Remote Sens. **2015**, *7*, 10400-10424; doi:10.3390/rs70810400

OPEN ACCESS

remote sensing

ISSN 2072-4292

www.mdpi.com/journal/remotesensing

Article

Land Cover and Crop Type Classification along the Season Based on Biophysical Variables Retrieved from Multi-Sensor High-Resolution Time Series

François Waldner [1,*], Marie-Julie Lambert [1], Wenjuan Li [2], Marie Weiss [2], Valérie Demarez [3], David Morin [3], Claire Marais-Sicre [3], Olivier Hagolle [3], Frédéric Baret [2] and Pierre Defourny [1]

[1] Earth and Life Institute, Université catholique de Louvain, 2 Croix du Sud, 1348 Louvain-la-Neuve, Belgium; E-Mails: julie.lambert@uclouvain.be (M.-J.L.); pierre.defourny@uclouvain.be (P.D.)

[2] INRA-EMMAH UMR 1114, 84914 Avignon, France; E-Mails: wenjuan.li@paca.inra.fr (W.L.); marie.weiss@paca.inra.fr (M.W.); baret@avignon.inra.fr (F.B.)

[3] CESBIO, UMR CNES-CNRS-IRD-UPS, 18 avenue Edouard Belin, 31401 Toulouse Cedex 4, France; E-Mails: alerie.demarez@cesbio.cnes.fr (V.D.); david.morin@cesbio.cnes.fr (D.M.); claire.marais-sicre@cesbio.cnes.fr (C.M.-S.); olivier.hagolle@cnes.fr (O.H.)

***** Author to whom correspondence should be addressed; E-Mail: francois.waldner@uclouvain.be; Tel.: +32-1047-3680.

Academic Editors: Sylvia Sylvander, Olivier Arino, Benjamin Koetz, Clement Atzberger and Prasad S. Thenkabail

Received: 31 May 2015 / Accepted: 4 August 2015 / Published: 13 August 2015

Abstract: With the ever-increasing number of satellites and the availability of data free of charge, the integration of multi-sensor images in coherent time series offers new opportunities for land cover and crop type classification. This article investigates the potential of structural biophysical variables as common parameters to consistently combine multi-sensor time series and to exploit them for land/crop cover classification. Artificial neural networks were trained based on a radiative transfer model in order to retrieve high resolution LAI, FAPAR and FCOVER from Landsat-8 and SPOT-4. The correlation coefficients between field measurements and the retrieved biophysical variables were 0.83, 0.85 and 0.79 for LAI, FAPAR and FCOVER, respectively. The retrieved biophysical variables' time series displayed consistent average temporal trajectories, even though the class variability and signal-to-noise ratio increased compared to NDVI. Six random forest classifiers were trained and applied along the season with different inputs: spectral bands, NDVI, as well as FAPAR, LAI and FCOVER, separately

and jointly. Classifications with structural biophysical variables reached end-of-season overall accuracies ranging from 73%–76% when used alone and 77% when used jointly. This corresponds to 90% and 95% of the accuracy level achieved with the spectral bands and NDVI. FCOVER appears to be the most promising biophysical variable for classification. When assuming that the cropland extent is known, crop type classification reaches 89% with spectral information, 87% with the NDVI and 81%–84% with biophysical variables.

Keywords: classification; biophysical variables; along the season; land cover; crop types

1. Introduction

As the number of satellites keeps increasing and the data become free of charge, the working environment is shifting to a data-rich environment. The integration of multi-sensor images in coherent time series and products becomes therefore of critical importance [1]. Besides, multi-sensor time series are a sensible solution to meet the requirements for agricultural monitoring, especially in areas affected by persistent cloud coverage at critical moments of the season [2,3]. A way forward for multi-sensor data integration is to merge time series at the biophysical variables level, e.g., using Leaf Area Index (LAI), because biophysical variables have a meaning that is independent from the sensors' characteristics. Biophysical variables from different sensors could be integrated to densify the time series, allowing a finer monitoring to better capture the rapid changes of vegetation, such as vegetative crop development. In principle, biophysical variables extracted from radar could then also be integrated into the time series [4].

Biophysical variables, such as LAI and the fraction of absorbed photosynthetic active radiation (FAPAR), can be derived from observations in the reflective solar domain [5]. These vegetation variables play a key role in several surface processes, including photosynthesis, respiration and transpiration. LAI is defined as half the total developed area of green elements per horizontal ground area unit [6]. The Normalized Difference Vegetation Index (NDVI) shows a logarithmic response to LAI: the relationship is tight at low range (LAI from 0–1.5) and scattered at medium range (LAI from 1.5–4); thereafter, NDVI saturates (LAI > 4) [7–11]. FAPAR, one of the main inputs in light use efficiency models [12], is defined as the fraction of radiation absorbed by the green vegetation elements in the 400–700-nm spectral domain under specified illumination conditions. In addition to LAI and FAPAR, the FCOVER, the fraction of green vegetation as seen from nadir, is requested by some users for vegetation monitoring, as well as for partitioning contributions between soil and vegetation within specific models for numerical weather prediction, regional and global climate modeling and global change monitoring [13]. FCOVER is independent from the illumination conditions as opposed to FAPAR, while showing sensitivity to a vegetation amount intermediate between FAPAR and LAI.

This papers aims at assessing the potential of multi-sensor structural biophysical variables (LAI, FAPAR, FCOVER) for land cover and crop type classification at high spatial (20 m) and temporal resolution with a focus on agriculture, as crops are dynamic and require dense observations. The use

of biophysical variables for classification is expected to better capture the vegetation dynamics, because they do not saturate at high NDVI values, yielding better class discrimination. Furthermore, they would foster the re-utilization of training data for classification, as they are not sensor-specific. The first part of the work focuses on extracting, integrating and evaluating the consistency of the retrieved high-resolution multi-sensor biophysical time series. Second, the potential of land/crop cover classification with the retrieved bio-physical variables along the season was assessed and compared to accuracies obtained with spectral information or NDVI directly. Particular focus was placed on the evolution of the accuracy of the different classifications along the season as information accumulates. Together, this permits answering the following research questions: (1) are biophysical variables more useful compared to traditional spectral and/or vegetation index-based inputs, and if not, what are the reasons for this underperformance; (2) are some biophysical variables more useful than others and why; (3) what are the most critical dates for class discrimination; and (4) how many observations are required within a given growing season for optimum classification accuracy.

2. Material

2.1. Study Site

The experiment was performed on the Joint Experiment for Crop Assessment and Monitoring (JECAM) site of South Midi Pyrenées in France over 4500 km^2 (Figure 1). The climatic zone is temperate with a mean annual precipitation of 650 mm. This study area is characterized by a clay loamy soil with a small slope (3%) toward the north. The typical field size varies around 23.5 ha. Main crop types are winter wheat and rapeseed. According to the crop calendar, sowing occurs in October and September for winter wheat and rapeseed and harvest in July and June, respectively. Broadleaved and needle-leaved forests cover the south of the study area.

2.2. Datasets

The remote sensing dataset is composed of 34 images acquired by two sensors: 16 SPOT-4 and 18 Landsat-8. The Landsat-8 data were obtained through the online Data Pool at the NASA Land Processes Distributed Active Archive Center (https://lpdaac.usgs.gov/get_data). The SPOT-4 imagery was obtained under the SPOT4/Take5 program. To increase the satellite observation frequency, the site was delimited by the overlap area of the area imaged during the SPOT-4 Take-5 experiment and two Landsat-8 scenes (Paths 198–199 and Row 30). The SPOT-4 Take-5 experiment consisted of lowering SPOT-4's altitude to a 5-day repeat cycle orbit to mimic the revisit cycle that ESA's Sentinel-2 mission will provide. SPOT-4 (20 m, 5-day revisit cycle) surface reflectance data from February 2013–May 2013 (Figure 2) were corrected from atmospheric effects, including adjacency effects and the terrain effects [14]. Clouds and associated shadows were removed from the original surface reflectance data with the Multisensor Atmospheric Correction and Cloud Screening (MACCS) processor [15]. MACCS is based on a multi-temporal method for cloud screening, cloud shadow detection, water detection, as well as for the estimation of the aerosol optical thickness. Landsat-8 (30 m, 16-day revisit cycle) surface reflectance data spanned from April–December 2013 (Figure 2). For the visible, near and short wave

infrared bands, Landsat data were also processed with MACCS. However, the processing was enriched thanks to the additional Landsat-8 spectral bands: (i) the 1.38-μm band enabled an enhanced detection of high and thin clouds; and (ii) the blue band provided an additional criterion to detect the aerosols, thanks to its quasi-constant relationship with the surface reflectances in the red above vegetation. The precision gain due to this criterion compensates for the precision loss due to the lower repeatability of Landsat-8 images. Besides, the thermal bands of Landsat-8 were not processed.

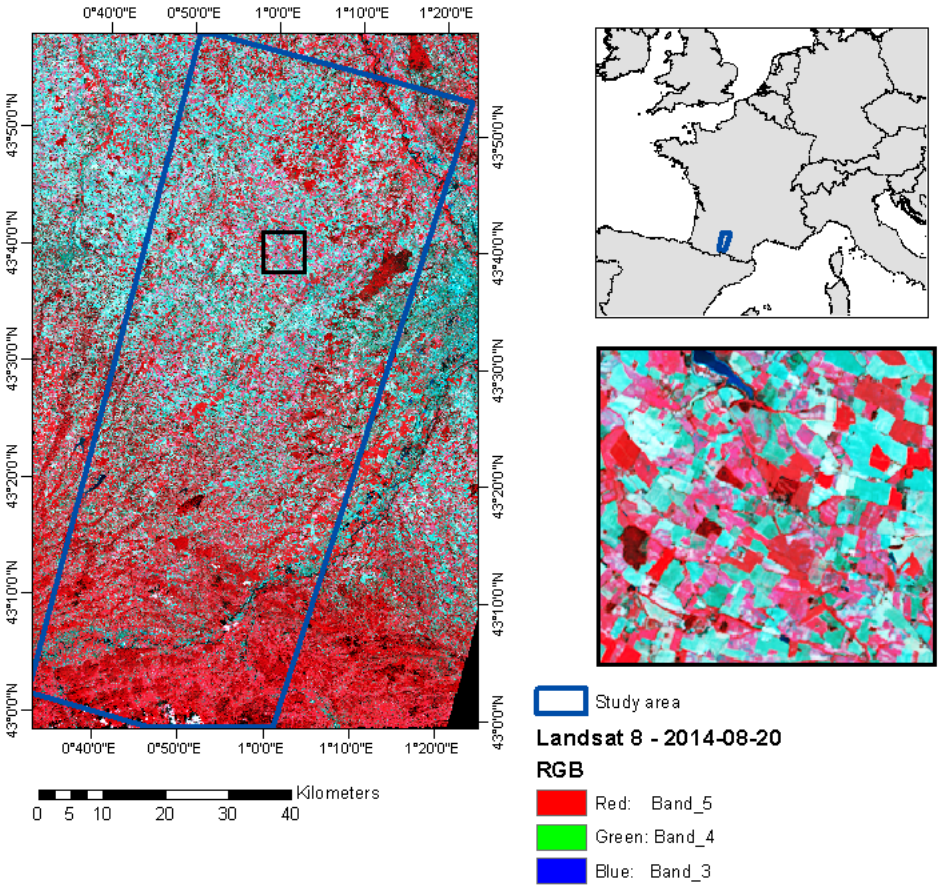


Figure 1. Location of the study site in southwest France. The blue box represents the study area. The background image is a Landsat-8 False Color Composite of 20 August 2014.

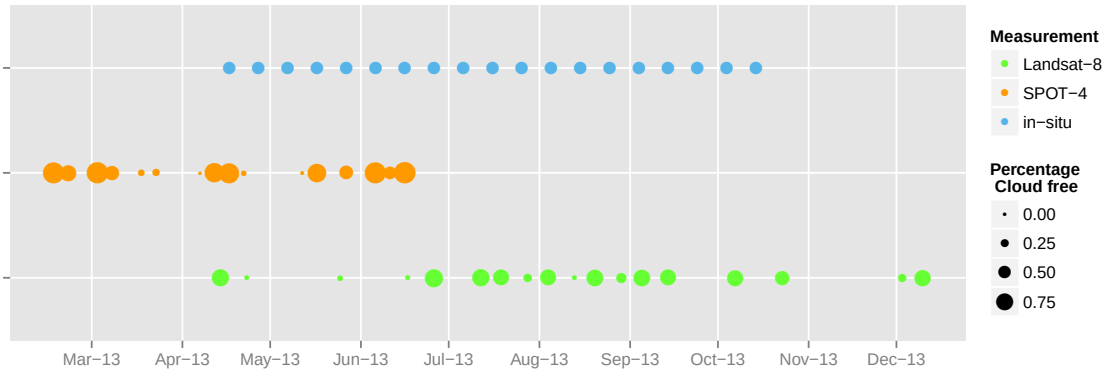


Figure 2. Temporal distribution of the images and their respective percentage of cloud-free area expressed as a fraction of the study area along the 2013 growing season.

Biophysical variable ground observations were collected in elementary sample units (ESU) (20 × 20 m^2) over wheat, corn and sunflower from 17 April–23 October 2013 every 10 days with hemispherical photographs (Nikon CoolPix 8400 camera with FC-E8 fisheye lens) [16]. LAI, FCOVER and FAPAR were derived from valid hemispherical photography inside an ESU using CAN-EYE software [17]. Its good performance to extract LAI from hemispherical pictures compared to destructive measurements has been demonstrated. Mougin *et al.* [18] showed a very good agreement between the two estimates with a correlation coefficient R^2 of 0.97 and a root mean square error of 0.26 (n = 99) in Sahelian rangelands. For crops, Demarez *et al.* [19] demonstrated a strong correlation between the estimates and the destructive measurements (root mean square error of 0.63 and correlation coefficient of 0.95). However, they also highlighted a systematic skew (y = 0.7810x) that leads to the well-known LAI underestimation in clumped canopies.

A reference dataset was built for the training of the classifiers and the validation of the maps based on a field visit for the crop classes and on visual interpretation for the other land cover classes. For the former, 616 fields were visited in 2013 and the five crop types described: winter wheat (164), rapeseed (49), corn (277), barley (26) and sunflower (100). For the latter, the ESA CCI Land Cover [20] and the Google Earth tool were employed to help identify 5 land cover types (water bodies, urban areas, broadleaved forest, needle-leaved forest and grassland). Half of the sample pixels was randomly selected as the training set, and the remaining half was selected as the validation set.

3. Methodology

3.1. Biophysical Variable Retrieval

Biophysical variable retrieval methods fall usually into four categories: parametric and non-parametric empirical relationships, radiative transfer model inversion and hybrid methods [21]. First, parametric regression methods define relationships between selected spectral information (NDVI, for example) and biophysical variables of interest. This implies choices on the spectral bands to select, on the index formulation and on the fitting functions [22]. Second, non-parametric regression methods build direct non-parametric algorithms from full optical spectral datasets based, for instance, on machine learning tools. For these two empirical approaches (parametric and non-parametric), a calibration phase relying on local *in situ* data is necessary, which decreases their generic capabilities [23]. In general, non-parametric regressions were found to outperform simpler parametric methods [21]. However, due to their arithmetic simplicity, empirical relationships are computationally economical, which is very valuable when addressing large areas or to meet processing time requirements of operational retrieval. The physically-based methods build on physical laws and well-known physical relationships between variables [11]. The inversion of a radiative transfer model (RTM) with remote sensing data allows retrieving biophysical variables in a generic fashion. Amongst the different existing methods, neural networks (NN) are widely used and often preferred to look-up table approaches or iterative optimization methods [24–26]. Finally, hybrid methods combine physical and statistical approaches to derive the training set and define the relationships between remote sensing data and biophysical variables, respectively [27].

In this work, neural networks were trained using GeoSAIL, a radiative transfer model [28], to retrieve LAI, FAPAR and FCOVER. The retrieval procedure includes three steps [16]: (1) the generation of a training database with GeoSAIL using canopy biophysical variables—leaf structure parameter, leaf chlorophyll concentration, dry matter content, equivalent water thickness, LAI, average leaf angle and hot spot size parameter—and soil brightness parameters to simulate theoretical reflectances; (2) the calibration of the NNs to inverse GeoSAIL by minimizing the root mean square error (RMSE); and (3) the application of NNs to measured SPOT-4 and Landsat-8 reflectances and geometries of observation in order to derive the biophysical variables of interest (LAI, FAPAR, FCOVER). Due to spectral specificities and geometries of the two different sensors, separate neural networks were generated. Regarding the SPOT-4, all four bands (green, red, near-infrared (NIR), short-wave infrared (SWIR)) were kept as for Landsat-8; only four bands (blue, green, NIR and SWIR-1) were selected because they are sensitive to the canopy characteristics, such as chlorophyll, water and dry matter. The unique geometrical configuration (VZA: view zenith angle; VAA: view azimuth angle) was used in the training and inversion. For Landsat-8, the view angle is always nadir, so all images have the same geometrical configurations, but for SPOT-4 images, each image has different view angles. Two thirds of the training dataset were used to train the NNs, while the remaining was used to validate the performance. For each biophysical variable, several neural network structures were tested, and the one that minimized the RMSE compared to the validation dataset was selected for each variable.

As biophysical variables' *in situ* measurements were only available for a limited number of classes (wheat, corn, sunflower), the temporal consistency of the retrieval was further investigated and compared to that of NDVI using the signal-to-noise ratio (SNR). The SNR can be estimated as [29,30]:

$$\mathrm{SNR}_i = \frac{\sigma^2_{signal,i}}{\sigma^2_{noise,i}} \tag{1}$$

where σ^2_i represents the variance of the signal or of the noise of the time series of a given pixel i; the larger the signal-to-noise ratio, the higher the temporal consistency and conversely. To estimate the signal, time series were smoothed with a cubic spline approach. The variance of the signal was computed on the smoothed values resulting from the cubic spline filter and the variance of the noise on the difference between the smoothed values and the observed values. The signal-to-noise ratio was computed from the 34 dates of the NDVI, LAI, FAPAR and FCOVER time series for all pixels belonging to the vegetation classes of the reference data. Thereby, it ensures that SNR statistics by land cover class are free of classification errors.

3.2. Land Cover Classification along the Season and Assessment

A plethora of classifiers have been developed for land cover mapping based on satellite imagery. Those classification methods range from unsupervised clustering algorithms (ISODATA or K-means [31]) to supervised classifiers. The latter encompasses parametric statistical algorithms, such as maximum likelihood [32], machine learning algorithms, such as artificial neural networks [33], decision trees [34,35], support vector machines [36,37] and ensembles of (machine learning) classifiers[38]. Several studies have demonstrated the higher accuracy and efficiency of machine learning algorithms especially with high dimensional datasets and large area mapping [39–41]. These algorithms are

particularly efficient and effective, because they do not assume any data distribution (e.g., normality) and yield generally to higher accuracy levels [42,43]. However, the number of hyper-parameters to adjust complicates their use and their automation [44]. In addition, some machine learning algorithms tend to overfit the data [34]. Machine learning ensemble classifiers have been developed to overcome this drawback: they rely on a community of base machine learning classifiers that generate multiple outputs, which are then combined according to specific decision rules. Ensemble learning circumvents the weaknesses of the base classifiers while keeping their strengths and, therefore, yields higher accuracy levels [45,46]. While assessing the effectiveness of random forest [47] (an ensemble of decision trees) for land cover classification, Rodriguez-Galiano *et al.* [48] demonstrated that random forest does not overfit and offers several advantages, such as: (1) the low number of user-defined hyper-parameters; (2) the estimation of the importance of variables (bands) for the general classification of the land cover categories and for the classification of each category by means of the Gini Index; and (3) its robustness to noise and training dataset size reduction. For these reasons, the random forest classifier was chosen for this study.

Object-based classifications propose additional discrimination parameters, such as texture, contextual information and shape (e.g., see Vieira *et al.* [49]). Even if segmentation is time consuming, object-based approaches offer a more generalized and more contiguous depiction of land cover, usually better matching the landscape perception of analysts [50]. The generalized appearance of the classes may account for an apparent preference over slightly better performing pixel-based classifications (e.g., [51]). Nevertheless, additional processing of pixel-based imagery (such as a filter), can also result in similar representations. The actual effects of object-based classifications seem to depend on the classes to be mapped [52], on the classification methods [53] and the areas considered [54]. Besides, the statistical significance of the method appears also highly variable [52,53,55,56]. As the additional discrimination power offered by object-based analysis was assumed to be similar for all input data (reflectances, vegetation indices or biophysical variables), pixel-based classifications were preferred.

Prior to classification, missing observations were filled in by linear interpolation between successive values, and Landsat-8 images were resampled to SPOT's grid using a nearest neighbor approach. The interpolation was achieved in a sensor-specific way for the spectral bands and NDVI, as all bands are not necessarily on-board both sensors and because their spectral responses differ. For the biophysical variables, the missing values were interpolated from both sensors simultaneously.

The selected output legend in terms of land cover and crop type distribution includes eleven classes: five crop types (barley, wheat, rapeseed, sunflower and corn), needle-leaved and broadleaved forests, grassland, urban areas and water bodies. As it was shown that the random forest out-of-bag error significantly underestimates error and can differ from independent error estimates [57], the reference data were randomly split into two parts: half of the samples was selected to train random forest classifiers along the season as the remote sensing data accumulates, while the other half served for independent classification accuracy assessment (Table 1). Six classifiers were trained using as input (1) all spectral bands; (2) NDVI; (3) LAI; (4) FAPAR; (5) FCOVER and (6) all three biophysical variables. For the sake of clarity, these classifications are henceforth referred to as Spectral-C, NDVI-C, LAI-C, FAPAR-C, FCOVER-C and BPV-C, respectively.

Table 1. Number of available training and validation pixels per class. The reference data were randomly split into two parts at the object level: half of the samples was selected to train random forest classifiers along the season as the remote sensing data accumulates, while the other half served for independent classification accuracy assessment.

Class	Training (n pixels)	Validation (n pixels)
Barley	1849	1609
Corn	16,670	18,180
Rapeseed	3755	4526
Sunflower	9393	7571
Winter Wheat	14,769	11,349
Broadleaved Forest	1645	2116
Needle-leaved Forest	1479	915
Grassland	8160	10,162
Urban	3108	1752
Water	592	434

The classification accuracy of the classifiers was assessed via quantitative accuracy indicators [58], including overall classification accuracy, producer's accuracy, user's accuracy and F_1-scores. In particular, the overall accuracy and the F_1-scores were computed for each classifier and at each date to study the evolution of the accuracy. To test the statistical significance of the differences in the classifications (LAI-C, FAPAR-C, FCOVER-C, BPV-C, NDVI-C), a z-test was applied. The z-test assesses the significance level of differences between the accuracy indicators of different classifications [59]:

$$z = \frac{\kappa_1 - \kappa_2}{\sqrt{\sigma_{\kappa_1} + \sigma_{\kappa_2}}} \quad (2)$$

where κ_1 and κ_2 represent the estimated kappa coefficients from two different classifications and σ_{κ_1} and σ_{κ_2} represent their corresponding variances [59–63]. Assuming a normal distribution, a statistically-significant difference is considered if $|z| > \alpha/2$, where $\alpha/2$ represents the cut-off value in the standard normal curve's upper tail. For α = 0.05, the difference can be declared significant at the 5% significance level if the z-test yields $|z| > 1.96$. Pairwise significance was assessed for all combinations of LAI-C, FAPAR-C, FCOVER-C, NDVI-C and BPV-C.

3.3. Importance of the Date and of the Length of the Time Series

Quantifying the importance of each image acquisition date permits evaluating the optimal image acquisition periods for classification. Amongst the possible alternatives [64–66], the mean decrease of the Gini Index (GI) was calculated and stored at each observation (34 dates) and for each classification. To assess the date importance, the random forest turns off one of the acquisition dates and keeps the others constant to evaluate the decline in accuracy using the Gini Index at each node of the random forest [47]. The mean decrease in GI measures the feature importance (acquisition dates in this case): the larger the decrease, the more important the variable. Four categories of importance were determined based on the quartiles of mean decrease Gini values: very high importance, high importance, medium importance and low importance.

Finally, the influence of the number of dates on the overall accuracy of the classification was investigated through a sensitivity analysis. For each number of dates ranging from 1–33, ten dates were systematically randomly selected from the complete FCOVER time series. For each random subset, a random forest classifier was trained, and the accuracy of the resulting maps was evaluated. The distribution of the overall accuracy within each sample and against samples indicates the pertinent number of dates to be used.

4. Results

4.1. Biophysical Variable Retrieval and Temporal Consistency

The retrieved biophysical variables show strong correlation with independent field measurements realized on corn, wheat and sunflower fields from April–23 October (Table 2 and see [16]). The overall RMSE is close to the Global Climate Observing System (GCOS) accuracy requirements, specifically ± 0.5 for LAI and ± 0.05 for FAPAR [67]. The strongest relationship and lowest bias were observed for corn (R^2 = 0.91, bias = 0.01) and the largest bias for wheat (bias = 0.11). LAI, FAPAR and especially FCOVER retrieved from neural networks are overestimated compared to field data (all crop types considered), as all biases are positive. This is probably due to the contribution of soil reflectance, clumping effects or field data uncertainties [16]. Similar results are expected for rapeseed and barley, but they cannot be ascertained in the absence of field data.

Table 2. Biophysical variable retrieval: validation with field data.

	LAI	**FAPAR (Black Sky)**	**FAPAR (White Sky)**	**FCOVER**
R^2	0.83	0.86	0.84	0.79
Bias	0.07	0.02	0.05	0.09
RMSE	0.49	0.1	0.12	0.15

Average temporal profiles for each land cover type were extracted (Figure 3) and show that they consistently integrate the two sensors and depict the expected temporal dynamics despite some outliers especially apparent in the FCOVER time series. In fact, regressions between SPOT-4 and Landsat-8 for two close dates (17 and 14 April, respectively) have a coefficient of determination of 0.61 for LAI, 0.65 for FAPAR and 0.63 for FCOVER, which confirms the good integration of the two sensors [16].

FCOVER of winter crops—barley, rapeseed and wheat—increases from the end of March until the end of May, and begins to decrease from June, which is consistent with the growing cycle. Despite some minor differences of the average values, their temporal profiles appear very similar. Summer crops—corn and sunflower—reach a maximum FCOVER at the end of August and the end of July, respectively. FCOVER of grassland rises from the end of February to the start of June and then declines slightly from June to the end of the year. Similar seasonal profiles can be observed for broadleaved and needle-leaved forests. From the temporal consistency viewpoint, FAPAR and LAI show profiles that are very much alike for all classes. The drop in NDVI for the needle-leaved forest at the beginning of the season might be due to spurious interpolation caused by high cloudiness. If the temporal dynamic appears consistent

for all biophysical variables, an enlargement of the variability of the time series compared to the NDVI is noticeable.

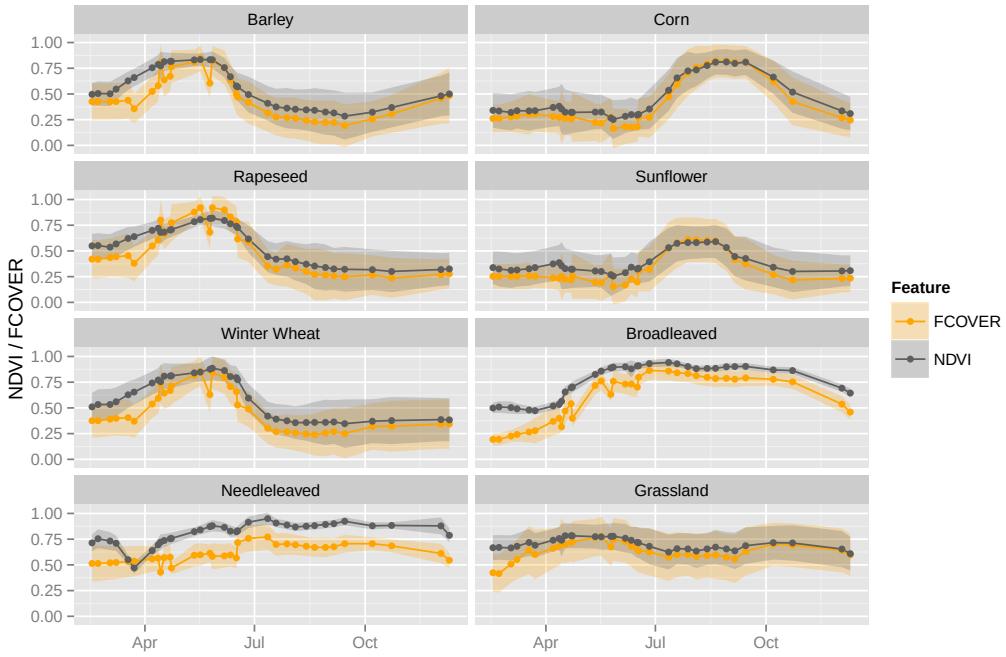


Figure 3. Temporal profiles of NDVI and FCOVER mean (± 1 standard deviation) computed from the reference sample set for selected classes.

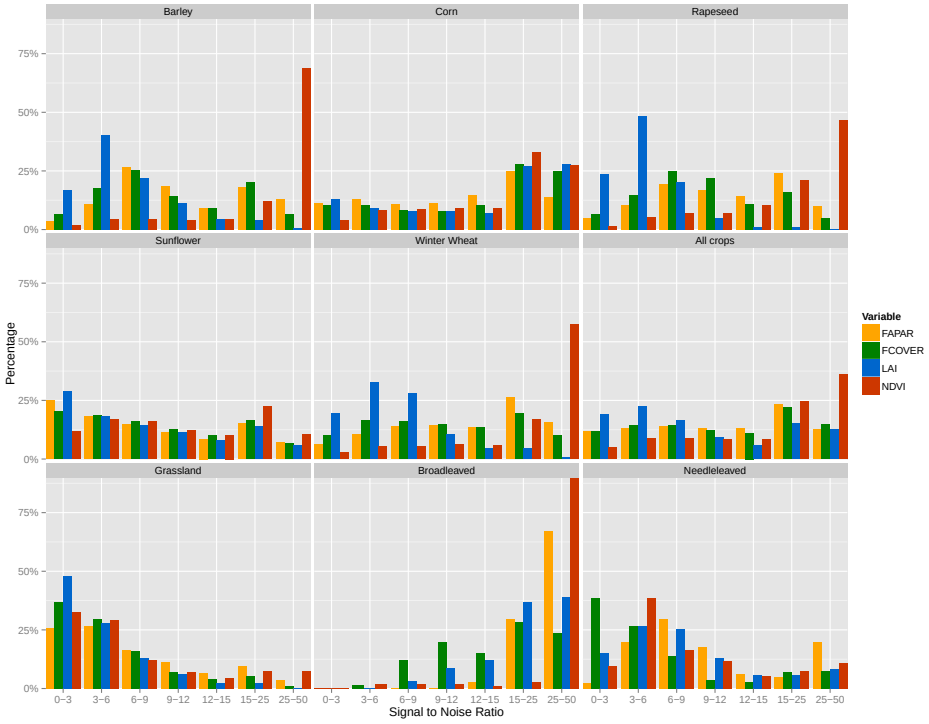


Figure 4. Distribution of the signal-to-noise ratio by land cover and crop types and by input data. NDVI shows higher ratios, which indicate more temporal consistency. Ratios between biophysical variables and NDVI particularly differ for the winter crop classes. Sunflower appears less affected by noise.

Distributions of the class-specific signal-to-noise ratio were extracted at the pixel level for all biophysical variables and the NDVI (Figure 4) and show contrasting distribution patterns: from uniform distributions (sunflower) to growing or decreasing exponential distributions (broadleaf forest and grassland) and a normal (barley) distribution. Winter crops tend to have high ratio values for NDVI, and yet, distributions of biophysical variables are associated with lower values, especially LAI. All biophysical variables, as well as the NDVI have similar distributions for summer crops and forest classes. Regardless of the class, the distribution of NDVI signal-to-noise ratio values is shifted towards higher values and, hence, shows more consistency than biophysical variables. This is particularly remarkable for the winter wheat and the barley classes. LAI is the noisiest BPV for crops with a signal-to-noise ratio of nine in comparison to 13 for FAPAR and FCOVER. The combined analysis of the temporal profiles and the signal-to-noise ratio reveals that the retrieval for corn is the most reliable.

4.2. Classification Results

Six random forest classifiers were trained along the season with the spectral bands (Bands-C), NDVI (NDVI-C) and three biophysical variables separately (LAI-C, FAPAR-C and FCOVER-C) and jointly (BPV-C). End of season maps show akin general patterns for the six classifiers. Nevertheless, a salt and pepper effect is more visible on maps derived from biophysical variables than on the NDVI-C and Bands-C maps (Figure 5). Accuracy indicators were extracted form confusion matrices; in particular, overall classification accuracies, producer's and user's accuracies, F_1-score and kappa coefficients were calculated for all classes (Table 3) and for crop types only (Table 4).

Bands-C provides the most accurate classification accuracy followed by NDVI-C and the biophysical variables (85% over 73%–82%). Less trivial is that NDVI (82%) outperformed the biophysical variables taken separately or jointly (about 75%). This is most probably because it is more efficient over urban, water and forest types. For all land covers, the users's and producers's accuracy indicators are generally in the range of 75%–95%.

In all cases, the barley class presents poor accuracy indicators, which might be explained by its spectral, structural and temporal similarity to wheat. One can also observe a drop in accuracy for the urban and water classes for the biophysical variable cases. This is expected, as those variables are less meaningful over such land cover classes. The accuracy of some classes, such as corn, was not significantly affected by the change in satellite input data. All classifications were found significantly different from one another (z-test > 1.96). These performances ought to be related to the temporal consistency of the input time series.

Assuming that the cropland extent is known, accuracy indicators were computed for the five crop types solely. All six classifications improved in overall accuracy, especially for those based on biophysical variables (+7%; see Table 3): the difference in accuracy is diminished two-fold. An overall accuracy of 84% is reached with FCOVER-C compared to 87% with NDVI-C (Table 4). FCOVER appears as the best biophysical variable for land cover and crop type classification with similar accuracies compared to BPV-C. The main confusion errors are encountered with classes such as barley-winter wheat (at least 20%), barley-grassland (at least 12%), barley-rapeseed (at least 9%) and corn-sunflower (at least 9%). Such classification errors were expected because of their similar temporal profiles. NDVI-C,

Bands-C and FAPAR-C show more difficulties to distinguish broadleaved and needle-leaved forests than FCOVER-C and LAI-C (29%, 34%, 48% of erroneous predictions). Corn is the crop type with the highest F_1-score for all classifiers and the lowest difference between classifiers.

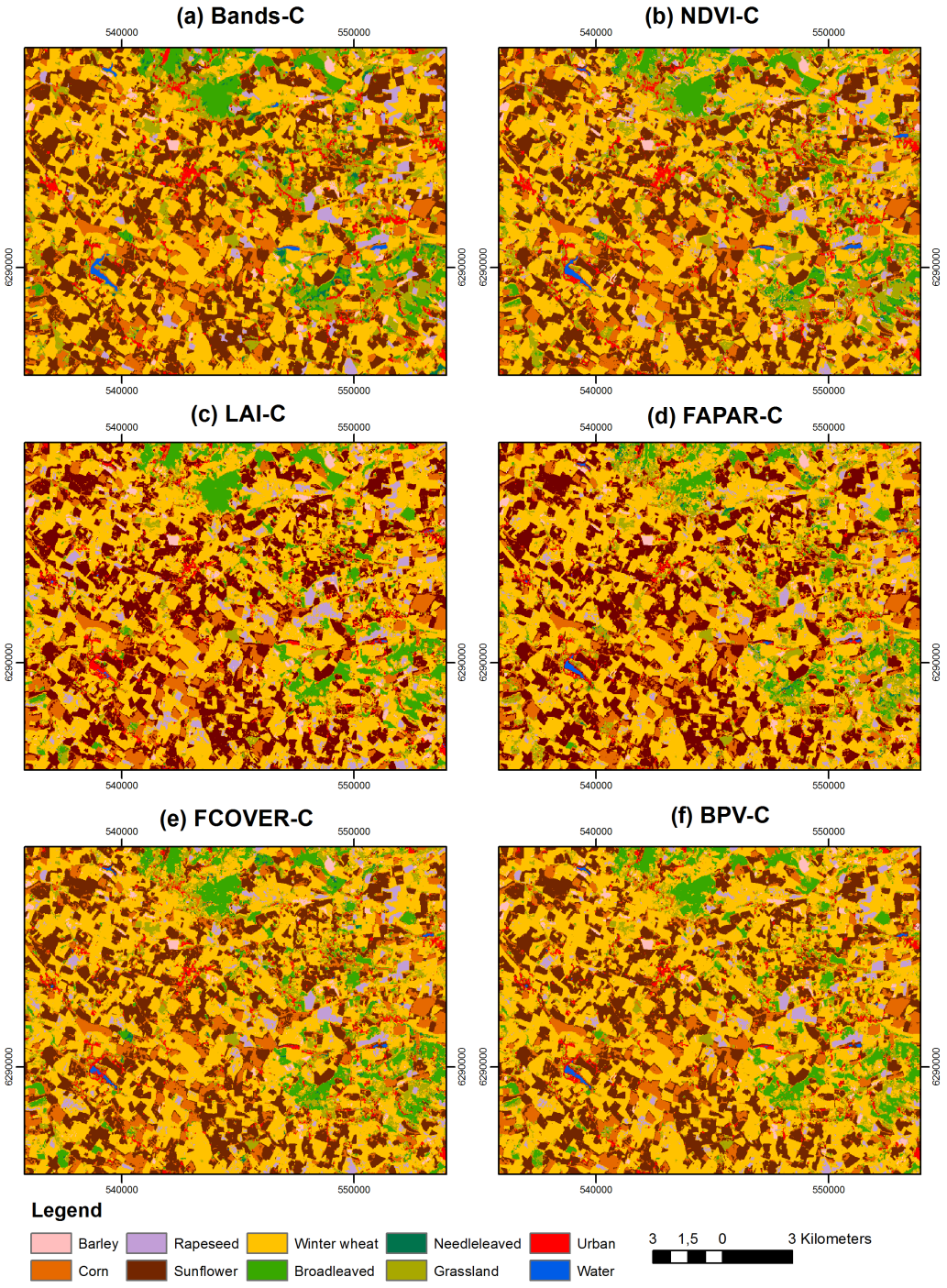


Figure 5. End-of-season maps for the six classifiers: (**a**) Band-C; (**b**) NDVI-C; (**c**) LAI-C; (**d**) FAPAR-C; (**e**) FCOVER-C; (**f**) BPV-C. Bands-C provides the most accurate classification accuracy, followed by NDVI-C and the biophysical variables. The pepper and salt is more visible on maps derived from biophysical variables than on the NDVI-C and Bands-C maps. This ought to be related to the lower SNR observed for biophysical variables.

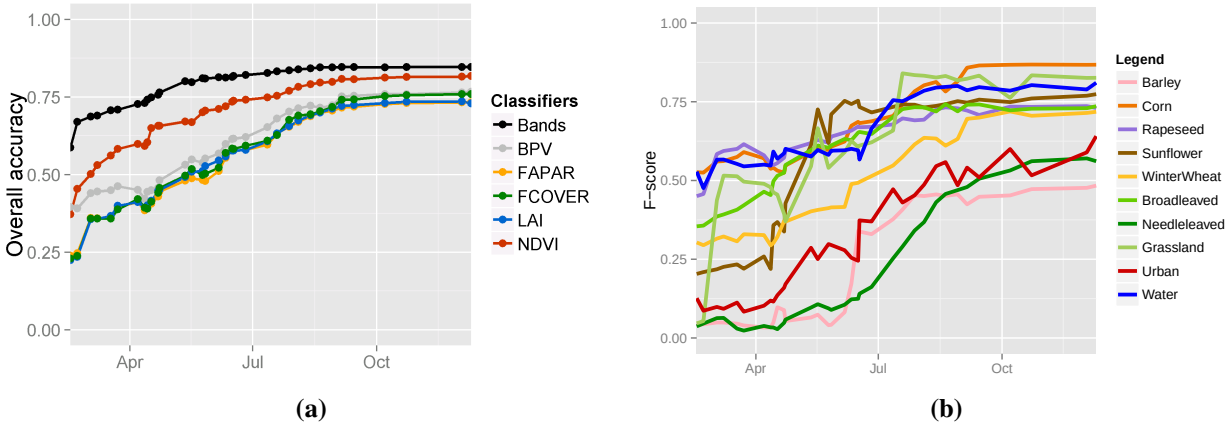
Table 3. Producer's accuracy (PA), user's accuracy (UA) and overall accuracy (OA) obtained at the end of the season with the six classifiers and for the eleven land cover and crop type classes, namely Barley, Corn, Rapeseed, Sunflower, Winter wheat (W. wheat), Broadleaved Forest (Broadleaved F.), Needleleaved forest (Needleleaved F.), Grassland, Urban and Water.

		Barley	**Corn**	**Rapeseed**	**Sunflower**	**W. wheat**	**Broadleaved F.**	**Needleleaved F.**	**Grassland**	**Urban**	**Water**	**OA**
Bands-C	PA	27.5	91.7	87.1	74.9	90.9	78.5	89.1	78.1	96.5	100	84.6
	UA	52.7	87.2	92.8	84.6	79.4	93.2	64.9	86	84.2	99.3	
NDVI-C	PA	42.2	90.3	72.7	72.4	86.5	84.4	77.8	76	90.9	98.9	81.8
	UA	59.9	86	89.7	80.8	77.7	92.8	70.4	79.2	73.3	99.8	
FAPAR-C	PA	38.2	89.3	49.8	70	78.1	35.6	52.9	69.4	54.6	95.9	73.2
	UA	51.9	81.9	67	74.7	64.5	86.3	50.7	71.3	75.8	94.8	
LAI-C	PA	44.7	89.4	64	70.6	73	23	61.9	69.2	58.3	36.2	73.0
	UA	49.7	81.6	69.1	74.9	66.2	76.4	86.1	69.2	64.3	43.1	
FCOVER-C	PA	36.2	89.9	67.2	68.6	80.1	58.6	49.5	71	51.7	96.1	75.9
	UA	58.1	83.2	83.3	76.8	68.9	81.1	93	81	66.5	20.3	
BPV-C	PA	43.3	90.3	72	70.9	81.1	49.2	68.4	70.1	45	97.9	76.7
	UA	54.8	83.5	83.7	72.8	67.4	91.7	99.4	76.2	74.6	71.4	

Table 4. Accuracy indicators obtained at the end of the season with the six input datasets for the agricultural classes only assuming that the the cropland is already known (producer's accuracy (PA), user's accuracy (UA) and overall accuracy (OA)).

		Barley	**Corn**	**Rapeseed**	**Sunflower**	**Winter Wheat**	**OA** (%)
	PA (%)	29.3	94.7	90.2	76.7	94.3	88.6
Bands-C	UA (%)	61.5	90.7	94.8	87.3	85.5	
	F_1-score	39.7	92.7	92.4	81.7	89.7	
	PA (%)	45.2	93.9	79.1	75.7	93.2	87.2
NDVI-C	UA (%)	67.8	89.9	93.2	84.8	84.3	
	F_1-score	54.3	91.9	85.6	80	88.5	
	PA (%)	41.0	93.1	52.6	71.9	84.4	80.9
FAPAR-C	UA (%)	58.7	88.1	69.2	82.6	74.1	
	F_1-score	48.3	90.5	59.8	76.9	78.9	
	PA (%)	47.9	92.9	67.7	72.9	79.3	81.6
LAI-C	UA (%)	57.1	88.8	70.3	80.1	77.5	
	F_1-score	54.1	90.8	68.9	76.3	78.4	
	PA (%)	38.7	94.1	71.1	70.7	89.2	84.2
FCOVER-C	UA (%)	68.5	87.9	84.3	86.1	78.3	
	F_1-score	49.4	90.9	77.1	77.6	83.4	
	PA (%)	45.1	93.9	74.9	72.5	86.2	84.3
BPV-C	UA (%)	63.5	88.7	85.0	82	80.3	
	F_1-score	52.8	91.2	79.6	76.9	83.1	

At the beginning of the season, Bands-C reaches an accuracy of 58% compared to 38% for BPV-c and NDVI-C (Figure 6a). Along the season, all classification schemes increase at a similar rate to reach a saturation plateau. The main difference between Bands-C and the other classifications is that it reaches a higher plateau (83%) and sooner (July). At the end of the season, Bands-C yields the best accuracy (85%) and delimits the upper accuracy bound: by reducing the spectral information to structural biophysical variables, the separability decreases. BPV-C outperforms NDVI-C when only one image is available (Figure 6). Even if NDVI saturates at high LAI values, it seems that the increased intra-class signal variability resulting from the retrieval prevents the classification based on biophysical variables to outperform those using simply the NDVI. By reducing the spectral information to structural biophysical variables, the separability decreases early in the season. Looking at class-specific accuracy measures (Figure 6b), one can distinguish three groups. The first group gathers land cover classes (corn, broadleaved forest, water, rapeseed), always well classified with respect to the reachable accuracy. The second group includes the classes requiring a longer time series to reach the accuracy level of the first group (winter wheat, grassland and sunflower). Finally, the third group concerns those classes that were not separable early in the season and whose accuracies flatten out at around 0.5 at the end of the season (urban, barley and needle-leaved forest).



(a) **(b)**

Figure 6. Evolution of the accuracy along the season. **(a)** Evolution of the overall accuracy over time. **(b)** Evolution of the F_1-score over time for the FCOVER classification.

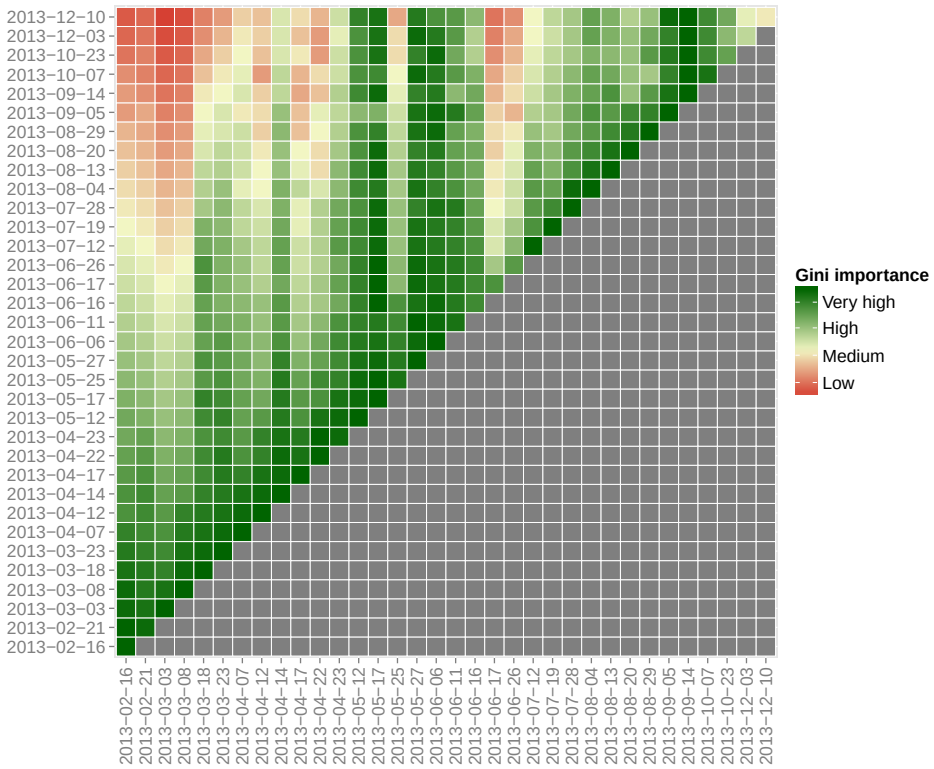
4.3. Importance of the Date and of the Length of the Time Series

To identify the key dates for classification, the Gini Index importance was extracted all along the season random forest classifier for FCOVER. The Gini coefficients were grouped into four categories of importance: very high, high, medium and low importance. For a specific date (y-axis), Figure 7a gives the Gini importance of all of the previous date until the specific date (x-axis) for the FCOVER classification. This demonstrates the impact of a new acquisition date on the importance of all previous dates. Note that green values correspond to high importance, whereas yellowish/reddish colors indicates lower importance. Two separability windows emerge: one in May–June and another in end of July–August. To confirm that, Gini Index importance was computed at the end of the season (last date: 10 December 2013) for all three variables, as well as for NDVI (Figure 7b). The two separability windows are identified and appeared consistent with the seasonality of the study area. These separability windows occurred at the time of maximum FCOVER for winter crops (minimum for summer crops) and the maximum of FCOVER for summer crops (minimum for winter crops). This provides evidence for supporting that the temporal trajectory matters more than the actual biophysical variable measurement. Besides, there was no correlation between the variable importance and its fraction of cloud-free pixels ($R^2 = 0.0032$).

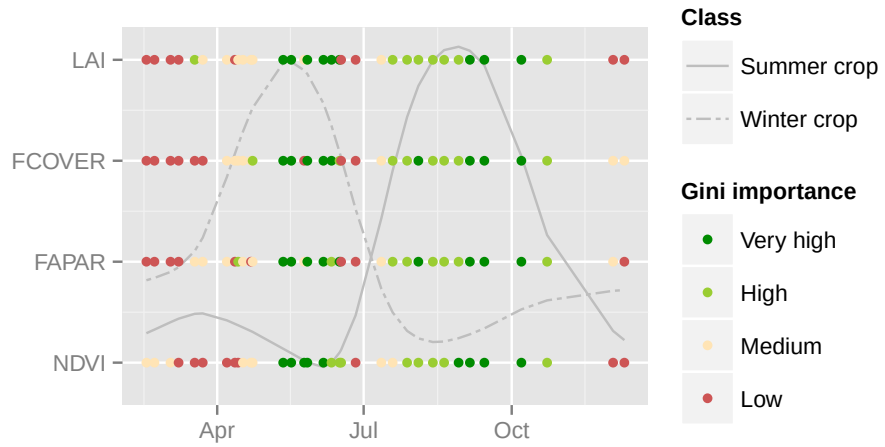
Classifications were then produced from the FCOVER time series using only the dates of the two first classes of Gini importance. When very high Gini importance classes are used (nine dates), the overall accuracy reaches 74.9%. When combining very high and high Gini importance (17 dates), the overall accuracy rises to 77.4%. Therefore, using the two most important classes of date importance outperforms the classification with all dates (75.9%), which coincides with that obtained on the overall accuracy evolution. This is most certainly due to a reduction of the noise by the selection of the best acquisition dates (quality of observation) or the most discriminant date of observation.

To evaluate the importance of the length of the time series, ten random date combinations with lengths ranging from 1–33 dates were selected and assessed (Figure 8). The distributions of the overall accuracy for different lengths of FCOVER time series underline that less than nine dates yields poor ($<$70%) and highly variable accuracy measures. On the contrary, with more than nine dates, the distribution tightens

and reaches an accuracy of less than 5% of the highest overall accuracy. It seems that adding dates affects the spread of the overall accuracy distribution itself more than the accuracy level.



(a)



(b)

Figure 7. Importance of the variables in FCOVER-Cs. (**a**) Importance of features (acquisition dates) along the season to the overall accuracy of classification. From bottom to top are images accumulated along the season, while from left to right, the importance of each new date is assessed. Note that the top line in this graph corresponds to the FCOVER line in Figure 7b. (**b**) Measures of the importance of the different dates in the classification accuracy. The FCOVER temporal profiles are presented with the dotted line for winter crop and with the continuous line for summer crop.

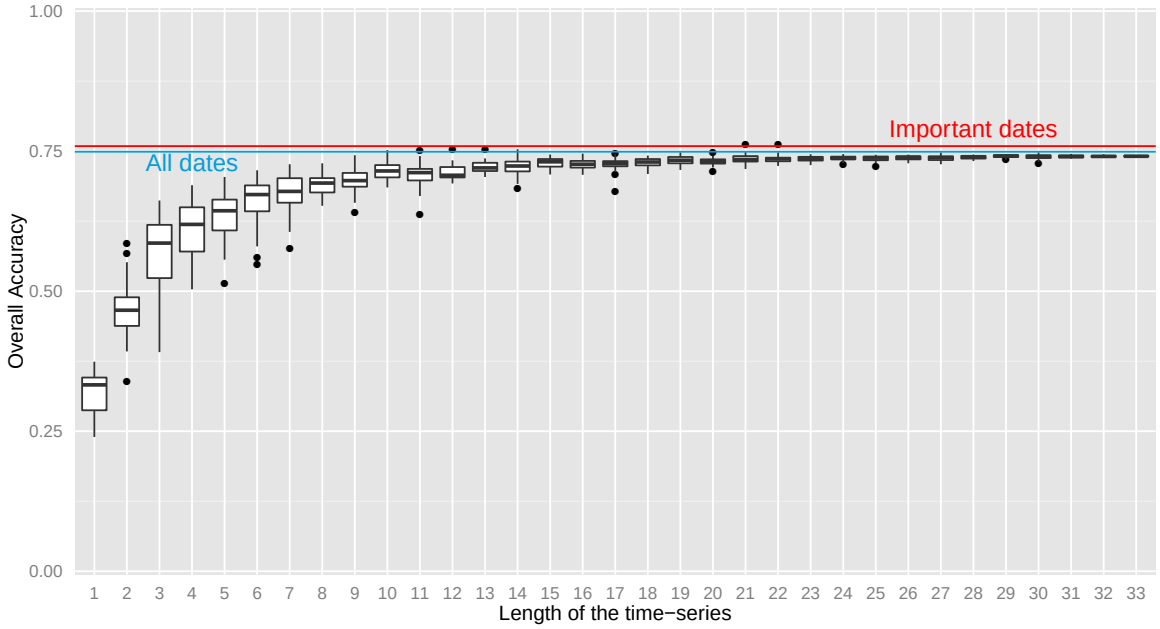


Figure 8. Impact of the number of variables on the FCOVER-C accuracy. For time series lengths ranging from 1–33, ten combinations were systematically and randomly selected. For each random subset, a classifier was trained and its accuracy evaluated. With more than nine dates, the increase in accuracy resulting from adding more dates flattens.

5. Discussion

The biophysical variable retrieval with neural network inversion was achieved with an RMSE of 0.49, 0.1 and 0.15 for LAI, FAPAR and FCOVER, respectively. These results are close to the GCOS accuracy requirements that recommend ±0.5 for LAI and ±0.05 for FAPAR. The range of 15%–20% is regarded by Baret *et al.* [68] as the currently achievable accuracy for LAI from remote sensing observations. Claverie *et al.* [69] inverted PROSAIL with neural networks on 105 very high resolution images (Formosat-2, 8 m) acquired in southwest France nearby the area of interest of this study: 90% of the LAI estimates and 78% of the FAPAR estimates were within the GCOS recommendation for accuracy (RMSE of 0.35 and 0.07, respectively). Combining Landsat-8 and HJ1 observations, Zhao *et al.* [70] retrieved LAI by means of look-up tables with an RMSE of 0.74 and 1.01, respectively. They demonstrated that averaging the LAI estimates in a 10-day LAI product after a pixel quality assessment improved the estimation (RSME = 0.42) compared to single-product estimates. The time series used in this study are similar to those that will be available when combining Sentinel-2 and Landsat-8 in terms of temporal frequencies. Sentinel-2 provides a great opportunity for global vegetation monitoring due to its enhanced spatial, spectral and temporal characteristics compared to SPOT [71], and its refined spectral bands could improve the retrieval accuracy. Richter *et al.* [72] retrieved LAI from simulated Sentinel-2 imagery and achieved crop-specific retrieval with an error of 16%. To reach a higher retrieval accuracy, Richter *et al.* [72] suggest that improvements could be achieved by employing spatial [73,74] and/or temporal information [75,76]. Other suggestions for improvements refer to the addition of prior

information on the type of canopy [77] to fine-tune the retrieval procedure or to select the retrieval strategy to apply [78].

From the classification accuracy point of view, classifications using spectral information systematically yielded the highest accuracies (84%). Results showed that using only structural biophysical variables, one could classify land cover with an accuracy of above 70% in August, 90% of the accuracy level reached with all of the spectral information. When considering that the spatial distribution of cropland is known, the performances of the classifications relying on biophysical variables reach 81%–84% compared to 88% for Bands-C. This ought to be related to the accuracy obtained (Kappa = 0.87) by Petitjean *et al.* [79] when classifying Formosat-2 time series in southwest France following a dynamic time warping approach. NDVI-C was found to consistently outperform classifications based on biophysical variables. Two main factors could explain these results. First, biophysical variables are less meaningful over certain land covers, which would consequently reduce class separability. However, for more dynamic vegetation classes, such as barley, the classification accuracy was improved. Second, the biophysical variable retrieval acted as an additional source of noise. Despite high RMSE and R^2, the larger variance of the biophysical time series and the signal-to-noise ratio analysis showed that the temporal consistency of the biophysical variables was reduced compared to NDVI. The retrieval effect on the temporal consistency was both class specific and variable specific. LAI was the variable most affected by noise and provided the classifications with the lowest accuracy. Despite an RMSE of 0.15, FCOVER appeared to have a high temporal consistency (low SNR) and reached in general the highest accuracy. For this reason, this study suggests that biophysical variable retrieval studies not only quantify the error in the inversion, but also provide measures of temporal consistency of the retrieved variables. Finally, classes with a high SNR ratio showed similar accuracy when classifying with spectral information, NDVI and biophysical variables. This is typically the case of corn, which had both a low RMSE and a high SNR.

The contribution of a specific date to the classification varies significantly along the season. These variations are related to phenology, rather than to cloud percentage. When considering the entire time series, critical dates occur at the peaks of the winter and summer crops. Focusing on the length of the time series regardless of the date, nine dates were sufficient to reach 95% of the highest accuracy. By reducing the spectral information to structural biophysical information, it appeared that classifiers trained on the variables are more prone to errors, especially early in the season. Denser time series would currently not improve the classification accuracy because (1) more subtle signal variations could not be captured with the current accuracy and consistency of the retrieval and (2) the added value after nine dates is marginal. However, if the temporal density of the time series improves jointly with its consistency, finer vegetation changes could be exploited to improve the classification. One could also consider to increase the density of the time series by combining optical data with radar (such as in [80] and [4]), which would be particularly useful in cloud-prone areas.

Including other biophysical variables, such as the plant pigments—chlorophyll, anthocyanin and carotenoids—could thus improve the classification accuracy. Leaf pigment concentration, such as chlorophyll, would add complementary information for the classification and, therefore, new possible windows of separability. All bands around the red-edge have been shown to be useful in assessing

vegetation condition, specifically canopy chlorophyll content. For example, Verrelst *et al.* [81] tested a machine learning regression algorithm for biophysical variable retrieval and reached the 10% precision required by end-users in the estimation of chlorophyll.

6. Conclusions

Using biophysical variables as a common unit to combine multi-sensor time series is a sensible way to increase the observation frequency and to foster the re-utilization of land cover and crop type ground data. Currently, more than ten international high spatial resolution resolution satellites, such as Landsat, SPOT, China-Brazil Earth Resources Satellite (CBERS) and the Huan Jing 1 (HJ1) satellite, could be combined to increase the temporal frequency. This paper proposes to exploit high spatial and temporal resolution multi-sensor time series for land cover classification with a focus on agricultural classes, as they are more dynamic. Based on a radiative transfer model and neural networks, LAI, white-sky FAPAR, black-sky FAPAR and FCOVER were retrieved with an RMSE of 0.49, 0.1, 0.12, 0.15 units, respectively. The retrieval was coherent between sensors and the average biophysical time series matching the temporal patterns of the considered classes. However, the pixel-level signal-to-noise ratio decreased for biophysical variables, especially LAI, compared to NDVI. Classifications with structural biophysical variables reached end-of-season overall accuracies ranging from 73%–77% compared to 84 and 82% with the spectral bands and NDVI, respectively. The additional noise observed in the biophysical variable time series substantially affected the classification accuracy: classes, such as corn, with a high signal-to-noise ratio and a low RMSE, achieved similar accuracies with biophysical variables or with spectral bands. FCOVER appears to be the most promising biophysical variable for land cover and crop type classification. Selecting observations at key phenological events (peaks of vegetation for winter and summer crops) and focusing on crop classes that are more dynamic and for which biophysical variables are more meaningful represent two ways of increasing the classification accuracy. Further developments of the method would include providing smoother biophysical variable time series, as well as integrating biophysical variables on the plant pigments, such as chlorophyll, jointly might improve the classification.

Acknowledgments

This research was carried out within the ImagineS (Implementing Multi-scale Agricultural INdicators Exploiting Sentinels) project in the FP7-SPACE-2012-1 Grant Agreement Number 311766. The Landsat data were obtained through the online Data Pool at the NASA Land Processes Distributed Active Archive Center (LP 440 DAAC), USGS/Earth Resources Observation and Science (EROS) Center, 441 Sioux Falls, South Dakota (https://lpdaac.usgs.gov/get_data). The SPOT4 imagery was obtained under the SPOT4/Take5 program. Imagery is copyrighted to CNES under the mention: "CNES 2013, all rights reserved. Commercial use of the product prohibited".

Author Contributions

François Waldner and Marie-Julie Lambert designed and carried out the study under the supervision of Frédéric Baret and Pierre Defourny. Olivier Hagolle provided corrected satellite images time-series.

Wenjuan Li, Marie Weiss and Frédéric Baret extracted the high resolution biophysical variables. All authors participated in the writing.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Hu, C.; Li, J.; Chen, N.; Guan, Q. An object model for integrating diverse remote sensing satellite sensors: A case study of Union Operation. *Remote Sens.* **2014**, *6*, 667–699.
2. 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.
3. Whitcraft, A.K.; Becker-Reshef, I.; Killough, B.D.; Justice, C.O. Meeting earth observation requirements for global agricultural monitoring: An evaluation of the revisit capabilities of current and planned moderate resolution optical earth observing missions. *Remote Sens.* **2015**, *7*, 1482–1503.
4. Wang, Y.C.; Feng, C.C.; Vu Duc, H. Integrating multi-sensor remote sensing data for land use/cover mapping in a tropical moutainous area in Northern Thailand. *Geogr. Res.* **2012**, *50*, 320–331.
5. Baret, F.; Weiss, M.; Lacaze, R.; Camacho, F.; Makhmara, H.; Pacholcyzk, P.; Smets, B. GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production. *Remote Sens. Environ.* **2013**, *137*, 299–309.
6. Chen, J.M.; Black, T. Defining leaf area index for non-flat leaves. *Plant Cell Environ.* **1992**, *15*, 421–429.
7. Asrar, G.; Fuchs, M.; Kanemasu, E.; Hatfield, J. Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat. *Agron. J.* **1984**, *76*, 300–306.
8. Baret, F.; Guyot, G. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sens. Environ.* **1991**, *35*, 161–173.
9. Richardson, A.; Wiegand, C.; Wanjura, D.; Dusek, D.; Steiner, J. Multisite analyses of spectral-biophysical data for sorghum. *Remote Sens. Environ.* **1992**, *41*, 71–82.
10. Duchemin, B.; Hadria, R.; Erraki, S.; Boulet, G.; Maisongrande, P.; Chehbouni, A.; Escadafal, R.; Ezzahar, J.; Hoedjes, J.; Kharrou, M.; *et al.* Monitoring wheat phenology and irrigation in Central Morocco: On the use of relationships between evapotranspiration, crops coefficients, leaf area index and remotely-sensed vegetation indices. *Agric. Water Manage.* **2006**, *79*, 1–27.
11. Dorigo, W.A.; Zurita-Milla, R.; de Wit, A.; Brazile, J.; Singh, R.; Schaepman, M.E. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. *Int. J. App. Earth Obs. Geoinf.* **2007**, *9*, 165–193.

12. McCallum, I.; Wagner, W.; Schmullius, C.; Shvidenko, A.; Obersteiner, M.; Fritz, S.; Nilsson, S.; Satellite-based terrestrial production efficiency modeling. *Carbon Balance Manag.* **2009**, *4*, 1–14.
13. Avissar, R.; Pielke, R.A. A parameterization of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology. *Mon. Wea. Rev.* **1989**, *117*, 2113–2136.
14. Hagolle, O.; Huc, M.; Villa Pascual, D.; Dedieu, G. A Multi-Temporal and Multi-Spectral Method to Estimate Aerosol Optical Thickness over Land, for the Atmospheric Correction of FormoSat-2, LandSat, VENμS and Sentinel-2 Images. *Remote Sens.* **2015**, *7*, 2668–2691.
15. Hagolle, O.; Dedieux, G.; Mougenot, B.; Debaecker, V.; Duchemin, B. Correction of aerosol effects on multi-temporal images acquired with constant viewing angles: application to Formosat-2 images. *Remote Sens. Environ.* **2008**, *112*, 1689–1701.
16. Wenjuan, L.; Weiss, M.; Waldner, F.; Demarez, V.; Hagolle, O.; Baret, F. Deriving LAI, FAPAR essential climate variables and FCOVER from SPOT and LANDSAT sensors : evaluation of the consistency and comparison with ground measurements. *Remote Sens.* **2015**, submitted.
17. Weiss, M.; Baret, F. *CAN-EYE V6.313 USER MANUAL*; INRA-EmmaH: Avignon, France, 2014.
18. Mougin, E.; Demarez, V.; Diawara, M.; Hiernaux, P.; Soumaguel, N.; Berg, A. Estimation of LAI, fAPAR and fCover of Sahel rangelands (Gourma, Mali). *Agric. For. Meteorol.* **2014**, *198*, 155–167.
19. Demarez, V.; Duthoit, S.; Baret, F.; Weiss, M.; Dedieu, G. Estimation of leaf area and clumping indexes of crops with hemispherical photographs. *Agric. For. Meteorol.* **2008**, *148*, 644–655.
20. Defourny, P.; Kirches, G.; Brockmann, C.; Boettcher, M.; Peters, M.; Bontemps, S.; Lamarche, C.; Schlerf, M.; Santoro, M. *Land Cover CCI : Product user Guide : Version 2*; Université catholique de Louvain: Louvain-la-Neuve, Belgium, 2014.
21. Verrelst, J.; Rivera, J.P.; Veroustraete, F.; noz Marí, J.M.; Clevers, J.G.; Camps-Valls, G.; Moreno, J. Experimental Sentinel-2 {LAI} estimation using parametric, non-parametric and physical retrieval methods - A comparison. *ISPRS J. Photogramm. Remote Sens.* **2015**, in press.
22. Rivera, J.P.; Verrelst, J.; Delegido, J.; Veroustraete, F.; Moreno, J. On the Semi-Automatic Retrieval of Biophysical Parameters Based on Spectral Index Optimization. *Remote Sens.* **2014**, *6*, 4927–4951.
23. Verrelst, J.; Muñoz, J.; Alonso, L.; Delegido, J.; Rivera, J.; Camps-Valls, G.; Moreno, J. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. *Remote Sens. Environ.* **2012**, *118*, 127–139.
24. De Wit, A.; d'Andrimont, R. *MOCCCASIN Comparaison of MERIS, MODIS and KMSS for Winter Wheat Growth Parameters Assessment*; Project Report for MOCCCASIN consortium; Alterra: Wageningen, The Netherlands, 2012.

25. Baret, F.; Hagolle, O.; Geiger, B.; Bicheron, P.; Miras, B.; Huc, M.; Berthelot, B.; Nino, F.; Weiss, M.; Samain, O.; Roujean, J.; Leroy, M. LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. Part 1 : Principles of the algorithm. *Remote Sens. Environ.* **2007**, *3*, 275–286.
26. Bacour, C.; Baret, F.; Béal, D.; Weiss, M.; Pavageau, K. Neural network estimation of LAI, fAPAR, fCover and LAIxCab, from top of canopy MERIS reflectance data: Principles and validation. *Remote Sens. Environ.* **2011**, *79*, 72–83.
27. Atzberger, C.; Darvishzadeh, R.; Immitzer, M.; Schlerf, M.; Skidmore, A.; le Maire, G. Comparative analysis of different retrieval methods for mapping grassland leaf area index using airborne imaging spectroscopy. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, in press.
28. Huemmrich, K. The GeoSail model: a simple addition to the SAIL model to describe discontinuous canopy reflectance. *Remote Sens. Environ.* **2001**, *3*, 423–431.
29. Fleming, S.W. Signal-to-noise ratios of geophysical and environmental time series. *Environ. Eng. Geosci.* **2010**, *16*, 389–399.
30. Duveiller, G.; Lopez-Lozano, R.; Cescatti, A. Exploiting the multi-angularity of the MODIS temporal signal to identify spatially homogeneous vegetation cover: A demonstration for agricultural monitoring applications. *Remote Sens. Environ.* **2015**, *166*, 61–77.
31. Jensen, J.R.; Garcia-Quijano, M.; Hadley, B.; Im, J.; Wang, Z.; Nel, A.L.; Teixeira, E.; Davis, B.A. Remote sensing agricultural crop type for sustainable development in South Africa. *Geocarto Int.* **2006**, *21*, 5–18.
32. Murthy, C.; Raju, P.; Badrinath, K. Classification of wheat crop with multi-temporal images: performance of maximum likelihood and artificial neural networks. *Int. J. Remote Sens.* **2003**, *24*, 4871–4890.
33. Civco, D.L. Artificial neural networks for land-cover classification and mapping. *Int. J. Geogr. Inf. Sci.* **1993**, *7*, 173–186.
34. Breiman, L.; Friedman, J.; Stone, C.J.; Olshen, R.A. *Classification and Regression Trees*; CRC press: Boca Raton, FL, USA, 1984.
35. Wardlow, B.D.; Egbert, S.L. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the US Central Great Plains. *Remote Sens. Environ.* **2008**, *112*, 1096–1116.
36. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 247–259.
37. Mathur, A.; Foody, G.M. Crop classification by support vector machine with intelligently selected training data for an operational application. *Int. J. Remote Sens.* **2008**, *29*, 2227–2240.
38. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140.
39. Hansen, L.K.; Salamon, P. Neural network ensembles. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 993–1001.
40. Huang, C.; Davis, L.; Townshend, J. An assessment of support vector machines for land cover classification. *Int. J. Remote Sens.* **2002**, *23*, 725–749.

41. Rogan, J.; Franklin, J.; Stow, D.; Miller, J.; Woodcock, C.; Roberts, D. Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. *Remote Sens. Environ.* **2008**, *112*, 2272–2283.
42. Foody, G.M. Land cover classification by an artificial neural network with ancillary information. *Int. J. Geogr. Inf. Syst.* **1995**, *9*, 527–542.
43. Friedl, M.A.; Brodley, C.E. Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* **1997**, *61*, 399–409.
44. Foody, G.M.; Mathur, A. A relative evaluation of multiclass image classification by support vector machines. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1335–1343.
45. Ghimire, B.; Rogan, J.; Miller, J. Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic. *Remote Sens. Lett.* **2010**, *1*, 45–54.
46. Kotsiantis, S.; Pintelas, P. Combining bagging and boosting. *Int. J. Comput. Intell.* **2004**, *1*, 324–333.
47. Breiman, L. Random forest. *Machine Learning* **2001**, *45*, 5–32.
48. Rodriguez-Galiano, V.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104.
49. Vieira, M.A.; Formaggio, A.R.; Rennó, C.D.; Atzberger, C.; Aguiar, D.A.; Mello, M.P. Object Based Image Analysis and Data Mining applied to a remotely sensed Landsat time-series to map sugarcane over large areas. *Remote Sens. Environ.* **2012**, *123*, 553–562.
50. Stuckens, J.; Coppin, P.; Bauer, M. Integrating contextual information with per-pixel classification for improved land cover classification. *Remote Sens. Environ.* **2000**, *71*, 282–296.
51. Dorren, L.K.; Maier, B.; Seijmonsbergen, A.C. Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification. *For. Ecol. Manage.* **2003**, *183*, 31–46.
52. Dingle Robertson, L.; King, D.J. Comparison of pixel-and object-based classification in land cover change mapping. *Int. J. Remote Sens.* **2011**, *32*, 1505–1529.
53. 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.
54. 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 agrosystems globally distributed using spatial and temporal high resolution time series. Remote Sens. **2015**, submitted.
55. Yan, G.; Mas, J.F.; Maathuis, B.; Xiangmin, Z.; Van Dijk, P. Comparison of pixel-based and object-oriented image classification approachesâĂŤa case study in a coal fire area, Wuda, Inner Mongolia, China. *Int. J. Remote Sens.* **2006**, *27*, 4039–4055.
56. Whiteside, T.G.; Boggs, G.S.; Maier, S.W. Comparing object-based and pixel-based classifications for mapping savannas. *Int. J. Appl. Earth Obs. Geoinf.* **2011**, *13*, 884–893.

57. Millard, K.; Richardson, M. On the Importance of Training Data Sample Selection in Random Forest Image Classification: A Case Study in Peatland Ecosystem Mapping. *Remote Sens.* **2015**, *7*, 8489–8515.
58. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57.
59. Foody, G.M. Classification accuracy comparison: hypothesis tests and the use of confidence intervals in evaluations of difference, equivalence and non-inferiority. *Remote Sens. Environ.* **2009**, *113*, 1658–1663.
60. Stehman, S. Estimating the kappa coefficient and its variance under stratified random sampling. *Photogramm. Eng. Remote Sens.* **1996**, *62*, 401–407.
61. Agresti, A. *Categorical Data Analysis*. John Wiley and Sons: Hoboken, NJ, USA, 2002.
62. Fleiss, J.L.; Cohen, J.; Everitt, B. Large sample standard errors of kappa and weighted kappa. *Psychol. Bull.* **1969**, *72*, 323.
63. Rosenfield, G.H.; Fitzpatrick-Lins, K. A coefficient of agreement as a measure of thematic classification accuracy. *Photogramm. Eng. Remote Sens.* **1986**, *52*, 223–227.
64. Nitze, I.; Barrett, B.; Cawkwell, F. Temporal optimization of image acquisition for land cover classification with Random Forest and {MODIS} time-series. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *34*, 136–146.
65. Ghosh, A.; Sharma, R.; Joshi, P. Random forest classification of urban landscape using Landsat archive and ancillary data: Combining seasonal maps with decision level fusion. *Appl. Geogr.* **2014**, *48*, 31–41.
66. Rodriguez-Galiano, V.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104.
67. GCOS. *Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC*; GCOS-138 (GOOS-184, GTOS-76, WMO-TD/No. 1523); WMO: Geneva, Switzerland, 2004; pp. 1–24.
68. Baret, F.; Neale, C.; Maltese, A. Biophysical vegetation variables retrieval from remote sensing observations. *Proc. SPIE* **2010**, *7824*, 17–19.
69. Claverie, M.; Vermote, E.F.; Weiss, M.; Baret, F.; Hagolle, O.; Demarez, V. Validation of coarse spatial resolution {LAI} and {FAPAR} time series over cropland in southwest France. *Remote Sens. Environ.* **2013**, *139*, 216 – 230.
70. Zhao, J.; Li, J.; Liu, Q.; Fan, W.; Zhong, B.; Wu, S.; Yang, L.; Zeng, Y.; Xu, B.; Yin, G. Leaf Area Index Retrieval Combining HJ1/CCD and Landsat8/OLI Data in the Heihe River Basin, China. *Remote Sens.* **2015**, *7*, 6862–6885.
71. Frampton, W.J.; Dash, J.; Watmough, G.; Milton, E.J. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 83–92.
72. Richter, K.; Hank, T.B.; Vuolo, F.; Mauser, W.; D'Urso, G. Optimal exploitation of the Sentinel-2 spectral capabilities for crop leaf area index mapping. *Remote Sens.* **2012**, *4*, 561–582.

73. Atzberger, C. Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models. *Remote Sens. Environ.* **2004**, *93*, 53–67.
74. Atzberger, C.; Richter, K. Spatially constrained inversion of radiative transfer models for improved LAI mapping from future Sentinel-2 imagery. *Remote Sens. Environ.* **2012**, *120*, 208–218.
75. Koetz, B.; Baret, F.; Poilvé, H.; Hill, J. Use of coupled canopy structure dynamic and radiative transfer models to estimate biophysical canopy characteristics. *Remote Sens. Environ.* **2005**, *95*, 115–124.
76. Bach, H.; Mauser, W. Methods and examples for remote sensing data assimilation in land surface process modeling. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1629–1637.
77. Combal, B.; Baret, F.; Weiss, M.; Trubuil, A.; Mace, D.; Pragnere, A.; Myneni, R.; Knyazikhin, Y.; Wang, L. Retrieval of canopy biophysical variables from bidirectional reflectance: Using prior information to solve the ill-posed inverse problem. *Remote Sens. Environ.* **2003**, *84*, 1–15.
78. Yin, G.; Li, J.; Liu, Q.; Fan, W.; Xu, B.; Zeng, Y.; Zhao, J. Regional Leaf Area Index Retrieval Based on Remote Sensing: The Role of Radiative Transfer Model Selection. *Remote Sens.* **2015**, *7*, 4604–4625.
79. Petitjean, F.; Inglada, J.; Gançarski, P. Satellite image time series analysis under time warping. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3081–3095.
80. Dusseux, P.; Corpetti, T.; Hubert-Moy, L.; Corgne, S. Combined Use of Multi-Temporal Optical and Radar Satellite Images for Grassland Monitoring. *Remote Sens.* **2014**, *6*, 6163–6182.
81. Verrelst, J.; Rivera, J.P.; Moreno, J.; Camps-Valls, G. Gaussian processes uncertainty estimates in experimental Sentinel-2 LAI and leaf chlorophyll content retrieval. *ISPRS J. Photogramm. Remote Sens.* **2013**, *86*, 157–167.