Status of Phenological Research Using Sentinel-2 Data: A Review

Gourav Misra 1,2,3,*, Fiona Cawkwell 2,3 and Astrid Wingler 1,3

1 School of Biological, Earth and Environmental Sciences, University College Cork, T12K8AF Cork, Ireland; astrid.wingler@ucc.ie
2 Department of Geography, University College Cork, T12K8AF Cork, Ireland; f.cawkwell@ucc.ie
3 Environmental Research Institute, University College Cork, T12K8AF Cork, Ireland
* Correspondence: gourav.misra@ucc.ie

Received: 10 July 2020; Accepted: 24 August 2020; Published: 26 August 2020

Abstract: Remote sensing of plant phenology as an indicator of climate change and for mapping land cover has received significant scientific interest in the past two decades. The advancing of spring events, the lengthening of the growing season, the shifting of tree lines, the decreasing sensitivity to warming and the uniformity of spring across elevations are a few of the important indicators of trends in phenology. The Sentinel-2 satellite sensors launched in June 2015 (A) and March 2017 (B), with their high temporal frequency and spatial resolution for improved land mapping missions, have contributed significantly to knowledge on vegetation over the last three years. However, despite the additional red-edge and short wave infra-red (SWIR) bands available on the Sentinel-2 multispectral instruments, with improved vegetation species detection capabilities, there has been very little research on their efficacy to track vegetation cover and its phenology. For example, out of approximately every four papers that analyse normalised difference vegetation index (NDVI) or enhanced vegetation index (EVI) derived from Sentinel-2 imagery, only one mentions either SWIR or the red-edge bands. Despite the short duration that the Sentinel-2 platforms have been operational, they have proved their potential in a wide range of phenological studies of crops, forests, natural grasslands, and other vegetated areas, and in particular through fusion of the data with those from other sensors, e.g., Sentinel-1, Landsat and MODIS. This review paper discusses the current state of vegetation phenology studies based on the first five years of Sentinel-2, their advantages, limitations, and the scope for future developments.

Keywords: Sentinel-2; phenology; short wave infra-red; red-edge; time series

1. Introduction

Accurate monitoring of vegetation phenology across space and time is imperative for understanding the impact of climate change on plants and animals. Many studies have reported the influence of climate on phenological events, and monitoring of the timing of these events is therefore considered one of the important indicators to study the impacts of climate change [1,2]. Since different species react to changes in climate differently, shifts in the timing of phenological events have the ability to desynchronize species interactions [3]. Phenology also influences ecosystem productivity, carbon cycling, succession, and migration strategies [4–7]. In addition, it has been classified as one of the essential biodiversity variables (EBV) and addresses goals 13 (climate action) and 15 (life on land) of the Sustainable Development Goals (SDG), which makes the study of phenological patterns and drivers crucial to plan adaptation and mitigation strategies [8,9]. The advancing of spring events, the lengthening of the growing season, the shifting of tree lines, the decreasing sensitivity to warming and the uniformity of spring across elevations are a few of the important evidences of trends in plant phenology [2,10–14]. Such has been the interest in the field that a two-fold increase has been observed...
in the published literature on the topic of plant phenology in the last decade (based on Scopus database search results).

Traditional approaches to monitoring phenology, also referred to as ground phenology (GP), involve recurring visual observations of phenological events by trained personnel and can be expensive and time consuming. GP of individual plants is recorded in the form of day-of-year observations of breaking of leaf buds, expansion of leaves, appearance of flowers, leaf senescence etc. [15,16]. In addition to professional networks, many citizen science networks exist across the world that strive to provide high quality observations of phenology on a voluntary basis [17,18]. The latest developments in technology allow fixed digital cameras (phenocams) to capture several images of vegetation on a daily basis to monitor plant growth from a close range [19,20]. Although they work on the same principle as satellite observations of phenology, their spatial coverage is, however, limited, and simultaneous monitoring of vegetation phenology across large areas can only be achieved through satellites [21]. Due to the requirement of repeated observations to study land surface phenology (LSP), most satellite-based phenology studies have relied on medium to coarse spatial resolution satellite sensors, such as MODIS, Landsat and AVHRR, or on fusing MODIS and Landsat imagery to improve the temporal and spatial resolution [22–25]. In contrast to GP observations, the phenophases derived from satellite imagery or LSP represents the timings of the changes observed in the reflectance from the vegetated features aggregated at the pixel scale of the sensor [26,27]. The phenological information has been modelled (satellite reflectance modelled as function of time, temperature or relative humidity, etc.) and extracted (thresholds, derivatives and moving averages) from time series of satellite reflectance data in various ways, and the resulting phenological metrics are known to be very sensitive to the choice of algorithm [26,28–31]. Apart from using the original reflectance values, information from more than one spectral band is often combined to enhance the vegetation properties within a satellite pixel. A list of frequently cited vegetation indices and their algorithms can be found in [32–34]. LSP metrics are typically estimated as the day of the year corresponding to the start of season (SOS), end of season (EOS) and peak of season (POS), when particular values of thresholds (or rates of change) on the ascending and descending portions, and the peak values on the reflectance or vegetation index (VI) time series are respectively reached [30,35]. The different LSP metrics have been linked to various GP observations with varying degrees of success and are particularly sensitive to canopy level changes, such as leaf unfolding and defoliation, rather than to events such as bud-break [36]. Despite their limitations and a lack of direct correspondence to GP, various LSP metrics have nonetheless been able to provide valuable insights into vegetation seasonality and their various controls at much larger scales (both temporal and spatial) [29,37–40].

Although optical remote sensing has shown promising results in detecting phenology, many optical sensors pose several shortcomings in terms of spatial and temporal resolution. For example, the longest freely available optical dataset from the Landsat missions has been found to suffer from frequent cloud cover and scan line errors (for Landsat-7) that inhibit its utility to study land surface phenology (LSP) and needs integration with higher resolution datasets to overcome such challenges [41]. Vrieling et al. (2018) [21] pointed out that, even with the higher spatial resolution and frequent revisit time, the data availability from Landsats-8 is limited due to its single orbit and is not fine enough to capture fast-occurring changes such as cropland green-up. With Landsat data being infrequent (with a repeat cycle of 16 days), requiring high computing resources to process and the challenge of filling gaps in Landsat-7, especially in cloudy conditions, other coarse resolution sensors, such as MODIS, AVHRR, SPOT-VEGETATION and MERIS, gained popularity since high quality datasets were readily available and easier to process [42,43]. However, many LSP studies conducted using data from these sensors, although providing frequent observations, suffered from mixed pixels and were therefore limited in their applicability over heterogeneous landscapes [44–46]. Although mixed pixels are unavoidable, a careful handling of pixel heterogeneity may help in quantifying the variability in the estimated LSP [47,48]. The combined datasets from different Landsat satellites (such as 7, 8 and potentially 9 once it is operational) could provide a nominal 8 day ground coverage [49], and are thus one of the best
choices for sampling data at a higher spatial and temporal resolution. However, given the constraints of cost and image acquisition during the commercialisation of Landsat from 1984 to 2008 [43], the freely available, wide-swath coarser resolution sensors, such as AVHRR and MODIS, have been the focus of much interest in the phenological community.

To overcome some of the limitations of these lower spatial resolution sensors, newer satellite data from constellations of identical sensors, such as Sentinel-2, RapidEye and Planet (with higher temporal and spatial resolution), could help improve both our estimates and understanding of LSP. The newest European Space Agency Sentinel-2 satellites, available from 2015 (S2A) with the second sensor (S2B) launched in 2017 to double the data acquisition, provide data at much finer spatial (10–20 m) and higher temporal (up to 5 days) resolutions, although currently over a shorter time series than Landsat and MODIS. The reader is referred to the Sentinel-2 mission website for detailed information on the specifications of its sensors [50]. The temporal resolution is one of the biggest advantages of S2 data, with improved sampling of the ground, and overlapping areas (for S2A and 2B orbits) achieving even higher temporal sampling, which is helpful especially in areas where cloud cover may be frequent [51]. It has been observed that, for the same area, Sentinel-2 can sample approximately ten times more pixels when compared to Landsat-8, owing to its better spatial resolution and higher revisit frequency from two sensors (up to four times improved temporal sampling of cloud-free observations), thereby providing improved capability for monitoring vegetation from space [52,53]. Early studies using simulated Sentinel-2 bands from hyperspectral data showed promising results in monitoring vegetation chlorophyll and landcover. Those studies also found Sentinel-2 bands to be scalable with many optical sensors, such as MODIS, VIIRS, MERIS, Landsat and RapidEye [54,55]. Similarly, using S2-like bands from other optical sensors, Veloso et al., 2017 [56] found the results to be highly correlated with fresh biomass and the green area index (GAI), and able to capture short-lived phenological stages, thereby enabling precise monitoring of crop development. With a high geometric and spectral consistency with Landsat, and additional information in the short wave infra-red (SWIR) and red-edge wavelengths, Sentinel-2 provides exciting opportunities to monitor and study plant phenology in much greater detail [57–59]. The interest in using Sentinel-2 data for phenological research has clearly gained momentum in the last five years, as can be seen in Figure 1.

This review paper evaluates the progress of phenology-related research in forests, crops and grasslands utilizing Sentinel-2 data after five years of its successful operation. This paper discusses the phenology of (a) forests (integration of multisensor data and the various vegetation indices used), (b) croplands (crop mapping and yield estimation using time series data), (c) grasslands (including monitoring of management practices) and (d) other vegetation cover classes (wetlands and urban areas). Additionally, the advances made in the application of the three unique narrow bands in the red-edge portion of the spectrum, the issues of cloud cover in optical sensors and gap filling of time series for phenological studies are analysed. Recent developments in the use of Sentinel-2 in phenological research are synthesized to identify remaining challenges and the scope for future research.
2. Phenology of Forests

2.1. Integration of Sentinel-2 Data with Other Imagery

The Sentinel-2 pair of sensors are becoming increasingly popular for phenological studies of forests due to their enhanced spatial and temporal resolutions that have helped overcome some of the limitations of optical remote sensing. Lange et al. (2017) [60] reported a Sentinel-2-based normalised difference vegetation index (NDVI) to be better correlated to ground-based sensors, as compared to NDVI estimated from MODIS for which green-up dates lagged by 8–14 days relative to Sentinel-2 and in situ measurements. Senescence as calculated from MODIS data also occurred 4–11 days earlier relative to Sentinel-2 and in situ data, with these variations attributed to the mixed pixel effects of the low resolution MODIS images. Bolton et al. (2020) [42] used harmonized Landsat-8 and Sentinel-2 (HLS) data to calculate LSP metrics at a spatial resolution of 30 m and temporal resolution of 1–4 days using a “multi-sensor land surface phenology (MS-LSP)” algorithm. This product showed strong correlations ($r = 0.9$) with phenocam data for 50% green-up dates for deciduous broadleaf forests, whereas very weak correlations ($r = 0.15$) were found for evergreen needle leaf forests. As is the case for other satellite sensors, it was a challenge to estimate evergreen forest phenology due to its very small seasonal amplitudes. Nevertheless, HLS data were able to capture seasonal variations in vegetation phenology at finer spatial resolutions than was possible previously. By contrast, Zhang et al. (2020) [61] found that the HLS data still contained insufficient observations for some locations as a result of the irregular temporal sampling of data and gaps due to the presence of snow and clouds in the field of view. Most of the gaps in the data were found to occur in spring (for forests), which limits the ability to detect the early stages of the growing season. The authors filled the gaps in the 30 m HLS with data from the much coarser resolution VIIRS (500 m) using a shape-matching technique, and were able to improve the detail within the time series. The green up, or start of season (SOS), dates estimated from the VIIRS-HLS fused data had a closer agreement, i.e., mean absolute deviation (MAD), of 8 days with the in situ USA-NPN dates than with HLS (MAD of 14 days). Pastick et al. (2018) [62] also pointed out that NDVI, SWIR, blue, green and NIR bands in Sentinel-2 and Landsat-8 are highly consistent
(R² > 0.90) in HLS data. A review of the literature on forest mapping reveals that such studies focused on the use of multi-temporal images representing different periods within the growing season to map broad forest types and their seasonality, rather than focusing on specific phenological stages [63–67]. The general principle is to use multi-temporal images to increase separability between forest types through temporal variations in seasonal growth curves or the time series curve geometry, as mentioned in Hanes et al. (2014) [26]. It is however essential to evaluate and normalise the differences that exist between different sensors, and even between the same sensor, when images are acquired under different viewing and illumination conditions, as is often the case when generating a dense time series of data. This effect can be much greater between multiple Sentinel-2 images than may be seen between Landsat images, due to the larger swath and viewing angle of S2, which permits the more frequent viewing of the same point on the ground surface [68]. In order to remove such artefacts from the data several approaches to normalise Bi-directional Reflectance Distribution Function (BRDF) effects have been suggested in the literature [69–72], and hence should be considered in phenological research.

2.2. Vegetation Indices for Phenological Research in Woody Species

The use of phenological information in the form of time series of vegetation indices has been successfully implemented by Marzialetti et al. (2019) [67]. The authors used monthly Sentinel-2 NDVI data and random forest models with hierarchical logic to discriminate coastal vegetation into dense and sparse vegetation of herbaceous and woody types. The trajectory of the NDVI values were found to possess critical phenological information at key stages that helped distinguish vegetation classes. The variability of NDVI values during the year, particularly its annual mean and its value in summer, helped discriminate dense woody vegetation, such as forests and shrubs (which have a high NDVI through the year), from other vegetation classes that experience greater changes in their greenness (or NDVI) through the year.

Nevertheless, although NDVI has been reported to be highly correlated with vegetation greenness and biomass, it performs poorly in identifying parts of the canopy, such as flowers, which are mostly non-green [73,74]. This issue was addressed by Chen et al. (2019) [73], who tested NDVI and a bloom index based on the visible bands to identify flowering in almond trees. They used multiscale time series data from sensors including a unmanned aerial vehicle (UAV), aerial photographs from ceres imaging, and satellite images from PlanetScope, Sentinel-2 and Landsat-8 with spatial resolutions ranging from 2 cm (UAV) to 30 m (Landsat-8). This study took advantage of the fact that the reflectance of the almond flowers was much higher in the visible spectrum (especially in the red and blue bands) in comparison to green leaves and soil. The authors developed an enhanced bloom index (EBI) from the red, green and blue bands to temporally characterize the flowering of plants including the start, peak and end of blooming, which was completely missed by the NDVI. The higher spatial resolution Sentinel-2-based EBI could identify flowers with more than twice the accuracy of Landsat-8 and had the highest correlation (r = 0.69) with the very high resolution ceres data (although this was available for only a single date). PlanetScope and Sentinel-2 were both found to be spatially and temporally suitable to capture the trajectory of almond flower blooming and its critical stages through the use of EBI. Although it is not surprising that the very high-resolution ceres dataset provides the maximum capacity in detecting blooming signals early in the season, it would be interesting to see if the flowers and other parts of the vegetation can be unmixed from coarser resolution data as mentioned in Ji et al. (2020) [75], who reported the capability of the visible bands and red-edge bands in discriminating the photosynthetic and non-photosynthetic fractions. Several studies have used spectral mixture analyses to identify photosynthetic and non-photosynthetic fractions [76–78] and vegetation indices to estimate concentration of plant pigments [79–81]. However, similar efforts are lacking to discriminate flowers from other parts of the plant and existing techniques have been applied mostly to high-resolution aerial or hyperspectral data [82–86]. Notably, the EBI developed by Chen et al. (2019) [73] was found to perform poorly in discriminating non-white flowers, and it may be that other band combinations are more suitable for different coloured petals.
Foliar nitrogen (N) varies between forest phenophases [87,88] and has been previously estimated using expensive hyperspectral sensors and commercially available multispectral images at high spatial resolution. With the advent of Sentinel-2, it is now possible to compare N estimation accuracies across different phenological stages much more easily and cheaply. Mutowo et al. (2020) [87] used Sentinel-2 images to estimate foliar N content across the start and end of the growing season with a random forest classifier to establish a relationship between the bottom of atmosphere (BoA) reflectances from all Sentinel-2 bands (except bands 1, 9 and 10) and spatial variation of foliar N content. The study revealed that the variability of N concentration is smaller at the start of the season, resulting in high measurement accuracies at the start of the season. However, if it is possible to monitor N concentration in the leaves of deciduous trees, this would be very useful for tracking the progress of autumn leaf senescence, which is closely linked to N recycling. Such approaches require additional research matching Sentinel-2 imagery to N concentration in a wider range of forest vegetation types and tree species to determine the relationship to N recycling and other senescence parameters. Studies using hyperspectral data have indicated the improved performance of simulated Sentinel-2 bands in comparison to Worldview-2, RapidEye and Landsat-8 datasets for estimation of nitrogen and its effect on yields [89,90]. However, a review of the literature points to the lack of usage of Sentinel-2 satellite data, which need to be explored further for their efficacy in estimating nitrogen concentration.

3. Phenology of Croplands

Analysis of cropland phenology is different from that of forests in several ways. Although for both it is important to differentiate background signals (e.g., from bare soil) from vegetation, analysis of time series curves for croplands can be particularly challenging due to multi-cropping, management practices, and standing water for rice crops. Additionally, the multi-modal vegetation index time series curves observed in areas with high cropping intensity, or more than one cropping season annually, are in sharp contrast to the unimodal time series curves of forests. Therefore, processing of time series data must be tuned accordingly and allow for decreases in vegetation index values indicating events such as harvest or slurry spreading [31]. While these issues present a challenge for any remotely sensed approach to crop phenology, the higher spatial, temporal and spectral resolutions of Sentinel-2 offer potential improvements.

3.1. Mapping of Crops Using Time Series Data

Several crop mapping studies used phenological information in the process of discriminating crop types. Such studies worked on the logic that the spectral response of different crops might be similar on any single day but the separability among crop types is greatly improved in a multi-temporal space [91]. This fact has been utilized by many studies for mapping crop types using Sentinel-2 data. Tian et al. (2019) [92] used Landsat 7–8 and Sentinel-2 images to increase the frequency of data for mapping winter crops. They reported that temporal compositing of minimum and maximum NDVI values at the peak and end of season could mitigate the need for a longer time series of data to discriminate winter crops from forests and spring crops. The improved spatial and temporal resolution of the combined Landsat and Sentinel-2 series successfully identified the winter crops in 96% of the pixels, in comparison to 82% when using MODIS data. Another issue with crop mapping in fragmented landscapes was highlighted by Mohammed et al. (2020) [93]. Though the models in the study were able to explain 89% of the variation in crop census data, they revealed that Landsat NDVI could not distinguish between cropland and grassland at its native 30-meter spatial resolution. The authors therefore proposed using red-edge bands from Sentinel-2 and the wet season radar backscatter from Sentinel-1 since they have the potential to distinguish crops. Studies [94,95] have reported the applicability of vegetation indices, such as enhanced vegetation index (EVI), NDVI and land surface water index (LSWI), from the high spatial resolution Sentinel-2 time series data in identifying rice cropped areas with high accuracy. Inclusion of phenological metrics, such as the SOS and EOS dates, in the time series curves from a
combined Landsat-7 and 8 and Sentinel-2 dataset further increased the mapping accuracy of rice crop from 78% to 93% [95].

Multi-sensor data from Sentinel-2, Landsat and MODIS were also used in combination by Pringle et al. (2018) [96] for a hierarchical classification of groups of crops. The authors reported an increase of almost 40% in the data availability since 2015 (coinciding with the launch of Sentinel-2A). This study revealed that the EVI curvature parameters including the rates of green-up and senescence are the most important predictors in the crop type models, and hence are most affected by the temporal sampling of data. The slow growth and faster decline of pulse crops during winter helped in their separation from cereals that exhibited the opposite pattern. Kussul et al. (2018) [97] used multi-temporal data from Sentinel-1, Sentinel-2 and Landsat-8 to map crops during and at the end of the growing season. This study highlighted the issue of cloud contamination in optical (Sentinel-2 and Landsat-8) data, which the authors addressed using radar data from Sentinel-1. An accuracy of 80% in mapping crops was obtained with only Landsat data, which improved to 85% on addition of Sentinel-1 data. Similar improvements were also seen in the early season crop maps from a combination of Sentinel-1 and -2 data, which led to an accuracy of 93.7% when used in combination. Only two cloud-free scenes from Sentinel-2 provided an accuracy of 87% in comparison to an accuracy of 92% from eight Sentinel-1 images covering the same period. Sun et al. (2019) [98] reviewed Sentinel-1 and Sentinel-2 data for mapping agricultural areas, showing that Sentinel-1 data for agricultural applications is limited by complex image data structures and limited bands as compared to optical data. Sentinel-2 was shown to be the most effective to map crops, providing an overall accuracy of 91% in comparison to 86% and 77% accuracies obtained from the use of L8 and S1 data, respectively.

A study by Sitokonstantinou et al. (2018) [99] revealed that Sentinel-2 is superior to Landsat-8 for smallholder agri-mapping and crop discrimination in Spain, providing accuracies close to 90% in comparison to 70% when using Landsat data. Though the authors did not find significantly different numbers of cloud-free images from Landsat and Sentinel-2, they identified the presence of additional red-edge bands in Sentinel-2 as being critical to the higher classification accuracies. In this study, from an analysis of variable importance in classification models based on machine learning approaches, such as random forests, it was found that Sentinel-2 images from May (emergence of inflorescence in winter crops) and July (senescence of winter crops and leafing of summer crop) as well as near infra-red (NIR), red-edge (RE) and short wave infra-red (SWIR) bands were the most important among temporal and spectral variables. The plant senescence reflectance index (PSRI) calculated from the red, green and NIR bands also appeared to be of particular value compared to NDVI and NDWI. The importance of including information on key phenological stages for mapping sugarcane crops was also raised by Wang et al. (2019) [100], who identified seedling, elongation and the harvest periods to be critical for maximum separability from other classes including forests. In comparison, use of single Sentinel-2 NDVI images provided sub-par performance, ranging from 23% for the seedling stage to 70% for the elongation stage. To maximize accuracy, cloud cover limitations were overcome by compositing images in each key stage. Similarly, Ottosen et al. (2019) [101] could discriminate winter and spring crops using NDVI time series information with a focus on the peak and harvest seasons. The peak of season was particularly important in identifying weed-infested pixels, where peaks after crop harvest (in a unimodal distribution) indicated weed growth. Additionally, they unmixed 10 m Sentinel-2 pixels with the CORINE land cover labelled as “non-irrigated and arable land” into seven phenologically distinct classes. In addition to phenological information from the two-band EVI (i.e., EVI2), Nguyen and Henebry (2019) [102] found that inclusion of other spectral band ratios improved accuracies for mapping crops such as wheat and maize. Although the EVI2 index performs similarly to the three band EVI, it lacks the blue band and is calculated using the red and NIR bands only [103]. It was found that the dates of green-up and the peak of season were most important in achieving high accuracies for all crop types. Moeini Rad et al. (2019) [104] also suggested using temporal information to discriminate rice from other crops due to its distinctly higher reflectance in the red spectrum at the harvest stage and
lower NIR reflectance in the transplanting period. The use of such temporally distinct phenological information led to mapping accuracies ranging from 88% to 98% for different rice cropping systems.

3.2. Estimation of Crop Yield

In addition to crop mapping, Sentinel-2 data have also been successfully adapted to predict crop development and yield. A study by d’Andrimont et al. (2020) [60] highlighted the advantage offered by the high temporal resolution of Sentinel-1 and -2 that makes them particularly suitable for “in-season crop forecasting”, as the different phenological events can be reliably captured. In this study, Sentinel-1 data in the VV polarisation mode were used for the analysis since this mode was established to be sensitive to the change of canopy due to the flowering of rape seed. Additionally, it was hypothesised that the normalised difference yellow index (NDYI) based on the blue and green bands was suitable for capturing the increasing yellowness during flowering due to strong absorption in the blue band by the flowers. The high reflectance in the green and red bands is then perceived as yellow (the colour of the flowers). This study revealed that the in situ data pertaining to the onset and end of flowering match with those captured by the Sentinel-1 radar backscatter in the VV mode. As regards to Sentinel-2, the in situ peak flowering date matched with the average peak of the NDYI. It is worthwhile to note that the peak flowering date as captured by Sentinel-1 was delayed by 5 days compared to that captured by Sentinel-2 due to the attenuation of the C band backscatter signal by the flowers. However, a strong correlation was found between the peaks as detected by both Sentinel-1 and Sentinel-2.

From a review of crop phenology studies based on optical remote sensing data, Duncan et al. (2015) [28] envisaged that Sentinel-2 has the potential to mitigate the limitations of existing optical sensors for crop yield estimates. Small farms, which are challenging to study with lower spatial resolution imagery due to the presence of multiple land cover types within a pixel, can now be monitored at the Sentinel-2 scale and with much greater confidence. Prey and Schmidhalter (2019) [90] used hyperspectral ground reflectance data along with Landsat-8, Quickbird, Rapid Eye, World View-2 and Sentinel-2 to estimate grain yield and grain nitrogen uptake. It was observed that Sentinel-2 with its narrower bands performed the best for both objectives. The NIR band was found to be best for studying grain yields, while the RE bands were highly correlated to nitrogen-related traits of plants. In a similar study, Ayu Purnamasari et al. (2019) [105] studied yield prediction and land suitability for the cassava crop in Indonesia, revealing that maximum correlation can be found between vegetation indices and green-up, and that remote Sentinel-2 data also helped to predict yields up to five months in advance of the harvest season. This study also showed that the yield was highly correlated with vegetation indices such as NDVI, soil-adjusted vegetation index (SAVI) and the inverted red-edge chlorophyll index (IRECI), and to the leaf area index (LAI) and fraction of photosynthetically active radiation (fAPAR). It was reported that the average yield of cassava was approximately 12 and 8 t/ha in the highly and moderately suitable areas for cultivation, respectively. While mapping winter wheat and barley using Sentinel-2 data, Nasrallah et al. (2018) [106] flagged the issue of mixed pixels that arises with low and medium resolution optical data. The authors showed that, for pixels of a single species, NDVI was highly correlated to above-ground biomass, and suggested a decision tree applied to the temporal profiles of NDVI could effectively discriminate wheat cropped areas and thus may help in monitoring crop yields, as mentioned in [107,108].

4. Phenology of Grasslands

4.1. Monitoring Seasonal Change in Grassland to Determine Management Practices, Invasion and Biomass Production

Grasslands provide valuable ecosystem services, and account for more than a third of all agricultural land on earth [55,109]. Characterising phenology of grasslands is particularly challenging since they are sensitive to both management practices (i.e., frequent grazing and cutting events) and weather [110]. Hill (2013) [55] noted that frequent observations are needed to capture changes
associated with management practices in grasslands, such as grazing and mowing. In his study, undertaken before Sentinel-2 was launched, Hyperion data were used to simulate observations from Sentinel-2, MODIS and VIIRS. Various vegetation indices and bands were used to study pigments (NDVI, carotenoid reflectance index (CRI), anthocyanin reflectance index (ARI), red–green ratio (RGR)), water content (normalised difference infra-red index (NDII)), soil and senescence (SWIR, PSRI) and biomass (soil adjusted total vegetation index (SATVI)). This study revealed that most of the simulated Sentinel-2 indices except ARI were highly correlated to MODIS and VIIRS with strong linear relationships. VIIRS and MODIS both lack a band at 700 nm that is critical for calculating ARI, and hence this had to be substituted with the closest available band, which might explain its poor correlation with Sentinel-2-based ARI data. The vegetation indices were used to predict grassland cover and nominal vegetation status (with respect to disturbances, i.e., degraded or pristine grasslands).

The Sentinel-2 constellation, with its high spatial and temporal resolutions, provides optimum conditions for detecting and quantifying mowing events, although its utility can be impeded by cloud cover given how quickly the grass regrows after mowing and also that long gaps in data can lead to failure in detecting disturbances. To overcome this problem, Griffiths et al. (2020) [111] presented a multi-sensor approach by using harmonised Landsat and Sentinel (HLS) data to map the frequency and timing of mowing events, with the underlying assumption that vegetation growth follows a climate-driven phenological trajectory and large deviations from the normal trajectory are a result of removal of biomass. They used NDVI data derived from 10-day HLS composites for detecting up to five mowing events per season. On evaluating the detection algorithm, they found that the HLS data allow for identification of mowing frequency and the timing of the detected mowing event with a high degree of inter-parcel variability. Pixels with only one mowing event were reported to have a higher probability of omission errors, since the application of cloud masks could potentially filter out valid observations in the time series data.

HLS data were also used by Pastick et al. (2020) [112] to map invasive annual grass species, which can be a major contributor to ecosystem degradation, and, therefore, it is imperative to monitor and map them for better management of the environment. Previously, invasive species were studied using MODIS and AVHRR to quantify their spatio-temporal variations [113]; however, the spatial resolutions of these sensors posed a drawback. The results showed better agreements for phenological metrics calculated from the expedited weekly MODIS (eMODIS) and integrated HLS data (with correlations of 0.55, 0.81 and 0.91 for SOS, peak of season and EOS timings, respectively). Furthermore, this study showed that the regression tree model using both spatio-temporal dynamics was better (correlation of 0.91 with eMODIS) than the strictly temporal information-based per-pixel approach (correlation of 0.87 with eMODIS) in constructing time series data. The invasive species maps were obtained at an enhanced resolution of 30 m, as compared to the much coarser resolution of 250 m for eMODIS data, and are better suited for decision makers on the ground.

The use of information from multiple sensors was also advocated by Wang et al. (2019) [110], who used Sentinel-1 and -2 and Landsat-8 data to predict the biomass of tall grazing grasses with high accuracy (correlation of 0.76 with in situ data). Their study reported an improvement of over 50% in the prediction accuracy of LAI and above-ground biomass (AGB) from a multiple-sensor approach as compared to Sentinel-1 alone. The phenology (i.e., timing of SOS, peak and EOS) estimated from the combined Landsat-8 and Sentinel-2 data could also explain around 92% of the variability observed in gross primary productivity (GPP) estimates from flux towers. Cerasoli et al. (2018) [89] used Landsat-8 and Sentinel-2-based vegetation indices (VIs) and bands to assess the combination that could best describe grassland phenology and GPP from annual peak to senescence. This study revealed that, although many of the bands in the two sensors are similar, Sentinel-2 performed better (with an improvement of over 10% for predicting GPP) than Landsat-8 in VI-based models due to the presence of additional red-edge bands. Additionally, information from the red-edge and SWIR was found to be critical for improved GPP estimates.
4.2. Matching Sentinel-2 with Phenocam Data in Grasslands

In contrast to satellites, phenocams provide close-range remotely sensed imagery at much higher spatial and temporal resolutions, and do not require atmospheric corrections [114]. Though phenocams currently typically operate in only a few spectral channels (e.g., three visible and one NIR band), they have been successfully used across a wide range of ecosystems, and are reported to provide better agreement with GP observations, more robust estimates of productivity, and an ability to track changes at the canopy level [115–119]. While their field of view is considerably reduced compared to satellites, multiple phenocam devices have emerged in the last decade, and have been used effectively to track vegetation changes. As with the satellite data, some pre-processing of the time series data is necessary before calculation of the phenological metrics, but gap-filling due to data loss is rarely required, unless the instrument fails, due to the high density of image acquisition [20]. It is important to acknowledge that the images acquired by satellites and phenocams cannot be directly compared with each other due to intrinsic differences in their sensors, fields of view and data format; however the phenological indicators derived from each can be equivalent. Gómez-Giráldez et al. (2020) [32] highlighted that, for phenology related research, a gap remains between observations made in the field and those acquired through satellites. Their proposed solution lies in integrating data from phenocams and satellites to minimise this scale gap. The authors used field measurements of soil moisture and meteorological conditions along with phenocam and Sentinel-2 data to study phenology and its relation to the hydrological behaviour of grasslands in an oak-grass savanna system. The green chromatic coordinate (GCC), computed using the red, green and blue bands, though not a direct measurement of chlorophyll content, has been shown to minimise the effect of light on images captured by cameras and to quantify the greenness as captured by phenocams. A high correlation (greater than 0.7) was obtained between the grassland’s phenological metrics (i.e., start of season (SOS), end of season (EOS) and peak of season (POS)), calculated from phenocam-based GCC and vegetation indices (VIs) from Sentinel-2, such as NDVI, EVI, SAVI and MERIS terrestrial chlorophyll index (MTCI). It was found that all VIs could estimate EOS with greater reliability than POS, with SOS estimates delayed by around 10 days for both Sentinel-2-based GCC and VIs as compared to that from phenocam-based GCC. Finally, the authors concluded that despite the additional information contained within more sophisticated VIs, NDVI was best suited for such analyses and its values most closely matched those obtained from the terrestrial photography GCC. However, no explanation was given regarding the reasons behind such improved performance of NDVI over other indices.

Zhou et al. (2019) [120] found a closer correspondence of Sentinel-2-based SOS for grasslands with phenocams in comparison to the one from Landsat-8. The authors found a 50% increase in the frequency of clear observations when using Sentinel-2 versus images available from Landsat-8 during days 90–120, which represents the SOS based on ground observations. In areas with high cloud cover, three to five cloud-free Sentinel-2 images were available as compared to one to two clear images from Landsat-8. In fact, Landsat-8 data were found to be unable to detect annual curves due to a lack of images through the growing season, and hence this sensor alone would be insufficient to extract seasonality for many pixels in the central USA. Within the Sentinel series, Stendardi et al. (2019) [121] found that the SOS and EOS estimated from the optical Sentinel-2 sensors were closer to those estimated from phenocams and were only delayed by 4 and 10 days, respectively. In comparison, the C-band microwave sensors on Sentinel-1 estimated the SOS and EOS with a delay of 10 and 20 days, respectively. Sentinel-2 was also better able to detect management practices, such as mowing events on mountainous slopes, primarily because the phenology detected from Sentinel-1 was less accurate due to the limitations of the microwave signal. The lack of stability in the VV backscatter at high elevations and the non-applicability of general concepts to detect phenology of pasture lands that had a flat time series curve of VH backscatter (i.e., no seasonal amplitude could be detected to extract phenological metrics) also limited the applicability of microwave data. A review of the literature revealed that, in contrast to optical data, the signals from both passive and active microwave sensors are highly dependent on the land cover and hence need to be processed separately for each class [121,122].
5. Phenology of Other Land Classes

5.1. Mapping of Wetland Vegetation

Sentinel-2 data have also been used to map the phenology of complex and ecologically important wetlands. Rupasinghe and Chow-Fraser (2019) [123] compared the classification accuracies of a support vector machine classifier using data from Sentinel-2, Landsat-7 and Landsat-8. This study demonstrated that, when mapping *Phragmites australis* (common reed), the overall accuracy was highest for Sentinel-2 across all seasons. The lack of red-edge bands and the omission of the Landsat SWIR bands was responsible for the reduction in accuracies (up to 3–8% across sites). In another study by Mahdianpari et al. (2019) [124], different types of wetlands were found to be mapped at an accuracy of 83% using multi-year composites of visible bands and vegetation indices derived from Sentinel-2 images for the summer season. It was found that bogs, fens and marshes that were poorly separable in the visible bands were discriminable using vegetation indices such as the NDVI and modified soil-adjusted vegetation index (MSAVI2). The combined use of Sentinel-1 and Sentinel-2 data further improved the discrimination of wetlands and led to overall accuracies of 88%. The improved performance of combined Sentinel-1 and Sentinel-2 time series data was also confirmed by Cai et al. (2020) [125], who found a 10% increase in overall accuracy of mapping different land covers in wetlands through the combined use of the data sets. The different land covers were found to have unique time series curves in both the optical and SAR domains, which led to improved discrimination of land cover types. The addition of phenological metrics such as SOS, EOS, length of season (LOS), maximum NDVI and the seasonal amplitudes led to an additional improvement of 4% in the overall mapping accuracy. A review of the literature suggests that the variability in NDVI and NDWI time series pairs is directly proportional to the separability between the various land covers in wetlands, and hence critical in improving classification accuracy.

5.2. Dealing with Mixed Pixels in Urban Areas

Urban areas are also significant reserves of vegetation that are affected by the differences in microclimate, artificial photoperiod and ground water status in comparison to rural and natural vegetated areas [126]. Tian et al. (2020) [47] used a fused Landsat-8 and Sentinel-2-based NDVI time series dataset to study the impacts of spatial resolution on the differences between urban and rural vegetation spring phenology, i.e., SOS. The authors reported that irrespective of the resolution, urban areas experience early SOS in comparison to rural areas. The coarser datasets however tend to overestimate the earliness in the urban areas, mostly due to higher mixing of different land covers in coarser pixels. The study suggested the use of high spatial resolution data in urban areas to prevent mixing of diverse time series profiles of NDVI. However, even at the highest resolution, pixels are rarely 100% homogenous within an urban fabric. Granero-Belinchon et al. (2020) [48] suggested a solution in this regard. The authors used a double logistic function to fit the time series data, and pixels with very high fit errors and very low VIs during maturity period were accurately identified to be mixed or contaminated with non-vegetation features. Additionally, the use of a local smoothing function i.e., Savitzky–Golay filter, on the identified “pure” pixels allowed the authors to correctly identify periods of stress or disturbance both intra- and inter-annually, which were tightly linked to an increase in temperature and decrease in soil water content in the summer season. However, the double logistic function fitted pure pixels completely missed such stress events, since this function results in smoother outputs than Savitzky–Golay and hence should be avoided when analysing such events. The effects of coarse spatial resolution on estimation of phenology in heterogeneous landscapes was evaluated by Zhang et al. (2017) [127], who suggested the pixel SOS to be driven more by the timing of the earliest vegetation to green-up than the later sub-pixel components. The authors suggested that the SOS at coarse resolutions is detected when around one-third of the pixel has experienced green-up; however, it is unknown whether this threshold is valid across different landscapes and for different methods of LSP estimation, especially in urban areas.
6. Current Developments in Using Sentinel-2 for Phenological Research

6.1. Performance of Sentinel-2 Red-Edge Bands in Phenological Research

In addition to better temporal and spatial resolutions, the red-edge bands on the Sentinel-2 sensors have been considered beneficial in studying urban vegetation phenology by Granero-Belinchon et al. (2020) [48]. The authors showed that, irrespective of the processing methods (double logistic function fitting or Savitzky–Golay smoothing of the time series), the normalised difference red edge (NDRE)-based time series (as compared to NDVI and normalised burn ratio (NBR)) shows greater stability (i.e., smaller standard deviations for pure vegetation pixels at a site) at the start of season and the peak of season. From a review of publications, it has emerged that most of the studies evaluated the temporal trajectory of red-edge bands for detecting plant pigments, discriminating different vegetation classes and estimating productivity. The lack of studies using red-edge bands to estimate different LSP metrics (i.e., SOS, EOS and POS, etc.) is quite evident.

The effectiveness of using Sentinel-2 red-edge bands has been tested in various applications, including detection of chlorophyll, tree species classification and biomass estimations in grasslands and crops [128–130]. The red-edge bands in Sentinel-2 were found to be comparable to MERIS and RapidEye, but band 6 (733–748 nm) is unique since it is situated at the longest wavelengths of the red-edge where the gradient remains high (the comparable band 10 of MERIS at 750–757.5 nm is situated slightly beyond the steep high gradient region). The Sentinel-2 bands 5 and 6 provide improved estimates of red-edge over the other sensors due to their optimal bandwidths and better spatial resolution [54]. The red-edge information from the MERIS based MTCI (operational from 2002–2012) was reported to have a robust relationship with gross primary productivity (GPP), but the data were limited by the coarse spatial resolutions of 300 m [131]. Lin et al. (2019) [33] also highlighted the suitability of red-edge-based vegetation indices from Sentinel-2 over the conventional non-red-edge indices (i.e., NDVI, EVI and the NIR reflectance of terrestrial vegetation (NIRv)). In this study chlorophyll red-edge and chlorophyll green indices, both based on red-edge bands, showed high correspondence (coefficient of determination \( r^2 \) ranging from 0.69 to 0.89) with in situ estimates of grassland GPP. However, it was found that the light use efficiency (LUE)-based GPP from MODIS rather than the red-edge-based vegetation indices (VIs) from Sentinel-2 were better suited for forested landscapes. The authors explained that GPP in forests is mainly driven by incident photosynthetically active radiation (PAR) rather than the chlorophyll content as detected by Sentinel-2 red-edge bands. In a study by Pinheiro et al. (2019) [132], the red-edge bands 5 (705 nm) and 6 (740 nm) on Sentinel-2 were also reported to capture differences in the density of tree canopies. Those red-edge bands in the wet season were particularly important to distinguish bushy grasslands from broadleaf scrubs.

Using pre-operational Sentinel-2 data, Immitzer et al. (2016) [51] showed that crop lands and tree species could be classified with high accuracy. From a random forest classification model, the red-edge (band 5) and SWIR (band 11 at 1610 nm) were found to be the two most important variables in comparison to other Sentinel-2 bands. Delegido et al. (2011) [58] used a simulated Sentinel band 4 (665 nm) in combination with the red-edge bands 5, 6 and 7 (705, 740 and 783 nm, respectively) from hyperspectral data to estimate chlorophyll content with improved accuracy. Additionally, this study showed that bands 4 and 5 can be used to calculate green LAI for crops. Sentinel-2 bands simulated from in situ spectroradiometers were also used by Clevers and Gitelson (2013) [133], who reported a highly linear relationship \( R^2 > 0.80 \) between the chlorophyll red-edge index (from red-edge bands 7 and 5) and the measured chlorophyll and nitrogen content of crops. The authors cited the higher sensitivity of red-edge bands due to their lower absorption by chlorophyll in the red-edge region as compared to the red band (when using NDVI data). Once operational, a study by Peng et al. (2017) [134] showed that canopy chlorophyll content in maize and soybean, as estimated from field observations, matched closely to that of Sentinel-2 calculated indices using the red-edge bands. The red-edge bands were found to be much more resistant to hysteresis driven by canopy structures compared to chlorophyll estimation using NDVI, EVI and simple ratio (SR). Arroyo-Mora et al. (2018) [135] found that at
the landscape scale, Sentinel-2 showed a high correlation of up to 0.73 with hyperspectral airborne imagery (compact airborne spectral imager (CASI)) when the latter was resampled spectrally and spatially to match that of Sentinel-2. The same study also demonstrated that, compared to NDVI or SR, the Sentinel-2 red-edge normalised vegetation index (RE-NDVI) was able to capture the greening trend most closely to that measured by field spectroscopy. Zarco-Tejada et al. (2018) [136] found that Sentinel-2 NDVI used in combination with the chlorophyll red-edge index could explain the variability in the temporal trajectory of changing vegetation conditions by about 88%. This helped establish distinct baseline conditions for monitoring vegetation health, and is therefore able to distinguish between healthy, recovering, and declining trees. Among a set of different indices evaluated in the study mentioned above (photochemical reflectance index (PRI), NDVI, transformed chlorophyll absorption index (TCARI)/soil adjusted vegetation index (OSAVI), chlorophyll red-edge and Maccioni index (Mac)), it was found that the chlorophyll red-edge had a close linear relationship with chlorophyll-a and -b. Similar observations were also made by Palchowdhuri et al. (2018) [91], who reported an improvement of over 10% in the overall accuracy from the multi-temporal NDVI-based mapping of crops when information from the red-edge bands was also included. This study revealed that the highly coincident spectral response of winter season cropped wheat and oat in the broad visible and NIR bands of high spatial resolution WorldView-3 were only separable in the narrow bands of Sentinel-2-based red-edge and NIR. By contrast, the red-edge bands were shown by Clevers et al. (2017) [137] to perform poorly in tracking the temporal changes in chlorophyll content in potato crops with both chlorophyll absorption and canopy scattering in the red-edge band 5 of Sentinel-2, and it is not possible to distinguish the contribution of each of the phenomena. In another study, Sun et al. (2019) [98] used data from several Sentinel-2 bands and indices to monitor colour changes and texture that indicate crop growth stage and crop density, respectively. The authors showed that the narrow NIR band 8a from Sentinel-2 revealed separability among crops for all growth stages, whereas the red-edge bands and other vegetation indices (i.e., triangular vegetation index (TVI), EVI, NDWI and normalised difference tillage index (NDTI)) could indicate crop separability only from the middle maturity, or mid growing season, onwards.

6.2. Overcoming Cloud Cover

When compared with other satellite and ground-based remote sensors, Sentinel-2 clearly offers spatial, temporal and/or spectral benefits. The high spatial resolution of 10–20 m, temporal resolution of 5 days or less, and multiple red edge bands makes Sentinel-2 quite promising in observing LSP. However, it does suffer from limitations for phenological studies, especially with respect to the unavoidable data gaps arising from cloud cover (which are a characteristic of optical sensors). For example, Duncan et al. (2015) [28] discussed the fact that high frequency cloud cover inhibits accurate estimates of SOS and POS, and Villa et al. (2018) [138] showed that the impact of clouds, and therefore missing observations, also affects estimates of EOS. This study revealed that identification of the SOS is more sensitive to cloud cover than POS or EOS because it tends to occur over a shorter period of time. Several successful initiatives involving data fusion have been highlighted, such as the HLS dataset to overcome the issue of data gaps arising due to cloud cover. The VENμS sensor with its Sentinel-2-like spatial and spectral characteristics but a two-day revisit also deserves to be mentioned. It has been reported to offer improved capabilities for use in conjunction with Sentinel-2, but its availability is limited to a few sites globally and hence offers little scope for greater use [139–141].

The accuracy of cloud masks is critical to minimise errors in processing of time series data from satellites. Faulty cloud masks have been shown to deviate the estimated start of season by up to 10 days [142] and hence require careful consideration in modelling of vegetation index time series curves. This is particularly important for the Sentinel-2 dataset, since Pastick et al. (2018) [62] showed that the cloud masks available with the HLS, which uses the popular LaSRC and Fmask algorithms, were of poor quality (accuracy of 76–89%) and had to be recalculated from Landsat-8 and Sentinel 2 data using decision trees (yielding an accuracy of 99%). The European Space Agency-based processor,
Sen2cor, is also known to contain several limitations, particularly in misclassifying urban areas and bright pixels as clouds, which further limits the accuracy of its scene classification layer [143,144]. Hence, more investigation is required in development of improved cloud masks, possibly through local refinement according to prevailing atmospheric conditions.


The issue of cloud cover that limits the number of usable optical images is a well-known challenge for phenology studies and affects all land covers. Studies have used different methods to overcome the limitations of frequent and/or long gaps (caused by cloud cover and snow) in optical time series data. Most notable among the gap-filling methods are the use of temporal information only or seasonal means [30,145,146], use of both spatial and temporal information [147,148], and other techniques in the temporal domain, such as Fourier transformation or harmonic analysis of time series data [149,150]. Data gaps in one sensor time series have also been filled using data from other sensors; for example Sentinel-2 series have been used in combination with Landsat 8 to produce a denser time series [61,70].

Vrieling et al. (2018) [21] analysed a time series of Sentinel-2 NDVI data and concluded that both the frequency and the timing of data gaps resulting from cloud cover are important in ensuring correct estimation of LSP metrics. For example, the presence of above-normal cloud cover in the spring season in slow-growing salt marshes of the Netherlands had less impact on the estimated LSP as compared to a rapidly greening-up agricultural plot. Jönsson et al. (2018) [65] modelled time series NDVI data using a double logistic function and filled temporal regions with long gaps using a “shape prior”, or in simple terms the averaged climatological shape (analogous to seasonal means of a time series). The authors reported an improvement in the SOS and EOS calculated from the use of shape priors for filling gaps in time series; this reduced the root mean square error (RMSE) between Sentinel-2 and daily MODIS-based LSP by more than half. Lessio et al. (2017) [57] reiterated the high consistency between Sentinel-2 and Landsat-8 bands with respect to their spatial and temporal characteristics, which could be used to map crops at higher resolutions. They found the NDVI and normalised difference water index (NDWI) pairs (from S2 and L-8) to be significantly correlated with each other, with negligible bias and systematic errors in geometric positioning of pixels. Differences observed between the indices were primarily due to the temporal mismatch in image acquisition, but those could be managed using a linear bias correction. This study thus revealed the immense scope in the conjunctive use of Sentinel-2 and Landsat-8 data to increase the temporal sampling of data and overcome the issue of gaps in optical time series data for phenological studies. In another study, Wu et al. (2018) [151] found that there were insufficient cloud-free images from the high spatial resolution Sentinel-2 sensors to identify the presence of cotton rot, and while there were enough clear MODIS images, its resolution of 250 m was inadequate. They fused both datasets to overcome their individual shortcomings and were able to simulate high fidelity Sentinel-2 data at much higher temporal frequencies. It was observed that the phenological curves of healthy and infected cotton differed significantly, and the value of peak NDVI, productivity, SOS and length of season were all reduced in the rot-affected cotton crop. Solano-Correa et al. (2020) [152] used a non-parametric regression technique to extract spatio-temporal phenological information from crop fields using Sentinel-2 NDVI to detect fallow lands and cropped pixels. In this study, accuracies in resulting maps were partly attributable to the higher temporal resolution that aided in obtaining multiple cloud-free images per month. Additionally, data from multiple sensors, such as VIIRS, Landsat and Sentinel-2, have been successfully used to fill gaps in data from the use of single sensors [53,61,153]. By contrast, temporal compositing of data covering key stages of crop growth has also been suggested to mitigate the need for very high frequency time series data to distinguish crops [92].

Though the evolution of different gap-filling techniques has led to large improvements in time series analysis of vegetation phenology, they can only be used as a surrogate for real data. Even within the HLS data set, Zhang et al. 2020 [61] found missing high quality information for periods of up to four months. Similarly, issues are expected to be encountered when analysing time series data for areas...
with extensive and frequent cloud cover [28]. Geostationary satellite sensors, such as the SEVIRI and Himawari-8, that provide multiple images at a sub-daily resolution, could help in further improving the identification of cloud-free data for accurate monitoring of land surface phenology [154,155]. Studies [155–158] evaluating the performance of geostationary datasets to construct time series of vegetation indices and estimate different LSP metrics have reported an increase of more than 50% cloud-free data in comparison to data from MODIS and VIIRS. Though improvements in SOS were marginal, the EOS estimated from the geostationary satellites were within days of the observed in situ dates, whereas MODIS-retrieved dates deviated by up to a month. The high temporal resolution of such geostationary satellites could thus be used in conjunction with Sentinel-2 sensors to model phenology effectively; however their low temporal resolution and limited viewing angle can limit their applicability for fragmented landscapes and at higher latitudes.

6.4. Further Prospects

The current HLS dataset covers the period starting from 2013 for L30 and 2015 for S30 products, as it includes only Landsat-8 and Sentinel-2 data. To date, little work has been reported on inclusion of Landsat-5 and 7 to further extend the historic time series back in time (though data from Sentinel-2 would be lacking for the period before 2015), which could provide more than 30 years of data starting from the 1980s. A study by Roy et al. (2016) [159] provides valuable insights into statistical techniques for normalising data from Landsat-5 and 7, which have different sensor characteristics. Such procedures in combination with data from high temporal frequency data from geostationary satellites (as discussed in Section 6.3) could further improve the modelling of land surface phenology.

With the Sentinel 2-C and 2-D sensors expected to be launched in 2021, to provide continuity of the mission, data from these new satellites, combined with the existing Sentinel-2 data (A and B, provided they are still in operation) could further help resolve issues associated with cloud cover. The Sentinel-2 suite of sensors is thus expected to provide up to 15 years of freely available time series data, which will facilitate more comprehensive studies of vegetation phenology and their climatic drivers. Similarly, with Landsat 9 (that has a similar design to Landsat-8) scheduled to be launched in 2021, the success of the HLS dataset could be further extended and improved. All of these data sets are expected to improve phenological studies by overcoming existing limitations.

In spite of the additional red-edge and SWIR bands available on the Sentinel-2 multispectral instruments with their improved vegetation species detection capabilities, there has been very little research on their efficacy to track vegetation cover and its phenology. For example, out of approximately every four papers that analyse NDVI or EVI derived from Sentinel-2 imagery, only one mentions either SWIR or the red-edge bands. However, there is a growing body of evidence that suggests that the information contained within the red-edge bands is critical for discriminating not only species but also their growth stages and health status. This review also brought to light that most Sentinel-2-based studies are concentrated on mapping crops and their seasonality. Although Sentinel 2 possesses capabilities to monitor phenological and other events across a range of different ecosystems as mentioned in this review, more attention is still needed to show enough confidence in the improved potential of Sentinel-2 in LSP studies.

The literature analysis in this review paper found Sentinel-2-based phenology research to be mainly skewed towards croplands, with other land covers lacking similar attention amongst peer-reviewed papers. Even within the croplands, the phenological metrics, such as the SOS, POS, and EOS, are not always calculated, and the studies mostly concentrate on mapping of crop types and discriminating croplands from other land uses and land covers. Along with other information on crops, such as cropping intensity, location, health, water use and productivity, crop phenological estimates are equally critical for ensuring world food security [160]. Additionally, information on crop phenology (i.e., the cropping calendar and the dates of key growth stages) are important in the agronomic management of crops [161]. Hence, more efforts are required for testing the efficiency of Sentinel-2 data in estimating phenophases of crops.
7. Conclusions

A growing body of work is being conducted using Sentinel-2 for both direct phenological assessment, such as changes in timing of phenological stages, and indirect applications, such as land cover mapping based on images from multiple dates through the growing season representative of the different phenological stages of the vegetation cover. The higher spatial, spectral and temporal resolution of Sentinel-2 is preferable to many of the other freely available optical datasets, since it has shown improved capabilities for vegetation mapping and estimating phenology. Sentinel-2 imagery used on its own in phenological studies has been proven to offer benefits over other sensors; however, its greatest strength lies in the ability to integrate the imagery with compatible sensors. This increases the opportunity to acquire cloud-free coverage and thus capture rapid, and short-lived, phenological events; however, as an optical sensor Sentinel-2 will always have limitations in areas with extensive and frequent cloud coverage. Given the success of the research done to date with data from within one year, or a small number of years, it would be envisaged that as the length of the Sentinel-2 record increases, this work can be expanded to explore climatological and human influences on phenology, as has been very successfully seen for the coarser resolution sensors, such as MODIS and AVHRR.

Author Contributions: G.M. wrote the first draft of the paper. F.C. and A.W. provided inputs and helped in extensive editing of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This project is funded under the EPA Research Programme 2014–2020 (PhenoClimate, 2018-CCRP-MS.54). The EPA Research Programme is a Government of Ireland initiative funded by the Department of Communications, Climate Action and Environment. It is administered by the Environmental Protection Agency, which has the statutory function of co-ordination and promoting environmental research.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

8. GEO BON. What are EBVs?—GEO BON. Available online: https://geobon.org/ebvs/what-are-ebvs/ (accessed on 30 June 2020).


51. Immitzer, M.; Vuolo, F.; Atzberger, C. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. Remote Sens. 2016, 8, 166. [CrossRef]


57. Lessio, A.; Fissore, V.; Borgogno-Mondino, E. Preliminary Tests and Results Concerning Integration of Sentinel-2 and Landsat-8 OLI for Crop Monitoring. *J. Imaging* 2017, 3, 49. [CrossRef]


60. Lange, M.; Dechant, B.; Rebmann, C.; Vohland, M.; Cuntz, M.; Doktor, D. Validating MODIS and Sentinel-2 NDVI Products at a Temperate Deciduous Forest Site Using Two Independent Ground-Based Sensors. *Sensors* 2017, 17, 1855. [CrossRef]


75. Ji, C.; Li, X.; Wei, H.; Li, S. Comparison of Different Multispectral Sensors for Photosynthetic and Non-Photosynthetic Vegetation-Fraction Retrieval. *Remote Sens.* 2020, 12, 115. [CrossRef]


141. Li, N.; Zhang, D.; Li, L.; Zhang, Y. Mapping the spatial distribution of tea plantations using high-spatiotemporal-resolution imagery in northern Zhejiang, China. *Fores* 2019, 10, 856. [CrossRef]


© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).