Evaluation of using sentinel-1 and -2 time-series to identify winter land use in agricultural landscapes
J. Denize, Laurence Hubert-Moy, J. Betbeder, Samuel Corgne, J. Baudry, E. Pottier

To cite this version:

HAL Id: hal-02042533
https://hal-univ-rennes1.archives-ouvertes.fr/hal-02042533
Submitted on 26 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Evaluation of Using Sentinel-1 and -2 Time-Series to Identify Winter Land Use in Agricultural Landscapes

Julien Denize 1,*, Laurence Hubert-Moy 2, Julie Betbeder 3, Samuel Corgne 2, Jacques Baudry 4 and Eric Pottier 1

1 Institute of Electronics and Telecommunications of Rennes IETR, UMR CNRS 6164, University of Rennes, 35000 Rennes, France; eric.pottier@univ-rennes1.fr
2 Littoral-Environnement-Télédétection-Géomatique LETG UMR 6554, University of Rennes, 35 000 Rennes, France; laurence.hubert@univ-rennes2.fr (L.H.-M.); samuel.corgne@univ-rennes2.fr (S.C.)
3 Internal Research Unit Forests & Societies, Centre de Coopération Internationale en Recherche Agronomique pour le Développement CIRAD, 34 398 Montpellier, France; julie.betbeder@cirad.fr
4 L’unité mixte de recherche Biodiversité, AGroécologie et Aménagement du Paysage UMR BAGAP, Institut National De La Recherche Agronomique, INRA, 35 000 Rennes, France; jacques.baudry@inra.fr
* Correspondence: julien.denize@univ-rennes1.fr

Received: 20 November 2018; Accepted: 20 December 2018; Published: 27 December 2018

Abstract: Monitoring vegetation cover during winter is a major environmental and scientific issue in agricultural areas. From an environmental viewpoint, the presence and type of vegetation cover in winter influences the transport of pollutants to water resources. From a methodological viewpoint, characterizing spatio-temporal dynamics of land cover and land use at the field scale is challenging due to the diversity of farming strategies and practices in winter. The objective of this study was to evaluate the respective advantages of Sentinel optical and SAR time-series to identify land use in winter. To this end, Sentinel-1 and -2 time-series were classified using Support Vector Machine and Random Forest algorithms in a 130 km² agricultural area. From the classification, the Sentinel-2 time-series identified winter land use more accurately (overall accuracy (OA) = 75%, Kappa index = 0.70) than that of Sentinel-1 (OA = 70%, Kappa = 0.66) but a combination of the Sentinel-1 and -2 time-series was the most accurate (OA = 81%, Kappa = 0.77). Our study outlines the effectiveness of Sentinel-1 and -2 for identify land use in winter, which can help to change agricultural practices.

Keywords: agricultural monitoring; earth observing sensors; multi-temporal classification; optical and SAR time-series; Random Forest algorithm; Support Vector Machine algorithm

1. Introduction

Monitoring vegetation cover during winter is a major environmental and scientific issue in agricultural areas. From an environmental viewpoint, the presence and type of vegetation cover in winter influences the transport of pollutants to waterbodies by reducing the loss of nitrates, nutrients, pesticides or sediment from agricultural fields [1,2]. Lack of vegetation cover acts as an accelerator when soils are bare after a main crop (e.g., maize, rapeseed), while catch crops act as an obstacle to transport [3]. In this context, identifying and characterizing winter land use is a major component of water quality restoration and sustainable management in agricultural landscapes [4]. From a methodological viewpoint, characterizing spatio-temporal dynamics of land use and land cover (LULC) at the field scale is challenging due to the diversity of farming strategies and practices in winter. Identifying winter land use remains a major scientific challenge for the remote sensing community. While optical remotely sensed data are used mainly to determine annual LULC [5], they have several limitations for identifying land use in winter.
Time-series of medium-resolution remotely sensed data have been shown to be useful for classifying LULC over large areas [6,7] or for mapping and monitoring bare soils during winter in intensive agricultural regions [8,9]. However, due to their insufficient spatial resolution, these time-series can detect only patches of bare soil for winter land use at the field scale ([10]. High and very high spatial resolution remotely sensed data are also widely used to discriminate the main crop rotations and land uses at the field scale [11,12]. Refs. [13,14] demonstrated the potential of very high resolution remotely sensed data to monitor land use during summer. [15,16] showed the ability of high and very high spatial resolution images (Landsat and IKONOS) to identify crop residues during winter with high accuracy (overall accuracy (OA) > 80%). However, soil surface and vegetation growth conditions vary daily, seasonally and among fields [17]. Capturing these variations, which is necessary to identify land use types, requires acquiring several remotely sensed images during winter.

Optical remotely sensed data are under-exploited for identifying winter land use, mainly because few cloud-free images are available in winter for monitoring intra-annual dynamics of crops. Synthetic-Aperture Radar (SAR) data provide a reliable solution to address the limitations of optical images because they are not sensitive to atmospheric conditions and can be acquired during the day or night [18]. Thus, time-series of SAR images can be acquired to study intra-annual changes in vegetation [19]. SAR images have been used extensively to map land use in winter, especially to identify bare soil and tillage practices [20,21]. SAR time-series can also be used to identify inter-crops and crop residues. Refs. [22,23] demonstrated that vegetation structure and phenology is directly related to the backscattering mechanisms that occur between the SAR signal and land surfaces.

However, while optical and SAR data should be complementary, few studies have evaluated the combined use of optical and SAR time-series to identify and characterize land use [5,24,25] and, to our knowledge, none has done so for land use in winter. The development of Sentinel-1 and -2 sensors, which acquire optical and SAR data with high spatial and temporal resolutions, provides interesting opportunities to monitor winter land use. Until now, few studies have evaluated the use of Sentinel-1 (SAR) and Sentinel-2 (optical) time-series for monitoring land use, either separately or combined [26,27]. Refs. [28,29] identified land use classes during summer using Sentinel-2 time-series, with OA > 91%. However, only three studies [28–30] have used Sentinel-2 data to monitor intra-annual land use changes, mainly because few cloud-free images are available during winter. Ref. [30] has shown the potential of Sentinel-2 time-series compared to Landsat-8 and SPOT-5 time-series for detecting changes in LULC, with better results obtained with Sentinel-2 than the other data. Refs. [28,29] demonstrated the ability of Sentinel-2 single-date images and time-series to map land use changes during the growing season, with OA > 90%. Several studies evaluated the use of Sentinel-1 data for identifying and monitoring land cover during summer, with OA > 80% [31,32]. Ref. [33] classified Sentinel-1 time-series to map winter vegetation in five quality classes (“bare soil” to “high quality”) using a deep-learning approach, with OA > 98%. Only two studies investigated the potential of the use of combined Sentinel-1 and -2 time-series to identify land cover. Refs. [34,35] demonstrated that the combined use of Sentinel-1 and -2 time-series increased the OA by 5–10 percentage points. Nevertheless, the evaluation of using Sentinel-1 and -2 time-series to monitor winter land use remains unexplored.

The aim of this study was to evaluate the ability to use Sentinel-1 and -2 time-series to identify winter land use. To this end, a time-series of nine Sentinel-1 and -2 images acquired during a single hydrological year in autumn, winter and early spring was processed, first separately and then combined. Optical and SAR parameters, such as backscattering coefficients and vegetation indices, were extracted first. These parameters were then used to perform classifications with Random Forest (RF) and Support Vector Machine (SVM) algorithms [36,37].

2. Study Site and Data

2.1. Study Site

The “Zone Atelier Armorique” study site, a long-term ecological research (LTER) site in the LTER-Europe and ILTER networks, is located south of the Bay of Mont-Saint-Michel (48°31'0" N, 1°31'30" W), France (Figure 1). It was established in 1993 to assess relationships between changes in farming practices, landscape dynamics and ecological processes related to biodiversity, water quality and climate [38]. This agricultural area covers ca. 130 km$^2$ and has a temperate climate with an annual mean temperature of 12 °C, minimum mean temperature for the coldest year of 8 °C, maximum mean annual temperature of 16 °C and mean annual precipitation of 650 mm. The site contains ca. 7000 agricultural fields surrounded by a hedgerow network; field size ranges from 0.1-65 ha, with a mean of 2.1 ha. The crop system is characterized by a single crop planted per field each year. The main annual crops are maize, wheat, rapeseed and barley. In winter, catch crops are grown to prevent nitrogen leaching, in the framework of the Nitrate Directive of the European Union [39] and some of the catch crops are fed to cattle. In certain areas, soils are not completely covered by vegetation. Hence, it is important to locate such areas to advise farmers how to implement the best management practices.

![Figure 1. Location of the “Zone Atelier Armorique” study site and ground surveys (RGB composite image from Sentinel-1 data, 2016, ©Copernicus data 2016).SE: Shannon Entropy.](image)

2.2. Field Data

Land Use Data

The winter land use types studied are winter crops, grasslands, catch crops, crop residues and bare soil (Figure 2). Winter crops (wheat, barley and rapeseed) cover 40% of the utilized agricultural area (UAA) (Table 1). Wheat and barley sown in October have similar plant structure and phenology. Barley is harvested at the end of June or beginning of July, whereas wheat is harvested mainly in the middle of July. Rapeseed is sown mainly in September and harvested at the beginning of July. At full development, rapeseed plants are twice as tall (ca. 1.8 m) as wheat and barley and their stems are intertwined, with no clear vertical structure [27]. Grasslands, which are considered to have a major influence in regulating water flows and nutrient cycling [40], cover ca. 30% of the UAA. They are
mown or grazed, which explains their similar plant structure but different phenology. Catch crops, sown after the main crops (August–October), cover 25% of the UAA, are diverse (e.g., oat, phacelia, mustard) and show different plant structure and phenology. Although bare soils have been banned for several years to avoid soil erosion and water pollutant flows, they cover an average of 5% of the UAA. Residues of annual crops (cereal stubble or maize stalks) cover parts of fields.

![Figure 2. Main land use types encountered in winter in the study site: (a) crop residues (maize stalks), (b) bare soils, (c) winter crops (winter barley), (d) grasslands and (e) catch crops (mustard).](image)

<table>
<thead>
<tr>
<th>Table 1. Land use classification.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Winter Land Use Types</strong></td>
</tr>
<tr>
<td>Winter Crops</td>
</tr>
<tr>
<td>Grasslands</td>
</tr>
<tr>
<td>Catch crops Fed to cattle</td>
</tr>
<tr>
<td>Catch crops Not used</td>
</tr>
<tr>
<td>Crop residues Cereal stubble</td>
</tr>
<tr>
<td>Bare soils</td>
</tr>
</tbody>
</table>

We conducted ground surveys every ten days between November 2016 and February 2017 on 257 crop fields to identify land use in order to calibrate and validate the classification of remote sensing data (Figure 1). Samples were randomly distributed throughout the study site; two-thirds (171 fields) were used for training and the other one-third (86 fields) for validation. The size of the inventoried fields ranged from 0.1–65 ha. The objects taken into account in this study for the object-oriented approach correspond to the 257 fields identified in situ from November 2016 to May 2017. We had a vector layer of field boundaries on which we reported the ground surveys.

**2.3. Satellite Imagery**

A series of nine optical Sentinel-2 and nine SAR Sentinel-1 images were acquired from autumn to spring (August 2016 to May 2017) from the data hub of the European Space Agency and Centre National d’Études Spatiales [41]. Sentinel-1 images were acquired in Single Look Complex (SLC) mode (delivered with VH and VV polarization states) with an incidence angle of 45–47°. The associated range and azimuth spatial resolutions were respectively 2.3 and 13.9 m. Sentinel-2 level 2A images were acquired (i.e., corrected for geometric and atmospheric effects) with a spatial resolution of 10–60 m.
and a spectral resolution of 13 bands. Characteristics of the optical and SAR images are summarized in Table 2.

**Table 2.** Characteristics of the Sentinel-1 Synthetic-Aperture Radar (SAR) and Sentinel-2 optical images. NIR: near infrared, SWIR: shortwave infrared.

<table>
<thead>
<tr>
<th>Sentinel-2 (Optical)</th>
<th>Sentinel-1 (SAR)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spatial resolution</strong></td>
<td>10–60 m</td>
</tr>
<tr>
<td><strong>Azimuth resolution</strong></td>
<td>2.3 m</td>
</tr>
<tr>
<td><strong>Band 1 (Coastal)</strong></td>
<td>-0.443 µm</td>
</tr>
<tr>
<td><strong>Band 2 (Blue)</strong></td>
<td>-0.490 µm</td>
</tr>
<tr>
<td><strong>Band 3 (Green)</strong></td>
<td>-0.560 µm</td>
</tr>
<tr>
<td><strong>Band 4 (Red)</strong></td>
<td>-0.665 µm</td>
</tr>
<tr>
<td><strong>Band 5 (Red Edge)</strong></td>
<td>-0.705 µm</td>
</tr>
<tr>
<td><strong>Band 6 (Red Edge)</strong></td>
<td>-0.740 µm</td>
</tr>
<tr>
<td><strong>Band 7 (Red Edge)</strong></td>
<td>-0.783 µm</td>
</tr>
<tr>
<td><strong>Band 8 (NIR)</strong></td>
<td>-0.842 µm</td>
</tr>
<tr>
<td><strong>Band 8A (NIR)</strong></td>
<td>-0.865 µm</td>
</tr>
<tr>
<td><strong>Band 9 (Water)</strong></td>
<td>-0.940 µm</td>
</tr>
<tr>
<td><strong>Band 10 (SWIR)</strong></td>
<td>-1.375 µm</td>
</tr>
<tr>
<td><strong>Band 11 (SWIR)</strong></td>
<td>-1.610 µm</td>
</tr>
<tr>
<td><strong>Band 12 (SWIR)</strong></td>
<td>-2.190 µm</td>
</tr>
<tr>
<td>Coverage</td>
<td>290 km</td>
</tr>
<tr>
<td></td>
<td>31-Oct-2016</td>
</tr>
<tr>
<td></td>
<td>30-Nov-2016</td>
</tr>
<tr>
<td></td>
<td>30-Dec-2016</td>
</tr>
<tr>
<td></td>
<td>19-Jan-2017</td>
</tr>
<tr>
<td></td>
<td>18-Feb-2017</td>
</tr>
<tr>
<td></td>
<td>9-Apr-2017</td>
</tr>
<tr>
<td></td>
<td>9-May-2017</td>
</tr>
</tbody>
</table>

### 3. Materials and Methods

The full dataset consisted of parameters derived from the optical and SAR image time-series. The method developed to identify winter vegetation cover from this dataset had three steps: (i) pre-processing the SAR and optical image time-series, (ii) selecting the most consistent parameters and (iii) classifying the optical and SAR time-series.

#### 3.1. Pre-Processing of Time-Series

**3.1.1. Pre-processing of Sentinel-1 Images**

Backscattering Coefficients

Sentinel-1 images were first radiometrically calibrated (Figure 3) using SNAP (Sentinel Application Platform) v5.0 software with the following equation [42]:

\[
value(i) = \frac{|DN_i|^2}{A_i^2},
\]

where DN is the digital number of each pixel (amplitude of backscattering signal) and A is the information necessary to convert SAR reflectivity into physical units provided in the Calibration Annotation Data Set in the image metadata. This equation transforms the DN of each pixel into a backscattering coefficient on a linear scale.
Figure 3. Pre-processing workflow for Sentinel-1 Synthetic-Aperture Radar (SAR) and Sentinel-2 optical images. SLC: Single Look Complex.

A refined Lee filter [43,44] was then applied with a window of $7 \times 7$ pixels to reduce speckle noise using SNAP v5.0 software. This window size was selected to decrease speckle noise while preserving a suitable spatial scale, which was necessary to ensure identification of winter land use. Sentinel-1 images were then geocoded using Shuttle Radar Topography Mission 3s data to correct topographic deformations. The accuracy of geometric correction was less than 10 m per pixel. A backscattering ratio was calculated by dividing $\sigma^0_{VH}$ by $\sigma^0_{VV}$. This ratio highlights differing scattering mechanisms of each target. The backscattering coefficients $\sigma^0_{VH}$ and $\sigma^0_{VV}$ and the backscattering ratio $\sigma^0_{VH} : \sigma^0_{VV}$ were then converted into decibels (db) using the following equation:

$$\sigma^0\,(\text{db}) = 10 \times \log_{10}(\sigma^0),$$

(2)

Polarimetric Parameters

A $2 \times 2$ covariance matrix ($C_2$) was first extracted from the scattering matrix $S$ of each SLC image using PolSARpro v5.1.1 software [45]. The elements of the matrix, which are independent of the polarimetric absolute phase [46], were directly geocoded using SNAP v5.0 software. A refined Lee filter was then applied using a window of $7 \times 7$ pixels to reduce speckle noise.

The second (C12) and fourth (C22) elements of the diagonal were compared to the backscattering coefficients. The SPAN, which corresponds to the total scattered power and the Shannon Entropy (SE), which corresponds to the sum of two contributions related to the intensity ($SE_i$) and the degree of polarization ($SE_p$) [46], were then calculated from the matrix. SE measures the disorder encountered in polarimetric SAR images using the following equation:

$$SE = \log(\pi^2 \sigma^2 |C_2|) = SE_i + SE_p,$$

(3)

A total of 13 polarimetric parameters were derived: the elements of the matrix, the SPAN, SE, SEi and SEp (Table 3).
Table 3. Parameters derived from Sentinel-1 Synthetic-Aperture Radar (SAR) and Sentinel-2 optical image time-series.

<table>
<thead>
<tr>
<th>Sentinel-2 Optical Parameters</th>
<th>Sentinel-1 SAR Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band 2-blue</td>
<td>Matrix element C11 decibels (C11 db)</td>
</tr>
<tr>
<td>Band 3-green</td>
<td>Matrix element C12 imaginary part (C12 img)</td>
</tr>
<tr>
<td>Band 4-red</td>
<td>Matrix element C12 real part (C12 rel)</td>
</tr>
<tr>
<td>Band 5-vegetation red edge</td>
<td>Matrix element C22 (C22)</td>
</tr>
<tr>
<td>Band 6-vegetation red edge</td>
<td>Matrix element C22 decibels (C22 db)</td>
</tr>
<tr>
<td>Band 7-vegetation red edge</td>
<td>Shannon entropy (SE)</td>
</tr>
<tr>
<td>Band 8-near infrared (NIR)</td>
<td>Shannon entropy Intensity ($SE_{int}$)</td>
</tr>
<tr>
<td>Band 8a-narrow NIR</td>
<td>Shannon entropy normalized ($SE_{norm}$)</td>
</tr>
<tr>
<td>Band 9-water vapor</td>
<td>Shannon entropy polarization normalized ($SE_{polar}$)</td>
</tr>
<tr>
<td>Band 10-shortwave infrared (SWIR) (cirrus)</td>
<td>Total scattered power (SPAN)</td>
</tr>
<tr>
<td>Band 11-SWIR</td>
<td>Backscattering coefficient VH ($\sigma_{0VH}$)</td>
</tr>
<tr>
<td>Band 12-SWIR</td>
<td>Backscattering coefficient VV ($\sigma_{0VV}$)</td>
</tr>
<tr>
<td>Normalized Difference Vegetation Index</td>
<td>VH/VV</td>
</tr>
<tr>
<td>(NDVI = $\frac{(\text{Band 8} - \text{Band 4})}{\text{Band 8} + \text{Band 4}}$)</td>
<td></td>
</tr>
<tr>
<td>Normalized Difference Water Index</td>
<td>$\text{NDWI} = \frac{\text{Band 8} - \text{Band 12}}{\text{Band 8} + \text{Band 12}}$</td>
</tr>
<tr>
<td>Soil Adjusted Vegetation Index</td>
<td>(SAVI = $\frac{\text{Band 8} - \text{Band 4}}{\text{Band 8} + \text{Band 4} + L} + (1 + L)$)</td>
</tr>
<tr>
<td>Leaf Area Index (LAI)</td>
<td>Fraction of photosynthetically active radiation (FAPAR)</td>
</tr>
</tbody>
</table>
| Fraction of vegetation cover (FCOVER) | }

3.1.2. Pre-processing of Sentinel-2 Images

Sentinel-2 images were pre-processed using the CNES Kalideos processing chain [47,48]. They were corrected for atmospheric disturbances, orthorectified and georeferenced based on the Universal Transverse Mercator (UTM) reference system (zone 30N). Assessment of Level-2A Sentinel-2 time-series, which were corrected from atmospheric and geometric effects from Sentinel-2 Level-1 time-series, was performed based on a visual interpretation of reference target reflectance (water, buildings, etc.).

Calculation of Vegetation Indices and Biophysical Parameters

Based on the literature, 2 vegetation indices, 1 water index and 3 biophysical parameters were computed using SNAP v5.0 software. The two vegetation indices were derived from the near infrared and red bands: The Normalized Difference Vegetation Index (NDVI; [49]) and the Soil Adjusted Vegetation Index (SAVI; [50]) that have demonstrated their relevance to study land use [51,52]. The Normalized Difference Water Index (NDWI; [53]) was calculated from the near and middle infrared bands often used to estimate water in vegetation [54]. Biophysical parameters (Leaf Area Index, fraction of photosynthetically active radiation and fractional vegetation cover) were calculated using the PROSAIL radiative transfer model [55,56] implemented in SNAP v5.0 software. These parameters describe the state of the vegetation cover and provide information on the density of green vegetation [57].

In total, the number of parameters to be processed was 297: 162 for Sentinel-2 data (18 parameters × 9 dates) and 135 for Sentinel-1 (15 parameters × 9 dates).

3.2. Processing Sentinel-1 and -2 Time-Series

The method for processing Sentinel-1 and -2 time-series we developed had three steps (Figure 4): (i) extracting features to yield one parameter value per field, (ii) reducing the optical and SAR
parameter dataset using a data-mining algorithm and (iii) classifying vegetation cover using RF and SVM algorithms.

Calculation of Vegetation Indices and Biophysical Parameters

Based on the literature, 2 vegetation indices, 1 water index and 3 biophysical parameters were computed using SNAP v5.0 software. The two vegetation indices were derived from the near infrared and red bands: The Normalized Difference Vegetation Index (NDVI; [49]) and the Soil Adjusted Vegetation Index (SAVI; [50]) that have demonstrated their relevance to study land use [51,52]. The Normalized Difference Water Index (NDWI; [53]) was calculated from the near and middle infrared bands often used to estimate water in vegetation [54].

Biophysical parameters (Leaf Area Index, fraction of photosynthetically active radiation and fractional vegetation cover) were calculated using the PROSAIL radiative transfer model [55,56] implemented in SNAP v5.0 software. These parameters describe the state of the vegetation cover and provide information on the density of green vegetation [57].

In total, the number of parameters to be processed was 297: 162 for Sentinel-2 data (18 parameters X 9 dates) and 135 for Sentinel-1 (15 parameters X 9 dates).

3.2. Processing Sentinel-1 and Sentinel-2 Time-Series

The method for processing Sentinel-1 and Sentinel-2 time-series we developed had three steps (Figure 4): (i) extracting features to yield one parameter value per field, (ii) reducing the optical and SAR parameter dataset using a data-mining algorithm and (iii) classifying vegetation cover using RF and SVM algorithms.

Figure 4. Image processing workflow applied to Sentinel-1 parameters, Sentinel-2 parameters and the combination of Sentinel-1 and Sentinel-2 parameters. SVM: Support Vector Machine, RF: Random Forest, OA: overall accuracy.

3.2.1. Feature Extraction

The parameters derived from Sentinel time-series (Table 3) were used to process two classification approaches: an object-based and a pixel-based approach. Hedgerows, which are considered noisy features in land use mapping, were removed from the images by applying a 5 m negative buffer around the 257 field boundaries observed on the study site. Then, two feature extraction sequences were performed: (i) for the object-based approach, the mean and median values of optical and SAR parameters were calculated at the field scale; (ii) for the pixel-based approach, 100 pixels were randomly selected from the 257 ground surveys for each winter land use class (5 classes × 100 pixels), which represents a total of 500 pixels.

3.2.2. Reduction of the Parameter Dataset

A two-step approach was used to remove noise and inconsistent data from the parameter datasets used for the classification. First, a correlation matrix was computed and a threshold of 0.95/−0.95 was applied to find the most correlated parameters (268 out of 297 parameters). Then, the
choice of parameters to be removed was made based on a current state-of-the-art to select the most relevant parameters for the study of inter land use [27,51,52]. At this stage, the dataset consisted of 144 parameters (90 optical and 54 radar parameters). In a second step, an analysis of the parameter relevance was performed using the RF significance function to further reduce the number of parameters to be classified. A measure of variable importance was provided for each candidate predictor using this heuristic method based on the Gini Index [36,58,59]. The break in the histogram was used as the threshold for selecting the more relevant parameters, that is, 15 optical and 15 SAR parameters (Figure 5). These 30 parameters were used for winter land use classifications.

The RF and SVM algorithms, which are supervised classification methods, were used to classify land use during winter 2016–2017. These algorithms were chosen for their consistently strong performance and the accuracy with which they classify LULC [60,61]. The RF algorithm is an ensemble algorithm that uses a set of classification and regression trees to make a prediction [36]. The package randomForest developed by [62] and implemented in R (v.3.3.2) was used to perform winter land use classifications. Two Random Forest parameters, namely the number of trees (ntree), which was created by randomly selecting samples from the training dataset [28] and the number of variables used for tree nodes splitting (mtry) were tuned and randomly determined using the tune function implemented in the randomForest package. For this study, the ntree parameter was set at 1,000, few articles having demonstrated that beyond the creation of 1,000 classification trees, the number of errors produced remains stable [63]. The SVM algorithm is based on statistical learning theory that aims to determine the location of decision boundaries that produce an optimal separation of classes [37]. In a two-class pattern-recognition problem in which classes can be separated linearly, the SVM selects the linear decision boundary that creates the greatest margin between the two classes. The margin is the sum of distances to the hyperplane from the closest points of the two classes [37]. Thus, it initially extracts the best linear boundary between two classes of the training set; however, it is not restricted to linear discrimination, since one of its main advantages is its extension to nonlinear discrimination via the kernel trick [19]. The package e1071 (v 1.7-0) developed by [64] and implemented in R (v.3.3.2)

**Figure 5.** Parameters that contributed the most to land use classifications, ranked by importance for (a) optical and (b) Synthetic-Aperture Radar image time-series from August 2016 to May 2017. See Table 3 for the definition of abbreviations.

3.2.3. Land Use Classification

The RF and SVM algorithms, which are supervised classification methods, were used to classify land use during winter 2016–2017. These algorithms were chosen for their consistently strong performance and the accuracy with which they classify LULC [60,61]. The RF algorithm is an ensemble algorithm that uses a set of classification and regression trees to make a prediction [36]. The package randomForest developed by [62] and implemented in R (v.3.3.2) was used to perform winter land use classifications. Two Random Forest parameters, namely the number of trees (ntree), which was created by randomly selecting samples from the training dataset [28] and the number of variables used for tree nodes splitting (mtry) were tuned and randomly determined using the tune function implemented in the randomForest package. For this study, the ntree parameter was set at 1,000, few articles having demonstrated that beyond the creation of 1,000 classification trees, the number of errors produced remains stable [63]. The SVM algorithm is based on statistical learning theory that aims to determine the location of decision boundaries that produce an optimal separation of classes [37]. In a two-class pattern-recognition problem in which classes can be separated linearly, the SVM selects the linear decision boundary that creates the greatest margin between the two classes. The margin is the sum of distances to the hyperplane from the closest points of the two classes [37]. Thus, it initially extracts the best linear boundary between two classes of the training set; however, it is not restricted to linear discrimination, since one of its main advantages is its extension to nonlinear discrimination via the kernel trick [19]. The package e1071 (v 1.7-0) developed by [64] and implemented in R (v.3.3.2)
was used to perform SVM classifications. A set of four SVM parameters (gamma, cost, degree, nu) was randomly tuned using the tune function integrated in the e1071 package. Then, several tests were carried out to determine the optimal kernel for winter use classification. At the end of the selection process the polynomial kernel was selected. Results obtained with RF and SVM algorithm were compared to evaluate their suitability for classifying vegetation cover into land use classes. Classification performance was estimated using a cross-validation test. The classification was applied to a varying subset of 257 fields (500 pixels): two-thirds (171 fields or 332 pixels) were used for training, one-third (86 fields or 168 pixels) for validation [19]. This process was repeated by changing the training/validation subsamples. Classification accuracy was assessed using OA and the Kappa index, which expresses the proportional decrease in error generated by the classification compared to the error of a completely random classification [65]. Finally, the vegetation cover used to map land use was classified using the algorithm with the highest OA. The classification process was tested for the two approaches: object-based and pixel-based. Sentinel-1 and Sentinel-2 parameters were classified separately and by combining optical and SAR data in the same dataset, that is, we realized a fusion at the lowest processing level (pixels level) referring to the merging of measured physical parameters [66]. The combination of Sentinel-1 and -2 parameters were then classified.

4. Results and Discussion

4.1. Importance of Optical and SAR Parameters for Identifying Winter Land Use

The correlation analysis computed with a threshold of 0.95/−0.95 was applied to find the most correlated parameters (268 out of 297 parameters), which resulted in the suppression of 153 parameters. However, correlation matrix was not displayed because the number of parameters (297) is too large and it would not be readable.

Figure 5 shows the parameter contribution to land use classifications, ranked by importance for optical and Synthetic-Aperture Radar image time-series. The break in the histogram (after band 5 in May for Sentinel-2; after polarized Shannon Entropy in April for Sentinel-1) was used as the threshold for selecting the most relevant parameters, that is, 15 optical and 15 SAR parameters.

The most important SAR parameter was the ratio VH/VV in May, which highlights crop growth, especially of winter crops [27]. Similarly, the backscattering coefficients were also important. Results show the importance of parameters derived from the May and April images due to their sensitivity to variations in double-bounce and volume-scattering mechanisms [67]. The backscattering coefficient VH calculated from the November image was also important due to its sensitivity to direct contributions from the ground and the canopy [68]. These parameters highlight the difference between bare soils and crops.

Concerning optical parameters, the two most important parameters were band 2 in December 2016 and SAVI in May 2017 with a Gini index above 70. One can notice that the best NDVI parameter came in 14th position, which is not in accordance with the current literature [52]. Results also show the importance of parameters derived from the May image, 11 parameters out of the 15 most important parameters being derived from this image. This is due to the fact that this period is the main phase of highly dynamic plant growth: vegetation peaks in May (the end of spring), when grasslands are easily detected (Figure 5).

The spectral distributions of two of the most pertinent parameters, that is, SAVI in May 2017 and VH/VV ratio in May 2017, were computed for the five classes of winter land use (Figure 6). SAVI was selected over Band 2 based on the current literature, which demonstrated the potential of this vegetation index to discriminate and characterize crop dynamics [28,54,69].

The Figure 6a shows the potential of SAVI to discriminate winter crops and grasslands from catch crops and crop residues in May -when the phenological stages and land use changes are most pronounced. Conversely, results highlight the difficulty to discriminate bare soils from the other classes, due to a high intra-class variance. To a lesser extent, Figure 6b shows the potential of the VH/VV
ratio to further separate winter crops and grasslands from catch crops and crop residues. However, like the SAVI parameter, the VH/VV ratio is not sufficient to separate bare soils from other classes due to a very high intra-class variance. These results are consistent with the existing literature, which demonstrated the importance of using backscattering coefficients alone ($\sigma_{VH}$ or $\sigma_{VV}$) or in combination (VH/VV) to identify land use [70].

![Graph](image)

**Figure 6.** Distribution of (a) the Soil Adjusted Vegetation Index (SAVI) and (b) the ratio VH/VV parameters for winter land use classes calculated from May 2017 images.

### 4.2. Winter Land Use Classification

Accuracy of the winter land use classification obtained from optical and SAR parameters varied significantly depending on (i) the classification approach (i.e., pixel-based or object-based), (ii) the classification algorithm (i.e., RF or SVM) and (iii) the time-series dataset (i.e., Sentinel-1, Sentinel-2 or a combination of both).

Results of the pixel-based and object-based approaches to identifying winter land use had similar accuracy for the Sentinel-2 dataset and the combined Sentinel-1 and -2 dataset. The object-based approach had a slightly higher OA (68–83%, Kappa = 0.64–0.77) than the pixel-based approach (55–81%, Kappa = 0.51–0.77) (Figure 7). Although this is consistent with other studies of LULC [71], few studies have presented advantages of using the object-based approach instead of the pixel-based approach to identify and characterize LULC [72,73]. In contrast, results obtained with the Sentinel-1 data showed the superiority of the object-based approach, due to the heterogeneity of SAR values within fields in winter.

![Graph](image)

**Figure 7.** Overall accuracy of (a) object-based and (b) pixel-based image classifications of winter land use, for the best Sentinel-1, Sentinel-2 and combined Sentinel-1 and -2 parameters using the Random Forest (RF) and Support Vector Machine (SVM) algorithms.
Concerning the classification techniques, the RF algorithm had higher OA than the SVM algorithm (median OA = 81% and 79%, respectively). The RF algorithm also had less variation in OA than SVM (72–83% (Kappa = 0.67–0.77) and 68–80% (Kappa = 0.64–0.76), respectively). The potential of SVM and RF algorithms for remote sensing studies has been widely demonstrated [59–61]. Our results show that the RF algorithm is slightly more accurate effective than the SVM algorithm, which is consistent with results of other studies [59,61,74].

This study evaluated the respective advantages of Sentinel optical and SAR time-series to identify winter land use. Classification was better using a combination of Sentinel-1 and -2 parameters (median OA = 81%, Kappa = 0.77) (Table 4), with OA ranging from 75–82% (Kappa = 0.68–0.77). Conversely, classifications based on either Sentinel-1 or Sentinel-2 parameters alone had OAs of 68–78% and 74–80%, respectively (Figure 7). While the results highlight the utility of Sentinel-1 and Sentinel-2 individually, they also emphasize that classification using the Sentinel-2 dataset always outperformed that using the Sentinel-1 dataset. The classification results highlight the advantages of using the combined Sentinel-1 and -2 datasets, with OA ranging from 68–83% (Kappa = 0.64–0.77). Therefore, our study confirms the effectiveness of Sentinel-1 and -2 time-series for identifying land use, as previous studies have demonstrated [26,27] and also shows the potential of using the combined Sentinel-1 and -2 datasets for this purpose. Additionally, the originality of this study is the identification of land use in winter.

Table 4. Median accuracy of winter land use classifications obtained for the best Sentinel-1, Sentinel-2 and combined Sentinel-1 and -2 parameters using the Random Forest (RF) and Support Vector Machine (SVM) algorithms. OA: overall accuracy, Kappa: Kappa index.

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Datasets</th>
<th>Object-Based Approach</th>
<th>Pixel-Based Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>OA</td>
<td>Kappa</td>
</tr>
<tr>
<td>RF</td>
<td>Sentinel-1</td>
<td>72%</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Sentinel-2</td>
<td>78%</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Sentinel-1 &amp; -2</td>
<td>81%</td>
<td>0.77</td>
</tr>
<tr>
<td>SVM</td>
<td>Sentinel-1</td>
<td>73%</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Sentinel-2</td>
<td>79%</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>Sentinel-1 &amp; -2</td>
<td>78%</td>
<td>0.75</td>
</tr>
</tbody>
</table>

The best classification, with OA of 81% and Kappa of 0.77, used an object-based approach and 30 parameters derived from a combination of Sentinel-1 and -2 parameters. Misclassification errors were observed between bare soils (under- and over-estimation rates of 64% and 94%, respectively) and the other classes (Table 5). This agrees with the study’s difficulty in discriminating bare soils from the other classes (Figure 6).

Table 5. Confusion matrix of the best winter land use classification obtained using a parameter dataset derived from a combination of Sentinel-1 and -2 time-series. Overall accuracy = 81%, Kappa index = 0.77.

<table>
<thead>
<tr>
<th>Catch crops</th>
<th>Winter crops</th>
<th>Grasslands</th>
<th>Crop residues</th>
<th>Bare soils</th>
<th>Commission Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catch crops</td>
<td>310</td>
<td>42</td>
<td>19</td>
<td>0</td>
<td>83</td>
</tr>
<tr>
<td>Winter crops</td>
<td>1</td>
<td>10</td>
<td>56</td>
<td>310</td>
<td>0</td>
</tr>
<tr>
<td>Grasslands</td>
<td>20</td>
<td>310</td>
<td>0</td>
<td>68</td>
<td>91.7 %</td>
</tr>
<tr>
<td>Crop residues</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>389</td>
<td>62</td>
</tr>
<tr>
<td>Bare soils</td>
<td>16</td>
<td>5</td>
<td>0</td>
<td>426</td>
<td>93.8 %</td>
</tr>
<tr>
<td>Omission errors</td>
<td>89.3 %</td>
<td>79.2 %</td>
<td>87.6 %</td>
<td>100 %</td>
<td>64.4 %</td>
</tr>
</tbody>
</table>

The spatial distribution of winter land uses mapped at the 1:100,000 scale from the best classification (Figure 8) shows that bare soils and crop residues covered less than 5% of the UAA, while a high percentage was covered with grasslands (30%) or winter crops (35%). In general, catch crops and winter crops were located on the largest fields, while bare soils and grasslands were located on the smallest fields.
Table 5. Confusion matrix of the best winter land use classification obtained using a parameter dataset derived from a combination of Sentinel-1 and -2 time-series. Overall accuracy = 81%, Kappa index = 0.77.

<table>
<thead>
<tr>
<th>Land Use Cover</th>
<th>Commission Errors</th>
<th>Omission Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catch crops</td>
<td>310</td>
<td>42</td>
</tr>
<tr>
<td>Winter crops</td>
<td>1</td>
<td>410</td>
</tr>
<tr>
<td>Grasslands</td>
<td>20</td>
<td>56</td>
</tr>
<tr>
<td>Crop residues</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Bare soils</td>
<td>16</td>
<td>7</td>
</tr>
</tbody>
</table>

The spatial distribution of winter land uses mapped at the 1:100,000 scale from the best classification (Figure 8) shows that bare soils and crop residues covered less than 5% of the UAA, while a high percentage was covered with grasslands (30%) or winter crops (35%). In general, catch crops and winter crops were located on the largest fields, while bare soils and grasslands were located on the smallest fields.

The distribution of membership probabilities associated with this classification indicates that accuracy decreased at the edges of the study site (Figure 9). Fields smaller than 1 ha had the lowest membership probability (0.47), while those larger than 10 ha had the highest (0.84), indicating that classification accuracy increased with field size. These results agree with the confusion matrix (Table 5), in which misclassification was greatest for bare soils.

![Land Use Cover Map](image_url)

**Figure 8.** Distribution of winter land use obtained using a parameter dataset derived from a combination of Sentinel-1 and -2 time-series. Classification was performed using a Random Forest algorithm.

![Membership Probabilities Map](image_url)

**Figure 9.** Map of the distribution of winter land use membership probabilities obtained using the Random Forest algorithm.
5. Conclusions

This study evaluated the respective advantages of using Sentinel optical and SAR time-series to identify winter land use, using SVM and RF algorithms with pixel-based and object-based approaches. Our study used high spatial and temporal resolutions of Sentinel data to identify land use types in winter and to our knowledge, this is the first time such a study has been undertaken.

Results show that winter land use can be identified accurately using combined Sentinel-1 and -2 time-series with a pixel-based approach using an RF algorithm. Analysis of using Sentinel-1 and -2 parameters to identify winter land use led to recommendations for extracting features when mapping winter land use. Results reveal the advantage of using backscattering coefficients alone or combined with the NDVI.

Our results also demonstrated limits of this approach to identifying winter land use in small fields, due to the spatial resolution of Sentinel sensors. Thus, future research could evaluate the use of very high spatial resolution optical and SAR images, such as ALOS-2 or TerraSAR-X data, to improve the accuracy of classifying land use types during winter. Better understanding of optical and SAR signal behaviours of different agricultural practices and environmental conditions would help to identify and monitor winter land use. This has important implications for developing sustainable agriculture that decreases the risk of transferring pollutants to the environment.


Funding: This research was funded by MESR (Ministry of Higher Education and Research of France), PHD grant 2016

Acknowledgments: This study was supported by the Kalideos project, funded by the CNES [75] and the Zone Atelier Armorique project. We would like to thank Napo N’Bohn and Marianne Balaresque for their help in the field.

References


11. Xu, D. Compare NDVI Extracted from Landsat 8 Imagery with That from Landsat 7 Imagery. Am. J. Remote Sens. 2014, 2, 10. [CrossRef]


73. Yan, G.; Mas, J.-F.; Maathuis, B.H.P.; Xiangmin, Z.; Van Dijk, P.M. Comparison of Pixel-Based and Object-Oriented Image Classification Approaches—A Case Study in a Coal Fire Area, Wuda, Inner Mongolia, China. Int. J. Remote Sens. 2006, 27, 4039–4055. [CrossRef]