



Article

Towards Optimising the Derivation of Phenological Phases of Different Crop Types over Germany Using Satellite Image Time Series

Abdelaziz Htitiou ¹, Markus Möller ¹, Tanja Riedel ², Florian Beyer ¹ and Heike Gerighausen ^{1,*}

¹ Julius Kühn Institute (JKI), Federal Research Centre for Cultivated Plants, Institute for Crop and Soil Science, Bundesallee 58, 38116 Braunschweig, Germany

² Julius Kühn Institute (JKI), Federal Research Centre for Cultivated Plants, Institute for Strategies and Technology Assessment, Stahnsdorfer Damm 81, 14532 Kleinmachnow, Germany

* Correspondence: heike.gerighausen@julius-kuehn.de

Abstract: Operational crop monitoring applications, including crop type mapping, condition monitoring, and yield estimation, would benefit from the ability to robustly detect and map crop phenology measures related to the crop calendar and management activities like emergence, stem elongation, and harvest timing. However, this has proven to be challenging due to two main issues: first, the lack of optimised approaches for accurate crop phenology retrievals, and second, the cloud cover during the crop growth period, which hampers the use of optical data. Hence, in the current study, we outline a novel calibration procedure that optimises the settings to produce high-quality *NDVI* time series as well as the thresholds for retrieving the start of the season (SOS) and end of the season (EOS) of different crops, making them more comparable and related to ground crop phenological measures. As a first step, we introduce a new method, termed UE-WS, to reconstruct high-quality *NDVI* time series data by integrating a robust upper envelope detection technique with the Whittaker smoothing filter. The experimental results demonstrate that the new method can achieve satisfactory performance in reducing noise in the original *NDVI* time series and producing high-quality *NDVI* profiles. As a second step, a threshold optimisation approach was carried out for each phenophase of three crops (winter wheat, corn, and sugarbeet) using an optimisation framework, primarily leveraging the state-of-the-art hyperparameter optimization method (Optuna) by first narrowing down the search space for the threshold parameter and then applying a grid search to pinpoint the optimal value within this refined range. This process focused on minimising the error between the satellite-derived and observed days of the year (DOY) based on data from the German Meteorological Service (DWD) covering two years (2019–2020) and three federal states in Germany. The results of the calculation of the median of the temporal difference between the DOY observations of DWD phenology held out from a separate year (2021) and those derived from satellite data reveal that it typically ranged within ± 10 days for almost all phenological phases. The validation results of the detection of dates of phenological phases against separate field-based phenological observations resulted in an *RMSE* of less than 10 days and an *R-squared* value of approximately 0.9 or greater. The findings demonstrate how optimising the thresholds required for deriving crop-specific phenophases using high-quality *NDVI* time series data could produce timely and spatially explicit phenological information at the field and crop levels.



Citation: Htitiou, A.; Möller, M.; Riedel, T.; Beyer, F.; Gerighausen, H. Towards Optimising the Derivation of Phenological Phases of Different Crop Types over Germany Using Satellite Image Time Series. *Remote Sens.* **2024**, *16*, 3183. <https://doi.org/10.3390/rs16173183>

Academic Editors: Clement Atzberger, Jochem Verrelst, Katja Berger and Egor Prikaziuk

Received: 31 July 2024

Revised: 23 August 2024

Accepted: 25 August 2024

Published: 28 August 2024



Copyright: © 2024 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 (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: crop phenology; Sentinel-2; *NDVI*; phenological metrics; time-series analysis

1. Introduction

Phenology is a word derived from the Greek word “phaino”, which means “to appear” or “to show yourself”, while its scientific definition is to study all biological events that characterize the life cycle of a plant or a crop (e.g., leaf emergence, flowering, dormancy) [1]. Due to its high importance, vegetation phenology has been the focus of interest in

various research domains. This concerns crop-related applications in particular, where accurate estimates of crop physiological growth phases or phenophases are crucial for many agronomic practices, such as irrigation and fertilization scheduling, agricultural weather or biodiversity index derivation, soil erosion, crop condition monitoring, and yield estimation or integrated pest management [2–9]. Crop phenophases were traditionally recorded by manual field observations and surveys, e.g., [10]. Although these types of data sets can accurately describe crop phenology, they are limited in their ability to provide continuous temporal and spatial phenological information on a large scale due to their time and resource-consuming nature [11]. Moreover, the large spatial variability of crop phenology, combined with the uneven spatial distribution of these limited manual ground observations, makes it challenging to fully characterise long-term crop phenology across a continuous spatial landscape [12]. On the other hand, the advent of satellite remote sensing has revolutionised the ability to obtain information about agricultural practices. Unlike traditional field surveys, remote sensing can considerably mitigate many of their limitations by providing wide spatial coverage, frequent temporal sampling, and improved operational configuration of modern satellites, thereby expanding the scope of crop phenological analysis well beyond what field surveys can achieve. Land surface phenology, or LSP, is the name given to phenology measurements obtained while using satellite-based observational data [13]. The detection of LSP itself basically relies on the observed variations of time series of remotely sensed vegetation indices (VIs), such as the normalized difference vegetation index (NDVI) [14], which has proven to be very effective in inferring the crop growth cycle and identifying the so-called LSP metrics, such as the start and end of growing season (SOS and EOS, respectively) [15]. Several studies have found that over croplands, LSP metrics can provide detailed information about crop calendar events (e.g., planting and harvesting dates) [11,16]. Due to the requirement of dense and repeated observations for LSP investigations, most satellite-based crop phenology studies have been based on moderate to coarse resolution sensors (e.g., Medium Resolution Imaging Spectrometer (MERIS), Moderate Resolution Imaging Spectroradiometer (MODIS), Advanced Very High Resolution Radiometer (AVHRR), Satellite Pour l'Observation de la Terre Vegetation (SPOT VGT)) as they provide frequent (almost daily) observations that are required during rapidly changing phenological stages [17–20]. However, at a spatial resolution of about 1 km, the pixels frequently include two or more types of crops and subsequently provide mixed VIs profiles, which might not be the best to fully characterise the spatial patterns of phenological variation relevant to a crop's important growth stages. Due to this, LSP has entered a new era with the availability of data from new missions such as Landsat 8 (L8) and Sentinel-2 (S2). Various studies have generated LSP metrics from the L8 satellite at 30 m resolution [21,22] but typically lack the temporal and sufficient spatial resolution to capture continuous crop development and reflect the local spatial heterogeneity of crop information. In this way, the Sentinel-2 mission can partially overcome this constraint by offering high spatial resolution of 10 m and a temporal resolution of either 10 days with one satellite or 5 days while using two-satellite constellation [23].

Although S2 boasts various enhanced design features compared to other low- to medium-spatial resolution satellites, in regions with persistent rain and clouds, these optical sensors are naturally impacted by cloud cover, aerosols, and other atmospheric contaminants [24]. This results in undesirable gaps and noise in current VIs products, which pose obstacles for subsequent high-level applications such as LSP analysis and extraction [25]. As a result, one of the essential processes that often precedes any LSP analysis is the construction of high-quality VIs time series using various filtering algorithms such as a Maximum-Value Composite (MVC) [26], the Savitzky–Golay (SG) filter [27,28], Asymmetric Gaussian (AG) function fitting [25], the best index slope extraction technique (BISE) [29], Double Logistic (DL) function fitting [30], HANTS [31,32], and the Whittaker smoother [33].

By smoothing the raw times series of VIs and producing high-quality VIs data, LSP detection can be effectively managed. However, accurate extraction of phenology, partic-

ularly for crops, is based on identifying valid crop phenological cycles [34]. Despite its theoretical simplicity, this task proves challenging in practice for several reasons. First, while natural vegetation shows a significant potential for periodicity in the variation of their *NDVI* seasonal pattern over the years, this is not true when aiming for a crop-specific analysis because multiple crop types with different phenological patterns are frequently planted in rotation on the same plot, resulting in the ability of a selected *NDVI* time series to demonstrate the feature of multiple peaks and cycles in a single year [35]. Second, the cultivation time period of a crop type is not always static but dynamically changeable from time to time in response to several variables such as climatic and environmental factors, crop varieties, and interannual variation of planting time, in addition to regional differences. Due to all of these factors, prior knowledge about crop cycle intervals may not be representative of the current crop being grown in a particular pixel; thus, the need arises to develop a robust pixel-based approach that can automatically and accurately determine the crop's growth cycle for each calendar year from the processed time series of VIs data.

Afterwards, the LSP parameters can be estimated. Currently, a variety of methods have been developed for the retrieval of phenological parameters from VIs time series. The main idea for most of these methods is to identify the turning points in the time series of satellite-derived VIs. However, they can be categorized into four broad categories: fixed and dynamic threshold, curvature and inflection, trend, and priori curve-based approaches [15]. Among these methods, the threshold-based approach has gained researchers' attention and has become the most widely used approach to estimate both SOS and EOS metrics due to its simplicity and efficiency in generating results within a reasonable time frame compared with other methods [15,36]. Threshold methods presume that a phenological stage starts when either the VIs reaches a fixed VIs value (fixed threshold method) or a given amplitude of the seasonal change in VIs values (dynamic threshold method) [37]. The latter has been widely used and confirmed to be superior to the fixed threshold method in comparative research because it is more flexible and depends on the magnitude of *NDVI* variations over the growing season [15]. However, it requires attempts to determine the appropriate and optimal percentage threshold, which limits its wide range of applications. Until now, only a few studies have attempted to determine the optimal threshold for estimating a crop's phenological events. For instance, ref. [36] has investigated the optimal thresholds for retrieving the SOS and EOS metrics of different crops using MODIS time series data. In another study illustrating a calibration procedure to systematically optimise the derivation of the phenometrics of four crops (green-up, heading, and senescence), ref. [38] also used MODIS data and a BISE filter combined with SG. A similar study in [39] analyzed different threshold levels (10%, 15%, 20%, and 25%) of SOS and EOS to map key stages of corn growth at a field scale with 30 m resolution, using fused surface reflectance data. With higher resolution, reference [40] similarly extracted S2 SOS/EOS from various amplitude thresholds of *NDVI*, *EVI2*, and a new index called the *PPI* (plant phenology index) value. They linked them to specific plant phenophases, finding that *PPI* SOS at 25% and *PPI* EOS at 15% provided the best matches with the phenological stages observed on the ground.

However, at this point, most crop phenology studies attempting to optimise phenological extraction are either limited to coarse satellite sensors, constrained to a few specific crop types, focused on small areas, or targeted to limited crop growth stages. These knowledge gaps have impeded our ability to connect high-resolution phenological metrics derived from *NDVI* to the stages of crop growth at the parcel level. Against this background, the primary aim of this study is to develop an innovative calibration procedure to accurately retrieve crop-specific phenophases at the field level of three major crops (winter wheat, corn, sugar beet) using high-quality reconstructed *NDVI* time series. Specifically, we seek to (1) introduce a robust method to filter the noise of the original *NDVI* time series for cloudy situations in Germany; (2) to determine the optimal dynamic thresholds for deriving the phenological metrics SOS and EOS based on the produced high-quality *NDVI* time series, making them more comparable and related to actual crop phenophases observations;

(3) to demonstrate the effectiveness of the selected dynamic thresholds by aligning the detected phenology measures with independently recorded phenophases observed on the ground; (4) to generate field-level crop-specific phenophase date maps for three federal states in Germany using the selected dynamic thresholds.

2. Study Area and Data

2.1. Study Area

This study was carried out over the federal territory of the north centre of Germany, which comprises the three German federal states of Lower Saxony (LS), Brandenburg (BB) and North Rhine-Westphalia (NW) (Figure 1). From a geological point of view, these federal states are part of the Northern Lowlands and Central Uplands of Germany [41]. With approximately 111,189 km², they comprise 31% of the territory of Germany and 32% of the agricultural area in Germany [42]. According to the most recent Köppen–Geiger classification [43], the climatic conditions range from warm and humid to continental with high average temperatures (more than 9 °C, on average) and a latitudinal precipitation gradient from an average of 823 mm/year in the west (NW and LS) to 564 mm/year in the extreme east (BB) [44]. These federal states were selected on the basis of the suitability and availability of the data.

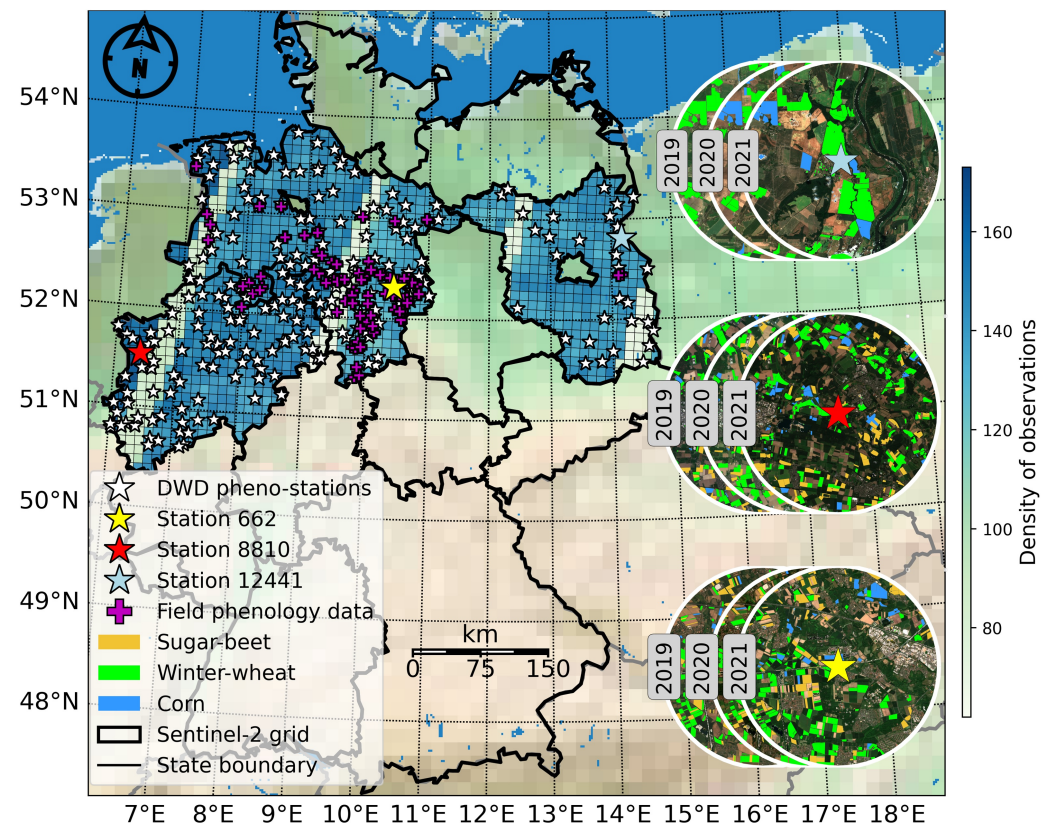


Figure 1. Overview of the study area spanning across three federal states in Germany. White stars indicate DWD phenological stations, and pink plus signs represent the phenological field sites used in this study. The black grid lines show the Sentinel-2 Analysis Ready Data (S2-ARD) acquisitions used in this study, with the background representing the number of observations in 2019 (as shown in the colour bar). Circles illustrate examples of the sampling design of agricultural field samples of three crops (winter wheat, corn, sugar beet) within a 5 km radius around three randomly selected DWD phenological stations for three years (2019, 2020, 2021).

2.2. Phenological In-Situ Data

Phenological data from 2019 to 2021 of three major crops (winter wheat, sugar beet and corn) were obtained from Germany's national meteorological service (DWD; in Ger-

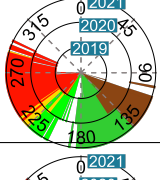
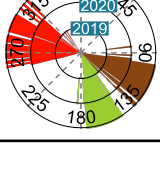
man: Deutscher Wetterdienst), which operates a phenological observation network across Germany with hundreds of stations, as shown in (Figure 1) [10]. Volunteer observers report 160 phenological phases of wild and cultivated plants immediately or annually, which are directly related to the main growth stages according to the Biologische Bundesanstalt, Bundessortenamt and CHemical Industry (BBCH) system [45]. Each crop was observed in a different number of stations, depending on the abundance and agrometeorological relevance of the respective crop type. And each phase of observation associated with a certain DWD station was considered representative of the crops that fall within a maximum radius of 5km of its location, as shown in Figure 1). It should be noted that the distribution of phenological stations is uneven, both across regions with varying environmental conditions and within regions with similar crop suitability. Moreover, the number and placement of observations fluctuate from year to year due to crop rotation practices [46].

DWD records from two years (2019–2020) were used to parameterize the phenological estimation, while the records for the year 2021 were used for the validation issue. Furthermore, we used additional in-situ BBCH information (2018–2022), with one portion collected from the “PatchCrop” experimental site managed by the Leibniz Centre for Agricultural Landscape Research [47] and another portion obtained through the PIAF database of the Plant Protection services of the German federal states where PIAF stands for the planning, information and evaluation system for trials in agriculture, viticulture, horticulture and agricultural research. PIAF is supervised by the Institute for Strategies and Technology Assessment at JKI.

These data will be used to further investigate the accuracy of phenological estimates, as they provide field-based information on the timing of all phases of BBCH estimations for several crops. Restricted by the availability of data, we utilise these data that are exclusively accessible for this study within the federal LS and BB states.

Table 1 lists all the phenological phases that were taken into consideration for the three main crops examined in this study based on DWD and BBCH, as well as a polar graph illustrating the distribution of each phase for each crop in our study areas from 2019 to 2021 as well as the number of in-situ observations.

Table 1. Phenological in-situ data collection summary, including crop types, DWD phases, BBCH Codes, and the number of observations for both BBCH and DWD. DWD Phase IDs are color-coded, and the DWD Phase Distribution column shows the distribution with polar graphs colored according to these IDs over three years.

Crop Type	BBCH Code	DWD Phase Description	DWD Phase ID	DWD Phase Distribution	N° of DWD Obs	N° of In-Situ BBCH Obs
Winter wheat	10	emergence	12		259	9
	31	shooting	15		206	15
	51	heading	18		264	2
	75	milk ripening	19		216	17
	87	yellow ripening	21		237	2
	99	harvest	24		286	27
Corn	10	emergence	12		416	11
	31	stem elongation	67		319	9
	75	milk ripening	19		274	4
	87	yellow ripening	21		227	2
	99	harvest	24		332	6
Sugar beet	10	emergence	12		125	7
	35	closed stand	13		120	6
	99	harvest	24		93	-

2.3. Agricultural Parcel Data

This study involves the use of a vector dataset that contains high spatio-temporal resolution information on agricultural parcels of the three studied crops from different years (2019–2021) in the study area indicated in Figure 1. The data set is collected and updated yearly through the Integrated Administration Control Systems (IACS) among the German Federal States. Access is limited due to privacy concerns in certain states. It consists mainly of the geometry of the agricultural fields and the types of crops planted in each cropping season. The parcel data for the three crops explored in this research and their respective years are outlined on the study area map (Figure 1).

2.4. Cloud Processing Platform and Satellite Data

In this paper, we take advantage of a cloud platform for accessing Earth observation data, on-demand processing, and the generation of a value-added product called CODE-DE (Copernicus Data and Exploitation Platform—Deutschland) [48]. This platform is a hybrid service that enables the provision of web access to a large number of satellite datasets and a geospatial database over Germany for viewing and downloading, while also offering cloud computing with virtual machines and network services in a highly scalable way. The entire workflows in this platform, including data preprocessing, modelling, and post-processing, were consistently implemented using Python programming language on a virtual machine provided on CODE-DE. The computing resources of this machine included 16 processor cores and 128 gigabytes of RAM.

The satellite data used in this study consisted of Copernicus Sentinel-2 analysis-ready datasets (ARD) that were built from S2 Level 2A [49], have already undergone a full pre-processing chain (atmospheric correction, cloud masking, etc.), and are consistently tiled into 4212 $10 \times 10 \text{ km}^2$ non-overlapping square tiles (Figure 1), whereas each tile is made up of a raster dataset that contains 10 atmospherically corrected cloud-masked spectral bands (Blue, Green, Red, RedEdge1, RedEdge2, RedEdge3, NIR 10 m, NIR 20 m, SWIR1, SWIR2) of the S2 sensor. Furthermore, each tile contains specific information that was gleaned from the scene classification layer provided in the Level 2 product regarding the percentage of pixels inside the tile that were impacted by clouds, snow, or other factors [50]. For the present study, a valid pixel filter was considered in the collection of S2-ARD data, in which scenes with less than 10% valid pixels were removed from the analysis. The S2-ARD image collection was used to create *NDVI* layers according to (1), where *NIR* and *RED* represent the near-infrared band and the red band reflectance values (B8 and B4 in S2, respectively). These layers were then used to create an annual time series of *NDVI* for each tile in the study region and further to extract crop-specific *NDVI* profiles for each field parcel from agricultural parcel data while using a 10 m internal buffer to reduce the influence of heterogeneous pixels at field boundaries. These calculated *NDVI* time series will serve as the basis for further processing and analysis.

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (1)$$

3. Methods

Our study focuses on integrating DWD and BBCH phenological observations, crop type data, and geometries with Sentinel-2 ARD time series. We begin by generating high-quality *NDVI* time series using a robust filtering approach, termed UE-WS, which combines upper envelope detection (UE) and Whittaker smoothing (WS) to filter out and replace erroneous data points before any phenological analysis. Next, we calibrate and determine optimal thresholds for the detection of different phenological stages. We then assessed the accuracy of the detection results using these chosen thresholds against independent ground observation data. Finally, we mapped the phenological detection outcomes at a regional scale. This flow chart shown in Figure 2 gives a general picture of the methodological approach used in this study.

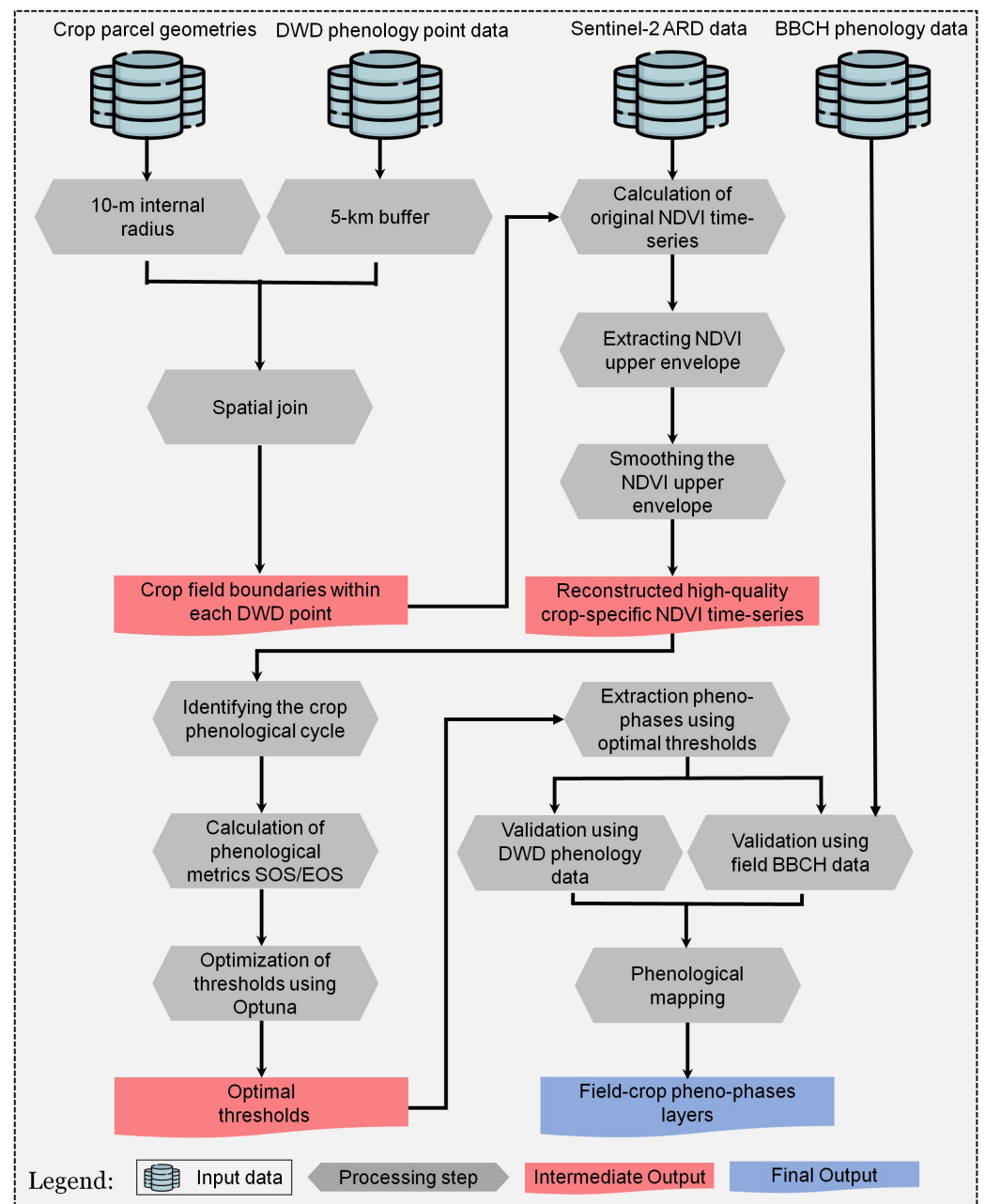


Figure 2. Flowchart illustrating the overall methodological workflow of this study.

3.1. High-Quality Crop-Specific NDVI Time-Series Reconstruction

To effectively reconstruct high-quality crop-specific *NDVI* time series profiles, two facts are noteworthy in this regard. First, *NDVI* is often underestimated for noises such as atmospheric effects. Second, vegetation growth typically remains stable, exhibiting gradual changes in the absence of disturbance. Even abrupt events such as crop harvesting can cause sudden decreases in the *NDVI* values, introducing fluctuations in the *NDVI* time series. However, occurrences where an observation shows a sharp drop shortly followed by a rapid increase, forming a deep V-shape, are quite rare. Therefore, instances of consecutive sudden decreases and increases in a time series are primarily considered noise. Considering these assumptions, the objectives of the proposed UE-WS method are to more effectively handle noisy observations, improve the accuracy of *NDVI* data reconstruction, and maintain the integrity of multi-temporal *NDVI* profiles so that they accurately reflect the natural growth stages and seasonal variations of the studied crops. There are two main steps to achieve these goals using the UE-WS method (Figure 2). We first extract the upper

envelope nodes of the *NDVI* time series in the backward and forward directions, then smooth the retrieved upper envelope to minimise the influence of residual noise.

3.1.1. Extracting *NDVI* Upper Envelope

Regarding the fact that most noise in *NDVI* time series is negatively biased, in the initial step, we attempt to reconstruct the upper envelope of the *NDVI* time series while preserving as much of the original data as possible using a number of strategies inspired by [51]. The first step involved the identification of all local maximum values across and starting from the first day of the *NDVI* time series. In this process, a local maximum is defined as a point position M_t with a value M_i that is the largest in the interval $(M_t - h, M_t + h)$. The search interval for the local maximum was set to be small in such a way that the M_i value is larger than its two direct neighbouring points. The second step is the identification of the other points that are highly likely to form the upper envelope of the *NDVI* time series but have not been recognised as local maxima in the first step. According to the assumptions mentioned above, a good quality *NDVI* point usually follows the gradual process of vegetation dynamics, implying that a sudden change in temporal values cannot be caused by natural conditions changes; thus points with rapid change were considered outliers, while those with a low rate of change were preserved to join the sets of local maxima to form the full upper envelope. The approach consists of using a very simple discriminant function that chronologically passes through the times series points and classifies the low-value noises based on the sets of local maxima found in the first step. The discriminant function of time (t) can be expressed according to Equation (2).

$$f(t) = \left(\frac{\sigma}{\sigma + 1} \right)^{\text{int}(t-i)} M_i \quad (2)$$

σ is the attenuation coefficient used to control the rate of decline in the discriminant function, and $t - i$ is the temporal distance between $NDVI_t$ and M_i . $NDVI_t$ is the original *NDVI* value of the point at time t , and M_i denotes the value of the last local maxima to time t . Selecting a proper attenuation coefficient σ is very important because choosing one too small can result in the identification of low *NDVI* points as envelope nodes and vice versa. Initially, the σ value is expressed as a percentage ranging from 1 to 100. To determine the optimal σ value, we tested various options for each crop due to their unique characteristics of the *NDVI* time series, which require different levels of filtering. We start with a reasonable range around 60, using a margin of ± 20 and an increment of 2, because 60 was identified as a strong hypothesis for the appropriate degree in [51]. Based on these tests and visual interpretations, a set of $\sigma = 40$, $\sigma = 50$, and $\sigma = 66$, respectively, for corn, winter wheat, and sugar beet appears to be appropriate. To obtain the rest of the upper envelope nodes, each $NDVI_t$ point from all the *NDVI* values was inserted as a new M_i in every iteration if it satisfied the following rule in Equation (3).

$$NDVI_t - f(t) \geq 0 \quad (3)$$

The points that did not meet this condition were excluded from the time series vector and substituted with non-data values. Subsequently, we filled the resulting gaps in the time series by linear interpolation between adjacent reliable values over time, resulting in a continuous time series vector. Finally, two analysis ensembles of the model state are derived in time t from the forward and backward analysis, respectively. In doing so, the function compares the values of the two curves in pairs and returns a final vector that contains the maximum values from each pair.

3.1.2. Smoothing the *NDVI* Upper Envelope

Although ordinary noises can be efficiently removed by extracting the upper envelope of the *NDVI* time series, under some circumstances, some sudden peaks, spikes, and fluctuations may continue, which could result in an erroneous performance in phenological

analysis. To eliminate such noise, the upper envelope *NDVI* time series generated in the first step was smoothed to a daily time step using the Whittaker filter [33], which is a computationally efficient reconstruction method for smoothing and filling gaps. This filter depends on a single adjustable parameter, λ , that penalises the time series roughness; a larger λ corresponds to an over-smoothing of the maximum curvature points. After trying several combinations, in this study a λ value equal to 100 was used, as it is adequate to obtain more convenient results.

3.2. Phenological Analysis

3.2.1. Identifying Crop Phenological Cycles

To automatically identify the valid phenological cycle corresponding to the *NDVI* profile for a given parcel, we proceed as follows.

1. We defined the time periods during which we expect that a specific crop should generally grow by analyzing the crop calendars in Germany for the previous years (Table 1).
2. All the local maxima were identified from the reconstructed *NDVI* time series corresponding to the crop window selected in step 1.
3. An ensemble of hyperparameters was used to filter out the erroneous maxima. The first one uses a specific threshold of 0.4 to remove the lower values because several studies suggested that the maximum *NDVI* values for crops generally exceed 0.4, as suggested in other studies [52]. We also used a search window of 30 days in this study. If multiple peaks can still be identified, we rely on the peak's prominence, which is a popular cue that indirectly quantifies how much a peak stands out from its surroundings.
4. Once the local maxima of our crop had been extracted, a similar approach was adopted to find out if the left and right minima could then be found for this corresponding phenological cycle.
5. Lastly, we employed the selected *NDVI* local maxima along with the left and right minima, and the area they span, to define the crop's phenological cycle.

3.2.2. Calculation of Phenological Metrics

To derive the timing of SOS and EOS phenological metrics in this study, we chose to follow the widely adopted dynamic threshold approach [37]. The basic principle of this method involves first normalising the data of the time series *NDVI* (which ranges from -1 to 1) into an $NDVI_{ratio}$ of the phenological cycle amplitude of *NDVI* (with values between 0 and 1). Subsequently, this $NDVI_{ratio}$ is used to establish dynamic thresholds for SOS and EOS retrieval. Specifically, when $NDVI_{ratio}$ exceeds or is below a certain threshold, the corresponding day of the year is determined as the SOS/EOS (see an example of different thresholds in Figure 3a). The $NDVI_{ratio}$ was calculated using Equation (4).

$$NDVI_{ratio} = \frac{NDVI_t - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \quad (4)$$

where $NDVI_t$ refers to the *NDVI* value on day t , $NDVI_{min}$ and $NDVI_{max}$ are the maximum and minimum *NDVI* values in the phenological cycle, respectively.

3.2.3. Optimal Parameterisation of Phenological Thresholds

Optimising the threshold in Equation (4) depends on the specific phenophase and the type of crop that is targeted. For example, a lower threshold that effectively identifies the emergence stage in winter wheat might overlook the heading stage of the same crop or could inaccurately detect the emergence phase of another crop. Therefore, it is crucial to calibrate and optimise these thresholds to accurately represent a particular phenological stage. This optimisation process ensures that the thresholds are well-suited to explain the targeted phenophase, such as emergence or heading, for a given crop type.

To achieve this, we have developed a two-step framework that optimises these thresholds using ground-phenological observations. This framework incorporates a novel optimisation method called Optuna [53] to obtain the optimal threshold for capturing a crop-specific phenological stage. The use of Optuna during the present study was due to its status as a next-generation optimisation framework that leverages advanced techniques such as Bayesian optimisation, gradient-based optimisation, and evolutionary algorithms for hyperparameter optimisation, making it suitable to tackle a variety of optimisation problems quickly and efficiently. The initial step in the proposed optimisation framework reduced the search space by identifying a promising range for the optimal thresholds for both SOS and EOS, examining values in intervals of 0.1 from 0 to 1. Subsequently, the second step fine-tunes and optimises the thresholds within the narrowed range determined in the first step, using steps of 0.01 from the initial threshold point -0.1 to the initial threshold point $+0.1$. This process is summarised in Figure 3b.

As mentioned in Section 2.2, the phenological stages linked to each DWD station are regarded as representative of the crops situated within a 5 km radius of the station's location. In light of this and to guarantee the calibration of each threshold using ground phenology observations from DWD, we computed the temporal differences Δt between the observed DOY of DWD and each estimated DOY of the phenophases for the parcels encompassed within the 5 km radius, according to the methodology introduced by [54] and as in Equation (5).

$$\Delta t = \Delta_{\text{obs}} - \Delta_{\text{est}} \quad (5)$$

Subsequently, the median values of the temporal difference Δt were calculated from the differences for each field within the station buffers for all investigated years. The median Δt values were used as measures of the accuracy of estimated phenophases. To demonstrate the principle of this optimisation process, we show a graphical representation of it for one DWD station in Figure 3. This procedure is repeated for all DWD stations, and the optimal threshold is returned when the lowest median Δt value is obtained between the DOY derived from the satellite and the observed DOY at all available DWD stations. Furthermore, the optimal threshold is compared between the SOS and the EOS to determine which yields the lower error in terms of Δt . This helps in judging which metric to use for each crop-specific phenophase.

3.2.4. Assessment of Pheno-Phases Detection Using Optimal Thresholds

To independently quantify the detection capability of the optimal thresholds obtained for each crop-specific phenophase, we adopted two schemes. The first scheme involved comparing crop phenology results at the DWD station level, while the second assessed the spatial consistency and details at the field level.

The first scheme involves assessing the performance of our optimal thresholds statistically using phenological observations from the hold-out year as a validation dataset (2021). This includes measuring the median of Δt following the same procedure previously described for the DWD data spanning 2019–2020. Additionally, to illustrate the significance of this optimisation process, a comprehensive analysis can be constructed, comparing the results of optimised and random thresholds with ground truth data. This visualisation highlights the importance of choosing the most suitable threshold, demonstrating how optimised thresholds lead to more accurate and reliable phenology retrievals.

For a more quantitative comparison with visual comparison, we aggregated in situ observations using the minimum bias approach proposed by [55,56], which involves selecting the in situ observation that has the minimum absolute difference from the reported DWD station. This method aims to minimise bias and ensure that the selected observations closely represent the conditions reported by the DWD station, enhancing the reliability of our analysis and visualisations. Moreover, the Empirical Cumulative Distribution Function (ECDF) and Kolmogorov–Smirnov (K-S) test were used to evaluate the accuracy of DOY phenophase estimations by comparing reference DOYs with DOYs predicted with minimum bias. This approach provides a visual and quantitative assessment of the

alignment between the two distributions [57]. The ECDFs allow for visual comparison, while the K-S test quantitatively measures the absolute maximum difference, yielding a p -value to indicate if the distributions differ significantly. This combined method not only allows for a detailed comparison but also helps to identify any systematic biases or inconsistencies in the prediction model, ensuring that the estimated DOYs are reliable and comparable to the actual reference DOYs [58].

The second scheme aimed to enhance confidence in the results by incorporating separate field observation BBCH data. Performance evaluation of these optimal thresholds was carried out by calculating the Root Mean Squared Error (RMSE), determination coefficient (R^2) and Mean Absolute Error (MAE) between the detection of phenology from the best threshold and the actual BBCH date value at each phenological stage and for each crop.

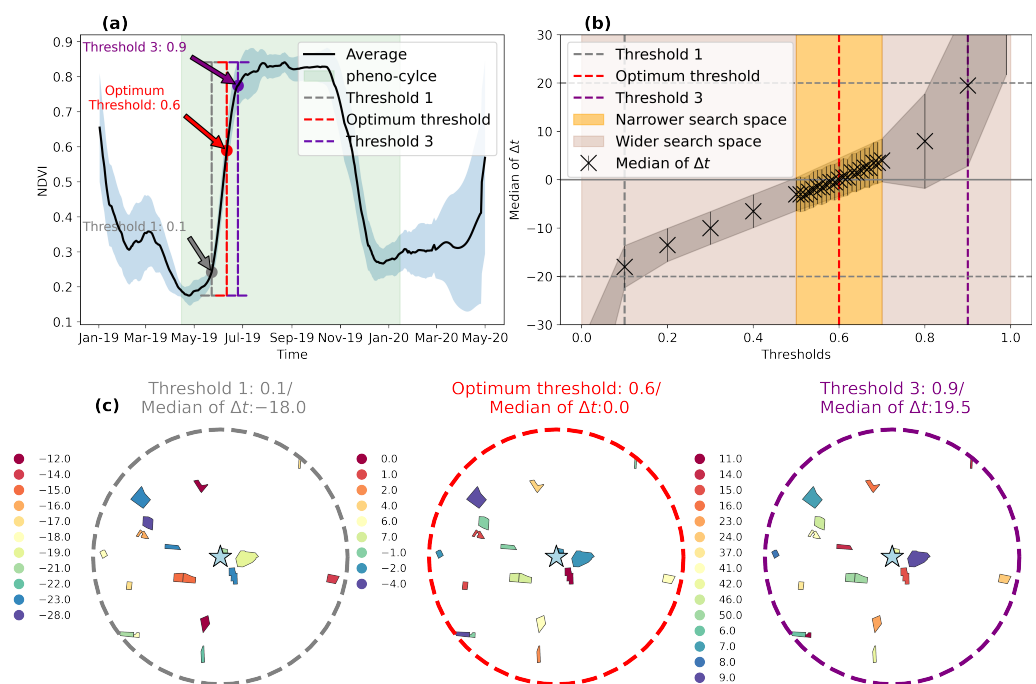


Figure 3. A graphical representation of the optimisation process. (a) The mean temporal profiles and standard deviation range of the $NDVI$ time series of all sugar beet fields in the DWD station 662 shown in Figure 1 with the different estimated DOYs of the closed stand phenophase according to three thresholds indicated by black, purple, and red dashed lines, corresponding to thresholds of 0.1, 0.9, and the optimal selected threshold, respectively. (b) Iterative optimisation of the median temporal difference error (Δt) of the closed stand phenophase of sugar beet estimation over thresholds from 0 to 1 using Optuna over all the fields in the DWD station 662 shown in (a). (c) Spatial representation of the temporal differences in days between the reported date from DWD station 662 (marked by the blue star) and the DOY estimates for all fields for the thresholds of 0.1, 0.9, and the optimal selected threshold.

4. Results

4.1. Sample Results of Reconstructed $NDVI$ Temporal Profiles and Identified Phenological Cycles

To assess how well our proposed approach performs in generating accurate $NDVI$ time series and reliably identifying crop phenological cycles, Figure 4a illustrates both the original and reconstructed $NDVI$ time series derived from our suggested methodology for three randomly sampled fields. The three samples were selected to be representative of the crop class studied within the 662 DWD station (cf. Figure 1), namely sugar beet, corn and winter wheat.

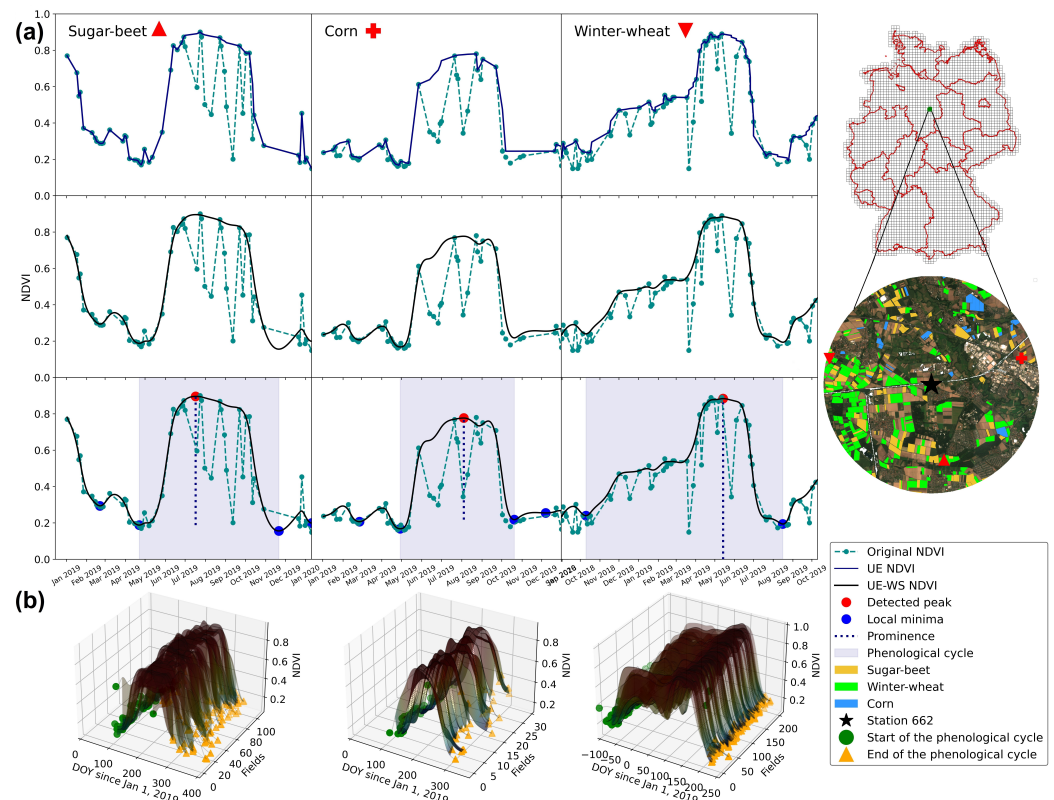


Figure 4. (a) Example temporal profiles of *NDVI* demonstrating the main stages of *NDVI* time series reconstruction using UE-WS for three crop samples at station 662. (b) Additionally, an example of identifying the phenological cycle for all crop parcels in DWD station 662 is illustrated in Figure 1.

Visual inspection of the three crops studied showed that the proposed UE-WS approach, which initially focuses on approximating the upper envelope of the *NDVI* time series and smoothing remaining noise, performed well in faithfully capturing distinct phenological seasonality from growth to dormancy for each crop and reconstructing high-quality *NDVI* time series that matched uncontaminated observations and adjusted contaminated ones.

Once the resulting high-quality *NDVI* time series from the compound approach UE-WS was obtained, we used the experimental setup described above in Section 3.2.1 to identify the phenological cycles of all parcels within DWD station 662 (Figure 4b). In general, the implemented approach for identifying phenological cycles proved effective in selecting appropriate start and end dates for candidate cycles for all crops and for almost all samples. The phenological cycles of the three crops, although different in time and date, have some similarities. For example, they are generally characterised by the beginning of the growing season, in which the *NDVI* generally rises sharply as the chlorophyll content increases in the tillering stage. In the following, period it no longer increases and in the ripening period, *NDVI* decreases as the plants dry out.

4.2. Optimisation of Phenological Phases

To determine the optimal crop-specific thresholds for each phenophase, we evaluated the performance of various thresholds by aligning the SOS and EOS dates with ground DWD phenological observations. This was carried out using the optimisation strategy detailed in Section 3.2.3, aiming to identify thresholds with the lowest error in a two-step process. Initially, it involved identifying the most suitable threshold interval, followed by a further refinement to detect the optimal threshold. The results of the optimisation procedure are shown in Figure 5.

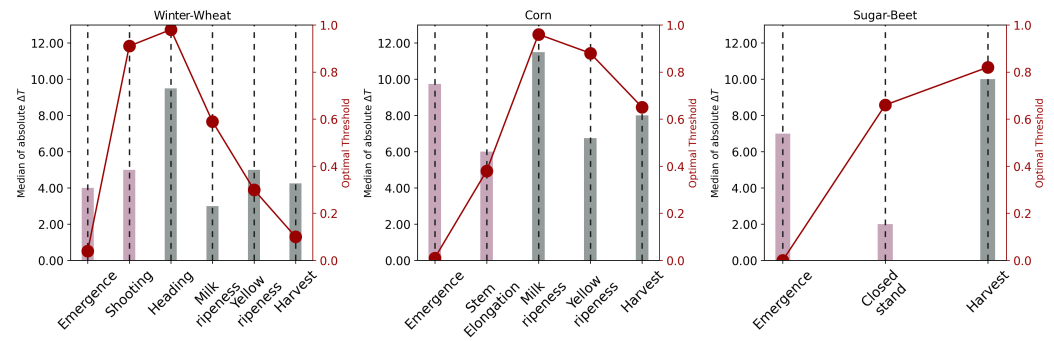


Figure 5. Bar plots representing the median of absolute Δt values across the obtained optimal $NDVI$ amplitude thresholds, ranging from 0 to 1 and indicated by a purple point. The pink bar represents the start of the season (SOS) and the grey bar represents the end of the season (EOS).

In general, optimal thresholds differed significantly across crops and phenophases, as well as depending on whether the metric for SOS or EOS was selected (Figure 5). Our findings show that early-stage phenophases were determined using SOS thresholds, while EOS thresholds proved to be more effective in identifying ripening and harvest stages. Regarding the detection of the heading stage for winter wheat, it is noted that typically, this stage is identified when the $NDVI$ reaches its maximum value which can be achieved using either SOS or EOS metrics, as both approaches capture high $NDVI$ values corresponding to the growth profile's peak. However, our findings indicate that the EOS metrics were particularly effective in identifying the heading stage. This distinction highlights the complementary roles of SOS and EOS metrics in the detection of phenophases throughout the growing season.

4.3. Validation of Phenological Estimates

The accuracy of the retrievals of phenophases using the optimal thresholds detected was further assessed against ground observation data from a separate year to assess the plausibility of the thresholds and the proposed approach. Figure 6 shows boxplots of the median of Δt between the phenophases derived from S2 using random thresholds (upper part) and the optimal ones (bottom part) compared to in situ measurements from DWD data in 2021 for the three crops under study. It can be observed that the phenological phases obtained using our proposed method and optimal thresholds correlate well with the phenophases reported by the DWD observations. This is in contrast to the results obtained using random or unoptimized thresholds. The interquartile range of the boxplots for optimum thresholds generally falls within the 10-day error range. Only three phenophases (emergence and milk ripeness of corn and harvest of sugar beet) have boxplots that extend slightly beyond the 10-day lines, yet their medians remain within this error range.

To better understand and evaluate the performance of thresholds obtained, we conducted an ECDF analysis. Figure 7 presents the comparisons of the ECDF graphs in different types of crops and phenophases, accompanied by the results of the K-S test. From the ECDF results, it is evident that the estimated DOYs align well with the reference minimum bias predicted DOYs, highlighting a strong agreement in the estimation approach. This is reflected in the K-S test values, which are considerably lower for almost all types of crops and phenophases investigated, except for the stem elongation phase of corn, which has a slightly higher KS value of 0.54. It should be noted that the lack of spatially explicit field locations, with only station locations available, may be the reason for influencing the error values.

To address this, we validated the phenological estimates derived using optimal thresholds against ground-truth field phenological observations. Figure 8 shows the scatterplots of the dates of phenophases obtained from field observations with those derived from satellite data. Each point represents a BBCH record of an agricultural field. For all crop phenophases considered here, the scatterplots show that the data points are closely aligned

with the 1:1 line, with most falling within the ± 10 days range. This distribution indicates that the optimal thresholds perform well in estimating field-level phenophase dates. The RMSE values are just under 10 days, specifically 9.39 days for winter wheat, 9.53 days for corn, and 9.69 days for sugar beet. Additionally, the R-squared values are relatively high, 0.97 for winter wheat, 0.96 for corn, and 0.89 for sugar beet, reflecting a strong correlation between predicted and actual phenophase dates.

The retained optimal thresholds were subsequently used to identify and map all crop-phenophase fields for the three years under study. Due to space constraints, Figure 9 illustrates the spatial distribution of the estimated DOY of only three phenological stages at the parcel level in the LS federal state.

In the various produced phenology maps either at the field level or state level, we can observe the differences in colour, indicating that the DOY to reach each phenological stage varies. There was less spatial variability observed for winter wheat and sugar beet with a certain continuity on the map, suggesting that most areas and their surrounding regions experience similar phenological stages. However, more spatial variability has been observed for corn, where this variation in space may be attributed to the difference in transplantation dates and management practices and also to soil and drainage properties. Exemplary temporal profiles for selected sample fields that have a BBCH observation, along with estimated phenophases, are also presented in Figure 9. This further illustrates the temporal effectiveness of our optimisation process in accurately identifying the correct phenophases.

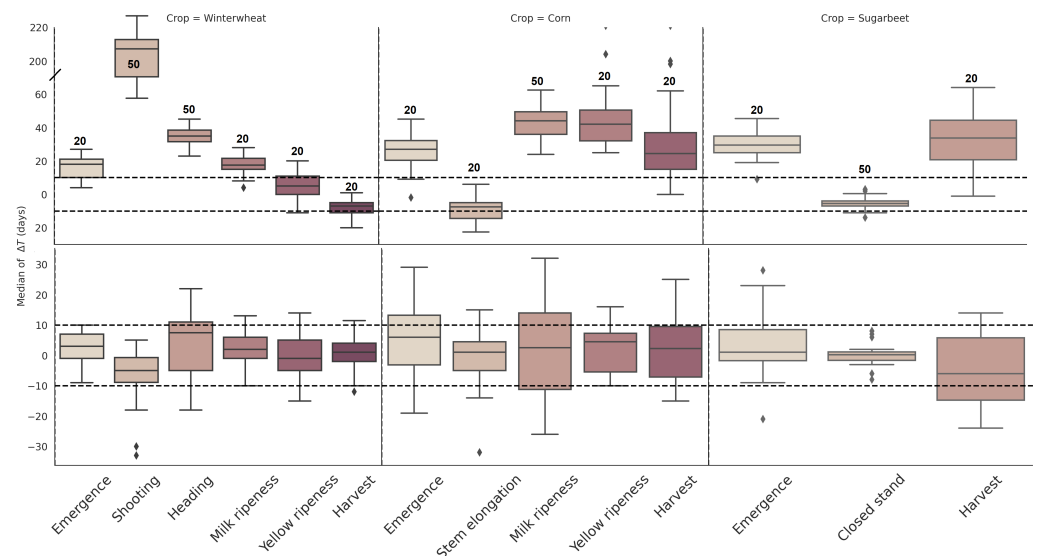


Figure 6. Boxplots of statistical agreement metrics median of Δt between satellite-derived phenophases dates and in-situ DWD dates across the three major crops for a separate year (2021). The top row of results was obtained using a fixed threshold approach, while the bottom set was derived after optimising thresholds.

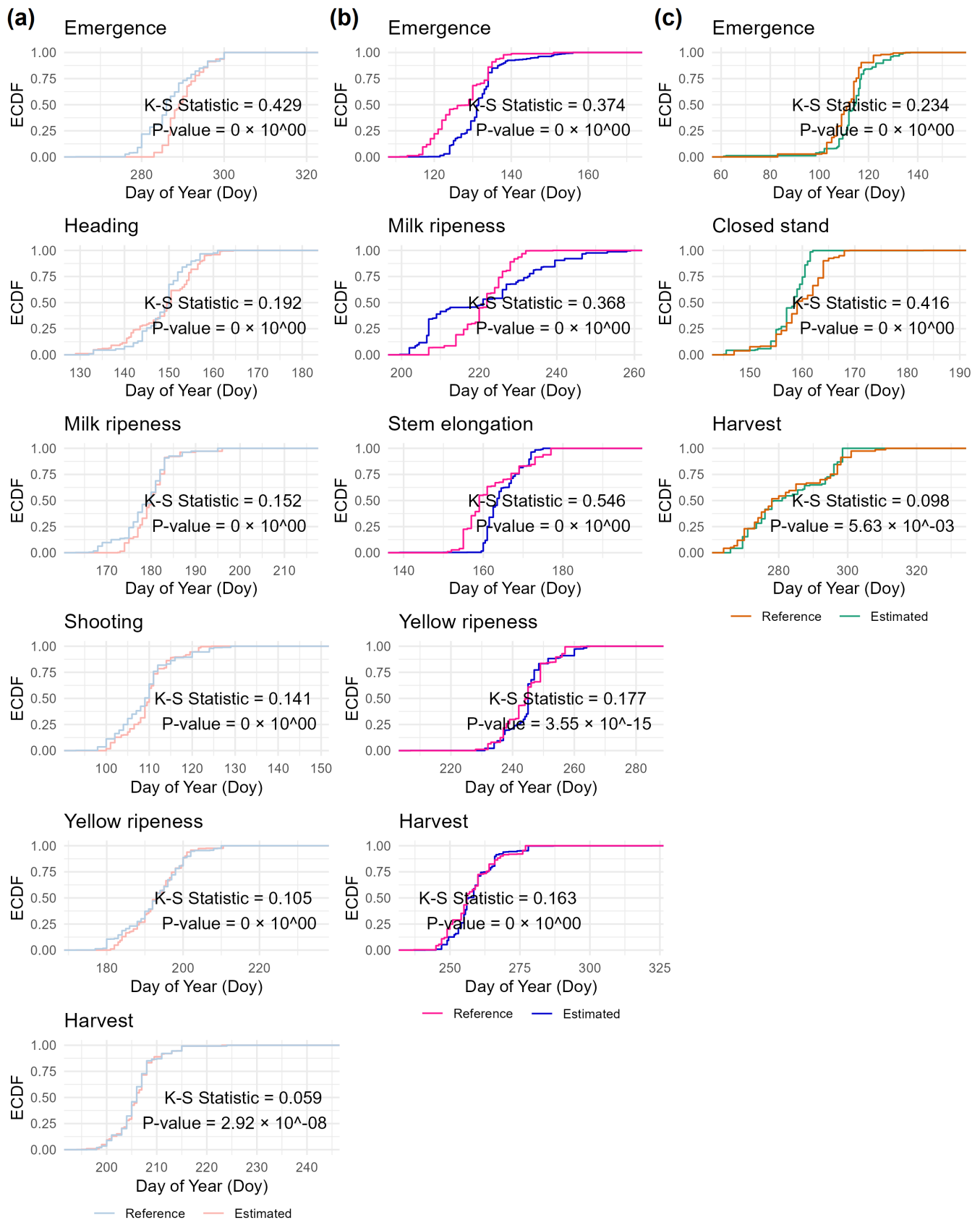


Figure 7. ECDF plot comparison of the reference DOYs with the aggregated minimum bias predicted DOYs for (a) winter wheat (b) corn, and (c) sugar beet.

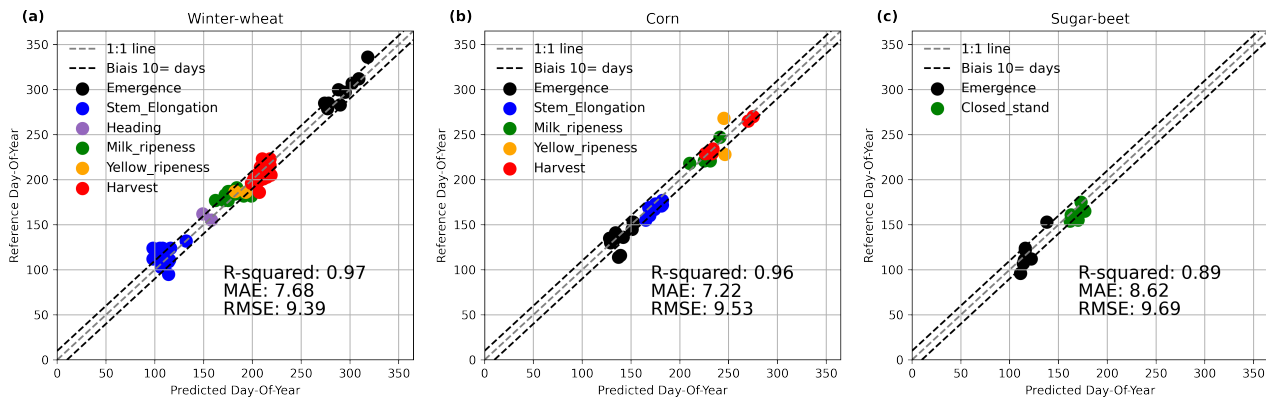


Figure 8. Comparison between field-referenced phenology and satellite-predicted phenological phases for (a) winter wheat (b) corn, and (c) sugar beet. The grey dashed lines represent the 1:1 lines, and the two black dashed lines are the deviation of ± 10 days from a perfect prediction.

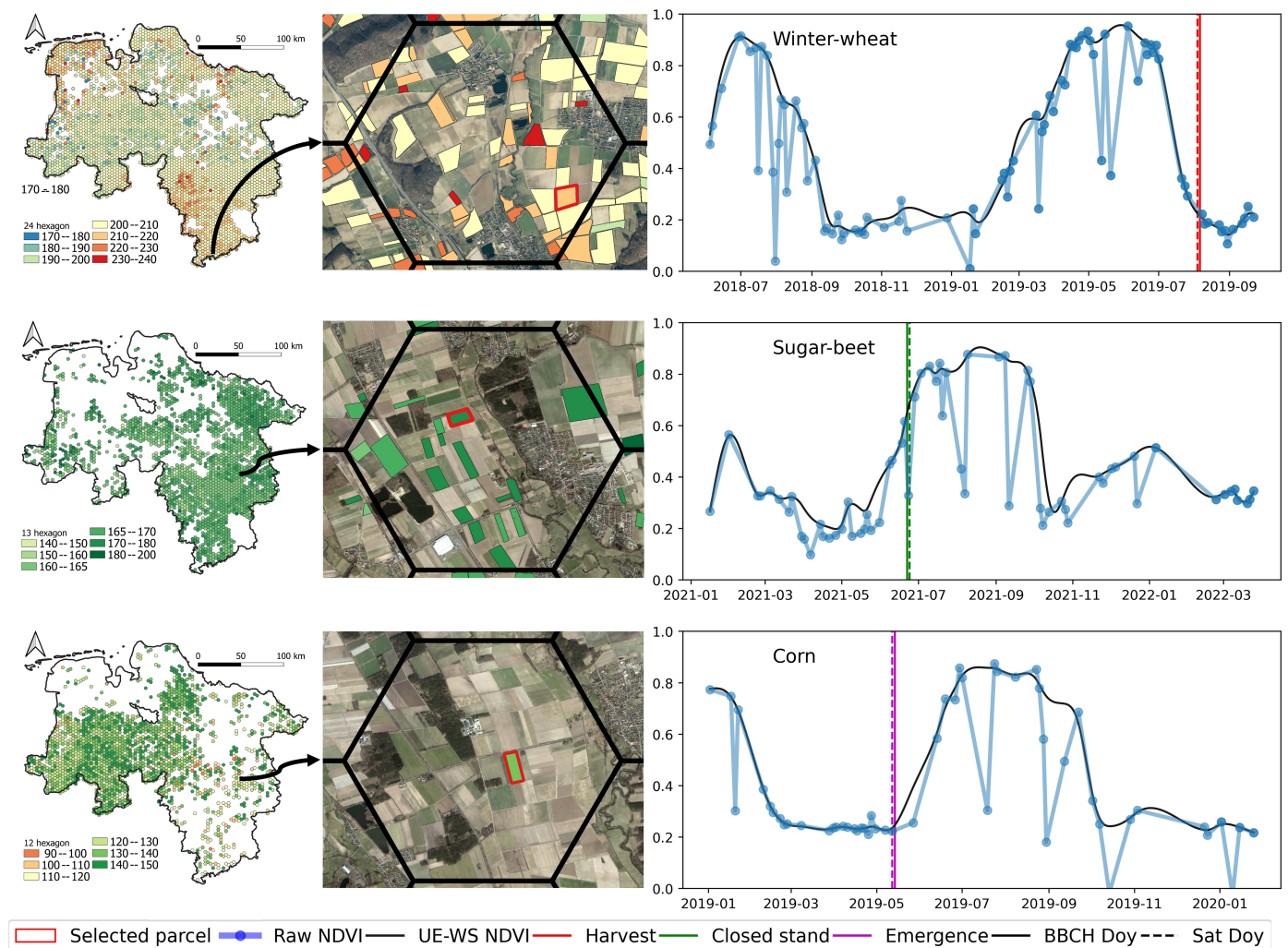


Figure 9. Predicted parcel-scale crop phenophases for three crop-specific stages (Harvest for winter wheat, Closed Stand for sugar beet, and Emergence for corn) aggregated into hex cells in the left panel. The middle and right panels display the locations of three example parcels (Each parcel is limited by a red outline) along with their NDVI temporal profiles. The vertical line marks the reference DOY, while the dashed line shows the satellite DOY estimates. Red indicates harvest, green represents closed stand, and magenta denotes emergence.

5. Discussion

5.1. NDVI Time Series Reconstruction

To ensure the reliability of optical remote sensing-based studies, particularly in cloudy and rainy regions, a basic step is to improve the data quality of *NDVI* time series by introducing a further pre-processing step that reduces noise caused by poor atmospheric conditions in the time series, and thus, key points within the growth cycle can be effectively captured [26]. In that light, the UE-WS method used in this study involves a robust approach to extract the upper envelope of the *NDVI* time series in order to remove outliers and then uses the WS filter to smooth random noise in essence. In our experiments, the sample results we present above clearly demonstrate the advantages of UE-WS when evaluated on the basis of general noise reduction under a variety of conditions, as illustrated in Figure 4a. This is not surprising, as the combination of another robust reconstruction method based on envelope detection and the SG filter has shown to be more stable for high temporal resolution *NDVI* time series reconstruction compared to conventional methods [51], and similar findings have also been reported in another study that proposed a two-level filtering approach for noise removal *NDVI* that employed the BISE and SG filter [59]. The question can be asked as to why a new type of approach might be needed since already existing ones have shown good performances. The answer is that it is not a matter of the nature of the method, but the type of problem for which it was designed, for instance, BISE and its improved method have proven effective at eliminating the influence of low-value noise caused by clouds and shadows; however, they were originally designed for daily and hypertemporal *NDVI* time series [60] and may perform poorly when there is a long-term, gradual decrease in *NDVI* as the one investigated in our study while using the S2 approach in a cloudy area such as Germany. Additionally, based on the reality that the phenology of vegetation is consecutive and gradual, upper envelope detection in other approaches [51] moves chronologically only forward along the time series *NDVI* while eliminating low-value noise and retaining high-value *NDVI* as envelope nodes. While, our approach provides the option to process the time series in both forward and backward directions, enhancing the flexibility and accuracy of extracting the upper envelope. Finally, the WS filter chosen to smooth the remaining speckles demonstrates high performance, as verified for time series reconstruction and vegetation phenology extraction [33,61].

5.2. Phenological Estimates and Optimisation

In this study, we proposed a full workflow for optimising the estimations of various crop phenophases using generated high-quality *NDVI* data. The process started with robust identification of the right phenological cycle of the corresponding crop field using the approach described in section 3.2 and illustrated samples in Figure 4b. Once the phenological cycles have been correctly identified, the phenological metrics (EOS/SOS) estimation becomes possible by setting a certain threshold on the time series. However, as has been discussed above, there is no single threshold that could be applied to all crop types, but in practice, the SOS/EOS threshold will depend on the crop type and targeted phenophases. Although the importance of such parameterization is high, most studies use thresholds that are mostly determined empirically, and there are surprisingly only a few studies that have tried to optimise the thresholds for phenology extraction [62]. Specifically, ref. [36] investigate the optimal thresholds for detecting crop phenophases by evaluating the retrieval accuracy of the SOS and EOS using different dynamic thresholds of VIs time series from MODIS data by steps of 5% from 0% to 35% for each of the crops. The same range of threshold values was tested in another study by [40] but on a finer scale using S2 phenology data. However, this approach is generally based on the estimation of vegetation phenology without focusing on specific crops. This contrasts with a study conducted in Germany [38] that obtained good accuracy in calibrating local *NDVI* thresholds from 0 to 1 with a step size of 0.01, but this was performed for only three crops and two phenostages per crop type with limited ground-truth data. Regarding all this, we propose a new optimisation framework that combines extensive ground-truth data from

DWD phenology observation and the high-quality produced *NDVI* data to calibrate the optimal thresholds using the Optuna framework in two steps from 0 to 1 for 14 phenostages and from three crops, as discussed in Section 3.2.3 and seen in Figure 3. Once the optimised thresholds have been derived, the assessment of satellite-based phenometrics using these optimized thresholds will be applied using two sources of ground-truth-observed phenology records: (1) a separate year of phenological DWD observations and (2) BBCH records from several ground-truth field measurements samples. In general, we found good agreement between satellite-based phenometrics and ground-truth observations for almost all the crop phenophases investigated (Figures 6 and 7). Given the differences in the study area, period, and methods used to compute crop phenophases and due to the scarcity of studies published to optimise *NDVI* thresholds, a direct comparison with other studies is not straightforward. However, some comparisons may be of value. In this context, our optimised thresholds seem to be consistent with some previous studies. The authors of [35] compare the phenology from the radar cross-polarization ratio VH/VV and from *S2 NDVI* with ground-truth phenological observations from DWD using three LSP metrics: peak season occurring when the function reaches 90% of the green-up amplitude (PS90), SOS extracted using the 50% thresholds (SOS50), and EOS extracted using the 50% thresholds (EOS50). Although the main objective of the study was not to compare several thresholds, it was found that in most cases, SOS50 was correctly associated with the early stages of BBCH (shooting, leaf development, and stem elongation). Except for wheat, where PS90 yields good results in predicting the BBCH stage (35, shooting), this would be close to the threshold we found with the optimisation in Figure 5 (90%). The study also found that the late secondary stages of the ripening stage for winter wheat and the harvest stage for corn were the closest stages to the EOS50 timing. This is in agreement with our findings, where our optimised thresholds vary around 50%. Ref. [38] also reported similar findings regarding, for instance, the green-up of (emergence) and the senescence phase (yellow ripeness) for winter wheat where the optimal thresholds found in this study that scored the lowest RMSE value are 0.1 and 0.41, respectively, which are close to the ones found in our study. However, this consistency was not the case for the emergence phase for sugar beet, in which [38] reported different findings. We believe that this slight difference could mainly depend on the nature of the sensor used, since we used a different sensor, and could also be based on the application of the step of identification of the cycle, which limits the cycle to a certain temporal domain; thus, the dynamic threshold approach that starts from the first day to the peak of the time series is different.

Also, another point to be noted is that the field-level crop-specific phenophase date maps generated, as shown in Figure 9, have the potential to significantly improve fine-scale agricultural research that was previously not applicable. These maps can enable a more precise timing of agricultural management measures, such as irrigation, fertilisation, and pest control, which are closely linked to specific phenophases, thus guiding agricultural production and strengthening food security [39]. Furthermore, by providing accurate field-level phenophase date maps and high-quality *NDVI* time series, this approach facilitates the classification of crop types, crop yield monitoring, and analysis of cropping system evolution to better support agricultural information management and food security [63,64].

5.3. Limitations and Way Forward

This study detected the phenophases of the crops using an optimised, robust, and unified approach. However, it still has some inherent limitations related to the data and methods, which are discussed in the following sections and could be attributed to several factors. First, the study covers only selected phenophases because the DWD only observed individual phenophases of the respective crops, rather than the entire growth cycle. In addition, biased or incorrectly reported information from DWD could lead to inaccuracies in representing the correct phenophases. This issue is common in volunteered geoinformation data, which, despite being a good source of free and abundant data, can be influenced by biases caused by volunteer observers, such as a lack of field identification

skills and geographic bias [65]. Specifically, ref. [66] found that 21% of the reported DWD phenological observations in Germany between 1959 and 2009 were questionable, with issues such as observations assigned to the wrong season or different phenological stages recorded on the same date for the same site and year. Therefore, adopting a new data filtering procedure could lead to filtering out these questionable observations and to increasing the accuracy of phenological observations [46,67]. Another consideration is that, according to the DWD instructions, each observation associated with a DWD station is considered representative of the crops within a maximum distance of 5 km from that location. However, some studies, such as [68], argue that the observations are based on at most four fields in the best cases. This means that the entry dates of the selected phenological stations could differ significantly when considering multiple fields for comparison, potentially worsening the results. Second, another limitation of the present study worth highlighting is the accuracy of parcel geometries of the crop types derived from IACS data, which is generally considered an excellent reference source. However, it can still contain errors due to inaccuracies in digitisation or false claims, and some farmers might not apply for subsidies [69]. Consequently, some outlier profiles might be attributed to a different type of crop, leading to incorrect detection of phenological estimations [54].

Building on the identified limitations, several key areas for future work and the way forward are crucial to advancing our approach and improving its performance: First, the UE-WS *NDVI* approach proposed in this study has already shown high performance in reducing noise and reconstructing high-quality *NDVI* profiles for phenological analysis. However, this *NDVI* time series might still be challenging for phenology detection if clouds persist over a field for an extended period or the frequency density of satellites is very low. This situation provides an excellent opportunity for spatiotemporal fusion to provide more informative data and support the current phenological analysis when temporally sparse Sentinel-2 observations occur [70] as well as the possibility to reconstruct and predict *NDVI* of cloud-contaminated images by establishing a relationship between SAR and optical images [71]. Second, although *NDVI* has been effective in estimating various crop-specific phenophases, there is a need to explore alternative vegetation indices that could address *NDVI*'s limitations, such as saturation. These alternative indices might offer more precise phenological information [40]. Third, while the study demonstrated strong regional and annual transferability of the thresholds obtained for most crops and phenophases within the available ground truth field measurements, limitations in reference data availability across different years and areas remain a bottleneck to performing a comprehensive assessment of this transferability, aiming to achieve high-precision estimates at the crop parcel level across Germany. Future research should focus on evaluating the performance of these thresholds in additional regions and years within the country. Additionally, more efforts are needed to collect more crop phenology data for robust validation and evaluation. This will support our approach of producing more integrated and comprehensive validation results and enhance our study's performance, applicability, and transferability.

6. Conclusions

Although many approaches have been proposed to retrospectively capture and map phenological phases, the calibration and optimisation of these phenophases that are specific to certain crop types and re-establishing their definition to be linked with specific agronomic stages of interest have rarely been addressed. This is mostly because it remains challenging to implement an optimised, robust, and unified approach to detect crop-specific phenophases on a regional scale from an optical remote sensing perspective in cloudy and rainy areas. In this context, reducing noise and reconstructing high-quality *NDVI* profiles for further time series analysis is a promising and effective approach to more reliably predict accurate crop phenophases. Using ARD Sentinel-2 time series as the core of this study, we proposed a compound approach based on envelope detection and a smoothing filter to robustly produce high-quality *NDVI* data and identify individual valid vegetation cycles before the phenological retrieval. The results revealed that the new method is effective for

reconstructing high-quality NDVI data to represent consistent phenological profiles of the different crops studied. Moreover, these high-quality NDVI time series data, along with extensive field information from collected in situ data and DWD phenology observations, were used together to calibrate and optimise the thresholds for extracting crop-specific land surface phenology. Our optimised thresholds provided enhanced capability in accurately estimating phenophases and achieved satisfactory accuracy that links more directly to in situ phenological observations. Moreover, the thresholds were used to predict field phenological information and thus illustrate more detailed spatial and temporal information. This study provides an advanced cloud platform-based phenology study case under different scenarios at landscape levels and provides technical support and application demonstration for field-level research, such as simulations of crop growth models and the analysis of the underlying reason for yield gaps at the field level.

Author Contributions: Conceptualization, A.H., M.M. and H.G.; methodology, A.H., M.M. and H.G.; software, A.H.; validation, A.H., T.R. and H.G.; formal analysis, A.H.; investigation, A.H.; data curation, F.B., T.R. and H.G.; writing—original draft preparation, A.H.; writing—review and editing, T.R., M.M. and H.G.; visualization, A.H.; supervision, H.G.; project administration, H.G.; funding acquisition, H.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research has received funding from the German Federal Ministry of Economic Affairs and Climate Action as part of the NaLamKI project under grant number 01MK21003E.

Data Availability Statement: No new data were created or analyzed in this study. Due to the sensitive nature and related data privacy regulations regarding phenological in-situ data from PIAF database is confidential and cannot be shared. Restrictions further apply to the availability of phenological in-situ data from patchCrop experimental site. Requests to access this data should be directed to Kathrin Grahmann from Leibniz Centre for Agricultural Landscape Research. All other data used in this study will be made available upon request.

Acknowledgments: We are very thankful to the Leibniz Centre for Agricultural Landscape Research (ZALF) for sharing data from the patchCROP experiment with us. The maintenance of the patchCrop infrastructure is supported by ZALF. We further thank the plant protection services of the German federal state for the provision of phenological data for validation of our results. We finally thank the operators of CODE-DE for providing a tailored EO cloud computing platform for public organisations.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lieth, H. *Purposes of a Phenology Book*; Springer: Berlin, Heidelberg, 1974; pp. 3–19. https://doi.org/10.1007/978-3-642-51863-8_1.
2. Möller, M.; Gerstmann, H.; Gao, F.; Dahms, T.C.; Förster, M. Coupling of phenological information and simulated vegetation index time series: Limitations and potentials for the assessment and monitoring of soil erosion risk. *Catena* **2017**, *150*, 192–205. <https://doi.org/10.1016/j.catena.2016.11.016>.
3. Strassmeyer, J.; Daehmlow, D.; Dominic, A.; Lorenz, S.; Golla, B. SYNOPSIS-WEB, an online tool for environmental risk assessment to evaluate pesticide strategies on field level. *Crop Prot.* **2017**, *97*, 28–44. <https://doi.org/10.1016/j.cropro.2016.11.036>.
4. Möller, M.; Doms, J.; Gerstmann, H.; Feike, T. A framework for standardized calculation of weather indices in Germany. *Theor. Appl. Climatol.* **2019**, *136*, 377–390. <https://doi.org/10.1007/s00704-018-2473-x>.
5. Abdi, A.M.; Carrié, R.; Sidemo-Holm, W.; Cai, Z.; Boke-Olén, N.; Smith, H.G.; Eklundh, L.; Ekroos, J. Biodiversity decline with increasing crop productivity in agricultural fields revealed by satellite remote sensing. *Ecol. Indic.* **2021**, *130*, 108098. <https://doi.org/10.1016/j.ecolind.2021.108098>.
6. Bucheli, J.; Dalhaus, T.; Finger, R. Temperature effects on crop yields in heat index insurance. *Food Policy* **2022**, *107*, 102214. <https://doi.org/10.1016/j.foodpol.2021.102214>.
7. Marrec, R.; Brusse, T.; Caro, G. Biodiversity-friendly agricultural landscapes – integrating farming practices and spatiotemporal dynamics. *Trends Ecol. Evol.* **2022**, *37*, 731–733. <https://doi.org/10.1016/j.tree.2022.05.004>.
8. Bucheli, J.; Conrad, N.; Wimmer, S.; Dalhaus, T.; Finger, R. Weather insurance in European crop and horticulture production. *Clim. Risk Manag.* **2023**, *41*, 100525. <https://doi.org/10.1016/j.crm.2023.100525>.
9. Riedesel, L.; Möller, M.; Horney, P.; Golla, B.; Piepho, H.P.; Kautz, T.; Feike, T. Timing and intensity of heat and drought stress determine wheat yield losses in Germany. *PLoS ONE* **2023**, *18*, e0288202. <https://doi.org/10.1371/journal.pone.0288202>.
10. Kaspar, F.; Zimmermann, K.; Polte-Rudolf, C. An overview of the phenological observation network and the phenological database of Germany's national meteorological service (Deutscher Wetterdienst). *Adv. Sci. Res.* **2014**, *11*, 93–99. <https://doi.org/10.5194/asr-11-93-2014>.

11. Gao, F.; Zhang, X. Mapping crop phenology in near real-time using satellite remote sensing: Challenges and opportunities. *J. Remote Sens.* **2021**, *2021*, 1–14. [https://doi.org/10.1016/s0034-4257\(02\)00096-2](https://doi.org/10.1016/s0034-4257(02)00096-2).
12. Song, Y.; Wang, J.; Yu, Q.; Huang, J. Using MODIS LAI data to monitor spatio-temporal changes of winter wheat phenology in response to climate warming. *Remote Sens.* **2020**, *12*, 786.
13. De Beurs, K.M.; Henebry, G.M. Land surface phenology, climatic variation, and institutional change: Analyzing agricultural land cover change in Kazakhstan. *Remote Sens. Environ.* **2004**, *89*, 497–509. <https://doi.org/10.1016/j.rse.2003.11.006>.
14. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150.
15. Zeng, L.; Wardlaw, B.D.; Xiang, D.; Hu, S.; Li, D. A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sens. Environ.* **2019**, *237*, 111511. <https://doi.org/10.1016/j.rse.2019.111511>.
16. Misra, G.; Cawkwell, F.; Wingler, A. Status of phenological research using Sentinel-2 data: A review. *Remote Sens.* **2020**, *12*, 2760. <https://doi.org/10.3390/rs12172760>.
17. Goward, S.N.; Tucker, C.J.; Dye, D.G. North American vegetation patterns observed with the NOAA-7 advanced very high resolution radiometer. *Vegetatio* **1985**, *64*, 3–14. <https://doi.org/10.1007/bf00033449>.
18. Jonsson, P.; Eklundh, L. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 1824–1832. <https://doi.org/10.1109/tgrs.2002.802519>.
19. Zhang, X.; Friedl, M.A.; Schaaf, C.B.; Strahler, A.H.; Hodges, J.C.; Gao, F.; Reed, B.C.; Huete, A. Monitoring vegetation phenology using MODIS. *Remote Sens. Environ.* **2003**, *84*, 471–475. [https://doi.org/10.1016/s0034-4257\(02\)00135-9](https://doi.org/10.1016/s0034-4257(02)00135-9).
20. Liu, Z.; Wu, C.; Liu, Y.; Wang, X.; Fang, B.; Yuan, W.; Ge, Q. Spring green-up date derived from GIMMS3g and SPOT-VGT NDVI of winter wheat cropland in the North China Plain. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 81–91. <https://doi.org/10.1016/j.isprsjprs.2017.05.015>.
21. Melaas, E.K.; Sulla-Menashe, D.; Gray, J.M.; Black, T.A.; Morin, T.H.; Richardson, A.D.; Friedl, M.A. Multisite analysis of land surface phenology in North American temperate and boreal deciduous forests from Landsat. *Remote Sens. Environ.* **2016**, *186*, 452–464. <https://doi.org/10.1016/j.rse.2016.09.014>.
22. Bolton, D.K.; Gray, J.M.; Melaas, E.K.; Moon, M.; Eklundh, L.; Friedl, M.A. Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. *Remote Sens. Environ.* **2020**, *240*, 111685. <https://doi.org/10.1016/j.rse.2020.111685>.
23. Graf, L.V.; Merz, Q.N.; Walter, A.; Aasen, H. Insights from field phenotyping improve satellite remote sensing based in-season estimation of winter wheat growth and phenology. *Remote Sens. Environ.* **2023**, *299*, 113860. <https://doi.org/10.1016/j.rse.2023.113860>.
24. Whitcraft, A.K.; Vermote, E.F.; Becker-Reshef, I.; Justice, C.O. Cloud cover throughout the agricultural growing season: Impacts on passive optical earth observations. *Remote Sens. Environ.* **2015**, *156*, 438–447. <https://doi.org/10.1016/j.rse.2014.10.009>.
25. Eklundh, L.; Jönsson, P. TIMESAT 3.1 Software Manual. Lund University: Lund, Sweden, 2012; pp. 1–82.
26. Holben, B.N. Characteristics of maximum-value composite images from temporal AVHRR data. *Int. J. Remote Sens.* **1986**, *7*, 1417–1434. <https://doi.org/10.1080/01431168608948945>.
27. Chen, J.; Jönsson, P.; Tamura, M.; Gu, Z.; Matsushita, B.; Eklundh, L. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sens. Environ.* **2004**, *91*, 332–344. <https://doi.org/10.1016/j.rse.2004.03.014>.
28. Savitzky, A.; Golay, M.J. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. <https://doi.org/10.1021/ac60214a047>.
29. Viovy, N.; Arino, O.; Belward, A. The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series. *Int. J. Remote Sens.* **1992**, *13*, 1585–1590. <https://doi.org/10.1080/01431169208904212>.
30. Beck, P.S.; Atzberger, C.; Høgda, K.A.; Johansen, B.; Skidmore, A.K. Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sens. Environ.* **2006**, *100*, 321–334. <https://doi.org/10.1016/j.rse.2005.10.021>.
31. Yang, G.; Shen, H.; Zhang, L.; He, Z.; Li, X. A moving weighted harmonic analysis method for reconstructing high-quality SPOT VEGETATION NDVI time-series data. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 6008–6021. <https://doi.org/10.1109/tgrs.2015.2431315>.
32. Cihlar, J. Identification of contaminated pixels in AVHRR composite images for studies of land biosphere. *Remote Sens. Environ.* **1996**, *56*, 149–163. [https://doi.org/10.1016/0034-4257\(95\)00190-5](https://doi.org/10.1016/0034-4257(95)00190-5).
33. Eilers, P.C. 2003. A Perfect Smoother. *Anal. Chem.* **2003**, *75*, 3299–3304. <https://doi.org/10.1021/ac034173t>.
34. Gray, J.; Sulla-Menashe, D.; Friedl, M.A. *User Guide to Collection 6 Modis Land Cover Dynamics (mcd12q2) Product*; NASA EOSDIS Land Processes DAAC: Missoula, MT, USA, 2019; Volume 6; pp. 1–8. Available online: https://lpdaac.usgs.gov/documents/1417/MCD12Q2_User_Guide_V61.pdf (accessed on 1 July 2023).
35. Meroni, M.; d’Andrimont, R.; Vrieling, A.; Fasbender, D.; Lemoine, G.; Rembold, F.; Seguini, L.; Verhegghen, A. Comparing land surface phenology of major European crops as derived from SAR and multispectral data of Sentinel-1 and-2. *Remote Sens. Environ.* **2021**, *253*, 112232. <https://doi.org/10.1016/j.rse.2020.112232>.
36. Huang, X.; Liu, J.; Zhu, W.; Atzberger, C.; Liu, Q. The optimal threshold and vegetation index time series for retrieving crop phenology based on a modified dynamic threshold method. *Remote Sens.* **2019**, *11*, 2725. <https://doi.org/10.3390/rs11232725>.
37. White, M.A.; Thornton, P.E.; Running, S.W. A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Glob. Biogeochem. Cycles* **1997**, *11*, 217–234. <https://doi.org/10.1029/97gb00330>.

38. Xu, X.; Conrad, C.; Doktor, D. Optimising phenological metrics extraction for different crop types in Germany using the moderate resolution imaging spectrometer (MODIS). *Remote Sens.* **2017**, *9*, 254. <https://doi.org/10.3390/rs9030254>.
39. Gao, F.; Anderson, M.C.; Zhang, X.; Yang, Z.; Alfieri, J.G.; Kustas, W.P.; Mueller, R.; Johnson, D.M.; Prueger, J.H. Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. *Remote Sens. Environ.* **2017**, *188*, 9–25. <https://doi.org/10.1016/j.rse.2016.11.004>.
40. Tian, F.; Cai, Z.; Jin, H.; Hufkens, K.; Scheifinger, H.; Tagesson, T.; Smets, B.; Van Hoolst, R.; Bonte, K.; Ivits, E.; et al. Calibrating vegetation phenology from Sentinel-2 using eddy covariance, PhenoCam, and PEP725 networks across Europe. *Remote Sens. Environ.* **2021**, *260*, 112456. <https://doi.org/10.1016/j.rse.2021.112456>.
41. Neue Anforderungen im europäischen Naturschutz: Das Schutzgebietssystem Natura 2000 und die FFH-Richtlinie der EU. *Nat. Landsch.* **1994**, *69*, 395–406.
42. Statistisches Bundesamt (Destatis) 2022 Land-und Forstwirtschaft, Fischerei: Bodenfläche nach Art der tatsächlichen Nutzung. Available online: <https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Landwirtschaft-Forstwirtschaft-Fischerei/Flaechennutzung/Publikationen/Downloads-Flaechennutzung/bodenflaechennutzung-2030510217005.html> (accessed on 1 July 2023).
43. Kotteck, M.; Grieser, J.; Beck, C.; Rudolf, B.; Rubel, F. World map of the Köppen-Geiger climate classification updated. *Meteorol. Z.* **2006**, *15*, 259–263. <https://doi.org/10.1127/0941-2948/2006/0130>.
44. Diers, M.; Weigel, R.; Leuschner, C. Both climate sensitivity and growth trend of European beech decrease in the North German Lowlands, while Scots pine still thrives, despite growing sensitivity. *Trees* **2023**, *37*, 523–543. <https://doi.org/10.1007/s00468-022-02369-y>.
45. Meier, U.; Bleiholder, H.; Buhr, L.; Feller, C.; Hack, H.; Heß, M.; Lancashire, P.D.; Schnock, U.; Stauß, R.; Van Den Boom, T.; et al. The BBCH system to coding the phenological growth stages of plants—history and publications. *J. Kult.* **2009**, *61*, 41–52.
46. Gerstmann, H.; Doktor, D.; Gläßer, C.; Möller, M. PHASE: A geostatistical model for the Kriging-based spatial prediction of crop phenology using public phenological and climatological observations. *Comput. Electron. Agric.* **2016**, *127*, 726–738. <https://doi.org/10.1016/j.compag.2016.07.032>.
47. Grahmann, K.; Reckling, M.; Hernandez-Ochoa, I.; Donat, M.; Bellingrath-Kimura, S.; Ewert, F. Co-designing a landscape experiment to investigate diversified cropping systems. *Agricultural Systems* **2024**, *217*, 103950. <https://doi.org/10.1016/j.agry.2024.103950>.
48. Benz, U.; Banovsky, I.; Cesarz, A.; Schmidt, M. CODE-DE Portal Handbook, Version 2.0. 2020. Available online: <https://code-de.org/en/> (accessed on 15 October 2021).
49. ESA. SENTINEL-2 User Handbook. Standard Document, Issue 1 Rev 2. 2015. Available online: https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook (accessed on 1 December 2022).
50. Main-Knorn, M.; Pflug, B.; Louis, J.; Debaecker, V.; Müller-Wilm, U.; Gascon, F. Sen2Cor for sentinel-2. In Proceedings of the Image and Signal Processing for Remote Sensing XXIII, Warsaw, Poland, 4 October 2017; pp. 37–48. <https://doi.org/10.1117/12.2278218>.
51. Liu, X.; Ji, L.; Zhang, C.; Liu, Y. A method for reconstructing NDVI time-series based on envelope detection and the Savitzky-Golay filter. *Int. J. Digit. Earth* **2022**, *15*, 553–584. <https://doi.org/10.1080/17538947.2022.2044397>.
52. Sakamoto, T.; Yokozawa, M.; Toritani, H.; Shibayama, M.; Ishitsuka, N.; Ohno, H. A crop phenology detection method using time-series MODIS data. *Remote Sens. Environ.* **2005**, *96*, 366–374. <https://doi.org/10.1016/j.rse.2005.03.008>.
53. Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A next-generation hyperparameter optimization framework. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; pp. 2623–2631. <https://doi.org/10.1145/3292500.3330701>.
54. Hoque, A. Variability of Wheat Phenology from Sentinel-1 and-2 Time Series: A Case Study for Brandenburg, Germany. Master's Thesis, University of Twente, Enschede, The Netherlands, 2022.
55. Ye, Y.; Zhang, X.; Shen, Y.; Wang, J.; Crimmins, T.; Scheifinger, H. An optimal method for validating satellite-derived land surface phenology using in-situ observations from national phenology networks. *ISPRS J. Photogramm. Remote Sens.* **2022**, *194*, 74–90. <https://doi.org/10.1016/j.isprsjprs.2022.09.018>.
56. Lobert, F.; Löw, J.; Schwieder, M.; Gocht, A.; Schlund, M.; Hostert, P.; Erasmi, S. A deep learning approach for deriving winter wheat phenology from optical and SAR time series at field level. *Remote Sens. Environ.* **2023**, *298*, 113800. <https://doi.org/10.1016/j.rse.2023.113800>.
57. Thas, O. *Comparing Distributions*; Springer Series in Statistics; Springer: New York, NY, USA; London, UK, 2010. OCLC: Ocn401153891.
58. Möller, M.; Koschitzki, T.; Hartmann, K.J.; Jahn, R. Plausibility test of conceptual soil maps using relief parameters. *CATENA* **2012**, *88*, 57–67. <https://doi.org/10.1016/j.catena.2011.08.002>.
59. Rahman, M.S.; Di, L.; Shrestha, R.; Eugene, G.Y.; Lin, L.; Kang, L.; Deng, M. Comparison of selected noise reduction techniques for MODIS daily NDVI: An empirical analysis on corn and soybean. In Proceedings of the 2016 Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Tianjin, China, 18–20 July 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–5. <https://doi.org/10.1109/agro-geoinformatics.2016.7577661>.
60. Li, S.; Xu, L.; Jing, Y.; Yin, H.; Li, X.; Guan, X. High-quality vegetation index product generation: A review of NDVI time series reconstruction techniques. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102640. <https://doi.org/10.1016/j.jag.2021.102640>.

61. Atkinson, P.M.; Jeganathan, C.; Dash, J.; Atzberger, C. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. *Remote Sens. Environ.* **2012**, *123*, 400–417. <https://doi.org/10.1016/j.rse.2012.04.001>.
62. Wang, S.; Chen, J.; Shen, M.; Shi, T.; Liu, L.; Zhang, L.; Dong, Q.; Wang, C. Characterizing spatiotemporal patterns of winter wheat phenology from 1981 to 2016 in North China by improving phenology estimation. *Remote Sens.* **2022**, *14*, 4930. <https://doi.org/10.3390/rs14194930>.
63. Bastidas, A.; Setiyono, T.; Dobermann, A.; Cassman, K.G.; Elmore, R.W.; Graef, G.L.; Specht, J.E. Soybean sowing date: The vegetative, reproductive, and agronomic impacts. *Crop Sci.* **2008**, *48*, 727–740.
64. Sakamoto, T.; Gitelson, A.A.; Arkebauer, T.J. MODIS-based corn grain yield estimation model incorporating crop phenology information. *Remote Sens. Environ.* **2013**, *131*, 215–231.
65. Sullivan, B.L.; Wood, C.L.; Iliff, M.J.; Bonney, R.E.; Fink, D.; Kelling, S. eBird: A citizen-based bird observation network in the biological sciences. *Biol. Conserv.* **2009**, *142*, 2282–2292. <https://doi.org/10.1016/j.biocon.2009.05.006>.
66. Siebert, S.; Ewert, F. Spatio-temporal patterns of phenological development in Germany in relation to temperature and day length. *Agric. For. Meteorol.* **2012**, *152*, 44–57. <https://doi.org/10.1016/j.agrformet.2011.08.007>.
67. Möller, M.; Boutarfa, L.; Strassemer, J. PhenoWin—An R Shiny application for visualization and extraction of phenological windows in Germany. *Comput. Electron. Agric.* **2020**, *175*, 105534. <https://doi.org/10.1016/j.compag.2020.105534>.
68. Harfenmeister, K.; Itzerott, S.; Weltzien, C.; Spengler, D. Detecting phenological development of winter wheat and winter barley using time series of sentinel-1 and sentinel-2. *Remote Sens.* **2021**, *13*, 5036. <https://doi.org/10.3390/rs13245036>.
69. Griffiths, P.; Nendel, C.; Hostert, P. Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. *Remote Sens. Environ.* **2019**, *220*, 135–151. <https://doi.org/10.1016/j.rse.2018.10.031>.
70. Htitiou, A.; Boudhar, A.; Benabdelouahab, T. Deep learning-based spatiotemporal fusion approach for producing high-resolution NDVI time-series datasets. *Can. J. Remote Sens.* **2021**, *47*, 182–197. <https://doi.org/10.1080/07038992.2020.1865141>.
71. Filgueiras, R.; Mantovani, E.C.; Althoff, D.; Fernandes Filho, E.I.; Cunha, F.F.d. Crop NDVI monitoring based on sentinel 1. *Remote Sens.* **2019**, *11*, 1441.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.