

Article

Scalable Parcel-Based Crop Identification Scheme Using Sentinel-2 Data Time-Series for the Monitoring of the Common Agricultural Policy

Vasileios Sitokonstantinou ^{1,*} , Ioannis Papoutsis ¹, Charalampos Kontoes ¹,
Alberto Lafarga Arnal ², Ana Pilar Armesto Andrés ² and José Angel Garraza Zurbano ² 

¹ Institute for Space Applications and Remote Sensing, National Observatory of Athens, I. Metaxa and Vas. Pavlou St, Penteli, 15236 Athens, Greece; ipapoutsis@noa.gr (I.P.); kontoes@noa.gr (C.K.)

² INTIA Tecnologías e Infraestructuras Agroalimentarias, Av. Serapio Huici, 22, 31610 Villava, Spain; alafarga@intiasa.es (A.L.A.); aarmesto@intiasa.es (A.P.A.A.); jgarraza@intiasa.es (J.A.G.Z.)

* Correspondence: vsito@noa.gr; Tel.: +30-2108-109-219

Received: 12 March 2018; Accepted: 7 June 2018; Published: 8 June 2018



Abstract: This work investigates a Sentinel-2 based crop identification methodology for the monitoring of the Common Agricultural Policy's (CAP) Cross Compliance (CC) and Greening obligations. In this regard, we implemented and evaluated a parcel-based supervised classification scheme to produce accurate crop type mapping in a smallholder agricultural zone in Navarra, Spain. The scheme makes use of supervised classifiers Support Vector Machines (SVMs) and Random Forest (RF) to discriminate among the various crop types, based on a large variable space of Sentinel-2 imagery and Vegetation Index (VI) time-series. The classifiers are separately applied at three different levels of crop nomenclature hierarchy, comparing their performance with respect to accuracy and execution time. SVM provides optimal performance and proves significantly superior to RF for the lowest level of the nomenclature, resulting in 0.87 Cohen's kappa coefficient. Experiments were carried out to assess the importance of input variables, where top contributors are the Near Infrared (NIR), vegetation red-edge, and Short-Wave Infrared (SWIR) multispectral bands, and the Normalized Difference Vegetation (NDVI) and Plant Senescence Reflectance (PSRI) indices, sensed during advanced crop phenology stages. The scheme is finally applied to a Landsat-8 OLI based equivalent variable space, offering 0.70 Cohen's kappa coefficient for the SVM classification, highlighting the superior performance of Sentinel-2 for this type of application. This is credited to Sentinel-2's spatial, spectral, and temporal characteristics.

Keywords: crop identification; support vector machines; random forest; common agricultural policy; Sentinel-2 MSI; Landsat-8 OLI; feature importance; multispectral image time-series

1. Introduction

In the next decades, the climate change and the estimated increase in the global population will have significant impact on the food sector [1]. In this context, increased agricultural productivity under environmentally friendly practices is an increasingly interesting topic and a top priority for the European Union, manifesting predominantly in the form of the Common Agricultural Policy (CAP) [2]. The administration and control of subsidy payments, associated with the relevant agrarian policies of the CAP, Cross Compliance (CC) and Greening, require systematic and timely knowledge of the agricultural land cover and land use [3]. Paying agencies are the Member State (MS) authorities responsible for the fair allocation of the CAP funds, and consequently for the realization of inspections to decide on the farmers' compliance to their CAP obligations. The farmers submit their subsidy

applications to their paying agency, declaring the location and the cultivated crop type for all owned parcels [4]. In turn, paying agencies are required to inspect at least 5% of declarations, via means of field visits and photointerpretation of Very High Resolution (VHR) and High Resolution (HR) satellite imagery (i.e., SPOT, Worldview, IKONOS) [4,5]. However, these methods are time-consuming, complex, and reliant on the skills of the inspector.

Contrariwise, automated remote sensing methods, suitably designed with respect to computational efficiency, large-scale coverage, and spatial resolution, can offer a monitoring alternative with reliable solutions [2]. In the context of the expected CAP 2020+ reform, there is a clear intention for simplification and a reduction of the administrative burden for both farmers and paying agencies. This study showcases how timely and accurate crop identification can assist the paying agency controls, suggesting an effective and efficient tool towards the required shift from sample inspections to large scale monitoring.

The freely available data and global coverage of Landsat satellites have established them, over the years, as one of the primary data sources for operational agriculture monitoring applications. Landsat data have been exploited in crop type mapping from as early as the 1980s, with Odenweller and Johnson (1984) using Vegetation Index (VI) time-series under threshold-based approaches, while in more recent studies their usage has evolved into non-parametric supervised classification schemes [2,6]. However, low (e.g., MODIS 250 m) and moderate (e.g., Landsat 30 m) spatial resolution data, with pixel sizes comparable to the parcel area, provide suboptimal thematic accuracy [7]. VHR (e.g., QuickBird and IKONOS) and HR (e.g., SPOT 5) satellite imagery has also been extensively used in crop classification studies [8,9]. VHR offers an excellent target to pixel ratio and allows for the appropriate exploitation of image processing features, such as texture, context, and tone [10]. Single-source missions that provide VHR imagery are often associated with high cost and long revisit times. However, the recently introduced small-satellite constellations, such as Planet's flock of Doves, can offer global coverage at high revisit frequency and unprecedentedly low costs. Nonetheless, the inexpensive, off-the-shelf sensor components used imply suboptimal noise resilience, radiometric performance, and cross-sensor uniformity [11].

Time-series of multispectral imagery that covers the agricultural cycle, from seeding to harvest, is necessitated for accurate crop identification [12]. Mature crops are more distinct and carry varying textures, water absorbency, and colors, having greater chances of being detected with remote sensing [13]. With respect to image processing, literature clearly points to object-based image analysis (OBIA) methods, assuming VHR and HR imagery is in use. Alternative pixel-based analyses often lead to misclassifications due to the canopy's spectral variability, bare soil background reflectance, and mixed pixels. Grouping pixels into spectrally consistent objects can alleviate such issues [10].

Non-parametric supervised classifiers such as Support Vector Machines (SVM), decision tree ensembles and artificial neural networks have been successfully employed in crop identification schemes [14,15]. In the last ten years, SVM and Random Forest (RF) classifiers have become increasingly popular in land cover mapping publications, doubling their appearances since 2012 [16]. Their success is attributed to effectively describing the nonlinear relationships between crops' spectral characteristics and their physical condition [17]. Both classifiers are resilient to noise and over-fitting and are therefore able to effectively cope with unbalanced data [16]. SVMs can ably handle small training datasets [18], while ensemble classifiers like RF are characterized by high computational efficiency [19]. These constitute ideal classification features for accurate and efficient crop type mapping and specifically for the purposes of the CAP monitoring, where data are big and ground truth information is scarce.

The Sentinel-2 mission introduces a paradigm shift in the quality and quantity of open access Earth Observation (EO) data, opening a new era for operational terrestrial monitoring systems, a fortiori in the agriculture sector. The mission offers unprecedented 10 m and 20 m spatial resolution data at a 5-day revisit time with twin satellites. This, however, introduces the notion of big data and the challenge of their efficient handling. EO data are freely and systematically received over very large

areas, which range beyond the limits of one region or a country, as they are nowadays required by the on-going and planned large scale initiatives and funded projects (e.g., GEOGLAM, GEO-CRADLE). In this regard, significant attention has been raised to Sentinels' potential for agriculture monitoring and particularly in the context of the CAP control scheme [20–23]. One indicative example would be the European Space Agency's (ESA) Sentinel-2 Agriculture project, aiming at the simplification of crop management. The system achieves an overall mapping accuracy of 85% for five major crop types, making up 75% of the regional agricultural zone [24]. Lebourgeois et al. (2017) have been among the first to use a scheme of mixed OBIA and RF approach for smallholding crop type classification, achieving 64.4% accuracy for 14 crop classes. In both publications, the authors tested their Sentinel specific methods, by simulating Sentinel-2 data using SPOT 5 and SPOT 4 (Take 5 experiment) as proxies. Lebourgeois et al. (2017) additionally incorporated Landsat-8 and VHR imagery, with the latter proving inconsequential in the classification process [7].

This study attempts to capitalize on the most promising research paradigms discussed above, with a clear operationalization potential. In this context, we seek to suggest, implement, and validate a crop mapping scheme of methods, enabling the identification of nine different crop types and thus provide information on farmer's compliance, with respect to their CAP requirements. A parcel-based Sentinel-2 MSI time-series approach is used for the construction of the feature space. The utilized imagery captures the growing stages of the various crop types in order to successfully discriminate between them. The local Land-Parcel Identification System (LPIS), which provides the geospatial input for crop delineation and local farmers' declarations, as part of their CAP subsidy applications, are employed for the object partitioning of the images and the supervised classifiers' training, respectively.

Thematic maps are produced for three different levels of crop nomenclature (crop type, crop family, and season of cultivation), employing separately a quadratic kernel SVM and a RF classifier. SVM and RF are two of the most popular non parametric classifiers in land cover mapping and their comparison has been an increasingly interesting topic in recent publications [16]. However, there is limited literature on their performance comparison under a Sentinel-2 based scheme, and even more under computational efficiency considerations that the Sentinel-induced big data shift demands. Differences between the two classifiers are evaluated in terms of execution time, classification accuracy, and relevance to the crop identification issue. Further, the size of the variable space and the importance of individual variables are quantified and accordingly evaluated. The proposed methodology is finally applied to Landsat-8 OLI and down-sampled Sentinel-2 MSI equivalent variable spaces, for the comparison of spatial, temporal, and spectral characteristics between the two sensors. Lastly, we performed a compliance check analysis for the CAP Crop Diversification requirement to showcase the scheme's capacity for effective decision making within the context of the control of CAP subsidies.

2. Data and Study Area

2.1. Site Description

The study area is located in the province of Navarra in northeastern Spain, as shown in Figure 1. It comprises of 9052 crop parcels and occupies approximately 215 km².

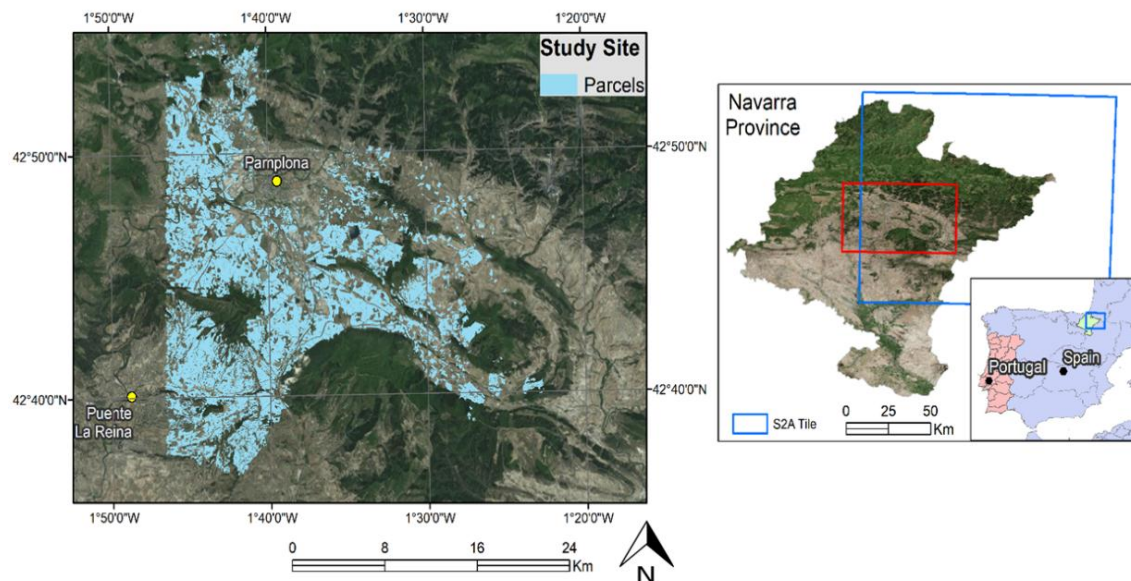


Figure 1. Study area located in the Spanish province of Navarra and the corresponding Sentinel-2A tile (S2A tile) covering it. The site refers to the agricultural zone surrounding the city of Pamplona.

The Pyrenees Mountains extend over the northern half of the province, as they stretch southward from France. Navarra's landscape is characterized by a mixed relief of watered valleys and forested mountains. Its agricultural yield is comparatively low compared to the rest of the country due to farm fragmentation; however, the fertilizer usage in the community is far greater than the national average [25]. In this study, crop type maps are produced for the nine main crop types found in the available dataset, and this entails soft wheat (9217 ha), barley (4835 ha), oats (1554 ha), corn (255 ha), sunflower (602 ha), vineyards (250 ha), broad beans (826 ha), and rapeseed (1000 ha). These refer to the lowest level of the nomenclature hierarchy, as shown in Figure 2, and explain nearly 90% of the total number of parcels (9052/10,274 parcels) that is found in the regional agricultural zone.

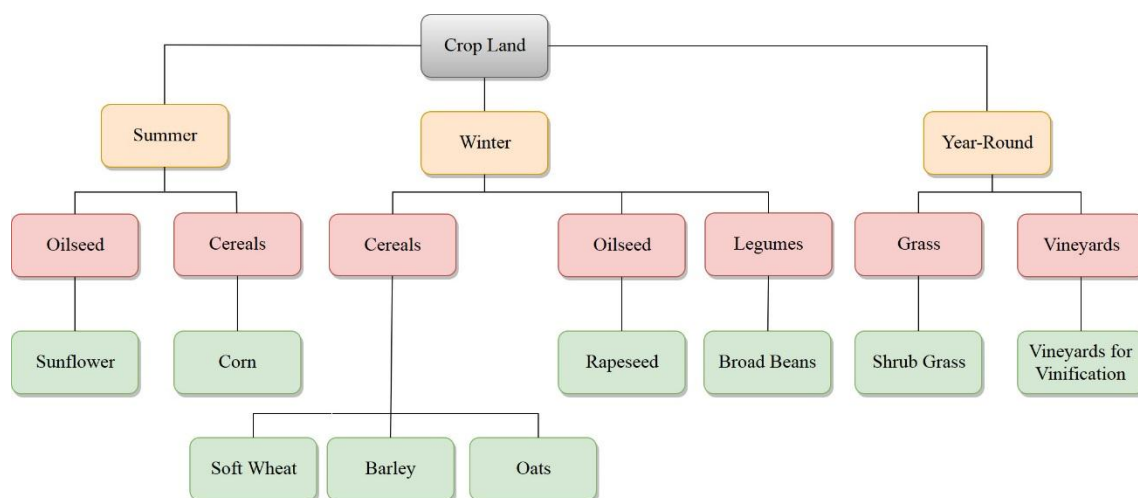


Figure 2. Crop type nomenclature adopted for Navarra, covering the nine main crop types that amount to almost 90% of parcels of the local agricultural zone. Crops are separated based on season, family, and type of cultivation.

The average parcel area in the study site is approximately 2 ha, which according to Lowder et al. (2016) belongs to the smallholding category [26]. This can be interpreted as 200 Sentinel-2 pixels

and less than 25 Landsat-8 pixels per parcel. The small average parcel size forms a challenging test site to showcase the capacity of methods and proves ideal for comparisons among sensors of different spatial resolution.

2.2. Datasets

The LPIS data used for the object partitioning of the time-series, along with the declared cultivated crop types, were made available by INTIA. INTIA is a public company, attached to the Department of Rural Development, Environment and Local Administration of Spain, which offers advanced services in the agri-food sector in Navarra. The provided dataset includes farmers' crop type declarations for 2016, family- and season-based grouping of the crop types, and the parcels' geospatial information in vector format. This unique dataset is provided by one of the direct end-users to the proposed scheme's application, which adds to the transparency of the process. INTIA additionally offered a timeline of the phenology stages for key Navarra crop types, which was instrumental in the appropriate selection of the image acquisition dates, but also for the logical assessment and interpretation of the results (Figure 3).

An analysis was performed (Supplementary Material) on the impact and feasibility of utilizing a dense time-series of Sentinel-2 imagery, regardless of the cloud coverage. The analysis showed that the usage of solely cloud-free imagery (less than 3% cloud coverage over the study site) that covers critical phenological stages was responsible for a near maximum crop type classification accuracy. On the other hand, the additional incorporation of cloudy imagery offered only a marginal increase in accuracy for a disproportionally larger cost in processing time. In this study, it was decided to use only cloud-free and nearly cloud-free imagery (<20% cloud coverage), which guaranteed an even temporal distribution of acquisitions and at the same time filtered out potentially redundant information. This way the data management and computational efficiency considerations, integral to the design of the proposed methodology, are appropriately showcased. This further supports the ultimate objective of the present study to be viewed in the context of an operational monitoring system for the control of farmers' compliance to their CAP obligations.

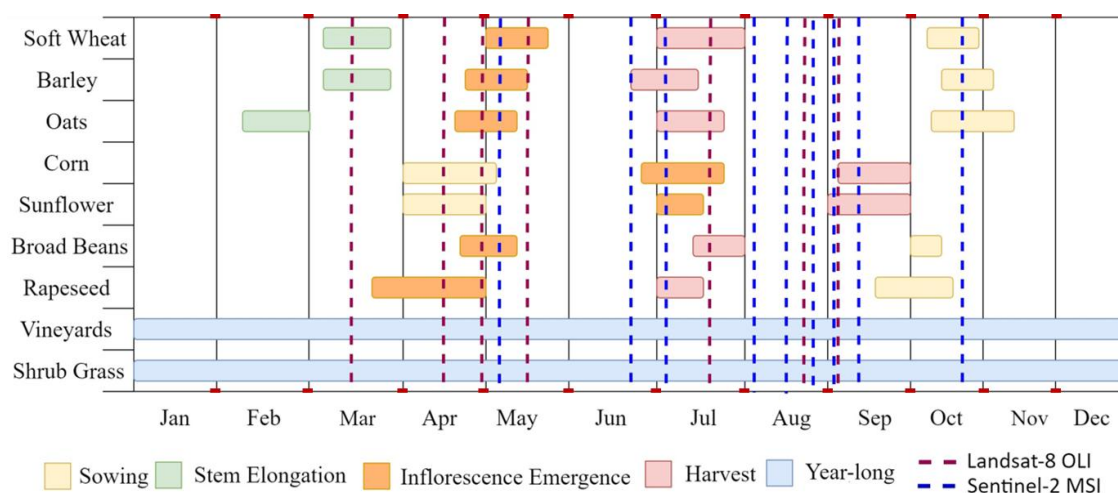


Figure 3. Timeline of phenology stages for key crops in Navarra, overlaid with the acquisition dates of Sentinel-2 images for 2016.

Sentinel-2 MSI's 10 m and 20 m resolution bands (B02–B08A & B11–B12) and Landsat-8 OLI's 30 m resolution bands (B02–B08), in the visible, Near Infrared (NIR), and Short-Wave Infrared (SWIR) parts of the spectrum, formed the Sentinel-based and Landsat-based feature spaces, respectively. Figure 3 depicts the nine selected cloud-free Sentinel-2A acquisitions (May to October) and the eight equivalent Landsat-8 scenes (March to September) for 2016.

3. Methods

Several machine learning algorithms of the supervised classification families of decisions trees, discriminant analysis, SVM, nearest neighbors, and tree ensembles were tested in a preliminary analysis. The analysis showed that the quadratic kernel SVM was the most accurate classifier, while RF came second best but with considerably improved computational efficiency. Hence, the implementation and comparison of these two classifiers was considered of great interest.

Crop type maps were produced via applying separately the SVM and RF classifiers to the Sentinel-2A MSI and Landsat-8 OLI imagery, under a parcel-based approach. An overview of the processing chain is shown in Figure 4, with a brief description of the steps in the following subsections.

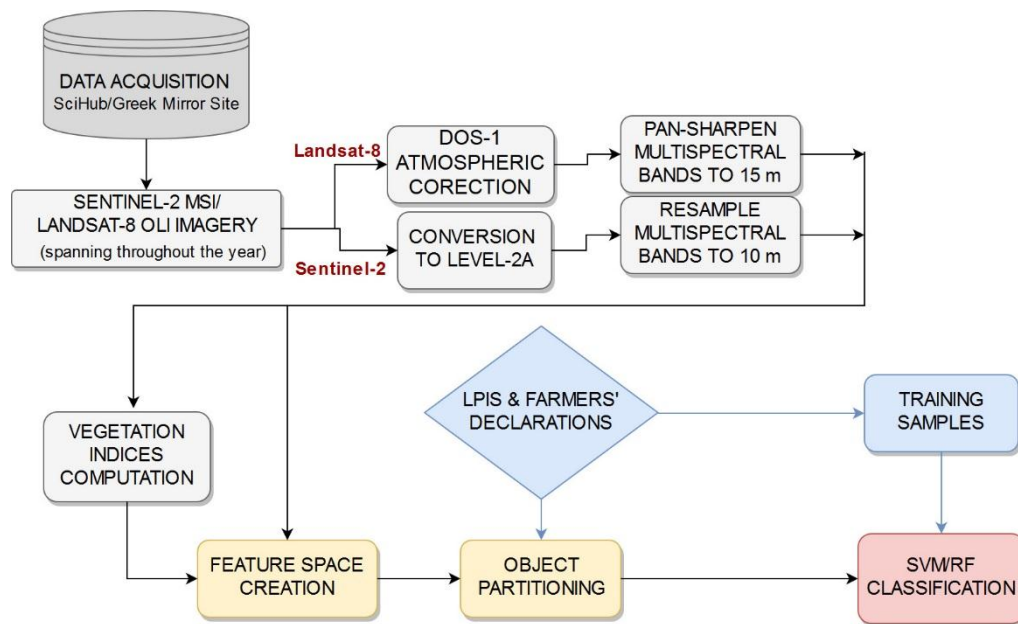


Figure 4. Flowchart of overall methodology: imagery acquisition, data pre-processing, feature space creation, crop type classification.

3.1. Pre-Processing

Sentinel-2A tiles were downloaded at Level 1C, which refers to $100 \times 100 \text{ km}^2$ ortho-images projected to cartographic geometry based on a Digital Elevation Model (DEM). Level 1C products are then transformed to Level-2A Bottom of Atmosphere (BOA) reflectances, using the Sen2Cor tool. Sentinel-2 bands of 20 m spatial resolution are resampled to 10 m spatial resolution, using the Nearest Neighbor (NN) resampling algorithm. Landsat-8 OLI imagery was acquired in GeoTiff format at Level 1TP, having been radiometrically calibrated and ortho-rectified. Image pixels have then been converted from digital numbers (DN) to radiances, to Top of Atmosphere (TOA) reflectances, using the Semi-Automatic Classification plugin in the geographic information system application QGIS [27]. DOS-1 atmospheric correction has been additionally applied, converting data to surface reflectances.

Pan-sharpening was then performed by merging the moderate resolution (30 m) multispectral data with the higher resolution panchromatic (15 m), providing multispectral imagery of higher resolution features. The employed technique implements a Brovey Transform, where the pan-sharpened values of the multispectral bands are calculated as in Equation (1) [27].

$$\text{MSpan} = \text{MS} \times \frac{\text{PAN}}{\text{I}} \quad (1)$$

where I is the intensity, while MS is the multispectral, and PAN is the respective panchromatic pixel values. Intensity weights are a function of the multispectral data and are defined as in Equation (2) [27].

$$I = \frac{0.42 \times \text{Blue} + 0.98 \times \text{Green} + 0.6 \times \text{Red}}{2} \quad (2)$$

3.2. Vegetation Indices

Vegetation indices (VIs) attempt to accentuate the vegetation signal, while diminishing soil background and solar irradiance contributions [28]. NDVI and Normalized Difference Water Index (NDWI) have been widely used in crop monitoring and crop mapping applications [7,29,30]. In this study, NDWI is used as defined by Gao (1996), utilizing the SWIR and NIR parts of the electromagnetic spectrum to correlate to plants' water content [31]. SWIR is sensitive to dry matter content, leaf structure, and water content, while NIR only accounts for the first two; hence the water content property is set apart. VIs can be fully exploited in multi-temporal schemes, since crop types' unique phenology calendars enable accurate crop discrimination. In this regard the Plant Senescence Reflectance Index (PSRI), which is defined as (Red-Green)/NIR, can form dissimilar spectral signatures for the different crop types, being particularly sensitive to their senescence phase [28].

3.3. Image Partitioning to Parcel Objects and Feature Space Creation

Image segmentation that makes use of the spatial, temporal, and spectral features of satellite imagery is a complex and computationally demanding process, requiring fine-tuning that depends on the region and involved crop types [24]. Consequently, the adoption of image segmentation for partitioning the image time-series into parcels would not have supported the considerations of operationalization, transferability, and scalability, which the proposed methodology aspires to respect. The study assumes the viewpoint of a paying agency that is in charge of the management and control of CAP payments. The LPIS is one essential computerized database that paying agencies use within their operations and are mandated to frequently update. It should be noted that the reliability of the LPIS data is dependent on the update frequency, given that the evolution of the field limits can be significant over the years. However, in an attempt to exploit the benefits of an object-based approach, while at the same time preserving the scheme's simplicity, it was decided that following a LPIS-based object partitioning approach is an acceptable trade off.

Table 1 lists the variables that make up the Landsat- and Sentinel-based datasets. The Sentinel-2 feature space, initially, comprises of all multispectral bands, for each individual scene employed (90 variables). In turn, VIs (NDVI, NDWI, and PSRI) for all temporal instances (27 variables) are calculated at pixel level and thereafter incorporated into the feature space (117 variables). In the same manner the Landsat-8 feature space includes a total of 63 variables, including the pan-sharpened multispectral data (42 variables) and the equivalent VIs (21 variables). Then the time-series of imagery is partitioned into parcels via utilizing the geometry of the LPIS vector data. Specifically, the pixel values that fall within the boundaries of the LPIS polygons are averaged, giving a single parcel value for every variable.

Table 1. Landsat-8 OLI and Sentinel-2 MSI variable spaces of multispectral band reflectances and vegetation indices; the pixels values are averaged for each parcel.

Type	Sentinel-2 MSI	Landsat-8 OLI
Reflectances	Bands: B02, B03, B04 (VIS), B08, B8A (NIR) of 10 m spatial resolution. B05, B06, B07 (Red-Edge) and B11, B12 (SWIR) of 20 m spatial resolution Mean values (per parcel) of 10 multispectral bands \times 9 image acquisitions (90 variables)	Bands: B02, B03, B04 (VIS), B05 (NIR), B06, B07 (SWIR) of 30 m spatial resolution. B08 (panchromatic) of 15 m spatial resolution Mean values (per parcel) of 6 pan-sharpened multispectral bands \times 7 image acquisitions (42 variables)
Vegetation Indices	Mean values (per parcel) of NDVI, PSRI, NDWI \times 9 image acquisitions (27 variables)	Mean values (per parcel) of NDVI, PSRI, NDWI \times 7 image acquisitions (21 variables)

3.4. Supervised Classification

We trained the classifiers based on a subset of the farmers' declarations. Validated ground truth data were available but limited in number, as such information can only be acquired via field inspections. Therefore, we have assumed that the farmers' declarations for 2016, as we received them, are valid in their majority. This is a fair assumption that is supported by INTIA inspectors, the workforce in charge of the field controls of the regional paying agency. In fact, cross-checking the validated data with the declarations showed that 99% of the farmers stated truthfully the cultivated crop type, out of 464 validated parcels.

The dataset was separated into randomly selected training and test sets. Multiple such splits were used in order to showcase the sensitivity of the models to the training data. Accordingly, the classification estimations were evaluated against the respective test subsets of the farmers' declarations. The number of training samples taken for each individual crop class was finalized at 20% of the total crop parcels of that class. When experimenting with samples higher than 20% for training, we encountered only a marginal increase in accuracy.

3.4.1. Support Vector Machines

Binary SVMs set a hyperplane by exploiting the feature information of each entity in the training set. This hyperplane would be the decision boundary in the classification process. Optimally, it is defined to maximize the distance between itself and the nearest training entity of any class (functional margin), with larger margins relating to lower generalization errors [18]. In this case, multiple binary classifiers for all different 36 class pairs (1-to-1 mapping) are combined to construct a single multi-class classifier.

Since the classification problem in this study is not linearly solvable, a quadratic kernel is used to transform the original feature space onto higher dimensions where the crop classes become linearly separable. However, the data are not perfectly separable and the algorithm allows for some misclassification in the training set, which is applied by the box constraint parameter (C). Specifically, higher box constraint values suggest stricter data separation. The model was built using the MATLAB function *fitcecoc*, where the kernel scale parameter was set to 'auto' mode for optimization. The box constraint parameter was selected via hyperparameter optimization, based on the minimization of the 10-fold cross-validation loss, as shown in Table 2. Thirty different objective evaluations were performed, of which an indicative subset is presented. The final selection was made under both cross-validation loss minimization and processing time considerations.

Table 2. Box constraint (C) parameter optimization using 10-fold cross-validation loss minimization.

# Evaluation	Box Constraint	10-Fold Loss (%)	Iteration Time (s)
1	1.048	5.97	35.29
2	736.8	8.54	286.29
3	46.36	7.29	177.32
4	220.22	8.32	689.06
5	998.40	8.34	450.62
6	0.0052	22.92	18.38
7	0.17	8.82	17.86

It is evident that a box constraint of approximately 1, as in the 1st Evaluation, provides the optimal 10-fold cross-validation loss, while it also proves superior in terms of processing efficiency, as compared to the evaluations of comparable cross-validation loss (i.e., Evaluations 2–5).

3.4.2. Random Forest

Ensemble classifiers, such as RF, increasingly gain popularity in EO applications using high spatial resolution imagery, as they can effectively manage large volumes of data [7]. RF starts with forming an

ensemble of standard decision trees, also known as weak learners. In a simple decision tree, the input entity is entered at the top and as it moves towards the bottom it gets grouped into progressively smaller subsets. The labels are then assigned to parcels based on the maximum number of votes among the ensemble of weak learners [19,32]. The RF model was built using the MATLAB function *fitensemble*, under the “Bag” method. Bagging refers to the repeated selection of random samples from the entirety of the training set, on which the weak learners train separately. Even though the estimates of a single decision tree would be noisy, the mean estimates over multiple decision trees are both unbiased and resilient to over-fitting [19]. In the same manner as for the SVM classification, the minimization of 10-fold cross-validation loss is used for the optimal hyperparameter tuning. Once again, thirty different objective evaluations were performed, of which an indicative subset is presented in Table 3.

Table 3. Number of weak learners (trees) and maximum number of splits optimization based on 10-fold cross-validation loss minimization.

# Evaluation	# of Weak Learners	Max # of Splits	10-Fold Loss (%)	Iteration Time (s)
1	479	1182	10.69	206.31
2	141	783	19.86	34.19
3	30	8925	11.12	14.70
4	781	2	33.87	11.81
5	11	3083	16.98	3.258

The hyperparameters “number of weak learners” and “maximum number of splits” were varied and evaluated against the minimization of the objective function. The third evaluation of 30 decision trees was selected as the optimal set of parameters. Nonetheless, it is observed that the first evaluation that uses 479 weak learners gives a lower 10-fold cross-validation loss, but this marginal increase in accuracy cannot excuse the additional processing effort required.

3.5. Accuracy Assessment

Classification performance was assessed according to producer’s accuracy (PA), user’s accuracy (UA), and Cohen’s kappa coefficient (K_c), computed as shown in Equations (3)–(5). PA is the ratio of correctly classified parcels over the total number of parcels for a ground truth class. Alternatively, UA is the ratio of correctly classified parcels for a given class to the total number of parcels predicted to belong to that class [33]. Statistical metric kappa is used to describe the overall classification accuracy and it is generally preferred over simple accuracy metrics, as it accounts for random agreement between truth and estimation [34]. Hence, K_c better showcases the capacity of both input data and classification techniques [35].

$$PA = \frac{n_{ii}}{n_{irow}} \quad (3)$$

$$UA = \frac{n_{ii}}{n_{icol}} \quad (4)$$

$$K_c = N \sum_{i=0}^r n_{ii} - \sum_{i=0}^r \frac{n_{icol} n_{irow}}{N^2} - \sum_{i=0}^r n_{icol} n_{irow} \quad (5)$$

where n_{ii} is the number of parcels correctly classified in a particular crop type class; N is the total number of parcels in the confusion matrix; r is the number of rows; n_{icol} and n_{irow} are the column (predicted class labels) and row (ground truth) total, respectively [36].

4. Results

4.1. Accuracy Performance

PA and UA results presented in Table 4 refer to the quadratic kernel SVM and RF classifications at the crop type nomenclature level. Since the dataset is strongly dominated by soft wheat parcels,

the overall accuracy would be respectively affected. Therefore, PA and UA better depict the classifiers performance for all individual crop classes. The values in Table 4 have been averaged over 20 iterations of random training sample splits.

Table 4. Producer’s accuracy and user’s accuracy for both Random Forest (RF) and Support Vector Machine (SVM) classifications for the crop type nomenclature level. The standard deviation of the mean accuracy values is also shown in parentheses.

Crop Type	RF PA (%)	RF UA (%)	SVM PA (%)	SVM UA (%)
Soft Wheat	94.89 (0.6)	82.88 (0.9)	95.60 (0.2)	91.49 (1.0)
Corn	96.35 (2.3)	93.29 (4.0)	92.56 (3.2)	96.48 (3.6)
Barley	81.81 (1.6)	87.78 (1.0)	90.10 (0.9)	90.43 (0.9)
Oats	44.38 (2.3)	90.36 (2.4)	78.57 (3.0)	89.59 (1.9)
Sunflower	85.69 (3.5)	92.03 (2.7)	85.36 (3.6)	95.45 (2.6)
Rapeseed	92.30 (1.7)	95.03 (1.4)	90.19 (1.8)	96.12 (0.7)
Broad Beans	85.41 (1.8)	92.41 (1.5)	86.83 (1.7)	95.08 (1.1)
Shrub Grass	77.93 (4.3)	76.90 (3.8)	84.07 (3.8)	83.48 (2.3)
Vineyards	86.15 (4.0)	83.72 (3.7)	87.52 (4.0)	83.48 (3.5)

Both classifiers provide excellent results with an overall accuracy 85.59% for RF and 91.30% for SVM. Nonetheless, according to the PAs, SVM considerably outperforms RF for the barley and oats ground truth classes. These are two of the most indicative classes as their proneness to misclassification can be physically interpreted and thus better describe SVM’s superiority. Barley, oats, and soft wheat are crops of the cereal family with strong physical and spectral resemblance. Even more, their similar cultivation calendars (Figure 3) minimize the discriminating qualities of the multi-temporal approach. The same holds true for shrub grass and vineyards, which are year-round crops, and thus their spectral profiles would not bear any solid temporal particularities. Shrub grass is also vegetation of a diverse nature, providing generic spectral signatures, and hence misclassifications are evident for both classifiers, particularly for RF. Finally, UA values identify SVM as the more reliable classifier, with very high accuracies for all crop types apart from shrub grass and vineyards. A clip of the SVM crop type classification map is shown in Figure 5, identifying both correctly and incorrectly crop type estimations.

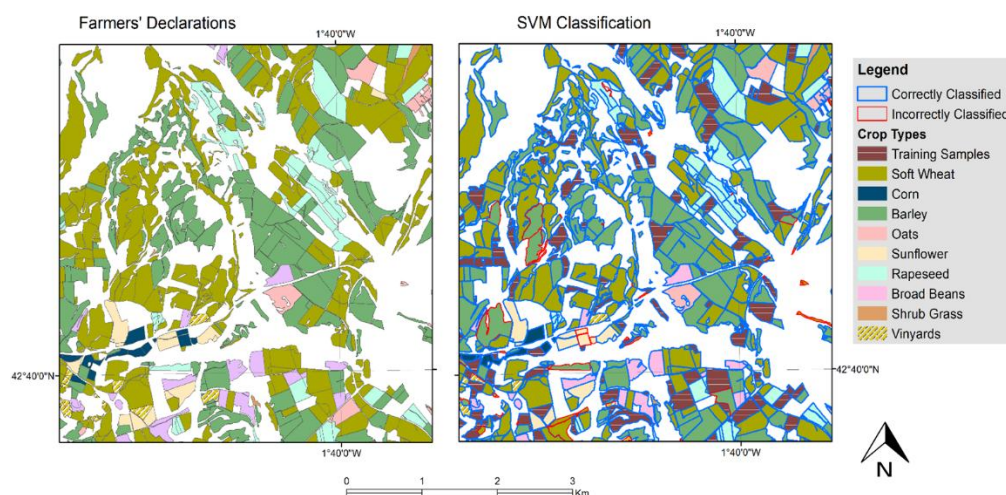


Figure 5. Farmers’ declarations and quadratic SVM classification at the lowest nomenclature level (crop type). Only a clipped part of the study site is shown, with correctly and incorrectly classified parcels also identified.

Performance differences between the two classifiers become even more apparent using K_c , as RF classification is subject to substantial random agreement. Maximum kappa coefficients of 0.87 and 0.78 for SVM and RF, respectively, showcase more reliably the dominance of the first classifier over the latter (Table 5). The listed accuracy values have been averaged over 20 classification iterations of randomly split training sets. The standard deviation of the accuracies for each classification scenario is also given. This is done to showcase the very low sensitivity of the overall accuracy results to the training set, with standard deviations varying in the range of 0.3–1%.

Table 5. Maximum Kappa coefficient values for all nomenclature levels; values are averaged over 20 classification iterations, with the standard deviation of the mean accuracy values (st.d.) also listed.

Nomenclature Level	Kappa RF	st.d. RF	Kappa SVM	st.d. SVM
Type	0.7834	0.0070	0.8721	0.0054
Family	0.8902	0.0030	0.9118	0.0053
Season	0.8856	0.0092	0.8729	0.0115

4.2. Feature Importance

Quantifying the importance of individual features of a large variable space enables the better understanding of the dataset itself. This would in turn provide the means for the physical interpretation of results and the proper evaluation of the input data. Importance weights for the entirety of the Sentinel-2 variable space are shown in Figure 6, with features being grouped into thematic divisions with respect to spectral characteristics and acquisition dates.

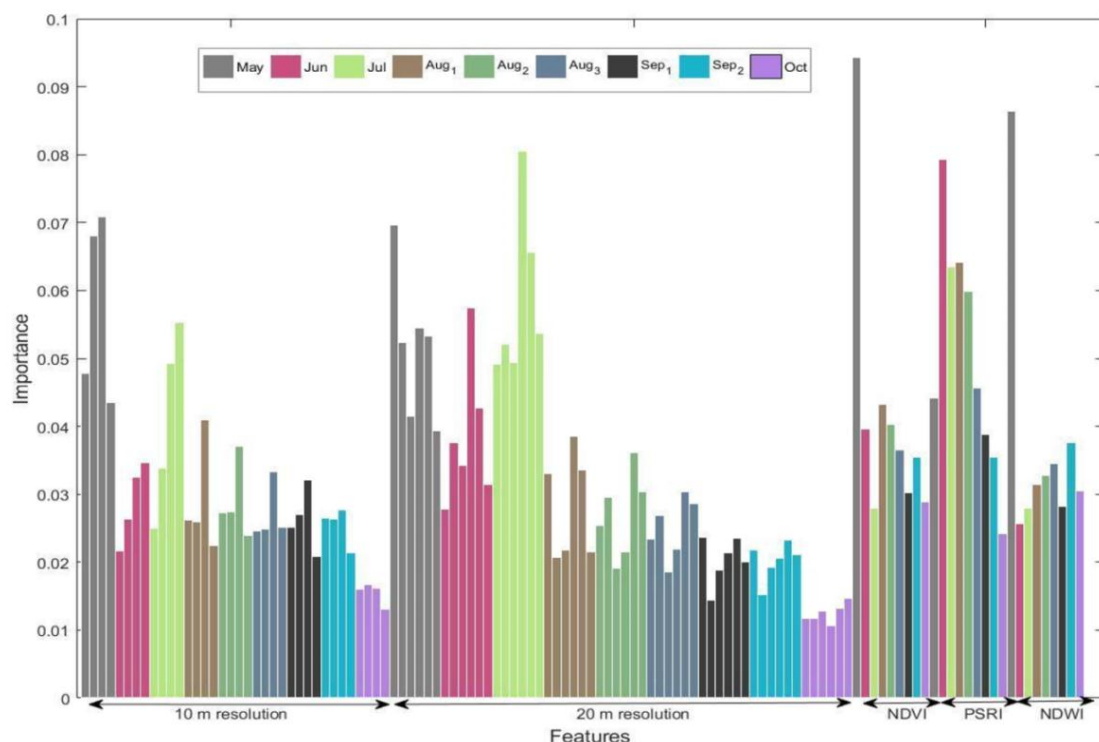


Figure 6. Feature importance bar graph of Sentinel-2 features, for the crop type classification. Features are grouped with respect to their spectral characteristics by position, and to their acquisition dates by color. The order for the 10 m resolution bands is B02-B03-B04-B08 and for the 20 m resolution bands is B05-B06-B07-B08A-B11-B12.

Feature importance, for the Sentinel-2 feature space, is calculated using the MATLAB function *ReliefF*. It is an extension of the original Relief algorithm suggested by Kira and Rendell (1992),

which estimates the quality of attributes with respect to the distinguishability they offer among classification entities near each other [37]. Further, the ReliefF extension enables feature importance estimation even in multi-class scenarios, whilst being more resilient to incomplete and noisy data [38]. The algorithm randomly selects entities (E_i) and thereafter searches for the k nearest neighbors coming from the same class, known as nearest hits (H_j), and the k nearest neighbors from different classes, known as nearest misses (M_j). Feature weights are estimated based on the E_i , H_j , and M_j values. The contribution of all hits and misses is averaged, with a contribution weight for each class of the misses being dependent on the prior probability of that class [39]. The function was set to search for 80 nearest neighbors, as trialing with larger values did not produce better data modeling.

May and July appear to attain the highest weights of importance. May coincides with the inflorescence emergence period for the winter crops, while July matches the winter crop senescence and summer crop leaf development (refer to Figure 3). PSRI is the most consistent of the VIs, having high weights of importance for nearly all scenes. In Figure 7 the highest ranked features, based on the feature importance estimations, are grouped into two thematic categories that relate to (a) month of image acquisition and (b) their spectral characteristics. Figure 7a quantifies the proportion of the top ranked features that stem from each month of image acquisition. In the same fashion, Figure 7b quantifies the proportion of top ranked features coming from (a) the visible and NIR bands of 10 m spatial resolution, (b) the vegetation red-edge and SWIR bands of 20 m resolution, and (c) the three VIs. The number of total top ranked features, for which the proportion values are calculated against, differ for each level of the nomenclature hierarchy. This refers to the least number of top ranked features that offers the first near-maximum classification accuracy; and amounts to 46, 25, and 21 features for the “type”, “family” and “season” classifications, respectively.

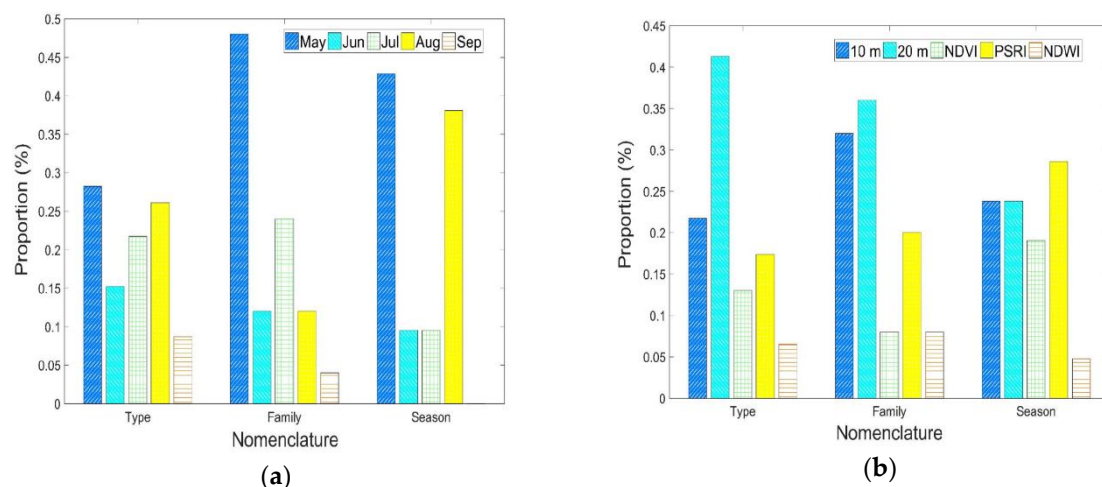


Figure 7. Proportion of feature contribution bar graphs for the Sentinel-2 scenario, calculated based on the feature importance values. The two figures are plotted for two thematic categories (a) month of image acquisition and (b) spectral information.

For the crop type level in Figure 7a, the most important features are distributed among May, June, July, and August. These months cover winter crops’ flowering to senescence stages and summer crops’ leaf development to flowering stages. Time-series of satellite imagery that covers the advanced phenological phases of both summer and winter crops is characterized by high discriminatory qualities. Furthermore, the image sensed during May proves of great significance in all nomenclature levels, explained by the dataset’s abundance in soft wheat and barley and thus in cereals and winter crops. In volume-wise skewed datasets, the dependence of variable importance in classes’ prior probability would affect the weights correspondingly. With respect to the spectral characteristics, it is the VIs and the bands of 20 m spatial resolution that contribute the most, despite the small volume of the

first and the lower spatial resolution of the latter. More specifically, what the red-edge, narrow-NIR, and SWIR data lack in spatial resolution, they gain in spectral resolution (narrower spectral windows) and reflectance profile distinctiveness among the different crop types.

Inspecting the higher levels of nomenclature in Figure 7b, crop season classification exposes considerable differences, with May and August being the main contributors in the discrimination process. Summer crops in May are germinating, while winter crops are already flowering or even ripening. On the other hand, summer crops in August are fully grown, whereas winter crops have already been harvested. Hence, the two months describe periods of maximal vegetation density difference between the two classes. This in turn explains the significant contribution of vegetation indices in crop season classification.

4.3. Cohen's Kappa Evolution with Increasing Number of Features

Based on the feature importance computation (Figure 6), the two classification algorithms are trained and tested for progressively larger feature spaces, as shown in Figure 8. The first point on the x-axis refers to a feature space comprising solely of the most important feature. For the second point onwards, features keep on populating the feature space, one by one, according to their importance ranking. It is observed that the total number of variables is large, with certain variables being strongly correlated and others being redundant by default.

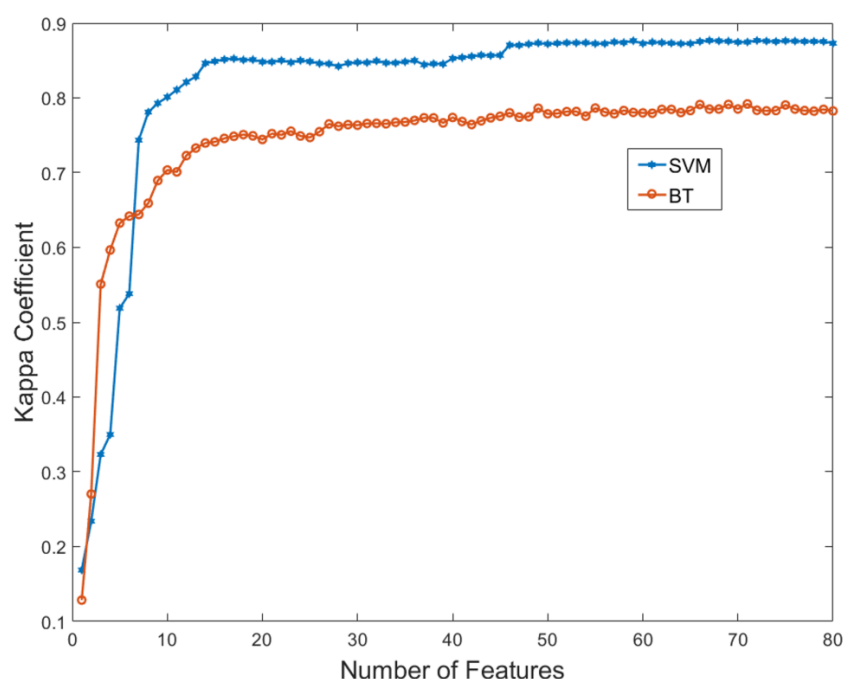


Figure 8. Kappa coefficient evolution for an increasing number of features for both SVM and RF crop type classifications.

In Figure 9 the SVM kappa coefficient evolution for all three levels of the nomenclature is depicted. All curves are almost monotonically increasing up until they reach a plateau, hence demonstrating the successful estimation of feature importance. Crop family and season of cultivation classifications reach their first optimal solution for spaces of fewer than 25 features, as indicated by the onset of the curves' flat region. On the other hand, crop type K_c evolution demands more than 45 to stabilize.

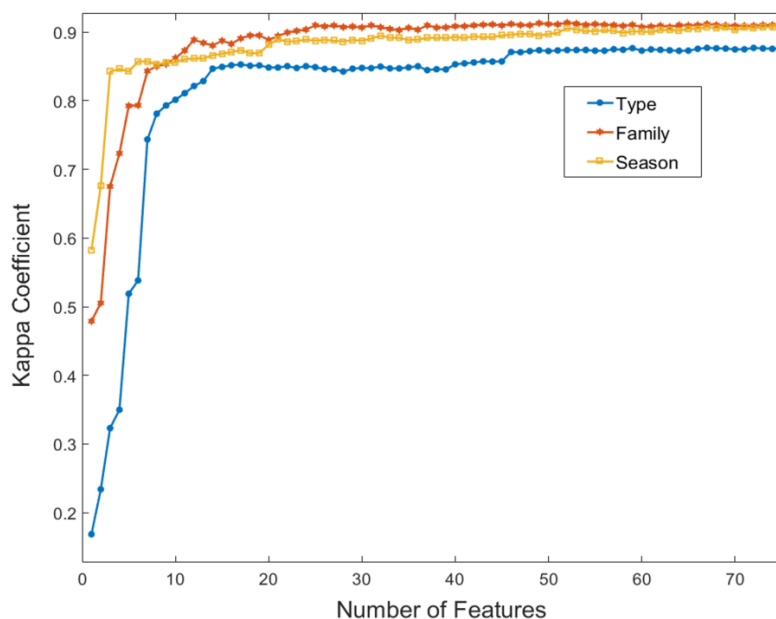


Figure 9. Kappa coefficient evolution for all levels of nomenclature using the SVM classifier.

Inspecting Figure 10, it is evident that RF performs exceptionally well for the crop family and season of cultivation classifications, with maximum K_c values comparable to the ones of SVM. However, crop type classification provides suboptimal performance. Introducing a larger number of spectrally similar classes, such as the constituent types of the cereal family, presents a challenge to the RF classification.

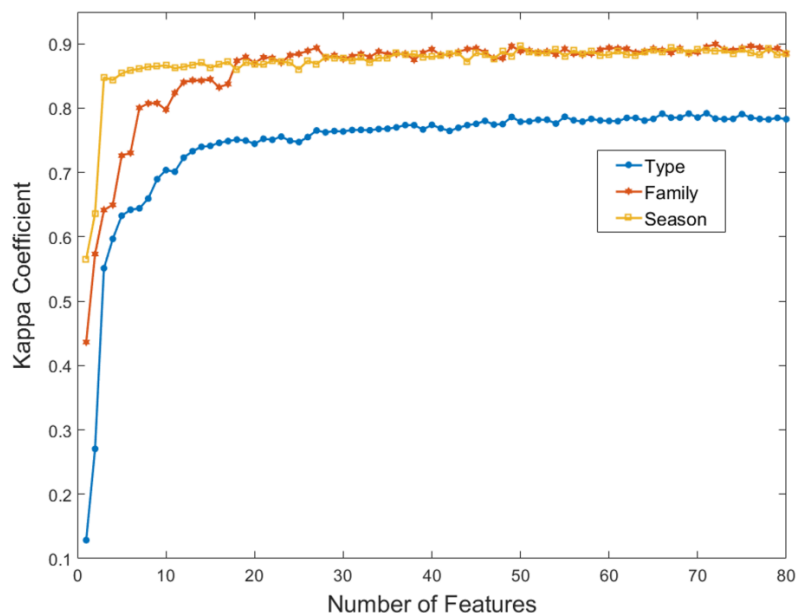


Figure 10. Kappa coefficient evolution for all levels of nomenclature using the RF classifier.

4.4. Execution Time Considerations

Sentinel's freely received high temporal and spatial resolution imagery demands big data handling, establishing computational efficiency as a key consideration. In this regard, RF and SVM classifiers are assessed in terms of both their training and implementation times. The values in Figures 11 and 12 have been averaged for five separate iterations.

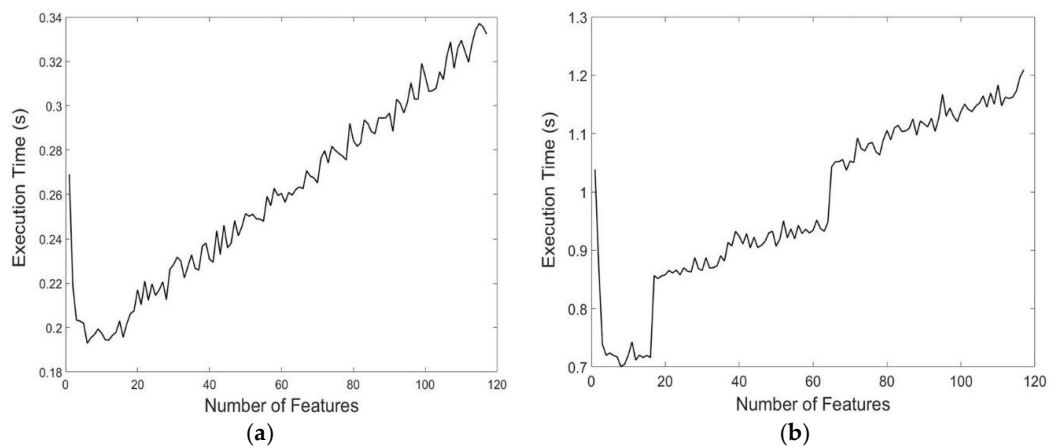


Figure 11. Execution time evolution with increasing number of features for the Sentinel-2 based RF crop type classification, showing in (a) the training time evolution and (b) the implementation time evolution.

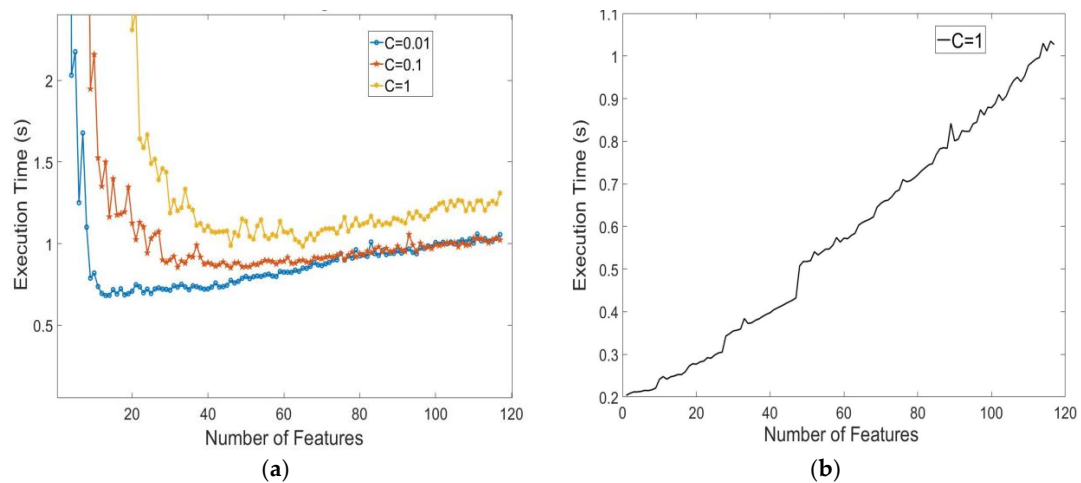


Figure 12. Execution time evolution with increasing number of features for the Sentinel-2 quadratic SVM crop type classification, showing in (a) the training time for three different C values and in (b) the implementation time for $C = 1$.

RF training time and both SVM and RF implementation times exhibit a linear relationship with an increasing number of features. The two distinct steps on the RF training curve are attributed to pre-processing overhead, evident in the MATLAB implementation of the algorithm. SVM training time curve acquires linearity only for feature spaces comprising of more than 45 features. Experimenting with lower box constraint values (0.01–0.1), thus forcing the optimizer to find larger margin hyperplanes, showed that training time is both low and linearly increasing, for even smaller feature spaces (Figure 12a). However, training accuracy is compromised as the misclassification cost constraint is relaxed. Inspecting Figure 8, the threshold of 45 features coincides with the onset of the classifiers accuracy evolution plateau. Since the constraint is strict, the optimizer takes longer to find a satisfying hyperplane for spaces of fewer than 45 features. All in all, higher execution times are evident for quadratic kernel SVM classifications, with differences strengthening as the number of variables increases (Table 6).

Table 6. Total RF and SVM (C = 1) execution times, including both training and implementation, as number of features increases (Sentinel-2 feature space).

Number of Features	Total Time RF	Total Time SVM	Difference
65	1.32 s	1.46 s	10.6%
90	1.42 s	1.80 s	26.7%
117	1.55 s	2.19 s	41.2%

5. Discussion

5.1. Sentinel-2 MSI and Landsat-8 OLI

We tested the performance of our processing workflow using Landsat data as input, as well. The aim is to highlight the performance improvement, if any, achieved with Sentinel-2 data, based on the unique spatial and temporal resolution, and the spectral characteristics in utilizing parts of the spectrum of key significance in discriminating between types of vegetation. The values in Table 7 have been averaged over 20 iterations of random training sample splits.

Table 7. Maximum K_c values for SVM crop type classifications based on the Sentinel-2 MSI, pan-sharpened Landsat-8 OLI and combined feature spaces.

Feature Space	Kappa RF	Kappa SVM
Landsat-8	0.5991	0.7011
Sentinel-2	0.8055	0.8830
Landsat-8 & Sentinel-2	0.8022	0.8908

Landsat crop-type classification proves to perform significantly worse than the Sentinel equivalent (Table 7). A combination of Landsat and Sentinel feature spaces marginally increases the overall accuracy and thus does not justify the higher complexity introduced. This is expected as the two feature spaces are highly correlated.

We used McNemar’s test to evaluate the statistical significance in the classification accuracy between (i) the two classifiers in the Sentinel scenario and (ii) Sentinel-Landsat pairs in the SVM scenario. This is an interpretable statistical test that quantifies the superiority between two thematic maps [40]. McNemar’s test is essentially a standardized normal chi-square statistic, computed from a two by two matrix based on correctly and incorrectly classified parcels in both classifications, as shown in Equation (6).

$$\chi^2 = \frac{(n_{ab} - n_{ba})^2}{n_{ab} + n_{ba}} \quad (6)$$

where n_{ab} is the number of parcels correctly classified by classifier one but incorrectly classified by classifier two; n_{ba} is the number of parcels correctly classified by classifier two but incorrectly classified by classifier one [36].

The difference between all pairs, shown in Table 8, proves to be statistically significant at a 99.99% confidence level. In this table, higher χ^2 values indicate better accuracy performance for the first item in the “pairs” column. These results further support the argument of SVM’s dominance over RF (pair 5) and similarly MSI’s dominance over OLI (pairs 1, 2 and 4). Relative χ^2 differences among the pairs can in ways determine the importance of parameters such as spatial resolution, spectral characteristics, and choice of classifier.

Table 8. McNemar’s test results for all different classification pairs (spatial resolution is defined in parentheses).

	Pairs		n_{ab}	n_{ba}	χ^2	p -Value	h
1	Sentinel (10 m)	Landsat (15 m)	205	1073	589.53	<0.0001	True
2	Sentinel (10 m)	Landsat (30 m)	207	1204	704.47	<0.0001	True
3	Landsat (15 m)	Landsat (30 m)	371	500	19.11	<0.0001	True
4	Sentinel (30 m)	Landsat (30 m)	322	988	338.54	<0.0001	True
5	Sentinel SVM	Sentinel RF	114	577	310.23	<0.0001	True

OLI’s data pan-sharpening and MSI’s data down-sampling, in pairs (1) and (4) respectively, attempt to reduce the effect of the spatial resolution difference and thus isolate and quantify the importance of other sensor attributes. The difference in accuracy is significant for both scenarios, which can be attributed to Sentinel-2’s superiority in spectral characteristics, specifically having four vegetation red-edge bands (B05, B06 and B07). The difference for pair (3) is statistically significant but, nonetheless, implies a marginal increase in accuracy. In other words, pan-sharpening proves indeed useful for Landsat-8 but cannot be compared with the higher spatial resolution, in which Sentinel-2 multispectral bands sense directly.

Moreover, pair (2) depicts the maximum difference between the two sensors, accounting for the impact of both spatial resolution and spectral characteristics. Sentinel’s superiority in temporal resolution could not be adequately exhibited for the present case study, as the cloud free images for the two sensors are comparable in both number and temporal span. Nevertheless, the sensors’ difference in temporal resolution is significant and even larger performance differences are expected in most of relevant scenarios.

5.2. Relevance of Methods

The overall performance of the proposed scheme is in accordance with the requirements of an operational agriculture monitoring system. Discrimination and characterization of the cultivated crop types, is of paramount importance in the development, conservation, and management of natural resources at both macro- and micro-scales [41]. The identification of crops, along with their distribution, management practices, and annual rotation schemes provide essential information for the implementation, control, and monitoring of agricultural policies and environmental measures imposed by the CAP [3].

Open access to the unmatched features of the Sentinel mission, in temporal and spatial resolution, and their efficient exploitation, signal a new era in the field. A revisit time of five days ensures the proper construction of imagery time-series, able for the consistent and timely monitoring of the agricultural landscape. Additionally, the alternative Landsat data, featuring a three-fold reduction in temporal resolution, can offer inadequate number of quality images, particularly in heavily clouded regions.

On the other hand, Sentinel-2’s high resolution multispectral data enable the successful employment of a parcel-based approach and the generation of parcel-specific thematic information. Besides, alternative VHR imagery is unrealistically costly for operational, large-scale, and consistent monitoring. The proposed methods, which effectively take advantage of the Sentinel-2 attributes, could ably function as the backbone for the EO-assisted compliance validation of CAP requirements, especially to what concerns the enhancement of transparency and the simplification of subsidy administration. The scheme was designed to directly assist the CAP paying agencies in their compliance inspections, by utilizing the data used (LPIS and farmers’ declarations) in their existing operations, to offer a monitoring alternative to their inefficient sample-based controls. The scheme is finally characterized by robustness, in the sense that it is largely independent of manual fine-tuning or case-specific optimizations and is thereby reproducible.

5.3. Relevance in Operational Scenarios

The notion of transferability, although addressed in the design process and the selection of input data, remains to be validated by applying the scheme to other agricultural landscapes in the European Union (EU). An evident issue of transferability is the absence of standardization and the dissimilarity of the LPIS data and farmers' declaration among the different EU countries. There are significant differences in the definition and description of cultivation practices and crop types. Dealing with this variability requires manual pre-processing and polishing of the input data and entails the appropriate definition of the crop types to be classified, merging or breaking them down to classes that describe spectrally coherent vegetation types.

In this study, it was shown that a time-series of multispectral imagery is required for accurate crop classification, as capturing the growing of crops exposes the most significant differences in their spectral signatures and thus enables their successful discrimination. Nonetheless, it was also exhibited that only a handful of images, appropriately distributed in time, is required to offer high thematic accuracy. The twin Sentinel-2 satellites offer a combined revisit time of five days, amounting to more than 70 images throughout the year. It is expected that in most cases this is a sufficient number of images to produce several quality cloud-free images within the year. However, in northern countries, where cloud-free imagery is scarce, Landsat-8 OLI data can be employed to enrich the feature space. Alternatively, fusion techniques with weather independent Sentinel-1 SAR data could also be explored.

When applying the scheme over large areas, the corresponding Sentinel-2 tiles can cover multiple adjacent satellite tracks, thus having different sensing dates. To overcome this issue, all tracks should be resampled with respect to a common temporal frame of reference, as proposed in [24]. Starting from the first image acquisition, a sampling step of five days, namely the Sentinel-2 revisit time, is set to create a grid of virtual sensing dates [24,42]. Then all images from all tracks are linearly interpolated to these predefined sample instances. The temporal interpolation was shown to marginally affect the classification accuracy [24].

Since the methodology introduced was designed to be fully transferable by making use of predominantly open access data and data provided by the targeted end-users, the concept of scalability in regional or national scales is of great interest. Scaling up comes against certain trade-offs, mainly regarding the computational complexity, overall accuracy performance, and geographic scale—which in ways can be addressed based on this study's results. Processing should be performed in adequately small regional extents, such as the one exhibited in the present study, in order to avoid the geographic variability of the crop type spectral signatures. Thereby, the difference in processing time between SVM and RF application, as presented in Section 4.4, can be thought of as linearly increasing in scaling up scenarios. SVM usage would indeed compromise computational economy, but not significantly. In that respect, accuracy is favored over computational complexity, as the relevant differences are less important. Nevertheless, in cases where the lowest level of nomenclature is not required, RF functions as an excellent alternative. It performs exceptionally for crop family and season of cultivation classifications, where it deals with fewer classes, of more distinct spectral profiles. All in all, it could be argued that both classifiers provide excellent results for a large number of classes, under an overall computationally efficient scheme.

Finally, the dependability of the methodology relies on the assumption that farmers' declarations are truthful. In this study, it was shown that declarations were correct in their vast majority (99%), which might not be the case for every relevant scenario. In [43] the authors have analyzed the effect of training class label noise for SVM (linear and radial basis function kernels) and RF classifiers that were applied on simulated vegetation profiles for 10 crop type classes. The results showed that SVM and particularly RF are robust for low noise levels (up to 20%), with a marginal decrease in the OA. The robustness of the RF can be attributed to its excellent generalization ability. Multiple uncorrelated weak learners are trained with a random subset of the training set and then decisions are made based on the majority vote of the ensemble, making the classifier resistant to overfitting [43]. All in all, it

can be argued that the original assumption is fair and that it can ultimately shape a robust scheme, even for cases where incorrect declarations amount to a considerable percentage of the training set.

5.4. Greening 1: Crop Diversification

Monitoring of compliance to CAP's Greening 1: Crop Diversification requirement is one example of the proposed scheme's direct application. Crop diversification entails the growing of different crop types based on farmers' total land area, aiming to improve biodiversity and reduce soil erosion. Farmers owning land of less than 10 ha are automatically exempted from the rule. If however, their land extends between 10 and 30 ha, at least two different crop types must be cultivated, with the main one not exceeding 75% of the total land. Similarly, arable land larger than 30 ha should involve at least three different crop types, with the main one occupying up to 75% and the main two occupying less than 95% of total land [44]. Based on the SVM crop type classification, 42.6% of farmers were exempted by having total arable land smaller than 10 ha, while only 2.2% were found to not comply with the requirement.

The Greening 1 requirement considers the total arable land owned by the farmer in order to decide on their compliance. Since our knowledge was limited to this particular dataset, we assumed that every farmer's total land is in fact encompassed within the borders of the study area. Therefore, the results are not representative of local farmers' compliance to Greening 1 but merely exhibit the ease of decision making based on the crop identification product.

6. Conclusions

Sentinel-2 data time series using a parcel-based quadratic SVM classification provided a successful crop identification product, reaching an overall accuracy higher than 0.87 Cohen's kappa, for the lowest level of the crop nomenclature hierarchy. RF classification provided comparable results for the family and season of cultivation nomenclature levels, while it underperformed for the crop type level. Also, Landsat-8 OLI imagery, of lower spatial resolution, resulted in inferior performance, validating the argument of Sentinel 2's dominance in smallholding agriculture mapping.

Accurate crop classification enables the monitoring of the CAP and allows for effective and efficient decision making on farmer compliance. The proposed methods were designed appropriately to be geographically independent and potentially scalable from the extent of a small region to national or even continental scales. In the same fashion, input data are deliberately kept least and freely available in order to achieve maximum transferability and ease of data retrieval.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-4292/10/6/911/s1>.

Author Contributions: V.S., I.P. and C.K. conceived and designed the methodology; V.S. and I.P. implemented the scheme and performed all relevant experiments; A.L.A., A.A.A. and J.G.Z. collected and provided the input data (LPIS, farmer declarations, validation dataset) and critically analyzed the results; V.S. wrote the paper.

Acknowledgments: Authors acknowledge the Copernicus Open Access Hub (<https://scihub.copernicus.eu/>), Hellenic National Sentinel Data Mirror Site (<https://sentinels.space.noa.gr/>) and the USGS for providing free access to Sentinel-2 MSI and Landsat-8 OLI images, correspondingly. We are also grateful to INTIA for supplying the farmers' declarations, LPIS data and phenology timeline. Finally, the authors would like to acknowledge the contribution of the section editor in chief Clement Atzberger and an anonymous reviewer, whose comments and suggestions were instrumental in strengthening the arguments of the study.

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

References

1. Fritz, S.; See, L.; You, L.; Justice, C.; Becker-Reshef, I.; Bydekerke, L.; Cumani, R.; Defourny, P.; Erb, K.; Foley, J.; et al. The Need for Improved Maps of Global Cropland. *Eos Trans. Am. Geophys. Union* **2013**, *94*, 31–32. [CrossRef]
2. Schmedtmann, J.; Campagnolo, M.L. Reliable crop identification with satellite imagery in the context of Common Agriculture Policy subsidy control. *Remote Sens.* **2015**, *7*, 9325–9346. [CrossRef]

3. Alganci, U.; Sertel, E.; Ozdogan, M.; Ormeci, C. Parcel-Level Identification of Crop Types Using Different Classification Algorithms and Multi-Resolution Imagery in Southeastern Turkey. *Photogramm. Eng. Remote Sens.* **2013**, *79*, 1053–1065. [CrossRef]
4. European Union. Regulation (EU) No. 1306/2013. Available online: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:347:0549:0607:EN:PDF> (accessed on 7 May 2018).
5. Eyraud, F.; Burger, A.; Åstrand, P.; Matteo, G.D. Community Image Data portal: Sharing licensed Earth observation data. *Int. J. Spat. Data Infrastruct. Res.* **2011**, *6*, 187–205. [CrossRef]
6. Odenweller, J.B.; Johnson, K.I. Crop identification using Landsat temporal-spectral profiles. *Remote Sens. Environ.* **1984**, *14*, 39–54. [CrossRef]
7. Lebourgeois, V.; Dupuy, S.; Vintrou, É.; Ameline, M.; Butler, S.; Bégué, A. A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated Sentinel-2 time series, VHRS and DEM). *Remote Sens.* **2017**, *9*, 259. [CrossRef]
8. Yang, C.; Everitt, J.H.; Fletcher, R.S.; Murden, D. Using high resolution QuickBird imagery for crop identification and area estimation. *Geocarto Int.* **2007**, *22*, 219–233. [CrossRef]
9. Turker, M.; Ozdarici, A. Field-based crop classification using SPOT4, SPOT5, IKONOS and QuickBird imagery for agricultural areas: A comparison study. *Int. J. Remote Sens.* **2011**, *32*, 9735–9768. [CrossRef]
10. 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. [CrossRef]
11. Houborg, R.; McCabe, M. High-Resolution NDVI from Planet’s Constellation of Earth Observing Nano-Satellites: A New Data Source for Precision Agriculture. *Remote Sens.* **2016**, *8*, 768. [CrossRef]
12. Vyas, S.P.; Oza, M.P.; Dadhwal, V.K. Multi-crop separability study of Rabi crops using multi-temporal satellite data. *J. Indian Soc. Remote Sens.* **2005**, *33*, 75–79. [CrossRef]
13. Jones, H.G.; Vaughan, R.A. Integrated Applications: Precision Agriculture and Crop Management. In *Remote Sensing of Vegetation: Principles, Techniques, and Applications*; Oxford University Press: Oxford, UK, 2011; ISBN 978-0199207794.
14. Foody, G.M. Supervised image classification by MLP and RBF neural networks with and without an exhaustively defined set of classes. *Int. J. Remote Sens.* **2004**, *25*, 3091–3104. [CrossRef]
15. Duro, D.C.; Franklin, S.E.; Dubé, M.G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sens. Environ.* **2012**, *118*, 259–272. [CrossRef]
16. Noi, P.T.; Kappas, M. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors* **2017**, *18*, 18.
17. Peña, J.; Gutiérrez, P.; Hervás-Martínez, C.; Six, J.; Plant, R.; López-Granados, F. Object-Based Image Classification of Summer Crops with Machine Learning Methods. *Remote Sens.* **2014**, *6*, 5019–5041. [CrossRef]
18. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 247–259. [CrossRef]
19. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
20. Valero, S.; Morin, D.; Inglada, J.; Sepulcre, G.; Arias, M.; Hagolle, O.; Dedieu, G.; Bontemps, S.; Defourny, P.; Koetz, B. Production of a dynamic cropland mask by processing remote sensing image series at high temporal and spatial resolutions. *Remote Sens.* **2016**, *8*, 1–21. [CrossRef]
21. Matton, N.; Canto, G.S.; 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. [CrossRef]
22. Inglada, J.; Vincent, A.; Arias, M.; Marais-Sicre, C. Improved early crop type identification by joint use of high temporal resolution sar and optical image time series. *Remote Sens.* **2016**, *8*, 362. [CrossRef]
23. Immitzer, M.; Vuolo, F.; Atzberger, C. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sens.* **2016**, *8*, 166. [CrossRef]
24. Inglada, J.; Arias, M.; Tardy, B.; Hagolle, O.; Valero, S.; Morin, D.; Dedieu, G.; Sepulcre, G.; Bontemps, S.; Defourny, P.; et al. Assessment of an operational system for crop type map production using high temporal and spatial resolution satellite optical imagery. *Remote Sens.* **2015**, *7*, 12356–12379. [CrossRef]

25. 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. [\[CrossRef\]](#)
26. Lowder, S.K.; Skoet, J.; Raney, T. The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Dev.* **2016**, *87*, 16–29. [\[CrossRef\]](#)
27. Congedo, L. *Semi-Automatic Classification Plugin Documentation*; ResearchGate: Berlin, Germany, 2015; p. 106. [\[CrossRef\]](#)
28. Hatfield, J.L.; Prueger, J.H. Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management practices. *Remote Sens.* **2010**, *2*, 562–578. [\[CrossRef\]](#)
29. Xiao, X.; Boles, S.; Liu, J.; Zhuang, D.; Frolking, S.; Li, C.; Salas, W.; Moore, B. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sens. Environ.* **2005**, *95*, 480–492. [\[CrossRef\]](#)
30. Huang, J.; Wang, H.; Dai, Q.; Han, D. Analysis of NDVI data for crop identification and yield estimation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4374–4384. [\[CrossRef\]](#)
31. 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. [\[CrossRef\]](#)
32. Lawrence, R.L.; Wood, S.D.; Sheley, R.L. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest). *Remote Sens. Environ.* **2006**, *100*, 356–362. [\[CrossRef\]](#)
33. Story, M.; Congalton, R.G. Accuracy assessment: A user's perspective. *Photogramm. Eng. Remote Sens.* **1986**, *52*, 397–399. [\[CrossRef\]](#)
34. Cohen, J. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46. [\[CrossRef\]](#)
35. Liu, C.; Frazier, P.; Kumar, L. Comparative assessment of the measures of thematic classification accuracy. *Remote Sens. Environ.* **2007**, *107*, 606–616. [\[CrossRef\]](#)
36. Petropoulos, G.P.; Kontoes, C.C.; Keramitsoglou, I. Land cover mapping with emphasis to burnt area delineation using co-orbital ALI and Landsat TM imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 344–355. [\[CrossRef\]](#)
37. Kira, K.; Rendell, L.A. A Practical Approach to Feature Selection. In Proceedings of the Ninth International Workshop on Machine Learning (ML92), Aberdeen, UK, 1–3 July 1992; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1992; pp. 249–256.
38. Kononenko, I. Estimating attributes: Analysis and extensions of RELIEF. In *Machine Learning: ECML-94 Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 1994; pp. 171–182.
39. Robnik-Šikonja, M.; Kononenko, I. Theoretical and Empirical Analysis of Relief and RRelief. *Mach. Learn.* **2003**, *53*, 23–69. [\[CrossRef\]](#)
40. Foody, G.M. Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 627–633. [\[CrossRef\]](#)
41. Li, Q.; Cao, X.; Jia, K.; Zhang, M.; Dong, Q. Crop type identification by integration of high-spatial resolution multispectral data with features extracted from coarse-resolution time-series vegetation index data. *Int. J. Remote Sens.* **2014**, *35*, 6076–6088. [\[CrossRef\]](#)
42. Inglada, J.; Vincent, A.; Arias, M.; Tardy, B.; Morin, D.; Rodes, I. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. *Remote Sens.* **2017**, *9*, 95. [\[CrossRef\]](#)
43. Pelletier, C.; Valero, S.; Inglada, J.; Champion, N.; Sicre, C.M.; Dedieu, G. Effect of Training Class Label Noise on Classification Performances for Land Cover Mapping with Satellite Image Time Series. *Remote Sens.* **2017**, *9*, 173. [\[CrossRef\]](#)
44. Greening. Available online: https://ec.europa.eu/agriculture/direct-support/greening_en. (accessed on 8 June 2018).

