Student thesis series INES nr 526

Time series analysis for spring barley phenology monitoring using Sentinel 2 - A case study of Southern-central Sweden

Pavlos Aslanis

2020
Department of
Physical Geography and Ecosystem Science
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

Pavlos Aslanis (2020).
Time series analysis for spring barley phenology monitoring using Sentinel 2 - A case study of Southern-central Sweden
Master degree thesis, 30 credits in *Subject of degree*
Department of Physical Geography and Ecosystem Science, Lund University

Level: Master of Science (MSc)

Course duration: *January* 2020 until *June* 2020

Disclaimer

Time series analysis for spring barley phenology monitoring using Sentinel 2 - A case study of Southern-central Sweden

---

Pavlos Aslanis

Master thesis, 30 credits, in *MSc in Geomatics*

Per-Ola Olsson
Dept of Physical Geography and Ecosystem Science

Johannes Albertson
Swedish University of Agricultural Sciences

Exam committee:
Lars Eklundh, Dept of Physical Geography and Ecosystem Science
Vaughan Phillips, Dept of Physical Geography and Ecosystem Science

Abstract

Monitoring crop phenology at parcel scale aligns with the concept of precision agriculture (PA) and can provide invaluable information to agronomic management systems. Satellite time series data are commonly deployed for detecting crop growth stages, while the recent advancements in remote sensing (RS) technologies such as the launch of Sentinel 2 (S2) are providing unprecedented opportunities for crop monitoring. In this thesis, the focus is on spring barley parcels with available in-situ crop growth stage measurements (recorded in Zadoks scale) located in south-central Sweden over the period of 2017 – 2019. More specifically, the aim was to detect three specific crop growth stages of spring barley that are crucial for applying external inputs (e.g. fungicides applications) named according to the Zadoks Scale: (i) first node detectable (31DC) (ii) flag leaf ligule just visible (39DC) and (iii) first spikelet of inflorescence just visible (51DC).

This thesis describes a simple empirical approach based on Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index 2 (EVI2) S2 time series. TIMESAT 4.0 was deployed to reconstruct the S2 NDVI and EVI2 trajectories using the Double Logistic (DL) smoothing function. Moreover, the available in-situ crop growth stage measurements were utilized to optimize the dynamic thresholds (% of the amplitude of the season) for detecting the different crop growth stages of interest (31DC, 39DC, 51DC). Two types of thresholds were conceptualized and optimized: the (i) local threshold and the (ii) global threshold. The optimal local threshold for each crop growth stage of interest refers to the threshold that created an agreement between the vegetation index (NDVI and EVI2) results and the in-situ measurements for each spring barley field individually. The global threshold refers to the optimal threshold that resulted in the smallest Root Mean Square Error (RMSE) between in-situ measurements and S2 derived results when applied on all the studied spring barley parcels.

The optimal local thresholds showed high variability especially for the crop growth stage 31DC, where the standard deviation (SD) of the local threshold values was 16.1% (NDVI) and 22.1% (EVI2). The variability of the optimal local thresholds showed a decreasing trend with latter crop growth stages, where for the stage 51DC the SD was 5.9% (NDVI) and 3.1% (EVI2). According to the results of the global threshold

optimization, the optimal thresholds for the stages 31DC, 39DC and 51DC based on NDVI were 74%, 92%, and 99% respectively, where for EVI2 the optimal global thresholds were 70% (31DC), 91% (39DC) and 99% (51DC). When applying the optimized global thresholds, the RMSE of the retrieved dates for the different crop growth stages against the in-situ measurements was smaller than 7.6 days (for both NDVI and EVI2). EVI2 consistently outperformed NDVI regarding all the crop growth stages of interest (31DC, 39DC, 51DC) where it resulted in lower RMSE (5.1 days, 4.8 days, 4.2 days) and higher coefficient of determination (R-square; 0.43, 0.45, 0.69) compared to the NDVI induced RMSE (6.9 days, 7.6 days, 7.2 days) and R-square (0.33, 0.27, 0.39). The results showed the feasibility of using S2 data in crop phenology studies and demonstrated its potential uses and inaccuracies regarding the detection of three specific crop growth stages of spring barley that are of interest in agronomic decision making.

Keywords: spring barley, crop phenology, crop growth stages, TIMESAT, double logistic, threshold optimization, Sentinel 2.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Per-Ola Olsson for the continuous and restless guidance during this project. Special thanks to Zhanzhang Cai for his assistance regarding the TIMESAT software. Many thanks to Vultus and Qiang Wang for giving me the opportunity to work on this topic. A thank you to Johannes Albertson and Robert Dinwiddie for providing the field data, and my deep appreciation to the field observers. Lastly, a big thanks to the teaching staff at the Department of Physical Geography and Ecosystem Science at Lund University and especially to Karin Larsson.

Table of Contents

1. Introduction

Precision agriculture (PA) is a sophisticated approach for monitoring the spatial and temporal variability of crop related parameters within a field (e.g. soil properties, nutrients, pests, etc.) in order to apply the right crop management practices explicitly where and when needed (Mondal and Tewari 2007). PA has been widely promoted for increasing crop yield and its quality while reducing adverse environmental effects (e.g. nutrient or pesticides leakage) and has been characterized as essential for the future of agriculture since it facilitates sustainable agronomic management (Whelan and McBratney 2000; Mulla 2013). PA enables precise crop monitoring throughout the growing season, and promotes timely and spatially accurate application of proper external input doses (herbicides, fungicides, fertilizers, etc.) which is crucial for maintaining healthy crops and for maximizing their profitability (Mulla 2013; Yang 2020). The importance of detecting the different crop growth stages for agronomic management is highlighted in many crop phenology studies (You et al. 2013; Zheng et al. 2016; Stendardi et al. 2019), where in order to apply external inputs efficiently it is crucial to refer to crop growth stages that exist in numerous crop growth stage scales (e.g. Zadoks scale; Tottman et al. 1979). As noted by Luo et al. (2020), in-situ measurements for monitoring crop growth stages are not time efficient and come with an economic cost that is higher than when utilizing technologies that can establish detailed and frequent methods for monitoring agricultural areas at regional scales.

Satellite remote sensing (RS) has been advertised as an indispensable data acquisition method for agriculture monitoring (Defourny et al. 2019). There have been tremendous advancements in RS technologies since the appearance of PA (Mulla 2013). These developments attracted many researchers to study the use of satellite RS techniques for monitoring crop phenology (Viña et al. 2004; Pan et al. 2015; Gao et al. 2017; Xu et al. 2017; Gomarasca et al. 2019; Huang et al. 2019; Nasrallah et al. 2019; d'Andrimont et al. 2020; Gao et al. 2020; Luo et al. 2020). Viña et al. (2004) and Yousfi et al. (2019) describe that monitoring the phenological cycle of crops can provide invaluable information about the crop growth stages. Furthermore, as noted by Jin and Eklundh (2014) when working with RS data the utilization of spectrally based equations known as Vegetation Indexes (VIs) is an effective way to monitor phenology. In recent decades, the most commonly used optical platforms for crop monitoring were offering either high spatial resolution (e.g. Landsat) or high temporal resolution (e.g. MODIS). However, the use of coarse or medium spatial resolution data (e.g. MODIS) is

insufficient for crop monitoring at parcel level, while satellites with low temporal resolution (e.g. Landsat) cannot provide sufficiently frequent coverage of the rapid crop growth cycle (Solano-Correa et al. 2017). In that scope, data with high spatial and temporal resolution are required to create dense and robust time series and ultimately for monitoring the phenological cycle of crops at field scale (Pan et al. 2015; Solano-Correa et al. 2017; Defourny et al. 2019).

Sentinel-2 (S2) is a mission managed by the European Space Agency (ESA) that provides invaluable data with a highly effective combination of spatial and temporal resolutions (Escolà et al. 2017). As a result, S2 has been labelled as an unprecedented improvement in optical RS sector for conducting near real-time crop monitoring of large agricultural areas at field level (Defourny et al. 2019) and thus, S2 has been the basis for a vast amount of agronomic applications (Escolà et al. 2017; Nasrallah et al. 2019; Stendardi et al. 2019; Belda et al. 2020; d'Andrimont et al. 2020). The Normalized Difference Vegetation Index (NDVI) is the most common VI used in PA and phenology studies (Viña et al. 2004; Gao et al. 2017; Yousfi et al. 2019; Zeng et al. 2020). However, the non-linear relationship between NDVI and Leaf Area Index (LAI) under dense vegetation has raised criticism by Huete et al. (2002). Out of the vast amount of VIs that can be used in phenology studies, the Enhanced Vegetation Index 2 (EVI2; Jiang et al. 2008) has been commonly utilized (Cai 2019; Zhang et al. 2020). EVI2 is a two-band version of the Enhanced Vegetation Index (EVI; Huete et al. 2002), with both of them exhibiting a higher sensitivity in dense canopies than NDVI (Huete et al. 2002; Jiang et al. 2008). While different VIs have been reported to have varying performance and results in phenology monitoring (Jiang et al. 2008; Gao et al. 2017; Huang et al. 2019; Zeng et al. 2020), many researchers have described the presence of noise caused by atmospheric contaminations as a common drawback of VIs (Hird and McDermid 2009; Wei et al. 2016; Zheng et al. 2016; Cai et al. 2017; Luo et al. 2020).

Hence, in phenology studies, the application of gap filling and noise reduction methods (i.e. smoothing functions) is required in order to reconstruct a continuous and smooth VI curve that can be used for extracting dates of phenological events (Wei et al. 2016; Cai et al. 2017; Xu et al. 2017; Belda et al. 2020; Zeng et al. 2020). Many researchers report that the use of satellite RS enables efficient crop growth monitoring, but they also underline the fact that the extracted phenological metrics, such as the start and end of growing season, do not refer precisely to crop growth stages (Delécolle et al. 1992; Gao et al. 2017; Liu et al. 2018; Huang et al. 2019). Zeng et al. (2020) describe the

efforts that aim to detect different crop growth stages as species-specific phenology studies, with the use of dynamic thresholds (percentage of the amplitude of the season; Gao et al. 2017; Gao et al. 2020) being among the most common methods for detecting phenological events (Pan et al. 2015; Huang et al. 2019). Zeng et al. (2020) explain that the optimization of dynamic thresholds based on in-situ measurements as performed in the study of Xu et al. (2017) corresponds to an empirical approach for detecting crop growth stages. Nonetheless, the optimization and comparison of crop phenological metrics with in-situ crop growth stage measurements is rare and challenging, where the availability of in-situ measurements is of high importance for studies that aim to detect specific crop growth stages (Xu et al. 2017; Luo et al. 2020; Zeng et al. 2020). Nevertheless, great efforts have been directed towards assessing the potential of S2 in crop growth stage detection. More specifically, Stendardi et al. (2019) and d'Andrimont et al. (2020) demonstrated the use of S2 time series for detecting the Start of the Season (SOS) and End of the Season (EOS) along with other broad phenological phases (e.g. flowering, heading) of different crops.

In this study, in-situ measurements related to crop growth of spring barley parcels (according to the Zadoks Scale) over the growing seasons of 2017 – 2019 were used. Such an archive of in-situ measurements allowed the focus on specific crop growth stages that are important in agronomic management instead of detecting the general phenological development of the crops (e.g. SOS) and ultimately, enabled this study to investigate the potential of S2 that comes with its high spatial and temporal resolutions along with its drawbacks and inaccuracies regarding the detection of crop stages.

1.1 Objectives and research question

This thesis aims to study the feasibility to monitor crop growth stages of spring barley in south-central Sweden, using S2 data. The objective is to assess the ability and potential use of S2 for the detection of three specific crop growth stages of spring barley that are of high importance in agronomic management: (i) first node detectable (ii) flag leaf ligule just visible and (iii) first spikelet of inflorescence just visible. The research questions this thesis will try to answer are:

- Can the aforementioned crop growth stages be detected in S2 data?
- Is NDVI or EVI2 more accurate in the detection of these crop growth stages?
- Is it possible to use a global threshold for a vegetation index to detect these stages?
- How much does the optimal local thresholds vary between fields?

2. Background

2.1 Crop phenology and crop growth stages in Zadoks scale

Studying the repetitive dynamic cycle of specific vegetation events like the leaf emergence, and flowering is known as phenology (You et al. 2013; Eklundh and Jönsson 2016; Zeng et al. 2020). Monitoring crop phenology has been characterized as vital for establishing an efficient agronomic management (Sakamoto et al. 2005; You et al. 2013; Xin et al. 2002; Liu et al. 2018). In order to optimize any agricultural practices, such as the application of external inputs (e.g. fungicides, nitrogen), it is required to have a standardized description of the different crop growth stages (Poole 2005). The above requirement was described decades ago, where Zadoks et al. (1974) underlined the need of creating a universal scale for describing the growth stages of cereals (e.g. spring barley). The scale developed by Zadoks et al. (1974) named Zadoks scale has been widely utilized in agriculture applications (Anderson et al. 1995). Tottman et al. (1979) provide a detailed explanation of Zadoks scale (Appendix Table S1) regarding the crop development stages of cereals, where, the crop phenological cycle is divided into 10 long lasting development phases called principal growth stages numbered with one-digit Decimal Code (DC) from 0 (germination) to 9 (ripening). Furthermore, in order to enable studies that require a higher level of detail, these stages are divided into shorter stages called secondary stages numbered with a two-digit DC that ranges from 00 to 99 with some being merged to avoid surplus classification of growth stages (Zadoks et al. 1974; Tottman et al. 1979). Figure 2.1 illustrates some characteristic crop growth stages of the spring barley phenological development along with their two-digit DC.

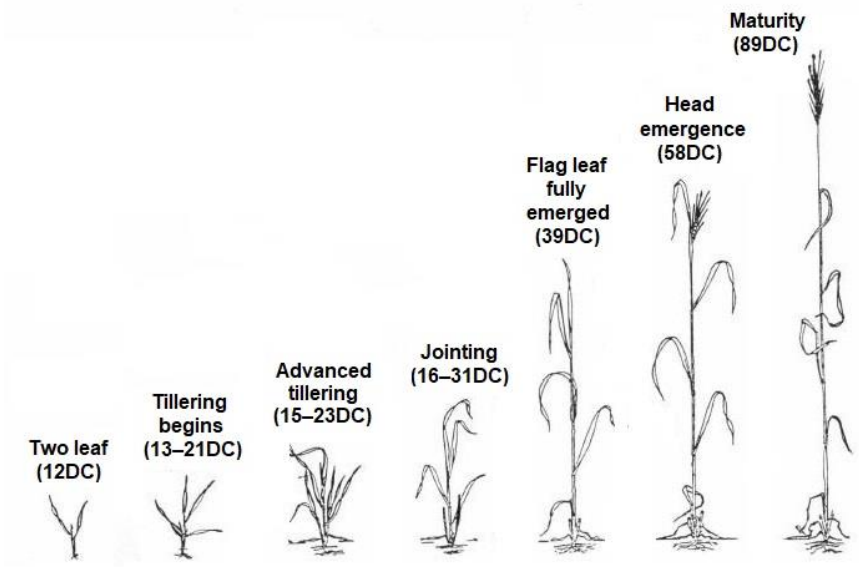

Figure 2.1 Cereal growth stages according to Zadok DC scale (modified from Anderson et al. 1995).

2.2 Fungicides application and crop growth stages

Fungal diseases are known for diminishing crop yield of spring barley crops by decreasing the number of ears along with the size and quality of the grain (Jenkyn 1974). Therefore, fungicides are widely used to counter fungal crop diseases by protecting crop plant matter, where the health state of specific parts of the plants like the three final leaves, the ear and the stems are having a great influence on crop yield of cereals (GRDC 2013). To maximize the protection of a leaf, fungicides should be applied after its full emergence and before the presence of disease signs, hence, it is crucial to monitor the emergence of the parts of the crops that have the biggest influence on crop yield and quality in order to ensure their health status (Poole 2005). Furthermore, a majority of fungicides are useful at certain crop growth stages only, where Nakajima (2010) informs that missing the optimal timing of fungicides application is a common problem in agricultural management which can have the same results as the omission of fungicides. Hence it is of high importance to assess the crop growth stages in order to apply fungicides at the suitable time, optimal amount and location which as an agricultural management procedure will ultimately increase crop profitability by minimizing yield loss and the amount of financial resources spent on fungicides (Newlands 2018). The emergence of the plant parts with the highest influence on crop yield and quality for cereals according to the Zadoks scale are falling between the crop growth stages of early stem elongation and ear emergence (30DC–59DC; Poole 2005). However, a leaf needs to be fully emerged to receive the maximum protection from fungicides application. The stage of the flag leaf being fully emerged (39DC) is common crop growth stage for applying fungicides, since at the stage 39DC the three most important leaves that drive grain production are fully emerged (Poole 2005), where additionally fungicides applications at stage 39DC are promoting spikelet survivability (AHDB 2020).

Fungicide applications scheduled once or twice during the growing season are usually sufficient for maintaining healthy spring barley crops (AHDB 2020). In the case of spring barley the double fungicide application should have the first dose performed within the period of late tillering (25DC–31DC) and the second dose between the stages of the flag leaf being fully emerged and the complete emergence of the ear (39DC–59DC) where for malting barley the latest timing of fungicide spraying should be the initiation of ear emergence (51DC). On the other hand, in the case of applying fungicides only once over the growing cycle of spring barley, the spraying should take

place at the stages around the flag leaf emergence (37DC – 39DC; AHDB 2020). As noted by McLean (2012) the double timing fungicide application schedule (at 31DC and 39DC) is more efficient than a single application, where fungicides applied at the initiation of ear emergence (51DC; first spikelet of inflorescence just visible) can promote a better yield but with questionable effect on the total crop profit.

According to Thorne (1966), the photosynthesis by the flag leaf (the last/top leaf) and ear is the biggest contributor to the crop yield of cereals, therefore, Jenkyn (1974) underline the need of protecting the top leaves (i.e. flag leaf, and the two leaves below the flag leaf; flag -1, flag -2) along with the ear of the spring barley plants. The flag leaf itself in barley is not that important as for other cereal crops (e.g. wheat), where for barley the lower leaves as the flag -1 and flag -2 are having a higher yield contribution, with their emergence taking place between the crop growth stages 30DC – 33DC (ADAMA, n.d.; Figure 2.2).

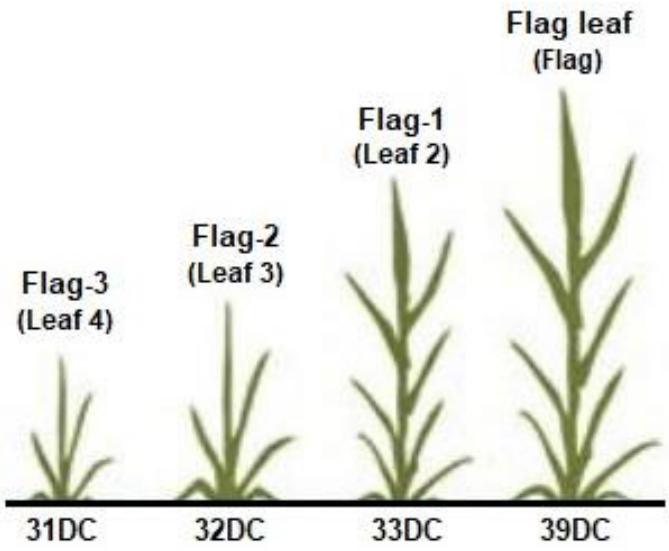

Figure 2.2 Crop growth stages of cereals along with important leaves for crop yield (modified from ADAMA n.d.).

Based on the above, it is apparent that the crop growth stages (Zadoks scale), (i) of first node detectable (31DC), (ii) flag leaf ligule just visible (39DC) along with the (iii) first spikelet of inflorescence just visible (51DC) comprise three crucial stages of the spring barley growth cycle for agronomic management (i.e. fungicides application).

2.3 Satellite based remote sensing and crop phenology

2.3.1 Satellite remote sensing and Vegetation Indices

The base of RS is the interaction of the electromagnetic radiation with the surface of the earth and the objects on it (Jones and Vaughan 2010). Chuvieco (2016) explains that the use of RS enables vegetation monitoring by providing information in different parts of the electromagnetic spectrum that can be used for highlighting specific properties of the vegetation canopy, such as the leaf's water content and structure along with its photosynthetic pigments. The wavelengths of the electromagnetic spectrum that are widely used in vegetation monitoring (Figure 2.3) are the visible light (0.4 – 0.7 μm; VIS), the near-infrared part (0.7 – 1.1 μm; NIR) and the short wave infrared part (1.3 – 2.5 μm; SWIR; Chuvieco 2016). The amount of received and reflected energy by a surface (e.g. vegetation) is expressing the reflectance properties of it, which differ throughout the EM spectrum resulting in a reflectance signature of that surface (Huete 2004; Chuvieco 2016). Figure 2.3 illustrates the reflectance properties (or spectral signature) of a green leaf in the VIS, NIR, and SWIR parts of the electromagnetic spectrum along with the plant properties that drive those changes.

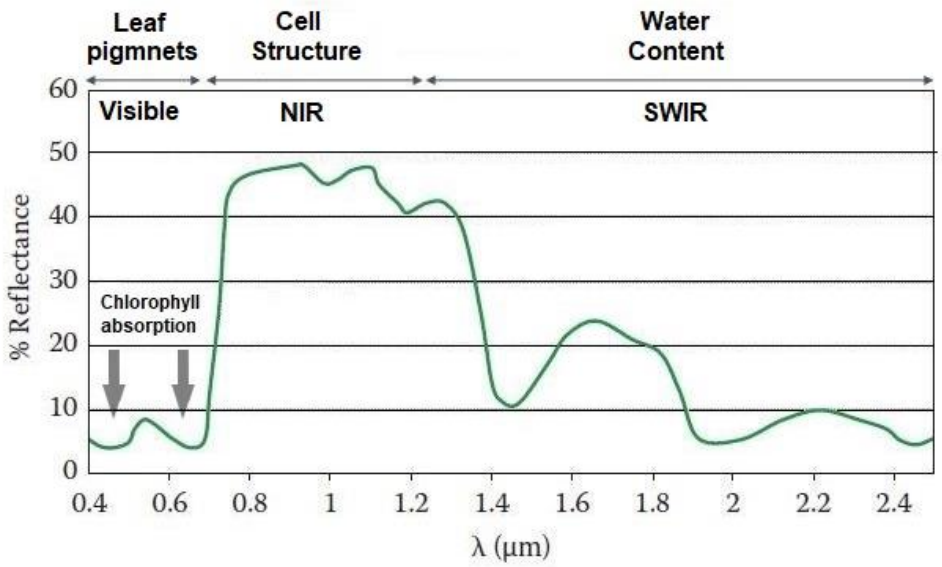

Figure 2.3 Reflectance properties of a green leaf in the visible, near-infrared, and shortwave near infrared (where λ is the wavelength; modified from Chuvieco 2016).

Leaf pigments comprise the main factor for having low reflectance in the VIS spectral region (e.g. chlorophyll in red, carotenoids in blue), in the NIR domain the internal cell structure of healthy leaves results in high reflectance properties, while the leaf water content is having high absorption in the SWIR region with peaks at 1.4 μm, 1,9 μm and 2,5 μm (Jones and Vaughan 2010; Chuvieco 2016). The large difference in reflectance

between the NIR (0.7 – 1.1 μm) and the red (around 0.65 μm) wavelengths has been widely exploited for the creation of spectrally based equations as a robust method for vegetation monitoring (Reed et al. 2009; Chuvieco 2016). These equations of two or more spectral bands also known as VIs are commonly used for assessing vegetation properties with the red and NIR being the most commonly used bands in VIs (Huete et al. 2002; Chuvieco 2016; Palchowdhuri et al. 2018).

A widely used VI in crop phenology studies is the Normalized Difference Vegetation Index (NDVI; Wei et al. 2016; Zheng et al. 2016; Zeng et al. 2020). NDVI is formulated as the ratio of the difference in NIR and red bands divided by their sum as a way to normalize for input noises (i.e. viewing effects). Despite its well described drawbacks (e.g. saturation under high LAI, soil disturbances) NDVI is known for capturing the spatial and temporal aspect of dynamic vegetation activities (Huete et al. 2002; Chuvieco 2016). The Enhanced Vegetation Index 2 (EVI2) developed by Jiang et al. (2008) is also a ratio based VI which uses the NIR and red bands and comprises a two band alternative of the Enhanced Vegetation Index (EVI; Huete et al. 2002). EVI is a three band VI, based on the NIR, red and blue bands and along with its two band alternative (EVI2), they are known for having the same performance especially under good quality observations, where both of them are able to capture variations in dense vegetation in contrast with the NDVI. According to, Jiang et al. (2008) the development of the two-band EVI2 was motivated by the fact that the blue band that was included in EVI mostly for aerosol noise normalization was lacking information about vegetation characteristics, while with the recent automation of atmospheric correction procedures (e.g. S2 LVL2A, etc.) the differences of these two indices that have origins in the exclusion of the blue band are minimized. Another important reason for the creation of the EVI2 was the absence of the blue band from many widely used satellite sensors (e.g. AVHRR; Jiang et al. 2008).

2.3.2 Satellite VI time series analysis for crop phenology

Since VI time series provide invaluable information about the crop growth stages (Pan et al. 2015; Araya 2017), their use has been the main RS approach for deriving crop phenological characteristics (Sakamoto 2018). However, the presence of noise caused by atmospheric effects (e.g. aerosols), cloud contaminations, and sensor and viewing effects (partly normalized with NDVI and EVI2) correspond to the main drawbacks of optical satellite time series products (Eklundh and Jönsson 2016; Sakamoto et al. 2005). The removal of noise is a necessary pre-processing step for acquiring smoothed time

series (You et al. 2013) and ultimately for extracting crop phenological metrics (Araya et al. 2018). In general, detecting crop growth stages from VI time series can be seen as a two-step procedure, where, first step is the reconstruction of a smooth continuous VI curve (smoothing and gap filling; Belda et al. 2020), and second step is the analysis of the smoothed time series for the extraction of phenological parameters that relate to the desired phenological phases (You et al. 2013; Gao et al. 2020).

Noise Removal and Gap Filling

Various approaches can be used for removing noise from satellite time series data with many comparison studies concluding that there is no universally superior method (Beck et al. 2006; Atkinson et al. 2012; Wei et al. 2016; Zheng et al. 2016; Cai et al. 2017; Belda et al. 2020). As noted by Eklundh and Jönsson (2016), the Maximum-Value Composite (MVC) proposed by Holben (1986) is a simplistic smoothing method that was widely adopted in numerous phenology studies based on data with high temporal resolution (Jönsson and Eklundh 2002; Xin et al. 2002; Jönsson and Eklundh 2004; Zheng et al. 2016). MVC refers to the selection of the highest VI value for every pixel of the satellite scene (single location) over a defined period of time (e.g. 8 days) that results in a single MVC image representing that period (e.g. decreased temporal resolution to 8 days; Holben 1986). A substitute of MVC, the Best Index Slope Extraction (BISE) developed by Viovy et al. (1992), has been also utilized in numerous phenology studies (White et al. 1997; Lange et al. 2017; Xu et al. 2017; Stendardi et al. 2019). BISE is based on a window that slides over the raw time series data in order to exclude negatively and positively biased noise according to a user defined range of accepted fluctuations in the VI values (Viovy et al. 1992).

In recent years, noise removal is additionally based on more sophisticated methods such as the fitting mathematical functions to time series of RS data (Eklundh and Jönsson 2016; Cai et al. 2017; Zeng et al. 2020). Zheng et al. (2016) focused on monitoring crop phenology and tested three of the most common fitting functions that are used in phenology studies: (1) the Savitzky–Golay filter (SG), (2) the Double Logistic function (DL), and (3) the asymmetric Gaussian function (AG). Cai et al. (2017), performed a comparison of these three smoothing methods and classified them in two broad categories with respect to their ability to maintain long or short term variations of the time series, named as global fitting (DL, AG) and local fitting methods (SG) respectively. DL as a global smoothing method is better at capturing the general trend

of crop properties, while SG as a local method maintains more of the original shape of the time series (Liu and Zhan 2016). Furthermore, Zeng et al. (2020) note that for selecting the most suitable smoothing method, it is needed to understand the input data along with the smoothing quality that was achieved prior to the extraction of the desired phenological metrics, since these factors have a great effect on the results. Figure 2.4 depicts an example of single year raw NDVI time series data (blue points) of spring barley, along with the quality marked acquisitions (marked with red circles) that were used for fitting the DL function (red line).

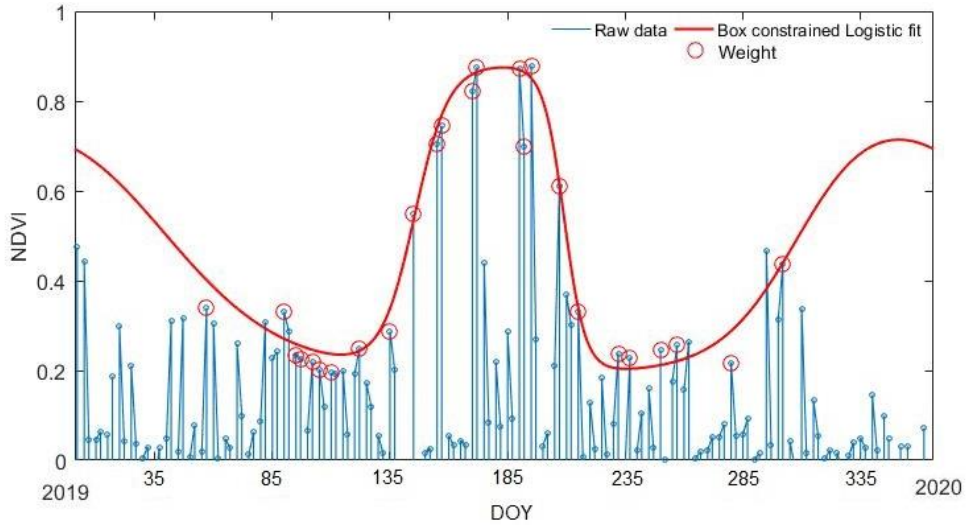

Figure 2.4 Example of single year raw NDVI time series data (blue points) of a single pixel of spring barley, showing the quality marked acquisitions (marked with red circles) used for the time series reconstruction with the use of the DL function (red line).

Extraction of phenological parameters

After reconstructing the smoothed VI time series (Figure 2.4), various methods can be used for detecting specific phenological events (e.g. fixed and dynamic threshold, moving average, derivative and inflection point based methods), where dynamic thresholds are among the most common approaches (Pan et al. 2015; Huang et al. 2019). The fixed and the dynamic threshold refer to the selection of a VI value for representing the desired phenological phase (Cai et al. 2017). A fixed threshold simply defines an absolute VI value for representing the desired phenological event. The dynamic thresholds are based on the characteristics of the VI time series (Zeng et al. 2020), such as the base level (the average of the minimum values on left and right side of the season peak) and the maximum seasonal amplitude (difference of maximum VI value and base level; Eklundh and Jönsson 2017). Figure 2.5 shows an example of single pixel NDVI time series of spring barley. The green dot along the DL fitted curve represents the use of fixed threshold (where NDVI equals to 0.4), and the red dot represents the use of dynamic threshold (40% of the amplitude of the season).

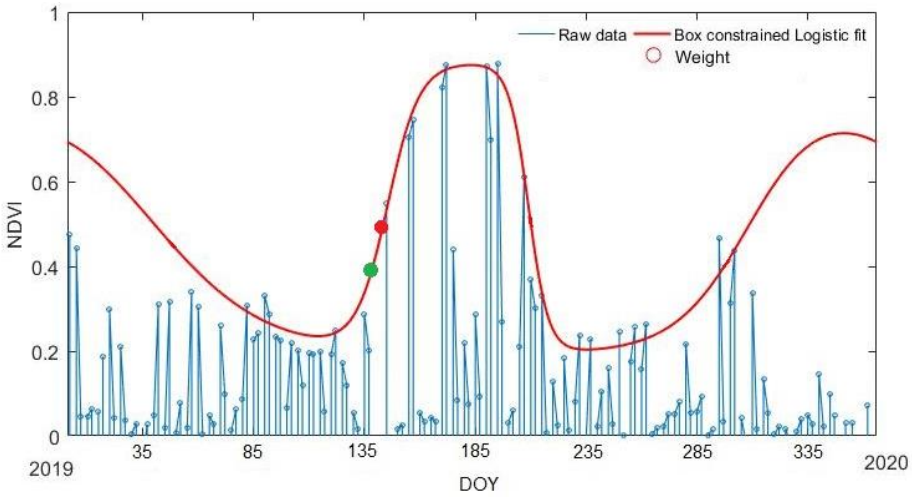

Figure 2.5 Example single year NDVI raw time series data (blue points) of spring barley (single pixel) along with the smooth NDVI curve (DL fit; red line). The green dot represents the use of fixed threshold (at 0.4 NDVI) and the red dot represents the use of dynamic threshold (at 40% of the seasonal amplitude).

Moreover, the procedures of smoothing noisy time series data and extracting phenological parameters have drawn the attention of many researchers, with several of time series processing software available (TIMESAT[1], Jönsson and Eklundh (2004); PhenoSat[2], Rodrigues et al. (2013); CropPhenology[3], Araya et al. (2018)). Liu and Zhan (2016) and Wei et al. (2016) used a 20% threshold (20% of the amplitude of the season) for denoting the Start of the Season (SOS) and the End of the Season (EOS) of crops in China. Xu et al. (2017), defined different threshold values for detecting different crop growth stages (e.g. heading, senescence) of various crop types (e.g. winter barley, winter wheat).

In this study, TIMESAT 4.0 was used, with the main motivation that new version of TIMESAT can handle the irregular time step of S2 time series. TIMESAT has been widely used in phenology studies (Xin et al. 2002; Jönsson and Eklundh 2004; Atkinson et al. 2012; Liu and Zhan 2016; Wei et al. 2016; Zheng et al. 2016; Cai et al. 2017; Huang et al. 2019; Zeng et al. 2020) since it provides a way of applying the most common methods for the reconstruction of RS time series (e.g. SG, DL, AG) while it also allows the use of user-defined dynamic thresholds for the extraction the desired phenological metrics (Zeng et al. 2020).

[1] http://web.nateko.lu.se/TIMESAT/timesat.asp
[2] https://www.fc.up.pt/PhenoSat/software.html
[3] https://cropphenology.wixsite.com/package

3. Data and methods

3.1 Study area

The study area consists of 21 spring barley fields that fall within three S2 tiles (33VUE, 33VVE, 33VXG) located in south-central Sweden (Figure 3.1). The selection of spring barley as the studied cereal crop was based on the availability of crop growth in-situ observations. Barley is known for its resilience to extreme climate conditions and its ability to thrive in high latitudes (Baik and Ullrich 2008). In Sweden barley is mainly used as a spring crop and can be found around the whole country but mostly at middle and lower latitudes (Swedish University of Agricultural Sciences 1996). According to the Swedish Board of Agriculture (Jordbruksverket) the total arable land in Sweden was approximately 2 500 000 hectares in 2019, where around 990 000 hectares (40%) were cultivated with cereals and more specifically 300 000 hectares (12%) were used for barley. Based on that, barley was the second most grown cereal crop in Sweden in 2019 (SCB 2019) where the majority of Sweden's barley production is used for livestock feed needs in the inland and part of it is used for malt production and export (Swedish University of Agricultural Sciences 1996).

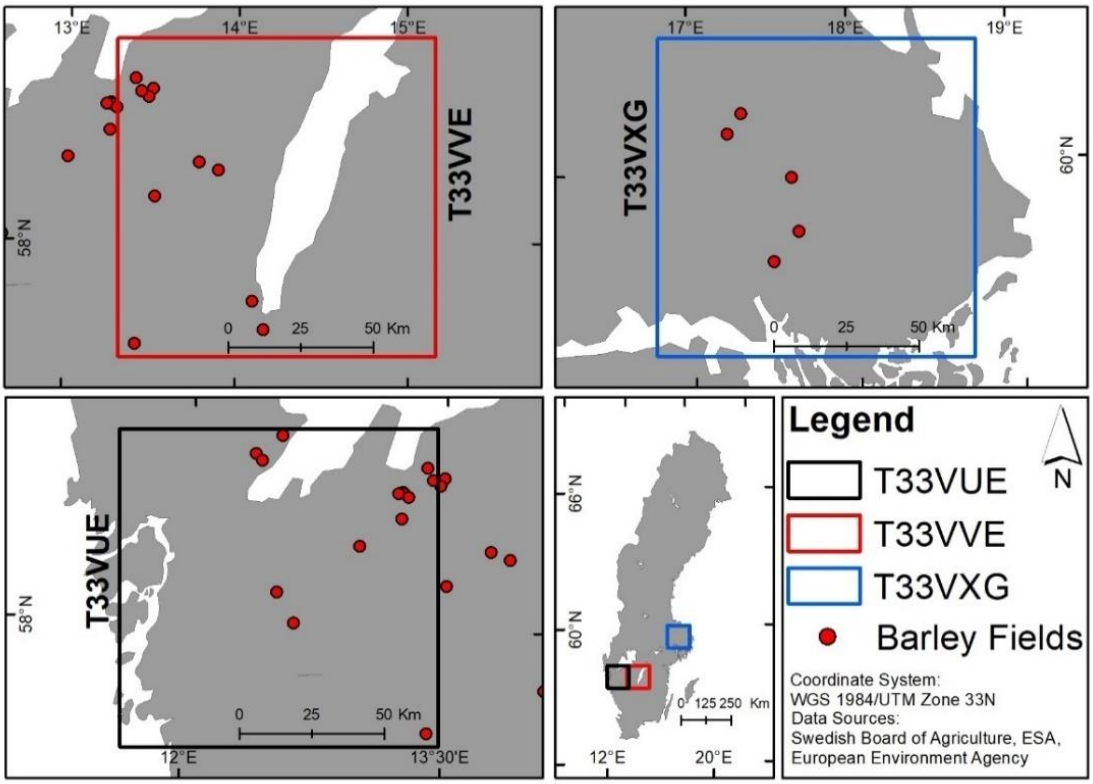

Figure 3.1 Study area map, showing the three S2 tiles (33VUE, 33VVE, 33VXG) that include the 21 barley fields that were sampled in 2017,2018, 2019.

The S2 tiles 33VUE, 33VVE (around 58°37´N, 11º33´E to 57°38´N, 15°10´E) and 33VXG (around 60°25´N, 16º45´E to 59°56´N, 18°48´E) include a considerable

percentage of the agricultural areas of central Sweden where each of the tiles covers an area of 10 000 km^2 with adjacent tiles (33VUE, 33VVE) having a 10 km overlap. For the years 2017, 2018, and 2019 the total number of spring barley fields with available in-situ measurements in the S2 tiles of interest was 5, 6 and 10 respectively.

3.2 Data

3.2.1 Sentinel-2 data

S2 is a multi-spectral satellite constellation that consists of S2A and S2B launched in 2015 and 2017 respectively and offers free products of high spatial resolution (60 to 10 meter). S2 has a 290 km swath width and provides data in 13 different spectral parts of the electromagnetic spectrum, ranging from the visible to the short wave infrared (SWIR) wavelength (a detailed description of the different bands can be found in the S2 user handbook published by ESA in 2015). The theoretical temporal resolution of the constellation is equal to 5 days at the Equator, with the middle latitudes having an overlap between adjacent acquisitions that results in a shorter revisit cycle (2–3 days; Escolà et al. 2017) but with inconsistent viewing conditions (European Space Agency, 2015).

All available S2 Level 2A (LVL2A) and Level 2Ap (LVL2Ap) scenes from the tiles 33VUE, 33VVE, 33VXG for the years 2017, 2018 and 2019 were downloaded on 26th February 2020 (through the Copernicus Open Access Hub[1] using Python in conjunction with the library Sentinelsat[2]) without applying a cloud coverage restriction. Both LVL2A and LVL2Ap refer to Bottom of Atmosphere (BOA) reflectance products that are georegistered in WGS84/UTM (European Space Agency, 2015). Therefore, it was not needed to perform any of the common image analysis preprocessing steps (e.g. atmospheric, geometric correction). Another feature of LVL2A/p scenes is that they include a scene classification product (SCL; generated based on atmospheric corrections procedures that ESA applies on these level of products) with 60 and 20 meter spatial resolution (European Space Agency, 2015). The LVL2Ap (pilot version of LVL2A) started being distributed in April 2017 and was replaced by its equivalent LVL2A in March 2018. Hence, scenes with an acquisition date before the 3rd of April 2017 were not included in the analysis, and additionally, the scenes before and after the

[1] https://scihub.copernicus.eu/
[2] https://pypi.org/project/sentinelsat/

23rd of March 2018 (LVL2Ap and LVL2A respectively) were treated as the same level of products.

In total, 1832 S2 scenes were downloaded with a total size of 0.9 TB. Table 3.1 shows the number of scenes that were used for each of the tiles and years. The year 2017 has a significantly lower number of scenes, which stems from the availability of LVL2A and LVL2Ap products since not all of the existing S2 scenes (e.g. available as LVL1C) have been converted by ESA to LVL2A/p (a complete description of the availability of LVL2A/p scenes for the year 2017 can be found in the work of Sudmanns et al. (2019)).

Table 3.1 Number of Sentinel 2 acquisitions that were used in the analysis per tile (33VUE, 33VVE, 33VXG) per year.

Year	33VUE	33VVE	33VXG	Total
2017	74	95	92	261
2018	222	282	262	766
2019	231	297	277	805

Table 3.2 The two bands of Sentinel 2 that were used in this study (source: ESA, 2015).

Band Number	Band Name	Central Wavelength (μm)	Spatial Resolution (m)
4	Red	0.665	10
8	Near infrared	0.842	10

Out of the 13 spectral bands that S2 provides, Band 4 (Red) and Band 8 (NIR) were utilized in this study for the calculation of the selected VIs (NDVI, EVI2), where both of them have 10 meter spatial resolution, Table 3.2 shows the basic characteristics of these bands. Moreover, the S2 scene classification product (SCL) with original spatial resolution of 20 meter was resampled to 10 meter (using the nearest neighbor method) and was then used as a quality indicator for the time series analysis.

3.2.2 In-situ measurements of crop growth stages

The in-situ crop growth stage measurements of the 21 spring barley fields for the period 2017-2019 were collected and provided by the Swedish Board of Agriculture (Jordbruksverket) in collaboration with the Swedish University of Agricultural Sciences (SLU). Xu et al. (2017) described the importance of recording the exact location when conducting crop in-situ measurements for reducing any uncertainties that can be caused by ground measurements of low spatial quality. That was apparent also in this study, when dealing with the vast amount of in-situ measurements that were

provided by Jordbruksverket and SLU. More specifically, a considerable part of the in-situ measurements did not allow the identification of the sampled spring barley parcel (e.g. one pair of coordinates was connected to more than one field, or the in-situ measurements were falling on buildings). Thus, it was not possible to include all the available ground measurements in the analysis, which in conjunction with the distribution of the S2 tiles, affected highly the number of the studied spring barley parcels (i.e. 21 parcels).

The spatial distribution of the sampled spring barley fields is illustrated in Figure 3.1 The in-situ measurements refer to 13 different varieties of spring barley (Tamtam, SW Makof, Severi, Salome, RGT Planet, Propino, Laureate, KWS Irina, Filippa, Ellinor, Dragoon, Anneli, and Anakin). The crop growth stage observations are based on the Zadok's scale (Appendix Table S1), where additionally, for 13 fields (out of the 21) the sowing date was also recorded.

Table 3.3 Example of the in-situ measurements protocol showing the field (field number) that was sampled, the year, the sampling dates and the corresponding observed crop growth stages (Zadoks scale).

Field Number & Year	Sample date	Crop Growth Stage (DC)	Field Number & Year	Sample date	Crop Growth Stage (DC)
#90, 2017	29/05/2017	12	#91, 2019	06/05/2019	12
	05/06/2017	21		13/05/2019	21
	12/06/2017	23		20/05/2019	30
	19/06/2017	31		27/05/2019	31
	26/06/2017	37		03/06/2019	37
	03/07/2017	49		10/06/2019	49
	10/07/2017	65		17/06/2019	59
	17/07/2017	73		24/06/2019	73
	24/07/2017	75		01/07/2019	83
	-	-		08/07/2019	83

Two examples of the utilized field measurements that correspond to 2 different spring barley parcels are shown in Table 3.3, where the number of the field and the year are shown along with the dates of measurements, and the observed crop growth stages according to the Zadoks scale. The time step of the in-situ observations is equal to one week for most of the observations. The first crop growth stage measurements was in most of the cases taken around May (crop growth stages ranging from 10DC to 25DC were recorded) which falls around a month after the sowing period (based on the sowing date of the observations) while the majority of the final measurements was performed around July (crop stages ranging from 70DC to 87DC were recorded). A limitation of the in-situ observations, is the fact that the crop growth stages that are of interest in this

study (i.e. first node detectable; 31DC, flag leaf ligule just visible; 39DC, and first spikelet of inflorescence just visible; 51DC) were not always recorded in the in-situ observations (e.g. Table 3.3; for the field number 90 in 2017 and 91 in 2019 the crop growth stages 39DC and 51DC are not present as measurements). In order to acquire the dates of the desired crop growth stages (31DC, 39DC, 51DC) linear interpolation was performed according to the Equation 1.

$$y = y_1 + (x - x_1) * \frac{(y_2 - y_1)}{(x_2 - x_1)}, \quad (1)$$

where y is the interpolated Day of Year (DOY) of the desired crop growth stage which is denoted as x (Zadoks scale), y_1 and y_2 are the DOY of the closest crop growth stages recorded respectively before and after the crop growth stage of interest, and x_1, x_2 are the crop growth stage recorded respectively before and after the desired crop growth stage.

In that way, for the field number 90 in 2017 in Table 3.3, the date of the 39DC was interpolated to be the 178th DOY (27th of June, 2017). As seen in Equation 1 and Table 3.3, for this interpolation the observations on the 26th of June, 2017 (DOY 177; 37DC) and on the 3rd of July, 2017 (DOY 184; 49DC) were used. The above method was performed for the majority of the fields, where the crop growth stage 31DC (first node detectable) was present in the measurements of 18 individual fields out of the 21 fields, the stage 39DC (flag leaf ligule just visible) was recorded for 2 of the fields, while the crop growth stage 51DC (first spikelet of inflorescence just visible) was recorded for 3 fields. Therefore, the interpolation for the crop growth stages 31DC, 39DC and 51DC was performed 3, 19 and 18 times respectively. Furthermore, for three of the fields it was not possible to perform linear interpolation in order to acquire the desired stages, in that case these fields were excluded from the analysis as they could be considered as non-informative for the purposes of this research. Table 3.4 shows the dates of the crop growth stages 31DC, 39DC and 51DC as obtained after the interpolation.

Table 3.4 The 3 crop growth stages of interest (31DC, 39DC, 51DC) for the spring barley fields sample (N=10), after the linear interpolation.

Field	Year	Tile	31DC (Date)	39DC (Date)	51DC (Date)
90	**2019**	33VUE	27/05/2019	06/06/2019	-
91	**2019**	33VUE	27/05/2019	04/06/2019	11/06/2019
60	**2019**	33VUE	27/05/2019	10/06/2019	15/06/2019
62	**2019**	33VUE	03/06/2019	13/06/2019	19/06/2019
92	**2019**	33VVE	03/06/2019	09/06/2019	19/06/2019
152	**2019**	33VVE	03/06/2019	12/06/2019	19/06/2019
62	**2018**	33VUE	03/06/2018	08/06/2018	13/06/2018
92	**2018**	33VUE	28/05/2018	02/06/2018	08/06/2018
61	**2017**	33VUE	05/06/2017	11/06/2017	17/06/2017
90	**2017**	33VXG	19/06/2017	27/06/2017	04/07/2017

The priority was to maintain the original location of the in-situ measurements, but some of the in-situ observations were falling on the very edge of the spring barley fields which can lead to an inhomogeneous pixel (mixel effect). However, the small sample size of spring barley fields did not allow the exclusion of every parcel with low quality in-situ measurements. Consequently, all in-situ observations that were located at field edges were identified in order to assess any major differences between the original location and close by pixels. In the case that the original pixel of the in-situ measurements was assessed to be inhomogeneous (mixed pixel), the location of the in-situ observations was moved 20-30 meters further into the spring barley field (that was performed only for the spring barley field number 61 in 2017; the NDVI time series of the new location and the original location are shown Figure 3.3 and Appendix Figure S3 respectively). Another case where the location of the in-situ observations was adjusted is the field 90 sampled in 2017, a detailed explanation for adjusting the original observed location of that spring barley parcel is given in Section 5.1 (Figure 5.1, Figure 5.2).

3.3 Method

The methodology in this study can be divided in four major stages (Figure 3.2): (1) the calculation and construction of single year VIs time series for 2017, 2018 and 2019, (2) reconstruction of the time series with the use of the Double Logistic function, (3) identifying optimal thresholds for detecting the crop growth stages of interest (31DC, 39DC, 51DC), and finally (4) the comparison of the extracted S2 (NDVI, EVI2) results against the in-situ measurements in order to assess and compare their accuracy.

Moreover, in this study, two types of dynamic thresholds were conceptualized, the optimal (i) local and (ii) global threshold. The (i) optimal local threshold for each spring barley parcel is the threshold that creates an agreement between the in-situ DOY and S2 (NDVI, EVI2) derived DOY for the different crop growth stages of interest (31DC, 39DC, 51DC). The (ii) optimal global threshold is the threshold that induces the highest agreement for the different crop growth stages of interest for all the studied spring barley parcels (when compared against the in-situ measurements).

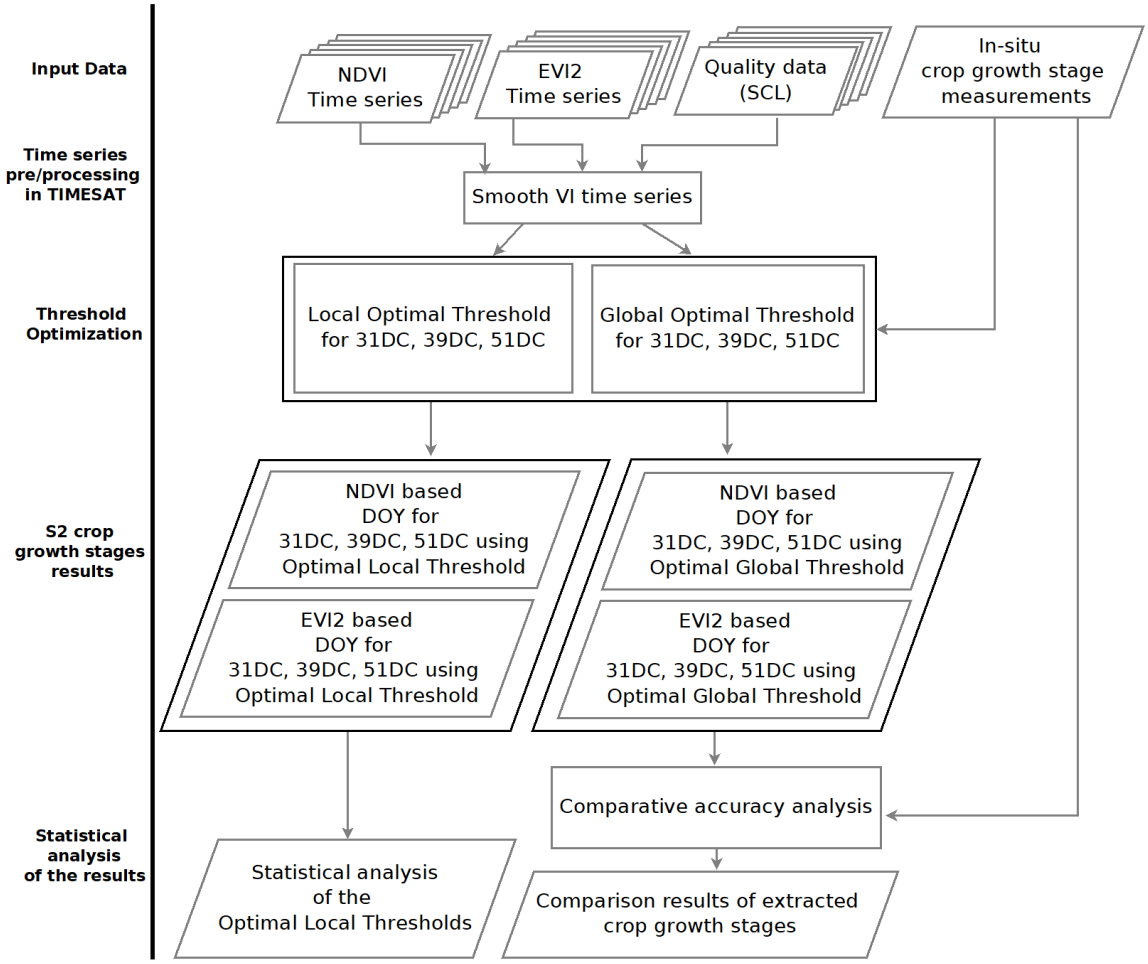

Figure 3.2 Basic steps of methodology workflow.

3.3.1 Preparation of VIs and quality data

Based on the time and storage limitations of this study, the analysis was limited to two of the widely used VIs in phenology studies, the NDVI and EVI2. First step was the calculation of NDVI and EVI2 for all scenes. NDVI was computed according to the Equation 2:

$$NDVI = \frac{(\rho_{NIR} - \rho_{RED})}{(\rho_{NIR} + \rho_{RED})}, \quad (2)$$

where ρ is the reflectance values in the respective band. NDVI as a VI is dimensionless and has a value range from -1 to 1, where bare soil has values around 0 and values of 1 expresses vigorous vegetation.

EVI2 is also dimensionless and its values increase as the amount of vegetation increases with a range from -1 to 1. EVI2 was computed according to the Equation 3:

$$EVI2 = \frac{G * (\rho_{NIR} - \rho_{RED})}{(\rho_{NIR} + C * \rho_{RED} + L)}, \quad (3)$$

where ρ is the reflectance values in the respective band, G equals to 2.5 and corresponds to a gain factor, C is equal to 2.4 representing aerosol resistant coefficient and L is equal to 1 and corresponds to the canopy background adjustment (Jiang et al. 2008).

The VIs calculations were performed in Python using the rasterio[1] library. In total 3664 VIs images were computed (a detailed description of the number of VIs per tile per year is shown in Table 3.1). Furthermore, the scene classification layer (SCL) that is available for every S2 scene (LVL1C and LVL2A/p) was used as a quality flag for the input data. It was necessary to first downscale the SCL data from the original spatial resolution (20 meter) down to 10 meter using the Nearest Neighbor resampling method in order to achieve pixel size consistency of the input data.

3.3.3 Reconstruction of VIs time series

TIMESAT v4.0 is capable of handling the irregular time step of S2 acquisitions, and thus, was used in this study in order to reconstruct the raw S2 time series. The double logistic function was used for the reconstruction of the VIs time series since it has been reported for having the ability to capture accurately the temporal trends of the rapidly

[1] https://pypi.org/project/rasterio/

changing crop dynamics without showing sensitivity to high levels of local variations (Beck et al. 2006; Liu and Zhan 2016). The Double Logistic function has the basis function in Equation 4.

$$g(t; x_1, \dots, x_4) = \frac{1}{1 + exp\left(\frac{x_1 - t}{x_2}\right)} - \frac{1}{1 + exp\left(\frac{x_3 - t}{x_4}\right)}, \qquad (4)$$

where, the parameter x_1 refers to the time of the left inflection point, x_3 refers to the time of the right inflection point, and x_2 , x_4 determine the time periods of increase and decrease at x_1, x_3 respectively. As explained by Jönsson et al. (2018) reconstructing time series data with unfiltered noise can lead to an inaccurate VI curve. To account for that, the aforementioned parameters of the DL function are constrained in range (box constrained DL) in order to reduce the effect of noisy signals. A further explanation of the box constrained DL can be found in the work of Eklundh and Jönsson (2017) where a detailed description of TIMESAT 3.3 functionalities is given.

TIMESAT allows the use of ancillary data/quality data (e.g. SCL) in order to exclude outliers by assigning weights based on noise levels (Eklundh and Jönsson 2017). The SCL (Scene classification Layer) of the S2 acquisitions was used in order to assign quality weights as a preprocessing filtering method before the reconstruction of the smooth VI curve. With the use of SCL data, only the pixels classified as vegetation (class 4 in SCL) and non-vegetated areas (e.g. bare soil; class 5 in SCL) were included in the analysis with a weight of 1. Pixels that were classified as snow, cloud, or shadows were given a weight of 0 and were thus excluded from the smoothing procedure. Moreover, TIMESAT compensates for the negatively biased noise that is present in various VIs by allowing the adaptation of the smoothed VI curve to the upper envelope using a multi-step iterative procedure (Eklundh and Jönsson 2017). TIMESAT is using two parameters for the adaption to the upper envelope, the number of iterations and the strength of the adaption. The definition of the parameters of the smoothing method can be seen as an empirical method which is bounded to the properties of the raw VI time series (Cai et al. 2017). For the reconstruction of the VIs time series the number of envelope iterations was set to three and the adaption strength was set as two (Table 3.5).

Additionally, with the use of the seasonal parameters TIMESAT enables the definition of the number of seasons to be detected within the studied time period. Because 10 (out of the 21) spring barley fields were showing a second season (double crop, or cover crop; Figure 3.3) with varying amplitude within the studied period, the seasonal

parameter for the reconstruction of the time series was set as 0 to force the generation of two seasons. Forcing a double season reconstruction resulted in a satisfying fit for the first season which corresponds to the phenological cycle of spring barley (Figure 3.3; year 2017).

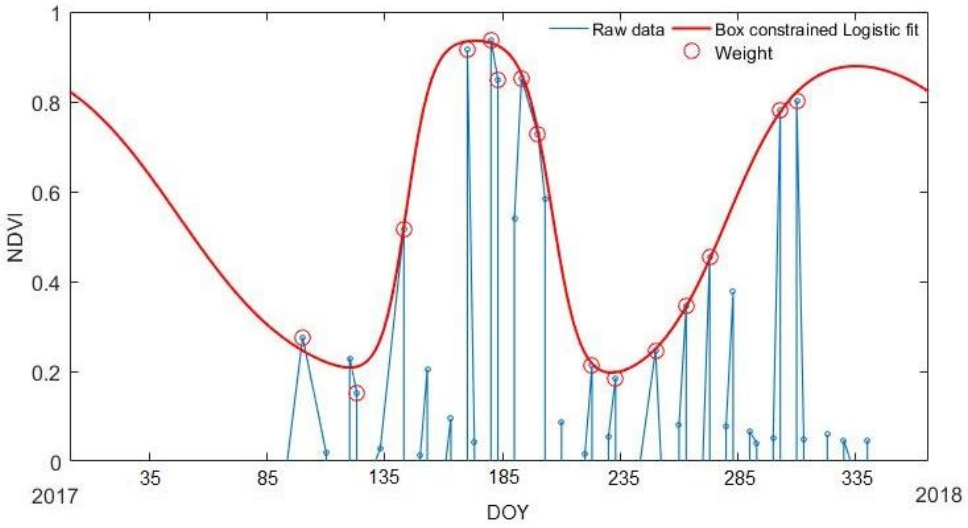

Figure 3.3 Example of NDVI time series of a pixel within a spring barley field (field number 61 in 2017) that was showing a double season. This pixel corresponds to the new location of the in-situ observations as moved in order to acquire a homogenous spring barley pixel (the NDVI time series of original location of the in-situ observations for that spring barley parcel in Appendix Figure S3).

Likewise the study of Gao et al. (2017), in order to deal with the common agricultural practice of crop rotation (i.e. different crop types cultivated at the same field in consecutive years), the analysis in this study was based on single year time series. The data preparation procedures for analyzing single year time series with TIMESAT have been described in detail by Eklundh and Jönsson (2017). The researchers inform that, when studying a single year (e.g. 2017), it is required to create artificial time series data for the year 2016, and 2018 which are identical to the studied year (i.e. 2017) and are placed in the timeline of 2016-2018 to enable the fit of the smoothing function in the middle year (i.e. 2017; Eklundh and Jönsson 2017). The settings used for the reconstruction of NDVI and EVI2 time series within TIMESAT 4.0 are shown in Table 3.5.

Table 3.5 Time series reconstruction settings as used in TIMESAT for the reconstruction of a smooth VIs curve.

Reconstruction settings as used in TIMESAT	
Smoothing Method	Box constrained Double Logistic function
Seasonal Parameter	0
Number of Iterations	3
Adaption Strength	2
Base Level	0.05

3.3.4 Thresholds for extracting the crop growth stages of interest (31DC, 39DC, 51DC)

As described in Section 3.3.4, two different types of thresholds (optimal local and global threshold) were conceptualized in order to acquire the desired phenometrics (phenological indicators) regarding the crop growth stages of interest named: (i) first node detectable (31DC), (ii) flag leaf ligule just visible (39DC), (iii) first spikelet of inflorescence just visible (51DC). The local threshold refers to the optimal threshold for each individual spring barley field (single pixel), while the global threshold was optimized to give the lowest error (RMSE; in number of days) when applied on all the studied spring barley fields. After the pixel level inspection of the location of the in-situ measurements it was decided that out of the 21 spring barley fields only 10 will be used for the optimal global and local threshold definition, since the single pixels studied for the 11 excluded fields did not allow the reconstruction of quality time series due to the common problems of optical RS time series (e.g. atmospheric contaminations; Figure 3.4, Figure 3.5, Appendix Figure S4, Figure S5).

This exclusion, was based on the fact that the crop growth stage 51DC (first spikelet of inflorescence just visible) as reported by Nasrallah et al. (2018) shall fall on the left side of the peak, with the peak of the season being during the flowering period (60 – 69DC). Thus, if the in-situ DOY for the stage 51DC was falling on the right side of the peak of the smoothed VI curve, then the fit of the DL function was considered to be inaccurate, and therefore, was excluded from the threshold optimization analysis.

Figure 3.4 depicts the NDVI time series of the location of the in-situ measurements (single pixel) for one of the spring barley parcels with available ground measurements (field number 93 in 2019). This field was excluded from the analysis since the achieved DL fit (Figure 3.4) did not allow the definition of reasonable thresholds that would result in a match between the in-situ DOY for the crop growth stages of interest (31DC, 39DC, 51DC) and the VIs extracted metrics.

More specifically, for the stage 51DC the in-situ DOY was falling on the right side of the smoothed VI peak, thus the achieved fit was considered as an outlier and was excluded from the threshold optimization analysis. Additionally, for that single pixel (Figure 3.4) the onset of the season (left side of the peak) does not have many quality observations (marked as red circles) to support an accurate fit, and thus, the fitted curve

and more specifically the green-up phase of the smoothed VI curve can be considered to be inaccurate as it is slightly shifted to the left.

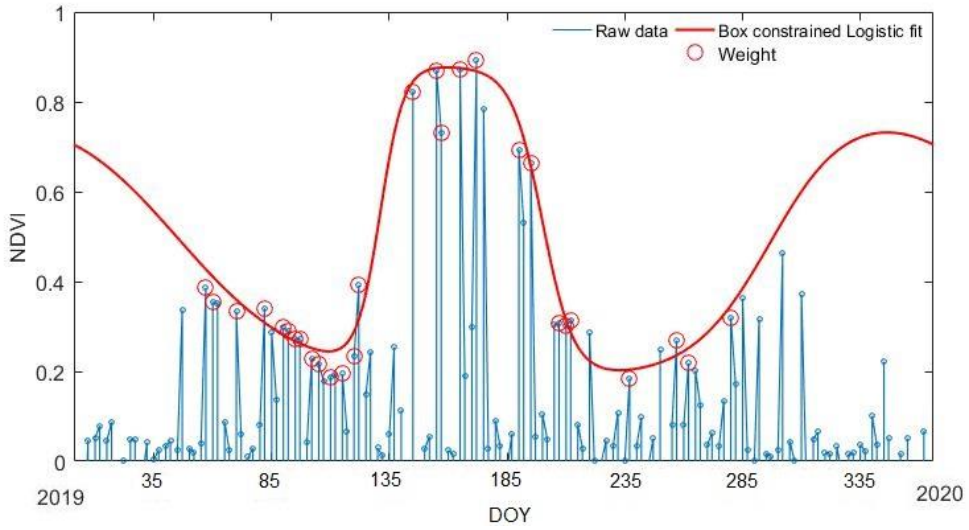

Figure 3.4 Example of NDVI time series of a pixel that represents the location of the in-situ measurements for one spring barley parcel (field number 93 in 2019). This pixel did not allow the reconstruction of a smooth VI curve. Quality observations that were used in the DL fitting procedure are marked in red circles.

Figure 3.5 illustrates the EVI2 time series for the same field (same pixel), where similar noise is present which results in a DL fit that is not able to depict accurately the phenological progress of that spring barley pixel.

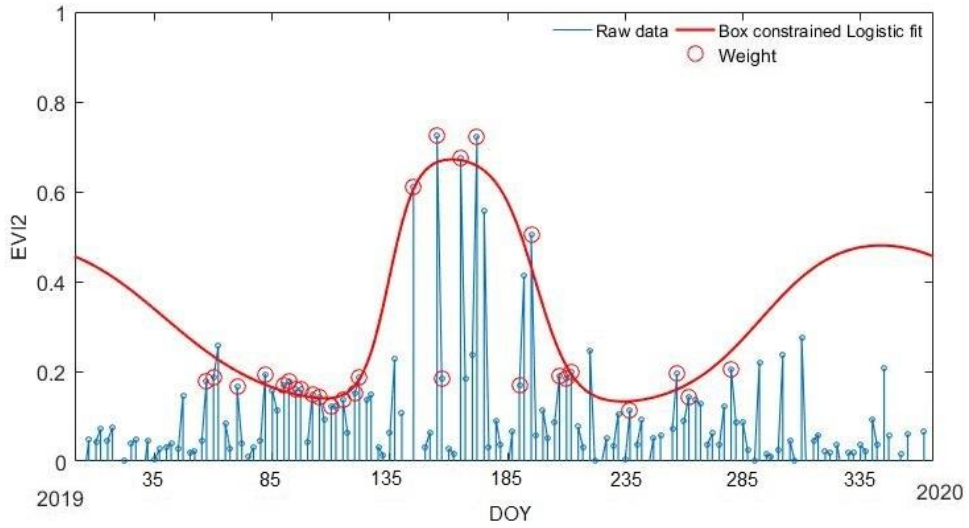

Figure 3.5 Example of EVI2 time series of a pixel that represents the location of the in-situ measurements for one spring barley parcel (field number 93 in 2019). This pixel did not allow for the reconstruction of a smooth VI curve. Quality observations that were used in the DL fitting procedure are marked in red circles.

Optimal local threshold definition

For the 10 studied spring barley fields a manual approach for the identification of the optimal local threshold (for each individual field) was performed. Each field was investigated at pixel level, where the single pixel representing the location of the in-situ measurements was examined to define the optimal local threshold for representing the desired crop growth stages (31DC, 39DC, 51DC). As described by Zeng et al. (2020), the optimal local threshold was defined as the percentage of the VI amplitude that creates an agreement between the in-situ DOY of the different crop growth stages (31DC, 39DC, 51DC) and the DOY extracted from the S2 data (NDVI, EVI2). The 10 spring barley fields were studied at pixel level (at the location of in-situ measurements; single pixel) and the optimal local threshold for each field was defined. For example, Figure 3.6 shows the NDVI time series of the location (single pixel) of the in-situ measurements of a spring barley parcel (field number 91 in 2019). According to the in-situ measurements the crop growth stage 39DC was recorded on the DOY 155, therefore, the optimal local threshold for that pixel was defined as 92% since it created an agreement between NDVI DOY and in-situ DOY.

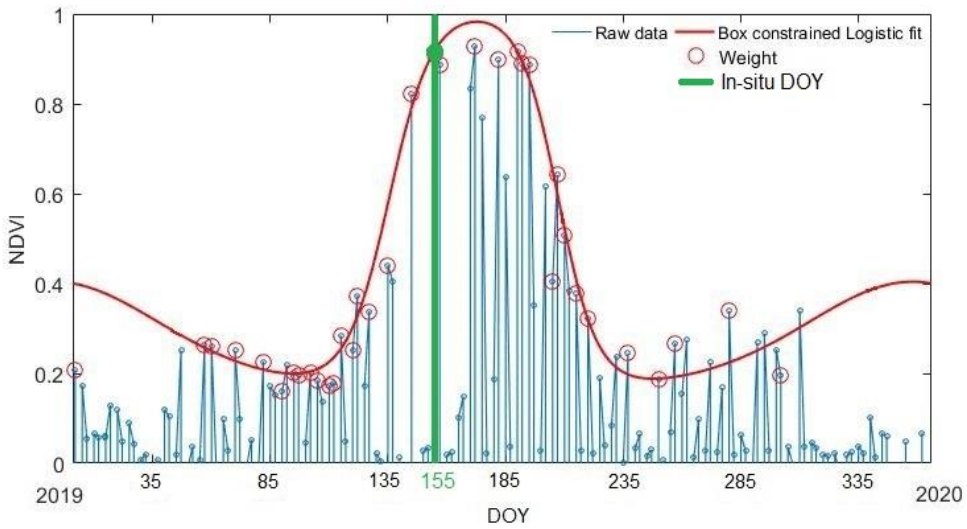

Figure 3.6 Example of NDVI time series of a pixel that represents the location of the in-situ measurements for one spring barley parcel (field number 91 in 2019). The achieved DL fit (red line) is shown along with the green point representing the crop growth stage 39DC at the optimal local threshold 92%. The green line represents the in-situ DOY for the crop growth stage 39DC.

Optimal global threshold definition

In order to assess the possibility of using a global threshold for detecting the crop growth stage of interest (31DC, 39DC, 51DC), the 10 spring barley parcels that allowed for a satisfying fit of the DL function were used as the sample (*N=10*) for the analysis. To find the optimal global threshold for each crop growth stage, different thresholds in

the range of 45% – 100% (incremented with a minimum step of 1%) were applied as in the study by Huang et al. (2019). In that way the optimal global threshold for each crop growth stage, was defined as the VI amplitude value that resulted in the highest accuracies between S2 derived DOY (NDVI and EVI2) and in-situ observed DOY for all of the studied spring barley parcels. The accuracy assessment of the S2 (NDVI and EVI2) results regarding the crop growth stages of interest was based on R-square (coefficient of determination) and Root Mean Square Error (RMSE; Equation 5, 6 respectively).

$$R^2 = \frac{cov(\hat{Y} - Y)^2}{var(\hat{Y})var(Y)}, \quad (5)$$

$$RMSE = \sqrt{\frac{\sum_i^N (\hat{Y}_i - Y_i)^2}{N}}, \quad (6)$$

where, $\hat{Y}_i$ corresponds to the S2 derived DOY for the crop growth stages of interest (31DC, 39DC, 51DC) for each spring barley parcel included in the sample denoted as N, Y_i refers to the DOY based on the in-situ measurements, $cov(\hat{Y} - Y)^2$ is the covariance between the S2 DOY and in-situ DOY, $var(\hat{Y})$ is the variance of the S2 DOY, and $var(Y)$ the variance of the in-situ DOY for the different crop growth stages.

R^2 expresses the amount of variation that is explained by the linear model and its value range between 0 for low performance models and 1 for models which are able to explain fully the variability of the dataset. The RMSE describes the average error (in days) between the S2 derived results and the in-situ observations (Huang et al. 2019, Liu and Huang 2019).

4. Results

4.1 The optimal local thresholds for detecting the crop growth stages of interest (31DC, 39DC, 51DC)

The NDVI and EVI2 single pixel time series of one spring barley field (field number 91 in 2019) are depicted in Figure 3.6 (NDVI) and Figure 4.1 (EVI2). For both VIs, the smoothed continuous VI curve that was reconstructed using the DL function is showed along with the raw data and the quality observations (based on S2 SCL data) marked as with the red circles. The achieved DL fit (Figure 3.6, Figure 4.1) was considered to depict the growing season of the spring barley parcel in a satisfactory way, where observations of high quality (marked in red circle) are supporting the DL fit. The optimal local NDVI threshold for that field (field number 91 in 2019) was defined as 79.8 % for the stage 31DC (resulting in the in-situ observed DOY 147), 92% for the stage 39DC (DOY 155; green point in Figure 3.6) and 97.4% for the stage 51DC (DOY 162).

For EVI2 the optimal local threshold for the aforementioned spring barley field (same exact pixel), equals to 64.5%, 84.2%, and 95% for the crop growth stages 31DC, 39DC (Figure 4.1) and 51DC respectively. In that way, the above thresholds created an agreement between the in-situ DOY and the EVI2 extracted phenometrics.

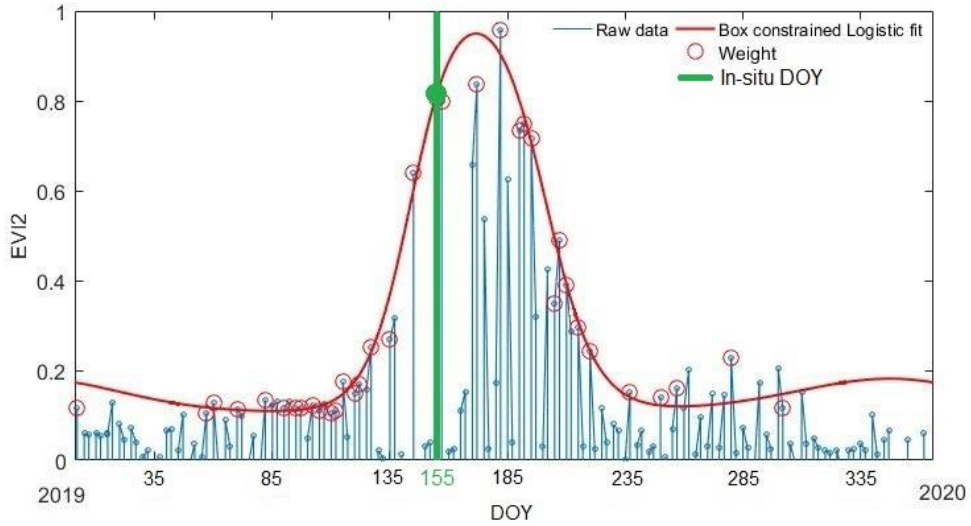

Figure 4.1 Example of EVI2 time series of a pixel that represents the location of the in-situ measurements for one spring barley parcel (field number 91 in 2019). The achieved DL fit (red line) is shown along with the green point representing the crop growth stage 39DC at the optimal local threshold 84.2%. The green line represents the in-situ DOY for the crop growth stage 39DC.

Table 4.1 shows the optimal local thresholds (NDVI, EVI2) for the crop growth stages (i) 31DC, (ii) 39DC, and (iii) 51DC for each spring barley field that allowed for an accurate fit of the DL function (*N=10)*. It is apparent that the optimal local thresholds

of the 10 spring barley fields show high variability (Table 4.1; Figure 4.2), where additionally, some optimal local thresholds (mostly for the stage 51DC) are very close to the peak of the season (100% threshold). More specifically, the variability of the optimal local thresholds for the 10 spring barley fields (in terms of standard deviation; SD) is showing a decreasing trend with latter crop growth stages (Figure 4.2, Table 4.1).

Table 4.1 Statistics of optimal local thresholds based on NDVI and EVI2 for the crop growth stages 31DC, 39DC, and 51DC for the 10 spring barley fields that were included in the analysis.

		Optimal Local Threshold (%)					
Field Number	**Sampling Year**	**NDVI 31DC**	**NDVI 39DC**	**NDVI 51DC**	**EVI2 31DC**	**EVI2 39DC**	**EVI2 51DC**
90	**2019**	53.9	83.5	-	46.7	80.5	-
91	**2019**	79.8	92	97.4	64.5	84.2	95
60	**2019**	63	97.7	99.9	54.5	94	99.3
62	**2019**	95	99	99.79	88.5	98.35	99.84
92	**2019**	95.65	98.75	99.998	76.7	89.1	99.5
152	**2019**	61	78.9	90	22	80.5	98.8
62	**2018**	62.3	76.2	87.7	67	79.8	90.2
92	**2018**	57	71	84.8	81.2	91.3	97.3
61	**2017**	90.5	97.3	99.58	73.5	87.5	96.4
90	**2017**	78.4	91.8	97.25	92.4	98.65	99.91
	Average	**73.6**	**88.6**	**95.1**	**66.7**	**88.3**	**97.3**
	SD	**16.1**	**10.4**	**5.9**	**21.1**	**7.1**	**3.1**
	Variance	**261.5**	**108.3**	**35.7**	**449.3**	**51.2**	**10.1**
	Maximum	**95.65**	**99**	**99.998**	**92.4**	**98.65**	**99.91**
	Minimum	**53.9**	**71**	**84.8**	**22**	**79.8**	**90.2**

The crop growth stages 39DC and 51DC based on EVI2 showed a smaller variability of optimal local thresholds compared to NDVI. However, the opposite applies for the crop growth stage 31DC, where EVI2 resulted in higher variability of optimal local thresholds compared to NDVI (SD = 21.1% and SD = 16.1% respectively).

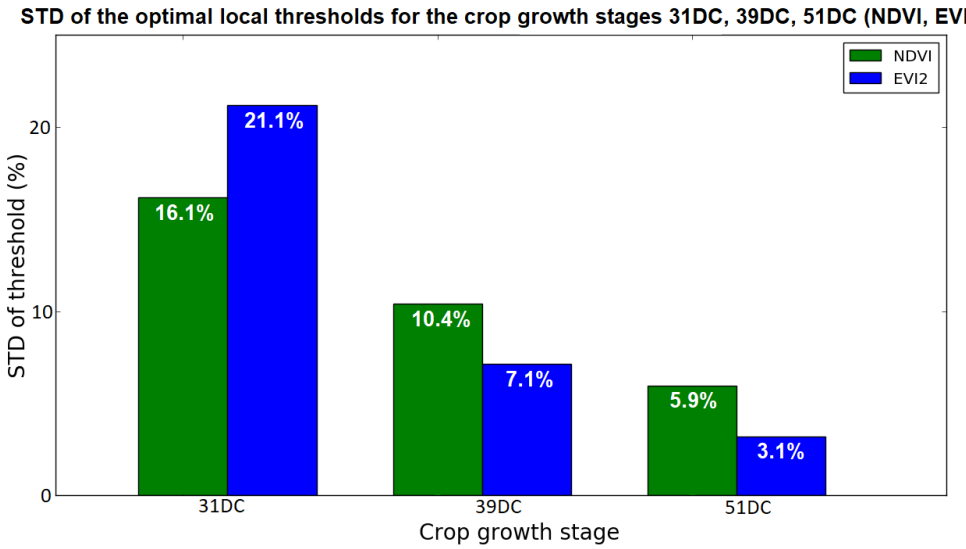

Figure 4.2 Standard deviation (SD) of the optimal local thresholds of the 10 spring barley fields regarding the crop growth stages 31DC, 39DC, and 51DC.

Such variability of the optimal local thresholds (unique threshold per field) can be explained by the quality of the DL fit, where for each of the barley fields the DL function is expected to have a varying reconstruction performance. Moreover, the decreasing variability of the optimal local thresholds with later crop growth stages also relates to the achieved DL fit, where it appears that the smoothed VI curve was more accurate the closer to the peak of the season (Figure 4.2). This could be possibly explained further by the higher probability of cloud coverage during the green up phase (May – early June) compared to the peak of the season (June – early July).

4.2 The optimal global threshold for detecting the crop growth stage – First node detectable (31DC)

For the crop growth stage 31DC (first node detectable) thresholds ranging from 45% – 90% were used in order to identify the threshold that induces the smallest RMSE against the in-situ observations of the 10 studied spring barley fields. The optimal global threshold for the crop growth stage 31DC was 74% for NDVI with a RMSE of 6.9 days, where for EVI2 the optimal global threshold was 70% inducing a RMSE of 5.1 days (Figure 4.3).

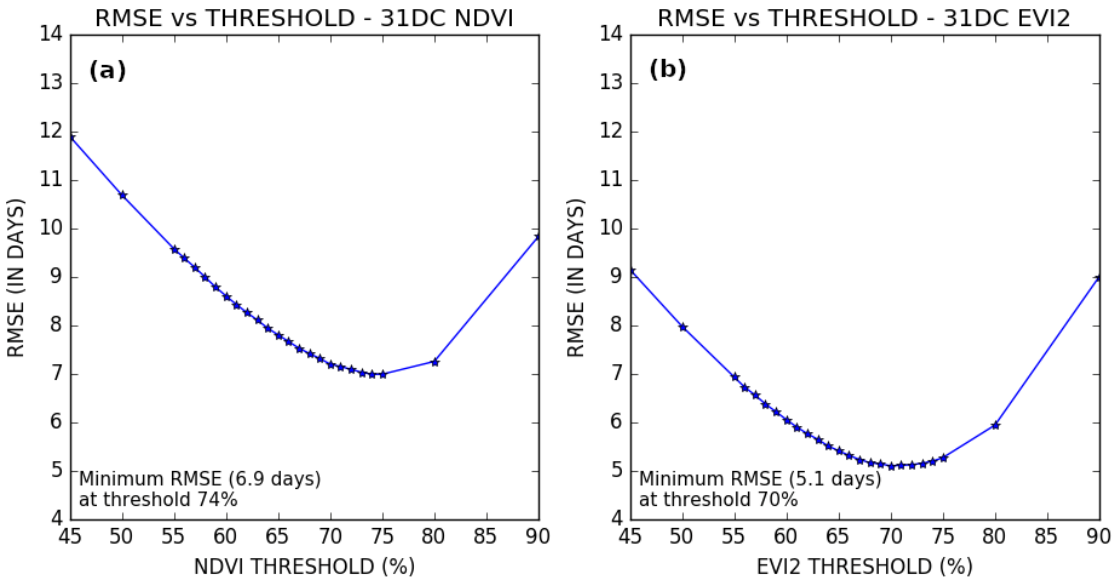

Figure 4.3 Incremented NDVI and EVI2 thresholds plotted against their induced RMSE (in days) for the crop growth stage 31DC (first node detectable). The optimal global threshold for NDVI (74%) induced a RMSE of 6.9 days (a), while the optimal global threshold for EVI2 (70%) induced a RMSE of 5.1 days (b).

According to the in-situ measurements, the crop growth stage 31DC for the studied spring barley parcels was recorded between DOY 147 and 170 (Table 4.2). The difference in days between the S2 derived DOY and the in-situ observed DOY for the crop growth stage 31DC varies from -13.2 to 6.4 days (range of 19.6 days; Table 4.2) for NDVI when applying the optimal global threshold (74%). For EVI2 (optimal global threshold 70%) the difference is smaller with a range between -8.2 and +7.2 days (range of 15.4 days; Table 4.2). The optimal global threshold for NDVI resulted in a R-square of 0.33 and a RMSE of 6.9 days (Figure 4.4), where the NDVI based DOY for the crop growth stage 31DC was between DOY 140.8 and 168.2.

Table 4.2 Studied fields with in-situ DOY for the crop growth stage 31DC (first node detectable), along with the S2 (NDVI, EVI2) derived DOY and RMSE when applying the optimal global threshold (74% NDVI; 70% EVI2) for the studied spring barley fields. The error in days of the global threshold for each studied spring barley field is based on VI DOY minus in-situ DOY.

Optimized Global Threshold for 31 DC (NDVI, EVI2)						
Field Number	**Sampling Year**	**In-situ DOY**	**NDVI 74% DOY**	**NDVI 74% Error (days)**	**EVI2 70% DOY**	**EVI2 70% Error (days)**
90	**2019**	147	153.1	6.1	153.5	6.5
91	**2019**	147	144.4	-2.6	149	2
60	**2019**	147	149.8	2.8	151.6	4.6
62	**2019**	154	140.8	-13.2	146.8	-7.2
92	**2019**	154	142.7	-11.3	151.4	-2.6
152	**2019**	154	160.4	6.4	161.2	7.2
62	**2018**	154	158.1	4.1	155.1	1.1
92	**2018**	148	154.1	6.1	144.4	-3.6
61	**2017**	156	149.9	-6.1	154.8	-1.2
90	**2017**	170	168.2	-1.8	161.8	-8.2
			NDVI RMSE	**6.9 days**	**EVI2 RMSE**	**5.1 days**

On the other hand, based on EVI2, using the optimal global threshold (70%) for the crop growth stage 31DC, resulted in a higher R-square (0.43) and lower RMSE (5.1 days; Figure 4.4), while the EVI2 observed DOY for the crop growth stage 31DC was between the DOY 144.4 and 161.8. In general, the offset of the S2 estimated phenological dates for the crop growth stage 31DC did not show any consistent pattern of underestimation or overestimation for both NDVI and EVI2 when compared to the in-situ measurements (Table 4.2, Figure 4.4).

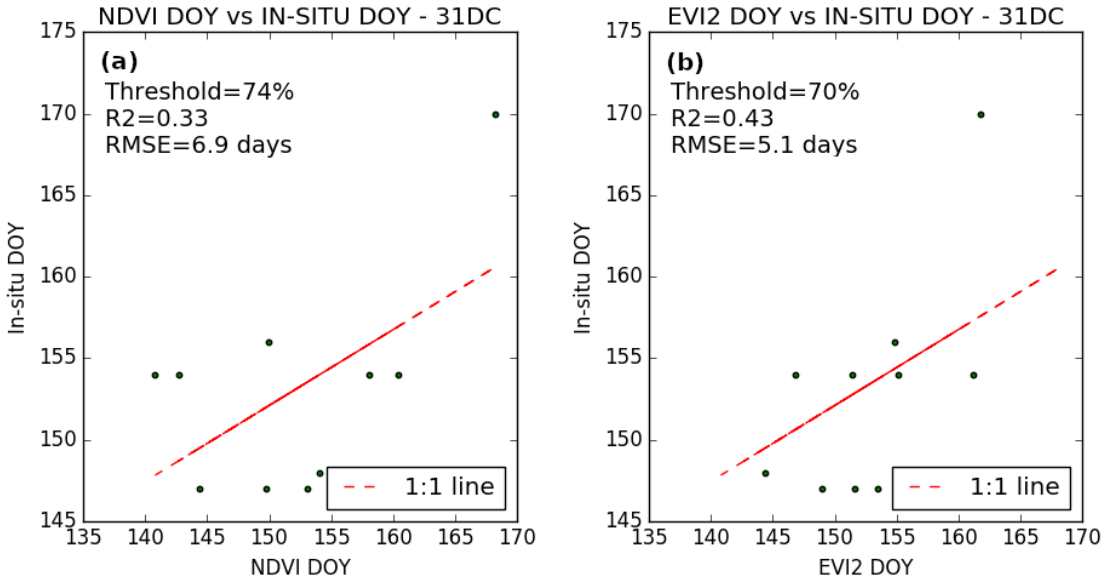

Figure 4.4 Comparison between detection of the crop growth stage 31DC (first node detectable) based on NDVI (a) and EVI2 (b; N=10). Scatterplot of NDVI derived DOY when applying the optimal global threshold (74%) and in-situ observed DOY (R2=0.33, RMSE = 6.9 days; a). Scatterplot of EVI2 derived DOY when applying the optimal global threshold (70%) and in-situ observed DOY (R2=0.43, RMSE = 5.1 days; b). The red line represents the 1:1 relationship.

Figure 4.5 and Figure 4.6 depict the resulted phenological maps regarding the crop growth stage 31DC for a studied spring barley parcel (field number 91 in 2019; NDVI and EVI2 respectively). The location of the in-situ measurements (single pixel) had an optimal local threshold of 79.8% for NDVI and 64.5% for EVI2 regarding the crop growth stage 31DC (resulting in DOY 147; Table 4.1). When subtracting the DOY of the local threshold from the DOY of the global threshold (Global threshold DOY minus Local threshold DOY) the difference in days for that pixel was -2.6 days for NDVI while the difference for EVI2 was +2 days (Figure 4.5b and Figure 4.6b respectively).

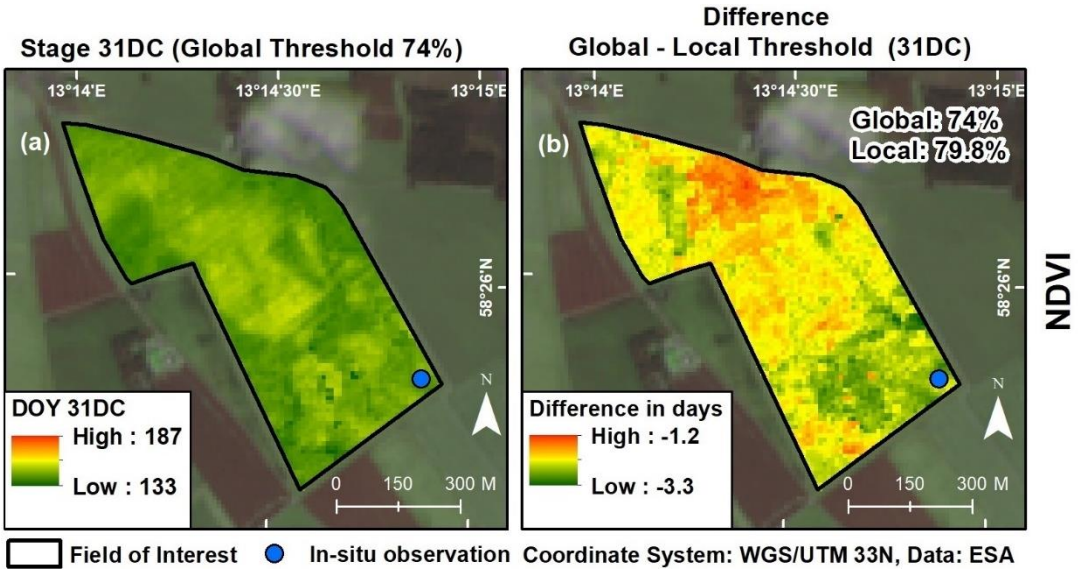

Figure 4.5 DOY for the crop growth stage 31DC derived from NDVI for the field number 91 in 2019 with a global threshold of 74% (a) and the difference in number of days between DOY based on the global minus local threshold (b).

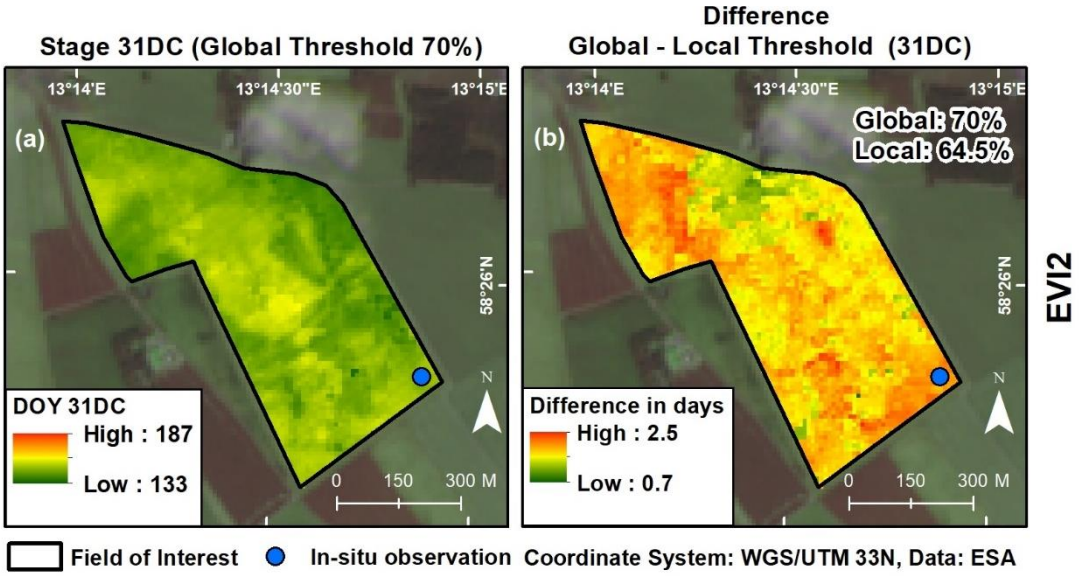

Figure 4.6 DOY for the crop growth stage 31DC derived from EVI2 for the field number 91 in 2019 with a global threshold of 70% (a) and the difference in number of days between DOY based on the global minus local threshold (b).

4.3 The optimal global threshold for detecting the crop growth stage – Flag leaf ligule just visible (39DC)

For the crop the crop growth stage 39DC (flag leaf ligule just visible) thresholds ranging from 70% – 99% were used in order to identify the threshold that induces the smallest RMSE against the in-situ observations of the 10 selected spring barley fields. The optimal global threshold for the crop growth stage 39DC (flag leaf ligule just visible) was 92% for NDVI with a RMSE of 7.6 days, where for EVI2 the optimal global threshold was 91% with a RMSE of 4.8 days (Figure 4.7)

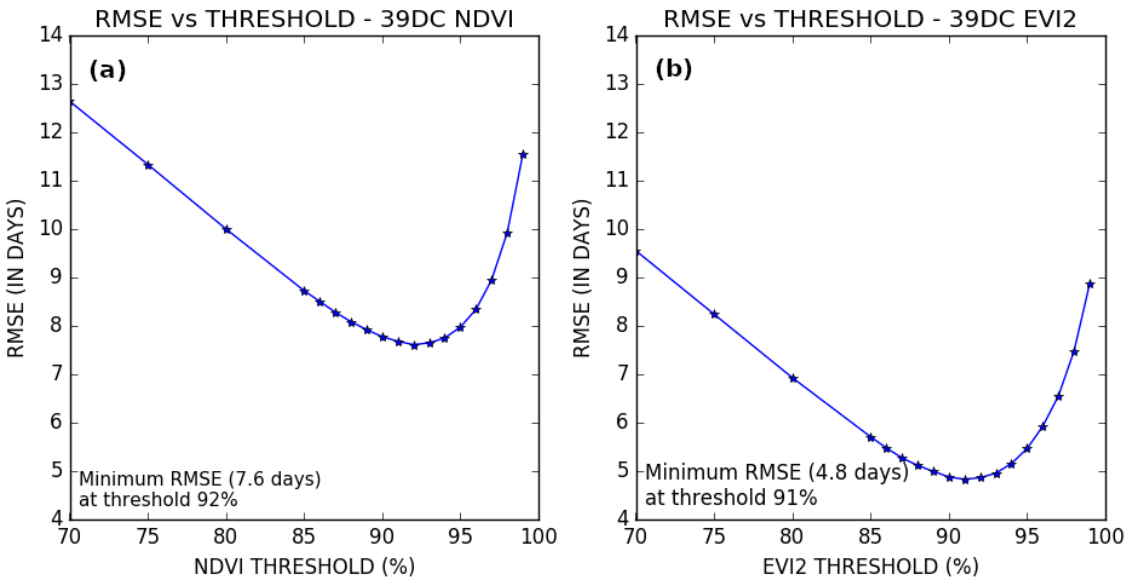

Figure 4.7 Incremented NDVI and EVI2 thresholds plotted against their induced RMSE (in days) for the crop growth stage 39DC (flag leaf ligule just visible). The optimal global threshold for NDVI (92%) induced a RMSE of 7.6 days (a), while the optimal global threshold for EVI2 (91%) induced a RMSE of 4.8 days (b).

According to the in-situ measurements the crop growth stage 39DC for the studied spring barley parcels was observed between the DOY 153 and 178 (Table 4.3). The difference in days between the S2 derived DOY and the in-situ observed DOY for the crop growth stage 39DC varies from -13.4 to +10.4 days (range of 23.8 days; Table 4.3) for NDVI when applying the optimal global threshold (92%). For EVI2 (optimal global threshold 91%) the difference is smaller with a range between -8.9 and +5.5 days (range of 14.4 days; Table 4.3). The optimal global threshold for NDVI resulted in a R-square of 0.27 and a RMSE of 7.6 days (Figure 4.8), where the NDVI based DOY for the crop growth stage 39DC was observed between DOY 150.6 and 178.2 (Table 4.3).

Table 4.3 Studied fields with In-situ DOY for the stage 39DC (flag leaf ligule just visible), along with the S2 (NDVI, EVI2) derived DOY and RMSE when applying the optimal global threshold (92% NDVI; 91% EVI2) for the studied spring barley fields. The error in days of the global threshold for each studied spring barley field is based on VI DOY minus in-situ DOY.

Optimized Global Threshold for 39 DC (NDVI, EVI2)						
Field Number	**Sampling Year**	**In-situ DOY**	**NDVI 92% DOY**	**NDVI 92% Error (days)**	**EVI2 91% DOY**	**EVI2 91% Error (days)**
90	**2019**	157	162.4	5.4	161.7	4.7
91	**2019**	155	155	0	158.9	3.9
60	**2019**	161	156.4	-4.6	159.4	-1.6
62	**2019**	164	150.6	-13.4	155.5	-8.5
92	**2019**	160	150.6	-9.4	161.2	1.2
152	**2019**	163	171.5	8.5	165.6	2.6
62	**2018**	159	166.4	7.4	164.5	5.5
92	**2018**	153	163.4	10.4	152.8	-0.2
61	**2017**	162	156.9	-5.1	163.9	1.9
90	**2017**	178	178.2	0.2	169.1	-8.9
			NDVI RMSE	**7.6 days**	**EVI2 RMSE**	**4.8 days**

Based on EVI2, the optimal global threshold (91%) for the crop growth stage 39DC resulted in higher R-square (0.45) and lower RMSE (4.8 days; Figure 4.8) compared to NDVI, while the EVI2 derived DOY for the crop growth stage 39DC was between the DOY 152.8 and 169.1 (Table 4.3). In general, likewise the results for the stage 31DC, the offset of the S2 estimated phenological dates for the crop growth stage 39DC did not show any consistent pattern of underestimation or overestimation for both NDVI and EVI2 when compared to the in-situ measurements (Table 4.3, Figure 4.8).

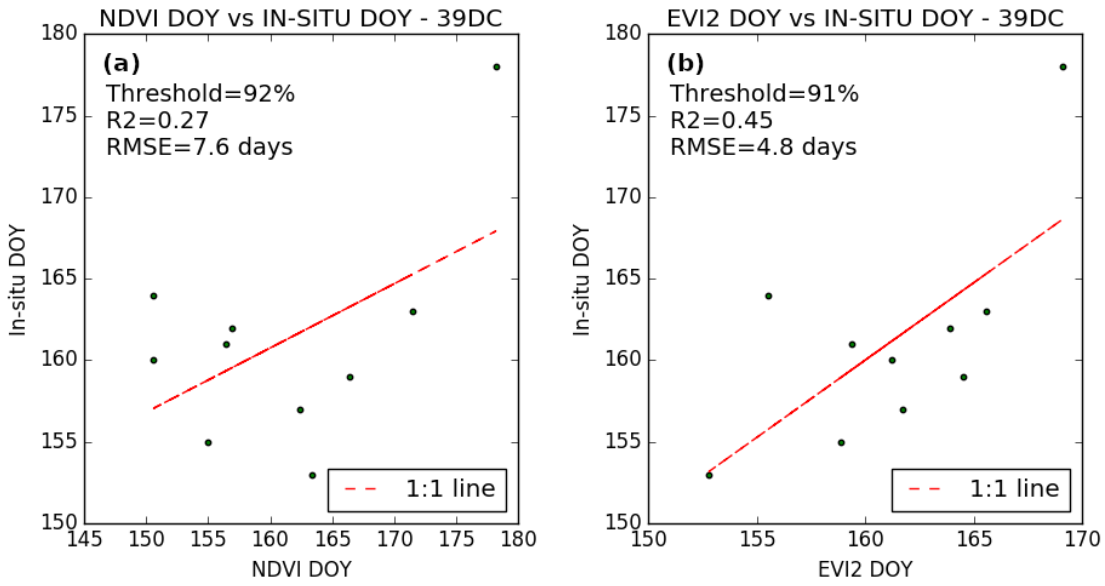

Figure 4.8 Comparison between detection of the crop growth stage 39DC based on NDVI (a) and EVI2 (b) for each spring barley field (N=10). Scatterplot of NDVI derived DOY when applying the optimal global threshold (92%) and in-situ observed DOY (R2=0.27, RMSE=7.6 days; a). Scatterplot of EVI2 derived DOY when applying the optimal global threshold (91%) and in-situ observed DOY (R2=0.45, RMSE= 4.8 days; b). The dashed line represents the 1:1 relationship.

Figure 4.9 illustrates the resulted phenological maps regarding the crop growth stage 39DC for one of the studied spring barley parcels (field number 91 in 2019) based on NDVI (a) and EVI2 (c). The location of the in-situ measurements (single pixel) had an optimal local threshold of 92% for NDVI and 84.2% for EVI2 regarding the crop growth stage 39DC (resulting in DOY 155; Table 4.1). When subtracting the DOY of the local threshold pixel value from the DOY of the global threshold (Global threshold DOY – Local threshold DOY) the difference in days for that pixel was 0 days for NDVI while the difference for EVI2 was +3.9 days (Figure 4.9b and Figure 4.9d respectively).

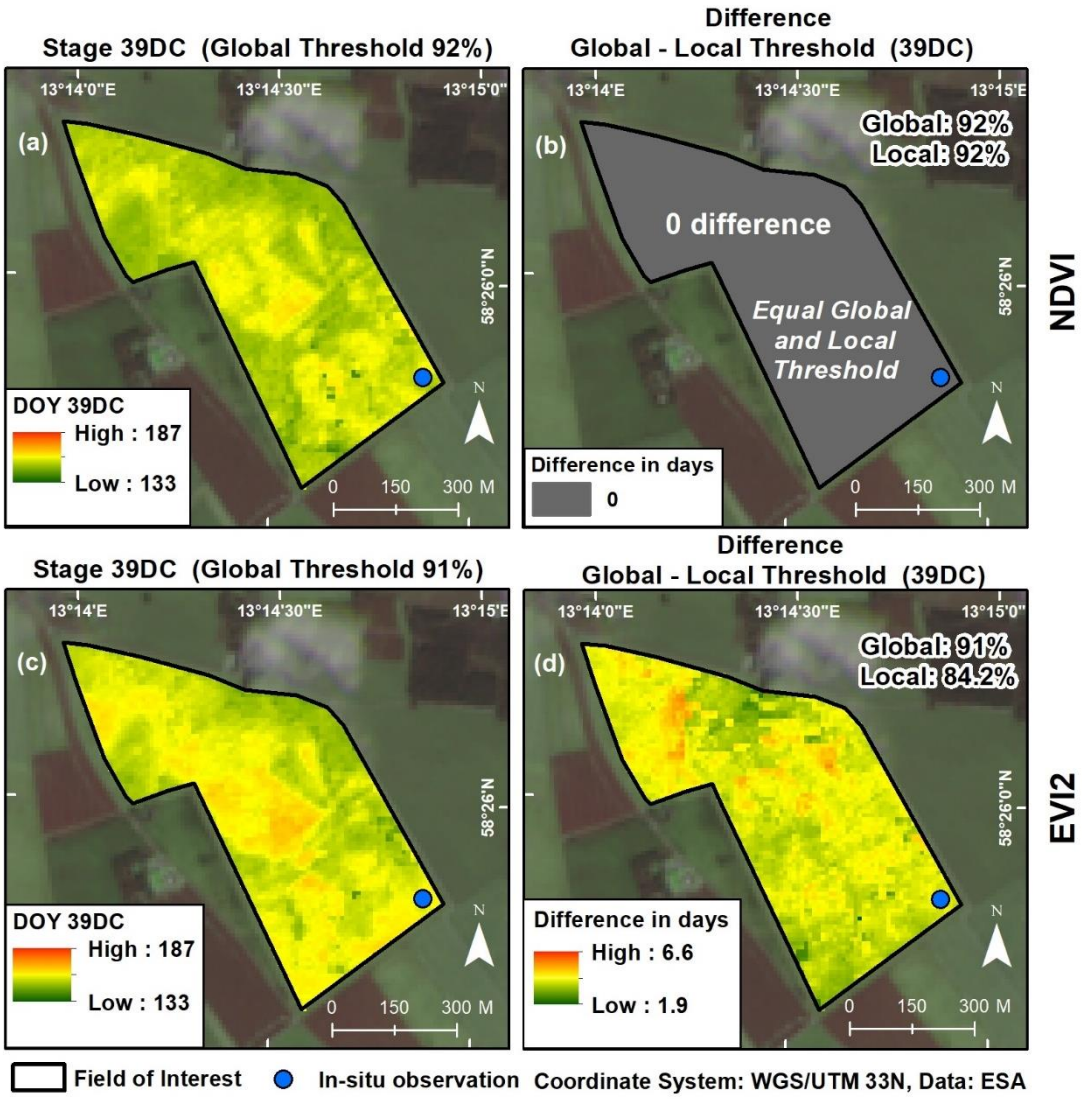

Figure 4.9 DOY for the crop growth stage 39DC derived from NDVI for the field 91in 2019 with a global threshold of 92% (a) and the difference in number of days between DOY based on the global minus local threshold (b), DOY for the crop growth stage 39DC derived from EVI2 with a global threshold of 91% (c) and the difference in number of days between DOY based on the global minus local threshold (d).

4.4 The optimal global threshold for detecting the crop growth stage – First spikelet of inflorescence just visible (51DC)

Regarding the crop growth stage 51DC (first spikelet of inflorescence just visible) thresholds ranging from 90% – 100% (peak of the season) were investigated in order to identify the threshold that results in the smallest RMSE against the in-situ measurements of the 10 studied spring barley fields. The optimal global threshold for both NDVI and EVI2 was 99% for the crop growth stage 51DC with a RMSE of 7.2 for NDVI and 4.2 days for EVI2 (Figure 4.10).

As seen in Figure 4.10, for the crop growth stage 51DC, the resulting RMSE shows an almost linear decreasing trend as the threshold value is approaching the peak of the season (100% threshold) followed by a rapid increase after the optimal global threshold

(99%). Consequently, when using a threshold value that is greater than 99% (i.e. close to the peak of the season when the DL function has a flatter shape) adjusting the threshold value will result in a greater change of RMSE than when using thresholds less than 99%.

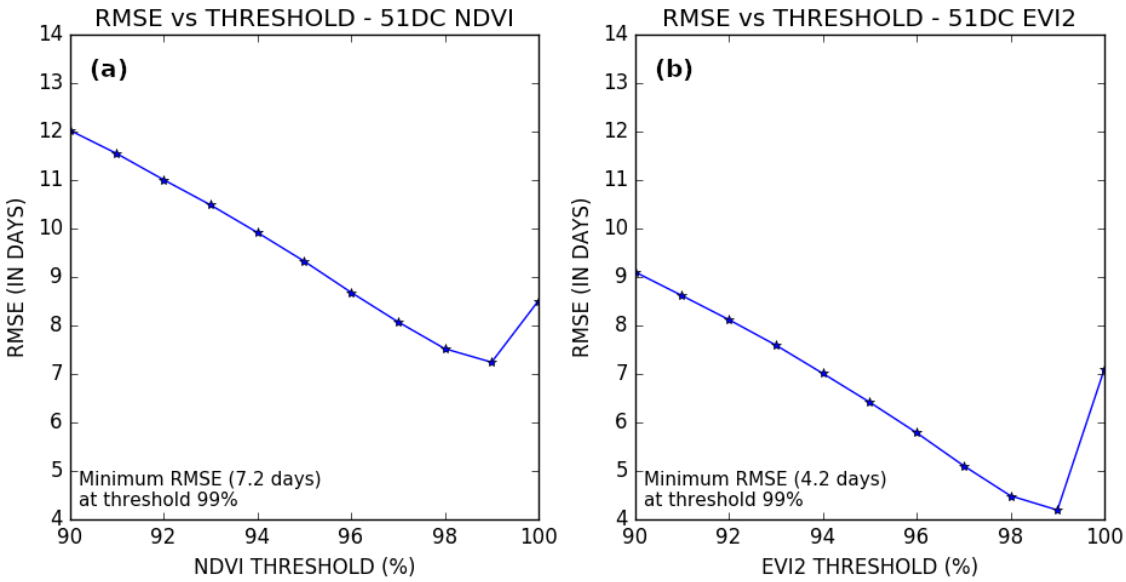

Figure 4.10 Incremented NDVI and EVI2 thresholds plotted against their induced RMSE (in days) for the crop growth stage 51DC (first spikelet of inflorescence just visible). The optimal global threshold for NDVI (99%) induced a RMSE of 7.2 days (a), while the optimal global threshold for EVI2 (99%) induced a RMSE of 4.2 days (b).

One of the studied spring barley fields did not include any in-situ observations past the stage 49DC (field number 90 in 2019), meaning that it was not possible to interpolate the DOY for stage 51DC. Therefore, for the crop growth stage 51DC the sample size is 9 spring barley fields (*N=9;* Table 4.4). According to the in-situ measurements, the crop growth stage 51DC for the studied spring barley parcels was observed between the DOY 162 and 185. The difference in days between the S2 derived DOY and the in-situ observed DOY for the crop growth stage 51DC varies from -9.1 to +12.2 days (range of 21.3 days; Table 4.4) for NDVI when applying the optimal global threshold (99%). For EVI2 (optimal global threshold 99%) the difference is smaller with a range between -5.9 and +7.1 days (range of 13 days; Table 4.4).

Table 4.4 Studied fields with in-situ DOY for the stage 51DC (first spikelet of inflorescence just visible), along with the S2 (NDVI, EVI2) derived DOY and RMSE when applying the optimal global threshold (99% NDVI; 99% EVI2) for the studied spring barley fields. The error in days of the global threshold for each studied spring barley field is based on VI DOY minus in-situ DOY.

Optimized Global Threshold for 51DC (NDVI, EVI2)						
Field Number	**Sampling Year**	**In-situ DOY**	**NDVI 99% DOY**	**NDVI 99% Error (days)**	**EVI2 99% DOY**	**EVI2 99% Error (days)**
90	2019	-	-	-	-	-
91	2019	162	165.9	3.9	167.2	5.2
60	2019	166	163	-3	165.5	-0.5
62	2019	170	164	-6	165.8	-4.2
92	2019	170	160.9	-9.1	168.9	-1.1
152	2019	170	179	9	170.3	0.3
62	2018	164	172.5	8.5	171.1	7.1
92	2018	159	171.2	12.2	162.9	3.9
61	2017	168	165.6	-2.4	171.4	3.4
90	2017	185	189.6	4.6	179.1	-5.9
			NDVI RMSE	**7.2 days**	**EVI2 RMSE**	**4.2 days**

The optimal global threshold for NDVI (99%) resulted in a R-square of 0.39 and a RMSE of 7.2 days (Figure 4.11a), where the NDVI derived DOY for the crop growth stage 51DC was observed between DOY 163 and 189.6 (Table 4.4). For EVI2 the optimal global threshold for the stage 51DC (99%) resulted in R-square of 0.69 and a RMSE of 4.2 days (Figure 4.11b), where the EVI2 derived DOY for the crop growth stage 51DC falls between DOY 162.9 and 179.1 (Table 4.4). As for the stages 31DC and 39DC, the offset of the estimated S2 phenological dates for the crop growth stage 51DC did not show any consistent underestimation or overestimation pattern for both NDVI and EVI2 when compared to the in-situ measurements (Table 4.4, Figure 4.11).

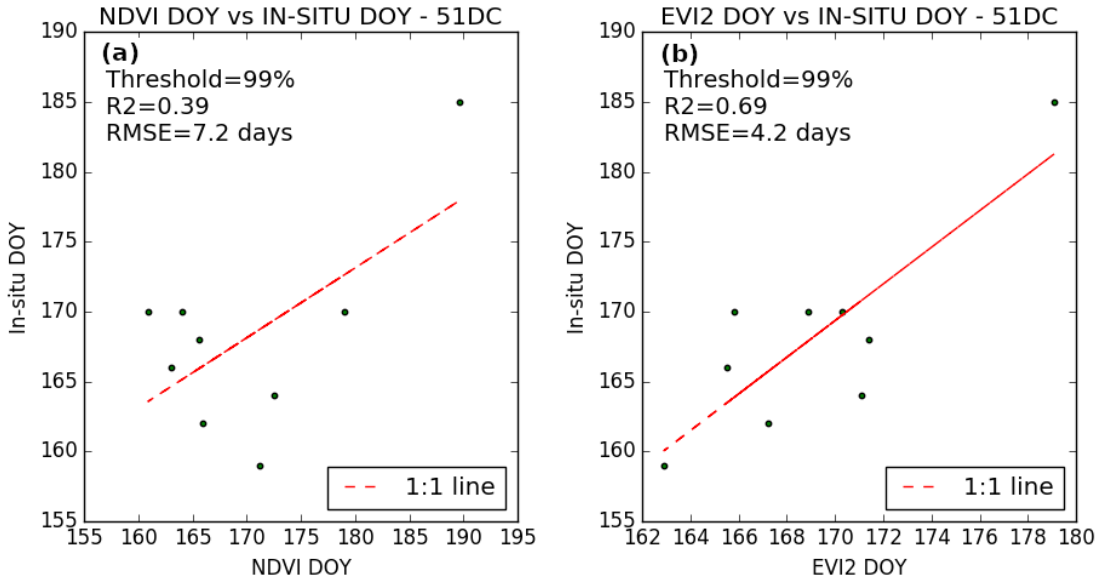

Figure 4.11 Comparison between detection of the crop growth stage 51DC based on NDVI (a) and EVI2 (b) for each spring barley field with in-situ measurements (N=9). Scatterplot of NDVI derived DOY when applying the optimal global threshold (99%) and in-situ observed DOY (R2=0.39, RMSE=7.2 days; a). Scatterplot of EVI2 derived DOY when applying the optimal global threshold (99%) and in-situ observed DOY (R2=0.69, RMSE = 4.2 days; b). The dashed line represents the 1:1 relationship.

Figure 4.12 and Figure 4.13 depict the phenological maps for a studied spring barley parcel (field number 91 in 2019) regarding the crop growth stage 51DC (NDVI and EVI2 respectively). The location of the in-situ measurements (single pixel) had an optimal local threshold of 97.4% for NDVI and 95% for EVI2 for the crop growth stage 51DC (resulting in DOY 162; Table 4.1). When subtracting the DOY of local threshold from the DOY of the global threshold (Global threshold DOY – Local Threshold DOY) the difference in days for that single pixel was +3.9 for NDVI and +5.2 days for EVI2. Another example of phenological maps regarding the crop growth stages of interest (31DC, 39DC, 51DC) for one of the studied spring barley parcels (field number 60 in 2019) is shown in Appendix Figure S1 (EVI2) and Figure S2 (NDVI).

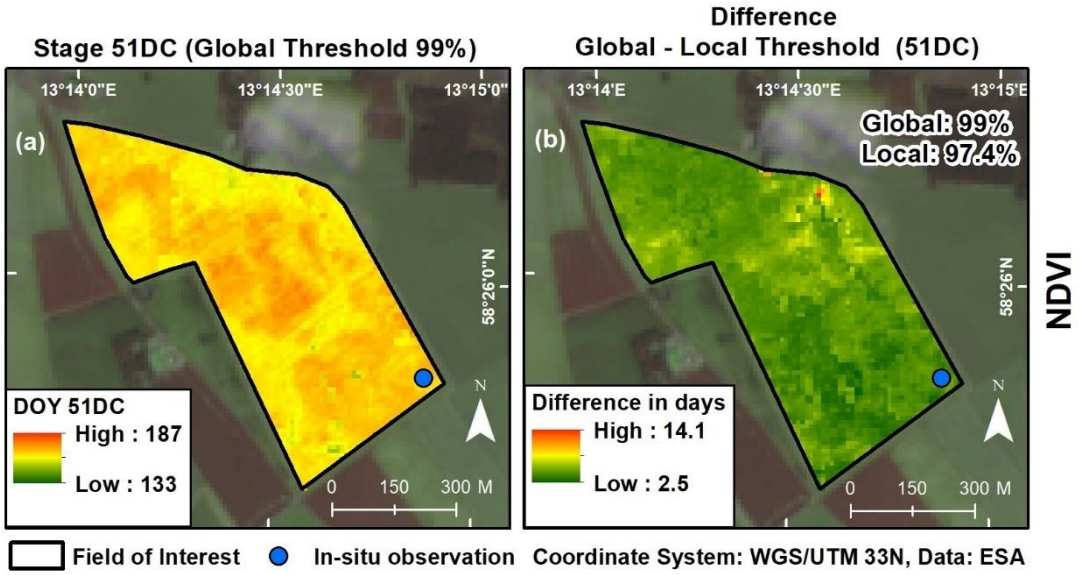

Figure 4.12 DOY for the crop growth stage 51DC derived from NDVI for field 91 in 2019 with a global threshold of 99% (a) and the difference in number of days between DOY based on the global minus local threshold (b).

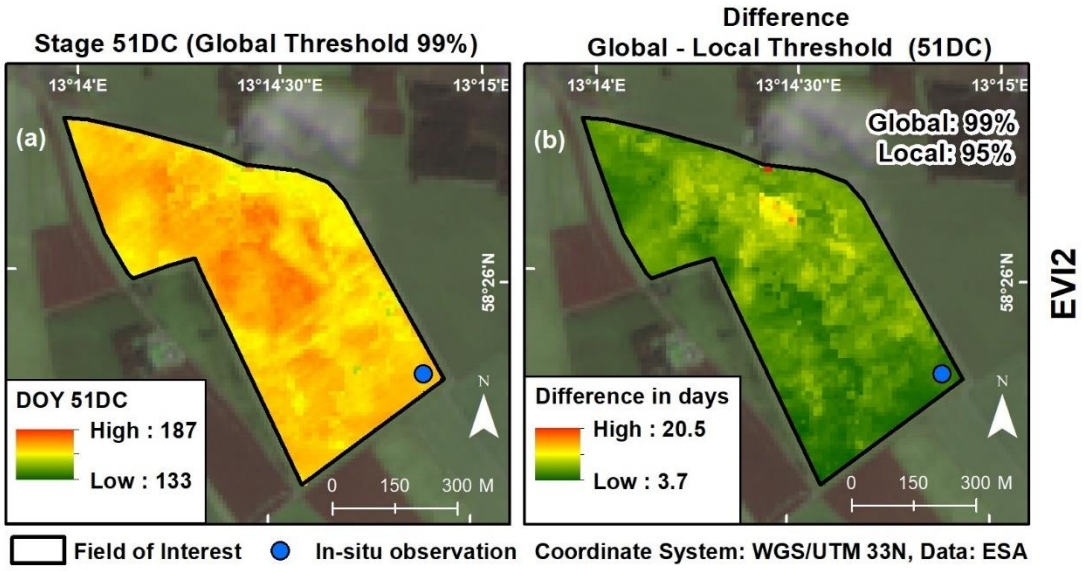

Figure 4.13 DOY for the crop growth stage 51DC derived from EVI2 for field 91 in 2019 with a global threshold of 99% (a) and the difference in number of days between DOY based on the global minus local thresholds (b).

4.5 Comparison of the retrieved crop growth stages based on NDVI and EVI2

After retrieving the crop growth stages of interest (31DC, 39DC, 51DC) based on NDVI and EVI2 data using the optimal global thresholds (defined in Section 3.3.4) and as also shown by the study of Huang et al. (2019) the comparison of the two VIs results indicate that NDVI is having slightly higher optimal (global) threshold values than EVI2 (Figure 4.14).

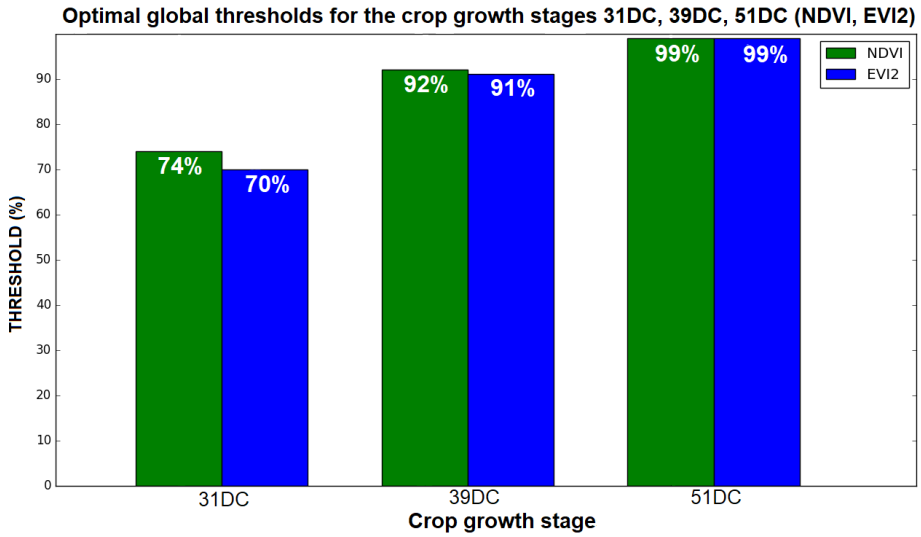

Figure 4.14 Values of the optimal global threshold as defined by the threshold optimization for NDVI (green) and EVI2 (blue) for deriving the crop growth stages of interest (31DC, 39DC, 51DC).

The difference in the values of the optimal global threshold is showing a decreasing trend with later crop growth stages. The highest difference was retrieved for the stage 31DC with the optimal global threshold being 74% and 70% for NDVI and EVI2 respectively. Regarding the crop growth stage 39DC, EVI2 resulted in lower optimal threshold (91%) compared to NDVI (92%), while for the crop growth stage 51DC the two VIs had the same optimal global threshold (99%; Figure 4.14).

EVI2 resulted in higher R-square values than NDVI when applying the optimal global thresholds for detecting the crop growth stages of interest (31DC, 39DC, 51DC). The highest R-square (0.69) was achieved with EVI2 for the crop growth stage 51DC, while the highest R-square achieved with NDVI was also for the crop growth stage 51DC. The lowest R-square values were found for the stage 39DC when using NDVI (R-square = 0.27), while for EVI2 the lowest R-square was induced for the crop growth stage 31DC (R-square = 0.43; Figure 4.15).

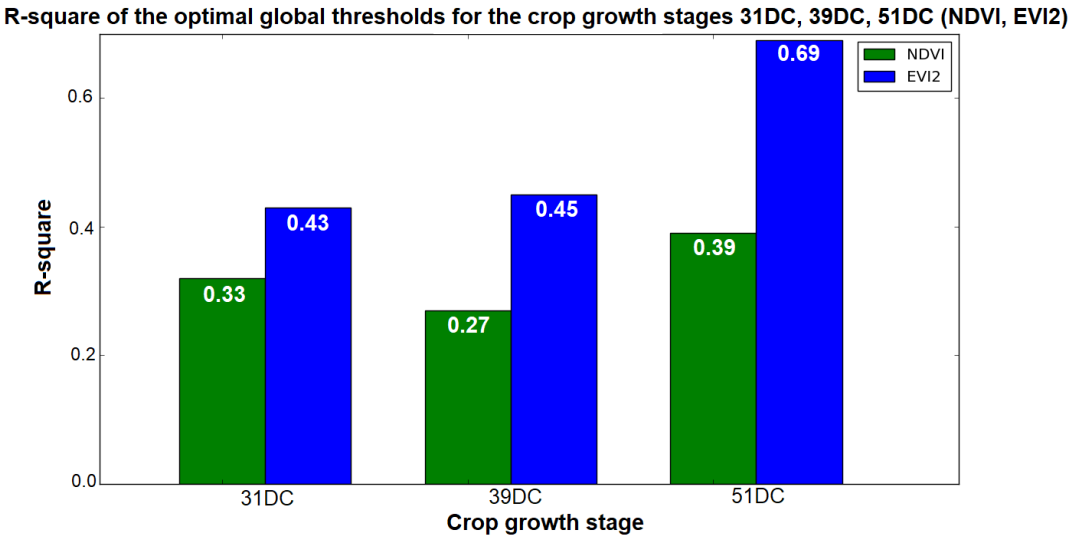

Figure 4.15 Optimized Global thresholds for the crop growth stages (31DC, 39DC, 51DC) and corresponding R-square values for NDVI and EVI2.

Regarding the RMSE of the optimal global threshold for NDVI and EVI2 for the different crop growth stages (Figure 4.16), the lowest RMSE was achieved for the stage 51DC using EVI2 (4.2 days), where the highest RMSE (7.6 days) was induced when using NDVI (optimal global threshold 92%) for the detecting the crop growth stage 39DC. The highest RMSE for EVI2 (5.1 days) was acquired for deriving the stage 31DC (optimal global threshold 70%), while the highest RMSE for NDVI was induced for the stage 39DC where the optimal global threshold was defined as 92%. The difference between NDVI and EVI2 induced RMSE for the different crop growth stages 31DC, 39DC, and 51DC was 1.8 days, 2.8 days, and 3 days respectively.

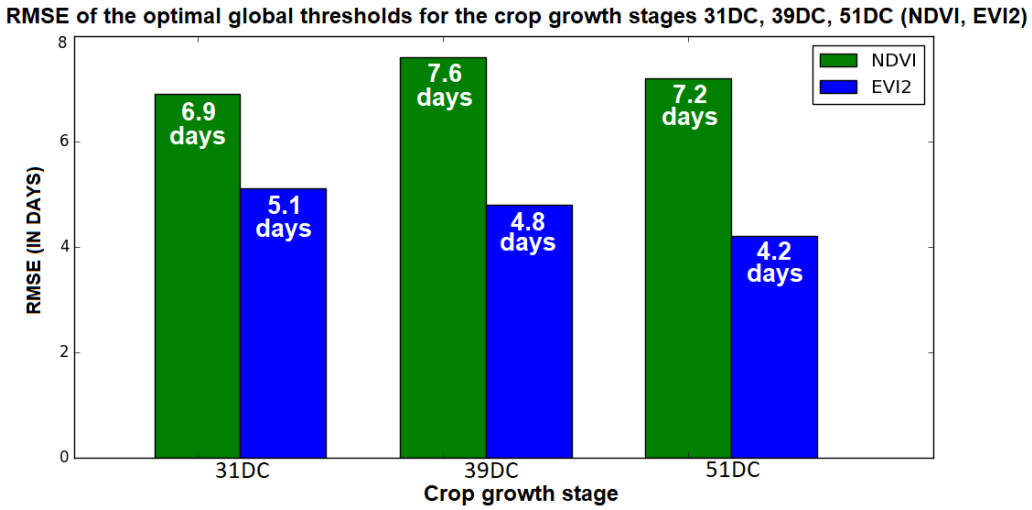

Figure 4.16 Optimized Global thresholds for the crop growth stages (31DC, 39DC, 51DC) and corresponding RMSE values for NDVI and EVI2.

In general, for all of the different crop growth stages (31DC, 39DC, 51DC) EVI2 resulted in higher R-square and smaller RMSE than NDVI when using the optimal global thresholds. These results describe the performance accuracies of the two VIs and suggest that EVI2 performs better than NDVI for detecting the crop growth stages 31DC, 39DC and 51DC of spring barley.

5. Discussion

5.1 Reconstruction of S2 time series and crop phenological metric extraction for crop growth stage detection.

In this study, the accuracies of S2 time series (NDVI, EVI2) for detecting the three crop growth stages of interest (31DC, 39DC, 51DC) were analyzed. The results of this study, suggest that EVI2 is more accurate than NDVI in detecting the crop growth stages 31DC (first node detectable), 39DC (flag leaf ligule just visible), and 51DC (first spikelet of inflorescence just visible). EVI2 resulted in a RMSE of 4.2 – 5.1 days for the different crop growth stages, while the RMSE for NDVI ranged between 6.9 – 7.6 days. Moreover, EVI2 performed better than NDVI in terms of R-square for all the studied crop growth stages (31DC, 39DC, 51DC).

The results, at the same time are depicting the inaccuracies caused by the common drawbacks of optical data (i.e. S2), that result from cloud coverage and other noise caused by particles of the atmosphere. Cloud coverage is drastically reducing optical data availability, while smaller noise (i.e. shadows, aerosol contamination) are also affecting the acquisition of dense high quality time series (Nasrallah et al. 2019; Stendardi et al. 2019; d'Andrimont et al. 2020; Gao et al. 2020; Zeng et al. 2020).

As explained in Section 3.3.4, noisy signals led to low quality S2 acquisitions during the growing season which did not allow for the inclusion of 11 spring barley fields in the studied parcel sample (out of the 21 spring barley parcels with available in-situ measurements). That was because the DL fit for these (11) fields was considered poor (Figure 3.4 NDVI; Figure 3.5 EVI2). This limited the sample size of spring barley fields that was used in the analysis down to 10 (*N=10*). The aforementioned disadvantages of S2 have been described in various studies (Vrieling et al. 2018; Nasrallah et al. 2019; Stendardi et al. 2019; Zeng et al. 2020) and were also apparent throughout the reconstruction of the S2 time series. Nonetheless, it must be noted, that S2 appears to be capable of monitoring crop phenology and its spatial variability at parcel scale. As reported by Nasrallah et al. (2019) and Xu et al. (2017), the time that cereal crops are reaching a specific crop growth stage is not the same within a field as a result of different conditions. This is obvious in the resulted phenological maps of the studied spring barley parcels (in the results section) where the within parcel variability in terms of DOY for reaching the crop growth stages 31DC, 39DC and 51DC is depicted (Figures 4.5, 4.6, 4.9, 4.12, 4.13, Appendix Figure S1). On the other hand, satellites

with a lower temporal resolution (e.g. Landsat) would provide sparser time series, while satellites with a daily revisit are usually having coarse spatial resolution (e.g. MODIS) and thus, do not enable monitoring of agricultural systems at parcel scale (Pan et al. 2015). Based on the above, it is apparent that factors like the availability and quality of data are having a great impact on the potential of S2 in detecting different crop growth stages (Gao et al. 2020).

Applying quality indicators on the raw time series data is a common pre-processing step (Zeng et al. 2020). Therefore, before reconstructing the VI time series with the use of the DL function, the Scene Classification Layer (SCL) was utilized as quality flag of the S2 data to filter out cloud contaminated observations. It is important to consider though, that such quality data (i.e. SCL) are also including errors and thus the time series data used in this study were far from noise free. An example of an inaccurate fit of the DL function that was induced by the misclassification of a pixel as good quality (i.e. vegetation) data when it was obviously contaminated by clouds is shown in Figure 5.1. After inspecting the SCL data for that pixel, it was assessed that moving the location of the in-situ measurements to the adjacent pixel (where according to the SCL it was classified as cloud and thus assigned a weight of 0) would result in a more visually satisfying fit of the DL function. That was performed in order to include the aforementioned field in the threshold optimization analysis. Figure 5.2 shows the EVI2 single pixel time series of the new location of the in-situ measurements for that spring barley field (field 90 in 2017).

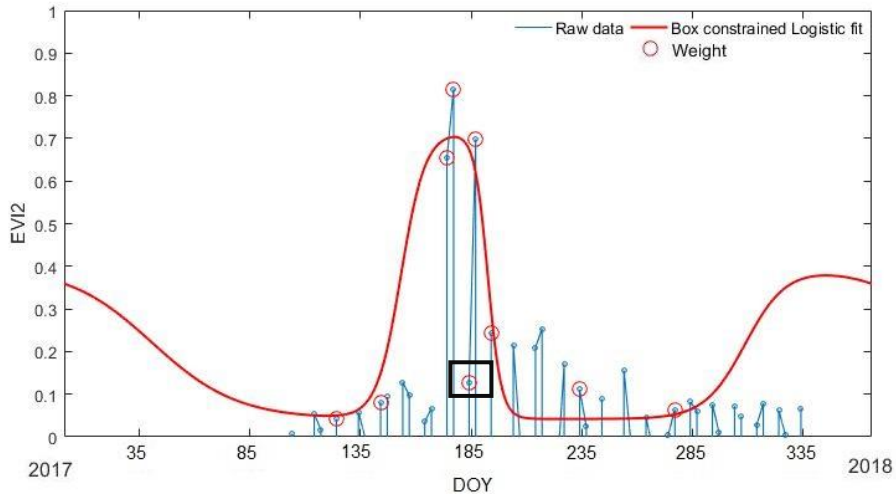

Figure 5.1 EVI2 time series of the original location of the in-situ observations for the field number 90 in 2017. Showing the cloud contaminated observation marked as good quality (in the black rectangle) which falls around the peak of the season.

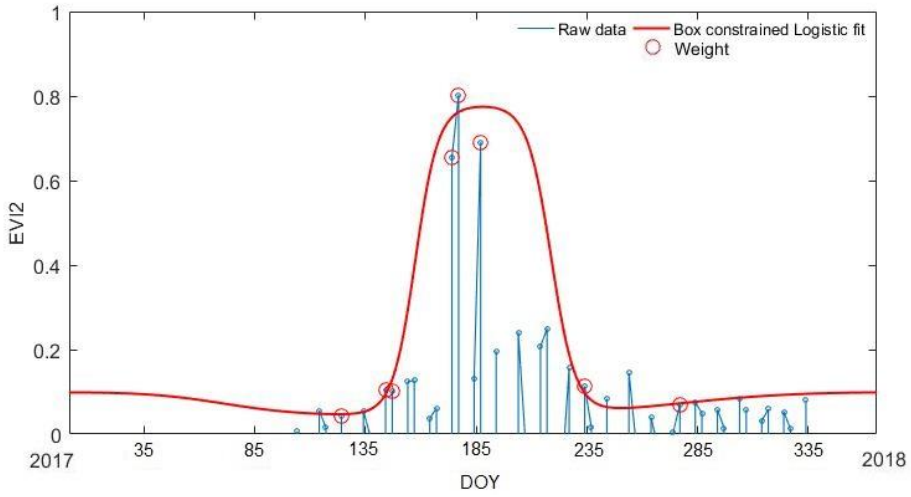

Figure 5.2 EVI2 time series of the new location (moved) of the in-situ observations for the field number 90 in 2017.

While other methods of cloud pre-filtering could have been adopted, such as the acquisition of S2 images based on cloud coverage percentage, it was decided that in that way it would not be possible to guarantee a reasonable exclusion of noisy S2 data from the time series. That is because, S2 scenes can have a high cloud coverage (e.g. 70%), but the agricultural parcels of interest (pixels within the parcel) could be cloud free. Moreover, the use of a range of weights (0 to 1) instead of binary weights (0 or 1) could have been applied for pre-filtering the raw time series data, but with questionable benefits with respect to the time and reasoning that is needed for selecting different weight values.

Based on the above, it is obvious that S2 has some strong drawbacks when used in phenology studies (i.e. cloud contamination). Therefore, several studies have focused on countering such problems through the combination of data (i.e. data fusion) from different optical and radar sensors (De Bernardis et al. 2016a; Gao et al. 2017; Stendardi et al. 2019). Data fusion methods, especially with the use of radar sensors (which are not affected by cloud contaminations) were found to effectively increase the amount of available observations throughout the growing cycle (De Bernardis et al. 2016a; Nasrallah et al. 2019; Stendardi et al. 2019). Therefore, it is of high interest for future studies to investigate the performance of such data fusion algorithms for acquiring denser time series, that can be used for extracting phenological metrics which relate to different crop growth stages on a more accurate basis.

Another factor that has a great influence on the extracted phenological results is the selection of the time series reconstruction method (Atkinson et al. 2012; Liu and Zhan 2016; Cai et al. 2017). Among the many smoothing methods that have been trialed in

various crop phenological studies (SG, AG, etc.) one might overperform others in different environments or crop types (Zeng et al. 2020). In this study, the selection of the DL function was based on the fact that DL is capable of capturing accurately the phenological cycle of crops while it requires the definition of fewer fitting parameters when compared to other smoothing functions (e.g. SG, AG; Liu and Zhan 2016; Wei et al. 2016; Cai et al. 2017).

Moreover, as reported by Cai et al. (2017), beside the selection of the smoothing method that is usually based on the characteristics of the data and the purposes of the study, the selection of the smoothing parameters is of high importance for reaching the maximum potential of a smoothing algorithm. Cai et al. (2017) tested a vast amount of different combinations of parameters and they described in detail how the selection of fitting parameters such as those that relate to the adaption of the reconstructed VI curve to the upper envelope can influence the accuracy of the results. Furthermore, they explained that such parameters are defined empirically and are known for having a great influence on the extracted phenometrics (e.g. DOY of different crop growth stages). The parameters used for the fitting of the DL function such as the (i) seasonal parameter, (ii) number of envelope iterations and (iii) adaption strength (Table 3.5) were thus based on the empirical judgement of the characteristics of the S2 time series (NDVI, EVI2). Therefore, potential directions for a future study could be to assess if different fitting settings when using the DL function would alter the results significantly, and also as Wei et al. (2016) and Liu and Zhan (2016) suggest, to further investigate the use of other reconstruction methods (e.g. SG, AG) in the concept of crop growth stage detection.

Additionally, it is to consider the reliability of the method used (i.e. dynamic thresholds) for the extraction of the phenological dates and how the results would differ if other methods were deployed. Even though the use of dynamic thresholds is among the most popular approaches for extracting phenological metrics (You et al. 2013; Pan et al. 2015), it is a simple and empirical method (Zeng et al. 2020). Huang et al. (2019) presented an enhanced version of the dynamic threshold in order to compensate for the effect of the different crop management practices in their study area. Thus, it is to consider in the future the use of other more sophisticated and flexible techniques for extracting crop phenological metrics that are based on the changing properties (e.g. maximum slope, inflection point) of the reconstructed VI trajectory (You et al. 2013; Zeng et al. 2020).

Lastly, as noted in the introduction section, RS data enable the approximation of the crop growth without providing any direct links to crop growth stages (Zeng et al. 2020). Consequently, it might not be possible to detect some specific crop stages (Gao et al. 2017) that correspond to smaller crop biophysical changes than the commonly studied phenological events (e.g. SOS, EOS). Therefore, in such cases, it could be more efficient to use other kinds of models (i.e. Growing Degree Days, GDD; which simulates the crop growth based on temperature data; De Bernardis et al. 2016b; Gao et al. 2017). Considering the above, future lines of questioning should be based on assessing which crop growth stages that are useful in agronomic management can be detected more accurately with the use of S2 data.

5.2 Validation of results with in-situ measurements

In order to assess the accuracy of the S2 results regarding the crop growth stages of interest (31DC, 39DC, 51DC), the S2 derived phenometrics were compared with the in-situ measurements of the spring barley parcels. It is important to note that, the crop growth stages of interest were not always recorded in the in-situ measurements protocol, and hence, it was needed to approximate the date of the crop growth stages of interest by performing linear interpolation (as described in Section 3.2.2), but since Zadoks DC is a non-linear scale (Wang and Engel 1998; Ahmed and Stockle 2017) such method is expected to introduce uncertainty in the results (during calibration and validation).

Other methods like the GDD could have been used as in the studies of d'Andrimont et al. (2020) and Gonzalez Piqueras et al. (2019), to acquire the dates of the desired crop growth stages more accurately. Even though, such methods could increase the reliability of the comparison results, they require additional data (e.g. temperature; De Bernardis et al. 2016b) and yet they also include uncertainties (Zeng et al. 2020) in addition to the fact that this kind of modelling was considered to be out of the scope of this thesis. Thus, and as suggested in the study of Xu et al. (2017), it is important to create a more accurate basis for crop phenology applications by focusing on recording the dates of different specific crop growth stages that are of high interest in agricultural management (e.g. 31DC, 39DC, 51DC). This can be achieved by adjusting the ground sampling schedule based on knowledge of the agricultural conditions within the study area (Xu et al. 2017), which could actually be very challenging (Nasrallah et al. 2019). Moreover, in-situ measurements always require the assumption that the individuals conducting the field observations are experienced and thus able to identify the crop

growth stages accurately (Xu et al. 2017). At the same time it is of high importance that the field observers follow consistent sampling rules (by taking the in-situ measurements within the field and not close to the edges of the parcel) which will ensure that the sampling location corresponds to a homogenous crop pixel.

5.3 Comparison of NDVI and EVI2 phenological metrics based on the threshold optimization

In general, NDVI showed higher values of optimal global thresholds compared to EVI2 for extracting the crop growth stages 31DC and 39DC (Figure 4.14) which is aligned with the findings of Huang et al. (2019), while for the stage 51DC the optimal global threshold was the same for both VIs. As Huang et al. (2019) note, the aforementioned differences in the optimal threshold values originate from the fact that NDVI captures changes in chlorophyll with higher sensitivity than EVI (where EVI2 is considered to have similar performance) which results in a higher season amplitude for NDVI compared to EVI2 (Figure 3.4, Figure 3.5). Moreover, the researchers note that sensors with high spatial resolution (i.e. S2) are capable of providing a homogenous time series signal (and not a mix of multiple land cover classes as happens with coarse resolution data; i.e. MODIS) which would ultimately increase the accuracy of the resulting NDVI phenological metrics. However, this is not depicted in the results of this study where EVI2 seems to outperform NDVI for all of the three crop growth stages of spring barley that are of interest in this study (31DC, 39DC, 51DC).

Overall, the RMSE for the different crop growth stages (31DC, 39DC, 51DC) is less than 7.6 days based on NDVI and less than 5.1 days for EVI2 (Figure 4.16). Nasrallah et al. (2019), defined the range of 6 days as a satisfactory offset, justified by the fact that crop growth stages will have a lasting period of a few days which will eventually reduce the actual shift of the S2 results. It is to question though if such an offset in days is viable for the timely application of external inputs (e.g. fungicides) in agricultural systems. Furthermore, the accuracies of NDVI and EVI2 varied for the different crop growth stages of interest of spring barley. EVI2 consistently resulted in higher R-square values and lower RMSE than NDVI (Figure 4.15 and Figure 4.16 respectively), which suggests that EVI2 performed better than NDVI regarding the detection of the crop growth stages 31DC, 39DC, 51DC of spring barley. This comes to an agreement with the hypothesis of Gao et al. (2017) where EVI2 is expected to detect crop growth stages that fall close to the peak of the season more accurately than NDVI. More specifically, the greatest difference in the accuracies of NDVI and EVI2 (R-square and RMSE;

Figure 4.15, Figure 4.16) was found for the crop growth stage 51DC (where the difference in RMSE was 3 days) followed by the crop growth stage 39DC, where both stages fall close to the peak of the season. The smallest difference in the accuracies of NDVI and EVI2 was found for the crop growth stage 31DC (where the difference in RMSE was 1.8 days). Such finding, could be possibly explained by the superiority of EVI2 over NDVI in capturing changes under dense vegetation (i.e. relationship with leaf area index; Jiang et al. 2008). Moreover, for both NDVI and EVI2, the offset of the estimated phenological dates for the different crop growth stages of interest (31DC, 39DC, 51DC) did not show any consistent pattern in terms of underestimation or overestimation when compared to the in-situ measurements (Figure 4.4, Figure 4.8, Figure 4.11).

Even though, it is challenging to make a comparison of phenological results derived from optical (i.e. S2) and radar data (i.e. S1; Stendardi et al. 2019) it is interesting to compare the results of this thesis with the study of Nasrallah et al. (2019) who used S1 data to monitor crop phenology. The researchers underline the fact that according to their results cereal crops that have reached the stage of germination early will reach the heading stage also early, while late germination leads to late heading. However, this is not true for the results of this thesis, where a spring barley parcel that was observed according to S2 to reach 39DC late would might reach 51DC earlier than a spring barley parcel that had reached 39DC earlier. Such a contradictory finding is strictly bounded to the achieved fit of the DL function, where the quality of the fit is known for having a great influence in the extracted phenometrics (Wei et al. 2016; Cai et al. 2017; Belda et al. 2020).

Moreover, as noted in the introduction section, and as the comparison results between NDVI and EVI2 suggest, different VIs are expected to give different phenological results (Jiang et al. 2008; Gao et al. 2017; Huang et al. 2019; Zeng et al. 2020). Additionally, as showed in the work of Sun et al. (2020), the red-edge bands of S2 are expected to improve the performance of VIs regarding crop monitoring, hence, it would be interesting to investigate in future work the use of NDVI and EVI2 or other VIs with the inclusion of other S2 spectral bands (e.g. red edge).

5.5 Near real-time crop growth stage detection

While phenology studies of precedented years is a widely explored domain (Gao et al. 2020), the concept of near real-time phenology monitoring, which refers to the detection of specific phenological events (e.g. crop growth stages) close to their occurrence (Zhang 2017), within the interval of ± 3 days (Liu et al. 2017; Liu et al. 2018) has drawn little attention (Zhang et al. 2012). Conversely, it is widely accepted that monitoring crop phenology in near real-time is beneficial for agronomic management (De Bernardis et al. 2016a; Liu et al. 2017; Liu et al. 2018; Defourny et al. 2019; Gao et al. 2020). Therefore, it is important to note that, the methodology deployed in this study is not capable of providing any information for the ongoing growing season in a timely fashion (i.e. in near real-time). As noted, in various phenology applications, workflows similar to the one utilized this study enable phenological studies for preceding years only since data from the entire growing season are required to fit a function (De Bernardis et al. 2016b; Gao et al. 2017; Liu et al. 2017; Gao et al. 2020; Zeng et al. 2020). According to Zhang (2017), near real-time monitoring of vegetation phenology using RS data is greatly hindered by the complications of noise filtering (e.g. atmospheric contaminations) when dealing with incomplete time series.

Nevertheless, great efforts have been directed towards conducting crop phenology monitoring in near real-time using different satellites (e.g. MODIS, Landsat, S2) with the combination of data from multiple sensors (i.e. data fusion) being extremely beneficial (Gao et al. 2017; Liu et al. 2018; Gao et al. 2020). More specifically, Liu et al. (2018) demonstrated the simulation of artificially generated VI (e.g. EVI) trajectories representing the increasing VI phase of the crop growth. Their method for detecting crop growth phases in near real-time is based on the work of Liu et al. (2017) and uses climatological phenology data (potential crop growth derived from historical MODIS data) and timely available satellite acquisitions (i.e. Visible Infrared Imaging Radiometer Suite; VIIRS). In a similar fashion, Olsson et al. (2016) utilized TIMESAT along with historical and timely available MODIS NDVI time series data to monitor in near real-time forest defoliation. For that, they used the DL function to reconstruct time series of two hundred MODIS pixels in order to create an average seasonal NDVI trajectory that represents a normal growing season. Furthermore, they explained that such methods might be also applicable on S2 data.

In that way, it could be possible to use historical MODIS data in combination with S2 data to create a VI trajectory (Olsson et al. 2016) that represents a typical growth cycle of barley. Additionally, the potential VI trajectory of spring barley could also be possibly derived by exploiting the rich spatial domain of S2. Meaning that instead of using the time domain (historical data) as in the case of sensors with coarse spatial resolution (e.g. MODIS), for S2 the larger amount of pixels might be able to provide the potential VI trajectory of spring barley which can be ultimately combined with timely available S2 acquisitions for detecting crop growth stages in near real-time. Based on the above, it is of high interest to investigate the capabilities of S2 in near real-time crop growth stage detection by deploying more sophisticated approaches along with its data fusion possibilities.

6. Conclusion

This study demonstrated the use of S2 time series data (NDVI, EVI2) for deriving the phenological cycle of spring barley and ultimately for detecting three specific crop growth stages that are of high interest in agronomic management, named according to the Zadoks scale as: (i) first node detectable (31DC), (ii) flag leaf ligule just visible (39DC) and first spikelet of inflorescence just visible (51DC). To achieve this aim, in-situ measurements were used to calibrate and validate the results of the (i) optimal local and (ii) global thresholds for the different crop growth stages of interest.

According to the results of the global threshold optimization, the induced RMSE and R-square values based on NDVI and EVI2 showed a varying agreement between the in-situ observations and derived phenometrics for the different crop growth stages of interest (31DC, 39DC, 51DC). More specifically the resulted RMSE when applying the optimal global threshold for NDVI was in the range 6.9 days – 7.2 days, where for EVI2 the induced RMSE ranged between 4.2 days – 5.1 days. It is challenging though to decide whether the resulted RMSE when applying the optimal global threshold for both VIs is in an acceptable range for applying fungicides (or other external inputs).

In general, the variability of the optimal local thresholds for both NDVI and EVI2 showed a decreasing trend with later crop growth stages. In detail, for the stages 39DC and 51DC NDVI showed higher SD (10.4%, 5.9% respectively) compared to EVI2 (7.1%; 39DC, 3.1%; 51DC). Conversely, for the stage 31DC EVI2 resulted in higher SD (21.1%) than NDVI (16.1%).

The overall results showed a varying agreement between the in-situ observations and the derived phenometrics based on NDVI and EVI2 for the different crop growth stages of interest (31DC, 39DC, 51DC) when applying the optimal global thresholds. Though, the extracted phenological dates associated with the optimal global thresholds based on the RMSE and R-square values seem to be moderately accurate when compared to the in-situ observations.

However, the uncertainty related to the in-situ measurements in conjunction with the small sample size of spring barley parcels that were used in the analysis does not allow any confident conclusions. Therefore, an attempt has to be made in order to record the exact dates of the crop growth stages of interest in the field protocol, along with the

inclusion of more parcels in such analysis in order to increase the reliability of the results.

In general, EVI2 consistently performed better than NDVI in terms of R-square and RMSE for all the different studied crop growth stages (31DC, 39DC, 51DC) when applying the optimal global threshold. The same applies in terms of variability of the optimal local threshold, except for the stage 31DC, where NDVI was better than EVI2. Thus, the results of this study suggest that EVI2 was generally more accurate in detecting the three crop growth stages of interest.

References

ADAMA n.d. Foliar Fungicide Guide. [https://www.adama.com/documents/1380147/1395527/Foliar+Fungicide+Guide+-+Cereals.pdf] Visited: 10th April 2020

Agriculture and Horticulture Development Board. 2020. Fungicide programmes for wheat and barley. [https://ahdb.org.uk/fungicide-programmes] Visited: 9th March 2020

Ahmed, M., and C. O. Stockle. 2017. *Quantification of Climate Variability, Adaptation and Mitigation for Agricultural Sustainability*. *Quantification of Climate Variability, Adaptation and Mitigation for Agricultural Sustainability*. Springer International Publishing. doi:10.1007/978-3-319-32059-5.

Anderson, P. M., E. A. Oelke, and S. R. Simmons. 1995. *Growth and Development Guide for Spring Barley*.

Araya, S. 2017. Package ‘ CropPhenology ’: 1–7.

Araya, S., B. Ostendorf, G. Lyle, and M. Lewis. 2018. CropPhenology: An R package for extracting crop phenology from time series remotely sensed vegetation index imagery. *Ecological Informatics* 46. Elsevier: 45–56. doi:10.1016/j.ecoinf.2018.05.006.

Atkinson, P. M., C. Jeganathan, J. Dash, and C. Atzberger. 2012. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. *Remote Sensing of Environment* 123. Elsevier Inc.: 400–417. doi:10.1016/j.rse.2012.04.001.

Baik, B. K., and S. E. Ullrich. 2008. Barley for food: Characteristics, improvement, and renewed interest. *Journal of Cereal Science* 48: 233–242. doi:10.1016/j.jcs.2008.02.002.

Beck, P. S. A., C. Atzberger, K. A. Høgda, B. Johansen, and A. K. Skidmore. 2006. Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sensing of Environment* 100: 321–334. doi:10.1016/j.rse.2005.10.021.

Belda, S., L. Pipia, P. Morcillo-Pallarés, and J. Verrelst. 2020. Optimizing Gaussian Process Regression for Image Time Series Gap-Filling and Crop Monitoring. *Agronomy* 10: 618. doi:10.3390/agronomy10050618.

Cai, Z. 2019. *Vegetation Observation in the Big Data Era : Sentinel-2 data for mapping the seasonality of land vegetation*.

Cai, Z., P. Jönsson, H. Jin, and L. Eklundh. 2017. Performance of smoothing methods for reconstructing NDVI time-series and estimating vegetation phenology from MODIS data. *Remote Sensing* 9: 20–22. doi:10.3390/rs9121271.

Chuvieco, E. 2016. *Fundamentals of Satellite Remote Sensing*. *Second Edition*. CRC Press. doi:10.1201/b18954.

d'Andrimont, R., M. Taymans, G. Lemoine, A. Ceglar, M. Yordanov, and M. van der Velde. 2020. Detecting flowering phenology in oil seed rape parcels with Sentinel-1 and -2 time series. *Remote Sensing of Environment* 239. Elsevier: 111660. doi:10.1016/j.rse.2020.111660.

De Bernardis, C., F. Vicente-Guijalba, T. Martinez-Marin, and J. M. Lopez-Sanchez. 2016a. Contribution to Real-Time Estimation of Crop Phenological States in a Dynamical Framework Based on NDVI Time Series: Data Fusion with SAR and Temperature. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9. IEEE: 3512–3523. doi:10.1109/JSTARS.2016.2539498.

De Bernardis, C., F. Vicente-Guijalba, T. Martinez-Marin, and J. M. Lopez-Sanchez. 2016b. Particle filter approach for real-time estimation of crop phenological states using time series of NDVI images. *Remote Sensing* 8: 5–10. doi:10.3390/rs8070610.

Defourny, P., S. Bontemps, N. Bellemans, C. Cara, G. Dedieu, E. Guzzonato, O. Hagolle, J. Inglada, et al. 2019. Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. *Remote Sensing of Environment* 221. Elsevier: 551–568. doi:10.1016/j.rse.2018.11.007.

Delécolle, R., S. J. Maas, M. Guérif, and F. Baret. 1992. Remote sensing and crop production models: present trends. *ISPRS Journal of Photogrammetry and Remote Sensing* 47: 145–161. doi:10.1016/0924-2716(92)90030-D.

Eklundh, L., and P. Jönsson. 2016. Timesat for processing time-series data from satellite sensors for land surface monitoring. In *Remote Sensing and Digital Image Processing*. doi:10.1007/978-3-319-47037-5_9.

Eklundh, L., and P. Jönsson. 2017. TIMESAT 3.3 with seasonal trend decomposition and parallel processing Software Manual. *Lund and Malmo University, Sweden*: 1–92.

Eklundh, L., M. Sjöström, J. Ardö, and P. Jönsson. 2012. High resolution mapping of vegetation dynamics from Sentinel-2. *European Space Agency, (Special Publication) ESA SP* 707 SP.

European Space Agency. 2015. SENTINEL-2 User Handbook. *Sentinel-2 User Handbook.*

Gao, F., M. Anderson, C. Daughtry, A. Karnieli, D. Hively, and W. Kustas. 2020. A within-season approach for detecting early growth stages in corn and soybean using high temporal and spatial resolution imagery. *Remote Sensing of Environment* 242. Elsevier: 111752. doi:10.1016/j.rse.2020.111752.

Gao, F., M. C. Anderson, X. Zhang, Z. Yang, J. G. Alfieri, W. P. Kustas, R. Mueller, D. M. Johnson, et al. 2017. Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. *Remote Sensing of Environment* 188: 9–25. doi:10.1016/j.rse.2016.11.004.

Gomarasca, M. A., A. Tornato, D. Spizzichino, E. Valentini, A. Taramelli, G. Satalino, M. Vincini, M. Boschetti, et al. 2019. Sentinel for applications in agriculture. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* 42: 91–98. doi:10.5194/isprs-archives-XLII-3-W6-91-2019.

Grains Research & Devellopment Corporation. 2013. Cereal Fungicides Fact Sheet: 1–4.

Hird, J. N., and G. J. McDermid. 2009. Noise reduction of NDVI time series: An empirical comparison of selected techniques. *Remote Sensing of Environment* 113. Elsevier Inc.: 248–258. doi:10.1016/j.rse.2008.09.003.

Holben, B. N. 1986. Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing* 7: 1417–1434. doi:10.1080/01431168608948945.

Huang, X., J. Liu, W. Zhu, C. Atzberger, and Q. Liu. 2019. The optimal threshold and vegetation index time series for retrieving crop phenology based on a modified dynamic threshold method. *Remote Sensing* 11. doi:10.3390/rs11232725.

Huete, A. R. 1988. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*. doi:10.1016/0034-4257(88)90106-X.

Huete, A., K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83: 195–213. doi:10.1016/S0034-4257(02)00096-2.

Huete, A. R. 2004. *Remote Sensing For Environmental Monitoring*. *Environmental Monitoring and Characterization*. Elsevier, Inc. doi:10.1016/B978-0-12-064477-3.50013-8.

Jenkyn, J. F. 1974. Effects of mildew on the growth and yield of spring barley: 1969-72. *Annals of Applied Biology* 78: 281–288. doi:10.1111/j.1744-7348.1974.tb01507.x.

Jiang, Z., A. R. Huete, K. Didan, and T. Miura. 2008. Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment* 112: 3833–3845. doi:10.1016/j.rse.2008.06.006.

Jin, H., and L. Eklundh. 2014. A physically based vegetation index for improved monitoring of plant phenology. *Remote Sensing of Environment*. doi:10.1016/j.rse.2014.07.010.

Jones, H., and R. Vaughan. 2010. Remote sensing of vegetation, principles, techniques, and applications. *Oxford: Oxford University Press*.

Jönsson, P., and L. Eklundh. 2002. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*. doi:10.1109/TGRS.2002.802519.

Jönsson, P., and L. Eklundh. 2004. TIMESAT - A program for analyzing time-series of satellite sensor data. *Computers and Geosciences* 30: 833–845. doi:10.1016/j.cageo.2004.05.006.

Lange, M., B. Dechant, C. Rebmann, M. Vohland, M. Cuntz, and D. Doktor. 2017. Validating MODIS and sentinel-2 NDVI products at a temperate deciduous forest site using two independent ground-based sensors. *Sensors (Switzerland)* 17. doi:10.3390/s17081855.

Liu, J., and P. Zhan. 2016. The impacts of smoothing methods for time-series remote sensing data on crop phenology extraction. *International Geoscience and Remote Sensing Symposium (IGARSS)* 2016-Novem: 2296–2299. doi:10.1109/IGARSS.2016.7729593.

Liu, J., and X. Huang. 2019. Evaluating crop phenology retrieving accuracies based on ground observations. *2019 8th International Conference on Agro-Geoinformatics, Agro-Geoinformatics 2019*. IEEE: 1–5. doi:10.1109/Agro-Geoinformatics.2019.8820703.

Liu, L., X. Zhang, Y. Yu, and W. Guo. 2017. Real-time and short-term predictions of spring phenology in North America from VIIRS data. *Remote Sensing of Environment* 194. Elsevier Inc.: 89–99. doi:10.1016/j.rse.2017.03.009.

Liu, L., X. Zhang, Y. Yu, F. Gao, and Z. Yang. 2018. Real-time monitoring of crop phenology in the Midwestern United States using VIIRS observations. *Remote Sensing* 10: 1–18. doi:10.3390/rs10101540.

Luo, Y., Z. Zhang, Y. Chen, Z. Li, and F. Tao. 2020. ChinaCropPhen1km: a high-resolution crop phenological dataset for three staple crops in China during 2000-2015 based on leaf area index (LAI) products. *Earth System Science Data* 12: 197–214. doi:10.5194/essd-12-197-2020.

McLean, M. 2012. Scald of Barley. Department of Environment and Primary Industries. [http://agriculture.vic.gov.au/agriculture/pests-diseases-and-weeds/plant-diseases/grains-pulses-and-cereals/scald-of-barley](Visited 20th April 2020)

Mondal, P., and V. K. Tewari. 2007. Present status of precision farming: A review. *International Journal of Agricultural Research*. doi:10.3923/ijar.2007.1.10.

Mulla, D. J. 2013. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering* 114. IAgrE: 358–371. doi:10.1016/j.biosystemseng.2012.08.009.

Nakajima, T. 2010. Fungicides Application against Fusarium Head Blight in Wheat and Barley for Ensuring Food Safety. In *Fungicides*. doi:10.5772/13680.

Nasrallah, A., N. Baghdadi, M. Mhawej, G. Faour, T. Darwish, H. Belhouchette, and S. Darwich. 2018. A novel approach for mapping wheat areas using high resolution sentinel-2 images. *Sensors (Switzerland)* 18: 1–23. doi:10.3390/s18072089.

Nasrallah, A., N. Baghdadi, M. El Hajj, T. Darwish, H. Belhouchette, G. Faour, S. Darwich, and M. Mhawej. 2019. Sentinel-1 data for winter wheat phenology monitoring and mapping. *Remote Sensing* 11. doi:10.3390/rs11192228.

Newlands, N. K. 2018. Model-Based Forecasting of Agricultural Crop Disease Risk at the Regional Scale, Integrating Airborne Inoculum, Environmental, and Satellite-Based Monitoring Data. *Frontiers in Environmental Science* 6: 1–16. doi:10.3389/fenvs.2018.00063.

Olsson, P. O., J. Lindström, and L. Eklundh. 2016. Near real-time monitoring of insect induced defoliation in subalpine birch forests with MODIS derived NDVI. *Remote Sensing of Environment* 181. Elsevier Inc.: 42–53. doi:10.1016/j.rse.2016.03.040.

Palchowdhuri, Y., R. Valcarce-Diñeiro, P. King, and M. Sanabria-Soto. 2018. Classification of multi-temporal spectral indices for crop type mapping: A case study in Coalville, UK. *Journal of Agricultural Science* 156: 24–36. doi:10.1017/S0021859617000879.

Pan, Z., J. Huang, Q. Zhou, L. Wang, and Y. Cheng. 2015. Mapping crop phenology using NDVI time-series derived from HJ-1 A / B data. *International Journal of Applied Earth Observations and Geoinformation* 34. Elsevier B.V.: 188–197. doi:10.1016/j.jag.2014.08.011.

Poole, N. 2005. *Cereal Growth Stages Guide*.

Reed, B. C., M. D. Schwartz, and X. Xiao. 2009. Remote sensing phenology status and the way forward. In *Phenology of Ecosystem Processes: Applications in Global Change Research*. doi:10.1007/978-1-4419-0026-5_10.

Rodrigues, A., A. R. S. Marcal, and M. Cunha. 2013. Monitoring vegetation dynamics inferred by satellite data using the pheno sat tool. *IEEE Transactions on Geoscience and Remote Sensing*. doi:10.1109/TGRS.2012.2223475.

Sakamoto, T. 2018. Refined shape model fitting methods for detecting various types of phenological information on major U.S. crops. *ISPRS Journal of Photogrammetry and Remote Sensing* 138. International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS): 176–192. doi:10.1016/j.isprsjprs.2018.02.011.

Sakamoto, T., M. Yokozawa, H. Toritani, M. Shibayama, N. Ishitsuka, and H. Ohno. 2005. A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment* 96: 366–374. doi:10.1016/j.rse.2005.03.008.

SCB. 2019. Use of agricultural land 2019 Final statistics.

Solano-Correa, Y. T., F. Bovolo, L. Bruzzone, and D. Fernandez-Prieto. 2017. Spatio-temporal evolution of crop fields in Sentinel-2 Satellite Image Time Series. *2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images, MultiTemp 2017*: 2–6. doi:10.1109/Multi-Temp.2017.8035236.

Stendardi, L., S. R. Karlsen, G. Niedrist, R. Gerdol, M. Zebisch, M. Rossi, and C. Notarnicola. 2019. Exploiting time series of Sentinel-1 and Sentinel-2 imagery to detect meadow phenology in mountain regions. *Remote Sensing* 11. doi:10.3390/rs11050542.

Sudmanns, M., D. Tiede, H. Augustin, and S. Lang. 2019. Assessing global Sentinel-2 coverage dynamics and data availability for operational Earth observation (EO) applications using the EO-Compass. International Journal of Digital Earth 0. Taylor & Francis: 1–17. doi:10.1080/17538947.2019.1572799.

Sun, Y., Q. Qin, H. Ren, T. Zhang, and S. Chen. 2020. Red-Edge Band Vegetation Indices for Leaf Area Index Estimation from Sentinel-2/MSI Imagery. *IEEE Transactions on Geoscience and Remote Sensing* 58: 826–840. doi:10.1109/TGRS.2019.2940826.

Swedish University of Agricultural Sciences. 1996. SWEDEN: COUNTRY REPORT TO THE FAO INTERNATIONAL TECHNICAL CONFERENCE ON PLANT GENETIC RESOURCES. *Food and Agriculture Organisation*: 1–25. doi:10.1039/C5LC00436E.

Thorne, G. N. 1966. Physiological aspects of grain yield in cereals. In The Growth of Cereals & Grasses (ed. F. L. Milthorpe & J. D. Ivins), pp. 88-105. London: Butterworths.

Tottman, D. R., R. J. Makepeace, and H. Broad. 1979. An explanation of the decimal code for the growth stages of cereals, with illustrations. *Annals of Applied Biology* 93: 221–234. doi:10.1111/j.1744-7348.1979.tb06534.x.

Viña, A., A. A. Gitelson, D. C. Rundquist, G. Keydan, B. Leavitt, and J. Schepers. 2004. Monitoring maize (Zea mays L.) phenology with remote sensing. *Agronomy Journal* 96: 1139–1147. doi:10.2134/agronj2004.1139.

Viovy, N., O. Arino, and A. S. Belward. 1992. The best index slope extraction (Bise): A method for reducing noise in NDVI time-series. *International Journal of Remote Sensing* 13: 1585–1590. doi:10.1080/01431169208904212.

Vrieling, A., M. Meroni, R. Darvishzadeh, A. K. Skidmore, T. Wang, R. Zurita-Milla, K. Oosterbeek, B. O'Connor, et al. 2018. Vegetation phenology from Sentinel-2 and field cameras for a Dutch barrier island. *Remote Sensing of Environment* 215. Elsevier: 517–529. doi:10.1016/j.rse.2018.03.014.

Wang, E., and T. Engel. 1998. Simulation of phenological development of wheat crops. *Agricultural Systems* 58: 1–24. doi:10.1016/S0308-521X(98)00028-6.

Wei, W., W. Wu, Z. Li, P. Yang, and Q. Zhou. 2016. Selecting the optimal NDVI time-series reconstruction technique for crop phenology detection. *Intelligent Automation and Soft Computing* 22: 237–247. doi:10.1080/10798587.2015.1095482.

Whelan, B. M., and A. B. McBratney. 2000. The "null hypothesis" of precision agriculture management. In *Precision Agriculture*. doi:10.1023/A:1011838806489.

White, M. A., P. E. Thornton, and S. W. Running. 1997. A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles* 11: 217–234. doi:10.1029/97GB00330.

Xin, J., Z. Yu, L. van Leeuwen, and P. M. Driessen. 2002. Mapping crop key phenological stages in the North China Plain using NOAA time series images. *International Journal of Applied Earth Observation and Geoinformation* 4: 109–117. doi:10.1016/S0303-2434(02)00007-7.

Xu, X., C. Conrad, and D. Doktor. 2017. Optimising phenological metrics extraction for different crop types in Germany using the moderate resolution imaging Spectrometer (MODIS). *Remote Sensing* 9. doi:10.3390/rs9030254.

Yang, C. 2020. Remote sensing and precision agriculture technologies for crop disease detection and management with a practical application example. *Engineering*. THE AUTHOR: 0–4. doi:10.1016/j.eng.2019.10.015.

You, X., J. Meng, M. Zhang, and T. Dong. 2013. Remote sensing based detection of crop phenology for agricultural zones in China using a new threshold method. *Remote Sensing* 5: 3190–3211. doi:10.3390/rs5073190.

Yousfi, S., J. Fernando Marin Peira, G. Rincón De La Horra, and P. V. Mauri Ablanque. 2019. Remote Sensing: Useful Approach for Crop Nitrogen Management and Sustainable Agriculture. *IntechOpen*. InTech. doi:10.5772/intechopen.89422.

Zadoks, J. C., T. T. Chang, and C. F. Konzak. 1974. A decimal code for the growth stages of cereals. *Weed Research* 14: 415–421. doi:10.1111/j.1365-3180.1974.tb01084.x.

Zeng, L., B. D. Wardlow, D. Xiang, S. Hu, and D. Li. 2020. A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sensing of Environment* 237. Elsevier: 111511. doi:10.1016/j.rse.2019.111511.

Zhang, X. 2017. *Land surface phenology: Climate data record and real-time monitoring*. *Comprehensive Remote Sensing*. Vol. 1–9. Elsevier. doi:10.1016/B978-0-12-409548-9.10351-3.

Zhang, X., J. Wang, G. M. Henebry, and F. Gao. 2020. Development and evaluation of a new algorithm for detecting 30 m land surface phenology from VIIRS and HLS time series. *ISPRS Journal of Photogrammetry and Remote Sensing* 161. Elsevier: 37–51. doi:10.1016/j.isprsjprs.2020.01.012.

Zheng, Y., B. Wu, M. Zhang, and H. Zeng. 2016. Crop Phenology Detection Using High Spatio-Temporal Resolution Data Fused from SPOT5 and MODIS Products. *Sensors* 16: 2099. doi:10.3390/s16122099.

Appendix

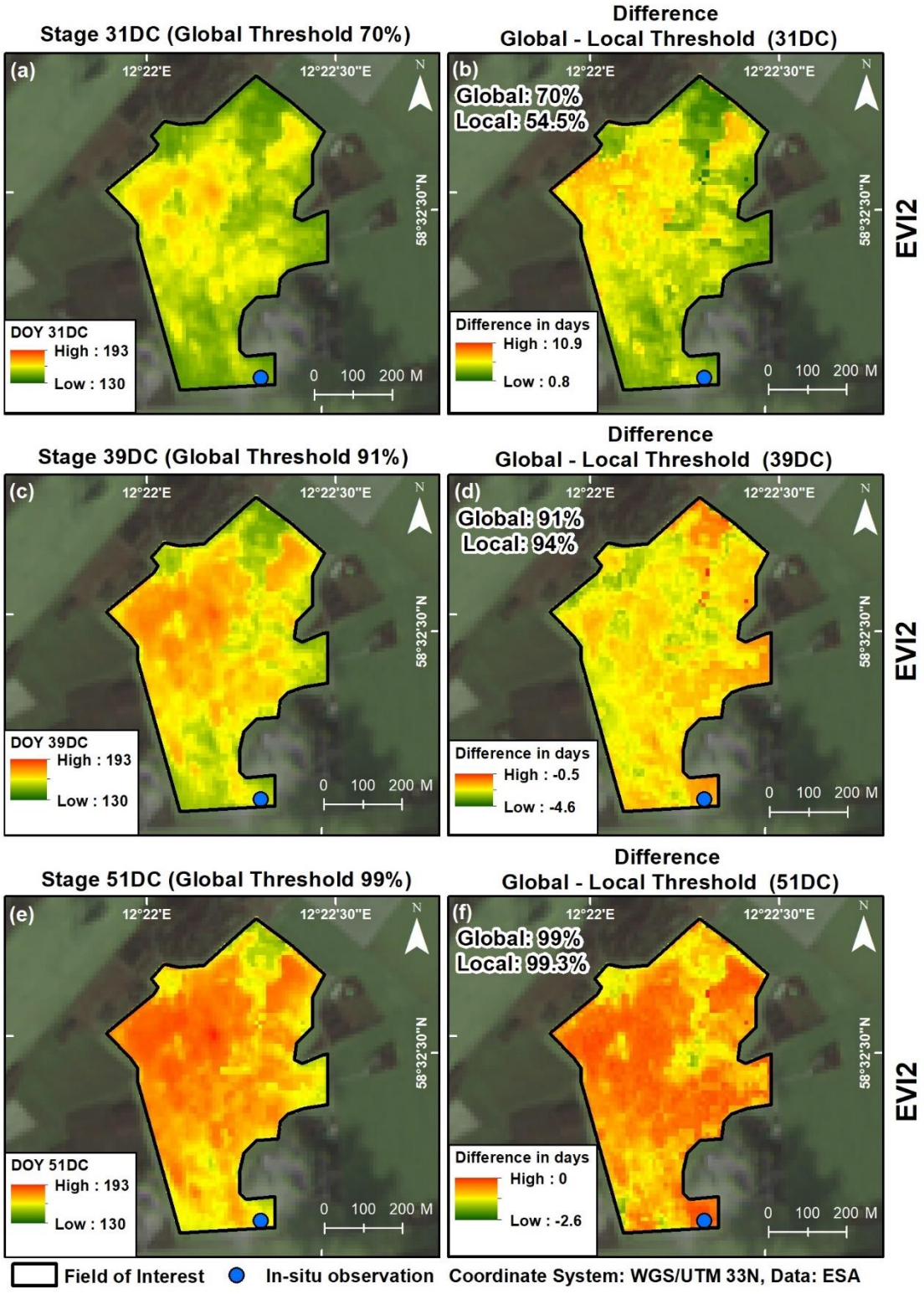

Figure S1 Example of a spring barley field (field number 60 in 2019) map showing the derived DOY from EVI2 for the crop stage 31DC when applying the global threshold (70%; a), and difference map in number of days based on the global minus local thresholds (b). DOY from EVI2 for the crop stage 39DC when applying the global threshold (91%; c), and difference map in number of days based on the global minus local thresholds (d). DOY from EVI2 for the crop growth stage 51DC when applying the global threshold (99%; e), and difference map in number of days based on the global minus local thresholds (f).

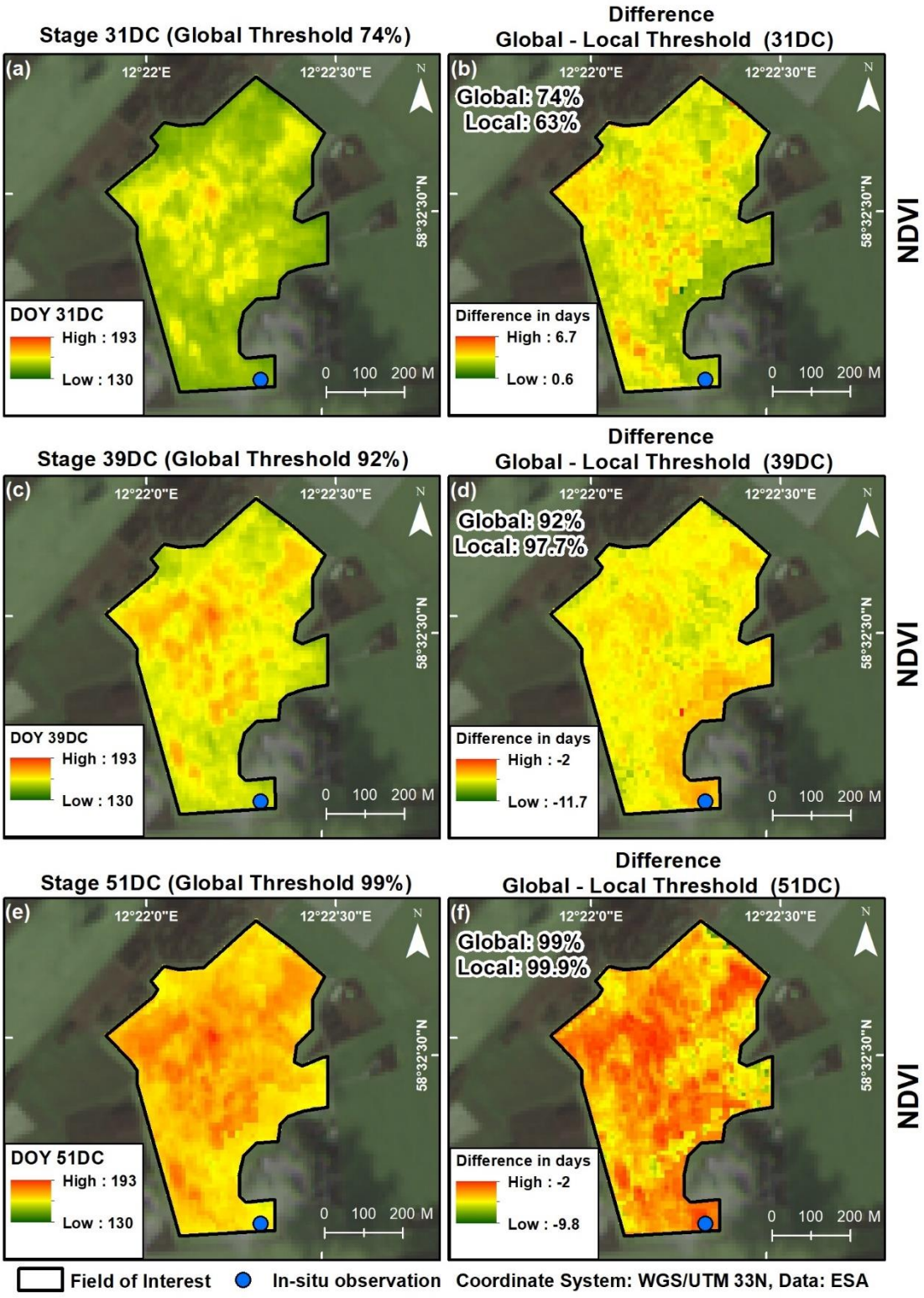

Figure S2 Example of a spring barley field (field number 60 in 2019) map showing the derived DOY from NDVI for the crop stage 31DC when applying the global threshold (74%; a), and difference map in number of days based on the global minus local thresholds (b). DOY from NDVI for the crop stage 39DC when applying the global threshold of (92%; c), and difference map in number of days based on the global minus local thresholds (d). DOY from NDVI for the crop growth stage 51DC when applying the global threshold (99%; e), and difference map in number of days based on the global minus local thresholds (f).

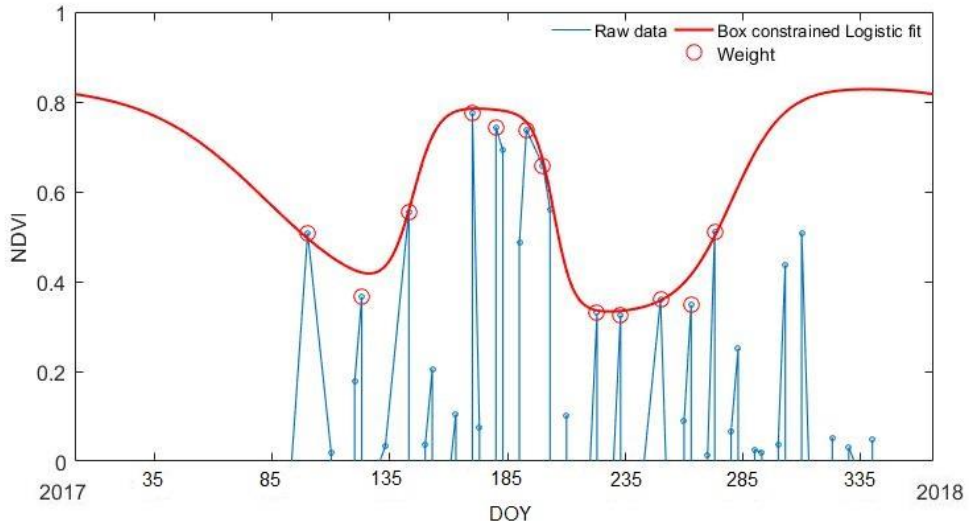

Figure S3 NDVI time series of a pixel that represents the location of the in-situ observations for a spring barley parcel (field number 61 in 2017) showing the effect of a mixed pixel. The NDVI time series of the new location is shown in Figure 3.3.

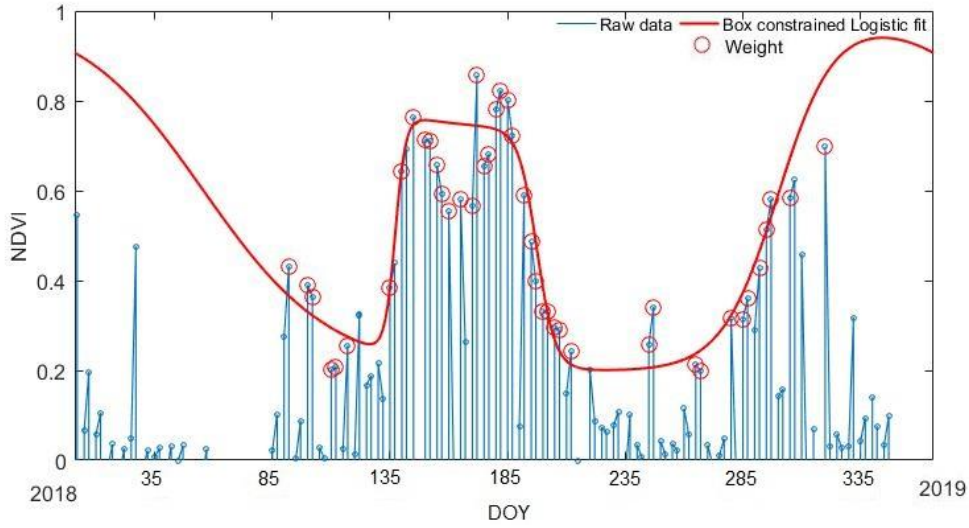

Figure S4 Example of NDVI time series of a pixel that represents the location of the in-situ measurements for one spring barley parcel (field number 91 in 2018). This pixel did not allow the reconstruction of a smooth VI curve. Quality observations that were used for the DL fit (red line) are marked in red circles.

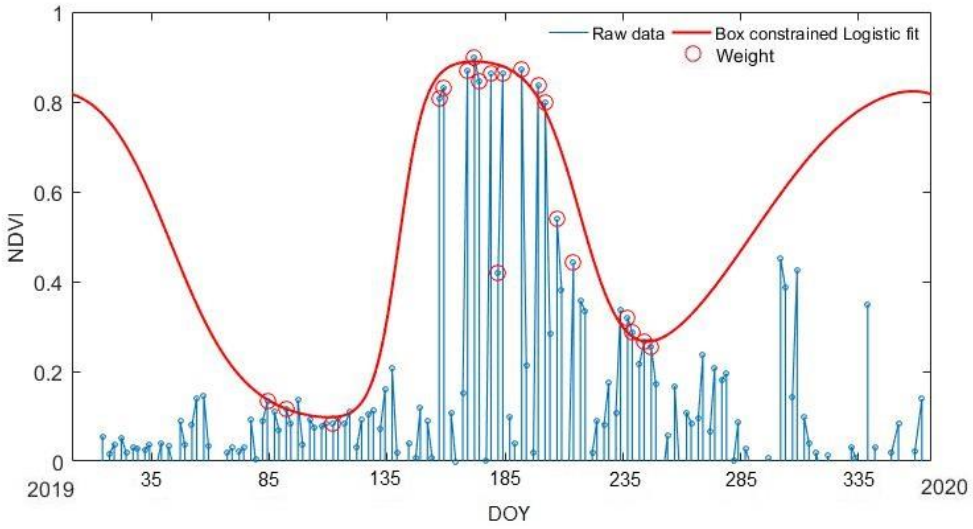

Figure S5 Example of NDVI time series of a pixel that represents the location of the in-situ measurements for one spring barley parcel (field number 217 in 2019). This pixel did not allow the reconstruction of a smooth VI curve. Quality observations that were used for the DL fit (red line) are marked in red circles.

Table S1 Zadoks scale for crop growth stages of cereals, with corresponding DC - decimal code (modified from Zadoks et al. 1974).

DC	Description	DC	Description	DC	Description	DC	Description
0	**Germination**	**3**	**Stem elongation**	**6**	**Anthesis**	**9**	**Ripening**
00	Dry seed	30	Pseudostem (leaf sheath) erection	60	—	90	—
01	Start of imbibition (water absorption)	31	First node detectable	61	Beginning of anthesis	91	Caryopsis hard (difficult to divide)
02	—	32	2nd node detectable	62	—	92	Caryopsis hard (not dented by thumbnail)
03	Imbibition complete	33	3rd node detectable	63	—	93	Caryopsis loosening in daytime
04	—	34	4th node detectable	64	—	94	Over-ripe, straw dead and collapsing
05	Radicle (root) emerged from caryopsis (seed)	35	5th node detectable	65	Anthesis half-way	95	Seed dormant
06	—	36	6th node detectable	66	—	96	Viable seed giving 50% germination
07	Coleoptile (shoot) emerged from caryopsis	37	Flag leafjust visible	67	—	97	Seed not dormant
08	—	38	—	68	—	98	Secondary dormancy induced
09	Leaf just at coleoptile tip	39	Flag leaf ligule just visible	69	Anthesis complete	99	Secondary dormancy lost
1	**Seedling Growth**	**4**	**Booting Process**	**7**	**Milk Development**		
10	First leaf through coleoptile	40	—	70	—		
11	First leaf unfolded	41	Flag leaf sheath extending	71	Caryopsis (kernel) water ripe		
12	2 leaves unfolded	42	—	72	—		
13	3 leaves unfolded	43	Boots just visibly swollen	73	Early milk		
14	4 leaves unfolded	44	—	74	—		
15	5 leaves unfolded	45	Boots swollen	75	Medium milk		
16	6 leaves unfolded	46	—	76	—		
17	7leaves unfolded	47	Flag leaf sheath opening	77	Late milk		
18	8 leaves unfolded	48	—	78	—		
19	9 or more leaves unfolded	49	First awns visible	79	—		
2	**Tillering**	**5**	**Infloresence emergence**	**8**	**Dough Development**		
20	Main shoot only	50	—	80	—		
21	Main shoot and 1 tiller	51	First spikelet of inflorescence just visible	81	—		
22	Main shoot and 2 tillers	52	—	82	—		
23	Main shoot and 3 tillers	53	1/4 of inflorescence emerged	83	Early dough		
24	Main shoot and 4 tillers	54	—	84	—		
25	Main shoot and 5 tillers	55	1/2 of inflorescence emerged	85	Softdough		
26	Main shoot and 6 tillers	56	—	86	—		
27	Main shoot and 7 tillers	57	3/4 of inflorescence emerged	87	Hard dough		
28	Main shoot and 8 tillers	58	—	88	—		
29	Main shoot and 9 or more tillers	59	Emergence of inflorescence completed	89	—		

References

Zadoks, J. C., T. T. Chang, and C. F. Konzak. 1974. A decimal code for the growth stages of cereals. *Weed Research* 14: 415–421. doi:10.1111/j.1365-3180.1974.tb01084.x.