phenological dates obtained from OCO2SIF time series, which provides more spatial and temporal coverage (2001-2020) than the FLUXNET GPP records. SoS was on average more dependent than EoS to temperature. The dependency of phenological dates to temperature was significantly lower when the limitation of radiation was high (Figure 4.1c). Regions with the highest radiation limitation during the SoS, where values of θ were lower than the 5th percentile (θ < 65), presented a lower temperature dependency (R = -0.50) than the rest of the regions (R = -0.63). Similarly, regions with the lowest radiation limitation during the EoS, defined as values greater than the 95th percentile (θ > 50), exhibited a greater temperature dependency (R = 0.23) than the rest of the regions (R = 0.16). In tropical dryland ecosystems, SoS and EoS were

dependent to the aridity index (the ratio of precipitation (P) to potential evapotranspiration (PET)). The dependency was negative during the SoS (R = -0.23) and positive during the EoS (R = 0.31) (Figure 4.1f), indicating that increasing aridity delays the SoS and advances the EoS and, thus, shortens the length of the growing season.

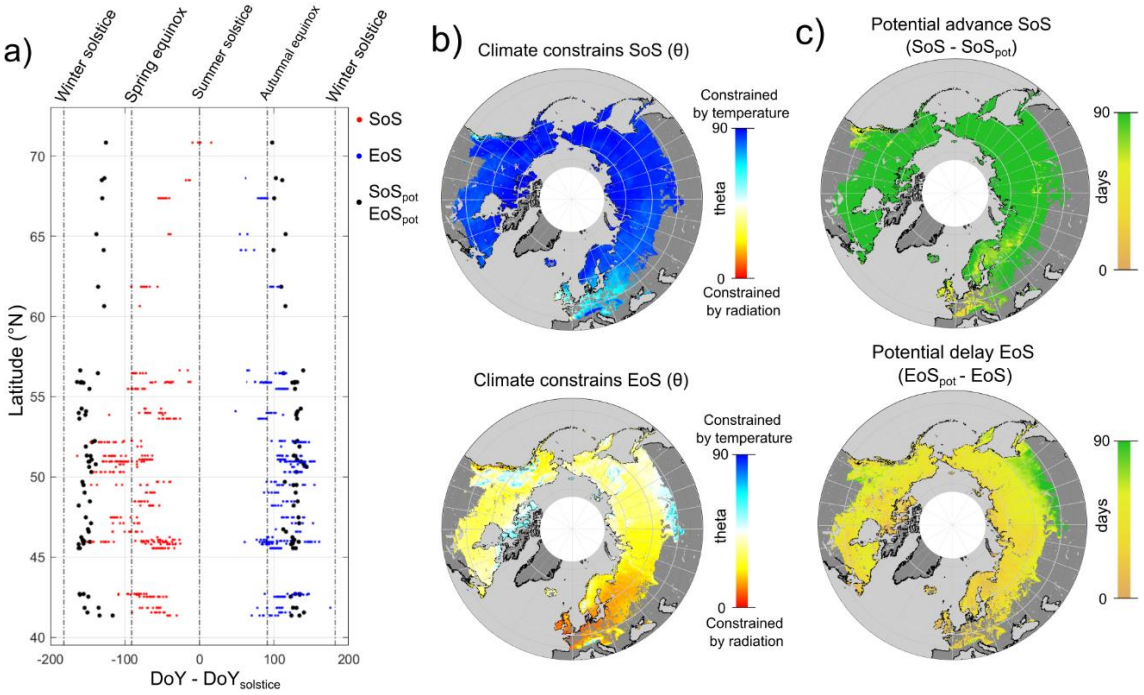


Figure 4.2. Potential lengthening of the growing season in temperate and cold regions of the Northern Hemisphere. a) Latitudinal gradients of the observed SoS and EoS, obtained from the GPP time series at the FLUXNET towers, and the mean potential start of season SoS_{pot} and mean potential end of season EoS_{pot}, which represent the dates when shortwave radiation is surpasses the 27 W m^{-2}. We offset the Day of Year such as the summer solstice ($DoY_{solstice}$) is the origin of the axis. b) Magnitudes of the constraints on SoS and EoS by radiation and temperature. The constraints (θ) on GPP are associated with the gradients of temperature and radiation; the temperature gradient increases with θ, and the radiation gradient decreases with θ. c) Potential advance of SoS and potential delay of EoS estimated using the ERA5-Land and OCO-2 SIF data sets for 2001-2020.

4.3.3. Potential lengthening of the growing season in temperate and cold regions of the Northern Hemisphere

We defined the dates when radiation could potentially become the climatic variable that fully restricted photosynthetic activity during spring onset and autumnal senescence as SoS_{pot} and EoS_{pot}, representing the days of the year when daily shortwave radiation reaches the 27 W m^{-2} threshold in spring and autumn. The 27 W m^{-2} threshold represents the lowest constraint boundary of radiation (Fig 1a). The current EoS was generally similar to the predicted EoS_{pot} based on the FLUXNET and gridded climatic data (Figure 4.2a, c) (mean difference EoS_{pot} – EoS

was 23.6 d for FLUXNET), indicating a strong constraint of radiation on GPP senescence in the Northern Hemisphere (Figure 4.2b) and suggesting that future climatic warming cannot delay the end of the photosynthetically growing season due to low levels of radiation in autumn. The observed SoS estimated using FLUXNET and gridded climatic data, in contrast, was far from reaching SoS_{pot} at the hemispheric scale (Figure 4.2c) (mean difference SoS_{pot} – SoS was -79.3 d for FLUXNET). SoS was only close to SoS_{pot} in Europe. These findings suggest that the advance of spring onset was overall not limited by radiation, and future warming can potentially lead to an advance of the spring onset of carbon uptake in evergreen needleaved forests, which respond promptly to favourable temperature and radiation conditions.

4.4. Discussion

Our results indicate that the most limiting factor determines the start and end of the photosynthetically active season. Carbon uptake by vegetation begins soon after conditions for growth become favorable and ceases when these conditions decline. Limiting factors are temperature and radiation in temperate and cold regions, and water availability in dryland ecosystems. We showed that climate constraint boundaries, which define the threshold under which photosynthesis is inhibited, can differentiate between active and restricted vegetation stages and, thus, determine the phenological dates across latitudes in evergreen needleaved forests. However, our findings were inconclusive regarding whether limiting factors determined the start and end of the growing season in deciduous forests.

The cold-avoidance strategy of deciduous forests may explain their exception to the law of the limiting factors in the phenology of the carbon uptake. Deciduous trees respond differently than evergreen conifers to cold winters. Most conifers in temperate and cold regions have a cold-tolerant strategy. Leaves in conifers resist cold, reject excessive radiation when temperatures are still low in the spring, and resume photosynthetic activity as soon as climatic conditions are favorable(Chang et al., 2021). In contrast, deciduous trees shed their leaves in autumn and leaf budbreak in spring (Chuine et al., 2013). This would account for our findings that deciduous trees remain dormant during the spring, even when photosynthesis conditions are favorable. The ecodormancy stage is only broken when a certain amount of heat is accumulated. Such heat requirements would explain the long difference between modelled SoS and flux-based SoS in deciduous forests. Thus, previous weather conditions affect the phenophases of deciduous

species, but not the carbon uptake phenology, which is determined by current weather conditions.

Temperature is the main limiting factor during spring onset, but temperature and radiation may have been limiting during autumn in temperate and cold regions of the Northern Hemisphere(Zhang et al., 2020). Temperature is seasonally asymmetrical, with lower temperatures before the summer solstice than after. That is because leafless temperate deciduous forests and snow-covered ecosystems have a high albedo in the springtime. Tundra FLUXNET sites provide key examples; daily mean air temperatures remained <2.5 °C just before the summer solstice, when annual radiation was at its peak. Because spring is colder than autumn, the growing season begins closer to the summer solstice than the senescence stage does. During the EoS, the radiation constraint on photosynthetic activity is increased because autumn temperatures remain high. This suggests that, as a result of future warming, radiation constraints will increase further in autumn, limiting the delay of the EoS.

The strongest radiation constraint was found in temperate Europe, where the maritime influence kept temperatures favourable for vegetation growth in autumn. In contrast, temperature mostly constrained carbon uptake phenology in high-elevated regions, which suggests that warming could potentially advance the SoS and delay the EoS in alpine ecosystems more than radiation-restricted vegetation. This further support the claims of spatial homogeneity in spring phenology(Liu et al., 2019) and the reduction of the altitudinal gradient in phenological dates as a result of global warming(Vitasse et al., 2018). With climatic warming, high-elevated vegetation will lengthen the growing season more than low-elevated vegetation, equalizing phenological dates across elevations.

Our findings have important implications for phenology modelling and the study of the global carbon cycle and the vegetation responses to climatic warming. Phenophase observations are straightforward to collect, which may account for their long records. Phenology research has had a major focus on the timing of phenophases in temperate deciduous forests(Piao et al., 2019). The PEP725, the most widely used in situ phenological database, mostly contains phenophase records of deciduous species in Central Europe. However, we showed that the climate of Europe has milder temperatures than other areas of the same latitude, exacerbating the limitation of radiation in both spring and autumn. Secondly, deciduous forests have high heat requirements for their release from dormancy that differ from other vegetation types,

which explains their different responses to favorable weather conditions during the spring onset. Lastly, the coverage of deciduous trees and mixed forests represents only 18% of the temperate and cold regions of the Northern Hemisphere (Supplementary Figure 4.1). All these factors indicate that forecasted shifts in phenology under climatic warming and projections of carbon sequestration by vegetation, based on the sole modelling of phenophases of deciduous trees in Central Europe, cannot be scaled to other biomes of the Northern Hemisphere. Our findings highlight the importance of extensive carbon uptake observations for the study of the global carbon cycle. Carbon uptake phenology is more difficult to track than tree phenophases, and only in situ (FLUXNET) and satellite (SIF) observations have been made extensively over the last two decades.

Radiation constraints explain the nonuniform seasonal temperature dependency during the SoS and EoS and across the temperate and cold regions, indicating divergent responses of vegetation phenology to future climatic warming. The dependency of the SoS and EoS to temperature was significantly lower in regions where radiation constraints were the highest. These results indicate that the sensitivity of vegetation phenology to temperature decreases with the increasing constraints of radiation. Our results show that the senescence stage has a low temperature dependency due to the constraints of radiation, and this might be a reason for the lower magnitude in the EoS delay than the SoS advance(Menzel et al., 2006). SoS could continue its advance with future warming, even if it is at a lower pace(Fu et al., 2015), although the advance of spring onset would be highly spatially variable depending on the local radiation limitations. In tropical dryland ecosystems, carbon uptake phenology is sensitive to the aridity index. This indicates that the length of the growing season would shorten with climatic warming, as the warming-induced increase in potential evapotranspiration will be greater than the increase in precipitation(Sherwood & Fu, 2014). The constraints imposed by radiation and water should be considered in modelling future phenological responses to climatic warming because the length of the growing season may be overestimated if temperature is the only climatic factor considered.

4.5. References

Chang, C. Y.-Y., Bräutigam, K., Hüner, N. P., & Ensminger, I. (2021). Champions of winter survival: Cold acclimation and molecular regulation of cold hardiness in evergreen conifers. *New Phytologist*, *229*(2), 675–691.

Chuine, I. (2010). Why does phenology drive species distribution? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1555), 3149–3160.

Chuine, I., de Cortazar-Atauri, I. G., Kramer, K., & Hänninen, H. (2013). Plant development models. In *Phenology: An integrative environmental science* (pp. 275–293). Springer.

Descals, A., Verger, A., Yin, G., & Penuelas, J. (2020). A threshold method for robust and fast estimation of land-surface phenology using Google Earth Engine. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

Eamus, D., & Prior, L. (2001). *Ecophysiology of trees of seasonally dry tropics: Comparisons among phenologies*.

Fu, Y. H., Zhao, H., Piao, S., Peaucelle, M., Peng, S., Zhou, G., Ciais, P., Huang, M., Menzel, A., Peñuelas, J., Song, Y., Vitasse, Y., Zeng, Z., & Janssens, I. A. (2015). Declining global warming effects on the phenology of spring leaf unfolding. *Nature*, *526*(7571), 104–107. https://doi.org/10.1038/nature15402

Jeong, S.-J., Schimel, D., Frankenberg, C., Drewry, D. T., Fisher, J. B., Verma, M., Berry, J. A., Lee, J.-E., & Joiner, J. (2017). Application of satellite solar-induced chlorophyll fluorescence to understanding large-scale variations in vegetation phenology and function over northern high latitude forests. *Remote Sensing of Environment*, *190*, 178–187.

Jolly, W. M., Nemani, R., & Running, S. W. (2005). A generalized, bioclimatic index to predict foliar phenology in response to climate. *Global Change Biology*, *11*(4), 619–632.

Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., Munger, J. W., O'Keefe, J., Schmid, H. P., Wing, I. S., & others. (2014). Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nature Climate Change*, *4*(7), 598–604.

Keenan, T. F., Richardson, A. D., & Hufkens, K. (2020). On quantifying the apparent temperature sensitivity of plant phenology. *New Phytologist*, *225*(2), 1033–1040.

Körner, C., & Basler, D. (2010). Phenology under global warming. *Science*, *327*(5972), 1461–1462.

Le Quéré, C., Raupach, M. R., Canadell, J. G., Marland, G., Bopp, L., Ciais, P., Conway, T. J., Doney, S. C., Feely, R. A., Foster, P., & others. (2009). Trends in the sources and sinks of carbon dioxide. *Nature Geoscience*, *2*(12), 831–836.

Liu, Q., Piao, S., Fu, Y. H., Gao, M., Peñuelas, J., & Janssens, I. A. (2019). Climatic warming increases spatial synchrony in spring vegetation phenology across the Northern Hemisphere. *Geophysical Research Letters*, *46*(3), 1641–1650.

Meng, L., Zhou, Y., Gu, L., Richardson, A. D., Peñuelas, J., Fu, Y., Wang, Y., Asrar, G. R., De Boeck, H. J., Mao, J., & others. (2021). Photoperiod decelerates the advance of spring phenology of six deciduous tree species under climate warming. *Global Change Biology*.

Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aasa, A., Ahas, R., Alm-Kübler, K., Bissolli, P., Braslavská, O., Briede, A., & others. (2006). European phenological response to climate change matches the warming pattern. *Global Change Biology*, *12*(10), 1969–1976.

Myneni, R. B., Keeling, C., Tucker, C. J., Asrar, G., & Nemani, R. R. (1997). Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature*, *386*(6626), 698–702.

Peaucelle, M., Janssens, I. A., Stocker, B. D., Ferrando, A. D., Fu, Y. H., Molowny-Horas, R., Ciais, P., & Peñuelas, J. (2019). Spatial variance of spring phenology in temperate deciduous forests is constrained by background climatic conditions. *Nature Communications*, *10*(1), 1–10.

Peñuelas, J., & Filella, I. (2001). Responses to a warming world. *Science*, *294*(5543), 793–795.

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

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., & Zhu, X. (2019). Plant phenology and global climate change: Current progresses and challenges. *Global Change Biology*, *25*(6), 1922–1940. https://doi.org/10.1111/gcb.14619

Richardson, A. D., Andy Black, T., Ciais, P., Delbart, N., Friedl, M. A., Gobron, N., Hollinger, D. Y., Kutsch, W. L., Longdoz, B., Luyssaert, S., & others. (2010). Influence of spring and autumn phenological transitions on forest ecosystem productivity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1555), 3227–3246.

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., & Toomey, M. (2013). Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agricultural and Forest Meteorology*, *169*, 156–173. https://doi.org/10.1016/j.agrformet.2012.09.012

Sherwood, S., & Fu, Q. (2014). A drier future? *Science*, *343*(6172), 737–739.

Singer, M. B., Asfaw, D. T., Rosolem, R., Cuthbert, M. O., Miralles, D. G., MacLeod, D., Quichimbo, E. A., & Michaelides, K. (2021). Hourly potential evapotranspiration at 0.1° resolution for the global land surface from 1981-present. *Scientific Data*, *8*(1), 1–13.

Sun, Y., Frankenberg, C., Wood, J. D., Schimel, D., Jung, M., Guanter, L., Drewry, D., Verma, M., Porcar-Castell, A., Griffis, T. J., & others. (2017). OCO-2 advances photosynthesis observation from space via solar-induced chlorophyll fluorescence. *Science*, *358*(6360).

Vitasse, Y., Signarbieux, C., & Fu, Y. H. (2018). Global warming leads to more uniform spring phenology across elevations. *Proceedings of the National Academy of Sciences*, *115*(5), 1004–1008.

Weiss, M., Baret, F., & Verger, A. (2014). BELMANIP2: Enhancement of the CEOS-BELMANIP ensemble of sites used for the validation of land products from medium resolution sensors. *Fourth International Symposium on Recent Advances in Quantitative Remote Sensing*.

Yin, G., Verger, A., Filella, I., Descals, A., & Peñuelas, J. (2020). Divergent estimates of forest photosynthetic phenology using structural and physiological vegetation indices. *Geophysical Research Letters*, *47*(18), e2020GL089167.

Zani, D., Crowther, T. W., Mo, L., Renner, S. S., & Zohner, C. M. (2020). Increased growing-season productivity drives earlier autumn leaf senescence in temperate trees. *Science*, *370*(6520), 1066–1071.

Zhang, Y., Commane, R., Zhou, S., Williams, A. P., & Gentine, P. (2020). Light limitation regulates the response of autumn terrestrial carbon uptake to warming. *Nature Climate Change*, *10*(8), 739–743.

Zohner, C. M., Benito, B. M., Svenning, J.-C., & Renner, S. S. (2016). Day length unlikely to constrain climate-driven shifts in leaf-out times of northern woody plants. *Nature Climate Change*, *6*(12), 1120–1123.

Chapter 5

Widespread drought-induced leaf shedding and subsequent decline of productivity in European temperate deciduous forests

Adrià Descals, Aleixandre Verger, Gaofei Yin, Iolanda Filella, and Josep Peñuelas

Under review in *Remote Sensing in Ecology and Conservation* (2022)

Abstract

Heatwaves and droughts are becoming more common and severe in Europe, causing changes in tree phenology, disrupting the sequestration of carbon, and causing tree mortality on a continental scale. The responses of leaf shedding to heatwaves and droughts remain uncertain, although temperate deciduous forests may shed their leaves if exposed to extreme heat and water stress. Little information, however, is available about the extent of early leaf shedding induced by drought, likely because it occurs in small forest patches and can be discriminated only during a few weeks. We used highly spatiotemporal Sentinel-2 data as evidence of widespread drought-induced early leaf shedding in Europe from 2017 to 2021. The 10-m resolution Sentinel-2 data identified early leaf shedding not detected by a low-resolution (500 m) MODIS sensor. Early leaf shedding was identified in 90 locations throughout Europe during 2017-2021, and its occurrence was linked to preceding anomalously high temperatures and arid conditions. Our results also indicated that mean summer NDVI decreased significantly in the years following early leaf shedding, suggesting a legacy decline in vegetation productivity. Our study demonstrates that decametric satellite data can be used to monitor the responses of forests to drought at the canopy level and indicates that early leaf shedding associated with heatwaves is more widespread and frequent across the continent than previously thought.

5.1. Introduction

Increases in temperature have lengthened the growing season over the last few decades (Menzel et al., 2006); leaf unfolding has advanced and leaf shedding has delayed in central Europe. These changes in vegetation phenology have increased the uptake of carbon by vegetation (Keenan et al., 2014). How vegetation phenology will respond to future climatic warming and to extreme climate, however, is unclear, particularly for leaf shedding, which remains insufficiently studied (Gallinat et al., 2015). Growing evidence suggests that heatwaves and droughts reduce the sequestration of carbon by vegetation (Bastos et al., 2020; Ciais et al., 2005), and heatwaves such as the European heatwave in 2003 have legacy effects and reduce carbon sequestration in the following years (Ciais et al., 2005). Drought is also an important cause of tree mortality in Europe, and drought is expected to worsen, potentially increasing tree mortality (Senf et al., 2020). Studies on the response of leaf shedding to heatwaves and droughts, however, have provided divergent results, and a deeper understanding of leaf shedding is needed to forecast atmospheric carbon levels and reduce uncertainties in climatic projections.

One line of research suggests that leaf shedding may be delayed in response to high temperatures in autumn and the presence of heatwaves and droughts. A study in North America using remotely sensed data indicated that drought stress delayed the start of dormancy (Xie et al., 2015). Another study found a positive correlation between the end of the growing season, estimated from satellite data, and a drought index in temperate biomes (Bórnez et al., 2021), and in situ measurements indicated that foliar longevity in deciduous trees was greater during the 2003 European heatwave than in previous years (Leuzinger et al., 2005). Manipulative experiments found that heat stress or dry air had no effect on the start of leaf shedding. (Mariën et al., 2021).

Another line of research suggests that the timing of leaf shedding in temperate deciduous forests cannot be delayed further, because photoperiod triggers leaf shedding in autumn (Way & Montgomery, 2015). Leaf shedding may even advance with climatic warming because shedding is linked to increased summer productivity (Zani et al., 2020), which is expected to increase with rising temperatures. Drought-induced early leaf shedding has been reported from local observations, some as far back as the 1913 drought in the United States of America (Kozlowski, 1976). Early leaf shedding in response to heatwaves and droughts, however, has

been poorly documented using remotely-sensed data, and only one recent study found satellite evidence of early leaf shedding during the 2018 European heatwave (Brun et al., 2020).

Land-surface phenology (LSP) is studied using the seasonality of indices of vegetation greenness and links observations from satellites with in situ observations of phenophases (Bornez et al., 2020). LSP studies have focused mainly on moderate- to low-resolution satellite data (spatial resolution >100 m), including data from MODIS (Zhang et al., 2003), AVHRR (Julien & Sobrino, 2009), and Proba-V (Bornez et al., 2020), among other satellite sensors. The spatial resolution of remotely sensed satellite data, however, has an impact on the estimation of phenological metrics (Hmimina et al., 2013), and low resolution may hamper the detection of early leaf shedding; LSP metrics at moderate resolutions generally delay the date of the end of the growing season (Hmimina et al., 2013), particularly in heterogeneous landscapes. The date of early leaf shedding from moderate-resolution satellite data may thus be underestimated or remain undetected when shedding occurs in small forest patches.

The recent launch of decametric-resolution satellites with short revisiting times enables the extraction of phenological data at the canopy scale. Sentinel-2 provides images at 10-m resolution. The revisiting time is 5 d at the equator but increases with latitude and can provide daily observations above 65° latitude (Descals et al., 2020). Previous studies have estimated LSP metrics at the canopy level on a continental scale (Bolton et al., 2020; Descals et al., 2020) and have linked LSP metrics with in situ observations of phenophases and PhenoCam time series (Tian et al., 2021). High-resolution satellite data can potentially improve the monitoring of future responses of forests to climate change (Hartmann et al., 2022); Sentinel-2 may enable the detection of early leaf shedding on a continental scale and for the years covered by the satellite mission, quantifying the advance of early leaf shedding.

A recent study explored the feasibility of using 10-m Sentinel-2 data for the detection of early leaf shedding during the 2018 European heatwave (Brun et al., 2020). Whether the occurrence will be limited to that specific heatwave or whether it will be more widespread and recurrent across the continent, however, is unclear. We studied the occurrence of early leaf shedding and its link to summer meteorological conditions in European temperate deciduous forests using Sentinel-2 data for widespread early leaf shedding across Europe during 2017-2021. We quantified the evidence of shifts in the dates of leaf shedding across the continent and compared them to low-resolution phenological data from MODIS. We inspected the legacy effects on the

mean normalised difference vegetation index (NDVI) in summer in the years after the early leaf shedding. Finally, we investigated the link between the dates of early leaf shedding and two climatic variables, temperature and aridity, to determine whether leaf shedding advances or delays as a response of summer meteorological conditions.

5.2. Methods

5.2.1. Study area

We investigated the occurrence of early leaf shedding in Europe for 2017-2021. The study area was limited to Europe because it is the only region covered by Sentinel-2 Level 2A for this period, with the rest of the globe only covered for 2019-2021. The study area covered the southern and central distribution of temperate deciduous forests in Europe, between 43 and 55°N and 3°W and 36°E, excluding the northeastern part of the boundary. This slightly longer coverage allowed a more robust temporal analysis that linked early leaf shedding and summer climatic conditions. The region experienced a severe heatwave and drought in central Europe in 2018 (Bastos et al., 2020) and in the Carpathians in 2019. The climate of the region is mostly temperate, but the southern part has a Mediterranean climate, where arid conditions in summer are more severe than in central Europe. This difference in climate accounts for the different composition of tree species in the study area; drought-tolerant deciduous species such as *Quercus faginea* and *Q. frainetto* grow under the Mediterranean climate in southern Europe, and common temperate deciduous species such as *Q. ruber* and *Fagus sylvatica* grow primarily at northern latitudes or highly elevated areas that have less severe arid conditions.

5.2.2. LSP estimation with Sentinel-2

We processed the Sentinel-2 Level-2A (surface reflectance) time series and estimated the end of the growing season (EoS) in the study area for 2017-2021. EoS was estimated using the threshold-based method proposed by (Descals et al., 2020). The method consists of five steps. 1) Invalid observations were masked using the Scene Classification Layer, which is included in the Sentinel-2 Level-2A product. We also masked observations that were flagged as clouds with a probability >65% in the cloud mask generated with the sentinel2-cloud-detector library of the European Space Agency (ESA). 2) We generated a time series of NDVI, i.e. the normalised difference between the near-infrared and red bands (bands 8 and 4 in Sentinel-2, respectively). 3) The NDVI time series at a revisiting time of 5 d were linearly interpolated at daily steps. 4) EoS

was estimated as the day when NDVI decreased in late summer and autumn, exceeding a dynamic threshold. The threshold was defined as 20% of the annual amplitude plus the annual minimum NDVI value. The 20% threshold represents low NDVI values that ensure the estimation of the leaf-shedding date. Values lower than the 20% threshold may lead to inconsistent estimates of EoS because of the noise in the NDVI time series during dormancy. 5) EoS was rejected if the gap between the closest valid observations before and after the EoS date was >15 d. We thereby ensured that only high-confidence estimates from dense Sentinel-2 time series were included in the analysis.

5.2.3. Detection of early leaf shedding

We identified pixels where early leaf shedding could potentially occur when the EoS estimated using Sentinel-2 data was earlier than 1 September. We used this definition because leaf shedding commonly occurs after this date in European deciduous forests. The same criterion was used in a previous study that detected early leaf shedding in central Europe (Brun et al., 2020). We masked potential false positives of early leaf shedding in these pixels. First, we removed pixels with vegetation that was affected by wildfires. Areas affected by summer wildfires also presented a decline in the NDVI time series, similar to the decrease in areas affected by early leaf shedding. We masked early-EoS pixels that overlapped with the FIRMS active-fire product (Giglio et al., 2016). Second, we masked all types of vegetation that were not deciduous broadleaf forests (DBFs). The DBF mask was obtained from the Copernicus Global Land Cover Layers (CGLS-LC100 Collection 3) at a resolution of 100 m for 2019 (Buchhorn et al., 2020), which included the class 'deciduous broadleaf closed forest (tree canopy >70%)'. We also used the ESA WorldCover 10-m v100, with the class 'Trees', to mask the remaining non-forest pixels. Finally, we applied a filter to remove isolated pixels that had early EoSs.

The occurrence of early leaf shedding was validated using a random sampling of pixels. We randomly sampled 150 points where Sentinel-2 potentially detected early leaf shedding during 2017-2021 (EoS date before 1 September). The points were sampled such that the minimum distance between two points was >10 km. The validation was done by visual interpretation of Sentinel-2 images. We visualised RGB true-colour Sentinel-2 images (bands 4, 3, and 2, corresponding to the red, green, and blue bands) and labelled the points as early leaf shedding if the RGB image indicated that the forest canopy was brownish between 20 August and 10 September. If a point was a false positive, we also labelled the cause of the false positive.

The Sentinel-2 EoS was compared to the EoS provided in the MCD12Q2 V6 Land Cover Dynamics product (Friedl et al., 2019). The aim of this comparison was to determine whether low resolution (MODIS product at 500-m resolution) could detect early leaf shedding. We used the MidgreenDown and the Dormancy metrics from the MCD12Q2 product, which represent the dates when the Enhanced Vegetation Index 2 (EVI2) last crossed 50 and 15% of the annual amplitude, respectively. The data were compared only at the points where early leaf shedding was confirmed by visual inspection for 2017-2019, which are the years that Sentinel-2 and MCD12Q2 product overlap. We also investigated whether the size of the forest affected by early leaf shedding influenced the MODIS EoS metrics. We hypothesised that the larger the affected area, the better MODIS EoS could identify early leaf shedding, resulting in a decrease in the difference between the Sentinel-2 EoS and MODIS EoS metrics. We inspected the difference in EoS between the two sensors as a function of the size of the affected area. The affected forest size was calculated as the number of 10-m pixels with early leaf shedding detected by Sentinel-2 that fell within a 500-m MODIS pixel.

5.2.4. Legacy effects of early leaf shedding

We investigated whether early leaf shedding affected the productivity of the forests in the years after the occurrence by identifying the trends in NDVI in the pixels where early leaf shedding was confirmed by visual interpretation. We extracted the mean NDVI for early summer (June and July), when NDVI in DBFs peaks during the growing season. We then analysed the trends in NDVI in the pixels where early leaf shedding was detected. We inspected whether the NDVI significantly decreased the year after the early leaf shedding. We used a two-sample *t*-test to determine if the mean summer NDVI differed significantly between years with a confidence interval >95%.

5.2.5. Assessment of drought impacts on leaf shedding

The relationship between early EoS and extreme climate was investigated using two climatic time series: air temperature and the aridity index. We used the air temperature 2 m above the surface provided by the ERA5-Land hourly data set (Muñoz-Sabater et al., 2021). The aridity index ($P\ PET^{-1}$) is the ratio of rainfall to potential evapotranspiration for a given time interval (Sherwood & Fu, 2014). The rainfall was obtained directly from the ERA5-Land hourly data set.

Potential evapotranspiration was calculated hourly using the method described by (Singer et al., 2021), which was based on the FAO Penman-Monteith equation.

We examined the relationships between temperature, aridity index, and the EoS date to identify two types of processes.

1) We investigated whether deciduous trees advanced the leaf shedding during short-term anomalous high temperatures and arid periods (i.e. low values of the aridity index) by first extracting the mean temperature and mean aridity before the date of early leaf shedding using various time intervals (15, 30, 60, 90, and 120 d). We then extracted the aridity for the same day of the year (DoY) but for other years during 2001-2021. Finally, we tested whether the temperature and aridity were significantly different in the year with an early EoS compared to 2001-2021. The data for temperature and aridity from 2001 to 2021 were normalised so that the mean was 0 and the standard deviation was 1. Assuming a Gaussian distribution, anomalous temperatures and aridities will have a normalised value differing substantially from 0.

2) We investigated whether deciduous trees delayed the shedding of leaves when the temperature and aridity were generally anomalously high throughout the growing season. For this case, we estimated the mean temperature and mean aridity for summer (June, July, August, and September) for 2001-2019, the period for which the MCD12Q2 V6 data were available. We then calculated the sensitivity of the MCD12Q2 EoS MidGreendown to mean summer temperature and mean summer aridity. The sensitivities to temperature and aridity were calculated as the linear slopes between EoS and the mean summer temperature and mean summer aridity, respectively, using an ordinary least-squares regression over all DBF pixels in the study area.

5.3. Results

5.3.1. Detection of early leaf shedding

We identified 90 points with early leaf shedding from 150 randomly selected locations in pixels where Sentinel-2 EoS was earlier than 1 September (Figure 5.1a; Supplementary Tables 5.1 and 5.2). Early leaf shedding was found throughout the study area and study period (2017-2021). Occurrence was highest in 2018 and 2019 (Figure 5.1b) and lowest in 2017. The insufficient

number of valid observations during 2017 was the primary cause of false positives in our method; a total of 17 points were incorrectly detected as early leaf shedding for 2017, mostly in central Europe (Figure 5.1a). Other causes of false positives were the inaccuracy of the DBF mask (14 points); many locations were in fact shrubland, which may be more prone to early leaf shedding than DBF forests. The early decrease in NDVI was caused by fires at 11 points, even though the FIRMS active-fire product was used to mask them, and forests were clear-cut at 7 points.

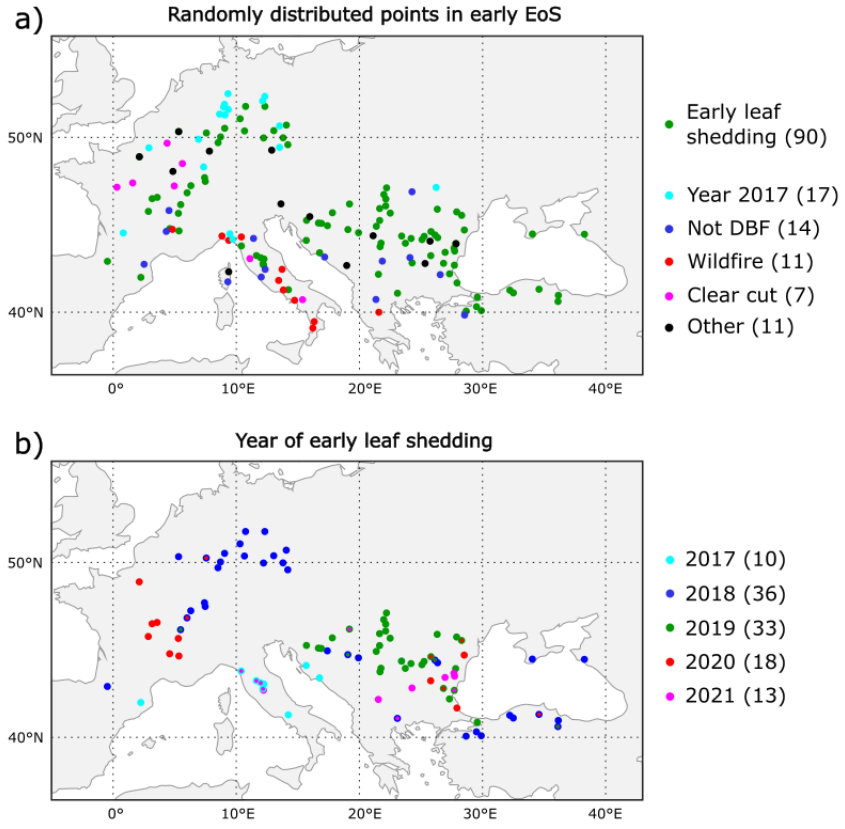


Figure 5.1. Spatial distribution of the validation points. a) One hundred and fifty points were randomly distributed in pixels where the end of the growing season (EoS), obtained with Sentinel-2, was earlier than 1 September. The colours of the points depict the labels assigned during the visual interpretation of Sentinel-2 images. Early leaf shedding was identified in 90 points. False positives appeared in the other points due to the lack of available data for 2017, inaccuracies in the mask for deciduous broadleaf forests (DBFs), the presence of wildfires and clear-cuts, and other reasons. b) The year of detection for the 90 points where early leaf shedding was identified with the visual interpretation of Sentinel-2 images. Early leaf shedding occurred in 110 site-years; some sites experienced early leaf shedding more than once during 2017-2021. The numbers in parentheses represent the number of occurrences.

The true-colour images at a resolution of 10 m identified the times before early leaf shedding (Figure 5.2a), after early leaf shedding (Figure 5.2b) and after all trees had entered dormancy (Figure 5.2c) at the 90 points where early leaf shedding was confirmed. The brown colour in the

RGB images depicts the areas where trees had coloured leaves or where their leaves were shed. Supplementary Figure 5.1 depicts forests that had been cleared or affected by fires, which are disturbances that were also identified visually in Sentinel-2 and discriminated against early leaf shedding.

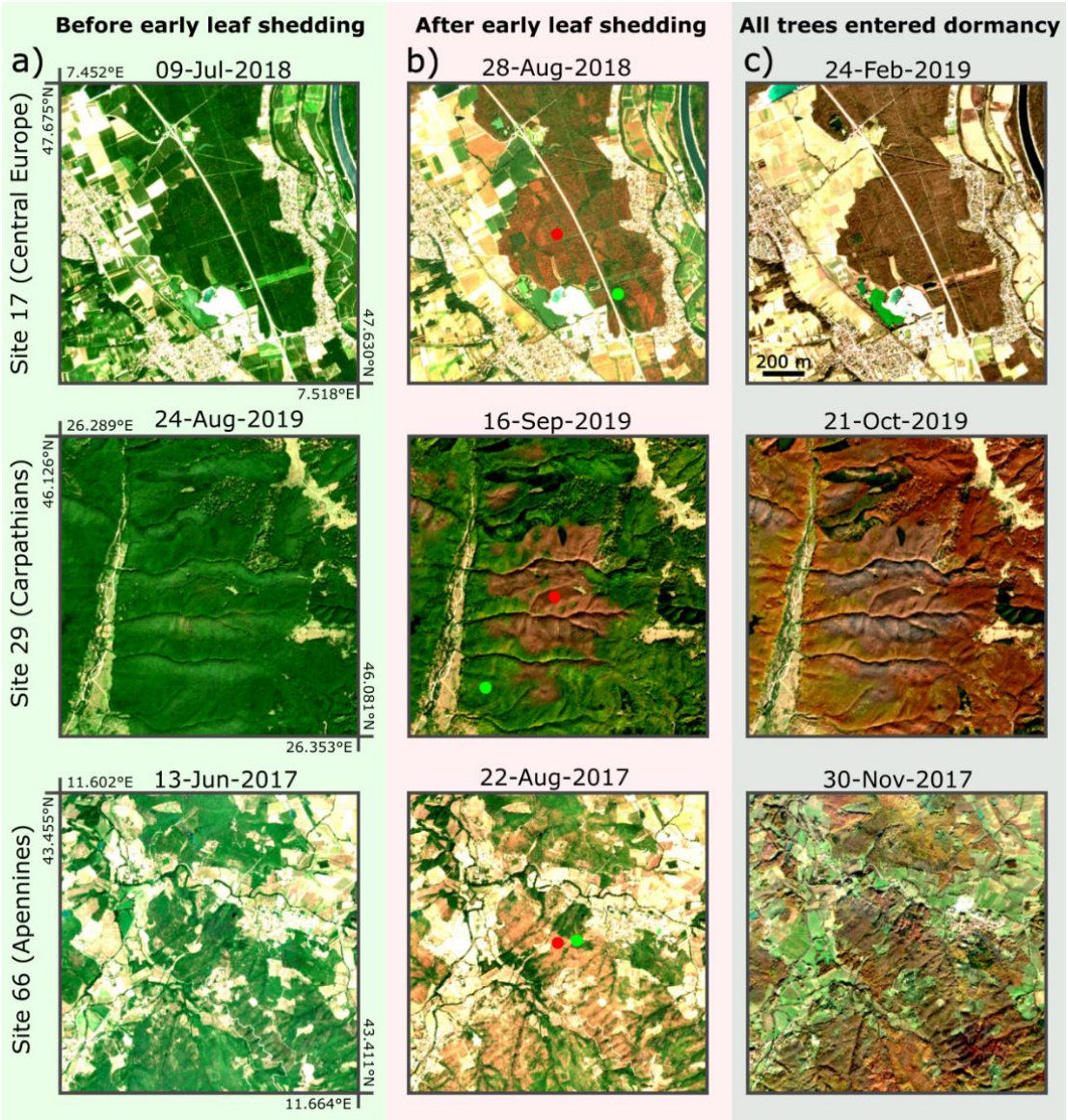


Figure 5.2. Sentinel-2 images for three sites and three times during the growing season: a) before early leaf shedding, b) after early leaf shedding, and c) after all trees have entered dormancy. The images show a true-colour composition; the channels correspond to Sentinel-2 bands 4 (red), 3 (green), and 2 (blue). Forests that have shed leaves prematurely in the growing season are depicted by brown in the middle images. The red and green points correspond to pixels whose time series are displayed in Figure 5.3.

The Sentinel-2 time series indicated that NDVI for a year with early leaf shedding decreased before 1 September (DoY = 244), substantially earlier than other years during 2017-2021 (Figure

5.3). The NDVI time series at two pixel locations, with early and normal leaf shedding, indicated that the vegetation at both locations had a similar seasonality, with similar EoSs. EoS, however, diverged for the year in which early leaf shedding was detected. The pixels with early leaf shedding had a sharp decrease in NDVI, with subsequent low NDVI values that were similar to those when the deciduous trees became dormant.

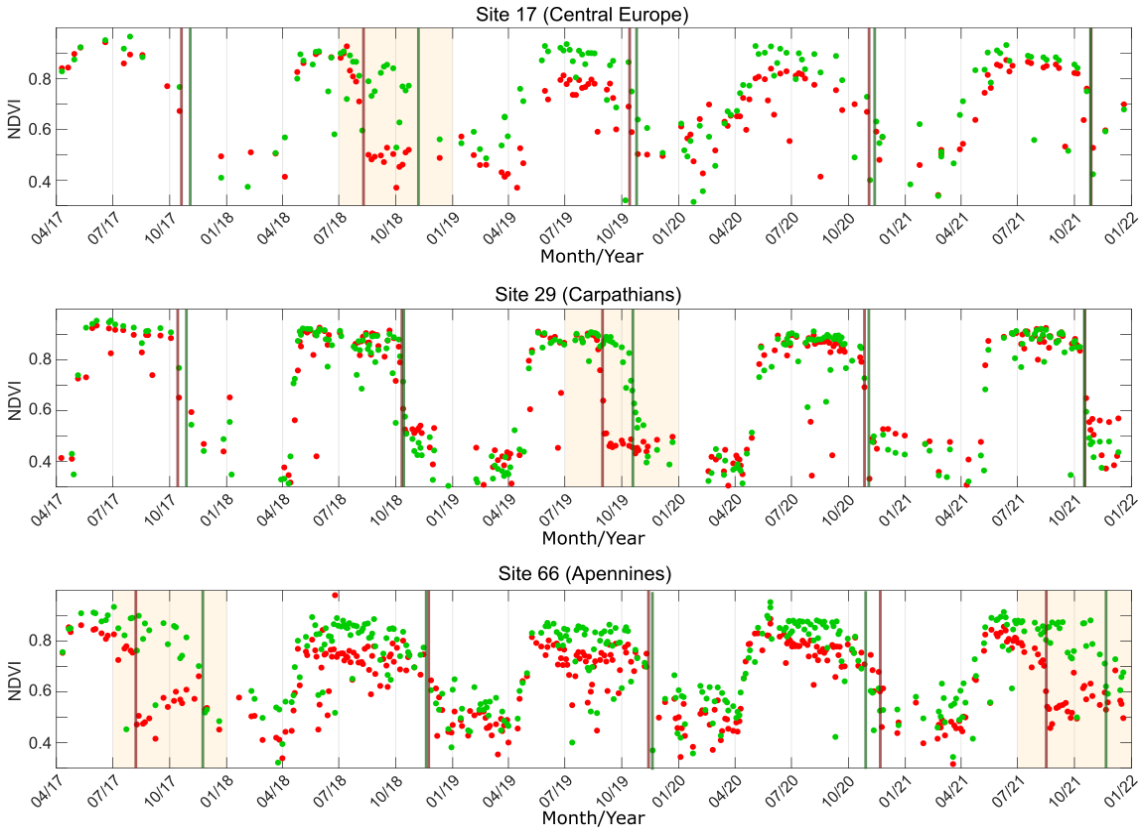


Figure 5.3. Time series of the normalised difference vegetation index (NDVI) for a canopy with early leaf shedding (red points) and a canopy that shed its leaves near the median of 2017-2021 (green points) at three sites (Figure 5.2). The time of the year when early leaf shedding was detected is indicated by the beige shading. The site locations are shown in Figure 5.2.

The magnitude of the early EoS was underestimated in satellite data with lower spatial resolution. MODIS underestimated the advance in the timing of early leaf shedding (Figure 5.4a). Sentinel-2 EoS was estimated using the 20% threshold, and the MODIS MidGreenDown and Dormancy metrics were calculated using a similar LSP method but using 50 and 15% thresholds, respectively. The Sentinel-2 EoS date should thus be between the MODIS MidGreenDown and Dormancy dates. Sentinel-2, however, generally detected early leaf shedding 9 [0, 24.25] d (median and interquartile range) before MODIS MidGreenDown and 71 [63.75, 85.25] d before MODIS Dormancy (Figure 5.4b). The differences in EoS between Sentinel-2 and MODIS could be

attributed to the use of a different satellite sensor or vegetation index (NDVI versus EVI2), but the same comparison for the year preceding the early leaf shedding indicated that Sentinel-2 EoS was between the MidGreenDown and dormancy dates (the difference was -23 [-47, -12.75] d and 12 [3, 27.5] d, respectively; Supplementary Figure 5.2), thus indicating that the difference in variable (NDVI versus EVI2) did not have an impact on the difference between Sentinel-2 and MODIS EoS.

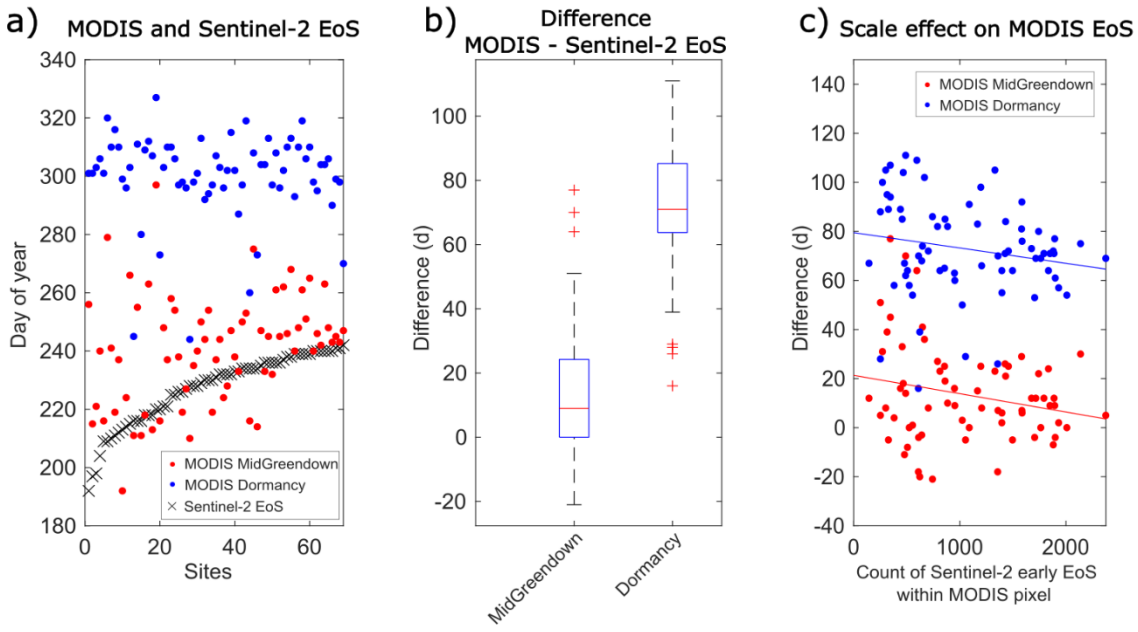


Figure 5.4. Comparison between the end of the growing season (EoS) obtained from MODIS and Sentinel-2 for 90 sites where early leaf shedding was detected. a) Representation of the Sentinel-2 EoS and the MODIS MidGreendown and Dormancy metrics per site. Sites are numbered with the Sentinel-2 EoS in ascending order. We show only 69 sites, which are the sites where the MODIS MidGreendown and Dormancy metrics were available. b) Boxplots showing the difference in distribution between the Sentinel-2 EoS and the MODIS EoS metrics. Red line depicts the median, blue box represents the interquartile range, and red crosses outside the whiskers are outliers (>2 and <-2 standard deviations). c) Effect of scale on the estimates of early leaf shedding in MODIS. The x-axis shows the number of pixels in which Sentinel-2 detected early leaf shedding within a 500-m MODIS pixel. The y-axis shows the difference between the Sentinel-2 EoS and the MODIS EoS metrics at each site.

The difference between the Sentinel-2 EoS metrics and the MODIS EoS metrics decreased as the size of the forest affected by early leaf shedding increased (Figure 5.4c), indicating that the EoS for both sensors became more similar as the effect grew in size. The slope, however, was only significant for the MODIS MidGreendown metric at a confidence interval of 95%. On average, the difference between Sentinel-2 EoS and the MODIS MidGreendown was close to zero when more than 2000 Sentinel-2 pixels presented early leaf shedding within a 500-meter MODIS pixel.

5.3.2. Legacy effects of early leaf shedding

The overall mean summer NDVI decreased in the year after the early leaf shedding (Figure 5.5). The two-sample *t*-test indicated that the decrease was significant (95% confidence interval) for early leaf shedding that occurred in 2017, 2018, and 2019. The decrease was not significant for 2020, but the difference was significant compared with the summer NDVI for 2017. These results suggest that summer NDVI decreased not only after the year of early leaf shedding but also in the years preceding it. This negative trend in NDVI was clear when early leaf shedding occurred in 2020 and 2021. The slope of this trend differed significantly from 0 for a confidence interval of 95%. The validity of this analysis, however, was hampered by the short time series; this analysis only included the five years of the Sentinel-2 data. Moreover, the NDVI did not decrease the year after early leaf shedding in some of the sites (see site 29 in Figure 5.3).

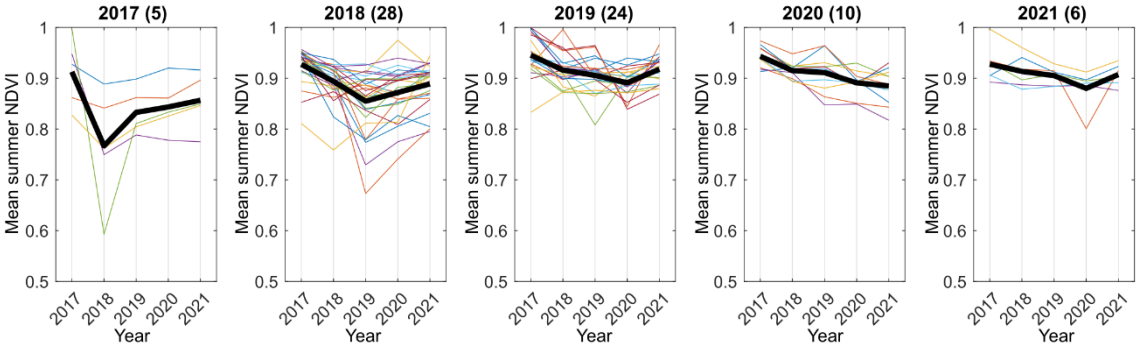


Figure 5.5. Mean summer normalised difference vegetation index (NDVI) at sites where early leaf shedding was observed during 2017-2021. The sites were grouped depending on the year of the occurrence. The titles show the year of disturbance and the number of sites in parentheses. Only sites with one year of early leaf shedding during 2017-2021 were considered. The thin lines represent the summer NDVI per site, and the thick line represents the mean summer NDVI across sites.

5.3.3. Assessment of drought impacts to leaf shedding

Early leaf shedding generally occurred at anomalously high temperatures and aridity (Figure 5.6). The normalised temperature and aridity 30 d before early leaf shedding were 0.41 [-0.01, 1.10] and -0.61 [-1.03, -0.23], respectively. The 30-d temperature was significantly higher and the 30-day aridity was significantly lower than the average for 2001-2021 for the same DoYs (p<0.05). When considering other aggregation periods (15, 60, 90, and 120 d before early leaf shedding), the mean temperature and mean aridity were also significantly different than the long-term averages (Supplementary Figure 5.3).

The analysis above considered aggregated mean temperature and aridity for all the sites where early leaf shedding was detected. The period under which arid conditions affect early leaf shedding, however, may differ locally. The area of early leaf shedding increased for specific years when conditions were driest (Figure 5.7), corroborating the finding in Figure 5.6, but the length of time under which the conditions were arid depended on the region. For example, a short period of arid conditions led to leaf shedding in the Carpathians (Figure 5.7b), while forests shed their leaves after an unusually long arid period in the Apennines and central Europe (Figure 5.7a and c).

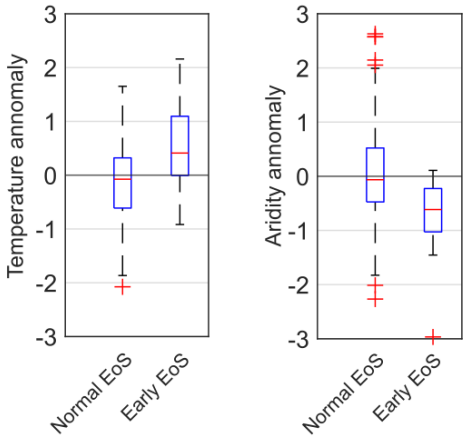


Figure 5.6. Anomalies of temperature and aridity (P PET^{-1}) at the 90 sites where early leaf shedding was identified. Temperature and aridity were calculated for a time lag of 30 d before early leaf shedding. The boxplots depict the temperature and aridity distributions normalised for 2001-2021; mean = 0 and standard deviation = 1. 'Early EoS' represents the normalised temperature and aridity for the year of early leaf shedding detected during 2017-2021, and 'normal EoS' represents the normalised temperature and aridity for the other years during 2017-2021. Red line depicts the median, blue box represents the interquartile range, and red crosses outside the whiskers are outliers (>2 and <-2 standard deviations).

The sensitivity of MODIS EoS was positive (0.86 d $°C^{-1}$) to summer mean temperature and negative (-6.12 d mm mm^{-1}) to summer mean aridity during 2001-2019, indicating that EoS was delayed with hotter and drier summers. The percentage of pixels with significant sensitivity was relatively low; 12.98 and 8.99% of the pixels covering DBFs had a significant sensitivity of EoS to temperature and aridity, respectively. The positive sensitivity to temperature and negative sensitivity to aridity were generally found throughout the study area except the northern coast of the Anatolian peninsula (Figure 5.8).

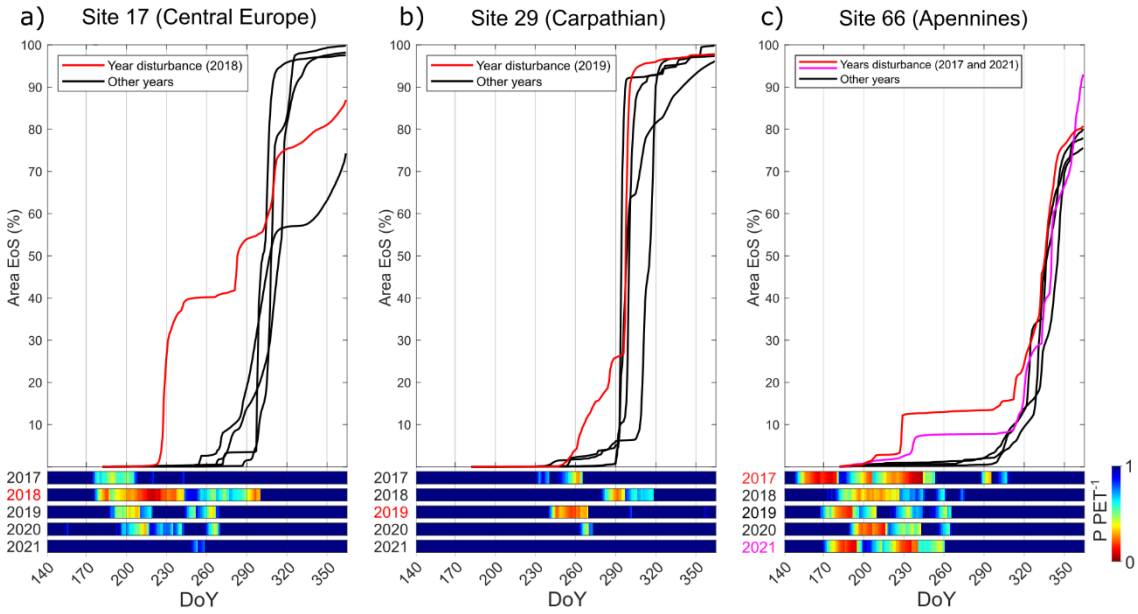


Figure 5.7. Accumulated area that reached the end of the growing season (EoS) at three sites. The accumulated EoS was calculated for an area of 10×10 km^2 at each site. Red lines correspond to years in which early EoS was detected. Black lines are other years during 2017-2021. The lower panels show the 30-d aridity (P PET^{-1}) for 2017-2021.

5.4. Discussion

Sentinel-2 was able to identify early leaf shedding at the continental scale during 2017-2021. Early leaf shedding occurred not only during extreme short-term heatwaves, such as the 2018 European heatwave, but in every year during 2017-2021 when meteorological conditions were less severe. The estimation of early EoS was possible due to the short revisiting times of the Sentinel-2A and -2B satellites. The MODIS EoS product greatly underestimated early leaf shedding. Low spatial resolution tends to overestimate the length of the growing season (Hmimina et al., 2013), which hampers the retrospective study of early leaf shedding before the era of Sentinel-2 or Landsat-7 and -8. The bias between Sentinel-2 and MODIS differed greatly among sites. Our results indicated that these differences depended on the varying sizes of the areas that experienced early leaf shedding. MODIS may thus detect early leaf shedding more accurately when it occurs in large areas, reducing the bias associated with the Sentinel-2 estimate. Early leaf shedding in European temperate deciduous forests, however, occurs mainly at a small scale and can be observed only during the last weeks of the growing season in late summer and early autumn. In contrast, tree mortality in coniferous trees can be easily detected with decametric-resolution satellite images in the following growing season (Meddens et al., 2013), which may account for why early leaf shedding in DBFs has been insufficiently studied in

remote-sensing studies, with only one study using Sentinel-2 data to cover this event (Brun et al., 2020).

Early leaf shedding was linked to unusually high temperatures and aridity in the days preceding the event. Most occurrences were in central Europe in 2018 and the Carpathians in 2019. These occurrences coincided in time and space with European heatwaves. The data from some sites also suggested that early leaf shedding stopped when aridity decreased (Figure 5.7), presumably because rain resumed. Our results indicated a link between early leaf shedding and arid conditions in summer, but the direct causes that trigger leaf shedding remain unknown. Deciduous broadleaved trees may simply shed their leaves to prevent excessive evapotranspiration (Bréda et al., 2006), which may help prevent the stems from desiccating, reducing the likelihood of death. This strategy does not apply to coniferous trees, where hydraulic failure occurs during arid conditions, leading to tree death (Arend et al., 2021). Pest infestations are another plausible cause of early leaf shedding. A previous study found that deciduous trees shed their leaves as a mechanism to control pest populations (Karban, 2007).

Satellite data identified a divergent response of EoS to drought in deciduous forests. The Sentinel-2 data indicated that leaves were shed prematurely in forest patches during the growing season, but the MODIS data indicated that deciduous forests generally delayed leaf shedding during warm and arid seasons. Similar results were found with low spatial resolution (Bórnez et al., 2021) and in situ observations (Leuzinger et al., 2005). Delayed shedding in an arid summer may involve a mechanism distinct from early leaf shedding. Heatwaves may induce early leaf shedding locally due to their brief duration but high severity of heat stress. Continual arid conditions during the summer, however, can reduce vegetation productivity (Ciais et al., 2005), which in turn could account for the delay in leaf shedding (Zani et al., 2020). Such persisting arid conditions throughout the growing season are not severe enough to cause early leaf shedding, which also suggests that early leaf shedding occurs when specific thresholds for heat and water stress are exceeded, but additional evidence is required to support this hypothesis.

Our results indicated that NDVI decreased the year after the early leaf shedding, indicating a legacy effect on carbon uptake in following years. A previous study found that carbon sequestration decreased after the 2003 European heatwave (Ciais et al., 2005). The effect of early leaf shedding on the annual carbon uptake nevertheless remains uncertain. Annual carbon

uptake may not be affected, because carbon uptake in spring increases due to early SoS and high spring temperatures (Bastos et al., 2020). Our findings indicate that future increases in droughts will negatively influence carbon uptake in temperate forests. More extreme droughts will lead to forests shedding their leaves prematurely, as we have demonstrated. These forests will cease to take up carbon very early in the growing season.

The detection of early leaf shedding was possible in some area, but we also found shortcomings in our method that hampered the estimation of EoS using the Sentinel-2 data. One shortcoming was the low availability of data in cloud-prone regions. The frequency of observations could be increased by including data from Landsat satellites, which would improve the confidence of EoS estimates, particularly for 2017 when only images from Sentinel-2A were available and the revisit time was 10 days. Our method of detection could be improved with more accurate forest masks and optimised algorithms for detecting clouds, as (Brun et al., 2020) have done for Central Europe. These improvements would reduce false positives, allowing for confident estimates of areas of early leaf shedding following the practices used by (Olofsson et al., 2014). Lastly, we found false positives associated with clear-cutting. These false positives are more difficult to detect automatically, and future algorithms for monitoring early leaf shedding should take this difficulty into account, particularly for central Europe where this harvesting practice is common.

Our study demonstrates the feasibility of using decametric satellite data to monitor forest responses to drought at the canopy level. Sentinel-2 acquisitions are freely delivered in near real-time with a resolution of 10 m, enabling the rapid detection and localisation of early leaf shedding. Our method may thus facilitate in situ observations by helping define the area to visit when an event occurs. These local observations would contribute to a better understanding of the factors that contribute to early leaf shedding. The causes of early leaf shedding are difficult to deduce from satellite data alone, and field observations are required to confirm the underlying mechanisms. Future research may also examine the decrease in NDVI in the years preceding early leaf shedding, which could be attributed to the effects of global warming on forest health in Europe.

5.5. References

Arend, M., Link, R. M., Patthey, R., Hoch, G., Schuldt, B., & Kahmen, A. (2021). Rapid hydraulic collapse as cause of drought-induced mortality in conifers. *Proceedings of the National Academy of Sciences*, *118*(16).

Bastos, A., Ciais, P., Friedlingstein, P., Sitch, S., Pongratz, J., Fan, L., Wigneron, J.-P., Weber, U., Reichstein, M., Fu, Z., & others. (2020). Direct and seasonal legacy effects of the 2018 heat wave and drought on European ecosystem productivity. *Science Advances*, *6*(24), eaba2724.

Bolton, D. K., Gray, J. M., Melaas, E. K., Moon, M., Eklundh, L., & Friedl, M. A. (2020). Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. *Remote Sensing of Environment*, *240*, 111685.

Bornez, K., Descals, A., Verger, A., & Peñuelas, J. (2020). Land surface phenology from VEGETATION and PROBA-V data. Assessment over deciduous forests. *International Journal of Applied Earth Observation and Geoinformation*, *84*, 101974.

Bórnez, K., Verger, A., Descals, A., & Peñuelas, J. (2021). Monitoring the Responses of Deciduous Forest Phenology to 2000–2018 Climatic Anomalies in the Northern Hemisphere. *Remote Sensing*, *13*(14), 2806.

Bréda, N., Huc, R., Granier, A., & Dreyer, E. (2006). Temperate forest trees and stands under severe drought: A review of ecophysiological responses, adaptation processes and long-term consequences. *Annals of Forest Science*, *63*(6), 625–644.

Brun, P., Psomas, A., Ginzler, C., Thuiller, W., Zappa, M., & Zimmermann, N. E. (2020). Large-scale early-wilting response of Central European forests to the 2018 extreme drought. *Global Change Biology*, *26*(12), 7021–7035.

Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L., & Smets, B. (2020). Copernicus global land cover layers—Collection 2. *Remote Sensing*, *12*(6), 1044.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogée, J., Allard, V., Aubinet, M., Buchmann, N., Bernhofer, C., Carrara, A., & others. (2005). Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*, *437*(7058), 529–533.

Descals, A., Verger, A., Yin, G., & Peñuelas, J. (2020). Improved estimates of arctic land surface phenology using Sentinel-2 time series. *Remote Sensing*, *12*(22), 3738.

Friedl, M., Gray, J., & Sulla-Menashe, D. (2019). MCD12Q2 MODIS/Terra+ Aqua Land Cover Dynamics Yearly L3 Global 500m SIN Grid V006. *NASA EOSDIS Land Processes DAAC*.

Gallinat, A. S., Primack, R. B., & Wagner, D. L. (2015). Autumn, the neglected season in climate change research. *Trends in Ecology & Evolution*, *30*(3), 169–176.

Giglio, L., Schroeder, W., & Justice, C. O. (2016). The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sensing of Environment*, *178*, 31–41.

Hartmann, H., Bastos, A., Das, A. J., Esquivel-Muelbert, A., Hammond, W. M., Martínez-Vilalta, J., McDowell, N. G., Powers, J. S., Pugh, T. A., Ruthrof, K. X., & others. (2022). Climate Change Risks to Global Forest Health: Emergence of Unexpected Events of Elevated Tree Mortality Worldwide. *Annual Review of Plant Biology*, *73*.

Hmimina, G., Dufrêne, E., Pontailler, J.-Y., Delpierre, N., Aubinet, M., Caquet, B., De Grandcourt, A., Burban, B., Flechard, C., Granier, A., & others. (2013). Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: An investigation using ground-based NDVI measurements. *Remote Sensing of Environment*, *132*, 145–158.

Julien, Y., & Sobrino, J. (2009). Global land surface phenology trends from GIMMS database. *International Journal of Remote Sensing*, *30*(13), 3495–3513.

Karban, R. (2007). Deciduous leaf drop reduces insect herbivory. *Oecologia*, *153*(1), 81–88.

Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., Munger, J. W., O'Keefe, J., Schmid, H. P., Wing, I. S., & others. (2014). Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nature Climate Change*, *4*(7), 598–604.

Kozlowski, T. (1976). Water supply and leaf shedding. *Soil Water Measurements, Plant Responses, and Breeding for Drought Resistance*, *4*, 191–231.

Leuzinger, S., Zotz, G., Asshoff, R., & Körner, C. (2005). Responses of deciduous forest trees to severe drought in Central Europe. *Tree Physiology*, *25*(6), 641–650.

Mariën, B., Dox, I., De Boeck, H. J., Willems, P., Leys, S., Papadimitriou, D., & Campioli, M. (2021). Does drought advance the onset of autumn leaf senescence in temperate deciduous forest trees? *Biogeosciences*, *18*(11), 3309–3330.

Meddens, A. J., Hicke, J. A., Vierling, L. A., & Hudak, A. T. (2013). Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date Landsat imagery. *Remote Sensing of Environment*, *132*, 49–58.

Menzel, A., Sparks, T. H., Estrella, N., & Roy, D. (2006). Altered geographic and temporal variability in phenology in response to climate change. *Global Ecology and Biogeography*, *15*(5), 498–504.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., & others. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data Discussions*, 1–50.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, *148*, 42–57.

Senf, C., Buras, A., Zang, C. S., Rammig, A., & Seidl, R. (2020). Excess forest mortality is consistently linked to drought across Europe. *Nature Communications*, *11*(1), 1–8.

Sherwood, S., & Fu, Q. (2014). A drier future? *Science*, *343*(6172), 737–739.

Singer, M. B., Asfaw, D. T., Rosolem, R., Cuthbert, M. O., Miralles, D. G., MacLeod, D., Quichimbo, E. A., & Michaelides, K. (2021). Hourly potential evapotranspiration at 0.1° resolution for the global land surface from 1981-present. *Scientific Data*, *8*(1), 1–13.

Tian, F., Cai, Z., Jin, H., Hufkens, K., Scheifinger, H., Tagesson, T., Smets, B., Van Hoolst, R., Bonte, K., Ivits, E., & others. (2021). Calibrating vegetation phenology from Sentinel-2 using eddy covariance, PhenoCam, and PEP725 networks across Europe. *Remote Sensing of Environment*, *260*, 112456.

Way, D. A., & Montgomery, R. A. (2015). Photoperiod constraints on tree phenology, performance and migration in a warming world. *Plant, Cell & Environment*, *38*(9), 1725–1736.

Xie, Y., Wang, X., & Silander, J. A. (2015). Deciduous forest responses to temperature, precipitation, and drought imply complex climate change impacts. *Proceedings of the National Academy of Sciences*, *112*(44), 13585–13590.

Zani, D., Crowther, T. W., Mo, L., Renner, S. S., & Zohner, C. M. (2020). Increased growing-season productivity drives earlier autumn leaf senescence in temperate trees. *Science*, *370*(6520), 1066–1071.

Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H., Hodges, J. C., Gao, F., Reed, B. C., & Huete, A. (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, *84*(3), 471–475.

Chapter 6

Local interpretation of machine learning models in remote sensing with SHAP: the case of global climate constraints on carbon uptake phenology

Adrià Descals, Aleixandre Verger, and Josep Peñuelas

Submitted to Remote Sensing of Environment (2022)

Abstract

Data-driven models using machine learning have been widely used in remote sensing applications such as the retrieval of biophysical variables and land cover classification. However, these models behave as a 'black box', meaning that the relationships between the input and predicted variables are hard to interpret. Recent regression models that downscale sun-induced fluorescence (SIF) with MODIS and weather variables are an example. The impact of weather variables on the predicted SIF in these models is unknown. The explanation of such weather-SIF relationships would aid in the understanding of climate-related constraints on carbon uptake phenology. Here, we used SHapley Additive exPlanations (SHAP) –a novel technique based on game theory– for explaining the contribution of input variables to the individual predictions in a machine learning model. We explored the capabilities of this technique with a weather-SIF model. The regression model predicted TROPOSIF measurements from ERA5-Land air temperature, shortwave radiation, and vapor-pressure-deficit (VPD) data. The SHAP values of the model were estimated at the start and end of the growing season for the entire globe. These values depicted the global constraints of the three climate variables on the photosynthetically active season and confirmed existing limitations regarding terrestrial carbon uptake with unprecedented spatial detail. Radiation was the limiting factor in tropical rainforest and VPD constrained the start and end of the growing season in tropical dryland ecosystems. In extra-tropical regions, temperature was the main limiting factor during the start of the growing season, but both temperature and radiation constrained photosynthesis at the end of the growing season. This technique may help future remote sensing studies that require the use of non-interpretable machine-learning regression models and explain how input variables contribute to the model prediction at the spatial and temporal scales.

6.1. Introduction

The field of vegetation phenology has gained attention recently, with the number of publications on phenology quintupling in the last two decades (Fu et al., 2020). This growing interest can be partially attributed to the link between changes in vegetation phenology and rising temperatures (Peñuelas et al., 2009), and the questions regarding how vegetation will respond to future climatic warming (Piao et al., 2019). Despite the growing interest in this field, phenology modelling has not received as much attention (Fu et al., 2020). Phenology modelling is important for understanding the factors influencing the timing of the onset and end of growing seasons and predict how vegetation will respond to future warming, which is critical for determining whether the growing season will continue to lengthen in the future, thus leading to further vegetation carbon uptake (Keenan et al., 2014).

The transition between the dormant and growing season and the climate factors determining it have been explained globally by models employing climate thresholds. (Jolly et al., 2005) calculated the growing season index (GSI), calculated with cut-off functions on three weather variables: temperature, vapor-pressure-deficit (VPD), and day length. These cut-off functions represent thresholds that were defined by expert knowledge and are constant for the entire globe. The GSI shows seasonal changes throughout the year and aims to replicate a spectral index (e.g., normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI)) or a biophysical variable (e.g., leaf area index (LAI)).

Other studies used regression models to fit climate datasets to vegetation indices or biophysical variables. Both standard machine learning regression –such as random forests (Li and Xiao, 2019)–, and deep learning (Ahmad et al., 2020) have been used given their ability to fit non-linear and non-parametric relationships between dependent and independent variables. This methodology predicts vegetation indices or biophysical variables, and the climate thresholds are, thus, defined empirically and more accurately than the cut-off functions in (Jolly et al., 2005). However, an important flaw in machine learning models is the lack of interpretability. Contrarily to (Jolly et al., 2005), the impact of the weather variables on the predicted outcome remains challenging in machine learning models. Recent regression models using vegetation indices and weather variables to downscale sun-induced fluorescence (SIF) are an example. For instance, the GOSIF product (Li and Xiao, 2019) uses a machine learning regression model to fit SIF with weather variables and EVI.

Machine learning models, however, do not allow for a clear understanding of the relationships between input and output variables – this refers to the impact of individual weather observations on the predicted SIF. A state-of-the-art local interpretation method for model explainability is SHapley Additive exPlanations (SHAP) (Lundberg and Lee, 2017), an approach that explains the contribution of input variables on the model output. SHAP has been used for understanding the risk of hypoxemia during anaesthesia (Lundberg et al., 2018), interpret the features that make an online product review helpful (Meng et al., 2020), understanding the pollutant removal mechanisms in wastewater treatment plants (Wang et al., 2022), or analysing large-scale biobank data for potential gene–gene and gene–environment interactions (Johnsen et al., 2021). This technique is, however, novel in remote sensing studies and might help to understand spatial and temporal impacts of geospatial data in machine learning models.

The aim of this study was to demonstrate the capability of a local interpretation technique to explain the correlation between geospatial gridded data and model predictions in a machine learning model. We used the case of weather-SIF models (Li and Xiao, 2019) to determine the global constraints of weather variables on vegetation activity. To achieve the objective of the study, we modelled TROPOSIF measurements using ERA5-Land temperature, shortwave radiation, and VPD. Then, we used SHAP to describe the effect of weather variables on SIF at the start and end of the growing season and, thus, determine the climate constraints on vegetation phenology. Finally, we discussed and validated the model interpretation with SHAP compared with the current understanding on carbon uptake dynamics. Finally, we discuss the consistency of the results obtained with SHAP considering the current understanding on carbon uptake dynamics.

6.2. Data

6.2.1. TROPOSIF global sun-induced fluorescence dataset

We used the TROPOSIF L2B product (Guanter et al., 2021), which provides non-gridded SIF measurements derived from observations in the 743–758 nm and 665–785 nm part of the spectrum. The observations were made by the TROPOMI sensor onboard Sentinel-5. The methodology that generates SIF uses a retrieval method that fits the top-of-atmosphere radiances with SIF training sets (Guanter et al., 2015). We used the data for all the product time coverage, which spans from May 2018 to April 2021. We used the SIF_745_corr, which represents corrected SIF in the 743–758 nm window. SIF observations that presented a cloud

cover greater than 50% were rejected. The L2B product already masks observations with cloud cover greater than 80%, a view zenith angle greater than 60°, and a solar zenith angle greater than 70°.

6.2.2. ERA5-Land hourly data

We used gridded climatic data from the ECMWF Re-Analysis data version 5 (ERA5-Land) hourly dataset (Muñoz-Sabater et al., 2021). The ERA5-Land is a reanalysis dataset that covers a period from 1950 to present. The data were produced by a combination of modelled data with observations collected across the globe and improves upon the ERA-5 since it has higher spatial resolution (about 9 km) at the same temporal resolution (1 hour). We used the near-surface air temperature (2m temperature) and the surface solar radiation downwards –the solar shortwave radiation that reaches the surface of the Earth. We also estimated VPD –atmospheric demand for evapotranspiration– using the ERA5 near-surface air temperature and dew point temperature (2m dew point temperature) as described by Barkhordarian et al. (2019).

6.3. Methods

6.3.1. Extraction of training pairs (TROPOSIF - ERA5) in Belmanip2 sites

We collected pairs of TROPOSIF measurements and ERA5 observations as training data. These data were used to train a machine learning model that predicted SIF from temperature, shortwave radiation, and VPD. We extracted the TROPOSIF and ERA5 data from the Belmanip2 sites for the period going from May 2018 to April 2021. The Belmanip2 consists of a collection of 445 sites of homogeneous areas that include the most representative land covers of the world (Weiss et al., 2014). We excluded bare soil, cropland, and other non-natural or non-vegetated land covers, which resulted in 233 sites (see location map of the Belmanip points in Supplementary Figure 6.1) including the following land covers: evergreen needleleaf forests (ENF), deciduous needleleaf forests (DNF), deciduous broadleaf forests (DBF), mixed forests (MX), closed shrublands (CSH), open shrublands (OSH), woody savannah (WSA), savannah (SAV), and grasslands (GRA). The land cover types were determined for each Belmanip2 site with the 'LC_Type1' layer of the MCD12Q1v6 product. The ERA5 data were extracted at hourly temporal resolution, and then aggregated daily. The TROPOSIF dataset provides daily non-gridded SIF measurements. We, thus, extracted the daily SIF observations that were located the closest to a Belmanip2 site. SIF observations more than 5 km away from a Belmanip2 site were rejected.

A total of 140,969 pairs of data were generated from the 233 sites for the May 2018 - April 2021 period.

6.3.2. Weather – SIF model

The regression model that fitted ERA5 data (air temperature, shortwave radiation, and VPD) to TROPOSIF was the Gradient Boosting (Friedman, 2001). Gradient Boosting is an ensemble model that uses decision trees as weak learners, where decision trees are trained sequentially by correcting the errors of a previously trained decision tree. The performance of the decision trees is improved using a loss function. We used Gradient Boosting because it is a common machine learning model used by the research community, it can easily capture non-linear and non-parametric relationships, and has very fast training and deployment times (Bentéjac et al., 2021). We trained the Gradient Boosting model with 75% of the data and kept the remaining 25% for validation. The accuracy metrics that we reported are the mean error (ME: difference between predicted minus observed), root-mean-squared error (RMSE), and the coefficient of determination (R^2).

We performed hyperparameter tuning to find the optimal parameters of the Gradient Boosting regression model. The hyperparameter tuning consisted of a random search for different combinations of hyperparameter values. The range of hyperparameters is depicted in Supplementary Table 6.1. For each combination of parameters, the RMSE of the model was evaluated with a 4-fold partition (75% training and 25% validation). The accuracy of the Gradient Boosting model was tested using 100 different hyperparameter combinations, and the model with the lowest root-mean-squared error (RMSE) was selected.

Additionally, we did a 4-fold cross-site validation that consisted of training with 175 sites (75% of the sites) and validating with the remaining 58 sites (25% of the sites). We reported the averaged accuracy metrics over the four folds.

6.3.3. Local interpretation with SHAP

The local interpretation of the Gradient Boosting was done with SHapley Additive exPlanations (SHAP) (Lundberg and Lee, 2017). SHAP is a state-of-the-art technique for machine learning explainability; it aims to explain the correlations between input and output variables in any machine learning model, in both regression and classification algorithms. SHAP is based on the

Shapley values (Shapley, 1953) of game theory and is categorized as a local interpretation technique –it explains the contribution of the input variables to individual model predictions. SHAP values represent the marginal effect of the input variables on the prior expectation of the model output. A negative SHAP value for a given input variable implies that the input variable has a decreasing effect on the model output, and vice versa; a positive SHAP value means that the input variable increases the model output. The greater the absolute value of SHAP, the greater the impact of the input variable on the predicted value. The model prediction is the prior expectation of the model plus the summatory of the SHAP values of all input variables. The mathematical formulation for SHAP is described in (Lundberg and Lee, 2017). In this study, we used the SHAP package in Python and used the approximation method for tree-based machine learning models (Lundberg et al., 2020).

6.3.4. Land Surface Phenology metric estimation

Phenological metrics were extracted from the predicted SIF time series at the global scale. We estimated the start of season (SoS) and end of season (EoS) using the Maximum Separation (MS) method (Descals et al., 2020). This is a threshold-based method that can effectively estimate phenological metrics without the need of time series pre-processing prior to the phenology extraction. These types of time series pre-processing include smoothing and interpolation techniques that are applied to improve the robustness of the phenology estimates. However, these pre-processing steps may produce a time series that differs from the original, resulting in biases in the phenology estimates. The Maximum Separation method can be applied directly to the original time series.

As any threshold-based method, the MS required a threshold value to calculate the SoS and EoS from the SIF time series. For each pixel, we defined a dynamic threshold, which represented 20% of the amplitude plus the minimum SIF value in the time series. The MS runs a moving window that calculates the proportion of observations that are above the threshold before and after the central day of the moving window. We determined a moving window size of 120 days (including the days before and after the central day). The moving window is applied for every day of the time series. SoS and EoS are defined as the days of the year when the difference in proportions (before minus after) reaches the minimum and maximum during the year. The implementation of the MS method is available in Python and in Google Earth Engine (Descals et al., 2020).

6.4. Results

The combination of hyperparameters that lead to the lowest RMSE in the validation dataset is shown in Supplementary Table 6.1. For these hyperparameters, RMSE was 0.21 mW m^{-2} sr^{-1} nm^{-1}, ME was -0.00 mW m^{-2} sr^{-1} nm^{-1}, and R^2 was 0.38. The accuracy metrics differed slightly depending on the land cover type (Supplementary Figure 6.2). The lowest accuracy was found in DBF (ME = 0.14 mW m^{-2} sr^{-1} nm^{-1}, RMSE = 0.35 mW m^{-2} sr^{-1} nm^{-1}), while the accuracy of the other land covers was close to the overall accuracy, with a minimal ME (ranging from -0.04 mW m^{-2} sr^{-1} nm^{-1} in OSH to 0.06 mW m^{-2} sr^{-1} nm^{-1} in MX) and similar RMSE (ranging from 0.15 mW m^{-2} sr^{-1} nm^{-1} in CSH to 0.23 mW m^{-2} sr^{-1} nm^{-1} in MX).

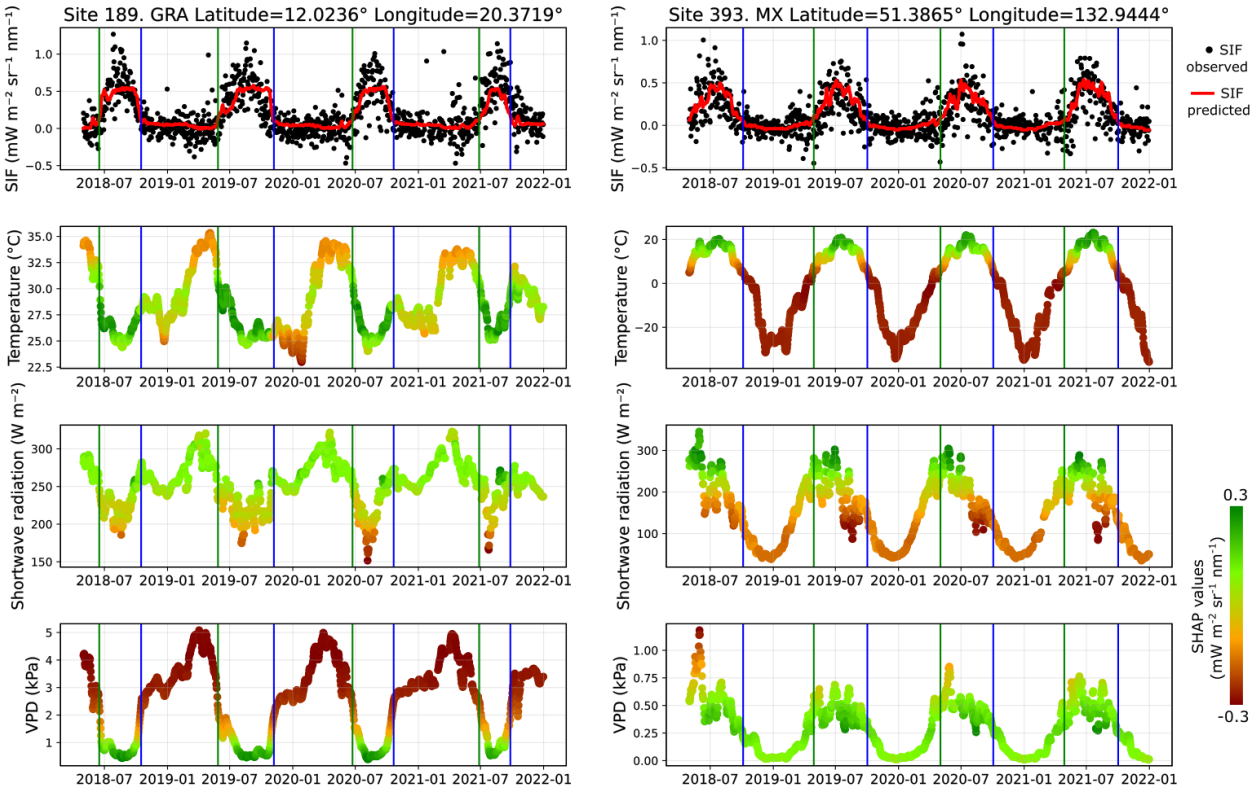


Figure 6.1. Time series of observed and predicted sun-induced fluorescence (SIF), air temperature, shortwave radiation, and vapor-pressure-deficit (VPD) in one grassland (GRA) and one mixed forest (MX) sites of the BELMANIP2 network. Vertical green lines depict the start of the growing season and blue lines represent the end of the growing season derived from predicted SIF time series. The observed SIF was extracted from the TROPOSIF dataset, while the predicted SIF was estimated with three climate variables using a machine learning regression model. Colours in the air temperature, shortwave radiation, and VPD time series depict the SHAP values. SHAP values indicate the impact of the input variables on the model mean SIF. Negative SHAP values mean that the input variable decreases the predicted SIF.

The model saturated the predicted values to 0.5 mW m^{-2} sr^{-1} nm^{-1} in observations with high SIF. Some observations above that value were underestimated, particularly in DBF and MX. Overall, the model fitted the data without substantial biases for SIF observations below 0.5 mW m^{-2} sr^{-1} nm^{-1}, which were the bulk of SIF observations. The cross-site validation did not differ substantially from the overall statistics except for the mean error. RMSE was 0.21 mW m^{-2} sr^{-1} nm^{-1}, ME was -7.22 mW m^{-2} sr^{-1} nm^{-1}, and R^2 was 0.37. Time series for two Belmanip2 sites exemplify the predicted SIF compared to the TROPOSIF measurements (Figure 6.1), and show that the model replicates the seasonality of the observed SIF, as also shown in the comparison between phenology metrics estimated with observed and predicted SIF (Supplementary Figure 6.3 and Supplementary Figure 6.4).

The time series of SHAP values represented the impact of the input climate variables on the predicted SIF. For instance, both low temperature and low shortwave radiation during winter were the most limiting factors in the Belmanip2 site 393 (Figure 6.1), located mid-latitude in a temperate climate. In Belmanip2 site 189 –a site in a dryland ecosystem– the SHAP values indicate that seasonal changes in VPD determined the SIF seasonality, with the growing season occurring when VPD values decrease to their annual minimum. These SHAP time series show the seasonal climate constraints throughout the year, and the constraints can be extracted at the start and end of the growing season. For example, in the site covering a dryland ecosystem, the SHAP values at the end of the growing season 2019 were 0.18 mW m^{-2} sr^{-1} nm^{-1} for air temperature, 0.03 mW m^{-2} sr^{-1} nm^{-1} for shortwave radiation and -0.26 mW m^{-2} sr^{-1} nm^{-1} for VPD. The low SHAP value for VPD means that this variable had a negative contribution on the prior expectation of the SIF model (0.13 mW m^{-2} sr^{-1} nm^{-1}), indicating that VPD was constraining vegetation activity at that moment of the year. The predicted SIF at the end of season was 0.08 mW m^{-2} sr^{-1} nm^{-1}, which is the result of adding the SHAP values (0.18 + 0.03 – 0.26 mW m^{-2} sr^{-1} nm^{-1}) to the prior expectation of the model (0.13 mW m^{-2} sr^{-1} nm^{-1}) (Supplementary Figure 6.5).

SHAP values for the Belmanip2 sites show that high temperature and shortwave radiation had an overall positive impact on modelled SIF, while higher VPD had a negative impact (Figure 6.2). VPD was the variable with the highest overall importance followed by temperature and shortwave radiation. The most extreme VPD values had a effect of approximately -0.3 and 0.3 mW m^{-2} sr^{-1} nm^{-1}, while the lowest and highest shortwave radiation had a lower effect, approximately -0.1 and 0.15 mW m^{-2} sr^{-1} nm^{-1}.

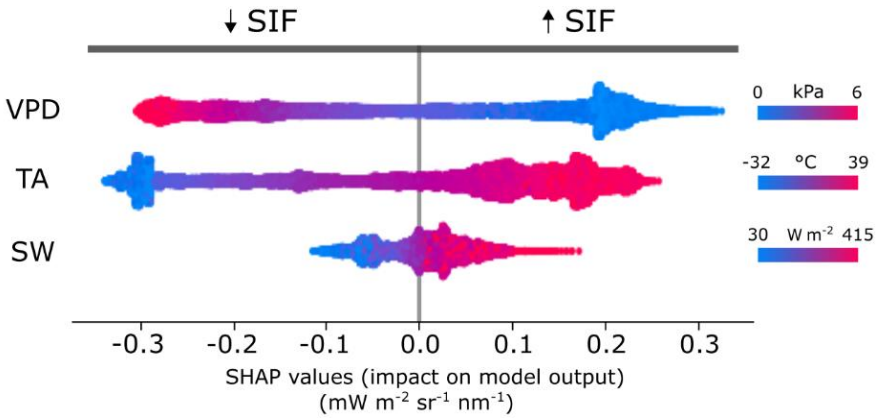


Figure 6.2. SHAP values of three input variables (air temperature (TA), shortwave radiation (SW), and vapor-pressure-deficit (VPD)) in a machine learning model that predicts sun-induced fluorescence (SIF). The SHAP values were estimated for 35,242 site-year observations of the Belmanip2 network (25% of the total observations were kept for model validation). SHAP values indicate the contribution of the input climate variables on the mean SIF. Negative SHAP values mean that the input variable decreases the predicted SIF and vice versa. The higher the absolute SHAP value, the higher the impact on predicted SIF. Color bars indicate the range of values (minimum to maximum) for each climate variable.

The impact of temperature, shortwave radiation, and VPD on predicted SIF differed spatially during the SoS and EoS. The maps of SHAP values at the SoS and EoS show that it was mostly in extratropical areas that temperature constrained SIF (Figure 6.3). However, VPD was the highest constraint in tropical dryland ecosystems, while radiation was the limiting factor in tropical rainforests. The impact of temperature, shortwave radiation and VPD differed in some regions depending on whether it was the start or the end of the season. The most prominent difference was observed in extratropical regions. Temperature was the only factor that constrained SIF during the SoS (except for Europe) but both temperature and secondarily radiation constrained SIF in extratropical regions during the EoS.

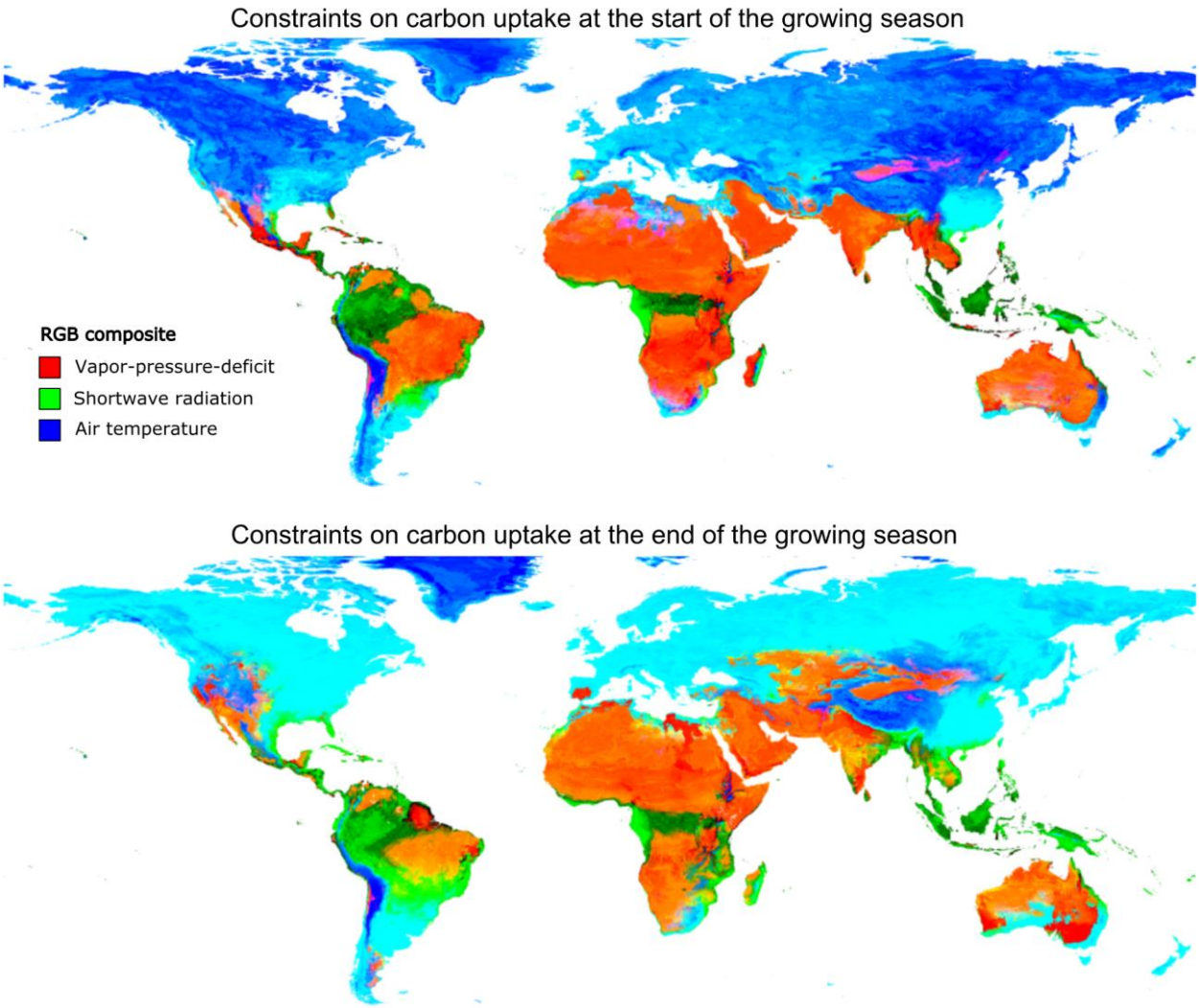


Figure 6.3. Maps of the climate constraints on carbon uptake at the start and end of the growing season. The maps represent an RGB composite of SHAP values. SHAP values were estimated from a machine learning model that fitted sun-induced fluorescence (SIF) with three climate variables: air temperature, shortwave radiation, and vapor-pressure-deficit. The SoS and EoS were estimated from daily averaged predicted SIF time series for the 2012-2021 period. The three climate variables were extracted from the ERA5-Land dataset and daily averaged for the 2012-2021 period. Low SHAP values indicate that the input variable decreases the average modelled SIF, suggesting the climate variable constrains carbon uptake. The maps depict the inverse of the SHAP values for illustration purposes (higher values indicate a greater SIF constraint).

6.5. Discussion

The results demonstrated the capabilities of SHAP in a case study that made use of geospatial climate data as input variables of a machine learning model. A weather-SIF model was trained on Belmanip2 sites using ERA5-Land and TROPOSIF measurements. SHAP values showed the

spatial and temporal impacts of the three climate variables –air temperature, shortwave radiation, and VPD– on SIF, indicating climate constraints on the carbon uptake dynamics.

The model showed good performance for replicating the vegetation seasonality even though it only considered weather variables as input. For low and medium values, the model had a minimal bias, but it struggled to predict high values (SIF >5 mW m^{-2} sr^{-1} nm^{-1}), which were largely underestimated. The underestimated SIF would be potentially corrected if a proxy for the fraction of Absorbed Photosynthetically Active Radiation (fAPAR) was included in the model. fAPAR might explain the variance in high SIF values for similar weather conditions. Proxies for fAPAR are the NDVI and EVI. Should a spectral index be used, the model would have a similar typology of input variables as the GOSIF product (Li and Xiao, 2019). GOSIF used EVI and weather variables (air temperature, VPD, and photosynthetically active radiation (PAR)) to predict Orbiting Carbon Observatory-2 (OCO-2) measurements. If we included a spectral index, however, the model would be less explainable because part of the predicted SIF would be attributed to changes in NDVI, which would conceal the marginal contributions of weather variables on the predicted SIF. Besides that, our model explanation aimed to understand the impact on SIF purely attributed to climate variables and, thus, providing accurate SIF predictions were less important than reproducing the SIF seasonality.

SHAP values confirmed previous finding on the spatial and temporal climate constraints on the vegetation activity. SHAP maps showed that VPD was the main factor limiting SIF in tropical dryland ecosystems at the start and end of the growing season. The growing season in these areas coincides with the rainfall season. High evaporative demand induces vegetation –mostly grasslands and sparse woody vegetation– into dormancy in the form of deep roots (Zhou et al., 2020). In tropical rainforests, however, both temperature and water are adequate for plants, and radiation was the only factor found to constrain SIF. This differs from the maps produced by (Jolly et al., 2005) which depicted tropical rainforests (Amazon and Central Africa) without any climate limitations. However, previous studies do suggest that radiation is a limiting factor in this biome, which supports our finding (Aguilos et al., 2018; Weber et al., 2009). In extratropical areas, temperature was the main constraint at the start and end of the growing season. Also, in extratropical areas, cold temperatures drive vegetation into dormancy, but we see a divergent constraint in terms of radiation. Overall, radiation was not limiting SIF during the SoS, except for Europe. This is due to the higher temperatures found in Europe compared to the rest of regions at similar latitudes. At the EoS, the radiation constraint was more generalized in the extratropical

areas of the Northern Hemisphere. This is consistent with recent findings (Zhang et al., 2020), which show that due to radiation constraints, rising temperatures will not increase autumn greening. The results obtained with SHAP are, thus, in line with the current understanding of carbon uptake dynamics.

SHAP proved to be a useful technique for explaining the correlations between SIF and climate factors that were captured by the machine learning model. The explainability of the ML models with SHAP, on the other hand, must be considered with caution. SHAP values show researchers which correlations machine learning has found, but these correlations do not necessarily imply a causality between input and output variables (Heskes et al., 2020). Expert knowledge is required to determine whether the correlations are coherent with the reality of the problem, and further research is required to determine whether causality exists. In our case study, we validated our findings with literature that supported the results revealed by SHAP maps. The capability of SHAP to explain spatially and temporally the predictions from geospatial gridded time series might assist remote sensing applications.

6.6. References

Aguilos, M., Hérault, B., Burban, B., Wagner, F., Bonal, D., 2018. What drives long-term variations in carbon flux and balance in a tropical rainforest in French Guiana? Agricultural and Forest Meteorology 253, 114–123.

Ahmad, R., Yang, B., Ettlin, G., Berger, A., Rodríguez-Bocca, P., 2020. A machine-learning based ConvLSTM architecture for NDVI forecasting. International Transactions in Operational Research.

Barkhordarian, A., Saatchi, S.S., Behrangi, A., Loikith, P.C., Mechoso, C.R., 2019. A recent systematic increase in vapor pressure deficit over tropical South America. Scientific reports 9, 1–12.

Bentéjac, C., Csörg\Ho, A., Martínez-Muñoz, G., 2021. A comparative analysis of gradient boosting algorithms. Artificial Intelligence Review 54, 1937–1967.

Descals, A., Verger, A., Yin, G., Peñuelas, J., 2020. A Threshold Method for Robust and Fast Estimation of Land-Surface Phenology Using Google Earth Engine. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14, 601–606.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Annals of statistics 1189–1232.

Fu, Y., Li, X., Zhou, X., Geng, X., Guo, Y., Zhang, Y., 2020. Progress in plant phenology modeling under global climate change. Science China Earth Sciences 63, 1237–1247.

Guanter, L., Aben, I., Tol, P., Krijger, J., Hollstein, A., Köhler, P., Damm, A., Joiner, J., Frankenberg, C., Landgraf, J., 2015. Potential of the TROPOspheric Monitoring Instrument (TROPOMI) onboard the Sentinel-5 Precursor for the monitoring of terrestrial chlorophyll fluorescence. Atmospheric Measurement Techniques 8, 1337–1352.

Guanter, L., Bacour, C., Schneider, A., Aben, I., van Kempen, T.A., Maignan, F., Retscher, C., Köhler, P., Frankenberg, C., Joiner, J., others, 2021. The TROPOSIF global sun-induced fluorescence dataset from the Sentinel-5P TROPOMI mission. Earth System Science Data 13, 5423–5440.

Heskes, T., Sijben, E., Bucur, I.G., Claassen, T., 2020. Causal shapley values: Exploiting causal knowledge to explain individual predictions of complex models. Advances in neural information processing systems 33, 4778–4789.

Johnsen, P.V., Riemer-Sørensen, S., DeWan, A.T., Cahill, M.E., Langaas, M., 2021. A new method for exploring gene–gene and gene–environment interactions in GWAS with tree ensemble methods and SHAP values. BMC bioinformatics 22, 1–29.

Jolly, W.M., Nemani, R., Running, S.W., 2005. A generalized, bioclimatic index to predict foliar phenology in response to climate. Global Change Biology 11, 619–632.

Keenan, T.F., Gray, J., Friedl, M.A., Toomey, M., Bohrer, G., Hollinger, D.Y., Munger, J.W., O'Keefe, J., Schmid, H.P., Wing, I.S., others, 2014. Net carbon uptake has increased through warming-induced changes in temperate forest phenology. Nature Climate Change 4, 598–604.

Li, X., Xiao, J., 2019. A global, 0.05-degree product of solar-induced chlorophyll fluorescence derived from OCO-2, MODIS, and reanalysis data. Remote Sensing 11, 517.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2020. From local explanations to global understanding with explainable AI for trees. Nature machine intelligence 2, 56–67.

Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. Advances in neural information processing systems 30.

Lundberg, S.M., Nair, B., Vavilala, M.S., Horibe, M., Eisses, M.J., Adams, T., Liston, D.E., Low, D.K.-W., Newman, S.-F., Kim, J., others, 2018. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. Nature biomedical engineering 2, 749–760.

Meng, Y., Yang, N., Qian, Z., Zhang, G., 2020. What makes an online review more helpful: an interpretation framework using XGBoost and SHAP values. Journal of Theoretical and Applied Electronic Commerce Research 16, 466–490.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., others, 2021. ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. Earth System Science Data Discussions 1–50.

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

Piao, S., Liu, Q., Chen, A., Janssens, I.A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., Zhu, X., 2019. Plant phenology and global climate change: Current progresses and challenges. Global change biology 25, 1922–1940.

Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence 1, 206–215.

Shapley, L.S., 1953. Stochastic games. Proceedings of the national academy of sciences 39, 1095–1100.

Wang, D., Thunéll, S., Lindberg, U., Jiang, L., Trygg, J., Tysklind, M., 2022. Towards better process management in wastewater treatment plants: Process analytics based on SHAP values for tree-based machine learning methods. Journal of Environmental Management 301, 113941.

Weber, U., Jung, M., Reichstein, M., Beer, C., Braakhekke, M., Lehsten, V., Ghent, D., Kaduk, J., Viovy, N., Ciais, P., others, 2009. The interannual variability of Africa's ecosystem productivity: a multi-model analysis. Biogeosciences 6, 285–295.

Weiss, M., Baret, F., Verger, A., 2014. BELMANIP2: Enhancement of the CEOS-BELMANIP ensemble of sites used for the validation of land products from medium resolution sensors., in: Fourth International Symposium on Recent Advances in Quantitative Remote Sensing.

Zhang, Y., Commane, R., Zhou, S., Williams, A.P., Gentine, P., 2020. Light limitation regulates the response of autumn terrestrial carbon uptake to warming. Nature Climate Change 10, 739–743.

Zhou, Y., Wigley, B.J., Case, M.F., Coetsee, C., Staver, A.C., 2020. Rooting depth as a key woody functional trait in savannas. New Phytologist 227, 1350–1361.

General discussion and conclusions

Satellite data is important for phenology studies, but limitations in the interpretation of time series, method and variable selection, and spatial resolution must be considered. The first two chapters tackle these uncertainties; Chapter 1 solves the problem of time series pre-processing in the estimation of phenology metrics and Chapter 2 shows the importance of high-spatial resolution for depicting vegetation dynamics at the canopy level. Other studies carried out during the thesis focused on other limitations of land surface phenology. For instance, we linked land surface phenology metrics with *in situ* measurements of phenophases, PhenoCam records, and carbon fluxes in deciduous trees (Bornez et al., 2020; Bórnez et al., 2020). We also demonstrated the divergent responses of vegetation indices during senescence stages of deciduous trees (Yin et al., 2020), indicating that vegetation indices might reflect leaf biomass or vegetation productivity and revealing the importance of variable selection.

The findings of Chapters 3-5 show that at the beginning and end of the growing season, different spatial constraints of temperature, light, and water availability exist, implying that vegetation phenology will respond differently to future climatic warming. These constraints might slow down the lengthening of the growing season in the future. This has important implications for the carbon cycle; future projections must consider that the capacity of terrestrial ecosystems to remove carbon from the atmosphere might be compromised in the 21st century. The finding in Chapters 3-5 were upscaled to the global scale with a local interpretation method. Moreover, the method presented in Chapter 6 tackles a methodological problem that broadly affects many remote sensing studies using machine learning models, and puts machine learning explainability into a new and key perspective for the remote sensing community. In the following paragraphs, we will briefly summarize the findings presented in each of the chapters as a general discussion and conclusions of the thesis.

We proposed a new land surface phenology method in Chapter 1, which had the advantage of allowing rapid processing of LSP maps on a global scale using a cloud-based platform. Furthermore, the method can handle raw time series without the use of preprocessing techniques like smoothing, gap filling, or interpolation. Other methods required time series pre-processing (Bornez et al., 2020) or the use of smoothed and interpolated products (Verger et al., 2016). These preprocessing techniques may distort the time series seasonality, whereas our method processes the data while maintaining the seasonality. Our method is more practical than

the standard threshold method because of the ease with which it can be applied to various satellite products. As in other chapters where we used our method to extract land surface phenology metrics, our study could help future phenology studies that require a fast and reliable method for extracting land surface phenology metrics.

Cloud-based platforms, Google Earth Engine in particular (Gorelick et al., 2017), also enabled the processing of land surface phenology metrics at the continental scale, as demonstrated in Chapter 2, in which we presented maps of the start and end of season for the Arctic. The chapter demonstrates that the high agreement between the phenology metrics estimated with Sentinel-2 and MODIS confirms the viability of LSP estimation due to the short revisit time of the Sentinel-2 time series (Bolton et al., 2020), even in cloudy regions. The data was further validated by the high level of similarity between the phenology metrics estimated using Sentinel-2 and MODIS and PhenoCam. When continuous gaps in the time series are present, however, the combination of Sentinel-2 and Landsat-8 is recommended. The chapter demonstrates how estimating land surface phenology at a 10-meter resolution allows for the study of canopy-level vegetation dynamics. This is crucial for comprehending the processes that take place in small forest patches, as shown in Chapter 4. Furthermore, 10-meter resolution is critical for accurately depicting the dynamics of vegetation in heterogeneous landscapes like crop-forest mosaics.

In Chapter 3, we found that spring onset of vegetation activity and soil thawing were closely associated in tundra and alpine grasslands, and weakly associated where soil froze intermittently. Deciduous forests at southerly latitudes required greater heat accumulation to unfold their leaves and initiate spring growth than tundra and needleleaved forests (Piao et al., 2019). The onset in greenness in needleaved forests, in particular boreal forests, was much closer to the end of the thawing period, further suggesting that this vegetation type starts the growing season when temperatures become favorable. The chapter also showed that spring onset rarely occured with negative soil temperatures, indicating that frozen soil acted as a major constraint on vegetation activity. However, the possible restriction in vegetation growth because of other climate constraints, such as insufficient levels of incoming radiation, was still undetermined but covered in Chapter 4.

Chapter 4 also confirms that carbon uptake by non-deciduous forests begins shortly after favorable growth conditions arise and ends when these conditions decline. The chapter confirmed that limiting factors of vegetation activity were temperature and radiation in

temperate and cold regions, and water availability in dryland ecosystems. Furthermore, we demonstrated that climate constraints differed between regions and seasonally. The strongest radiation constraint was found in temperate Europe, where the maritime influence maintained temperatures that were favorable for vegetation growth, especially in the autumn. Temperature, on the other hand, largely constrained carbon uptake phenology in most of the Northern Hemisphere, particularly in high-elevated regions. These findings suggest that future warming may advance the SoS, but at a slower rate. As previously reported (Zhang et al., 2020), radiation exerts a constraint on photosynthesis, which could prevent the EoS from advancing significantly with future warming. Carbon uptake phenology in tropical dryland ecosystems, on the other hand, is sensitive to aridity, suggesting that as the climate warms, the length of the growing season will shorten.

Water stress may also play an increasing role in leaf shedding in temperate deciduous forest under global warming. Chapter 5 demonstrates that early leaf shedding was linked to unusually high temperatures and aridity in the days preceding the event. Most occurrences were in central Europe in 2018, as previously reported (Brun et al., 2020), and the Carpathians in 2019, although early leaf shedding was detected in all years of the Sentinel-2 time series across Europe. Higher temperatures mean higher evaporative demands (Sherwood & Fu, 2014), which seems to induce the early leaf shedding and disrupt leaf senescence in deciduous forests. The results suggest that future increases in droughts will shrink the length of the growing season and negatively influence carbon uptake from forests.

In the last chapter, a novel remote sensing technique (Lundberg & Lee, 2017) confirmed worldwide constraints on carbon uptake phenology (Jolly et al., 2005). VPD was the key factor restricting SIF in tropical dryland ecosystems during the start and end of the growing season, according to the SHAP maps. Tropical rainforests have availability to both temperature and water, but radiation was the limiting factor for SIF. SHAP values also corroborated the findings in Chapter 4 about the divergent radiation constraints at the start and end of the season in extratropical Northern Hemisphere regions.

Phenology research that makes use of satellite remote sensing may benefit from our recent developments in land surface phenology estimation as well as the SHAP technique that we have proposed. Our research demonstrates that decametric satellite data can be used to successfully monitor vegetation phenology and the responses of forest canopies to drought at the canopy

level. Our findings have important implications for phenology modeling, as well as the study of the global carbon cycle and the responses of vegetation to climatic warming.

References

Bolton, D. K., Gray, J. M., Melaas, E. K., Moon, M., Eklundh, L., & Friedl, M. A. (2020). Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. *Remote Sensing of Environment*, *240*, 111685.

Bornez, K., Descals, A., Verger, A., & Peñuelas, J. (2020). Land surface phenology from VEGETATION and PROBA-V data. Assessment over deciduous forests. *International Journal of Applied Earth Observation and Geoinformation*, *84*, 101974.

Bórnez, K., Richardson, A. D., Verger, A., Descals, A., & Peñuelas, J. (2020). Evaluation of vegetation and proba-v phenology using phenocam and eddy covariance data. *Remote Sensing*, *12*(18), 3077.

Brun, P., Psomas, A., Ginzler, C., Thuiller, W., Zappa, M., & Zimmermann, N. E. (2020). Large-scale early-wilting response of Central European forests to the 2018 extreme drought. *Global Change Biology*, *26*(12), 7021–7035.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27.

Jolly, W. M., Nemani, R., & Running, S. W. (2005). A generalized, bioclimatic index to predict foliar phenology in response to climate. *Global Change Biology*, *11*(4), 619–632.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*.

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., & Zhu, X. (2019). Plant phenology and global climate change: Current progresses and challenges. *Global Change Biology*, *25*(6), 1922–1940.

Sherwood, S., & Fu, Q. (2014). A drier future? *Science*, *343*(6172), 737–739.

Verger, A., Filella, I., Baret, F., & Peñuelas, J. (2016). Vegetation baseline phenology from kilometric global LAI satellite products. *Remote Sensing of Environment*, *178*, 1–14.

Yin, G., Verger, A., Filella, I., Descals, A., & Peñuelas, J. (2020). Divergent estimates of forest photosynthetic phenology using structural and physiological vegetation indices. *Geophysical Research Letters*, *47*(18), e2020GL089167.

Zhang, Y., Commane, R., Zhou, S., Williams, A. P., & Gentine, P. (2020). Light limitation regulates the response of autumn terrestrial carbon uptake to warming. *Nature Climate Change*, *10*(8), 739–743.

Supplementary Material

Chapter 1

A threshold method for robust and fast estimation of land-surface phenology using Google Earth Engine

Supplementary Table 1.1. Vegetation indices used in the extraction of Land Surface Phenology (LSP) metrics.

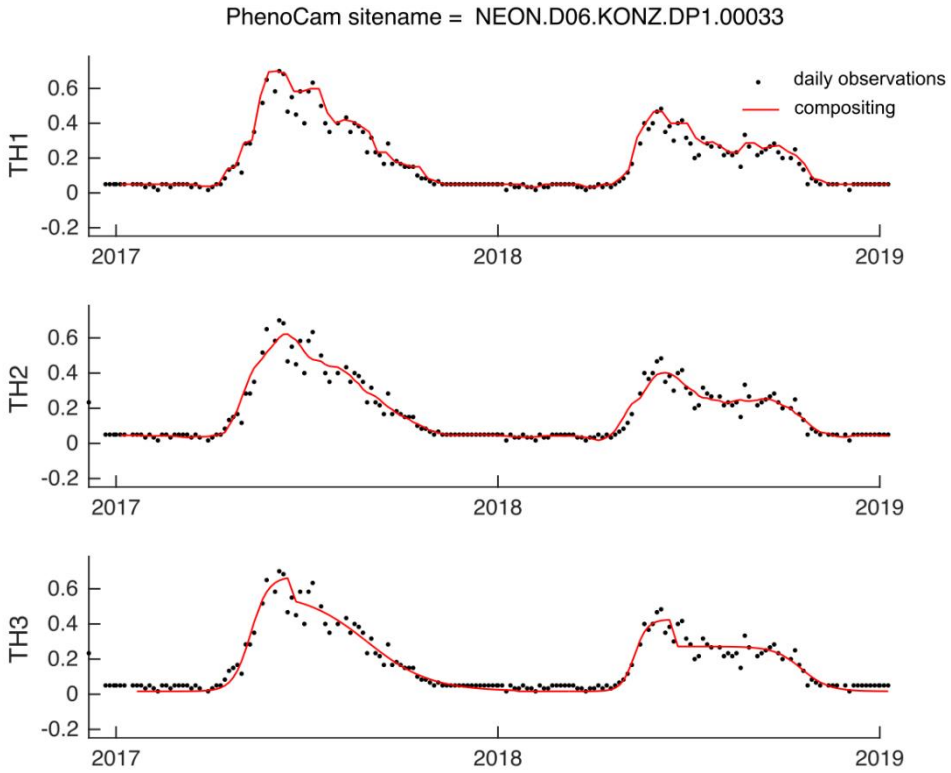
Vegetation Index	Formula
Normalized Vegetation Difference Index (NDVI)	$NDVI = (NIR - Red) / (NIR + Red)$
Enhanced Vegetation Index (EVI)	$EVI = Green \times (NIR - Red) / (NIR + 6 \times Red - 7.5 \times Blue + 1)$
Green Chromatic Coordinate (GCC)	$GCC = Green / (Green + Red + Blue)$

Supplementary Table 1.2. Site name, time coverage and land cover of the PhenoCam sites used in the study. The land cover types are evergreen needleaved forests (ENF), deciduous broadleaved forests (DBF), mixed forests (MX), open shrublands (OSH), woody savannah (WSA), savannah (SAV), grasslands (GRA), temporary crops (CRO1), and cropland/natural vegetation mosaics (CRO2).

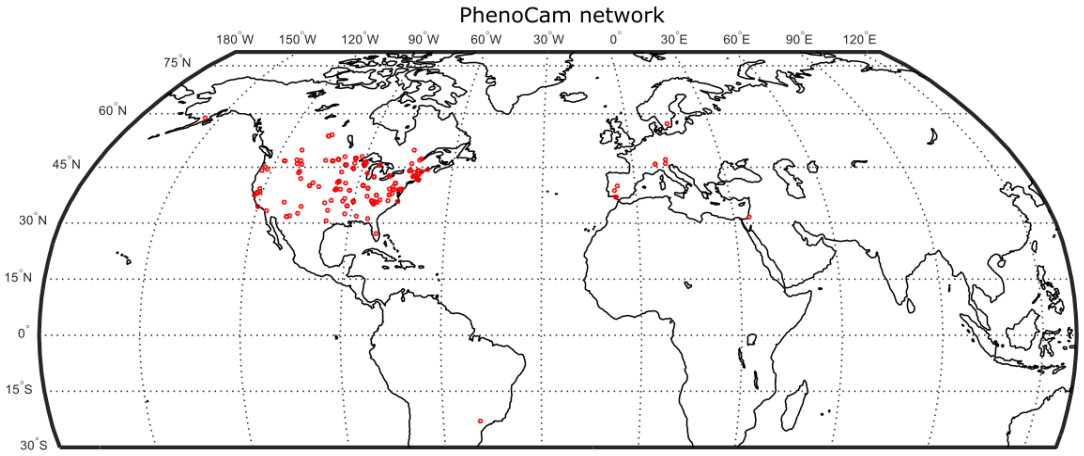
Site name	Year start	Year end	Land Cover	Site name	Year start	Year end	Land Cover	Site name	Year start	Year end	Land Cover
NEON.D01.BART.DP1.00033	2016	2018	MF	harvard	2008	2018	MF	spruceA0P21SH	2017	2018	MF
NEON.D01.BART.DP1.00042	2016	2018	MF	harvardbarn2	2012	2018	MF	spruceT0P06SH	2017	2018	MF
NEON.D01.HARV.DP1.00033	2016	2018	MF	harvardbarn	2011	2018	MF	spruceT0P06	2015	2018	MF
NEON.D01.HARV.DP1.00042	2016	2018	MF	harvardblo	2009	2018	MF	spruceT0P19ESH	2017	2018	MF
NEON.D02.BLAN.DP1.00033	2016	2018	CRO2	harvardfarmnorth	2015	2018	CRO2	spruceT0P19E	2015	2018	MF
NEON.D02.SCBI.DP1.00033	2016	2018	DBF	harvardfarmsouth	2015	2018	MF	spruceT2P11ESH	2017	2018	MF
NEON.D02.SERC.DP1.00033	2016	2018	DBF	harvardgarden	2016	2018	MF	spruceT2P11E	2015	2018	MF
NEON.D03.JERC.DP1.00042	2016	2018	WSA	harvardhemlock2	2015	2018	MF	spruceT2P2OSH	2017	2018	MF
NEON.D05.TREE.DP1.00033	2016	2018	MF	harvardhemlock	2010	2012	MF	spruceT2P2O	2015	2018	MF
NEON.D05.TREE.DP1.00042	2016	2018	MF	harvardlph	2010	2011	MF	spruceT4P04ESH	2017	2018	MF
NEON.D05.UNDE.DP1.00033	2016	2018	MF	hawbeckereddy	2015	2018	MF	spruceT4P04E	2015	2018	MF
NEON.D06.KONA.DP1.00042	2016	2018	GRA	humnokericea	2015	2018	CRO1	spruceT4P13SH	2017	2018	MF
NEON.D06.KONZ.DP1.00033	2017	2018	GRA	humnokericec	2015	2018	CRO1	spruceT4P13	2015	2018	MF
NEON.D06.KONZ.DP1.00042	2017	2018	GRA	innsbruck	2013	2018	ENF	spruceT6P08	2015	2018	MF
NEON.D07.GRSM.DP1.00033	2017	2018	MF	intervale	2015	2016	MF	spruceT6P16ESH	2017	2018	MF
NEON.D07.GRSM.DP1.00042	2017	2018	MF	jasperridge	2012	2017	WSA	spruceT6P16E	2015	2018	MF
NEON.D07.ORNL.DP1.00033	2016	2018	MF	jernort	2014	2018	OSH	spruceT9P10ESH	2017	2018	MF
NEON.D07.ORNL.DP1.00042	2016	2018	MF	joycekilmer	2006	2013	DBF	spruceT9P10E	2015	2018	MF
NEON.D07.WALK.DP1.20002	2016	2018	DBF	juncabalejo	2016	2018	CRO1	spruceT9P17	2015	2018	MF
NEON.D08.LENO.DP1.00033	2016	2018	MF	kansas	2012	2018	CRO2	stjones	2015	2018	CRO2
NEON.D09.DCFS.DP1.00033	2017	2018	CRO1	kelloggcorn2	2015	2018	CRO2	sweetbriar	2014	2018	DBF
NEON.D09.NOGP.DP1.00033	2017	2018	GRA	kelloggcornsoy2	2015	2018	CRO2	sweetbriargrass	2016	2018	CRO2
NEON.D09.NOGP.DP1.00042	2017	2018	GRA	kelloggmiscanthus	2015	2018	CRO2	sylvania	2015	2018	MF
NEON.D09.PRLA.DP1.20002	2017	2018	CRO1	kendall	2012	2018	GRA	teddy	2010	2018	GRA
NEON.D09.WOOD.DP1.00033	2016	2018	CRO1	keplerltar	2018	2018	CRO2	tfforest	2016	2018	MF
NEON.D10.ARIK.DP1.20002	2016	2018	GRA	kingmanfarm	2016	2018	MF	thompsonfarm2N	2009	2010	MF
NEON.D11.CLBJ.DP1.00033	2017	2017	GRA	konza	2012	2018	GRA	tonzi	2011	2018	SAV
NEON.D11.OAES.DP1.00033	2017	2018	GRA	lacclair	2014	2018	MF	torgnon-ld	2012	2018	ENF
acadia	2007	2017	MF	laclaflamme	2014	2018	MF	torgnon-nd	2013	2018	GRA
alligatorriver	2012	2018	MF	laurentides	2013	2018	MF	torrepalacio	2017	2018	OSH
arbutuslakeinlet	2015	2018	MF	lethbridge	2011	2018	GRA	turkeypointdbf	2012	2018	DBF
archboldbahia	2017	2018	CRO2	lostcreek	2015	2018	MF	turkeypointenf02	2012	2018	CRO2
arsltarmdcr	2017	2018	CRO2	luckyhills	2013	2015	OSH	turkeypointenf39	2012	2018	MF
arsmnswanlake1	2015	2018	CRO1	macleish	2017	2018	MF	turkeypointenf74	2012	2018	MF
arsmorris1	2017	2018	CRO1	mammothcave	2002	2003	DBF	twitchellalfalfa2	2016	2018	CRO1
arsmorris2	2017	2018	CRO1	mandanh5	2015	2018	CRO1	twosfpr	2018	2018	GRA
asa	2010	2014	MF	mandani2	2016	2018	CRO1	uiefmaize	2008	2018	CRO1
ashbottoms	2015	2017	GRA	manilacotton	2016	2018	CRO1	uiefmiscanthus	2008	2018	CRO1
ashburnham	2012	2016	MF	marcell	2014	2018	MF	uiefprairie	2008	2018	CRO1
asuhighlands	2016	2018	CRO2	marena	2012	2018	GRA	uiefswitchgrass	2008	2018	CRO1
bartlettir	2008	2016	MF	mayberry	2010	2018	WSA	umichbiological2	2008	2014	MF
bbc1	2015	2018	MF	mead1	2016	2018	CRO1	umichbiological	2008	2014	DBF
bbc2	2015	2018	MF	mead2	2016	2018	CRO1	upperbuffalo	2006	2007	DBF
bbc5	2015	2018	MF	mead3	2016	2018	CRO1	usgseros	2014	2017	CRO1
bbc7	2015	2018	MF	meadpasture	2016	2018	CRO1	usmpj	2013	2018	OSH
bouldincorn	2017	2018	CRO1	missouriozarks	2012	2018	DBF	usof1	2017	2017	CRO1
boundarywaters	2006	2007	MF	montebondonegrass	2015	2018	GRA	usof2	2017	2017	CRO1
bozeman	2016	2018	GRA	monture	2001	2010	CRO1	usof3	2017	2018	CRO1
bullshoals	2013	2018	DBF	morganmonroe2	2017	2018	DBF	uwmfieldsta	2013	2018	CRO2
burdettericea	2015	2018	CRO1	morganmonroe	2008	2018	DBF	vaira	2011	2018	WSA
burdettericec	2015	2018	CRO1	nationalelkrefuge	2015	2018	CRO1	warrenwilson	2016	2018	CRO2
butte	2009	2018	GRA	ninemileprairie	2015	2017	CRO1	westpond	2012	2018	CRO1
cafboydnorthltar01	2017	2018	CRO1	niwot3	2015	2018	ENF	willamettepoplar	2015	2018	CRO1
cafboydsouthltar01	2018	2018	CRO1	niwot5	2016	2018	ENF	willamettewheat	2015	2016	CRO1
cafcookeastltar01	2017	2018	CRO1	northattleboroma	2012	2018	CRO2	willowcreek	2012	2018	DBF
canadaOA	2011	2016	MF	oakridge1	2008	2009	DBF	wolfesneckfarm	2017	2018	MF
canadaOBS	2011	2018	ENF	oakridge2	2008	2016	DBF	woodshole	2011	2018	MF
canadaoa2	2016	2018	MF	oakville	2014	2018	CRO1	woodstockvt	2015	2016	MF
chibougamau	2008	2011	ENF	oregonMP	2011	2018	ENF				
coaloilpoint	2008	2012	OSH	pace	2017	2018	DBF				
columbiamissouri	2006	2007	DBF	portal	2017	2018	OSH				
coville	2010	2014	OSH	quickbird	2014	2015	GRA				
coweeta	2011	2016	MF	robinson2	2017	2018	DBF				
cperuvb	2015	2018	GRA	robinson	2017	2018	DBF				
dollysods	2003	2014	DBF	rosemount	2008	2008	CRO1				
donanafuenteduque	2017	2018	WSA	rosemountcons	2017	2018	CRO1				
drippingsprings	2001	2009	WSA	rosemountconv	2017	2018	CRO1				
dukehw	2013	2018	MF	rosemountg21	2015	2017	CRO1				
eastend2	2017	2018	CRO1	rosemountnprs	2015	2018	CRO1				
esalb	2016	2018	CRO1	russellsage	2013	2018	MF				
eslm1	2014	2018	SAV	segabluechute	2018	2018	GRA				
eucflux	2018	2018	ENF	shahariya	2011	2013	CRO1				
forbes	2017	2018	SAV	sherman	2014	2015	CRO1				
gatesofthemountains	2001	2009	ENF	smokypurchase	2003	2006	DBF				
grandrivergrass	2015	2018	CRO1	snakerivermn	2010	2010	MF				
grandteton	2015	2018	GRA	spruceA0EMI	2016	2018	MF				
greenridge1	2016	2018	DBF	spruceA0P07	2016	2018	MF				

Supplementary Table 1.3. Statistics used for the comparison between Land Surface Phenology (LSP) metrics generated with MODIS and PhenoCam. The LSP metrics are compared pairwise (x_i-y_i), where x_i represents a LSP metric extracted from PhenoCam and y_i is the same LSP metric extracted from MODIS. The LSP metrics are the Start of Season (SoS) or End of Season (EoS).

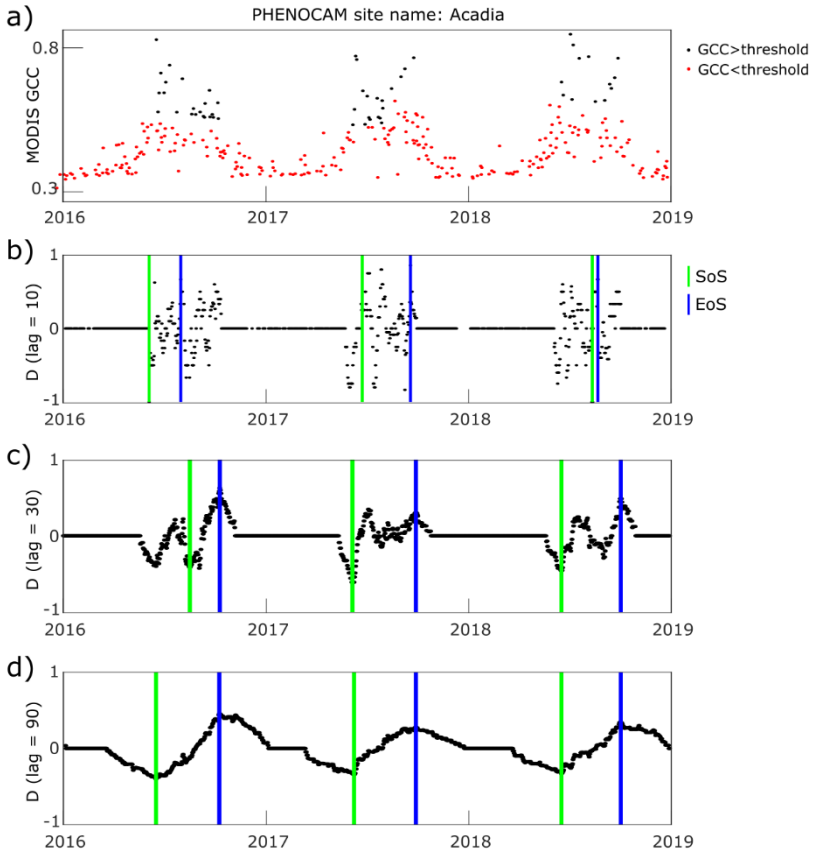
Statistic	Formula
Mean error (ME)	$ME = \frac{1}{n}\sum_{i=1}^{n} x_i - y_i$
Root mean squared error (RMSE)	$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (x_i - y_i)^2}{n}}$



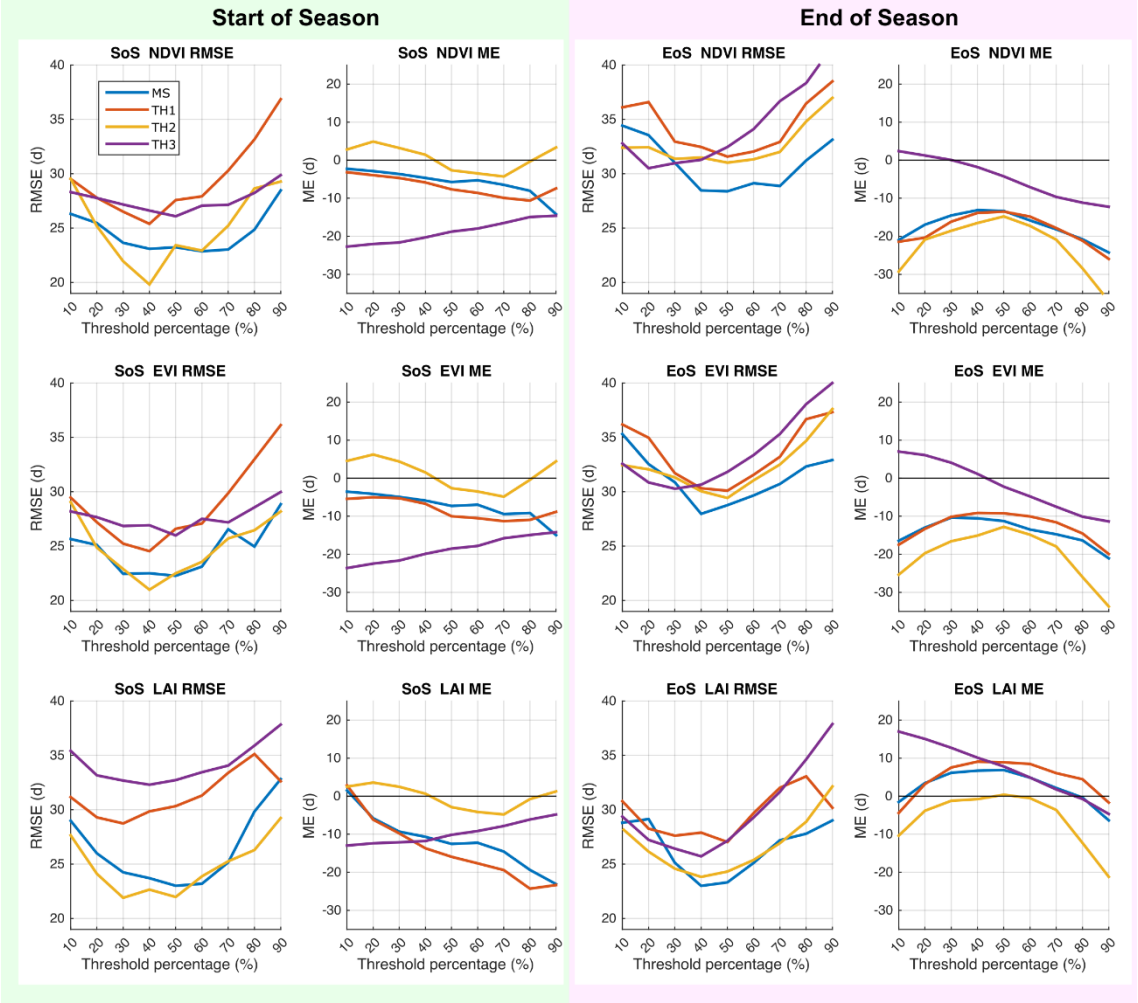
Supplementary figure 1.1. Example of three types of time series preprocessing for Land Surface Phenology estimation. TH1 and TH2 use a moving-window approach; TH1 uses an 8-day composite and then interpolates the composites to daily observations, while TH2 uses an outlier rejection algorithm that excluded low values, a smoothing step with the Savitzky–Golay filter, and a linear interpolation. TH3 fits a logistic function to the time series.



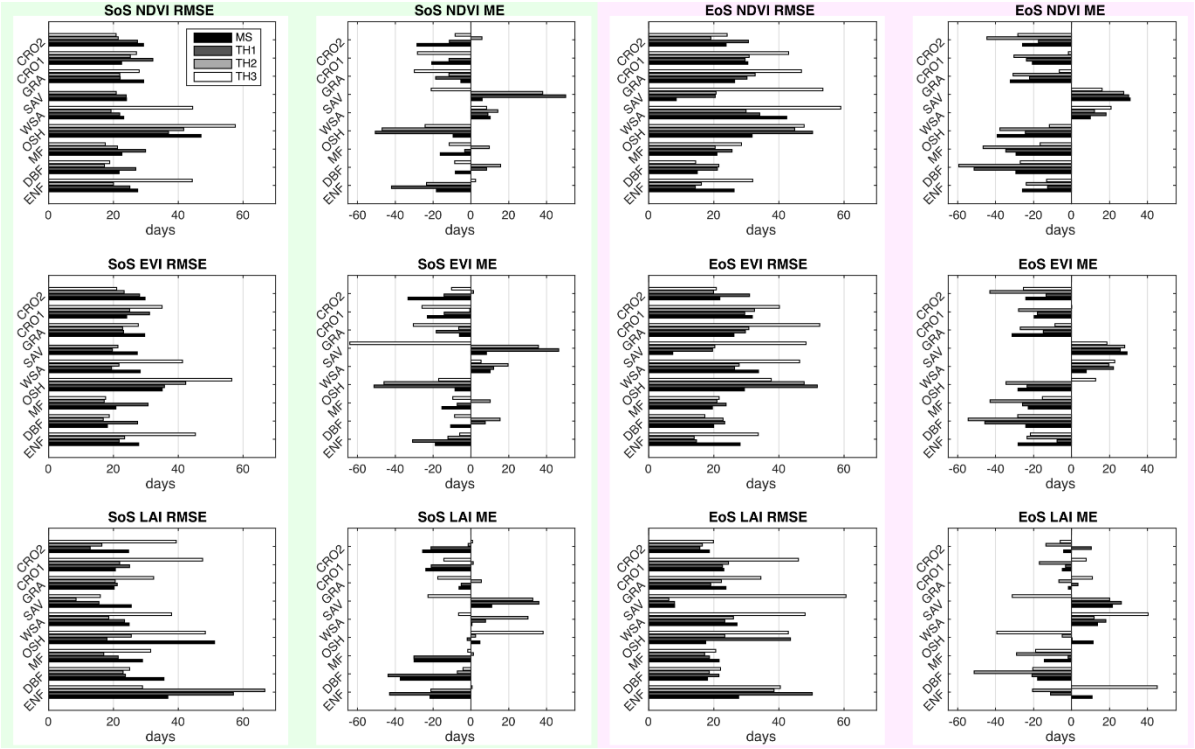
Supplementary figure 1.2. Location map of the PhenoCam sites selected in the study.



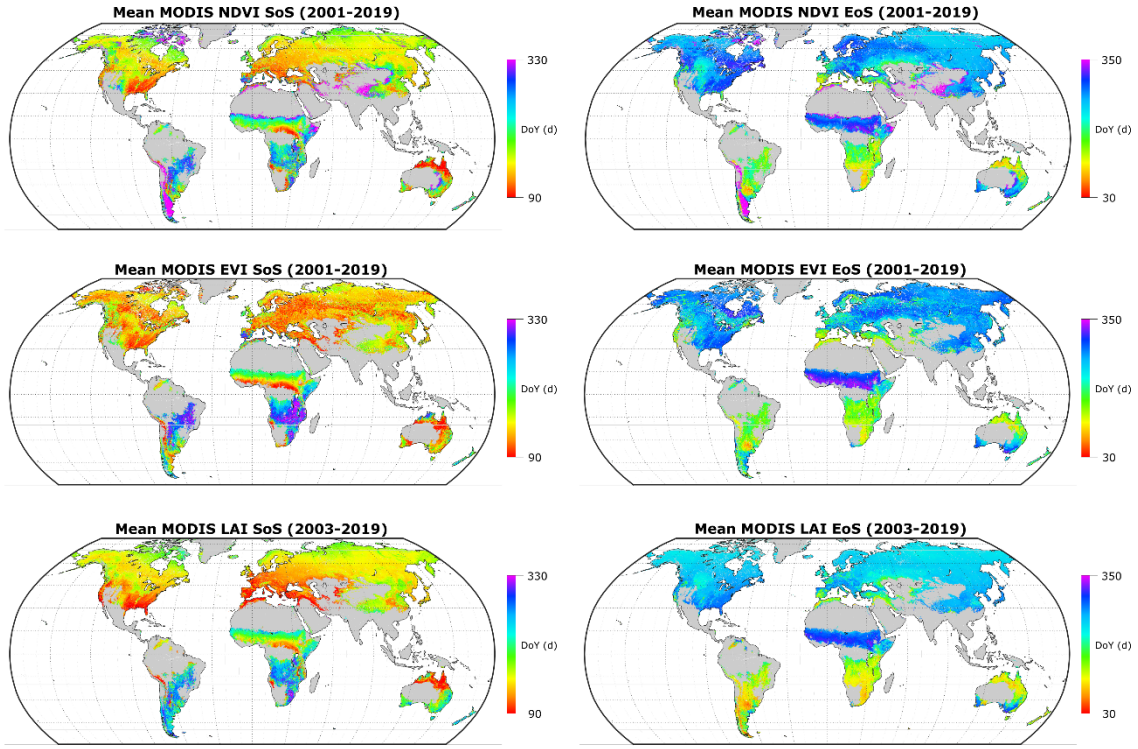
Supplementary figure 1.3. Impact of window size on the maximum separation method for the Acadia PhenoCam deciduous forest site. The time series in a) shows the daily MODIS Green Chromatic Coordinate Index (GCC) for 2016-2018. Red dots in this panel depict the observations that lay below a dynamic threshold of 50% of the amplitude. The time series in b), c), and d) show the differences in the proportion of dormant observations for different window sizes (semi-period of 10, 30, and 90 day).



Supplementary figure 1.4. Root mean squared error (RMSE) and mean error (ME) obtained from the comparison of the LSP metrics generated with the PhenoCam dataset and the MODIS normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), and leaf area index (LAI) for a range of threshold values. The LSP metrics are the Start of Season (SoS) and the End of Season (EoS).



Supplementary figure 1.5. Root mean squared error (RMSE) and mean error (ME) obtained from the comparison of the LSP metrics generated with the PhenoCam dataset and the MODIS normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), and leaf area index (LAI). The comparison is based on the land cover of the Phenocam sites: evergreen needleleaf forests (ENF), deciduous broadleaf forests (DBF), mixed forests (MX), open shrublands (OSH), woody savannah (WSA), savannah (SAV), grasslands (GRA), temporary crops (CRO1), and cropland/natural vegetation mosaics (CRO2). The LSP metrics are the Start of Season (SoS) and the End of Season (EoS).



Supplementary figure 1.6. Maps of the mean start and end of season (SoS and EoS) estimated using the Maximum Separation method and a dynamic threshold of 50% of the amplitude for years 2001-2019. The SoS and EoS are shown as the Day of Year (DoY). The Maximum Separation method was applied to the MODIS NDVI, EVI, and LAI with a window size of radius equal to 30 days. The maps were processed fully in Google Earth Engine.

Supplementary Material

Chapter 2

Improved estimates of Arctic Land Surface Phenology using Sentinel-2 time series

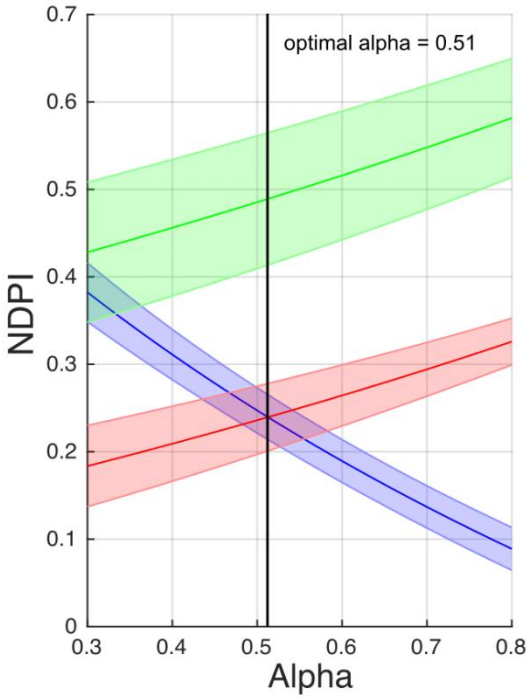
Estimation of the optimal alpha in Normalized Difference Phenology Index (NDPI)

Normalized difference phenology index (NDPI), defined as eq. 5, adjusts the snow observations to the same values of bare soil and corrects the break in the time series due to the snow-to-vegetation transition. Since the *alpha* parameter was originally set according to the spectral response of MODIS, we re-estimated the optimal *alpha* parameter for Sentinel-2.

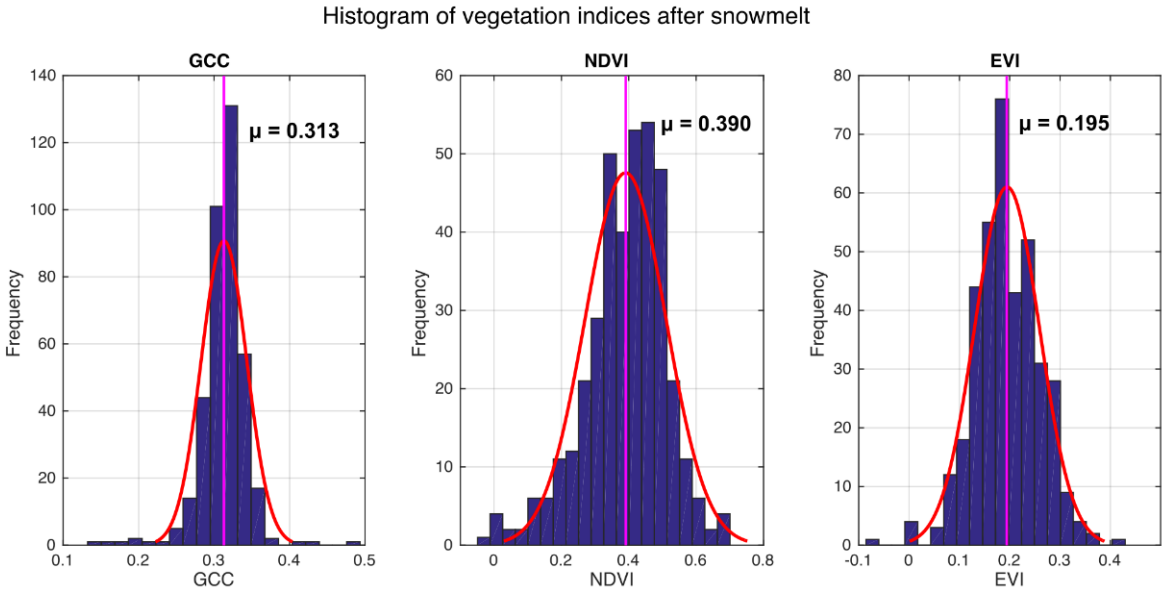
The optimal alpha was originally set so that the NDPI estimated 32 typical spectral responses to show the maximum separability between vegetation and snow-bare soil. Here, instead of using typical spectral responses from the Aster Spectral Library, we estimated the spectral responses empirically in snow, vegetation during the dormant period, and vegetation during the peak growth (Supplementary Figure 1). First, we randomly distributed 400 points covering Sentinel-2 pixels that did not present long discontinuities in the time series (maximum gap < 10 days). Then, we extracted the Red, near infrared (NIR), and shortwave-infrared (SWIR2) values of the following observations: a) The last Sentinel-2 snow observation (reflected in the scene classification layer (SCL)) before the growing season, b) the first Sentinel-2 snow-free observation of the year, and c) the observation with the maximum normalized difference vegetation index (NDVI) value. Values in a) corresponded to snow, b) corresponded to bare soil or vegetation during the dormant period, and c) corresponded to vegetation during the maximum growth activity. Finally, we searched for the *alpha* that showed the best separability between vegetation during the maximum growth, and the snow and dormant vegetation.

Supplementary Table 2.1. Site name and coordinates of the PhenoCam sites used in the study.

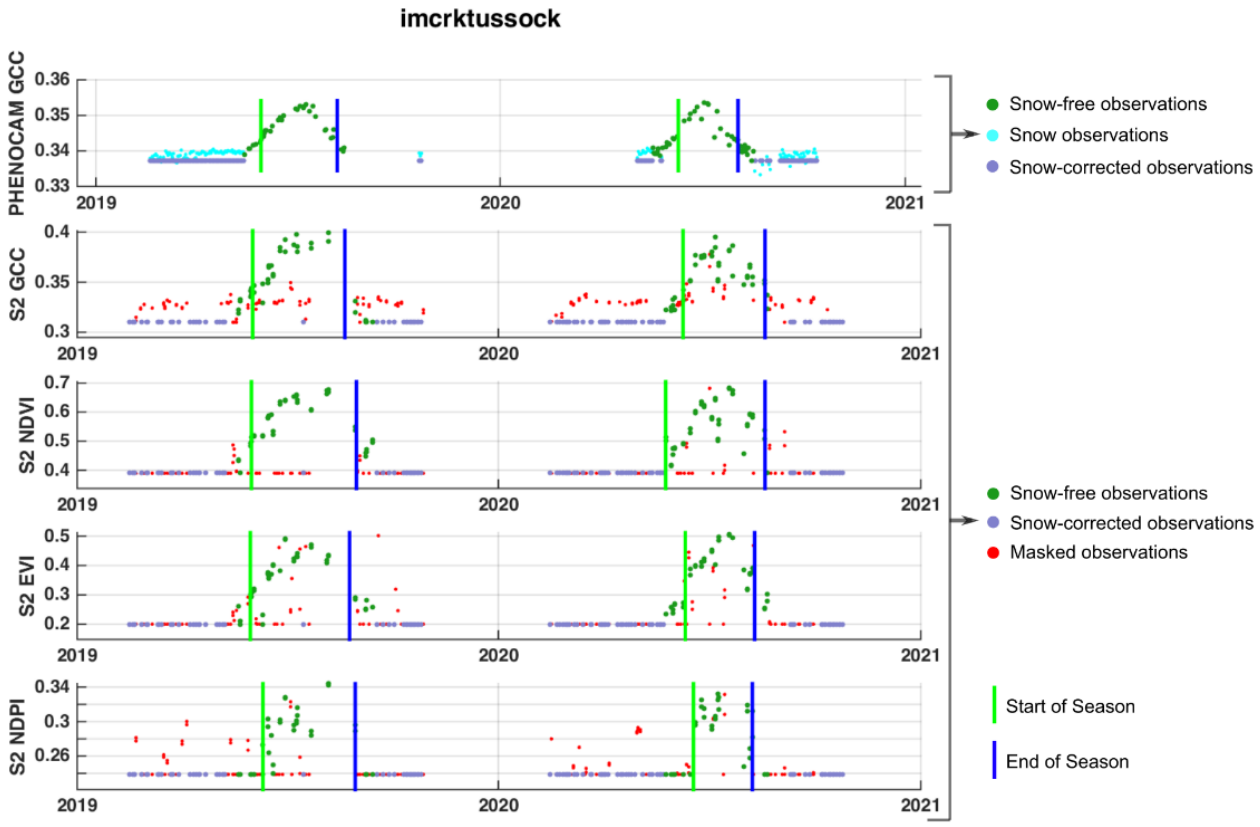
PhenoCam site name	Latitude	Longitude
imcrktussock	68.6063°N	149.3041°W
NEON.D18.TOOL.DP1.00033	68.6611°N	149.3705°W
NEON.D19.HEAL.DP1.00033	63.8757°N	149.2133°W



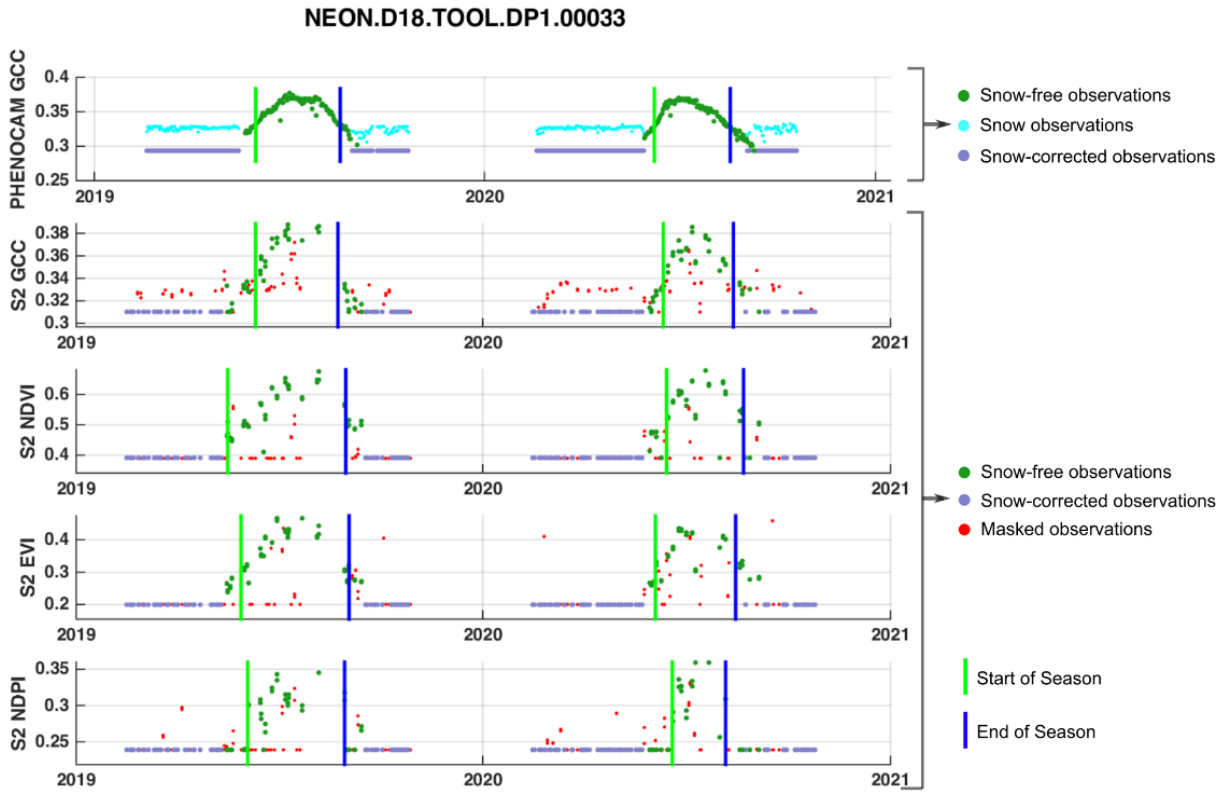
Supplementary figure 2.1. Relationship between the normalized difference phenology index (NDPI) derived from Sentinel-2 data and the alpha parameter (eq. 5) for three stages in the tundra time series: vegetation during the maximum growth (green), vegetation after snowmelt (red), and snow observations (blue). The central line is the mean of 400 samples and the error bar represents one standard deviation. The optimal alpha value that provides the maximum separability between productive vegetation and vegetation during the dormant period and snow is 0.51 (black line).



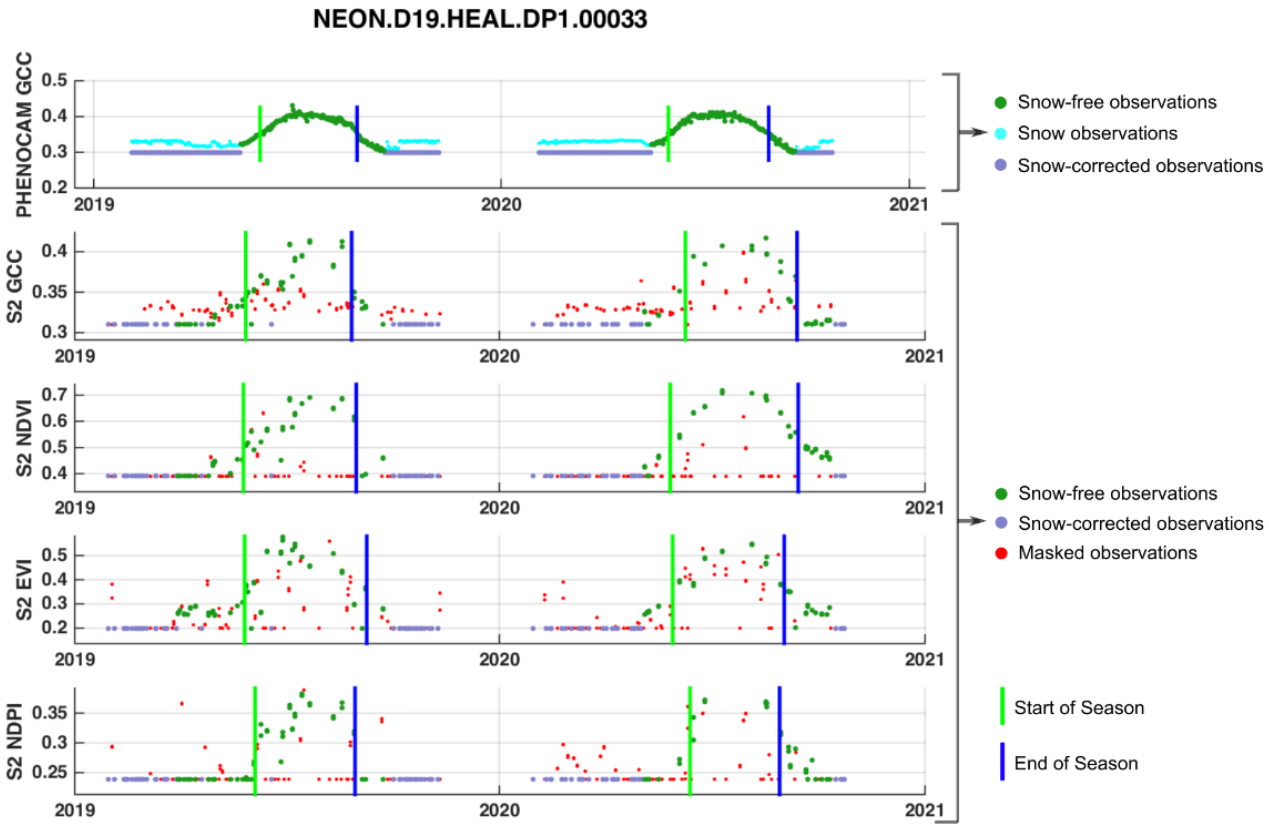
Supplementary figure 2.2. Histogram of Sentinel-2 green coordinate chromatic (GCC) index, normalized difference vegetation index (NDVI), and enhanced vegetation index (EVI) values observed after the last snow observation. The histogram represents the NDVI value of vegetation during the dormant period. This analysis only considers the time series with a gap that did not exceed 10 days.



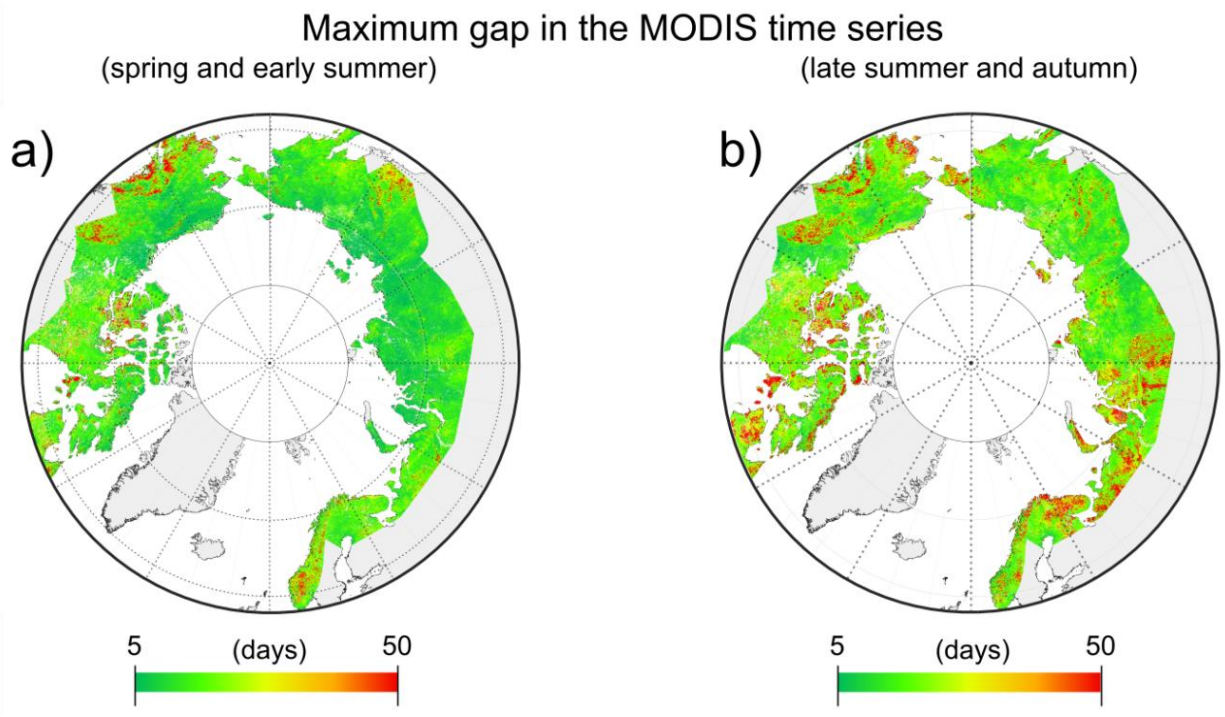
Supplementary figure 2.3. Time series of green coordinate chromatic (GCC) index of the PhenoCam site *imcrkrtussock* and four vegetation indices extracted for the site location from the Sentinel-2 level 2A for the years 2019 and 2020, and the associated land surface phenology metrics estimated with the threshold method. The vegetation indices were the GCC, normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), and normalized difference phenology index (NDPI).



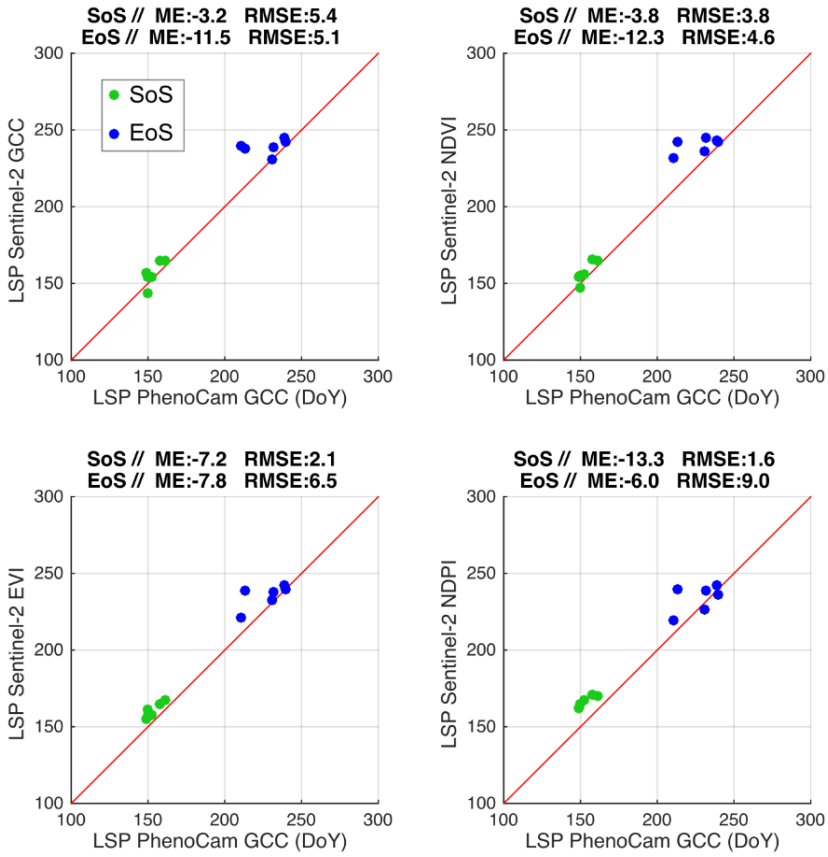
Supplementary figure 2.4. Time series of green coordinate chromatic (GCC) index of the PhenoCam site *NEON.D18.TOOL.DP1.00033* and four vegetation indices extracted for the site location from the Sentinel-2 level 2A for the years 2019 and 2020, and the associated land surface phenology metrics estimated with the threshold method. The vegetation indices were the GCC, normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), and normalized difference phenology index (NDPI).



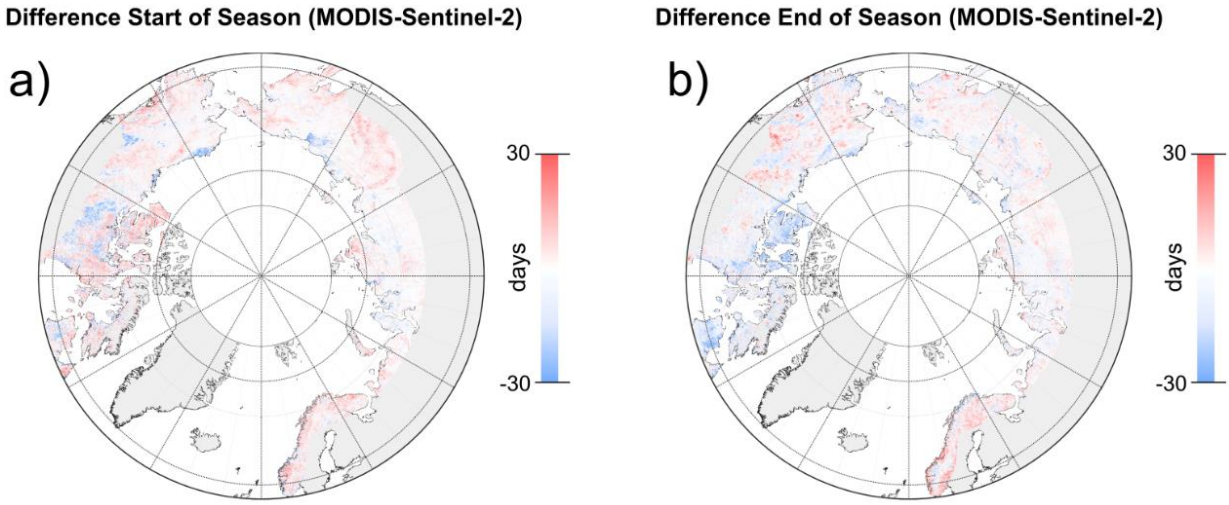
Supplementary figure 2.5. Time series of green coordinate chromatic (GCC) index of the PhenoCam site *NEON.D19.HEAL.DP1.00033* and four vegetation indices extracted for the site location from the Sentinel-2 level 2A for the years 2019 and 2020, and the associated land surface phenology metrics estimated with the threshold method. The vegetation indices were the GCC, normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), and normalized difference phenology index (NDPI).



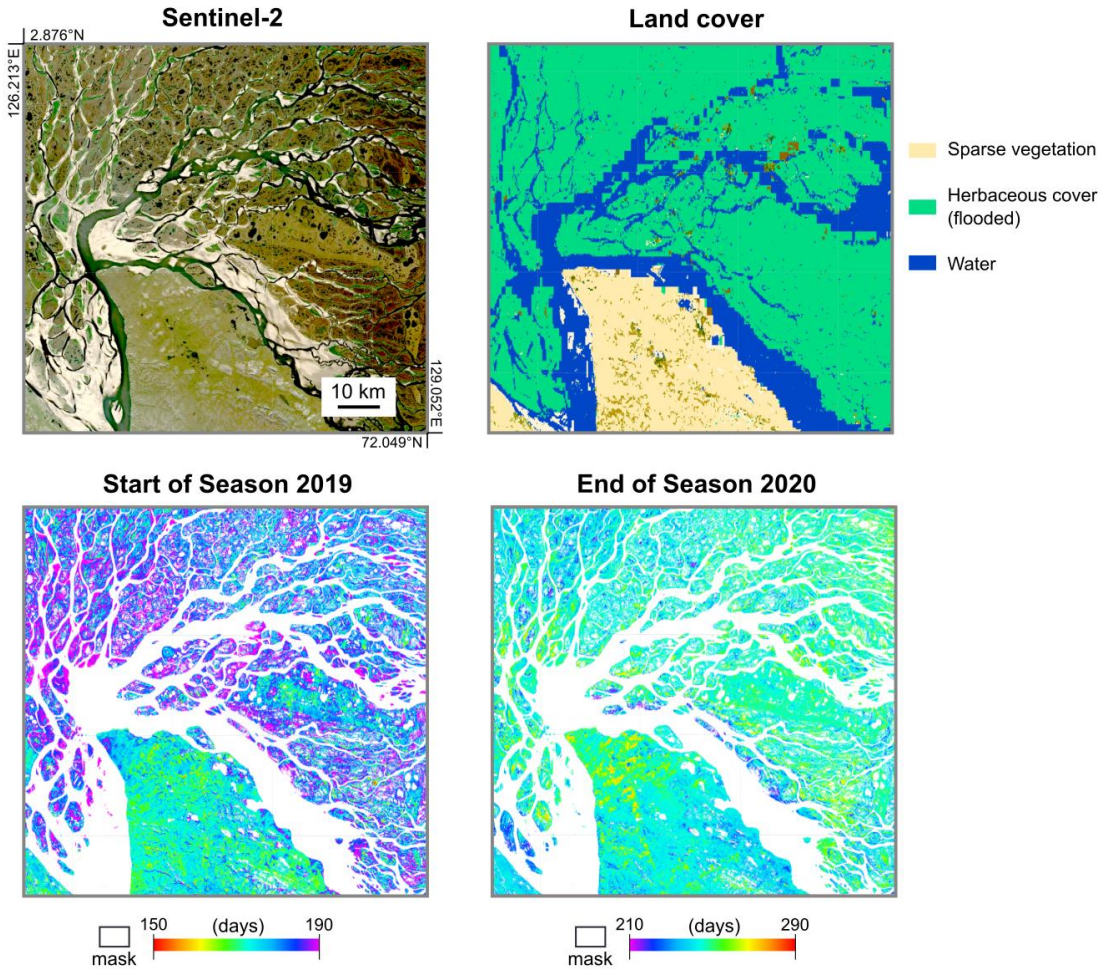
Supplementary figure 2.6. Maximum discontinuity in the surface spectral reflectance of Moderate Resolution Imaging Spectroradiometer (MODIS) (MOD09GAv6) time series after cloud masking for (a) spring and early summer (1 May to 15 July) and (b) late summer and autumn (15 July to 30 September) 2019.



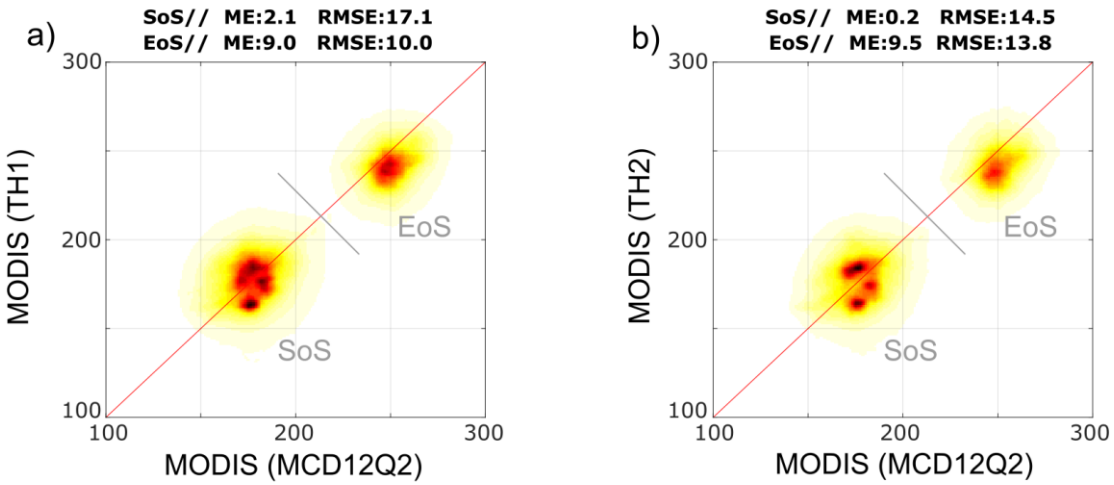
Supplementary figure 2.7. Comparison of start of season (SoS) and end of season (EoS) dates between PhenoCam and four vegetation indices estimated with Sentinel-2. The comparison was performed for three PhenoCam sites in the tundra biome for the years 2019 and 2020. The vegetation indices were the green chromatic coordinate (GCC) in PhenoCam and the GCC, normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), and normalized difference phenology index (NDPI) for Sentinel-2. The phenology metrics were extracted with a 50% threshold method after time series smoothing. The bias between PhenoCam and Sentinel-2 is reported with the mean error (ME) and the accuracy with the root mean squared error (RMSE).



Supplementary figure 2.8. Difference between start of season (a) and end of season (EoS) (b) phenology metrics extracted with the 50% threshold method applied to the non-smoothed Sentinel-2 and MODIS enhanced vegetation index (EVI) time series for the Arctic in 2019. The phenology metrics estimated with Sentinel-2 were resized to the spatial resolution of MODIS (500 m).



Supplementary figure 2.9. Differences in the vegetation dynamics in the delta of the Lena River. The Sentinel-2 image is a true composition (R: Red, G: Green, and B: blue) taken on 3 August 2019. The land cover is the GlobCover 2009.Changes in the land cover dynamics, primarily in the spring growth onset, depended on the vegetation type; greening occurs in the shrublands in the mainland earlier than the herbaceous cover in the delta.



Supplementary figure 2.10. Comparison between the phenology metrics extracted with the 50% threshold method applied to the non-smoothed (a) and smoothed (b) enhanced vegetation index (EVI) MODIS time series derived from MOD09GAv6 surface reflectance product and the bands 'MidGreenup_1' and 'MidGreendown_1' in the MCD12Q2v6 phenology product for 2018.

Supplementary Material

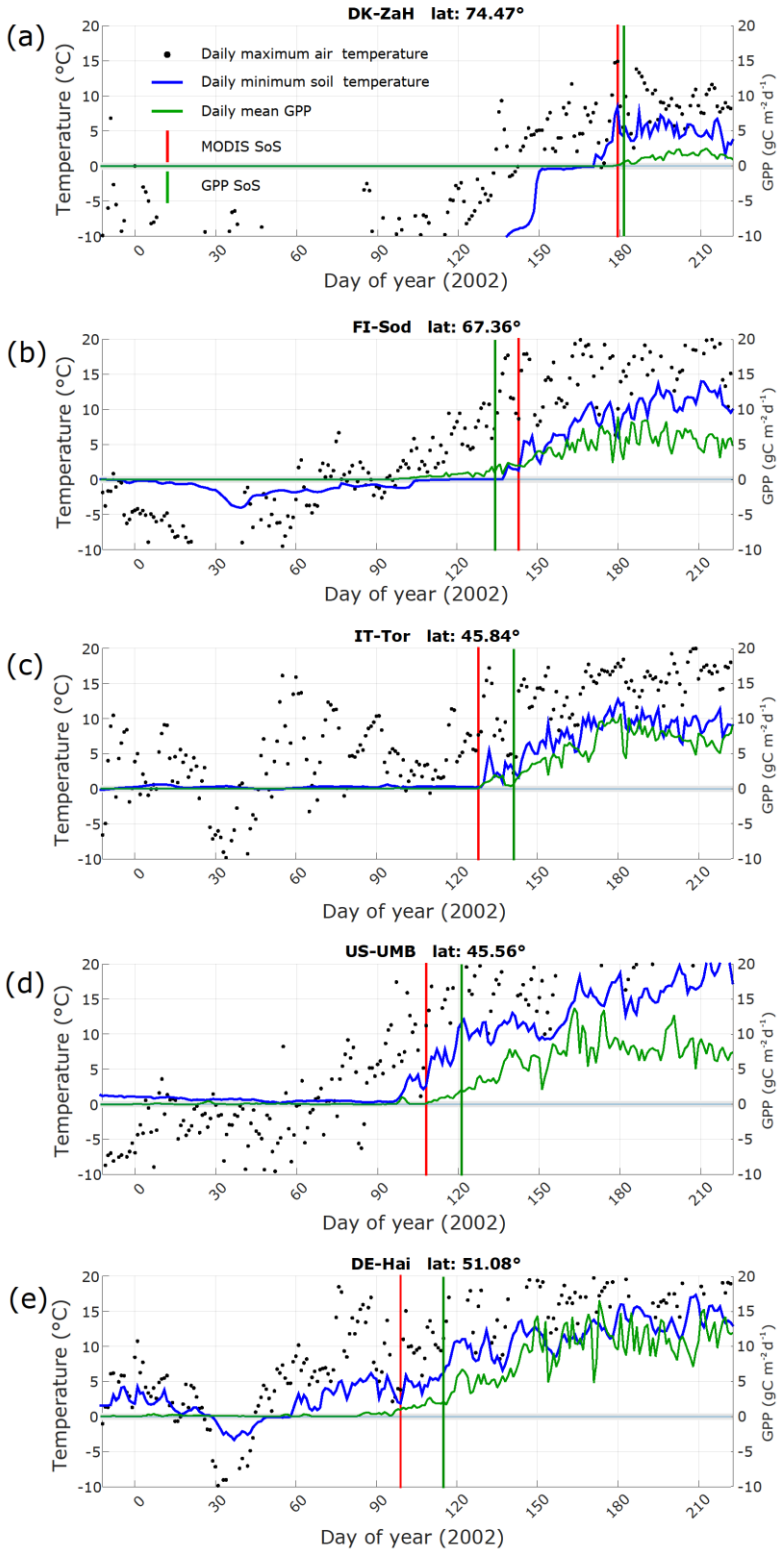
Chapter 3

Soil thawing regulates the spring growth onset in tundra and alpine biomes

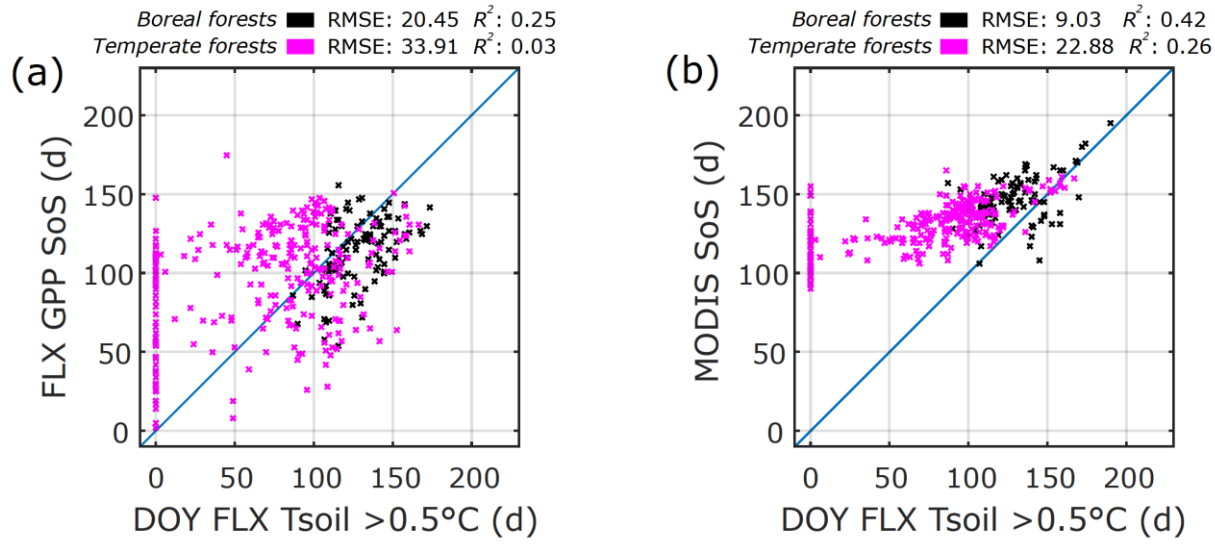
Supplementary Table 3.1. Information for the 78 FLUXNET sites.

Site ID	IGBP	Latitude (°)	Longitude (°)
CA-Man	ENF[1]	55.8796	-98.4808
CA-NS1	ENF	55.8792	-98.4839
CA-NS2	ENF	55.9058	-98.5247
CA-NS3	ENF	55.9117	-98.3822
CA-NS4	ENF	55.9144	-98.3806
CA-NS5	ENF	55.8631	-98.485
CA-Oas	DBF[2]	53.6289	-106.1978
CA-Obs	ENF	53.9872	-105.1178
CA-Qfo	ENF	49.6925	-74.3421
CA-SF1	ENF	54.485	-105.8176
CA-SF2	ENF	54.2539	-105.8775
CA-TP1	ENF	42.6609	-80.5595
CA-TP2	ENF	42.7744	-80.4588
CA-TP3	ENF	42.7068	-80.3483
CA-TP4	ENF	42.7102	-80.3574
CA-TPD	DBF	42.6353	-80.5577
CH-Dav	ENF	46.8153	9.8559
CN-Ha2	WET[6]	37.6086	101.3269
CN-HaM	GRA[7]	37.37	101.18
CZ-BK1	ENF	49.5021	18.5369
DE-Hai	DBF	51.0792	10.453
DE-Lkb	ENF	49.0996	13.3047
DE-Lnf	DBF	51.3282	10.3678
DE-Obe	ENF	50.7867	13.7213
DE-Tha	ENF	50.9624	13.5652
DK-NuF	WET	64.1308	-51.3861
DK-Sor	DBF	55.4859	11.6446
DK-ZaF	WET	74.4814	-20.5545
DK-ZaH	GRA	74.4733	-20.5503
FI-Hyy	ENF	61.8474	24.2948
FI-Let	ENF	60.6418	23.9595
FI-Sod	ENF	67.3624	26.6386
FR-LBr	ENF	44.7171	-0.7693
IT-Col	DBF	41.8494	13.5881
IT-Isp	DBF	45.8126	8.6336
IT-La2	ENF	45.9542	11.2853
IT-Lav	ENF	45.9562	11.2813
IT-MBo	GRA	46.0147	11.0458
IT-Ren	ENF	46.5869	11.4337
IT-Ro1	DBF	42.4081	11.93
IT-Ro2	DBF	42.3903	11.9209
IT-SR2	ENF	43.732	10.291
IT-SRo	ENF	43.7279	10.2844
IT-Tor	GRA	45.8444	7.5781
JP-MBF	DBF	44.3869	142.3186
NL-Loo	ENF	52.1666	5.7436
NO-Adv	WET	78.186	15.923
NO-Blv	SNO[3]	78.9216	11.8311
RU-Che	WET	68.613	161.3414
RU-Cok	OSH[4]	70.8291	147.4943
RU-Fyo	ENF	56.4615	32.9221
RU-Sam	GRA	72.3738	126.4958
RU-Tks	GRA	71.5943	128.8878
RU-Vrk	CSH[5]	67.0547	62.9405
SE-St1	WET	68.3542	19.0503
US-Atq	WET	70.4696	-157.4089
US-GLE	ENF	41.3665	-106.2399
US-Ha1	DBF	42.5378	-72.1715
US-Ivo	WET	68.4865	-155.7503
US-MMS	DBF	39.3232	-86.4131
US-Me1	ENF	44.5794	-121.5
US-Me2	ENF	44.4523	-121.5574
US-Me3	ENF	44.3154	-121.6078
US-Me4	ENF	44.4992	-121.6224
US-Me5	ENF	44.4372	-121.5668
US-Me6	ENF	44.3233	-121.6078
US-NR1	ENF	40.0329	-105.5464
US-Oho	DBF	41.5545	-83.8438
US-Prr	ENF	65.1237	-147.4876
US-UMB	DBF	45.5598	-84.7138
US-UMd	DBF	45.5625	-84.6975
US-WCr	DBF	45.8059	-90.0799
US-Wi0	ENF	46.6188	-91.0814
US-Wi1	DBF	46.7305	-91.2329
US-Wi2	ENF	46.6869	-91.1528
US-Wi3	DBF	46.6347	-91.0987
US-Wi4	ENF	46.7393	-91.1663
US-Wi8	DBF	46.7223	-91.2524

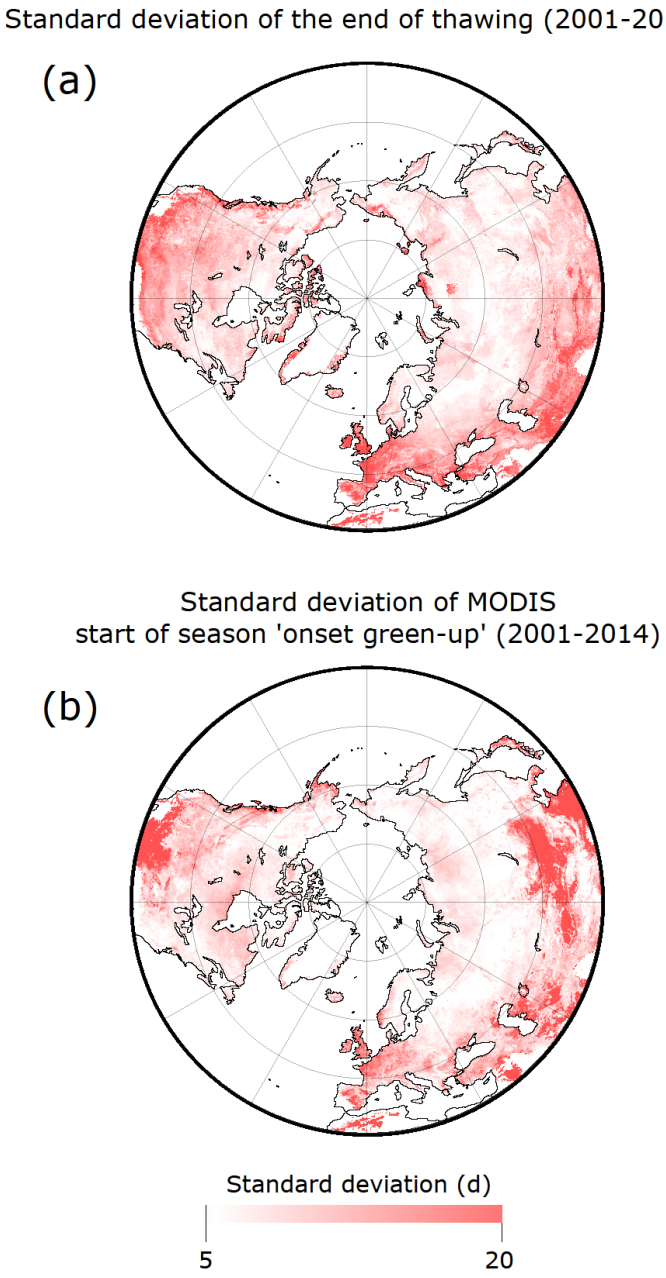
[1] Evergreen needleleaf forests
[2] Deciduous broadleaf forests
[3] Snow and ice
[4] Open shrublands
[5] Closed shrublands
[6] Permanent wetlands
[7] Grasslands



Supplementary Figure 3.1. One year of measurements at five FLUXNET sites representative of (a) tundra, (b) boreal forest, (c) alpine grassland, and (d) and (e) temperate forests with seasonally frozen soil. Soil temperature was measured at a depth of 2 cm, GPP SoS was estimated with the GPP time series of the flux towers, and MODIS SoS is the 'Onset Greenness Increase' of the MCD12Q2 V5 Land Cover Dynamics product.



Supplementary Figure 3.2. Scatter plots of spring onset (gross primary productivity (GPP) start of season (SoS) (a) and MODIS SoS ('MidGreenup1' band in MCD12Q2v6 product) (b)) and the last day before the growing season using soil temperature >0.5 °C. Point colors represent boreal forests (black) and temperate forests (magenta) for both deciduous and needleaved forests.



Supplementary Figure 3.3. Maps of standard deviation of end of thawing (EoT) (a) and standard deviation of MODIS start of season (SoS) (b) for the period 2001-2014. EoT was estimated using the GLDAS dataset as the last day of the dormant period with soil temperatures <0.5 °C. MODIS SoS is the 'Onset Greenness Increase' layer in the MCD12Q2 V5 Land Cover Dynamics product. We resized MODIS SoS to the spatial resolution of GLDAS. The map projection is the Lambert Azimuthal Equal-Area.

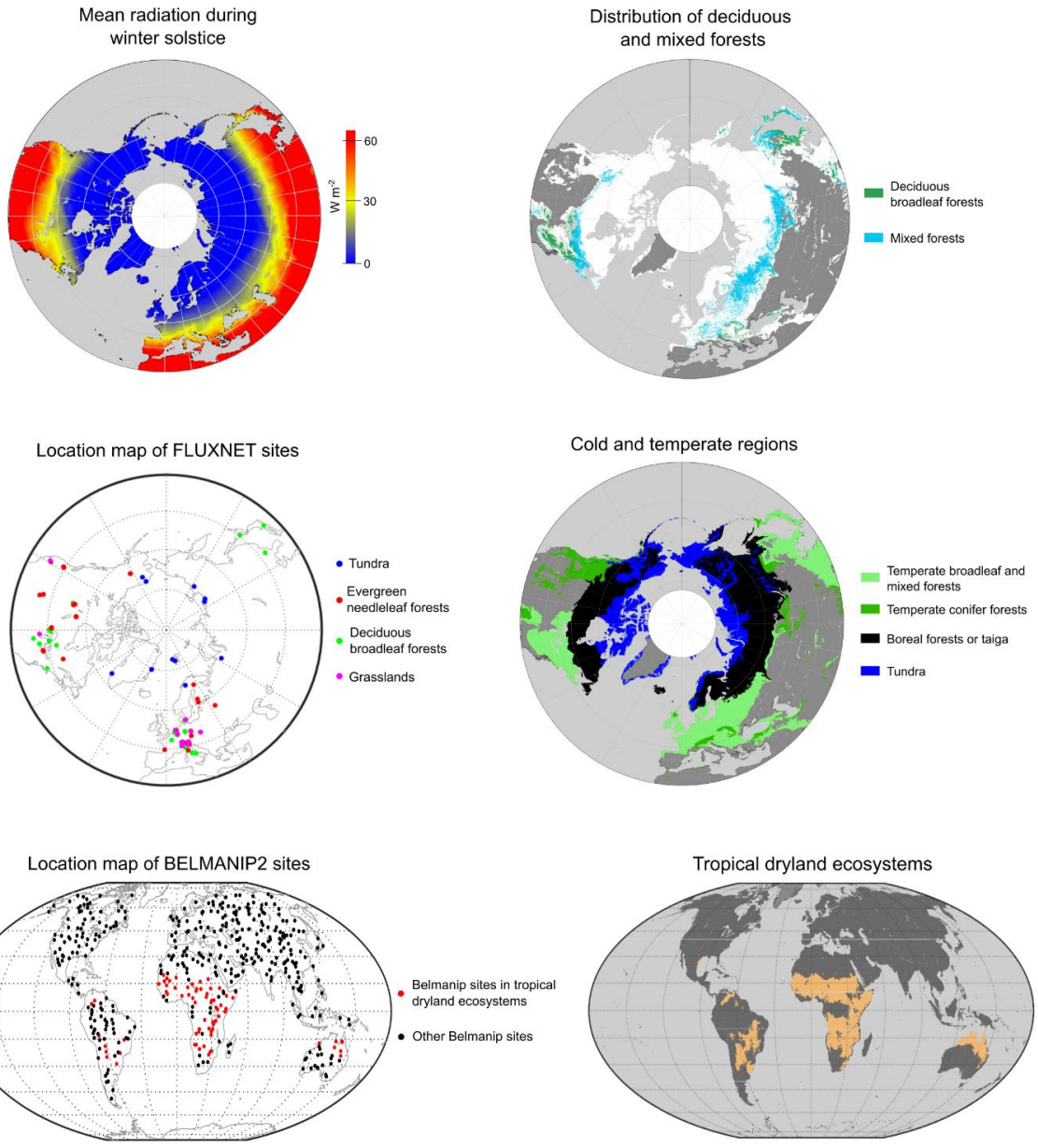
Supplementary Material

Chapter 4

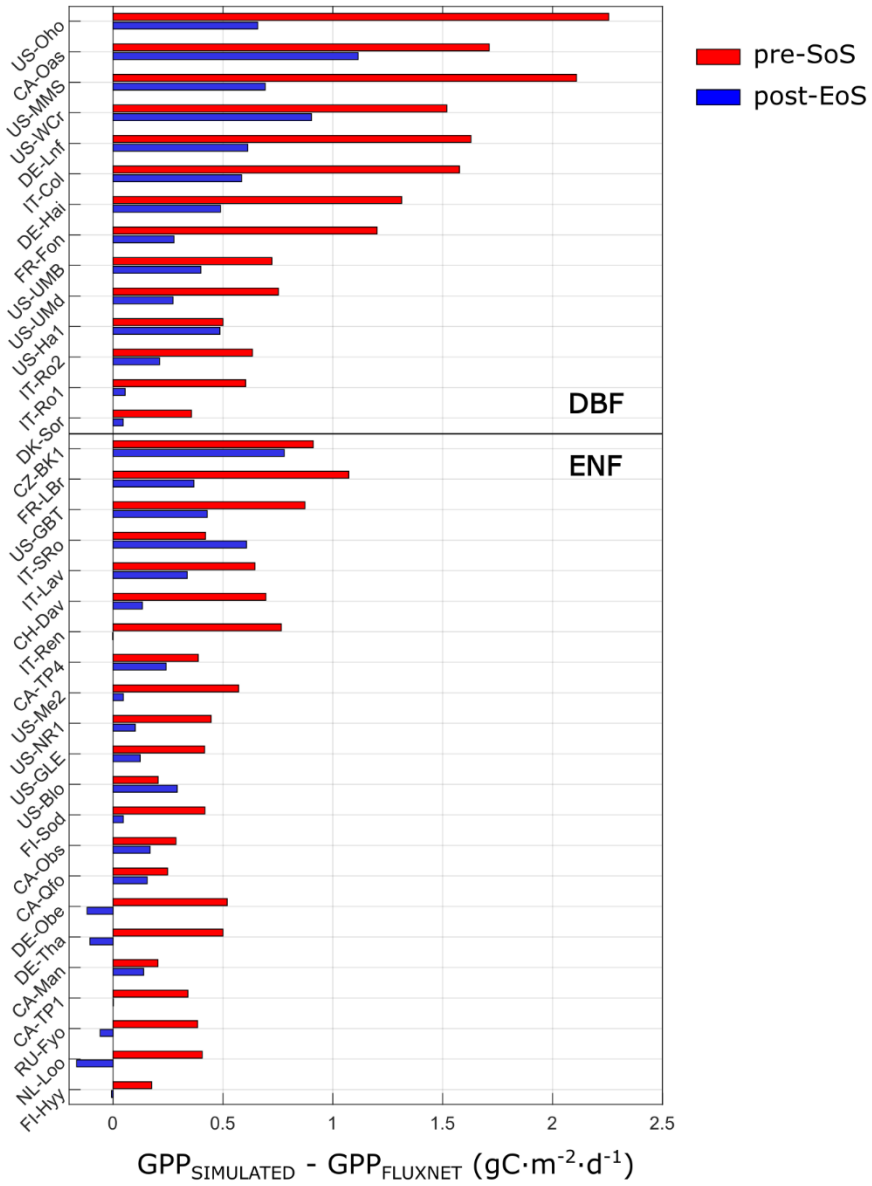
Climate-constrained boundaries cause nonuniform responses of vegetation phenology to climatic warming

Supplementary Table 4.1. Parameters of A, B, and C of the constraint boundary for different GPP thresholds, and overall accuracy (OA) obtained by binary classification (dormant versus growing stage) from climatic time series in the FLUXNET sites for different vegetation types; tundra, evergreen needleaved forests (ENF), deciduous broadleaved forests (DBF), mixed forests (MX), and grasslands (GRA). Values between parenthesis in A and B represent the 95% confidence interval.

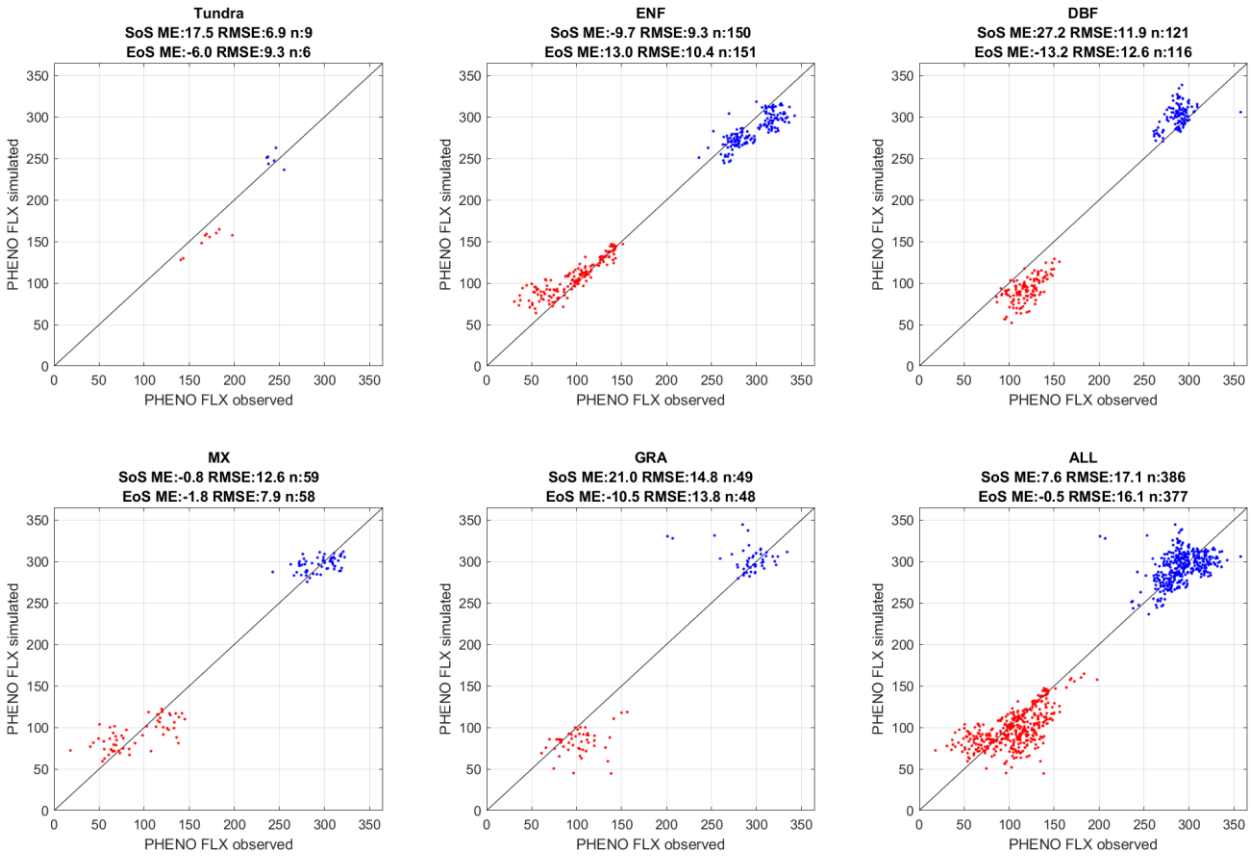
GPP threshold	A	B	C	OA (%) (all)	OA (%) Tundra	OA (%) ENF	OA (%) DBF	OA (%) MX	OA (%) GRA
1	1.2 (0.0, 2.1)	25 (1, 41)	150	85.54	85.56	89.66	84.16	86.34	79.55
1.5	2.7 (1.0, 3.6)	28 (7, 43)	150	84.99	85.17	88.77	85.05	86.41	79.19
2	2.5 (2.0, 3.4)	27 (11, 54)	350	84.74	85.79	87.89	85.98	87.35	78.51
2.5	4.9 (3.2, 5.6)	27 (5, 46)	350	84.21	85.87	85.84	88.01	87.42	77.94
3	5.0 (4.2, 7.3)	42 (22, 59)	350	83.97	85.75	85.86	87.85	87.34	77.55



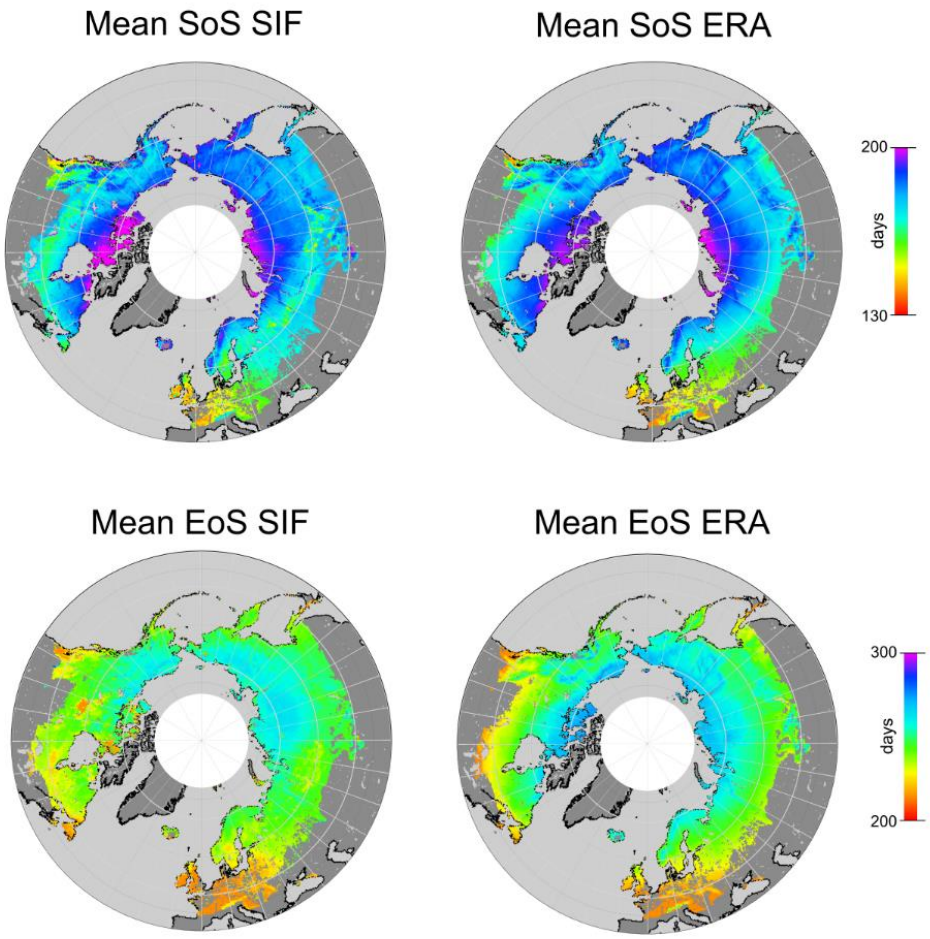
Supplementary figure 4.1. Mean incoming shortwave radiation during the winter solstice, estimated as the mean radiation from the Day of Year 345 to 365 and aggregated over the 2000-2020 period, locations of the FLUXNET sites used in the study, and distribution of deciduous broadleaved forests and mixed forests extracted from the MODIS Land Cover product, and biomes of the temperate and cold regions of the Northern Hemisphere extracted from the RESOLVE ECOREGIONS dataset.



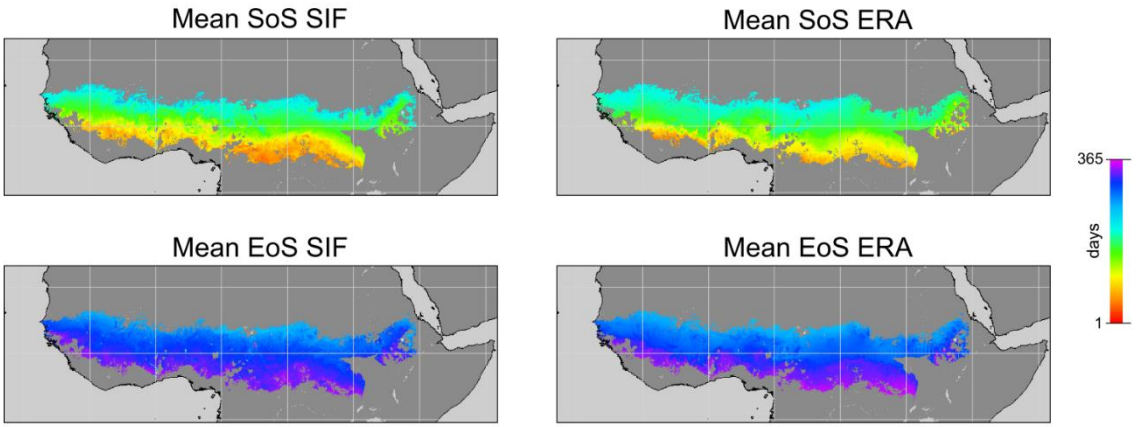
Supplementary figure 4.2. Mean difference between simulated GPP and observed GPP during pre-SoS (15 days before the start of season) and post-EoS (15 days after the end of season) at the FLUXNET sites classified as deciduous broadleaved forests (DBF) and evergreen needleaved forests (ENF) in temperate and cold regions of the Northern Hemisphere.



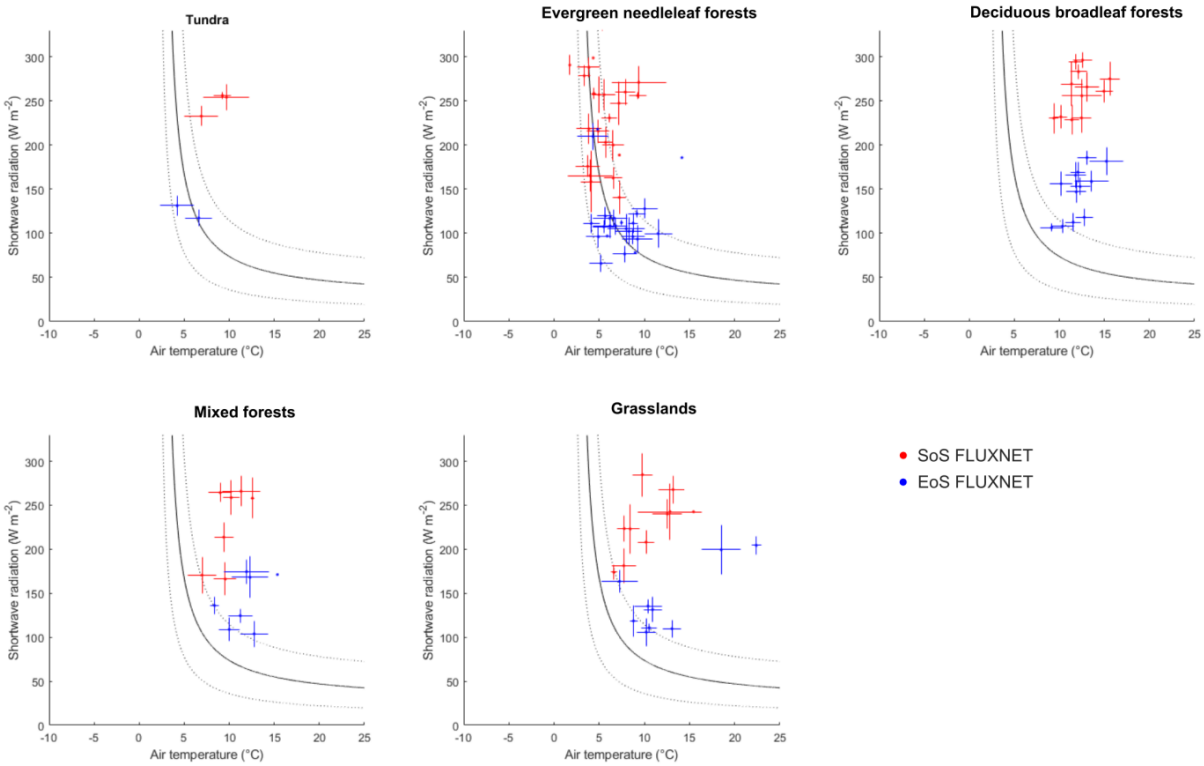
Supplementary figure 4.3. Comparison between observed and simulated start of the growing season (SoS) (in red) and end of the growing season (EoS) (in blue) at the FLUXNET towers in temperate and cold regions of the Northern Hemisphere. Observed SoS and EoS was obtained from GPP time series from the FLUXNET towers. Simulated SoS and EoS was obtained from ERA5-Land temperature and radiation time series. We reported the mean error (ME), the root mean squared error (RMSE), and the number of observations (n).



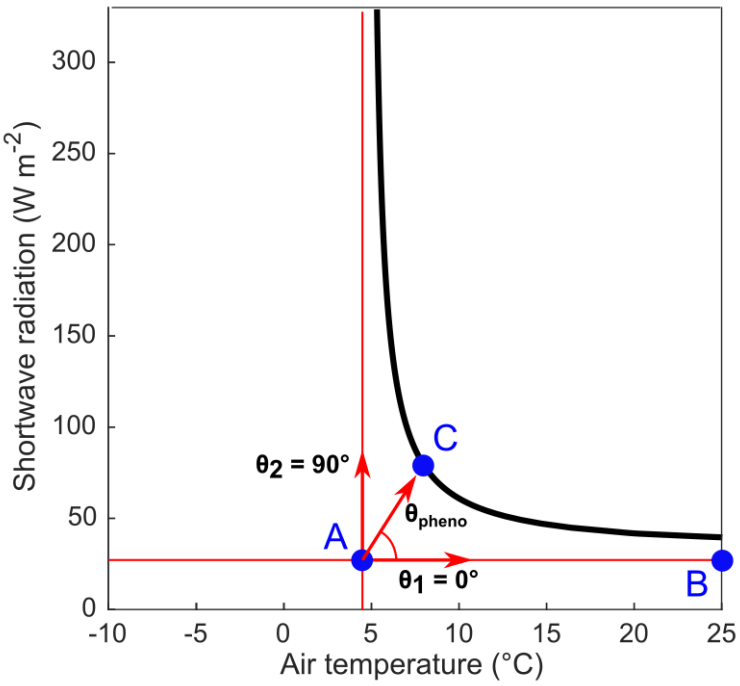
Supplementary figure 4.4. Comparison of the mean start of the growing season (SoS) and end of the growing season (EoS) extracted from the OCO-2 SIF time series and SoS and EoS estimated using the air temperature and amount of shortwave radiation from the ERA5-Land records during 2001-2020 in temperate and cold regions of the Northern Hemisphere.



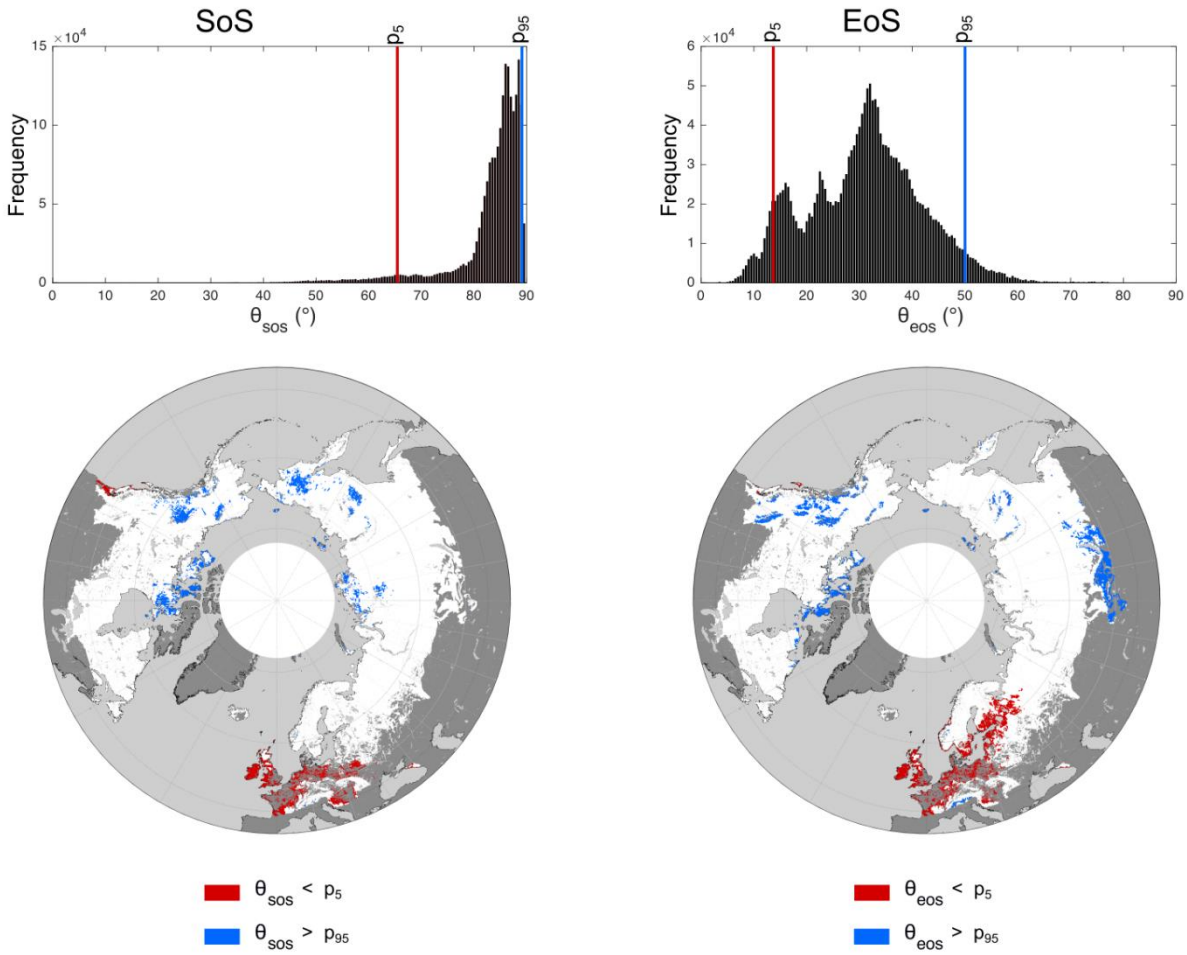
Supplementary figure 4.5. Comparison of the mean start of the growing season (SoS) and end of the growing season (EoS) extracted from the OCO-2 SIF time series and SoS and EoS estimated using air temperature, shortwave radiation, and soil moisture from the ERA5-Land records during 2001-2020 in the Sahel.



Supplementary figure 4.6. Temperature and radiation at the start and end of carbon uptake at the FLUXNET sites for various vegetation types in temperate and cold regions of the Northern Hemisphere. Errors bars indicate two standard deviations.



Supplementary figure 4.7. Graphical representation of the estimation of θ. θ_1 and θ_2 represent the highest constraints on GPP by radiation and temperature, respectively. The radiation constraint during a phenological event (θ_{pheno}) is estimated with the points A and B, with fixed coordinates defined by the asymptotes (red lines) of a rational function (black line), and the point C with variable coordinates depending on the temperature and radiation during the phenological event.



Supplementary figure 4.8. Histogram of the climatic constraints, depicted by θ, during the start of the growing season (SoS) and the end of the growing season (EoS) for the temperate and cold regions of the Northern Hemisphere. High values of θ represent a high restriction of photosynthetic activity by temperature, and low values of θ represent a high restriction by radiation. The pixels in red depict the areas with the highest radiation constraint (< 5th percentile of θ (p5)) and the pixels in blue depict the areas with the highest temperature limitation on the phenological dates (> 95th percentile of θ (p95)).

Supplementary Material

Chapter 5

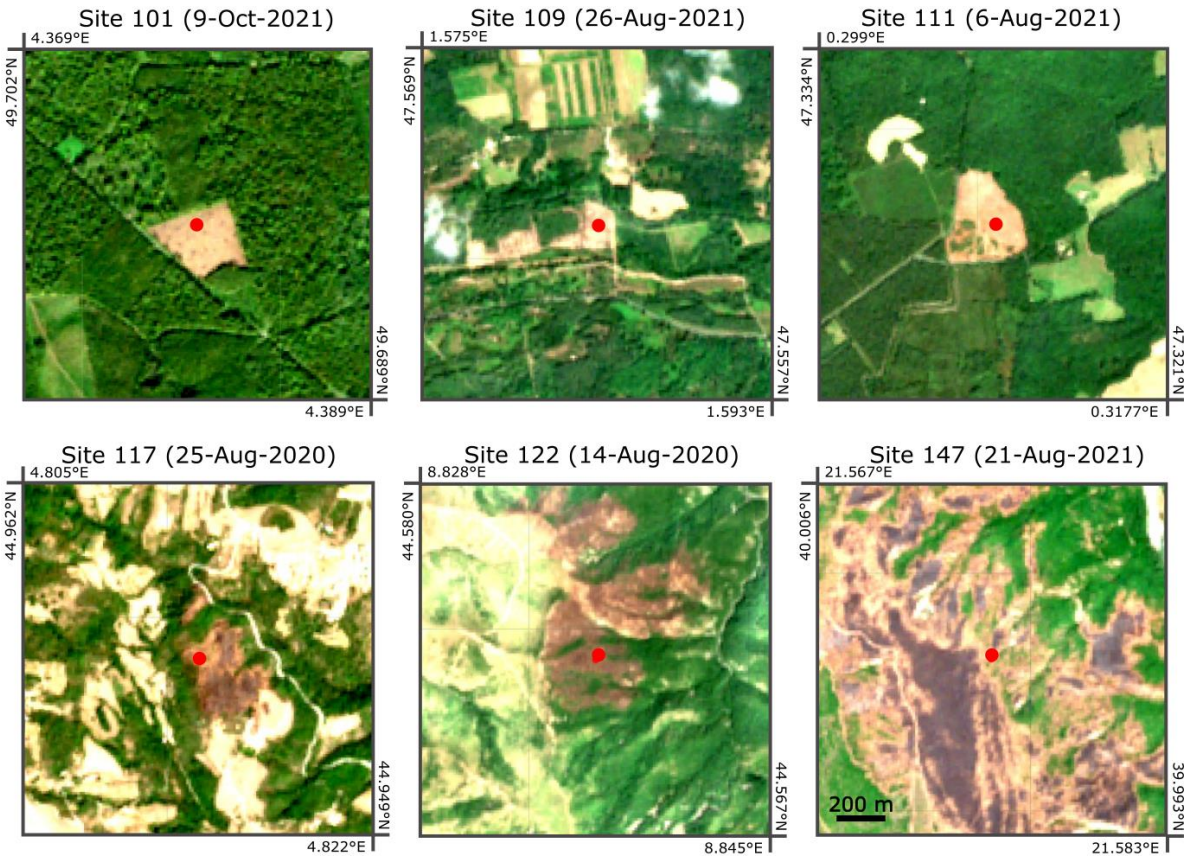
Sentinel-2 land-surface phenology indicates widespread drought-induced early leaf shedding and subsequent legacy decline of productivity in European temperate deciduous forests

Supplementary Table 5.1. List of the 90 randomly distributed points where early leaf shedding was detected with Sentinel-2 for 2017-2021.

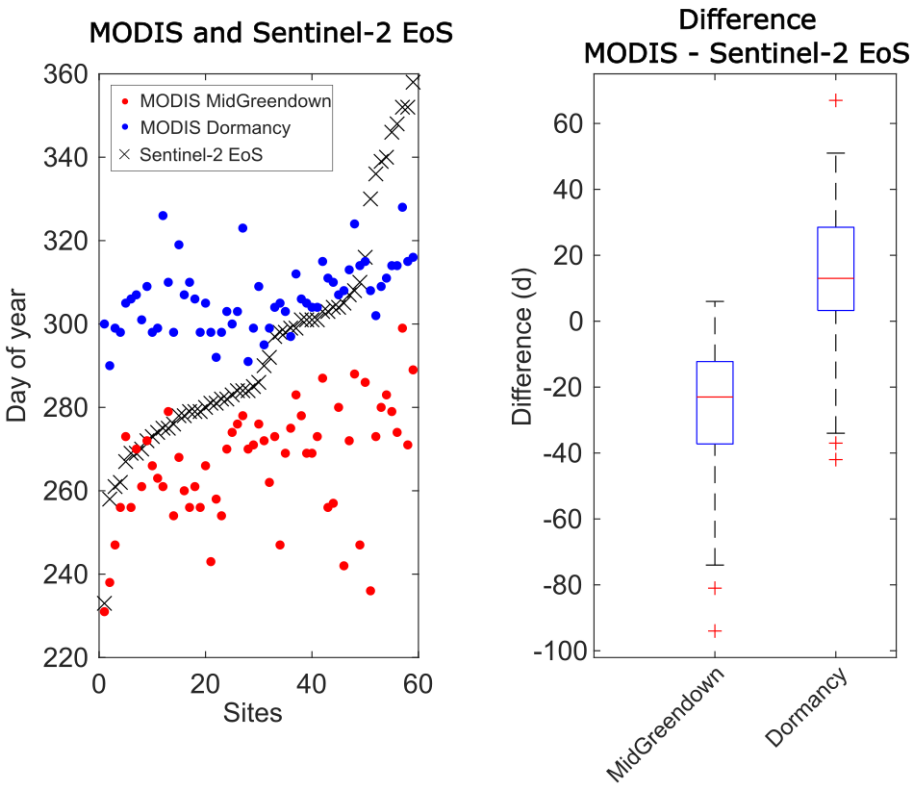
Site	Latitude (°N)	Longitude (°)	Year of early leaf shedding
1	51.5884	10.7581	2018
2	51.5877	12.3224	2018
3	50.9620	10.3089	2018
4	50.6364	14.0646	2018
5	50.4716	9.0572	2018
6	50.3485	13.0455	2018
7	50.3343	10.6697	2018
8	50.3008	5.3208	2018
9	50.2299	7.5697	2020
10	50.0257	8.7236	2018
11	49.9785	13.8007	2018
12	49.9705	12.2161	2018
13	49.7223	8.5197	2018
14	49.6157	14.1869	2018
15	48.9773	2.1278	2020
16	47.8527	7.4244	2018
17	47.6473	7.4800	2018
18	47.4120	6.3112	2018
19	47.2932	22.2024	2019
20	47.0194	6.0057	2018
21	46.9266	21.9809	2019
22	46.7666	3.5761	2020
23	46.6947	3.1664	2020
24	46.6744	22.1416	2019
25	46.3974	19.1924	2019
26	46.3621	5.4843	2018
27	46.2886	22.1328	2019
28	46.1408	21.6416	2019
29	46.1030	26.3210	2019
30	45.9805	2.8621	2020
31	45.9517	27.8966	2019
32	45.9042	17.8029	2019
33	45.8832	22.4997	2019
34	45.8708	5.2992	2020
35	45.7571	28.2998	2019
36	45.4741	15.7036	2019
37	45.4660	21.6428	2019
38	45.3205	16.6713	2019
39	45.3033	16.9093	2019
40	45.1663	17.3987	2018
41	45.1338	21.3708	2019
42	44.9974	4.5977	2020
43	44.9464	19.0650	2018
44	44.9208	28.5164	2020
45	44.8707	5.3461	2020
46	44.8300	25.8105	2019
47	44.7648	19.9258	2018
48	44.6907	34.0465	2018
49	44.6758	38.2729	2021
50	44.6346	26.1390	2019
51	44.5737	23.4262	2019
52	44.5559	25.2572	2019
53	44.4794	26.3529	2018
54	44.4325	24.1928	2019
55	44.3662	25.0815	2019
56	44.3198	15.6831	2017
57	44.1616	21.7527	2019
58	44.1541	23.7302	2019
59	44.1385	27.8190	2019
60	43.9994	10.4104	2021
61	43.9550	21.6769	2019
62	43.8725	27.6829	2021
63	43.6960	27.7317	2021
64	43.6246	26.9566	2021
65	43.6006	16.7657	2017
66	43.4329	11.6335	2021
67	43.4260	25.7916	2020
68	43.3017	11.9891	2017
69	43.2290	12.2487	2017
70	43.0914	-0.4635	2018
71	42.9999	24.2759	2021
72	42.9738	26.8285	2019
73	42.9463	12.1702	2021
74	42.8685	12.2148	2021
75	42.8623	27.7090	2019
76	42.3447	27.3153	2019
77	42.3173	21.5422	2021
78	42.1342	2.2482	2017
79	41.7993	27.9238	2020
80	41.4131	34.5849	2018
81	41.3851	14.2311	2017
82	41.3579	32.2147	2018
83	41.1924	32.4993	2018
84	41.1708	23.0888	2021
85	41.0436	36.1558	2018
86	40.9326	29.5742	2019
87	40.6652	36.1244	2018
88	40.3432	29.5074	2018
89	40.0996	29.8976	2018
90	40.0756	28.6682	2018

Supplementary Table 5.2. List of the 60 randomly distributed points where false positives of early leaf shedding were detected with Sentinel-2 for 2017-2021. The column 'Comment' categorises the cause of the false positives; low availability of data in 2017 (2017), presence of clear-cutting ('Clear-cut'), land cover is not deciduous broadleaved forest (Not DBF), presence of wildfires (Wildfire), and other reasons (Other).

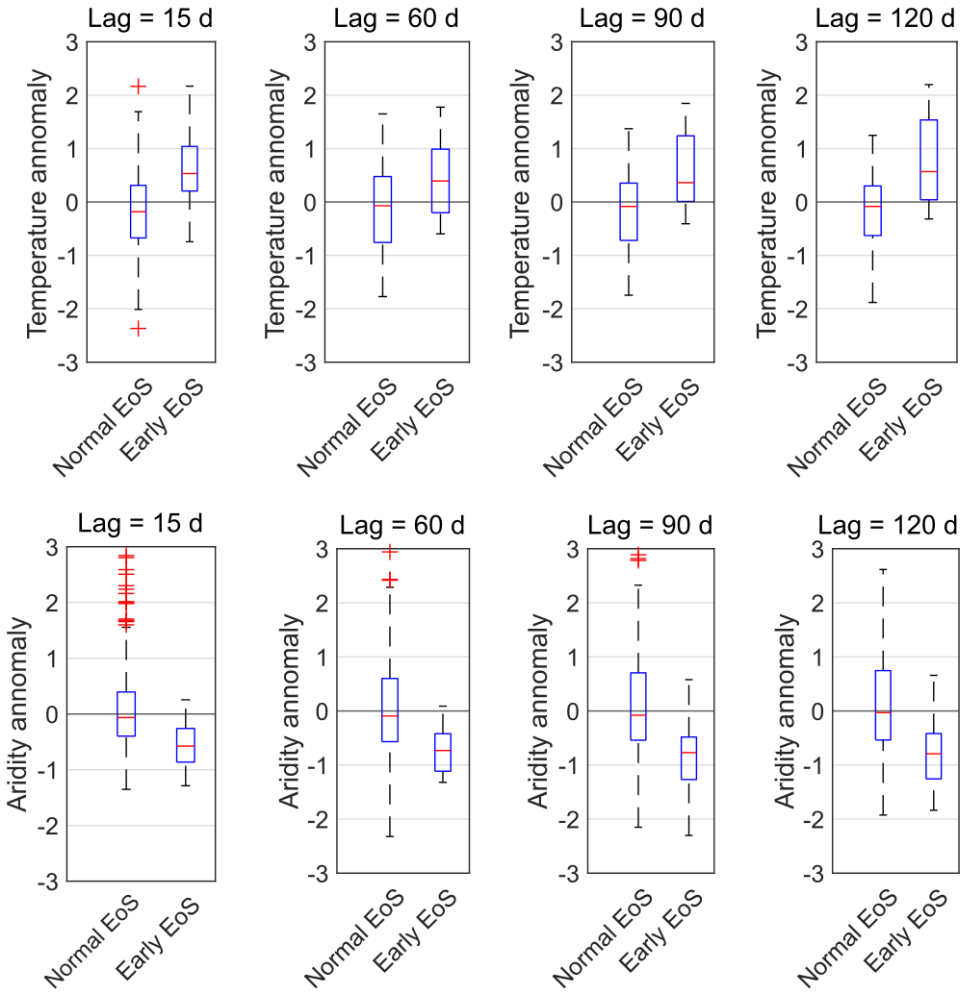
Site	Latitude (°N)	Longitude (°)	Year of detection	Comment
91	52.2266	9.3249	2017	2017
92	52.0932	12.3139	2017	2017
93	51.8630	12.1392	2017	2017
94	51.6949	9.0444	2017	2017
95	51.5550	8.9713	2017	2017
96	51.4324	9.3457	2017	2017
97	51.1987	8.6428	2017	2017
98	51.1564	9.0813	2017	2017
99	50.5972	13.5134	2017	2017
100	49.8991	6.9337	2017	2017
101	49.6955	4.3789	2021	Clear-cut
102	49.4756	13.5091	2017	2017
103	49.4529	2.9018	2017	2017
104	49.3265	12.8750	2017	Other
105	49.2748	7.8171	2017	Other
106	48.6121	5.6215	2020	Clear-cut
107	48.4334	7.3438	2017	2017
108	48.1914	4.8521	2020	Other
109	47.5630	1.5840	2021	Clear-cut
110	47.3978	4.9607	2020	Clear-cut
111	47.3280	0.3084	2021	Clear-cut
112	47.3191	26.2285	2017	2017
113	47.0771	24.2676	2018	Not DBF
114	46.4033	13.6250	2021	Other
115	46.0303	4.5338	2020	Not DBF
116	45.6843	15.9663	2019	Other
117	44.9556	4.8136	2020	Wildfire
118	44.8400	4.3134	2020	Not DBF
119	44.7497	0.8191	2017	2017
120	44.7042	9.4756	2017	2017
121	44.5904	21.1123	2019	Other
122	44.5737	8.8363	2020	Wildfire
123	44.5212	10.4102	2017	Wildfire
124	44.4344	11.3968	2021	Not DBF
125	44.3835	9.7173	2017	2017
126	44.3244	9.3869	2017	Wildfire
127	44.2741	25.7336	2019	Other
128	43.3443	17.1742	2017	Not DBF
129	43.3128	24.1000	2019	Not DBF
130	43.2511	11.1039	2018	Clear-cut
131	43.1036	21.8572	2019	Not DBF
132	42.9643	25.3288	2020	Other
133	42.9168	2.5136	2019	Not DBF
134	42.8441	18.9664	2021	Other
135	42.6159	13.7179	2021	Wildfire
136	42.6121	12.3443	2021	Not DBF
137	42.4682	9.3828	2021	Not DBF
138	42.2957	26.5618	2019	Not DBF
139	42.1534	12.0261	2021	Not DBF
140	41.9480	13.4390	2017	Wildfire
141	41.8651	9.3125	2020	Not DBF
142	41.3689	13.8233	2017	Wildfire
143	41.3597	32.8321	2018	Other
144	40.7879	21.3406	2019	Not DBF
145	40.7749	15.3810	2017	Clear-cut
146	40.7377	14.7326	2019	Wildfire
147	39.9993	21.5753	2021	Wildfire
148	39.8220	28.5393	2018	Not DBF
149	39.4118	16.3277	2017	Wildfire
150	39.0044	16.2175	2017	Wildfire



Supplementary Figure 5.1. Six validation points where Sentinel-2 detected an early end of the growing season (before 1 September). The upper images show a true-colour Sentinel-2 composition of clear-cutting, and the lower images show forests that were affected by wildfires.



Supplementary Figure 5.2. Comparison between the end of the growing season (EoS) obtained from MODIS and Sentinel-2 for 90 sites. The comparison was for the year before premature leaf shedding was detected. The left panel shows a representation of the Sentinel-2 EoS and the MODIS MidGreendown and dormancy metrics. Sites are numbered with the Sentinel-2 EoS in ascending order. We show only 69 sites, which are the sites where the MODIS MidGreendown and dormancy metrics were available. The right panel shows boxplots of the difference in distribution between the Sentinel-2 EoS and the MODIS metrics. Red line depicts the median, blue box represents the interquartile range, and red crosses outside the whiskers are outliers (>2 and <-2 standard deviations).



Supplementary Figure 5.3. Anomalies of temperature and aridity (P PET^{-1}) at the 90 sites where early leaf shedding was identified. Temperature and aridity were calculated for four intervals: 15, 30, 60, 90, and 120 d before early leaf shedding. The boxplots depict the distributions of temperature and aridity normalised for 2001-2021; mean = 0 and standard deviation = 1. 'Early EoS' represents the normalised temperature and aridity for the year of early leaf shedding detected during 2017-2021, and 'normal EoS' represents the normalised temperature and aridity for the other years during 2017-2021.

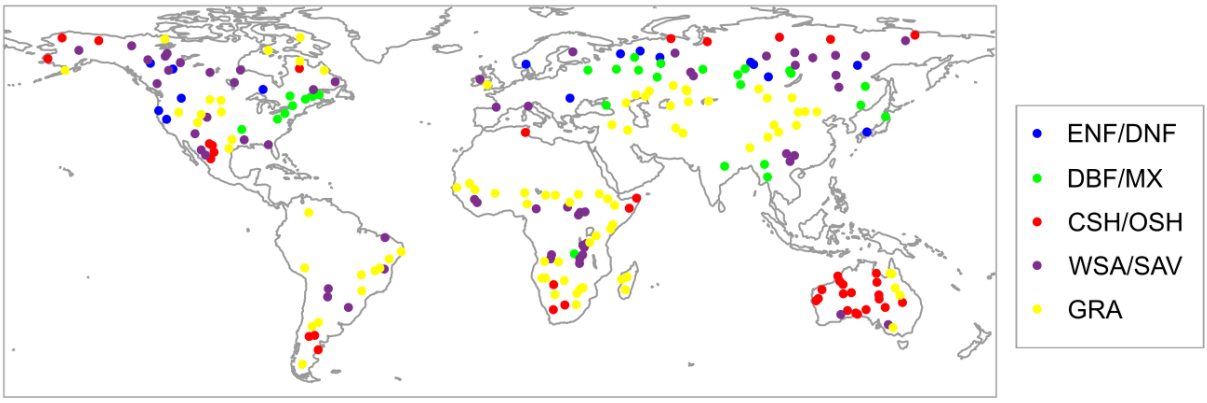
Supplementary Material

Chapter 6

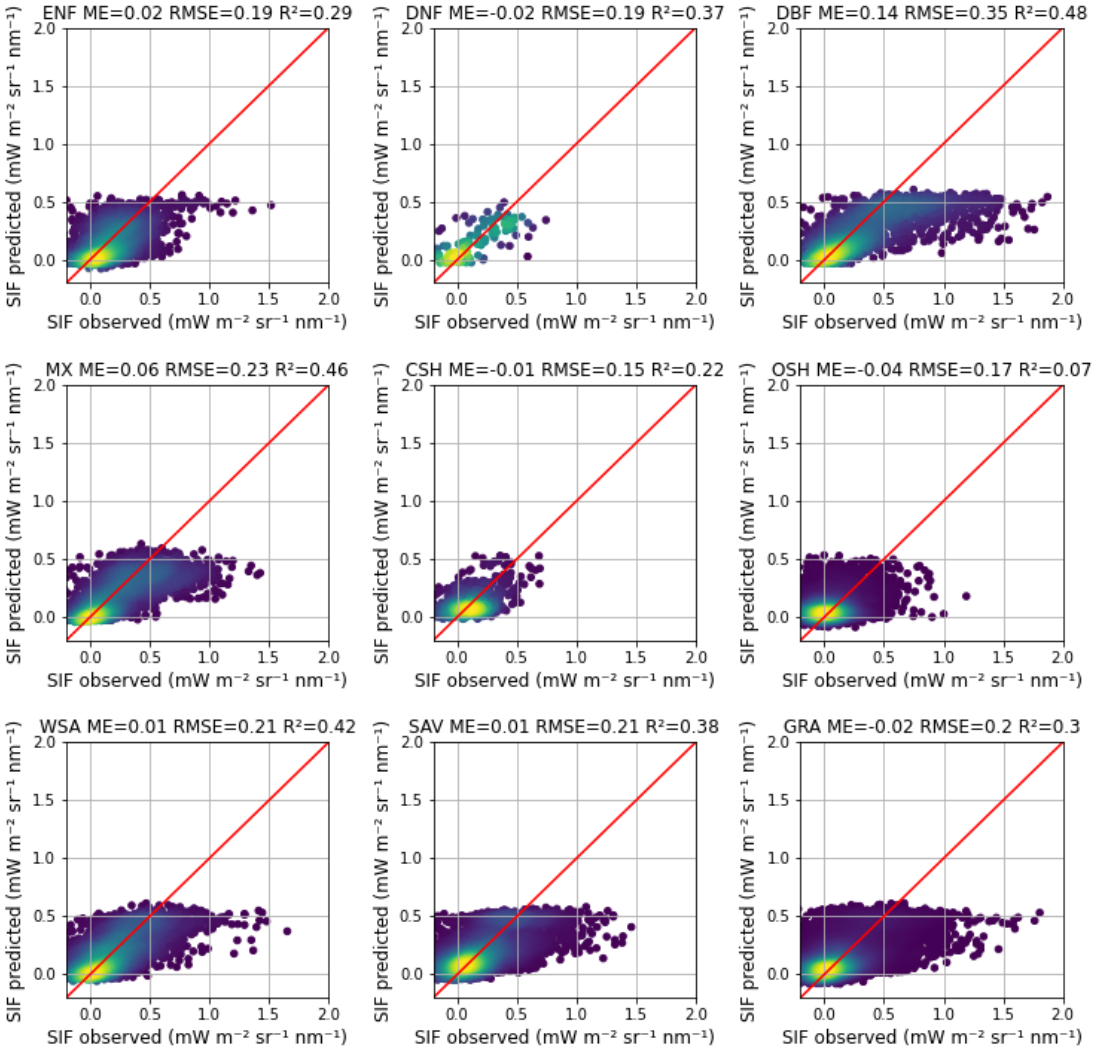
Local interpretation of machine learning models in remote sensing with SHAP: the case of global climate constraints on carbon uptake phenology

Supplementary Table 6.1. List of Gradient Boosting hyperparameters that were evaluated using a random search. The accuracy of the Gradient Boosting model was evaluated with 100 random combinations of these hyperparameters and the model with the lowest root-mean-squared error (RMSE) was selected. The combination of hyperparameters that lead to the lowest RMSE is included in the column 'Best parameter'.

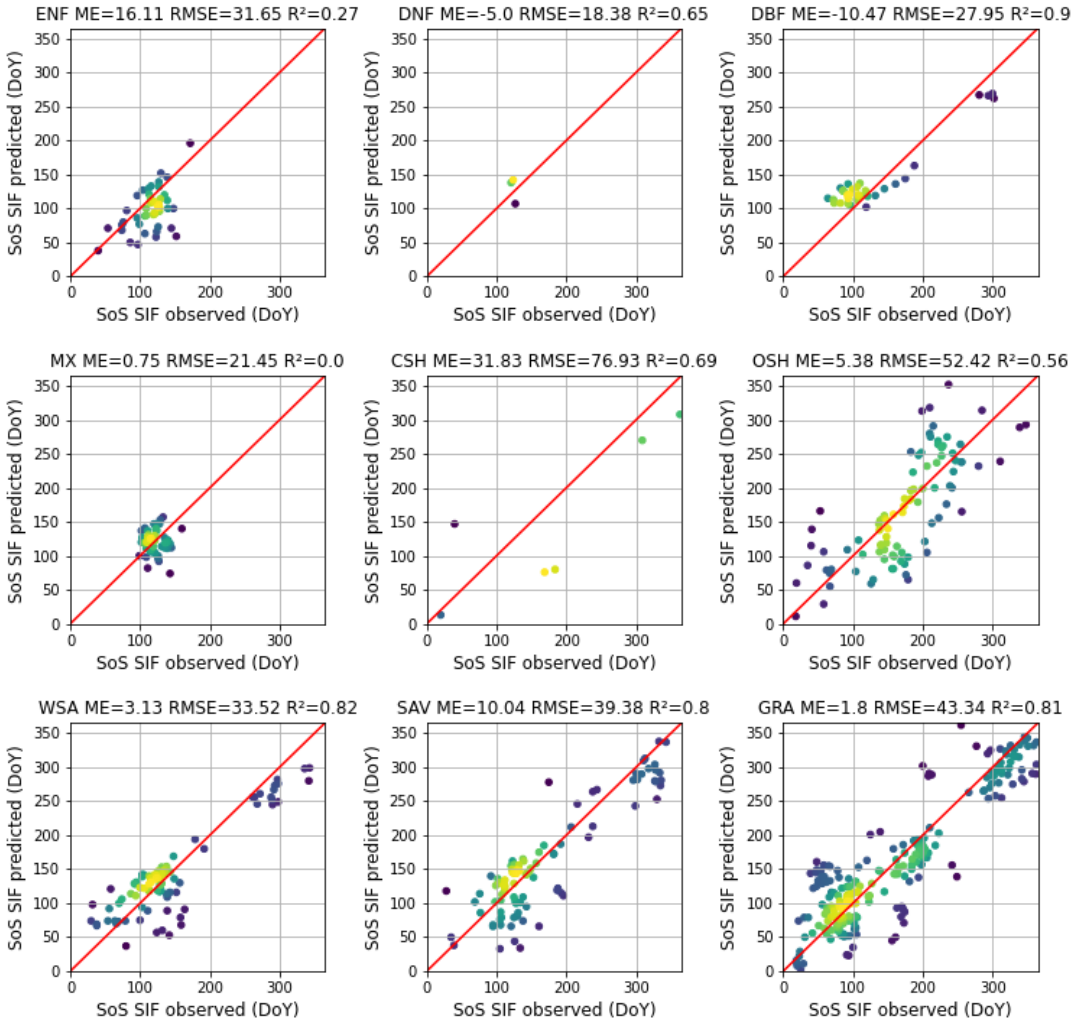
Hyperparameter	Values	Best parameter
max_depth	[3, 4, 5, 7, 10]	3
learning_rate	[0.005, 0.01, 0.05, 0.1]	0.05
subsample	[0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]	0.7
colsample_bytree	[0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]	0.8
colsample_bylevel	[0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8]	0.4
n_estimators	[200, 300, 400, 500, 600, 700, 800, 900]	300



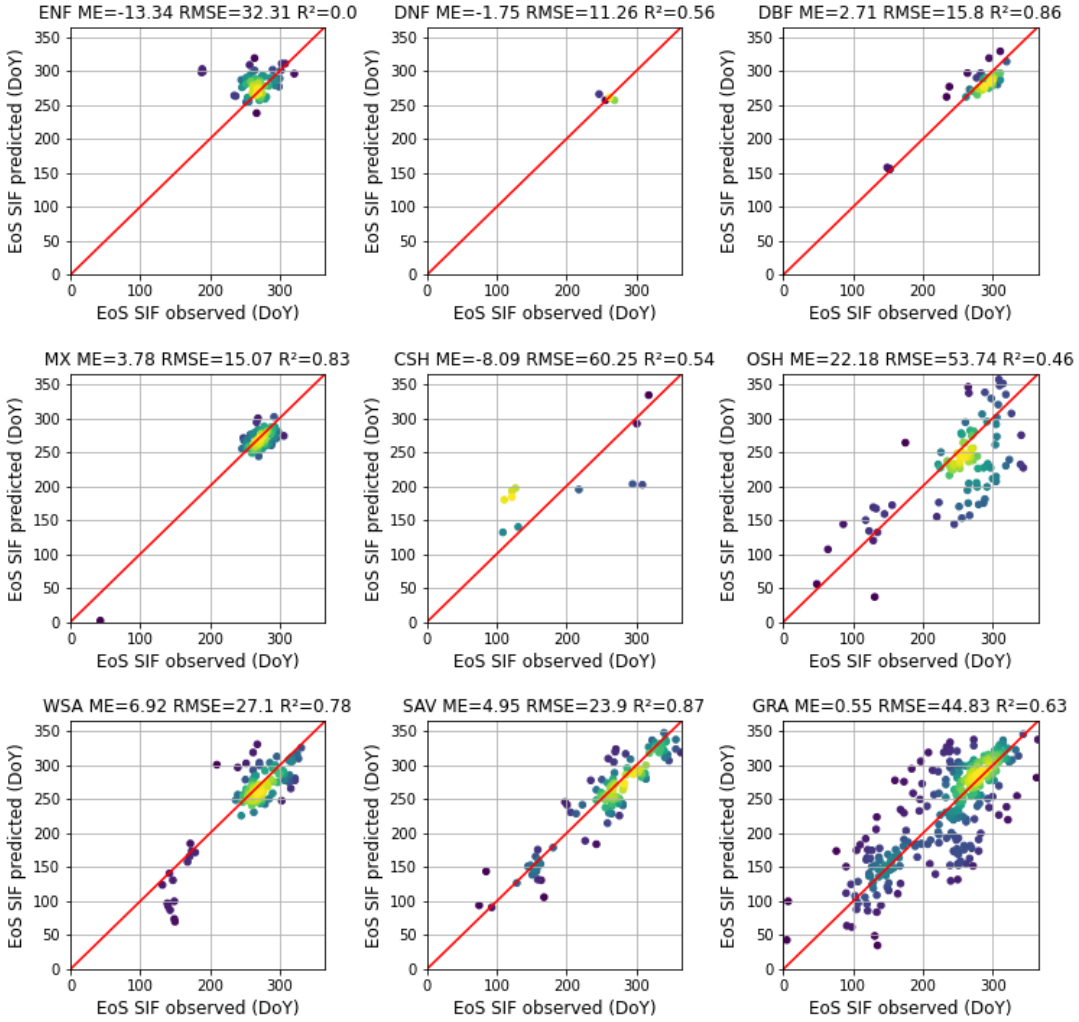
Supplementary figure 6.1. Location map of the 233 Belmanip2 sites. Colours depict vegetation cover types: evergreen needleleaf forests (ENF), deciduous needleleaf forests (DNF), deciduous broadleaf forests (DBF), mixed forests (MX), closed shrublands (CSH), open shrublands (OSH), woody savannah (WSA), savannah (SAV), and grasslands (GRA).



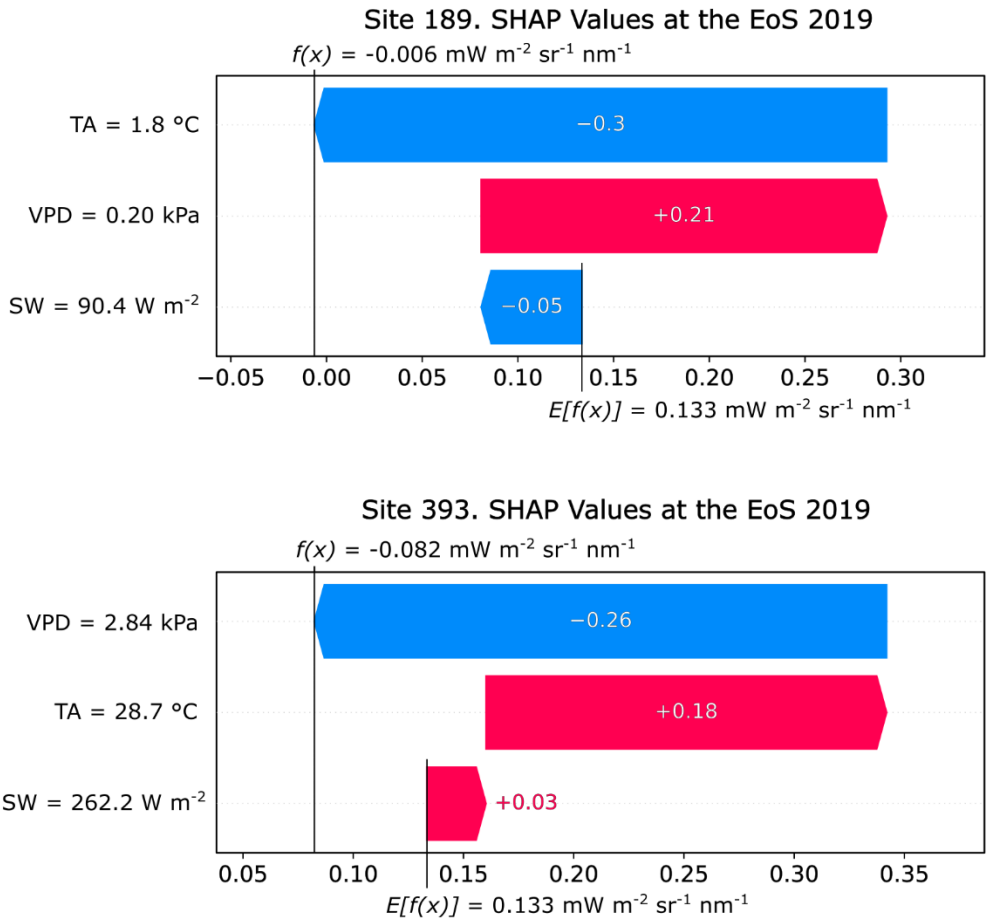
Supplementary figure 6.2. Comparison between observed and predicted sun-induced fluorescence (SIF) in 233 sites of the Belmanip2 network. The observed SIF was extracted from the TROPOSIF dataset. The predicted SIF was estimated from a machine learning model that fitted the observed SIF with three ERA5-Land variables: air temperature, shortwave radiation, and vapor-pressure-deficit. Each panel shows the comparison for a different land cover type: evergreen needleleaf forests (ENF), deciduous needleleaf forests (DNF), deciduous broadleaf forests (DBF), mixed forests (MX), closed shrublands (CSH), open shrublands (OSH), woody savannah (WSA), savannah (SAV), and grasslands (GRA). The tile of each panel reports the mean error (ME), root-mean-squared error (RMSE), and the coefficient of determination (R^2).



Supplementary figure 6.3. Comparison between the start of the growing season estimated from time series of observed and predicted sun-induced fluorescence (SIF) in 233 sites of the Belmanip2 network. The time series of observed SIF were extracted from the TROPOSIF dataset, and the time series of predicted SIF was estimated from weather-SIF model. The measurements indicate the Day of Year (DoY) at which the growing season begins for the years 2018 to 2021. The metrics were estimated using a threshold-based method: SoS was defined as the date for which SIF reaches a 20% of the annual amplitude before the maximum. Each panel shows the comparison for a different land cover type: evergreen needleleaf forests (ENF), deciduous needleleaf forests (DNF), deciduous broadleaf forests (DBF), mixed forests (MX), closed shrublands (CSH), open shrublands (OSH), woody savannah (WSA), savannah (SAV), and grasslands (GRA). The tile of each panel reports the mean error (ME), root-mean-squared error (RMSE), and the coefficient of determination (R^2).



Supplementary figure 6.4. Comparison between the end of the growing season estimated from time series of observed and predicted sun-induced fluorescence (SIF) in 233 sites of the Belmanip2 network. The time series of observed SIF were extracted from the TROPOSIF dataset, and the time series of predicted SIF was estimated from weather-SIF model. The measurements indicate the Day of Year (DoY) at which the growing season ends for the years 2018 to 2021. The metrics were estimated using a threshold-based method: EoS was defined as the date for which SIF reaches a 20% of the annual amplitude after the maximum. Each panel shows the comparison for a different land cover type: evergreen needleleaf forests (ENF), deciduous needleleaf forests (DNF), deciduous broadleaf forests (DBF), mixed forests (MX), closed shrublands (CSH), open shrublands (OSH), woody savannah (WSA), savannah (SAV), and grasslands (GRA). The tile of each panel reports the mean error (ME), root-mean-squared error (RMSE), and the coefficient of determination (R^2).



Supplementary figure 6.5. Waterfall plots depicting the contribution of three climate variables (air temperature (TA), shortwave radiation (SW), and vapor-pressure-deficit (VPD)) on the model output *f(x)* in one grassland (GRA) and one mixed forest (MX) sites of the BELMANIP2 network. The prediction is done for the end of the season (EoS) for year 2019. The output of the model is the prior expectation of the model *E[f(x)]* plus the summatory of SHAP values. Red arrows indicate a positive contribution and blue arrows indicate a negative contribution of the input variable to the predicted value.

Acknowledgments

Vull agrair a Josep Peñuelas i Aleixandre Verger per la direcció de la tesi, per introduir-me en el món de la recerca amb una visió aplicada a problemes globals, i a la vegada donar-me llibertat a l'hora de triar els temes de recerca. Agrair també a tots els coautors per col·laborar en els articles.

Als companys que ens relacionàvem al voltant del despatx C5b/048, per les converses interessants. A tots els companys del CREAF pel bon ambient.

Gràcies a bases de dades de Fluxnet i PhenoCam, i la NASA i l'ESA per la seua política de dades lliures. Però en especial, gràcies a Google Earth Engine per permetre'ns accedir i processar aquestes dades satel·litals amb poques linees de codi, per a qualsevol lloc del món, i des de qualsevol ordinador.

Al meu poble, per fer-me aprendre la perseverança que és necessària per fer una tesi.

I com no, a l'Anna i la Krufka. L'estima per part de la família ha sigut essencial durant la tesi.