Article

A Crop Group-Specific Pure Pixel Time Series for Europe

Christof J. Weissteiner [1,](ORCID), Raúl López-Lozano [2], Giacinto Manfron [1] (ORCID), Grégory Duveiller [1] (ORCID), Josh Hooker [1], Marijn van der Velde [1] (ORCID) and Bettina Baruth [1] (ORCID)

[1] European Commission, Joint Research Centre (JRC), 21027 Ispra, Italy; giacinto.manfron@ec.europa.eu (G.M.); Gregory.Duveiller@ec.europa.eu (G.D.); josh.hooker@ec.europa.eu (J.H.); marijn.van-der-velde@ec.europa.eu (M.v.d.V.); bettina.baruth@ec.europa.eu (B.B.)

[2] French National Institute for Agricultural Research (INRA), UMR EMMAH, CEDEX 9, 84914 Avignon, France; raul.lopez-lozano@inra.fr

[*] Correspondence: Christof.Weissteiner@ec.europa.eu or mail@weissteiner.eu

check for updates

Abstract: Long timeseries of Earth observation data for the characterization of agricultural crops across large scales are of high interest to crop modelers, scientists, and decision makers in the fields of agricultural and environmental policy as well as crop monitoring and food security. They are particularly important for regression-based crop monitoring systems that rely on historic information. The major challenge lies in identifying pixels from satellite imagery that represent pure enough crop signals. Here, we present a data-driven semi-automatic approach to identify pure pixels of two crop groups (i.e., winter and spring crops and summer crops) based on a MODIS–NDVI timeseries. We applied this method to the European Union at a 250 m spatial resolution. Pre-processed and smoothed, daily normalized difference vegetation index (NDVI) data (2001–2017) were used to first extract the phenological data. To account for regional characteristics (varying climate, agro-management, etc.), these data were clustered by administrative units and by year using a Gaussian mixture model. The number of clusters was pre-defined using data from regional agricultural acreage statistics. After automatic labelling, clusters were filtered based on agronomic knowledge and phenological information extracted from the same timeseries. The resulting pure pixels were validated with two different datasets, one based on high-resolution Sentinel-2 data (5 sites, 2 years) and one based on a regional crop map (1 site, 7 years). For the winter and spring crop class, pixel purity amounted to 93% using the first validation dataset and to 73% using the second one, averaged over the different years. For summer crops, the respective values were 61% (91% without one critical validation site) and 72%. The phenological analyses revealed a clear trend towards an earlier NDVI peak (approximately −0.28 days/year) for winter and spring crops across Europe. We expect that this dataset will be useful for various applications, from crop model calibration to operational crop monitoring and yield forecasting.

Keywords: pixel purity; crop classification; crop monitoring; Gaussian mixture model; phenology; remote sensing; timeseries

1. Introduction

Information on crop condition and yield at continental extents is of high importance to decision makers in the fields of agricultural and environmental policy as well as food security [1]. Such information is of even higher value if provided at regional (usually sub-national) scale. Region-wise, yield forecasts can lead to higher accuracies because the analyses can be based on more locally

appropriate predictors. For regional remote sensing-based crop monitoring, frequently based on regression analysis, the availability of crop-specific data (crop masks) assumes a major role [2]. Nonetheless, remote sensing-dependent crop models are often run over large geographic extents with only an approximate knowledge of crop distribution. This is due to the limitations in the spatial and temporal resolution of the remotely sensed and ground data [3], adding uncertainty and lowering model accuracy. Such crop model simulations could be considerably improved if fed with adequately geo-located and pure crop specific land cover data for every growing season [4,5].

In view of the increasing and increasingly sophisticated use of Earth observation (EO) data as model input, the correct geo-localization of crops becomes increasingly important. Large-scale crop monitoring is usually run by (inter)governmental organizations or NGOs. Eight of the most prominent of such organizations were interviewed by researchers [6] on how they viewed their data and model inputs. Six out of eight organizations identified the data gaps for their crop-type maps as "critically important" or "critical". Frequently, surrogates for such crop type maps are derived from downscaled statistics of larger administrative areas. The geographic distribution of crops in those products is often spatially re-allocated based on proxies. Four products created by such downscaling methodologies were compared by researchers [7]: They found large discrepancies among them which suggests the generally low robustness of such approaches. The eight surveyed crop monitoring organizations [6] viewed crop calendars as crop model input even more critically than crop masks. Local planting and harvesting dates for different crop types are known to exert important influence on crop monitoring and yield estimates. The need for reliable large-scale crop-type maps and phenological data calls for respective research in this direction.

Remote sensing is widely used to try to fill these data gaps [8–15]. In much of this research, crops are identified at only regional extents and for only one or a few years. Here, we aimed to employ remote sensing on a continental level and on a thoroughly multi-annual basis. Our analysis period extended from 2001 to 2017, and our area of interest covered the 28 member countries of the European Union (EU).

However, long-period multi-annual identification of crops or crop groups at the continental scale poses serious challenges, since it needs to deal with variability in climatic conditions. Reference data are commonly used in these types of analysis. Such approaches require large amounts of reference data to cover each year and climatic zone and whose collection is labour- and cost-intensive.

These difficulties, in particular the scarcity of available reference data, might be the main reason for why multi-annual spatially explicit identification of pure crop-or crop group-specific land cover at continental extents is largely absent, although several approaches have been undertaken. Indeed, significant work has been done in the United States, where the National Agricultural Statistics Service (NASS), an agency of the United States Department of Agriculture (USDA), each year publishes the Cropland Data Layer (CDL). This dataset, derived at a 30 m spatial resolution, is a highly valued source for crop acreage estimates [16]. The CDL, created with a machine learning approach, is, in addition to remote sensing data (mainly Landsat, previously also AWiFS), heavily based on ground reference data, collected by local county offices. The CDL data were used by Massey et al. [17] in an approach with parallels to our own. They developed an automated decision tree classification to map the dominant crop types across the United States using a MODIS NDVI (250 m) timeseries.

Other data made available by USDA are the Common Land Units (CLUs), described by the same authors and created to delineate the field boundaries of registered US farmland. Similar field data also exist in Europe, collected within the LPIS (Land Parcel Identification System) of the EU member countries [18], used as the identification system of agricultural blocks or parcels within the Integrated Administration and Control System (IACS) to manage the implementation of the EU Common Agricultural Policy. However, only some of these data are made publicly available by their respective EU member states, and then often only after harvest. The LPIS systems across the different countries of the EU are not harmonized in either their methodologies or their output qualities [19]. Most LPIS systems provide land-use parcels of the size of larger blocks or cadastral parcels, often

occupied by more than a single crop, including non-productive areas. Only recently (as of 2017) has delineation of the effectively cropped land parcel containing a single crop become mandatory. This is referred to as Geospatial Aid Application (GSAA), but again is not suited for retrospective analyses as envisaged here.

In order to overcome the low availability of reliable and harmonized reference data, especially historic data, the idea was to work with as little reference data as possible. If self-parameterization does not lead to meaningful results, the system should be simplified. Obviously, this reduces classification depth and requires the creation of meaningful crop group classes, merging crops with similar properties. Such a parsimonious approach would facilitate the implementation of an operational and highly independent system which could be run on large scales.

The objective of this work was to fill these gaps and to identify and extract yearly sets of pure pixel signals. The datasets should adequately represent pre-defined crop groups for the time span 2001–2017 at the regional level, minimizing the negative effect of mixed pixels. The approach should fulfil the requirements for automation, be cost-effective, reliable, and robust.

2. Materials and Methods

2.1. Data

Since our objective was to provide a retrospective archive of pure crop pixels back to the year 2001, the choice of remote sensing data fell on the MODIS sensor (Moderate resolution imaging spectroradiometer) [20,21]. This sensor has been active since the beginning of this period and allows us to avoid data discontinuity issues. The MODIS provides a high temporal observation density and has, among available medium resolution sensors, (e.g., SPOT-VEGETATION, Satellite pour l'observation de la Terre) a relatively high spatial resolution (250 m, in 2 bands). This is expected to image Europe's farmland complexity appropriately. Because of the limited spectral details at this resolution, the focus was placed predominantly on the temporal domain.

The MODIS sensor is described in detail by the respective NASA (National Aeronautics and Space Administration) website [22]. Here, we used daily reflectance data in the red and near-infrared range at 250 m spatial resolution (bands 1 and 2). Data of other bands were indirectly used (through provided flags) for internal atmospheric correction or cloud masking.

The specific products we used for this work were MOD09GQ.006 and MYD09GQ.006 from the Terra and Aqua platforms, respectively, provided by NASA LP DAAC at the USGS EROS Center [23]. For the present application, collection 6 was used. Both datasets are daily, global, atmospherically corrected surface reflectance (L2G) products at 250 m resolution. The MODIS data are provided as gridded products in the Sinusoidal projection. In addition to MODIS bands 1 and 2, quality rating, observation coverage, and observation number are also provided in the products. The product was used in conjunction with the MOD09GA (Terra) or MYD09GA (Aqua) (500–1000 m resolution) where additional important quality and viewing geometry information is stored. In addition to these MODIS products, the following further datasets were used

1. Agro-statistical data at varying levels of administrative units (NUTS-0 to NUTS-2, Nomenclature of territorial units for statistics), collected by member states and/or EUROSTAT (European Statistical Office). These data were modified and harmonized to reflect the average coverages of Utilized Agricultural Areas (UAAs) per administrative units (up to NUTS-2) for the years of 2006–2015 [24]. These data were used to extract the number of crop group clusters per region.
2. Satellite data of Sentinel-2A and 2B (MSI sensor) for five different sites in Europe, and downloaded from the European Space Agency (ESA) Copernicus Open Access Hub [25]. High-resolution Sentinel-2 data were processed at 10 m resolution to extract the NDVI [26].
3. The Castile and Leon Crops and Natural Land Map (Spanish acronym MCSNCyL) [27], a GIS-based crop coverage product stemming from a publicly available source for the region Castile and Leon (CyL) in Spain [28,29]. Yearly LPIS datasets were used, together with other land cover data, as

training data for a high-resolution remote sensing data classification using machine learning. The described product was used for validation of the years 2011–2017.

4.　CORINE land cover 2012, version 18.5 [30], as a raster map of 250 m resolution, generated from the original vector data (with thematic accuracy >85%) by a "maximum combined area" approach. A European land use/land cover map, to distinguish arable land (irrigated and non-irrigated) from other land use/land cover classes.

5.　SRTM (Shuttle Radar Topography Mission), a radar-based digital elevation dataset (version 4) at 90 m resolution, which was void filled and complemented by other DEMs [31]. These data were used to mask highlands.

6.　Farm Heterogeneity Index (FHI), a spatial, highly detailed indicator expressing the field heterogeneity of farmland in Europe [32], used for explanatory reasons in the validation.

All computations and calculations were done with Google Earth Engine [33] and R [34].

2.2. Method

Figure 1 provides an overview of the workflow followed to detect pure crop group pixels from the MODIS timeseries across the EU. Pre-processing of the data, where first masks and filters were applied, was followed by smoothing and then by the extraction of the phenology. Finally, the phenology data were clustered and labelled into crop groups, taking stock of existing statistical data and basic agronomic knowledge. A final filtering step, aimed at minimizing noise, was applied at the end of this process.
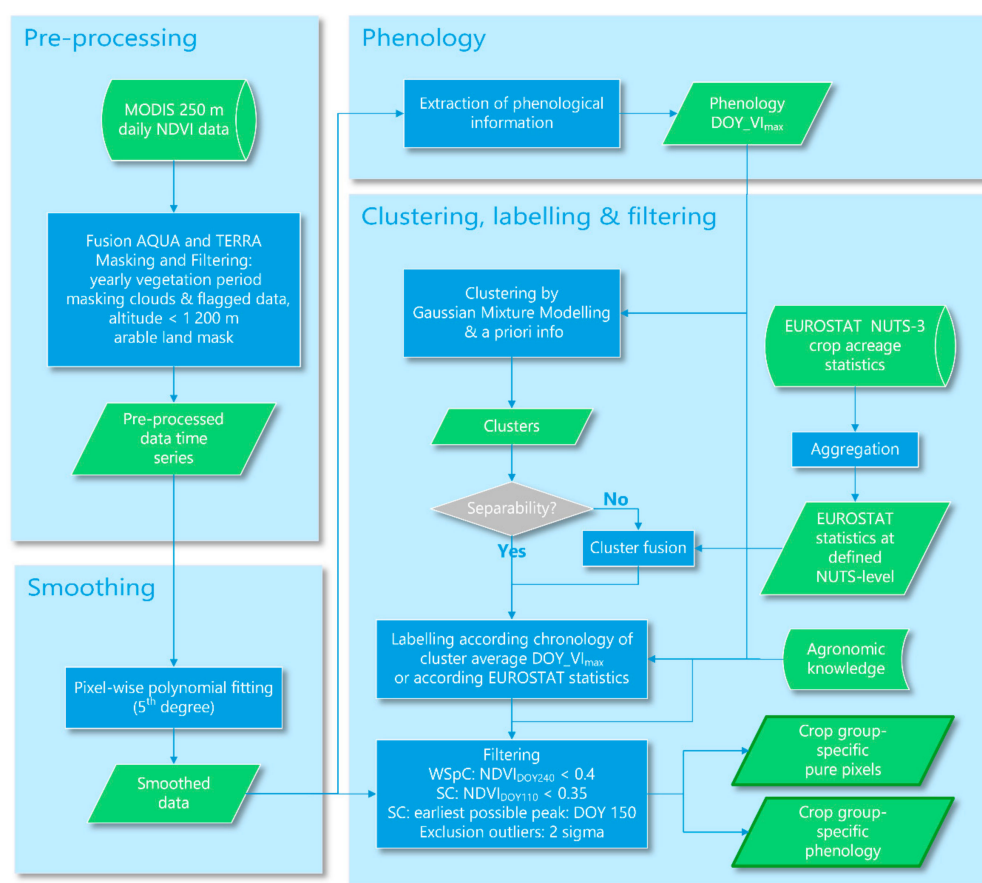

Figure 1. Flow diagram of pre-processing, phenology and noise removal, clustering, and labelling procedures.

2.3. Pre-Processing

The MODIS MOD09GQ.006 (Terra) and MYD09GQ.006 (Aqua) daily surface reflectance L2G global data at 250 m spatial resolution were merged to increase the number of observations. We limited the observation time period to the supposed growing cycle after winter dormancy of the main agricultural crops in Europe, lasting from 1st of March to 31st of August. The archive comprised the years of 2001–2017. The geographic extent was limited to the 28 EU member countries as of the end 2017. The map in Figure 2 depicts the study area. Both MODIS products (Aqua and Terra) provided surface reflectance (ρ) in the red (RED) and near infrared (NIR) electromagnetic spectrum, with which the NDVI was calculated:

$$NDVI = \frac{\rho NIR - \rho \text{RED}}{\rho NIR + \rho \text{RED}} \tag{1}$$
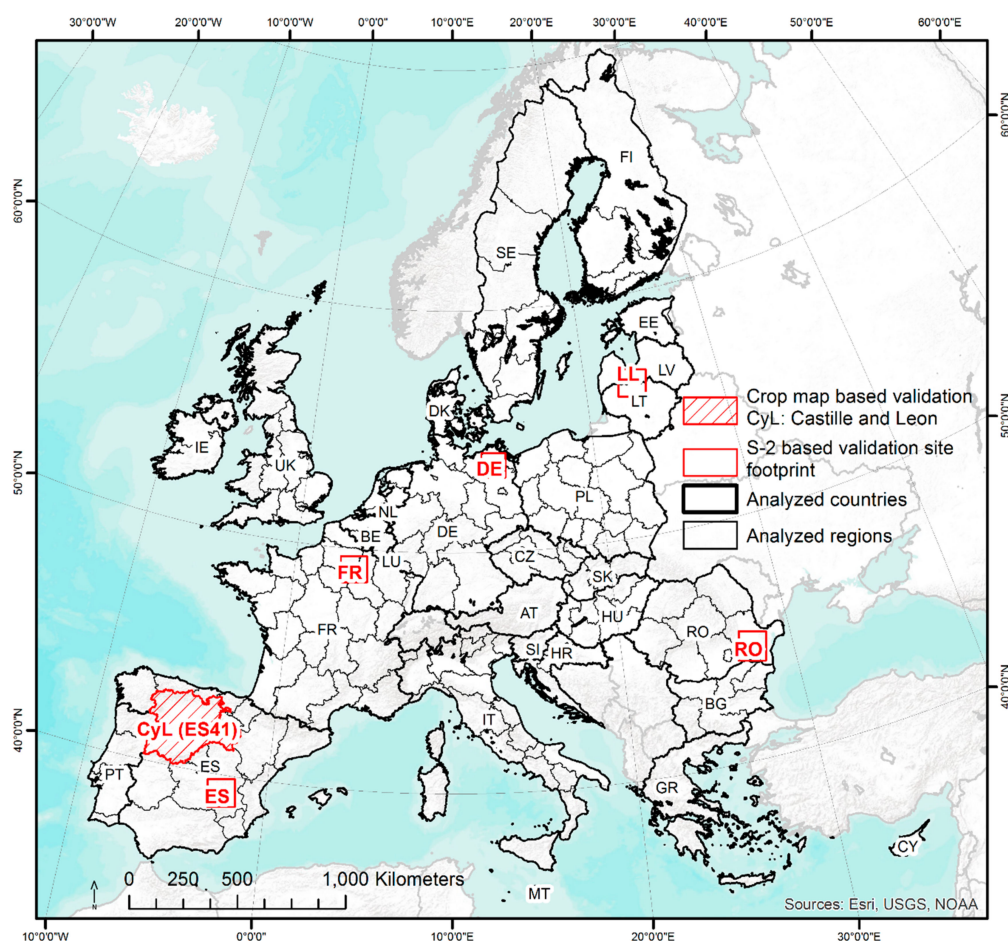

Figure 2. Study area and the analyzed NUTS regions. Validation sites in red: five Sentinel-2-based sites (S-2 footprints) and the region Castile and Leon (CyL, NUTS-region ES41) with a crop map-based validation. Analyzed countries and regions for pure pixel extraction outlined with their black borders.

The MOD09GQ and MYD09GQ both provide reflectance band quality, coded as a 16 bit unsigned integer data type. A detailed description of the quality flags is provided in the MODIS user guide [35]. The rules applied to create a positive quality mask are reported in the Supplementary Materials, Table S1.

In addition to the surface reflectance quality band, the MOD09GA and MYD09GA (500–1000 m resolution) products provide important quality and viewing geometry information, coded as quality flags, and often referred to as state QA flags. These were used to define a positive quality mask

(described in the Supplementary Materials, Table S2). The mask was applied to the group of 16 pixels at 250 m resolution within each 1 km grid cell.

Additional masks were also used, such as the SRTM digital elevation model (keeping values < 1200 m a.s.l.) and the CORINE land-use land cover product of the year 2012, version 18 [30], where non-irrigated and irrigated arable land was selected (CORINE classes 211 and 212, respectively).

2.4. Smoothing and Extraction of Phenology

The final NDVI value was calculated by fitting a 5th degree polynomial to each NDVI pixel timeseries between 1st of March and 31st of August of each year. The 5th degree polynomial was chosen to appropriately map multi-peak NDVI timeseries on arable land. This smoother was empirically tested on the data and found to be adequate.

To maintain consistency in the coverage and quality of phenological data that are otherwise reported in only a few areas of Europe [36], the NDVI timeseries itself was used to extract phenology. The timing of NDVI peak (DOY_VImax), expressed as DOY, considers both spectral and temporal EO-based features of the land cover. This way it provides a sort of synthesis of both domains and might therefore be considered an information-rich indicator. As documented in scientific literature for wheat and maize, the timing of NDVI peak (DOY_VImax) occurs around the booting or heading date [37,38], from shortly before flowering to the time of flowering (silking of maize [39]).

The analyzed growing season was extended by one month for the countries Hungary (HU), Romania (RO), and Bulgaria (BG), where important phenological events can still take place in September (late NDVI peaks possible). The DOY_VImax values occurring before DOY 115 were excluded for the countries Ireland (IE) and United Kingdom (UK) to avoid misinterpretations caused by early greening.

Phenological trend magnitude was calculated according to Sen [40], with trend significance following the Mann–Kendall test [41]. Regions with on average less than 100 observations/year were not considered for trend analysis.

2.5. Clustering, Labelling, and Noise Filtering

The DOY_VImax values from the smoothed NDVI timeseries were extracted and considered as a population of statistical samples. For further processing in R, the data were re-projected from the native sinusoidal projection to CORINE's projection, EPSG code 3035 (European Petroleum Survey Group), by a nearest neighbor resampling.

Pixels were region-wise and year-wise clustered by Gaussian mixture modelling (GMM), as implemented in the R package mclust [42]. This technique was recently successfully applied by Skakun et al. [43] for winter crop mapping in the US state of Kansas and in Ukraine and was considered appropriate for our goal.

$$p\left(\vec{x}\right) = \sum_{i=1}^{k} \varphi_i N\left(\vec{\mu}_i, \Sigma_i\right) \tag{2}$$

The assumption of GMM (Equation (2)) is that an overall sample population is composed of *k* normally distributed subpopulations (components $Ni \ldots k$), resulting in a probability density function $p(x)$. Each component of a mixture has its own mean (μ_i) and variance (σ^2_i) (or vector $\vec{\mu}_i$ and covariance Σ_i for the multivariate case). Additionally, each component is defined by a component weight φ_i, which is constrained to sum to 1 for all components. The component weights are learnt during the unmixing process; the parameters of the GMM are estimated using an expectation–maximization (EM) algorithm. The EM is an iterative technique for maximum likelihood estimation of probabilistic models [44]. The variable used for clustering is DOY_VImax.

The resulting clusters normally correspond to crop groups. If DOY_VImax occurs early in the year, the cluster will be winter and spring crops (WSpCs, small-grained cereals, and rapeseed), if it occurs later in the year, it will be summer crops (SCs, sugar beet, potato, sunflower, maize, soybean). The classification to crop groups relies merely on the chronological order at which the created clusters

(maximum 2) reach their NDVI peak. If only a single cluster was detected, this cluster is assumed to represent the crop group with the highest crop acreage according to EUROSTAT statistics.

The resulting clusters were checked for their separability. The average DOY_VImax values of different clusters within the same region are not allowed to be closer than 10 days, as this is no longer considered separable. If they are closer, the clusters are merged, and the number of clusters is reduced to 1.

Depending on which crop group a pixel is assigned to, threshold-based filters for a set of agronomically-based rules were applied to ensure a high degree of pixel purity. These filters require that, at a certain time during the growing cycle, NDVI timeseries values, which remain linked to the classified cluster pixels, pass certain thresholds. In particular, for the class WSpCs, the NDVI value on DOY 240 must stay at a value lower than 0.4, since crops are by then harvested or at a low level of chlorophyll, reflected in a low NDVI. In contrast, at this stage, high NDVI values do typically occur for SCs. For SCs, instead, the NDVI is required to be lower than 0.35 on DOY 110, assuming that no significant biomass is accumulated by that day, and that this, in turn, is reflected by a low chlorophyll content and a low NDVI value, whereas WSpCs would instead be identified by higher NDVI values at this time. For SCs, it was also required that the NDVI peak did not occur earlier than DOY 150. The mutual exclusion at the selected DOYs was expected to be robust over all of the area of interest, since values were chosen generously, leaving some margin for the event of exceptional years. Additionally, a two-sigma criterion (±2 SD) was applied to each class population, to filter atypical values or outliers, as often occurs in mixed pixels.

The input samples for each GMM model were typically limited to sub-national administrative regions (NUTS-units). The cluster building process relies on self-parameterization requiring input of only the number of clusters to build. Given this individual region-wise processing, site-specific characteristics of the climate, phenology, and agro-management regime did not require any prior normalization, and the results (clusters) were comparable with respect to the individual characteristics intrinsically contained within them.

The choice of NUTS-level determines the size of its spatial elements. An adequate choice aims at maximizing crop-relevant regional homogeneity (e.g., climate and phenology) but also at confining the total number of regions, guaranteeing at the same time a certain critical number of samples per region. An additional criterion for the choice of NUTS-level was a harmonization of the spatial element size. The chosen NUTS-levels for analysis (ranging from NUTS-level 0 to 2) of the countries are reported in Table 1 and are visualized on the map in Figure 2.

Table 1. NUTS-levels for analysis.

NUTS Level	Countries
0	AT, CY, EE, GR, LT, LU, LV, MT, SI
1	BE, DE, NL, UK
2	BG, CZ, DK, ES, FI, FR, HR, HU, IE, IT, PL, PT, RO, SE, SK

We employed evidence-based data to define the number of clusters. The EUROSTAT reports, for the European administrative units of usually larger unit sizes (NUTS-1 or NUTS-2), the crop acreages of the main agricultural crops. National statistics agencies report these data often at even finer resolution. Researchers [24] have collated and elaborated these data by homogenizing, disaggregating, and completing it in such a way that the crop acreage statistics are available for all EC member countries at NUTS-2 level back to at least the year 2006. These crop acreage statistics serve to define the expected number of clusters. After grouping the reported crop acreages per NUTS analysis level (see Table 1) according to the grouping procedure described in Table 2, each crop group's relative importance (RI) was estimated as the crop area share averaged over the years of 2006–2017 of a NUTS analysis unit. If a crop group's RI was larger than or equal to a threshold of 10%, the group was considered for analysis, otherwise it was neglected.

Table 2. Identified crops or crop groups.

Crop Group Code	Crop Group Acronym	Crop Group	Contained Main Agricultural Crops
1	WSpCs	Winter and spring crops	Soft and durum wheat, winter and spring barley, triticale, rye, oats, rapeseed
2	SCs	Summer crops	Sugar beet, potato, sunflower, maize, soybean

2.6. Validation

The validation of the results focused on pixel purity, assessed by high-resolution satellite data. Crop group-specific pixel purity was calculated as the average coverage of the pure MODIS pixel extent by pixels from classified high-resolution satellite data of the same class. Two independent validation approaches based on different validation datasets were undertaken. Both datasets rely on high-resolution remote sensing data (between 10 and 20 m), assumed to provide sufficient detail and reliability for the assessment of pixel purity.

2.6.1. Sentinel-2-Based Validation

The Sentinel-2 (S-2)-based validation approach was pursued to counter the lack of high-quality reference data suitable for validation. The processed reference data are an expert-based binary classification of S-2 high-resolution remote sensing data. They provided the classes WSpCs and SCs, following a highly similar although simpler process to the one presented here. This approach was applicable to any site within the area of interest, and a sufficient number of well-chosen sites may represent an even more representative database for validation than sporadically available data sources elsewhere. The five sites (100×100 km, corresponding to S-2 tiling grids, see Figure 2) in Spain (ES), France (FR), Germany (DE), Romania (RO), and the border area between Latvia and Lithuania (LL) all represent regions with a high arable land share and the presence of both WSpCs and SCs. Validation site DE has a high share of rapeseed. Agro-climatically, the sites represent the majority of possible climates, ranging from low latitude, very warm and dry, to high latitude, very cold and humid climates.

In order to produce high-resolution reference maps, 418 Sentinel-2 level 1C (top-of-atmosphere reflectance) acquisitions were atmospherically corrected to (L2A) top-of-canopy (TOC) reflectance products by the simplified model for atmospheric correction algorithm (SMAC) [45], using MODIS-based aerosol optical density (product: MYD04_3k, collection 6) and the default SMAC values for ozone and water vapor. This algorithm provided a good relation between processing time and output quality. Scene cloud contamination was estimated following the methodology implemented in the Sen2cor processor [46], where for a pixel, a minimum of 15% cloud presence probability was required to be classified as cloud. Additionally, Sen2cor's cloud shadow radiometry-based binary mask was applied. The NDVI was calculated from band 4 (red) and band 8 (near infrared) at 10 m spatial resolution with 5 day temporal resolution. The resulting timeseries was smoothed using a Whittaker interpolation approach [47,48]. The use of a different smoother compared to the MODIS processing was not considered problematic, as the crop group separation based on DOY_VImax is considered highly robust with typically large margins between crop groups. The timeseries for a current growing season was extracted on arable land according to CORINE land cover and for pixels accounting for at least 10 cloud-free observations within the timeseries. After classifying DOY_VImax values by a threshold-based approach, the frequency and, hence, the pixel share of WSpCs and SCs within each MODIS pixel were calculated. The classification rules to extract the two groups, WSpCs and SCs, were scene-histogram based, separating two sub-populations of DOY_VImax at the lowest frequency of their scene-histogram. In single cases, this general rule did not lead to a meaningful separation and the dividing DOY_VImax value was defined based on agronomic knowledge.

2.6.2. Crop Map-Based Validation

The second reference dataset was a classified dataset with detailed land use information and, in particular, crop classes covering a large and important agricultural region of Spain (NUTS-2 region ES41, see Figure 2) for the years of 2011–2017. These data offer a high degree of data independence, being generated by a third party and based on an independent data source. The data is referred to as Castile and Leon crops and natural land map [27]. The sensors Deimos-1 (2011–2016), Landsat 8 (2013–2016) and Sentinel-2 (2016 until now) provide the predictors for the classification. The training data were extracted from LPIS (and more recently from GSAA), the GIS of the Integrated Administration and Control System for Common Agricultural Policy. Spain's LPIS follows the cadastral parcel as reporting geometry. The overall classification accuracy of these data is 82% on average, being generally much higher in crop classes than in natural land [28,29].

3. Results

The results were obtained at the chosen NUTS-level (Table 1) for all EU member countries except Malta and Finland, and for all years from 2001 to 2017. Due to the insufficient areas of arable land, the countries Malta and Finland and other single NUTS regions were automatically excluded from the analysis (e.g., Corsica in France or Berlin in Germany).

3.1. Pure Pixel Results

The pure pixel results are shown in Figure 3 over the European extent. In order to visualize all the years simultaneously, we depicted the average share of a crop group's pure pixel density over arable land for the whole timeseries of 2001–2017 on a grid size of 10 by 10 km. The legend's color ramp indicates the share of the specific crop group density in relation to the sum of both crop group densities. The shares had a range of 0–1 and were complementary for WSpCs and SCs. It is important to note that the results depicted below should not be confused with a fully exhaustive crop classification, such as a wall-to-wall classification. An example of the real distribution of pure pixels in a single year is visible in more detail in Figure 4. The clustering and labelling were performed on a subset of arable land pixels, those of highest purity. Although its sample size should, in general terms, be in line with the proportional area covered by this crop group on a regional extent, this is, especially for a particular year, not necessarily the case.
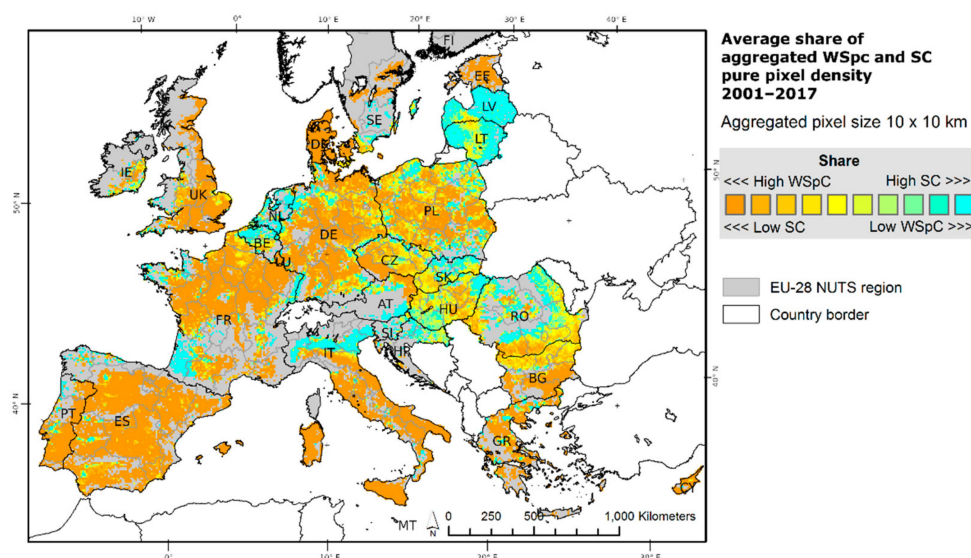

Figure 3. Average share of the WSpCs and SCs pure pixel density from 2001–2017 over arable land, aggregated to a pixel size of 10 × 10 km.
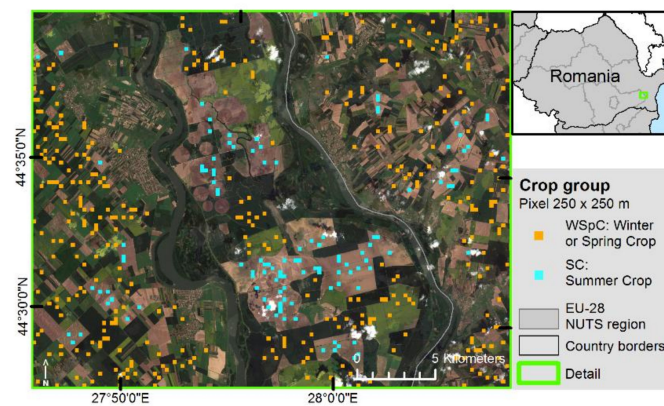
Figure 4. Details of a single year (2017) for a subset of the Romanian validation site (RO), showing the real size of pure pixels (250 × 250 m), overlaid on a Sentintel-2 RGB maximum value composite for May 2017.

Looking at the results at a regional level, a small-scale map reveals more geographic details in the last column of Figures 5 and 6. In this case, the examples are from Aragón in Spain and from the Sud-Est region in Romania. The first column shows the probability density function of the pure pixels, underlined by the respective crop group color. The second column shows the NDVI timeseries of the crop groups; both graphs are shown as examples and for a limited number of years (as rows).
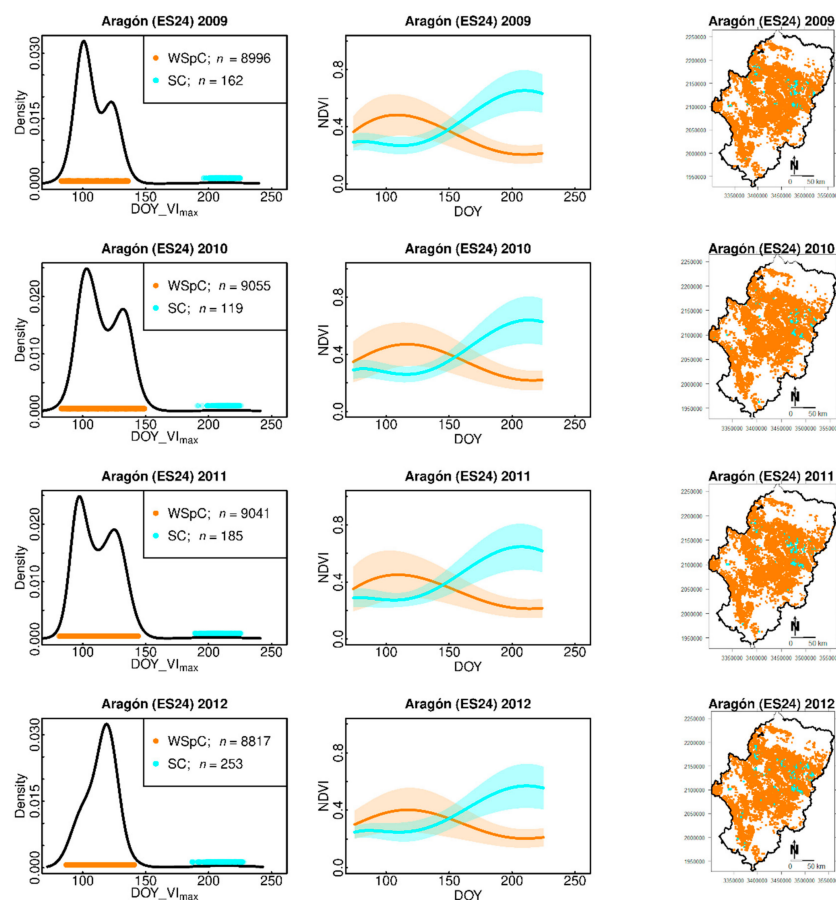

Figure 5. Probability density function of DOY_VImax (first column), NDVI timeseries (2nd column), and geographic distribution (3rd column) of the two crop group's pure pixels for the region Aragón (Spain) from 2009–2012 (by row). Winter/spring crops (WSpCs, orange), summer crops (SCss, cyan). The timeseries depicts mean run ± 1 SD.
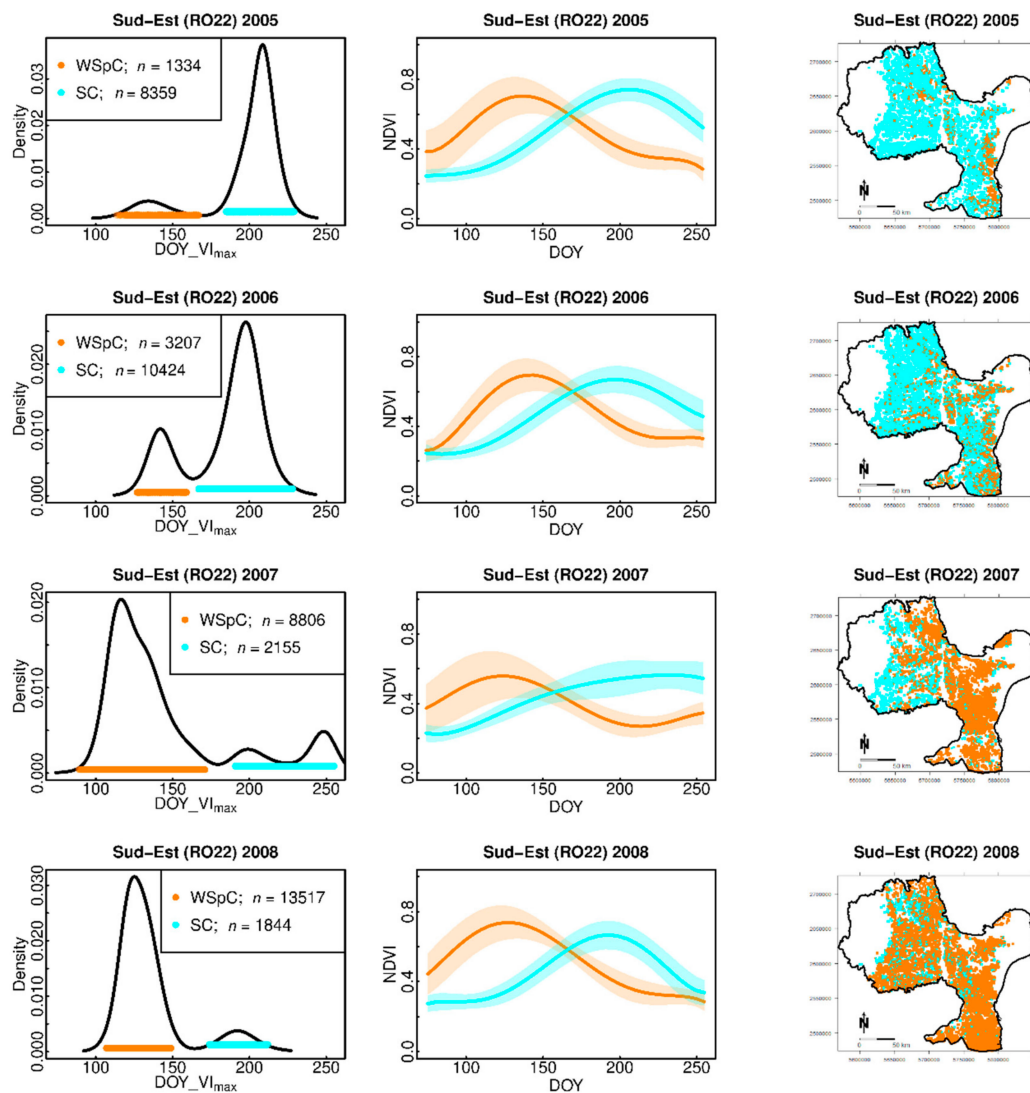
Figure 6. Probability density function of DOY_VImax (first column), NDVI timeseries (2nd column), and geographic distribution (3rd column) of the two crop group's pure pixels for the region Sud-Est (Romania) in 2005–2008 (by row). Winter/spring crops (WSpCs, orange), summer crops (SCs, cyan). The timeseries depicts mean run ± 1 SD.

The algorithm's ability to deal with varying phenological expression across several years can be well observed for the case of Aragón (Figure 5). For example, the probability density function clearly shows an earlier and later peak of WSpCs in the years of 2009–2011, which are, however, both classified correctly as WSpCs, as is the single peak classification in 2012 conflating the winter and spring crop signal. These graphs show the versatility of the algorithm, able to deal correctly with different situations such as the consequences of inter-annual differences in phenology. Around DOY 200, a few SCs were detected (their number is depicted in the graph's legend), yet extracted correctly, as the classified timeseries in the middle column indicates. The geographic distribution reveals high stability for both crop groups across the years shown.

Figure 6 shows a contrasting example for the Sud-Est region of Romania over the period of 2005–2008. Here it is clear that the algorithm, thanks to its robustness, can deal with large inter-annual variability in phenology when detecting WSpCs and SCs. Note that the DOY_VImax of WSpCs in 2007 was advanced by about 3 weeks, compared to 2006, recognizable from the probability density function (1st column) and the timeseries (middle column). In 2007, moreover, SCs were composed of two sub-clusters with distinct DOY_VImax peaks (see 1st column), which were both classified correctly as

SCs. It can also be seen that the group-wise number of samples of pure pixels can change considerably among years, a fact that can have several reasons, the main two being (i) real acreage changes and (ii) changing plant conditions expressed by the NDVI. Both reasons might also occur at the same time, as was the case in this example.

3.2. Validation of Pure Pixels

Pixel purity is closely related to the algorithmic behavior of the filtering, which we briefly explore here, since it helps to understand the full complexity of pure pixel quality. As already mentioned, pixels not fulfilling the multiple NDVI thresholds and the two-sigma criterion were filtered out after running the GMM. Assessing preference or unevenness in the pure pixel filtering process is of high interest, since the pure pixel representativeness and comparability might be compromised. The exclusion of samples (pixels) according to these defined filters was measured with an exclusion rate (ER) = number of excluded samples/number of original samples. The ER was generally very high, with a median of 89% (ER interquartile range: 15%, 5th, and 95th percentile: 49% and 99%, respectively). A slight tendency to retain more samples was observed in Mediterranean countries, in particular in Spain (ER median: 59%), which is probably related to a higher spatial separation of the two crop groups' geo-locations; SCs are typically irrigated spots within wider WSpCs areas. Lower exclusion rates were observed also for the Czech Republic (ER median: 75%) and Slovakia (ER median: 77%), which is probably due to the larger patches of the same crop group, compared to other countries. Analytically, and over the European continental extent, the ER was compared to the farm heterogeneity index (FHI). The correlation (linear-log model) of ER and averaged FHI regional values yielded $R = 0.50$, indicating a positive correlation of the ER with heterogeneous farmland (Figure 7). This suggests that a more homogeneous agricultural landscape does generally result in less noise and, consequently, in less data exclusion by filtering.
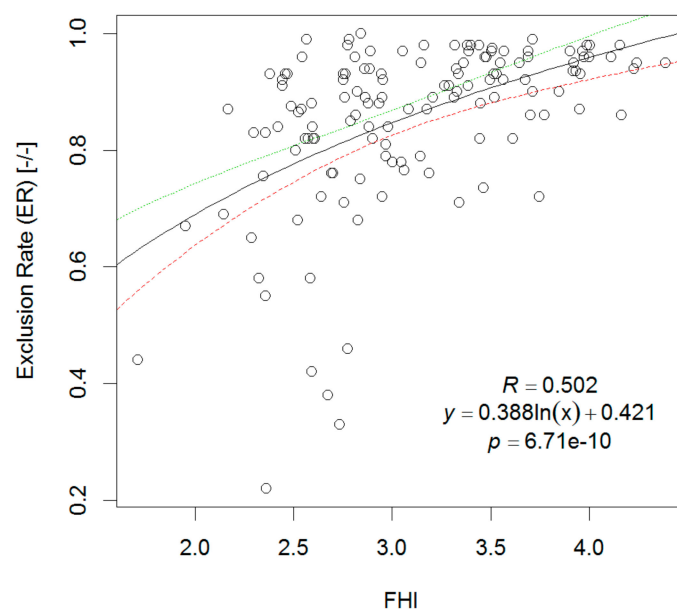

Figure 7. Exclusion rate (ER) versus farmland heterogeneity index (FHI), compared as the regional medians and means, respectively, as colored lines; the confidence interval was at level 0.95.

Pixel purity is closely related to the algorithmic behavior of the filtering, which we briefly explore here, since it helps to understand the full complexity of pure pixel quality. As already mentioned, pixels not fulfilling the multiple NDVI thresholds and the two-sigma criterion were filtered out after running the GMM. Assessing preference or unevenness in the pure pixel filtering process is of high interest, since the pure pixel representativeness and comparability might be compromised. The exclusion of

samples (pixels) according to these defined filters was measured with an exclusion rate (ER) = number of excluded samples/number of original samples. The ER was generally very high, with a median of 89% (ER interquartile range: 15%, 5th, and 95th percentile: 49% and 99%, respectively). A slight tendency to retain more samples was observed in Mediterranean countries, in particular in Spain (ER median: 59%), which is probably related to a higher spatial separation of the two crop groups' geo-locations; SCs are typically irrigated spots within wider WSpCs areas. Lower exclusion rates were observed also for the Czech Republic (ER median: 75%) and Slovakia (ER median: 77%), which is probably due to the larger patches of the same crop group, compared to other countries. Analytically, and over the European continental extent, the ER was compared to the farm heterogeneity index (FHI). The correlation (linear-log model) of ER and averaged FHI regional values yielded $R = 0.50$, indicating a positive correlation of the ER with heterogeneous farmland (Figure 7). This suggests that a more homogeneous agricultural landscape does generally result in less noise and, consequently, in less data exclusion by filtering.

Figure 8 shows a subset of the validation site in Romania in 2017, where the identified pure pixels are overlaid on the Sentinel-2-based validation data.
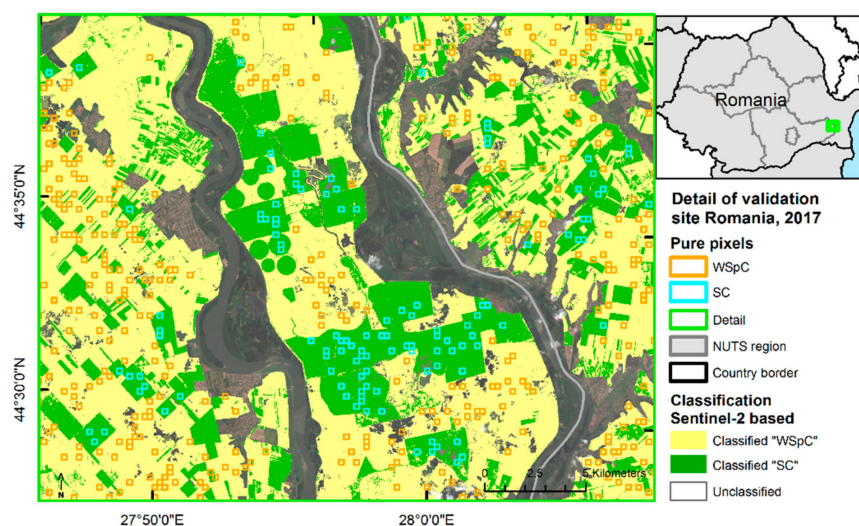

Figure 8. Details of the validation site in Romania: pure pixels (250 m spatial resolution) overlaid on the Sentinel-2-based classification.

Class-wise pixel purities for five validation sites in 2016 and 2017, based on Sentinel-2 high-resolution data (10 m resolution), are provided in the upper part of Table 3. According to this validation scheme, pixel purities for WSpCs were on average 93%. For SCs, the pixel purities were less homogeneous than for WSpCs, while pixel purities were still larger than 80% at the validation sites in Romania (RO) and Spain (ES) as well as in Germany (DE) and France (FR) for 2017; they were significantly lower in 2016 at the latter two sites and were particularly low for Lithuania and Latvia (LL). However, the low pixel purities at these sites should not be over-interpreted, since the year 2016 was particularly compromised by low pixel availability (see the number of samples reported in Table 3) due to the high cloudiness in the S-2 data which, especially for SCs, highly compromised the results. Regardless of this low data availability, the temporal vegetation profiles of WSpCs and SCs in extremely late maturing areas can be easily confused, as they resemble each other to a high degree. Lithuania and Latvia (LL) in particular represent one of the most extreme cases in Europe of this late maturation; cereals are harvested almost together with potatoes (which are a very prominent SCs there). Due to the resulting confusion, SCs are overestimated there, leading to low pixel purities. In DE, instead, the low pixel purity for SCs in 2016 was also due to the slight underestimation of S-2-based reference SCs, caused by a later DOY_VImax of WSpCs that year, resulting in a later division threshold value among crop groups.

Table 3. Pixel purity for each validation site.

Site	NUTS Name/NUTS code/S-2 Graticule	Year	*n* Samples		Purity (%): Mean	
			WSpCs	SCs	WSpCs	SCs
RO	Sud Muntenia/	2016	5830	985	89.51	98.15
	RO31/35TNK	2017	3831	1671	91.15	97.41
ES	Castilla La Mancha/	2016	2743	69	91.73	80.29
	ES42/30SWJ	2017	1959	53	88.71	87.64
DE	Mecklenburg-Vorpommern/	2016	410	63	98.5	59.9
	DE8/32UQE	2017	1548	188	98.15	90.57
FR	Champagne-Ardenne/	2016	190	10	98.77	54.1
	FR21/31UEQ	2017	1278	58	94.26	80.12
LL	Border Latvia + Lithuania/	2016	1014	274	98.97	13.77
	LV+LT/34VFH	2017	3288	2445	97.83	21.63
Weighted average, all years, all S-2 regions					92.80	60.57
Weighted average, all years, all S-2 regions without LL					91.52	90.77
CyL (ES)	ES41	2011	18,986	561	74.49	73.42
		2012	19,119	925	71.88	72.24
		2013	17,571	284	72.82	77.41
		2014	18,566	578	72.85	74.77
		2015	17,474	932	74.03	72.30
		2016	19,417	263	74.25	76.31
		2017	18,225	677	73.29	61.70
Weighted average, all years, region CyL					73.37	71.76

Turning to the pixel purity validation based on the crop map of Castille and Leon (CyL), Spain, in 2011–2017 (lower part of Table 3), generally lower pixel purities were found compared to the S-2-based ones, and the results were rather more homogeneous.

3.3. Phenology-Related Results

We extracted crop-specific phenological data at the locations of identified pure pixels. The data are available for all years, and some examples of these data are presented in Figure 9, where the average timing of the NDVI peak (average DOY_VImax of 2001–2017) is depicted for WSpCs (A) and SCs (B) across Europe by averaging the values per region. Figure 9C,D shows the DOY_VImax trends over the time period 2001–2017 for WSpCs and SCs, respectively.

The difference between the geographic patterns of WSpC- and SC-based DOY_VImax values (Figure 9A,B) is striking: while WSpCs followed a mostly latitude-based and, therefore, climatic gradient, with earliest DOY_VImax values in the south and latest in the north, the pattern for SCs was much more heterogeneous across Europe.

The crop-specific trends in the DOY_VImax values over the period 2001–2017 are shown in Figure 9C,D. The WSpCs trend was negative for 90% of the analyzed regions over the area of interest (total average −4.76 days or −0.28 days/year), meaning that the NDVI peak was occurring earlier in large parts of Europe, which is in line with other research on climate change effects for wheat [49]. For SCs, the trends were generally less obvious (total average +0.92 days, +0.05 days/year), but large areas were dominated by either positive trends (Central Europe) or negative trends (South and South East Europe).
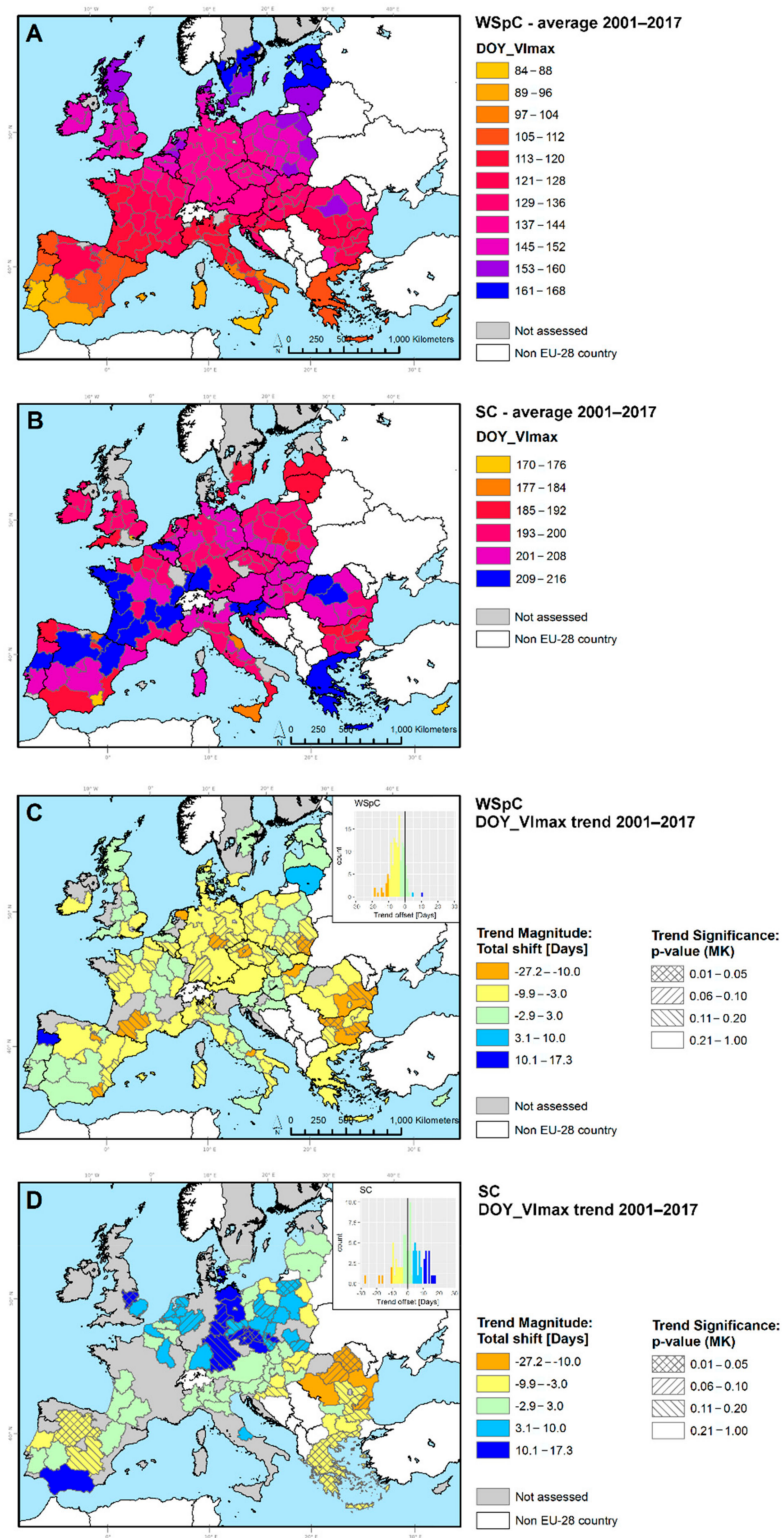
Figure 9. Average DOY_VImax values of the years 2001–2017 for WSpCs (**A**) and SCs (**B**), averaged by region. Magnitude and significance of DOY_VImax trends from 2001–2017 for WSpCs (**C**) and SCs (**D**), averaged per region.

4. Discussion

The motivation to consider crop groups was that single crops are typically difficult to separate on an operational basis with this kind of remotely sensed data. Despite a loss in classification depth,

grouping is still effective if looked at from the point of view of data utilization. The majority of these grouped European agricultural crops appear very similar in terms of their phenological life cycle, based on NDVI [50]. While this similarity is particularly noticeable with respect to the timing of maximum vegetation index, it also applies for the majority of crops to large parts of (or the entire) crop life cycle. Therefore, these crop groups can have a high potential to act as surrogates for single crops contained within them.

This methodology was applied separately on each individual region and year across the EU, minimizing problems related to regional characteristics, climatic inter-regional and inter-annual variability, or distinct local agronomic practices. Despite independent region-wise data treatment, the transitions at regional borders were generally smooth and seamless. However, a more abrupt change was observable at the border of Latvia and Estonia, which was due to the overestimation of SCs in Latvia (and Lithuania). For this reason, we found the lowest purity values in this area, where the crop group-specific DOY_VImax were temporally too close to each other and, therefore, the crop groups could not be discriminated. In such regions, an additional variable not based on NDVI peak timing could improve the results.

The intensity of a specific crop group might differ from year to year. Such differences may but not necessarily mean that the cultivated area has changed. They often result from climatic events, favoring or hampering events of crop growth (dry or wet periods, late or early sowing, etc.) which, for a specific crop group, might be reflected in a lower or higher presence of pure pixels. Most permanently irrigated areas, usually SCs, remain highly stable throughout the years, such as those used to cultivate summer crops in warm and dry areas of Spain, southern France, Italy, Greece, and Bulgaria. These areas are typically visible as small spots within larger WSpCs areas.

We found lower exclusion rates for Spain and some other countries (CZ, SK). We concluded that all the explanations for these observations point in the same direction: larger patches lead to less noise and, consequently, to a lower filtering rate. This theory is corroborated by an analysis of regional data, where regions in eastern Germany, with their known large field structures, were found to be slightly less prone to data filtering. Accordingly, in this analysis, the pixels show a slight preference for more homogeneous farmland, often hand in hand with larger agricultural patches.

As a rule of thumb, we can say that a crop group suffering from unfavorable growth conditions is represented by a smaller number of samples, due to the more inherent noise, which is eventually filtered (e.g., stemming from soil signals). This can be observed for example in Figure 6. Looking at the NDVI timeseries in 2007 and 2008 for SCs (middle column), lower than usual NDVI (2007) or an earlier NDVI peak (2008) suggest sub-optimal plant conditions, resulting in lower sample volumes same as for 2005–2006. Highly vigorous and dense stands, instead, produce less noise, resulting, generally, in a higher number of samples. Once again, this means that a change in crop group acreage cannot unambiguously be deduced from the sample volume. This should be taken into account when reading the pure pixels' geographic distribution (Figure 6, last column). Such changes in sample frequency might also occur solely due to the plant status. The question, if whether the sampling preference for vigorous crops leads to significant changes on the averaged NDVI run per crop group, is legitimate. Apart from a lower representativeness due to the fewer samples, it could have an important effect, especially if crop conditions are not homogeneous within a region. Then, vigorous crops would tend to rule out poorer ones of the same category. This could happen in particular in larger regions with stands in either good or poor crop status, or in regions with a high climatic gradient, leading to advanced and delayed crop phenologies. In dry years, in particular, irrigated SCs could be overrepresented if present together with non-irrigated SCs in the same region. It becomes clear that the choice of region size (NUTS-level) is highly important; on the one hand, its size should be large enough to guarantee a regular run of the algorithm, providing enough samples; while on the other hand, it should be small enough to limit plant or phenological heterogeneities caused by, for example, climate or topography.

Other factors that contribute to pixel impurity might include the optical properties of the MODIS sensor. The observation geometry of MODIS plays an important role, which is as a wide field sensor

(swath of 2330 km) prone to the negative effects of possible high off-nadir viewing angles towards the margins of the swath. A related consequence of this is the increasing incongruence, along with rising viewing angle between the actual observation footprint and the gridded pixel in the processed image. While the grid area geometry of an observation point (once filled with values, becoming a pixel) is set to a constant depending on the spatial resolution, the observation footprint area of a single optical element of the sensor's optical array increases primarily as a function of viewing angle. This can create high noise levels depending on the degree of land cover heterogeneity in the immediate neighborhood of an observation point. Such noisy data then leads to mixed signals, being filtered out with the algorithm as presently described, but also causing an increase in the omission error. This is discussed in the scientific literature [51] in more detail.

The applied filters might locally exclude double cropping systems, which in Europe, due to the climatic constraints, are primarily systems of one main crop and a cover crop. These systems are more intensively practiced in Central–Western Europe [52]. Depending on the type of analysis being conducted with pure pixels, this should be considered, although one would generally only expect minor effects.

In terms of pixel purity, we found rather different values for the S-2-based and crop map-based validation. There might be several reasons for the lower purities in the CyL site. Firstly, the CyL land cover map is a more sophisticated product than the simple S-2-based DOY_VImax grouping into two crop groups; there is more detail in the CyL land cover map, which, compared to S-2-based grouping, tends to identify more impurities. The CyL land cover map relies on spatially more detailed information, while the S-2-based validation data are created on an arable land mask according to CORINE land cover, which in turn was created with a large minimum mapping unit of 25 ha. Secondly, the CyL validation dataset itself contains inherent uncertainties: the dataset is reported with an overall accuracy of 84% [29]. Some classes, however, do not reach this accuracy and, therefore, provide, when used as reference data, a potential source of misclassification.

The strength of the crop map dataset (CyL) is its independence in terms of base data and methodology, whereas the S-2-based validation dataset, although of higher spatial resolution, is derived using a highly similar technique to the one used to derive the pure pixels, in this way perhaps overestimating the real purity values. Considering the drawbacks (e.g., the inherent uncertainty) of the crop map validation of CyL on the one hand, and the methodological similarity and simplicity of the S-2-based validation dataset on the other, we assumed that the real purity values lie somewhere in among their resulting values. Additional crop map-based validation data sources and more S-2-based validation years, which we did not obtain or were not of the desired quality, could have improved the reliability of the pixel purity assessment.

A comparison of our results with other research is difficult, since there is little available for large-scale assessments and on a multi-annual basis. Massey et al. [17] identified comparable crop classes such as corn–soybean, potato, and wheat–barley and provided one-year's (2008) user's accuracy (UA, complement of commission error). User's accuracy is a better basis of comparison for pixel purities than producer's accuracy or overall accuracy, since the pure pixels are based on a selective procedure where an omission error assessment is not meaningful. However, a direct comparison of UA and pixel purity needs to be interpreted in a qualitative way. Massey et al. [17] found UA values for the crop classes corn–soybean, potato, and wheat–barley of 77.4%, 96.2%, and 74.5%. Our pixel purities on the S-2 validated sites were on average 92.8% and 60.6% (91.5% and 90.8% without LL) for WSpCs and SCs, respectively, and for the crop map-based validation, 73.4% and 71.8%, respectively. If we assume our purity values to lie somewhere in between the S-2 and crop map-based validation, we can deduce that we achieved for WSpCs clearly higher purity values than the reported UA. Similarly, we found higher purity values for SCs than the reported UA, at least on the sites not too far north if we consider that the influence of the potato class, despite its high UA, is very low compared to the corn—soybean class (potato covers < 0.5% of total US cropland). We can therefore assume that our pixel purities on the majority of sites were most probably higher in quality than the reported results of

Massey et al. [17], both for WSpCs and SCs. This is of even greater importance considering that we did not use any field data and that the European agricultural landscape was probably more complex than the US-American pendant. On the other hand, it was not unexpected, since we focused on the selection of pure pixels and not on a full wall-to-wall map. Skakun et al. [41] identified winter crops without field reference data by a GMM-based within-season approach using MODIS NDVI (250 m) in multiple years in Kansas and in one year in the Ukraine. They achieved impressive UA values of, on average, 95.7% in Kansas and 85.1% in the Ukraine. It has to be noted in our favor that both test sites exhibit large agricultural fields and a high winter crop proportion, which is expected to facilitate the classification compared to the more complex European reality. We indeed achieved similar qualities, for example, on the test site in Eastern Germany which is known for its larger field sizes.

For certain applications, some of the mentioned limitations were less relevant, since the data were averaged over the area or time, for example, the creation of crop group-specific regional average maps (e.g., the presented results of phenology, Figure 9) which can assume a wall-to-wall character. However, we can derive much more details from the pure pixel timeseries, as shown in Figure 3, where we visualized a map of the average share of WSpCs and SCs pure pixel density of all analyzed years with a 10 × 10 km grid cell size. This map provides a glimpse of what is possible with these results. If we assume that the geographic crop distribution on the ground is not subject to major changes, then an average geographic crop group distribution within a grid cell is quite well represented by the density of pure pixels over multiple years, i.e., within a 6 to 8 year time window. In this case, the distorting effect of single years, in which the density of crop group-specific pure pixels might not represent so well the typical acreage, are likely to level out. To normalize inter-regional differences, such as stronger sampling in particular NUTS-regions (e.g., a larger field size), it is advisable to calculate the crop group-specific share of aggregated pure pixel densities which is indeed shown in Figure 3.

Such a multi-year pure pixel density map, although considered static as it expresses a preference for a crop group over multiple years, may still present, in many cases, a pragmatic crop mask proxy if annual crop masks are unavailable or insufficient in terms of density. Assuming no particular crop group-specific preference in pure pixel sampling, it can be considered a probabilistic crop mask of the considered time period.

The derived data may also help to fill an important data gap in crop modelling and monitoring systems related to phenology. Spatialized knowledge of crop group-specific occurrence of the NDVI peak, often related to the occurrence of heading or flowering, presents an important contribution to crop calendars (Figure 9 A,B)

The smooth gradient of WSpCs (Figure 9A) is, to a high degree, due to the fact that this crop group is usually rainfed, hence heavily constrained by climatic conditions. This results from water-extensive management, with earlier sowing dates in water-limited areas (predominantly the south), in turn, leading to smaller DOY_VImax values. Towards cooler and more humid zones, management decisions are increasingly determined by temperature-based criteria, leading to the latest sowing dates in the coldest areas (in predominantly the north), hence leading to later DOY_VImax values. The SCs instead (Figure 9B) often rely on irrigation in drier regions, which allows a certain detachment from natural climatic constraints such as rainfall or soil humidity. This enables longer photosynthetic assimilation periods to achieve higher yields, as this is practiced, for example, with dedicated maize varieties. If not irrigated, analogous water and temperature-based criteria determine the management of SCs. The high versatility and suitability of many WSpCs leads to a wider DOY_VImax range than for SCs, whose cultivation in cooler zones is somewhat more limited.

Interpreting the trends in DOY_Vimax (Figure 9C,D), we assume that farmers in drier areas are adapting to climate change by moving the usual cultivation patterns ahead of time to cope with later water scarcity or excessive heat. In more humid areas, where water availability is less limited, farmers instead might change to higher yielding cultivars with longer growing cycles such as later maize maturity groups. A similarly positive DOY_VImax trend can be expected for areas where irrigation is

intensified, leading to longer growing periods and, therefore, later NDVI peaks [53] as evidenced in Figure 9D for Andalusia, Spain [54,55].

5. Conclusions

The aim of this work was to map crop group-specific pure pixels for the time period 2001–2017, as, currently, there is no any available crop-specific geo-localized product at the European level. The results presented here are a first step in filling this gap and provide the basis for further inter-annual crop group-specific analyses. The main result was a timeseries of crop group-specific pixels of high purity. In addition, as a first application of the new data, crop group-specific phenological data were identified and published. The data can be made available to the research community upon reasonable request.

Our almost exclusively data-driven approach delivers information on crop groups across the continental extent with high robustness. The validations in several regions of Europe reveal reasonable pixel purities: by applying the S-2-based remote sensing validation approach, pixel purities for WSpCs range in all site-year combinations around 90% (and are frequently above this) but are, on average, lower for SCs. The Spanish validation site Castille and Leon (CyL), where a crop map-based validation was applied, provided more balanced values of around 73% and 72% purity for WSpCs and SCs, respectively. Despite some areas of lower performance (in particular the Baltics, for reasons described above), the data are considered of high value for numerous crop group-specific applications, and its quality is in line with existing research. Some class confusion was detected in more northern regions where classes were harder to separate due to the more similar phenological expression of WSpCs and SCs. This confusion occurred especially in colder regions, such as in the Baltic countries, where WSpCs show highly delayed crop development thereby approaching the SCs' temporal profiles. The algorithm offers a possibility to bypass problem areas by visual quality checking and then subsequent substitution of badly classified data by data based on models from neighboring regions of the same year.

Our phenological analyses showed a clear trend towards an earlier NDVI peak timing for WSpCs, which might be interpreted as some kind of response to climate change. For SCs, often irrigated, distinct trends were observed but differed depending on region.

The algorithm presented requires a very low level of parameterization. The method requires few a priori definitions, based on agronomic knowledge, and regional statistical data on typical absolute or relative crop acreages. No field data are required, and the approach relies mainly on the MODIS NDVI timeseries, from which decisive phenology metrics are extracted. This is of some value, since field data across such large extents in space and time are usually not available or very cumbersome and costly to collect.

There is no need for stratification or biogeoclimatic regionalization to account for climatic gradients. Analyses were instead applied on individual regions, which is highly advantageous.

The approach works on a pixel and regional level, where the region is defined by statistical territorial administrative units (NUTS). The regional scale is freely adjustable, assuming that crop acreage data are available for the chosen level of analysis. However, the analysis level represents a trade-off so as (i) not to run into problems with intra-unit climate heterogeneity and related phenology shifts while, (ii) at the same time, providing enough samples for the analysis.

The process runs in a semi-automatic way and can be extended to include coming years as well as other regions. For northern regions, an additional variable might need to be considered to help discriminate the otherwise overlapping NDVI peak timings of WSpCs and SCs. Only data in the red and infrared spectral range were used (MODIS), albeit with a high revisit rate, and this is therefore easily interchangeable with other sensors, similar indicators, or biophysical variables (e.g., fAPAR) of even finer spatial resolution, as they increasingly become available at higher temporal resolutions.

A known limitation is the classification depth, but the method could be complemented with satellite data of higher spatial and/or spectral resolution and an extended methodology to increase the number of classes. Indeed, such data and an additional processing step could be integrated on top of

this approach, taking advantage of a first crop group separation. For some applications, such as for inventories, the fact that this result is not providing a wall-to-wall classification but a selection of pure pixels of two crop groups presents a drawback. However, these results are considered an important first contribution, and further steps towards an exhaustive map can be undertaken.

The number of pixels is considered sufficient to provide a representative picture of the whole area of interest. As such, the methodology provides the basis for further research or modelling, especially for crop group-specific biomass or acreage estimation. A slight tendency towards samples on more homogeneous farmland was observed as well as a tendency towards more vigorous crops which, for specific aspects, might be considered a limitation. This could be problematic for regions with heterogeneous farm structures (i.e., where small- and large-scale structured agriculture is present in the same region) but might be alleviated by reducing the region size to achieve an increase in intra-unit homogeneity.

In summary, these results have the potential to contribute important information to ongoing research regarding crop group-specific analyses over large extents, in both retrospective and future directions. Applying a crop-specific parameterization, this research can contribute, for example, to crop biomass estimates, crop yield forecasts, or crop residue analyses. Obviously, for estimates of the ongoing crop season, for example, via statistical regressions to historic area-based yields, in-season estimates are needed, which we envisage as a next step, applying extensions to this algorithm. Finally, a historic series of pure pixels allows for an a posteriori cropland classification, of interest to, for example, soil carbon, soil erosion, and greenhouse gas modelers, as well as to agricultural landscape researchers touching on issues such as agricultural biodiversity. Another perspective, which takes advantage of this research, is the integration of the results presented here in the operational MARS Crop Yield Forecasting System (MCYFS), run by the Directorate General Joint Research Centre (DG JRC) of the European Commission.

References

1. Van der Velde, M.; Biavetti, I.; El-Aydam, M.; Niemeyer, S.; Santini, F.; van den Berg, M. Use and relevance of European Union crop monitoring and yield forecasts. *Agric. Syst.* **2019**, *168*, 224–230. [CrossRef]
2. López-Lozano, R.; Duveiller, G.; Seguini, L.; Meroni, M.; García-Condado, S.; Hooker, J.; Leo, O.; Baruth, B. Towards regional grain yield forecasting with 1km-resolution EO biophysical products: Strengths and limitations at pan-European level. *Agric. For. Meteorol.* **2015**, *206*, 12–32. [CrossRef]
3. Waldner, F.; Fritz, S.; Di Gregorio, A.; Defourny, P. Mapping Priorities to Focus Cropland Mapping Activities: Fitness Assessment of Existing Global, Regional and National Cropland Maps. *Remote Sens.* **2015**, *7*, 7959–7986. [CrossRef]
4. McNairn, H.; Champagne, C.; Shang, J.; Holmstrom, D.; Reichert, G. Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 434–449. [CrossRef]

5. Zhang, Y.; Chipanshi, A.; Daneshfar, B.; Koiter, L.; Champagne, C.; Davidson, A.; Reichert, G.; Bédard, F. Effect of using crop specific masks on earth observation based crop yield forecasting across Canada. *Remote Sens. Appl. Soc. Environ.* **2019**, *13*, 121–137. [CrossRef]

6. Fritz, S.; See, L.; Bayas, J.C.L.; Waldner, F.; Jacques, D.; Becker-Reshef, I.; Whitcraft, A.; Baruth, B.; Bonifacio, R.; Crutchfield, J.; et al. A comparison of global agricultural monitoring systems and current gaps. *Agric. Syst.* **2018**, *168*, 258–272. [CrossRef]

7. Anderson, W.; You, L.; Wood, S.; Wood-Sichra, U.; Wu, W. An analysis of methodological and spatial differences in global cropping systems models and maps: A comparative analysis of global cropping systems models. *Glob. Ecol. Biogeogr.* **2015**, *24*, 180–191. [CrossRef]

8. Atzberger, C.; Formaggio, A.R.; Shimabukuro, Y.E.; Udelhoven, T.; Mattiuzzi, M.; Sanchez, G.A.; Arai, E. Obtaining crop-specific time profiles of NDVI: The use of unmixing approaches for serving the continuity between SPOT-VGT and PROBA-V timeseries. *Int. J. Remote Sens.* **2014**, *35*, 2615–2638. [CrossRef]

9. Atzberger, C.; Rembold, F. Mapping the Spatial Distribution of Winter Crops at Sub-Pixel Level Using AVHRR NDVI Time Series and Neural Nets. *Remote Sens.* **2013**, *5*, 1335–1354. [CrossRef]

10. Chang, J.; Hansen, M.C.; Pittman, K.; Carroll, M.; DiMiceli, C. Corn and Soybean Mapping in the United States Using MODIS Time-Series Data Sets. *Agron. J.* **2007**, *99*, 1654–1664. [CrossRef]

11. Durgun, Ö.Y.; Gobin, A.; Van De Kerchove, R.; Tychon, B. Crop Area Mapping Using 100-m Proba-V Time Series. *Remote Sens.* **2016**, *8*, 585. [CrossRef]

12. Duveiller, G.; Baret, F.; Defourny, P. Remotely sensed green area index for winter wheat crop monitoring: 10-Year assessment at regional scale over a fragmented landscape. *Agric. For. Meteorol.* **2012**, *166–167*, 156–168. [CrossRef]

13. Ozdogan, M. The spatial distribution of crop types from MODIS data: Temporal unmixing using Independent Component Analysis. *Remote Sens. Environ.* **2010**, *114*, 1190–1204. [CrossRef]

14. Quarmby, N.A.; Townshend, J.R.G.; Settle, J.J.; White, K.H. Linear mixture modelling applied to AVHRR data for crop area estimation. *Int. J. Remote Sens.* **1992**, *13*, 415–425. [CrossRef]

15. Rembold, F.; Maselli, F. Estimation of Inter-Annual Crop Area Variation by the Application of Spectral Angle Mapping to Low Resolution Multitemporal NDVI Images. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 55–62. [CrossRef]

16. Boryan, C.; Yang, Z.; Mueller, R.; Craig, M. Monitoring US agriculture: The US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program. *Geocarto Int.* **2011**, *26*, 341–358. [CrossRef]

17. Massey, R.; Sankey, T.T.; Congalton, R.G.; Yadav, K.; Thenkabail, P.S.; Ozdogan, M.; Sanchez Meador, A.J. MODIS phenology-derived, multi-year distribution of conterminous U.S. crop types. *Remote Sens. Environ.* **2017**, *198*, 490–503. [CrossRef]

18. Sagris, V.; Wojda, P.; Milenov, P.; Devos, W. The harmonised data model for assessing Land Parcel Identification Systems compliance with requirements of direct aid and agri-environmental schemes of the CAP. *J. Environ. Manag.* **2013**, *118*, 40–48. [CrossRef]

19. Bertaglia, M.; Milenov, P.; Angileri, V.; Devos, W. *Cropland and Grassland Management Data Needs from Existing IACS Sources*; Publications Office of the European Union: Luxembourg, 2016; p. 63.

20. Wardlow, B.D.; Egbert, S.L.; Kastens, J.H. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sens. Environ.* **2007**, *108*, 290–310. [CrossRef]

21. Wardlow, B.D.; Egbert, S.L. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sens. Environ.* **2008**, *112*, 1096–1116. [CrossRef]

22. National Aeronautics and Space Administration (NASA): Moderate Resolution Imaging Spectroradiometer (MODIS). Available online: http://modis.gsfc.nasa.gov (accessed on 11 November 2019).

23. Vermote, E.; Wolfe, R. MOD09GQ MODIS/Terra Surface Reflectance Daily L2G Global 250 m SIN Grid V006 (Data Set). Available online: https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MOD09GQ/ (accessed on 11 November 2019).

24. Cerrani, I.; López Lozano, R. *Algorithm for the Dissagregation of Crop Area Statistics in the MARS Crop Yield Forecasting System*; European Commission, DG Joint Research Centre: Luxembourg, 2017; p. 21.

25. Copernicus Open Access Hub. Available online: https://scihub.copernicus.eu/dhus/ (accessed on 11 November 2019).

26. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W.; Harlan, J.C. *Monitoring the Vernal Advancement of Retrogradation of Natural Vegetation*; NASA/GSFC: Greenbelt, MD, USA, 1974.

27. MCSNCyL Castile and Leon Crops and Natural Land Map (MCSNCyL). Available online: http://www.mcsncyl.itacyl.es/en/inicio (accessed on 10 March 2018).

28. Del Blanco Medina, V.; Nafría García, D.A. Mapa de cultivos y superficies naturales de Castilla y León. Available online: https://web.archive.org/web/20161022230301/http://www.congreso2015aet.com/web/descargas/XVI_Congreso_AET_libro_actas_BAJA.pdf (accessed on 10 March 2018).

29. Paredes Gómez, V.; Del Blanco Medina, V.; Bengoa, J.L.; Nafría García, D.A. Accuracy assessment of a 122 classes land cover map based on Sentinel-2, Landsat 8, and Deimos-1 images and ancillary data. In Proceedings of the Observing, Understanding and Forecasting the Dynamics of Our Planet, Valencia, Spain, 22–27 July 2018.

30. EEA. Corine Land Cover (CLC) 2012. *Version 18.5.1*. Available online: https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012?tab=metadata (accessed on 10 March 2018).

31. Jarvis, A.; Reuter, H.I.; Nelson, A.; Guevara, E. *Hole-Filled SRTM for the Globe Version 4*; CGIAR: Montpellier, France, 2008.

32. Weissteiner, C.J.; García-Feced, C.; Paracchini, M.L. A new view on EU agricultural landscapes: Quantifying patchiness to assess farmland heterogeneity. *Ecol. Indic.* **2016**, *61*, 317–327. [CrossRef]

33. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]

34. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2015.

35. Vermote, E.F.; Roger, J.C.; Ray, J.P. MODIS Surface Reflectance User's Guide. Collection 6; Version 1.4. 2015. Available online: https://www.google.com.hk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=2ahUKEwjS29Wn397lAhUMyYsBHcEIDWwQFjABegQIBRAC&url=http%3A%2F%2Fmodis-sr.ltdri.org%2Fguide%2FMOD09_UserGuide_v1.4.pdf&usg=AOvVaw2vwPDJeHMQLn5blvGXA1VO (accessed on 1 May 2015).

36. Kaspar, F.; Zimmermann, K.; Polte-Rudolf, C. An overview of the phenological observation network and the phenological database of Germany's national meteorological service (Deutscher Wetterdienst). *Adv. Sci. Res.* **2015**, *11*, 93–99. [CrossRef]

37. Manfron, G.; Delmotte, S.; Busetto, L.; Hossard, L.; Ranghetti, L.; Brivio, P.A.; Boschetti, M. Estimating inter-annual variability in winter wheat sowing dates from satellite timeseries in Camargue, France. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *57*, 190–201. [CrossRef]

38. Pan, Z.; Huang, J.; Zhou, Q.; Wang, L.; Cheng, Y.; Zhang, H.; Blackburn, G.A.; Yan, J.; Liu, J. Mapping crop phenology using NDVI time-series derived from HJ-1 A/B data. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *34*, 188–197. [CrossRef]

39. Seo, B.; Lee, J.; Lee, K.-D.; Hong, S.; Kang, S. Improving remotely-sensed crop monitoring by NDVI-based crop phenology estimators for corn and soybeans in Iowa and Illinois, USA. *Field Crop. Res.* **2019**, *238*, 113–128. [CrossRef]

40. Sen, P.K. Estimates of the Regression Coefficient Based on Kendall's Tau. *J. Am. Stat. Assoc.* **1968**, *63*, 1379–1389. [CrossRef]

41. Kendall, M.G. *Rank Correlation Methods*, 4th ed.; Charles Griffin: London, UK, 1975.

42. Fraley, C.; Raftery, A.E.; Murphy, T.B.; Scrucca, L. *Mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation*; Department of Statistics, University of Washington: Seattle, WA, USA, 2012.

43. Skakun, S.; Franch, B.; Vermote, E.; Roger, J.-C.; Becker-Reshef, I.; Justice, C.; Kussul, N. Early season large-area winter crop mapping using MODIS NDVI data, growing degree days information and a Gaussian mixture model. *Remote Sens. Environ.* **2017**, *195*, 244–258. [CrossRef]

44. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum Likelihood from Incomplete Data via the EM Algorithm. *J. R. Stat. Soc. Ser. B (Methodol.)* **1977**, *39*, 1–38.

45. Rahman, H.; Dedieu, G. A simplified method for the atmospheric of satellite measurements in the solar spectrum. *Int. J. Remote Sens.* **1994**, *15*, 123–143. [CrossRef]

46. Richter, R.; Lous, J.; Müller-Wilm, U. Sentinel-2 MSI–Level 2A Products Algorithm Theoretical Basis Document. *ESA SP* **2012**, *46*, 1–72, Volume B (ATCOR), chapter 3. 2010.

47. Eilers, P.H. A perfect smoother. *Anal. Chem.* **2003**, *75*, 3631–3636. [CrossRef] [PubMed]

48. Whittaker, E.T. On a new method of graduation. *Proc. Edinb. Math. Soc.* **1923**, *41*, 63–75. [CrossRef]

49. Ren, S.; Qin, Q.; Ren, H. Contrasting wheat phenological responses to climate change in global scale. *Sci. Total Environ.* **2019**, *665*, 620–631. [CrossRef]

50. Waldner, F.; Lambert, M.-J.; Li, W.; Weiss, M.; Demarez, V.; Morin, D.; Marais-Sicre, C.; Hagolle, O.; Baret, F.; Defourny, P. Land Cover and Crop Type Classification along the Season Based on Biophysical Variables Retrieved from Multi-Sensor High-Resolution Time Series. *Remote Sens.* **2015**, *7*, 10400–10424. [CrossRef]

51. Duveiller, G.; Lopez-Lozano, R.; Cescatti, A. Exploiting the multi-angularity of the MODIS temporal signal to identify spatially homogeneous vegetation cover: A demonstration for agricultural monitoring applications. *Remote Sens. Environ.* **2015**, *166*, 61–77. [CrossRef]

52. SAPM. Survey on Agricultural Production Methods. Available online: https://ec.europa.eu/eurostat/statistics-explained/index.php/Survey_on_agricultural_production_methods (accessed on 23 August 2019).

53. Viña, A.; Gitelson, A.A.; Rundquist, D.C.; Keydan, G.P.; Leavitt, B.; Schepers, J. Monitoring Maize (*Zea mays* L.) Phenology with Remote Sensing. *Agron. J.* **2004**, *96*, 1139–1147.

54. Expósito, A.; Berbel, J. Agricultural Irrigation Water Use in a Closed Basin and the Impacts on Water Productivity: The Case of the Guadalquivir River Basin (Southern Spain). *Water* **2017**, *9*, 136. [CrossRef]

55. Sampedro Sánchez, D.; del Moral Ituarte, L. Tres décadas de política de aguas en Andalucía. Análisis de procesos y perspectiva territorial. *Cuadernos Geográficos* **2014**, *53*, 36–67.