



"Production of a Dynamic Cropland Mask by Processing Remote Sensing Image Series at High Temporal and Spatial Resolutions"

Valero, Silvia ; Morin, David ; Inglada, Jordi ; Sepulcre Canto, Guadalupe ; Arias, Marcela ; Hagolle, Olivier ; Dedieu, Gérard ; Bontemps, Sophie ; Defourny, Pierre ; Koetz, Benjamin

Abstract

The exploitation of new high revisit frequency satellite observations is an important opportunity for agricultural applications. The Sentinel-2 for Agriculture project S2Agri (<http://www.esa-sen2agri.org/SitePages/Home.aspx>) is designed to develop, demonstrate and facilitate the Sentinel-2 time series contribution to the satellite EO component of agriculture monitoring for many agricultural systems across the globe. In the framework of this project, this article studies the construction of a dynamic cropland mask. This mask consists of a binary "annual-cropland/no-annual-cropland" map produced several times during the season to serve as a mask for monitoring crop growing conditions over the growing season. The construction of the mask relies on two classical pattern recognition techniques: feature extraction and classification. One pixel- and two object-based strategies are proposed and compared. A set of 12 test sites are used to benchmark the methods and algorithms with regard ...

Document type : *Article de périodique (Journal article)*

Référence bibliographique

Valero, Silvia ; Morin, David ; Inglada, Jordi ; Sepulcre Canto, Guadalupe ; Arias, Marcela ; et. al. *Production of a Dynamic Cropland Mask by Processing Remote Sensing Image Series at High Temporal and Spatial Resolutions*. In: *Remote Sensing*, Vol. 8, no.55, p. 1-21 (11 January 2016)

DOI : 10.3390/rs8010055

Article

Production of a Dynamic Cropland Mask by Processing Remote Sensing Image Series at High Temporal and Spatial Resolutions

Silvia Valero ^{1,*}, David Morin ¹, Jordi Inglada ¹, Guadalupe Sepulcre ², Marcela Arias ¹, Olivier Hagolle ¹, Gérard Dedieu ¹, Sophie Bontemps ², Pierre Defourny ² and Benjamin Koetz ³

Received: 3 June 2015; Accepted: 16 December 2015; Published: 11 January 2016

Academic Editors: Anton Vrieling, Yoshio Inoue and Prasad S. Thenkabail

¹ CESBIO-CNES, CNRS (UMR 5126), IRD, Université de Toulouse, 31401 Toulouse Cedex 9, France; morind@cesbio.cnes.fr (D.M.); jordi.inglada@cesbio.eu (J.I.); marcela.arias@cesbio.cnes.fr (M.A.); olivier.hagolle@cnes.fr (O.H.); gerard.dedieu@cesbio.cnes.fr (G.D.)

² Earth and Life Institute, Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium; guadalupe.sepulcre@outlook.com (G.S.); sophie.bontemps@uclouvain.be (S.B.); pierre.defourny@uclouvain.be (P.D.)

³ ESRIN D/EOP-SEP, European Space Agency, Via Galileo Galilei, 00044 Frascati, Italy; Benjamin.Koetz@esa.int

* Correspondence: silvia.valero@cesbio.cnes.fr; Tel.: +33-561-55-85-19

Abstract: The exploitation of new high revisit frequency satellite observations is an important opportunity for agricultural applications. The Sentinel-2 for Agriculture project S2Agri (<http://www.esa-sen2agri.org/SitePages/Home.aspx>) is designed to develop, demonstrate and facilitate the Sentinel-2 time series contribution to the satellite EO component of agriculture monitoring for many agricultural systems across the globe. In the framework of this project, this article studies the construction of a dynamic cropland mask. This mask consists of a binary “annual-cropland/no-annual-cropland” map produced several times during the season to serve as a mask for monitoring crop growing conditions over the growing season. The construction of the mask relies on two classical pattern recognition techniques: feature extraction and classification. One pixel- and two object-based strategies are proposed and compared. A set of 12 test sites are used to benchmark the methods and algorithms with regard to the diversity of the agro-ecological context, landscape patterns, agricultural practices and actual satellite observation conditions. The classification results yield promising accuracies of around 90% at the end of the agricultural season. Efforts will be made to transition this research into operational products once Sentinel-2 data become available.

Keywords: cropland mapping; satellite image time series; Sentinel-2; dynamic classification; Random Forests

1. Introduction

Operational global agricultural monitoring products will play an important role in addressing some challenges that the world will have to face in the coming years. The importance food security increases the relevance of mapping the world’s agricultural landscapes [1]. At the global scale, monitoring food supply and demand is important for providing early warning on potential price spikes or food shortages. Accordingly, several international efforts aim at improving the timely mapping of crops grown, and to monitor crop growth.

Remote sensing satellite imaging has significantly contributed to the monitoring of agricultural areas [2]. Optical satellite images are a valuable resource for gathering information on crops over

large areas with a high revisit frequency. However, mapping crop areas through the analysis of multi-temporal satellite images is not a straightforward problem. A reliable global crop monitoring system needs satellite data acquired on a regular basis with broad coverage. Unfortunately, obtaining a series of cloud free images at high spatial resolution over an extended period of time is challenging [3]. Snow can also hamper earth surface observations during long periods of time.

Mapping crops across the great diversity of the world's landscapes is complex. Cropland areas can be composed of different crop types and field sizes [4,5]. Moreover, the spectral response of cropland areas varies significantly among the crop phenological stage and the cropping calendars applied around the world [6]. Individual crop types show variability in their spectral response because of climate and soil conditions, and differences between practices from country to country, such as the use of fertilizers or the human practices, are yet another contributing factor. In addition, cereals can easily be confused with natural grasslands at some periods of the cycle.

Satellite-based monitoring of agriculture started in the early 1970s as described by Becker-Reshef *et al.* [7]. In most local to regional scale mapping efforts, cropland has not been a single land cover class, but was contained within mosaic classes. This is typical for global land cover products, such as GLC2000 [8], GlobCover 2005/2009 [9,10], GLCShare [11], and MODIS Land Cover [12], which do not specifically target the agricultural component of the landscape. Even the most recent and more precise ESA Climate Change Initiative (CCI) Land Cover products [13] obtained from a multi-year multi-sensor approach still consider croplands as a land cover class. Cropland is a land use category rather than a type of land cover. Accordingly, a land cover typology is not well suited to the detection of cropland areas.

Furthermore, croplands are often included in mosaics or mixed classes, where crops are classified with natural vegetation such as grassland and shrubland. Therefore, the cropland mask obtained by using this mixed land cover class is not very useful in agricultural applications. Nevertheless, a few global crop maps have been produced recently [4,14], with an emphasis on water management: the global map of rainfed cropland areas (GMRCA) [15] and the global irrigated area map (GIAM) [16]. However, their coarse spatial resolution (10 km) does not meet the needs of operational applications and they suffer from large uncertainties [17] especially in complex farming systems in Africa. The notable limitations of cropland maps in Africa has given rise to the development of several studies and products such as the GEOLAND-2 SATCHMO product [18], which uses 10 × 10 km Landsat images. The goal of this product is to construct cropland maps at 30-m spatial resolution over many countries in order to deliver statistics on cropped areas at national scale.

Many programs have been set up by agricultural agencies to provide regular local and global agricultural monitoring [19]. One of the leading sources of global information on food production and food security is the Global Information and Early Warning System (GIEWS) established by the FAO [20]. In the same context, one of the oldest and longest running global crop monitoring activities is carried out by the USDA Foreign Agricultural Service (FAS) [21]. The Monitoring Agricultural ResourceS (MARS) [22] and Global Monitoring of Food Security (GMFS) [23] programs are operated by the European Commission, while the Institute of Remote Sensing and Digital Earth (RADI) within the Chinese Academy of Science is responsible for the Chinese Crop Watch Program [24].

One of the most important efforts in the last decade has been the Global Agricultural Monitoring (GLAM) Project [7] developed by NASA, UMD (University of Maryland) and USDA. Another important initiative has been the GEO Global Agricultural Monitoring initiative (GEOGLAM) [25], which aims to strengthen the international community's capacity to produce and disseminate relevant, timely and accurate agricultural information.

The global extent of cropland at 250 m spatial resolution is obtained by using multi-year MODIS (MODERate Resolution Imaging Spectroradiometer) time series and thermal data [14]. Since crop monitoring needs a high revisit frequency, NASA's MODIS sensor has been widely used [12,26–28] for the purpose. The use of this coarser sensor has proved suitable for identifying major crops over large areas of low diversity. However, its coarse resolution does not have the ability to provide crop

specific information with an adequate spatial resolution for small, heterogeneous croplands [29–31]. This drawback comes from subpixel heterogeneity, which is more obvious where individual fields are smaller than individual pixels [29]. Considering these limitations, new efforts are in progress to develop a Global Cropland Area Database at 30 m resolution (GCAD30) through Landsat and MODIS Data Fusion for the years 2010 and 1990 [32].

The upcoming Sentinel-2 mission offers new opportunities for agriculture monitoring on a regional to global scale. New possibilities for crop monitoring are offered by its 10–20 m spatial resolution, its 5-day revisit frequency, its global coverage and its compatibility with the Landsat missions. In this context, the Sentinel-2 for Agriculture (Sen2-Agri) project S2Agri has been recently launched by ESA with the aim of providing validated algorithms to derive Earth observation products that are relevant to crop monitoring for the agricultural user community at national and international level. Development and benchmarking of algorithms will lead to a range of products dedicated to agricultural monitoring [33]. Selected algorithms will be implemented in an open-source processing system, which will be validated and demonstrated in collaboration with national and international users.

The dynamic binary cropland mask (*i.e.*, crop/non-crop) that is introduced in this article is one example of these products being developed within the Sen2-Agri project for operational use. It distinguishes between annual cropland and other land cover classes. From the first dates of the season, the goal of this product is to forecast the cropland mask for the end of the season. Accordingly, the binary cropland mask is updated every time a new image is available during the agricultural season. At the end of the season, the cropland mask will contain the regions where at least one crop has been planted as the season progresses (not including grasslands, perennial crops or woody vegetation). Although the methodology may be used anywhere in the world, it requires local *in-situ* data and therefore a large amount of work to produce it at country—not to mention global—scale.

In the framework of the Sen2-Agri project, this work explores different processing chains in order to construct the dynamic cropland mask presented here. These systems rely on the two classic pattern recognition techniques of feature extraction and supervised classification.

Three different classification strategies are studied and proposed here in order to assess the differences between pixel- and object-based methodologies: Pixel-Based (PB), Object-Based with a Pre-filtering task (P-OB) and Object-Based with a Post-filtering task (OB-P). The first strategy studied is a pixel-wise classification system that does not consider information about the spatial structures of the image. In it, the classification map is constructed by taking the temporal properties of individual pixels into account. In contrast, the two object-based classification methodologies take the spatial relationship among pixels into account. For both object-based classification methodologies, the spatial information is included in the classification system by using a segmentation result of the original image. In the first case, the segmentation result is used to filter the image by using a pre-processing task. In the second, the segmentation result is used in a post-processing task, which applies a majority voting decision algorithm.

To evaluate the three classification strategies, a test phase to develop, tune and validate the generation of cropland mask products was performed at 12 sites around the world, in preparation for operational expansion to the global level once Sentinel-2 data are systematically acquired. These experiments rely on the use of SPOT4-Take5 data (<https://spot-take5.org>), complemented by LANDSAT 8 as a proxy for Sentinel-2 [34]. For test sites that were not covered by SPOT4-Take5, RapidEye images (<http://www.satimagingcorp.com/satellite-sensors/other-satellite-sensors/rapideye/>) were used. The test sites are described in detail in [33], and cover a range of agricultural systems intended to be representative of the global diversity of agricultural landscapes and satellite conditions. The sites are part of the Joint Experiment of Crop Assessment and Monitoring (JECAM) (<http://www.jecam.org>) and are distributed across the globe. The three proposed classification systems are described in Section 3. Experimental results are reported in Section 5 and conclusions are drawn in Section 6.

2. Sentinel-2 Proxy Data

Simulations of Sentinel-2 time series [34] were obtained by using SPOT4-Take5 (or RapidEye in the absence of SPOT4-Take5) images complemented by LANDSAT-8 data. A weighted linear interpolation technique was used to fill the gaps and all the data were resampled to the spatial resolution of 20 m.

The distribution of the 12 selected test sites all around the world ensured a coverage of a diversity of landscape patterns, different agricultural practices and several satellite observation conditions as the presence of clouds. The list of sites covered Europe (France, Belgium, Ukraine, Russia), Africa (Morocco, Madagascar, Burkina Faso, South Africa), Asia (China, Pakistan), North America (USA) and South America (Argentina). A detailed description of the test sites, demonstrating how they represent the diversity of cropping conditions, can be found in [33]. The cropland mask construction was evaluated at different time intervals (3 months, 6 months, 9 months and 12 months) during the season. For each test site, the interval of time was computed from the first available image. Consequently, as Figure 1 shows, the number of available images not fully cloudy (90% of clouds) for each interval of time was different.

		SPOT4 - Landsat 8									RapidEye - Landsat 8		
		Mo	Mad	Ar	Be	Ch	So	Uk	US	Fr	Ru	Pa	Bu
3 m		17	13	13	6	11	15	11	19	10	6	6	3
6 m		31	27	24	9	23	30	25	33	18	9	11	6
9 m		36	33	28	10	27	35	27	37	22			
12 m		40	36	31	11	28	37	28	41	23		14	9

Figure 1. Number of images with less than 90% cloud cover for the different Test Sites. Mo: Morocco, Mad: Madagascar, Ar: Argentina, Be: Belgium, Ch: China, So: South Africa, Uk: Ukraine, US: United States (Maricopa), Fr: France, Ru: Russia, Pa: Pakistan, Bu: Burkina Faso.

For the various test sites, the cloud coverage on the complete image time series was quite different as shown in Figure 2. Therefore, the number of images and the quality of the image time series was not homogeneous across the test sites. For instance, for the Argentina and Maricopa (USA) sites a large number of images was available, whereas for the Pakistan and Russian sites this was much less. The number of images of Burkina Faso was also small because no images were available during the rainy period.

		Percentage of images covered by clouds											
		Mo	Mad	Ar	Be	Ch	So	Uk	US	Fr	Ru	Pa	Bu
% of pixels		41.28	65.23	20.19	35.89	34.19	16.42	38.22	6.77	49.17	51.72	18.88	51.72

Figure 2. Percentage of images covered by clouds. Mo: Morocco, Mad: Madagascar, Ar: Argentina, Be: Belgium, Ch: China, So: South Africa, Uk: Ukraine, US: United States (Maricopa), Fr: France, Ru: Russia, Pa: Pakistan, Bu: Burkina Faso.

Besides the satellite imagery, *in-situ* data provided by site managers were used. All the test sites belonged to the JECAM network except for the US and Pakistan data set. JECAM is an initiative to reach a convergence of approaches, and develop monitoring and reporting protocols and best practices for a variety of global agricultural systems. The standardized field data JECAM protocol has been used to collect field data. It was composed of crop pixels (wheat, sunflower, maize, *etc.*) and no-crop pixels (urban areas, water surfaces, grassland, *etc.*). A single label (crop or no-crop) described the pixel class during the complete image time series (even if the crop appeared in the last months). Figure 3 shows the number of samples used in the benchmarking evaluation for all the test sites studied.

	Number of Samples for evaluating the different Test Sites											
	Mo	Mad	Ar	Be	Ch	So	Uk	US	Fr	Ru	Pa	Bu
Crop	28168	628	45248	1997559	6066	96320	228700	673008	113416	863290	558	1122
No Crop	1850731	798	33613	1138725	10875	21251	25954	1843987	201984	74484	859	331

Figure 3. Number of pixels from Sentinel proxy data, having a spatial resolution of 20 m, used in the validation of cropland masks. Mo: Morocco, Mad: Madagascar, Ar: Argentina, Be: Belgium, Ch: China, So: South Africa, Uk: Ukraine, US: United States (Maricopa), Fr: France, Ru: Russia, Pa: Pakistan, Bu: Burkina Faso.

Depending on the site, the number of samples composing the *in-situ* data varied widely. In particular, the number of samples in Madagascar, Pakistan and Burkina Faso was quite small.

3. Dynamic Global Cropland Mask Processing Chain

The crop mask consisted of a binary map separating annual cropland areas and other areas. Annual cropland was defined as a piece of land with a minimum area of 0.25 ha, actually sowed/planted and harvestable at least once within the year following the sowing date. This binary map was produced during the whole agricultural season, to serve as a mask for monitoring crop growing conditions. Its accuracy was expected to increase as long as additional images were integrated into the development process. To do this, the dynamic classification system shown in Figure 4 was proposed, in which the cropland area was extracted progressively after each real-time data acquisition captured at instant of time T . At the end of the season, the cropland mask contained the regions where at least one crop was planted during the year.

Figure 4 shows that two forms of input data are required in the classification system at each time T . The first input corresponds to a supervised classifier, which is trained by using an image time series acquired over the complete year preceding the season. This classification model is needed in order to progressively update the cropland mask as each new acquisition arrives. The second type of input data corresponds to the image time series acquired from the beginning of the season until the instant of time in question.

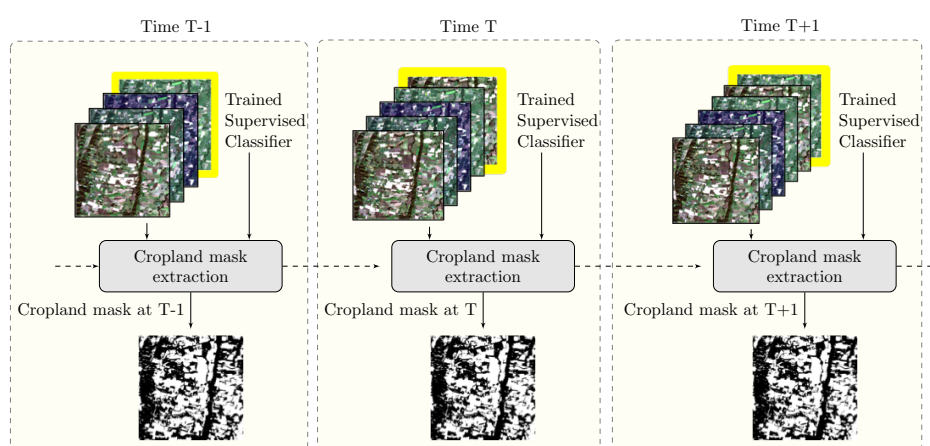


Figure 4. The proposed dynamic classification system.

The satellite data corresponded to a gap-filled image time series without missing data. In order to estimate the value of the missing pixels, a weighted linear interpolation was used. The no-data values were determined thanks to available missing value masks, which show pixels affected by clouds, cloud shadows or saturation effects. The weights used for the linear interpolation were computed by measuring the temporal linear distance between the interpolated pixel and the sampling valid pixels.

Using the flowchart described in Figure 4, three different strategies were proposed to perform the cropland mask extraction step: Pixel-Based (PB), Object-Based with a Pre-filtering task (P-OB) and Object-Based with a Post-filtering task (OB-P).

3.1. Pixel-Based Supervised Classification (PB)

For this first approach, the processing chain design is illustrated in Figure 5, which should be read from left to right. Three different steps, detailed in the following sections, compose this flowchart: the computation of several spectral indices, a feature extraction step and the classification step. The classification was performed by the Random Forest classifier [35]. The input data corresponded to the total number of patterns computed in the feature extraction step. In order to train the classifier, ground truth data containing crop and no-crop samples was mandatory in our system. In the framework of the Sen2-Agri Project [33], alternative, unsupervised approaches have been also studied [36], which do not rely on *in-situ* training data.

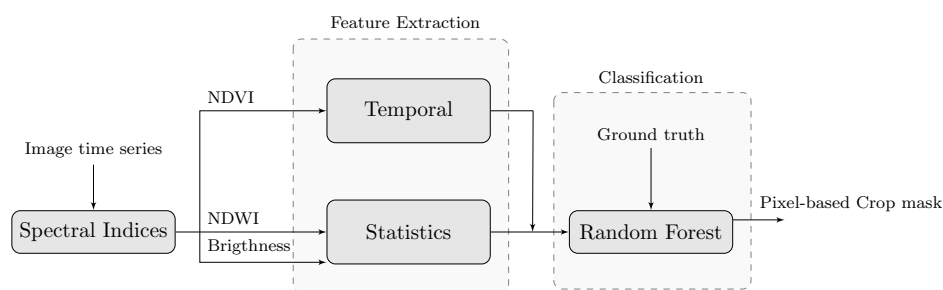


Figure 5. Supervised classification strategy for pixel-based approach.

3.1.1. Computation of the Spectral Indices

Three well-known spectral indices were used: The Normalized Difference Vegetation Index (NDVI) [37], the Normalized Difference Water Index (NDWI) [38] and the Brightness index, which is the modulus of the spectral vector computed as

$$\text{Brightness} = \sqrt{G^2 + R^2 + NIR^2 + SWIR^2} \quad (1)$$

where G , R , NIR and $SWIR$ are the green, red, near-infrared and shortwave infrared radiance values. The interest of using multi-temporal NDVI phenological profiles has been extensively confirmed in the context of crop identification [14,28]. The use of NDWI and Brightness provided complementary information improving the discrimination between crop and no-crop areas.

3.1.2. Feature Extraction

As input to the Random Forest classifier we extracted an input feature set from the image time series. The goal is to use these multi-year metrics as independent variables for the cropland mapping analysis. Metrics characterizing crops over a complete year, such as image time series reflectances or the Normalized Difference Vegetation Index (NDVI) [26,28,39], are traditionally used for crop monitoring. Specifically aimed at detecting crops at the global scale, a set of 39 multi-year MODIS metrics incorporating four MODIS land bands, NDVI (Normalized Difference Vegetation Index) and thermal data was introduced by Pittman *et al.* [14]. These NDVI metrics are computed on NDVI profiles calculated for specific crop classes during a complete year and they consist of the mean values of some reflectance bands or the mean NDVI value. The main problem of these features is that they do not take account of the fact that the NDVI profiles have specific shape due to the crop's phenological cycle. This shape can be modeled by a mathematical function such a double logistic function [40]. Following this crop model, crop NDVI profiles have strong transitions due to greenness and the onset of senescence, which are not usually found in other natural vegetation areas.

Considering the above drawbacks, our goal was to define features that could be computed dynamically as the satellite acquisitions became available. For instance, the construction of a global cropland mask implies the use of different satellite footprints that have been acquired at different (but close) dates. Also, the study of a large area implies that some crop characteristics, e.g., the peak of greenness, occur in different time periods. Accordingly, it was considered that the proposed crop features could not depend directly on the acquisition dates. Therefore, some features, such as the NDVI value of a pixel at a specific instant of time, were not considered as relevant. Also, the date of the maximum NDVI value was not considered as a discriminative feature since it could show strong variability in a given satellite footprint. Concerning the real-time constraint, the proposed features should be computed iteratively during the complete year without waiting for the end of the crop season.

Bearing this in mind, the computation of two different categories of features was proposed according to the computed spectral indices (as shown in Figure 5). In the case of NDVI, 17 temporal features characterizing the evolution of a crop profile were proposed, and five statistical features were computed for each of the other two spectral indices.

Table 1. NDVI Time Features proposed as an input feature set for Random Forest classifier.

Notation	Description
$NDVI_{max}$	Maximum value
$NDVI_{mean}$	Mean value
$NDVI_{std}$	Standard deviation value
$NDVI_{Dif-max}$	Maximum difference in a sliding temporal neighborhood having a size w (default value $w=2$)
$NDVI_{Dif-min}$	Minimum difference in a sliding temporal neighborhood having a size w (default value $w=2$).
$NDVI_{Dif-Dif}$	Difference between $NDVI_{Dif-max}$ and $NDVI_{Dif-min}$ value estimating the transition jump
$NDVI_{max}$	Maximum mean value in a sliding temporal neighborhood having a size w (default value $w=2$).
$NDVI_{max-Lg}$	Length of the flat interval containing the peak area associated to $NDVI_{max}$ value.
$NDVI_{max-Sur}$	Surface of the flat interval area containing the peak value associated to $NDVI_{max}$
$NDVI_{Pos-Sur}$	Maximum surface of the first positive derivative interval
$NDVI_{Pos-Lg}$	Length of the first positive derivative interval associated to $NDVI_{Pos-Sur}$ value
$NDVI_{Pos-Rt}$	Rate of the first positive derivative interval associated to $NDVI_{Pos-Sur}$
$NDVI_{Neg-Sur}$	Maximum surface of the first negative derivative interval
$NDVI_{Neg-Lg}$	Length of the first negative derivative interval associated to $NDVI_{Neg-Sur}$ value.
$NDVI_{Neg-Rt}$	Rate of senescence of the first negative derivative interval associated to $NDVI_{Neg-Sur}$
$NDVI_{Pos-Tr}$	Flag detecting if there is a bare soil transition before the positive derivative
$NDVI_{Neg-Tr}$	Flag detecting if there is a bare soil transition after the negative derivative

NDVI Time Features: Global patterns are often observed in crop NDVI signatures, describing the three most important crop stages: the onset of greenness, a maturity period and the onset of senescence. To exploit this information, phenological features extracted from NDVI time series have been proposed in the literature. One popular tool for extracting phenological features is the TIMESAT software [41]. Unfortunately, these features cannot be used with our classification system for several reasons. The first is that TIMESAT computes features using time series with regular temporal sampling (the time between two consecutive image acquisitions is the same). The second important limitation is that these features are computed on time series data representing at least one complete year. This requirement is mandatory since the features are computed after fitting the data with a well-known double logistic regular function describing a regular crop model. In our case, this could not be done since our goal was to extract crop features that could be computed dynamically from the first image acquisitions. For this reason, the temporal features presented in Table 1 related to the phenological cycle of the crops were used.

Brightness and NDWI statistical features: The aim of incorporating these features was to better discriminate the land cover classes not included in cropland areas. These features were based on the computation of the global statistics on the spectral index profiles. These features are presented in Tables 2 and 3.

Table 2. Brightness Statistical Features.

Notation	Description
BRIGHTNESS_{\max}	Maximum value
BRIGHTNESS_{\min}	Minimum value
$\text{BRIGHTNESS}_{\text{mean}}$	Mean value
$\text{BRIGHTNESS}_{\text{std}}$	Standard deviation value
$\text{BRIGHTNESS}_{\text{median}}$	Median value

Table 3. NDWI Statistical Features.

Notation	Description
NDWI_{\max}	Maximum value
NDWI_{\min}	Minimum value
$\text{NDWI}_{\text{mean}}$	Mean value
NDWI_{std}	Standard deviation value
$\text{NDWI}_{\text{median}}$	Median value

3.1.3. Supervised Classification: Random Forest

In the context of crop mapping, the supervised Decision Tree classifier used by Wardlow *et al.* [28] and Johnson *et al.* [42] has shown that MODIS-NDVI features can produce good accuracy levels. A very similar strategy is presented in Pittman *et al.* [14], who apply a set of global classification tree models using a bagging (Bootstrap AGGregatING) methodology [43]. Following the same philosophy, Random Forests (RF) were proposed by Breiman [35] in order to improve bagging results. The effectiveness of the RF classifier for the generation of Land Cover maps has been demonstrated by various studies [44–48].

3.2. Object-Based Supervised Classification with a Pre-Filtering Task (P-OB)

Pixel-based procedures analyze the information provided by each individual pixel, without taking the spatial or contextual information related to the pixel of interest into account. If the neighborhood information is not used, these classification techniques tend to produce salt and pepper noisy results that contain many wrongly classified pixels. In order to solve the limitations of local decisions taken by these methods, object-based image analysis techniques have been proposed [49–51].

Object-based image analysis focuses on analyzing images at the object level instead of working at the pixel level. Objects are image regions composed of a groups of pixels that are similar to one another according to a property measure. The fact that the local properties from neighboring pixels are considered explains the significant improvement that can be introduced by these methods. The general procedure for object-based image analysis is that first an image is segmented into homogenous regions according to a predefined criterion. Then, the image is classified by assigning each object to a class based on features and criteria set by the user.

Various studies have compared pixel and object-based image analysis aiming at classifying remote sensing images (Gao *et al.* [52] or Cleve *et al.* [53]). An important consideration for the evaluation of the two approaches is the image spatial resolution, which refers to the size of the smallest feature that can be detected. In the case of coarser spatial resolution images, object-based image analysis shows no great advantages over the pixel-based one. In general, object-based approaches are particularly suitable for high resolution images and lead to more robust and less noisy results. This can be easily understood since objects are not made up of several pixels in the case of low spatial resolution images.

Because low spatial resolution imagery has been used for most crop mapping exercises, and low spatial resolution imagery is not well-suited to object-based classification approaches.

However, some interesting object-based works have been proposed for cropland mapping, e.g., by Lobo *et al.* [49,54,55]

In the object-based classification method presented here, our first goal was to detect the agricultural field boundaries defining the spatial relationship of the image pixels. To do that, a segmentation strategy of the entire image was proposed. The construction of an image segmentation result is acknowledged to be a difficult task that depends on the final application. Here, the work focused on obtaining regions representing homogeneous agricultural areas having the same temporal features.

The resulting homogeneous regions were then used to reduce the intensity variation between one pixel and its surrounding neighborhood. This was carried out by a pre-processing task that corresponded to a mean filtering. Finally, the features obtained by applying the feature extraction task to the filtered image were used on the RF classification task. These different steps are shown in the flowchart of Figure 6, which can be divided in three steps: the construction of a segmentation of the original image, the pre-processing filtering and the feature extraction steps.

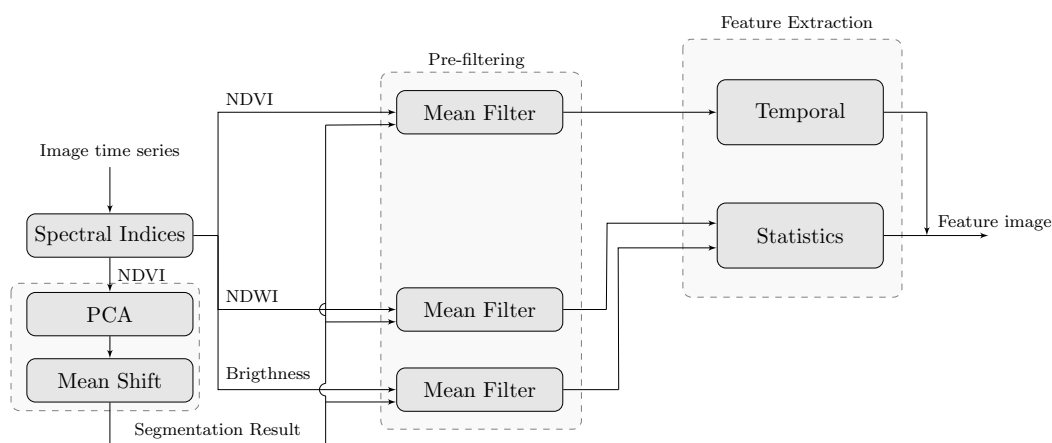


Figure 6. The proposed object-based classification system with a pre-filtering task.

The pre-processing filtering was done by replacing each pixel value with the mean value of its neighbors, including itself, as shown in Figure 6. The segmentation result was obtained by concatenating two classical techniques: Principal Component Analysis [56] and the Mean-Shift segmentation algorithm [57]. The input data required for these methods is the NDVI image time series, which was chosen given our crop monitoring application. The purpose was to group pixels representing agricultural areas having similar temporal NDVI profiles, which is very common for pixels belonging to the same crop field. In a similar context, the interest of using NDVI profiles to partition the image into coherent groups has already been shown [58,59].

Image Segmentation

Principal Components Analysis [56] is a multivariate statistical technique that can be used for data compression or time series evaluation. The work introduced by Eastman *et al.* (1993) [60] showed how the use of principal components could satisfactorily extract the information on temporal vegetation changes. In our case, PCA was computed in order to characterize important NDVI transitions on crop profiles.

The Mean Shift (MS) segmentation algorithm [57] is a nonparametric clustering technique, found in some classification works, such as [61]. In this work, the MS segmentation algorithm was applied to PCA space, obtained by applying the PCA transformation to the NDVI time series.

The algorithm requires the definition of three parameters: the spatial bandwidth h_s , the spectral bandwidth (or range information) h_r , and the minimum number of pixels in a region N_{min} (*i.e.*, regions containing less than N_{min} pixels will be eliminated and merged into the neighboring region).

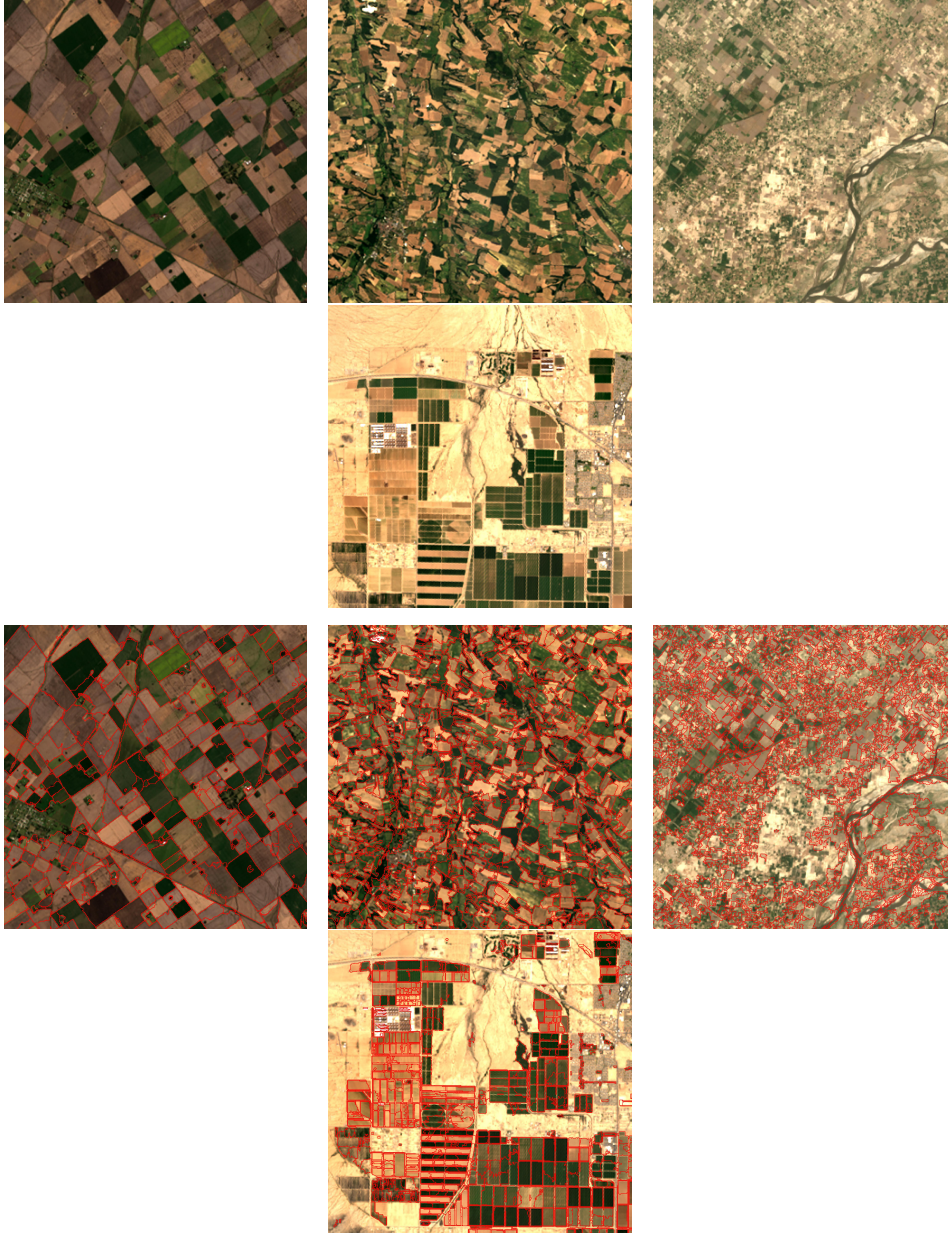


Figure 7. MS segmentation results obtained for different agricultural landscapes. The MS segmentation algorithm was applied to the six first principal components for all the sites. The same set of MS parameters was used for the four examples : $h_s = 10$, $h_r = 0.65$ and $N_{min} = 10$. From left to right, the first row contains the Argentina, France, Pakistan, Maricopa (US) true color RGB compositions. The second row contains the MS segmentation results obtained.

Figure 7 shows four MS segmentation results obtained with different agricultural landscapes. In the case of the Argentina test site, the cropland area is composed of large rectangular crop fields. In contrast, the agricultural landscape of France, shown in the second row, is composed of small crop fields. The third row shows the Pakistan test site where the landscape is composed of heterogeneous and textured areas. The last row shows Maricopa (US), which is composed of large, isolated crop

areas. The results obtained show that most of the crop fields were correctly detected as individual objects of the image.

3.3. Object-Based Supervised Classification with Post-Filtering Task (OB-P)

A different object-based approach was proposed here in order to reduce pixel-wise classification noise known as salt-and-pepper. The main difference compared to the last approach was that the spatial information was introduced as a post-filtering task as shown in Figure 8. The goal was to reduce the noise of the classification map obtained in Section 3.1 by applying a majority vote regularization. Compared to the pixel-wise classification system, two additional steps were included: first, the obtention of a segmentation result from the original image, which was performed by following the strategy presented in Section 3.2 and, second, a post-filtering task corresponding to majority vote filtering.

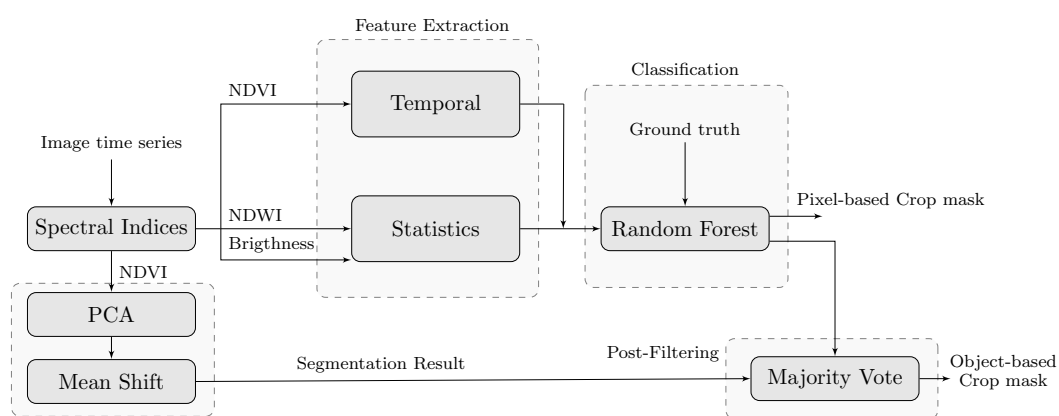


Figure 8. The proposed object-based classification system with a post-processing task.

The voting ensemble proposed here carried out a spatial regularization on the pixel-wise classification obtained by an RF pixel-wise classifier. The regularization process tried to improve the cropland mask results by making more homogeneous crop areas. The spatial neighborhoods defined by the segmentation result were used to delimitate the region to which the majority vote decision was applied. For each region, all the pixels voted for the class label, and the majority vote among the class labels was then assigned to all pixels of the region.

4. Validation Approach

The dynamic time cropland masks were computed on all the test sites by the three supervised classification systems proposed. To do this, the reference data was split into 2 disjoint subsets, the training set and the validation set. In order to avoid correlation between training and testing data as far as possible, two samples belonging to the same object were placed in the same subset. The first subset was composed of 1/3 of polygons and was used for training the RF classifier. Despite working with a binary classification, this subset contained the representation of all the possible classes (maize, grassland, water, etc.). The second subset contained the rest of the *in-situ* data, which was used for the validation of the algorithms.

To avoid highly correlated results concerning the splitting task of the *in-situ* data, 10 random splits were made for each site. The mean of the 10 experiments was then compared for the three different classification strategies. The supervised classifier learnt on the polygons not used in validation. From this data set, 1000 crop and 1000 no-crop samples were selected for the RF classifier training. In the case of Madagascar, Burkina Faso and Pakistan, the number of the training samples used was lower, given to the limited number of *in-situ* samples. Specifically, the numbers of training samples used for each class in Madagascar, Burkina Faso and Pakistan were 314, 165 and 279

respectively. Concerning the object-based approaches, the MS segmentation algorithm was applied to the first six principal components ($Nb_{PCA} = 6$) for all the sites. The same set of MS parameters was used: $h_s = 10$, $h_r = 0.65$ and $N_{min} = 10$.

Different classical measures were used to compare the binary cropland masks obtained at the different time intervals. The first quality measures used to evaluate crop and no-crop areas were the Precision and the Recall ratios. Precision was the ability of the classifier not to label as positive a sample that was negative. Recall was the ability of the classifier to find all the positive samples. In addition to these two ratios, Overall Accuracy (OA) [62] was computed in order to globally evaluate the fraction of crops and no-crops correctly classified. Figure 9 shows how these measures were computed.

		Predicted Class	
		Crop	No Crop
Actual Class	Crop	True Positive (TP)	False Positive (FP)
	No Crop	False Negative (FN)	True Negative (TN)

$$Recall = \frac{TP}{FN + TP}$$

$$Precision = \frac{TP}{FP + TP}$$

$$Overall\ Accuracy = \frac{TP + TN}{TP + TN + FN + FP}$$

Figure 9. On the left, the 2x2 confusion matrix displaying the number of correct and incorrect predictions made by the classifier. On the right, equations of the classical quality measures used in the following.

Finally, the Kappa coefficient [63] was also computed to globally evaluate the detection of crop and no-crop areas. This measure evaluated the relationship between beyond chance agreement and expected disagreement.

5. Experimental Results and Discussion

5.1. Cropland Area Validation

The three classification strategies evaluated were: Pixel-Based (PB), Object-Based with a Pre-filtering task (P-OB) and Object-Based with a Post-filtering task (OB-P). The results obtained are shown in Figure 10, where each table displays the results found for a specific test site.

Globally, croplands were detected with a high degree of accuracy. In most of the sites, the detection of the crop samples was higher than 90% at the end of the season. As expected, the detection of crops became better when the number of images increased along the agricultural season. The best results were obtained with the good quality data sets, like Maricopa or South Africa. In contrast, at sites where persistent cloud cover prohibited the collection of many good-quality images, like in Burkina Faso or Pakistan, the accuracy was significantly lower. Furthermore, the test sites that had been trained with a poor training data set due to limited fieldwork (Madagascar, Pakistan and Burkina Faso) obtained low accuracy for distinguishing crops. These low values can also be explained by the complex Madagascar, Pakistan and Burkina Faso landscapes composed of small fields.

A significantly low precision ratio can be observed for Morocco. In this case, a significant number of crop samples were detected as no-crop. This can be mainly explained by the analysis of the *in-situ* data. For this test site, one problem was that the presence of mixed crop fields was not negligible. Hence, these crop fields were detected as no-crop areas. Another problem of the Morocco test site

was that polygons describing the no-crop samples were very large and they could contain some false no-crop samples.

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	22.3	23.9	23.7	82.9	80.1	72.6
6 m	31.2	30.5	27.8	96.2	94.8	85.2
9 m	34.7	33.1	30.4	96.5	96.4	85
12 m	32.6	31.9	34.7	96.8	96.8	96.4

(a) Morocco

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	78.3	73.3	72.5	67.8	46.3	57.6
6 m	73.7	71.8	68.8	74.3	63.3	71.6
9 m	71.9	74.4	72.7	79.5	79.1	86.9
12 m	72.2	71.9	73.7	80.8	82.5	87.1

(b) Madagascar

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	91.1	69.7	92	89	89.7	90.2
6 m	90.8	74.3	90.9	88.3	89	89.8
9 m	90.5	76.6	91.2	90.4	90.7	91.3
12 m	89.8	81.6	90.5	92	91.8	92.4

(c) Argentina

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	93.3	91.8	92	94.3	95	94.6
6 m	96.8	97.3	97.6	95.3	93.1	93.2
9 m	97.2	97.7	97.8	95.2	92.8	93.1
12 m	97.4	97.9	97.9	95.1	92.7	93

(d) Belgium

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	78.7	57.3	87.3	93.9	79.8	96.3
6 m	93.2	74.9	94.3	93.9	90.2	93.8
9 m	92.9	74.3	93.2	95.2	88.8	94.6
12 m	92.7	88.2	93	95.3	88.8	94.5

(e) China

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	96.8	92	98.1	90.9	88.1	92.6
6 m	96.2	89.1	98.6	91.9	88.6	93.4
9 m	96.4	90.9	98.9	94	92.7	94.7
12 m	94.2	92.2	96.4	96.6	95.3	96.1

(f) South Africa

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	94.9	94.1	95.1	88	92.3	88.2
6 m	97.2	95.9	97.3	95.6	96.5	96.9
9 m	97.2	96.1	97.5	96.1	96.4	97.3
12 m	97.2	96	97.5	96.1	96.4	97.3

(g) Ukraine

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	94.1	93.7	93.9	66.4	68	65.9
6 m	96	95.1	96.9	91.9	91.7	90.4
9 m	96	95.4	95	95.4	94.5	94
12 m	95.2	94.7	96.3	96.6	95.4	94.9

(h) Maricopa

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	62.2	48.7	66.3	79.8	80.4	81.1
6 m	83.8	83.1	86.4	91.5	91.2	92.4
9 m	87.6	88.9	88.9	92	91.9	92.8
12 m	88.6	90.1	90.6	92.2	91.9	92.5

(i) France

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	94.4	93.8	94.4	84.5	95.1	89.3
6 m	97.7	97	97.9	89.3	92.2	86.1

(j) Russia

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	56	49.8	47	74.3	62	63.3
6 m	69.7	56.3	54.4	92.6	86.4	93.3
8 m	70.7	60.6	63.3	93.3	84.7	89.4

(k) Pakistan

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	87.2	84.6	87.3	71.3	73.8	75.9
6 m	92	90.3	91.5	82.3	83.4	82.4
8 m	92.4	89.8	91.6	85.1	85.9	86.4

(l) Burkina Faso

Figure 10. Precision and Recall evaluation for the cropland areas for the different data sets in 12 sites (a–l). Three methodologies are compared: Pixel-Based (PB), Object-Based with a Pre-filtering task (P-OB) and Object-Based with a Post-filtering task (OB-P)

Although the number of images was quite large, the results obtained for Madagascar were not very satisfactory because of the significant presence of clouds. Also, the crop fields on this test site were very small and the spatial resolution of 20 m was not good enough to detect accurate crop data. The good results for Belgium, despite the small number of images, came from the good quality of the ground truth data. Furthermore, the crop fields composing Belgium's landscape had good delimitation boundaries and homogeneous reflectances. The poor precision results of Pakistan can be explained by the small number of training samples, the poor quality of the complete data set and the lower image quality due to high atmosphere aerosol content. The significant complexity of the landscape (see Figure 7) explains why the results obtained by object-based approaches were poorer than the pixel-based results. In the case of Burkina Faso, object-based approaches did not really improve results obtained at pixel level. In this case, it was confirmed that the crop fields were usually very small and frequently covered by tree shadows areas. In the case of object-based approaches, the post-filtering task improved the results with respect to the approach using the pre-filtering step.

5.2. No-Crop Area Validation

The cropland area was evaluated following the same strategy as in the last section. The results obtained are shown in Figure 11. Looking at the evolution of the results with time, it can be seen that the results improved during the year. However, the Recall and Precision improvements achieved with increasing time were less significant than in the case of cropland areas. This was to be expected since no-crop samples do not change their labels during the agricultural season.

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	99.7	99.7	99.6	94.5	94.8	95.6
6 m	99.9	99.9	99.8	95.6	95.1	95.8
9 m	99.9	99.9	99.8	96.2	95.7	96.2
12 m	99.9	99.9	99.9	95.9	95.4	96

(a) Morocco

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	77.9	69.6	72.5	85.6	86.4	82.1
6 m	80.8	76.9	77	78.2	80.6	73.4
9 m	84.1	84.6	89.1	74.8	76.4	73.7
12 m	85.4	86.7	89.6	74.4	72.3	75.3

(b) Madagascar

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	85.8	67.4	87.3	88	41.8	89.2
6 m	85.1	71.8	86.8	87.7	54	87.8
9 m	87.4	78.6	88.5	87	59.3	88
12 m	89.2	83.9	89.6	85.9	70.1	86.8

(c) Argentina

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	89.9	90.7	90.1	88.1	85.1	85.5
6 m	91.9	88.7	89	94.5	95.5	96
9 m	91.8	88.4	88.8	95.2	96.2	94.6
12 m	91.7	88.5	88.7	95.5	94.6	96.6

(d) Belgium

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	96.1	57.8	97.6	83.5	47.6	88.3
6 m	96.6	91.1	96.5	96.2	70.6	96.8
9 m	97.3	85.2	97	95.9	70	96.1
12 m	97.4	93.8	96.9	95.8	93.1	96

(e) China

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	57.9	55.9	63.2	87.4	65.2	92.6
6 m	69.7	53	75.7	84.2	51.4	94
9 m	75.7	63.9	79.5	84.3	56.8	94.8
12 m	82.3	72.8	81.5	73.5	63.5	83.3

(f) South Africa

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	37.4	45.3	39.2	58.3	48.7	60.1
6 m	67.7	69.3	75.7	75.5	63.6	76.3
9 m	69.6	68.9	77	75.5	65.7	77.9
12 m	69.6	69.4	77	75.5	64.7	77.9

(g) Ukraine

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	88.6	89.1	88	98.3	98.1	97.8
6 m	97	96.9	96.3	98.5	98.1	98.8
9 m	98.3	97.9	97.7	98.4	98.2	97.6
12 m	98.7	98.2	98.1	98.1	97.9	98.6

(h) Maricopa

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	86.7	83	87.9	72.6	51.5	76.1
6 m	95	94.8	95.6	90.1	89.6	91.7
9 m	95.5	95.4	95.9	92.8	93.6	93.4
12 m	95.6	95.5	95.8	93.4	94.4	94.7

(i) France

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	20	41.9	24.2	41.3	26.6	38.7
6 m	39	46.3	34.3	76.3	66.5	79.2

(j) Russia

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	90.7	89.3	93.5	51.9	36.5	28.9
6 m	93.8	86	92.7	73.2	52.7	43.7
8 m	94.4	86	91.4	74.3	62	63.3

(k) Pakistan

	Precision			Recall		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	40.2	40.2	43.7	63.7	53.4	61.6
6 m	56.5	56.5	56.4	74.9	68.2	72.7
8 m	61.2	59.9	62.7	75.4	65.7	71.8

(l) Burkina Faso

Figure 11. Precision and Recall evaluation for the no-crop areas for the different data sets in 12 sites (a–l). Three methodologies are compared: Pixel-Based (PB), Object-Based with a Pre-filtering task (P-OB) and Object-Based with a Post-filtering task (OB-P)

At the end of the season, most test methods accurately identified no-crop areas. In general, the best results were obtained by PB and OB-P. OB-P tended to provide better results for the data sets having good/moderate data quality. The precision of no-crop areas in Russia was quite poor for all the quality measures. The small number of images covering only 6 months was not enough to characterize no-crop areas. In fact, all data sets composed of a small number of images led to significant difficulty in detecting no-crop areas. Similarly, the small number of training samples used in Madagascar, Pakistan and Burkina Faso explains the poor results obtained for these test sites.

	Kappa			OA		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	31.9	33.2	31.3	94.3	94.6	95.3
6 m	43.2	41.9	36.8	95.6	95	95.6
9 m	46.9	45.1	39.1	96.2	95.7	96.1
12 m	44.8	44	46.5	95.9	95.5	96

(a) Morocco

	Kappa			OA		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	54.1	33.8	40.8	78.1	70.4	72.2
6 m	52.2	44.9	44.7	76.4	74.1	72.9
9 m	53.7	55.9	59.6	77.4	78.8	80.1
12 m	54.6	54.6	61.5	77.7	77.9	81

(b) Madagascar

	Kappa			OA		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	76.8	32.2	79.2	88.6	69.2	89.8
6 m	75.7	43.7	77.4	88.1	73.9	57.4
9 m	77.6	51	79.3	89.1	77	89.9
12 m	78.4	62.9	79.5	89.5	82.2	90.1

(c) Argentina

	Kappa			OA		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	82.7	81.2	81	92.1	91.4	91.3
6 m	89.3	87.2	87.8	95	94	94.2
9 m	89.7	87.4	87.8	95.2	94	94.3
12 m	89.8	87.4	87.9	95.2	94.1	94.3

(d) Belgium

	Kappa			OA		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	74.4	29.9	83.4	87.2	59.5	91.2
6 m	89.9	60.3	90.7	95.4	77.6	95.7
9 m	90.6	58.8	90.5	95.7	77	95.6
12 m	90.6	81.7	90.2	95.6	91.5	95.5

(e) China

	Kappa			OA		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	60.7	46.9	67.7	86	81.9	88.6
6 m	70.1	39.3	79.7	90.5	81.8	93.5
9 m	74.6	50.8	83.1	92.2	86.3	94.8
12 m	72.3	60.9	77.8	92.2	89.4	93.8

(f) South Africa

	Kappa			OA		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	36.6	39	38.4	85	87.8	85.3
6 m	66.8	61.7	72	93.6	93.1	94.8
9 m	68.2	62.7	74	94	93.2	95.3
12 m	68.2	62.3	74	74	93.2	95.3

(g) Ukraine

	Kappa			OA		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	71.2	72.2	70	89.7	90	89
6 m	91.6	90.8	91.1	96.7	96.4	96.5
9 m	94.1	93	92	97.7	97.2	96.7
12 m	94.3	93.1	93.9	97.7	97.2	97.6

(h) Maricopa

	Kappa			OA		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	49.2	28.1	54.5	75.1	61.9	77.9
6 m	80	79.1	82.9	90.6	90.2	92
9 m	83.8	84.8	85.4	92.5	93	93.2
12 m	84.8	85.9	86.7	93.5	93	93.9

(i) France

	Kappa			OA		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	17.7	22	21.7	81	89.6	85.2
6 m	45.4	48.6	40.8	88.2	90.2	85.5

(j) Russia

	Kappa			OA		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	20.9	24.5	20.9	67.6	58.1	55.4
6 m	62.3	36.1	33.5	81	66.2	63.4
8 m	64	43.8	49.5	81.9	70.9	73.9

(k) Pakistan

	Kappa			OA		
	PB	P-OB	OB-P	PB	P-OB	OB-P
3 m	29	24.6	32.6	69.7	69.3	72.7
6 m	51	47.6	49.6	80.7	80	80.3
8 m	55.7	49.3	54.9	83	81.4	83.2

(l) Burkina Faso

Figure 12. Kappa and Overall Accuracy (OA) evaluation for the cropland masks for the different data sets in 12 sites (a–l). Three methodologies are compared: Pixel-Based (PB), Object-Based with a Pre-filtering task (P-OB) and Object-Based with a Post-filtering task (OB-P)

5.3. Global Evaluation

The Kappa and Overall Accuracy (OA) evaluations were performed in order to evaluate the cropland masks in a global way. The results are shown in Figure 12. Accuracies generally higher than 0.8 after 6 months and generally higher than 0.9 at the end of the year are observed. As expected, the test sites covered by a limited number of images and a small number of training samples show the poorest results. The poor Kappa value for the Morocco data set stemmed from the poor precision value obtained for the cropland areas. Despite the crop detection problem, a high OA value was obtained due to the unbalanced number of crop/no-crop samples shown in Figure 3. The good data sets, such as Maricopa, obtained very good results even after only the first three months.

The results obtained by PB and OB-P are quite similar. However, PB seems to give better numerical results especially for data sets having a small number of images, significant cloud cover and a small number of *in-situ* data samples. In contrast, looking at the cropland mask images of Figure 13, it can be seen that OB-P gives less noisy results. The visual evaluation confirms that OB-P has removed the small isolated pixels that do not belong to the cropland mask. Concerning the P-OB approach, numerical and visual results prove that this approach obtains the worst results.

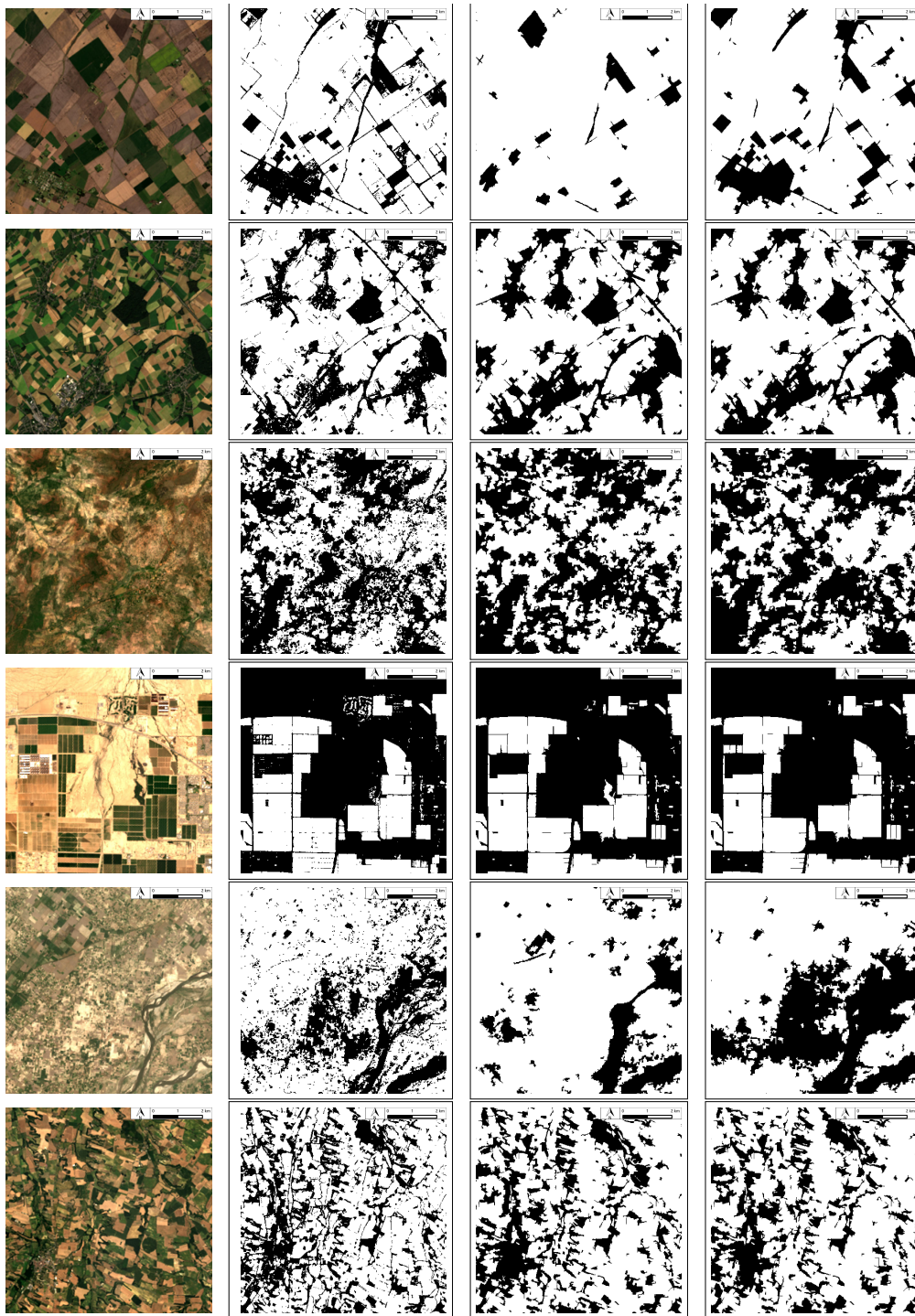


Figure 13. From top to bottom: Cropland mask results obtained after 12 months for test sites in Argentina, Belgium, BurkinaFaso, Maricopa, Pakistan, and France. From left to right: true color RGB composition, PB cropland results, P-OB cropland results and OB-P cropland results. Crops appear in white whereas no-crop areas appear in black.

6. Conclusions

In the framework of the Sen2-Agri project, this study aimed to develop a dynamic classification system that could be applied globally when enough *in-situ* data are available. The goal was to use Sentinel-2 proxy data to construct accurate maps of the extent of croplands in different

countries, which will work when Sentinel-2 data becomes operationally available. The cropland masks constructed contain plots of land with a minimum area of 0.25 ha, actually sowed/planted and harvestable at least once within the year following the sowing date (not including grasslands or woody vegetation). From the first dates of the season, the goal was to forecast the cropland mask for the end of the season.

Three different classification strategies were proposed and compared on 12 different test sites spread across the globe. Different agricultural landscapes were studied that were characterized by differences in cloud coverage, the presence of snow or the availability of *in-situ* data.

The classification approaches studied relied on the same feature extraction and supervised classification tasks. However, they presented different strategies in order to explore the advantage of incorporating the spatial information during the classification decision. Temporal vegetation and spectral indices were used in the feature extraction step to derive crop features.

The resulting set of crop features, computed dynamically during the agricultural season was proposed for the characterization of crop profiles. The set of features showed its effectiveness for crop classification. The supervised Random Forest classifier was used and it demonstrated its good performance. The spatial information was incorporated into two of the three proposed classification systems by using a segmentation result. The same methodology for partitioning the original image in order to look for similar cropland areas was proposed for all the sites studied.

The results clearly showed that the cropland extent accuracy increased when the number of images increased through the agricultural season. The real-time classification results yielded very promising accuracies (around 80% in the middle of the season) achieving an accuracy around 90% at the end of the season. As expected, the importance of the quantity and the quality of images to achieve real-time crop detection was confirmed. The impact of missing observations in some periods of the year was also found to be a serious limiting factor. Working with a supervised classification system also showed that the number of the *in-situ* data and their quality could have a significant effect on the classifier.

Comparisons of the different approaches showed that the object-based classification system using a pre-filtering task gave the worst results. In contrast, the pixel-based quantitative results were very similar to the results obtained by the Object-Based classification system using a Post-filtering task (OB-P). However, visual assessment of the cropland maps obtained showed less noisy results when the OB-P approach was used. One of the main limitations of the OB-P classification system was the detection of small crop fields, which was not always possible using the spatial resolution of 20 m. On the other hand, the results obtained by this last object-based methodology are expected to improve when real Sentinel-2 data is used. This is explained by the fact that the spatial resolution of SPOT4-Take5 and LANDSAT-8 (20 and 30 m), which were used to provide proxy data, is lower than future Sentinel-2 images.

Besides the spatial resolution, Sentinel-2 will also increase the temporal resolution and the number of spectral bands including the red-edge spectrum. Hence, future works will focus on the inclusion of these new sources of information in the OB-P classification system.

Acknowledgments: This work has been conducted within the framework of the Sentinel-2 for Agriculture (Sen2Agri) project, funded by the Data User Element of the European Space Agency. The Take5 initiative (SPOT4 and RapidEye) was sponsored by CNES (the French Space Agency) and ESA. The authors would like to thank the JECAM site managers who provided field data: Diego de Abelleira (Argentina), Aline Léonard, Damien Jacques (Belgium), Agnes Begué (Burkina Faso, Madagascar), Wu Bingfang, Miao Zhang, Zheng Yang (China), Claire Marais-Sicre, Jean-François Dejoux, Valérie Demarez (France), Michel Lepage (Morocco), Ijaz Ahmad Bhutta (Pakistan), Igor Savin (Russia), Terry Newby (South Africa), Nataliia Kussul, Andrii Shelestov (Ukraine) and Andrew French (USA).

Author Contributions: Silvia Valero and David Morin are the principal authors of this manuscript. Silvia Valero, David Morin, Jordi Inglada and Marcela Arias designed and implemented the processing chain and processed the data. Guadalupe Sepulcre, Olivier Hagolle, Gérard Dedieu, Sophie Bontemps and Pierre Defourny participated in discussions during the system design and evaluation and gave valuable methodological advice. Benjamin Koetz reviewed the system design and the results. All authors have been involved in the writing of the manuscript.

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

References

1. Thenkabail, P.S. Global Croplands and their Importance for Water and Food Security in the Twenty-first Century: Towards an Ever Green Revolution that Combines a Second Green Revolution with a Blue Revolution. *Remote Sens.* **2010**, *2*, 2305–2312.
2. Pinter, P.J.; Hatfield, J.L.; Schepers, J.S.; Barnes, E.M.; Moran, M.S.; Daughtry, C.S.T.; Upchurch, D.R. Remote sensing for crop management. *Photogram. Eng. Remote Sen.* **2003**, *69*, 647–664.
3. Whitcraft, A.K.; Vermote, E.F.; Becker-Reshef, I.; Justice, C.O. Cloud cover throughout the agricultural growing season: Impacts on passive optical earth observations. *Remote Sens. Environ.* **2015**, *156*, 438–447.
4. Fritz, S.; See, L.; McCallum, I.; You, L.; Bun, A.; Moltchanova, E.; Duerauer, M.; Albrecht, F.; Schill, C.; Perger, C.; *et al.* Mapping global cropland and field size. *Glob. Chang. Biol.* **2015**, doi:10.1111/gcb.12838.
5. Duveiller, G.; Defourny, P. A conceptual framework to define the spatial resolution requirements for agricultural monitoring using remote sensing. *Remote Sens. Environ.* **2010**, *114*, 2637–2650.
6. Whitcraft, A.K.; Becker-Reshef, I.; Justice, C.O. Agricultural growing season calendars derived from MODIS surface reflectance. *Int. J. Digit. Earth* **2014**, DOI:10.1080/17538947.2014.894147.
7. Becker-Reshef, I.; Justice, C.; Sullivan, M.; Vermote, E.; Tucker, C.; Anyamba, A.; Small, J.; Pak, E.; Masuoka, E.; Schmaltz, J.; *et al.* Monitoring Global Croplands with Coarse Resolution Earth Observations: The Global Agriculture Monitoring (GLAM) Project. *Remote Sens.* **2010**, *2*, 1589–1609.
8. Bartholomé, E.M.; Belward, A.S. GLC2000: A new approach to global land cover mapping from Earth Observation data. *Int. J. Remote Sens.* **2005**, *26*, 1959–1977.
9. Bontemps, S.; Defourny, P.; van Bogaert, E.; Arino, O.; Kalogirou, V.; Perez, J.R. *GlobCover 2009: Products Description and Validation Report*; European Spatial Agency and Université Catholique de Louvain: Frascati, Italy, 2011.
10. Defourny, P.; Bicheron, P.; Brockman, C.; Bontemps, S.; van Bogaert, E.; Vancutsem, C. The first 300 m global land cover map for 2005 using ENVISAT MERIS time series: A product of the GlobCover system. In Proceedings of the 33th International Symposium of Remote Sensing of Environment, Stresa, Italy, 4–8 May 2009.
11. Latham, J.; Cumani, R.; Rosati, I. Bloise, M. Global Land Cover SHARE (GLC-SHARE) Database Beta-Release Version 1.0. 2014. Available online: http://www.glc.org/downs/prj/glcshare/GLC_SHARE_beta_v1.0_2014.pdf (accessed on 31 May 2015).
12. Friedl, M.A.; Mclver, D.K.; Hodges, J.C.F.; Zhang, X.Y.; Muchoney, D.; Strahler, A.H.; Woodcock, C.E.; Gopal, S.; Schneider, A.; Cooper, A.; *et al.* Global land cover mapping from MODIS: Algorithms and early results. *Remote Sens. Environ.* **2002**, *83*, 287–302.
13. Bontemps, S.; Defourny, P.; Brockmann, C.; Herold, M.; Kalogirou, V.; Arino, O. New Global Land Cover mapping exercise in the framework of the ESA Climate Change Initiative. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 22–27 July 2012.
14. Pittman, K.; Hansen, M.C.; Becker-Reshef, I.; Potapov, P.V.; Justice, C.O. Estimating global cropland extent with multi-year MODIS data. *Remote Sens.* **2010**, *2*, 1844–1863.
15. Biradar, C.M.; Thenkabail, P.S.; Noojipady, P.; Li, Y.; Dheeravath, V.; Turrall, H.; Velpuri, M.; Gumma, M.K.; Gangalakunta, O.R.P.; Cai, X.L.; *et al.* A global map of rainfed cropland areas (GMRC) at the end of last millennium using remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11*, 114–129.
16. Thenkabail, P.S.; Biradar, C.M.; Noojipady, P.; Dheeravath, V.; Li, Y.J.; Velpuri, M.; Gumma, M.; Reddy, G.P.O.; Turrall, H.; Cai, X.L. A Global irrigated area map (GIAM) using remote sensing at the end of the last millennium. *Int. J. Remote Sens.* **2009**, *30*, 3679–3733.
17. Portmann, F.T.; Siebert, S.; Döll, P. MIRCA2000-Global Monthly Irrigated and Rainfed Crop Areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Glob. Biogeochem. Cycles* **2010**, doi:10.1029/2008GB003435.

18. Brink, A.B.; Bodart, C.; Brodsky, L.; Defourney, P.; Ernst, C.; Donney, F.; Lupi, A.; Tuckova; K. Anthropogenic pressure in East Africa monitoring 20 years of land cover changes by means of medium resolution satellite data. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *28*, 60–69.
19. Atzberger, C. Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. *Remote Sens.* **2013**, *5*, 949–981.
20. Global Information and Early Warning System (GIEWS) on Food and Agriculture. Food and Agriculture Organization (FAO). Available online: <http://www.fao.org/giews/english/index.htm> (accessed on 18 December 2015).
21. United States Department of Agriculture (USDA). Foreign Agricultural Service (FAS). Available online: <http://www.fas.usda.gov> (accessed on 31 May 2015).
22. Baruth, B.; Royer, A.; Genovese, G.; Klisch, A. The use of remote sensing within the MARS crop yield monitoring system of the European Commission. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *36*, 935–941.
23. Global Monitoring of Food Security (GMFS) Program of the European Union. Available online: <http://www.gmfs.info> (accessed on 18 December 2015).
24. Wu, B.; Gommès, R.; Zhang, M.; Zeng, H.; Yan, N.; Zou, W.; Zheng, Y.; Zhang, N.; Chang, S.; Xing, Q.; Van der Heijden, A. Global Crop Monitoring: A Satellite-Based Hierarchical Approach. *Remote Sens.* **2015**, *7*, 3907–3933.
25. Becker-Reshef, I.; Justice, C.; Doorn, B.; Reynolds, C.; Anyamba, A.; Tucker, C.J.; Korontzi, S. NASA's contribution to the Group on Earth Observations (GEO) Global Agricultural Monitoring System of Systems. *NASA Earth Obs.* **2009**, *21*, 24–29.
26. Wardlow, B.; Egbert, S.; Kastens, J. 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.
27. Shao, Y.; Lunetta, R.; Ediriwickrema, J.; Liames, J. Mapping cropland and major crop types across the Great Lakes Basin using MODIS-NDVI data. *Photogramm. Eng. Remote Sens.* **2010**, *75*, 73–84.
28. 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.
29. Löw, F.; Duveiller, G. Defining the spatial resolution requirements for crop identification using optical remote sensing. *Remote Sens.* **2014**, *6*, 9034–9063.
30. Leroux, L.; Jolivot, A.; Bégué, A.; Seen, D.L.; Zoungrana, B. How Reliable is the MODIS Land Cover Product for Crop Mapping Sub-Saharan Agricultural Landscapes. *Remote Sens.* **2014**, *6*, 8541–8564.
31. Hannerz, F.; Lotsch, A. *Assessment of Land Use and Cropland Inventories for Africa*; Centre for Environmental Economics and Policy in Africa, University of Pretoria: Pretoria, South Africa, 2006.
32. Global Cropland Area Database at 30 m Resolution (GCAD30). Available online: <https://earthdata.nasa.gov/our-community/community-data-system-programs/measures-projects/global-cropland-area-database> (accessed on 18 December 2015).
33. Bontemps, S.; Arias M.; Cara, C.; Dedieu, G.; Guzzonato, E.; Hagolle, O.; Inglada, J.; Matton, N.; Morin, D.; Popescu, R.; *et al.* Building a data set over 12 globally distributed sites to support the development of agriculture monitoring applications with Sentinel-2. *Remote Sens.* **2015**, *7*, 16062–16090, doi:10.3390/rs71215815.
34. Hagolle, O.; Sylvander, S.; Huc, M.; Claverie, M.; Clesse, D.; Dechoz, C.; Lonjou, V.; Poulain, V. SPOT-4 (Take5): Simulation of Sentinel-2 time series on 45 large sites. *Remote Sens.* **2015**, *7*, 12242–12264, doi:10.3390/rs70912242.
35. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *40*, 5–32.
36. Matton, N.; Sepulcre Canto, G.; Waldner, F.; Valero, S.; Morin, D.; Inglada, J.; Arias, M.; Bontemps, S.; Koetz, B.; Defourny, P. An Automated Method for Annual Cropland Mapping along the Season for Various Globally-Distributed Agrosystems Using High Spatial and Temporal Resolution Time Series. *Remote Sens.* **2015**, *7*, 13208–13232.
37. Rouse, J.W.; Hass, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. In *Third Earth Resources Technology Satellite-1 Symposium- Volume I: Technical Presentations*; NASA: Washington, DC, USA, 1973; pp. 309–317, NASA SP-351.
38. Gao, B.-C. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* **1996**, *58*, 257–266.

39. Jakubauskas, M.E.; Legates, D.R.; Kastens, J.H. Crop identification using harmonic analysis of time-series AVHRR NDVI data. *Comput. Electr. Agric.* **2002**, *37*, 127–139.
40. Zhang, X.; Friedl, M.A.; Schaaf, C.B.; Strahler, A.H.; Hodges, J.G.F.; Gao, F.; Reed, B.C.; Huete, A.R. Monitoring vegetation phenology using MODIS. *Remote Sens. Environ.* **2003**, *84*, 471–475.
41. Jonsson, P.; Eklundh, L. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 1824–1832.
42. Johnson, D.M.; Mueller, R. The 2009 cropland data layer. *Photogramm. Eng. Remote Sens.* **2010**, *76*, 1201–1205.
43. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140.
44. Gislason, P.O.; Benediktsson, J.A.; Sveinsson, J.R. Random Forests for land cover classification. *Pattern Recog. Lett.* **2006**, *27*, 294–300.
45. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104.
46. Long, J.A.; Lawrence, R.L.; Greenwood, M.C.; Marshall, L.; Miller, P.R. Object-oriented crop classification using multitemporal ETM+ SLC-off imagery and random forest. *GISci. Remote Sens.* **2013**, *50*, 418–436.
47. Inglada, J.; Tardy, B.; Valero, S.; Arias, M.; Hagolle, O.; Dedieu, G.; Bontemps, S.; Sepulcre G.; Defourny, P. Algorithm selection for the operational production of crop maps in the frame of the Sentinel-2 Agriculture project. In Proceedings of the 4th International symposium on Recent Advances in Quantitative Remote Sensing, Valencia, Spain, 22–26 September 2014.
48. Kamusoko, C.; Gamba, J.; Murakami, H. Mapping Woodland Cover in the Miombo Ecosystem: A Comparison of Machine Learning Classifiers. *Land* **2014**, *3*, 524–540.
49. Lobo, A.; Chic, O.; Casterad, A. Classification of Mediterranean crops with multisensor data: Per-pixel versus per-object statistics and image segmentation. *Int. J. Remote Sens.* **1996**, *17*, 2385–2400.
50. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16.
51. Hay, G.; Castilla, G. Object-based image analysis: Strengths, weakness, opportunities and threats. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2006**, XXXVI-part 6, 4–5.
52. Gao, Y.; Mas, J.F. A Comparison of the Performance of Pixel Based and Object Based Classifications over Images with Various Spatial Resolutions. *Online J. Earth Sci.* **2008**, *2*, 27–35.
53. Cleve, C.; Kelly, M.; Kearns, F.R.; Moritz, M. Classification of the wildland-urban interface: A comparison of pixel- and object-based classifications using high-resolution aerial photography. *Comput. Environ. Urban Syst.* **2008**, *32*, 317–326.
54. Peña-Barragán, J.M.; Ngugi, M.K.; Plant, R.E.; Six, J. Object based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sens. Environ.* **2011**, *115*, 1301–1316.
55. De Wit, A.J.W.; Clevers, J. Efficiency and accuracy of per-field classification for operational crop mapping. *Int. J. Remote Sens.* **2004**, *25*, 4091–4112.
56. Pearson, K. On lines and planes of closest fit to systems of points in space. *Phil. Mag.* **1901**, *2*, 559–572.
57. Comaniciu, D. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 603–619.
58. Lhermitte, S.; Verbesselt, J.; Jonckheere, I.; Nackaerts, K.; van Aardt, J.A.N.; Verstraeten, W.W.; Coppin, P. Hierarchical image segmentation based on similarity of NDVI time series. *Remote Sens. Environ.* **2008**, *112*, 506–521, doi:10.1016/j.rse.2007.05.018.
59. Kennedy, R.E.; Yang, Z.; Cohen, W.B. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. *Remote Sens. Environ.* **2010**, *114*, 2897–2910.
60. Eastman, J.R.; Fulk, M. Long sequence time series evaluation using standardized principal components. *Photogramm. Eng. Remote Sens.* **1993**, *59*, 1307–1312.
61. Petitjean, F.; Kurtz, C.; Passat, N.; Gançarski, P. Spatio-temporal reasoning for the classification of satellite image time series. *Pattern Recognit. Lett.* **2012**, *33*, 1805–1815.

62. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437.
63. Carletta, J. Assessing agreement on classification tasks: The kappa statistic. *Comput. Linguist.* **1996**, *22*, 249–254.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).