Unsupervised crop anomaly detection at the parcel-level using optical and SAR images: application to wheat and rapeseed crops

Florian Mouret, Mohanad Albughdadi, Sylvie Duthoit, Denis Kouamé, Hervé Poilvé, Guillaume Rieu, Jean-Yves Tourneret

Unsupervised crop anomaly detection at the parcel-level using optical and SAR images: application to wheat and rapeseed crops⋆

Florian Mouret[a,b,*], Mohanad Albughdadi[a], Sylvie Duthoit[a], Denis Kouamé[c], Hervé Poilvé[d], Guillaume Rieu[a] and Jean-Yves Tourneret[b]

[a] *TerraNIS, 12 Avenue de l'Europe, 31520 Ramonville-Saint-Agne, France*

[b] *University of Toulouse / IRIT-INP-ENSEEIHT / TéSA, 2 Rue Charles Camichel, 31000 Toulouse, France*

[c] *University of Toulouse / IRIT-UPS, 118 Route de Narbonne, 31062 Toulouse Cedex 9, France*

[d] *AIRBUS Defence & Space, 5 rue des satellites, 31400 Toulouse, France*

ARTICLE INFO

Keywords:
Crop anomaly detection
Parcel-level analysis
Sentinel 1 & 2 images
Wheat crops
Rapeseed crops
Unsupervised crop monitoring
Isolation Forest
One-Class SVM
Local Outlier Factor

ABSTRACT

This paper proposes a generic approach for crop anomaly detection at the parcel-level based on unsupervised point anomaly detection techniques. The input data is derived from synthetic aperture radar (SAR) and optical images acquired using Sentinel-1 and Sentinel-2 satellites. The proposed strategy consists of four sequential steps: acquisition and preprocessing of optical and SAR images, extraction of optical and SAR indicators, computation of zonal statistics at the parcel-level and point anomaly detection. This paper analyzes different factors that can affect the results of anomaly detection such as the considered features and the anomaly detection algorithm used. The proposed procedure is validated on two crop types in Beauce (France), namely, rapeseed and wheat crops. Two different parcel delineation databases are considered to validate the robustness of the strategy to changes in parcel boundaries.

⋆This document is the results of the research project funded by TerraNIS SAS.
*Corresponding author
florian.mouret@terranis.fr / florian.mouret@irit.fr (F. Mouret)

1. Introduction

Remote sensing images are widely used to study vegetation status and more specifically to monitor crop development. The link between these images, in particular optical and synthetic aperture radar (SAR), and crop vegetation has been investigated in many studies (Bannari et al., 1995; Wegmuller and Werner, 1997; Daughtry et al., 2000; Moran et al., 2000; Weiss et al., 2020). Optical images have been used in various applications requiring for instance land-cover classification and biophysical parameter estimation (Verrelst et al., 2015; Gómez et al., 2016). Optical images are convenient to use and interpret. However, they are sensitive to cloud coverage and atmospheric conditions, which is a known issue in remote sensing. SAR images provide more information on the vegetation structure (Kumar et al., 2013; Betbeder et al., 2016; Abdikan et al., 2016; Khabbazan et al., 2019) and are available at any day time and for any cloud coverage. However, SAR images are challenging to interpret, *e.g.*, due to the presence of speckle noise, which may limit their readability. Recent reviews have mentioned the potential interest of using SAR images for vegetation monitoring (McNairn and Shang, 2016; Liu et al., 2019). Combining SAR and optical images is interesting because both types of images are complementary. Their joint use has also been encouraged by the large amount of free data provided by Sentinel-1 (S1) and Sentinel-2 (S2) satellites from the European Space Agency (ESA). Interesting remote sensing problems that have been addressed using optical and SAR images include crop classification (Hedayati and Bargiel, 2018; Denize et al., 2018) and change detection (Prendes et al., 2015a,b,c).

This paper addresses the problem of anomaly detection for crop monitoring at a parcel-level using optical and SAR images. It aims at meeting the farmers' need for automated crop monitoring, *e.g.*, for optimization of agricultural practices (disease or fertilization management for instance). The proposed anomaly detection strategy makes use of recent improvements in terms of computational power and access to near-real time (NRT) satellite images. To our knowledge, it is the first time that crop monitoring is formulated as an anomaly detection problem. At this point, it is worth mentioning the work conducted in (Albughdadi et al., 2017), which was devoted to detecting crop anomalies using agronomic indicators extracted from multispectral images. However, the algorithm studied in (Albughdadi et al., 2017) was based on the combination of missing data reconstruction and classification and not on anomaly detection methods.

In the literature, a similar problem solved using anomaly detection techniques concerns the mon-

itoring of wide agricultural areas (*e.g.*, at the country-level). A motivation for this problem is the automated detection of vegetation areas that are potentially in critical status. It is particularly suited for international or national agencies interested in monitoring food security that have a need for early warning of food crises, drought monitoring, etc. Many investigations have been made using MODerated Resolution Imaging Spectroradiometer (MODIS) images. MODIS images have a low spatial resolution and a high temporal resolution, which make them adapted to global vegetation monitoring. Another interest with MODIS data is that they have been made available since 2000. As a consequence, very long time series are available for the analysis of vegetation. Various methods have been introduced in the literature to address the problem of detecting abnormal vegetation areas using time series constructed from the MODIS Normalized Difference Vegetation Index (NDVI). Most approaches aim at modeling NDVI time series using historical data and detecting potential anomalies affecting these time series by comparing new observations with their corresponding predicted values. Common approaches are based on a parametric model for the time evolution of NDVI such as the double logistic or symmetric Gaussian models (Atzberger and Eilers, 2011a; Beck et al., 2006). Smoothing techniques can help to have more reliable time series and to work in NRT (Atzberger and Eilers, 2011b; Hird and McDermid, 2009; Klisch and Atzberger, 2016; Meroni et al., 2019). Various other approaches such as a decomposition of NDVI time series using season trend models (J. Verbesselt, 2012; Zhou and Tang, 2016), Seasonal Autoregressive Integrated Moving Average (SARIMA) models (Zhou et al., 2016) and prediction models such as extended Kalman filters (Sedano et al., 2015) have also been considered for analyzing the time evolution of NDVI. When looking at the literature on anomaly detection (Chandola et al., 2009; Pimentel et al., 2014), it appears that all the aforementioned methods aim at detecting contextual anomalies. In the case of agricultural monitoring, time is a contextual attribute and looking for contextual anomalies means finding an abnormal time period within the time series associated with the analyzed area. As defined in (Chandola et al., 2009), one way to deal with contextual anomalies is to model a normal temporal behavior and use this dynamic model to detect anomalies. In that case, it is worth noting that having long time series is mandatory to model normal behavior.

Two challenges appear when trying to adapt contextual anomaly detection methods (adapted to agricultural monitoring at the country-level) to the specific case of crop monitoring (and precision agriculture in general). The first challenge is the coarse spatial resolution of MODIS images, which is not adapted to work at the parcel-level. This problem is easily addressed with the new generation

of satellites such as S1 and S2 from the Copernicus program of the European Space Agency (ESA). The fine spatial and temporal resolution of S1 and S2 images allows working at the parcel-level with a high revisit frequency. The second challenge concerns the detection strategy to be used for crop anomaly detection. Contextual anomaly detection seems not particularly adapted to this problem for multiple reasons. First of all, historical modeling can be challenging because of crop rotation. There is also a lack of historical data for Sentinel satellites since the first S2 satellite was launched in 2015. Finally, it is important to note that crop anomaly detection requires to detect the most abnormal parcels within a season rather than detecting inter-annual abnormalities. All these reasons motivate the need to investigate new approaches for anomaly detection dedicated to crop monitoring, which is the main objective of this work.

Point anomaly detection has been receiving an increasing interest in the literature (as outlined in the review paper (Chandola et al., 2009) and (Pimentel et al., 2014)). Unlike contextual anomaly detection, point anomaly detection aims at comparing each instance of the dataset (here, each parcel) to the rest of the data to find the most different instances, that are isolated from the majority of the observed data. Point anomaly detection can be considered without a large amount of historical data. Moreover, multiple indicators providing information about the crop development can be used to detect point anomalies (not only NDVI), since no model of the normal behavior is required for the detection. Finally, note that point anomaly detection techniques are not able to detect "generalized" anomalies that could be defined as anomalies affecting the majority of parcels, such as drought. However, such techniques are able to detect the parcels that are the most (or less) affected by the so called "generalized" anomalies, which is more interesting for crop monitoring. All these reasons motivated our choice of using point anomaly detection techniques for crop monitoring.

This paper investigates the performance of point anomaly detection techniques for crop monitoring using SAR and optical images acquired using S1 and S2 sensors. It is organized as follows. Section 2 presents briefly the study area and the data used in this work. Section 3 introduces a parcel-based anomaly detection procedure for crop monitoring. Each step of this procedure is provided with practical implementation details. Section 4 validates the proposed approach on two different crop types (wheat and rapeseed crops) with a detailed analysis of the detection results. Moreover, two different parcel delineations are considered to confirm the robustness of the approach method to changes in the parcel boundaries. Finally, some conclusions and future work are reported in Section 5.

2. Study area and data

2.1. Study area

The analyzed area is located in the Beauce region in France covered by the T31UCP S2 tile (whose center is located approximately at 48°24'N latitude and 1°00'E longitude). Fig. 1 shows the tile location and the wide studied area. This area was chosen due to its richness of large cereal crop fields such as wheat and rapeseed. One season of wheat crops (2016/2017) and one season of rapeseed crops (2017/2018) are considered to show that the proposed method is robust to changes in crop types.

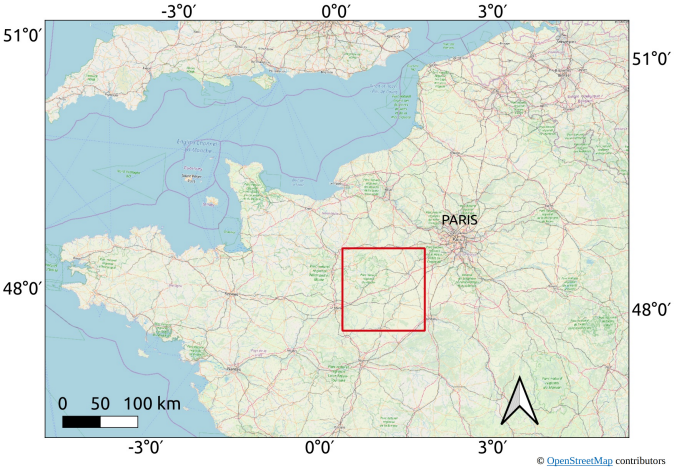


Fig. 1: Tile T31UCP S2 considered in this work located in the Beauce area (near Paris) and delimited by the red box.

2.2. Parcel data

Two different parcel data are used in the study. The first part of the analysis (Sections 4.3 and 4.4) is conducted on a total of 2218 rapeseed parcels (2017/2018 growing season) and 3361 wheat parcels (2016/2017 growing season). The parcels discarded during the preprocessing steps detailed in what follows are not counted here. In this first part, parcel delineations are defined using a customer database coming from a partner of the TerraNIS company, located in Toulouse, France. Using this customer database allows us focusing on crop anomalies rather than detecting anomalies related to information coming from the database (such as the reported crop type, reported field delineation, etc.) even if some errors in the parcel boundaries are still present. In Section 4.5, we will validate the robustness of the proposed method to changes in the parcel boundaries using another parcel delineation system for the rapeseed growing season. To that extent, 2118 parcel delineations resulting from the French Land

Parcel Identification System (LPIS) will be investigated. The French LPIS is also known as *Registre Parcellaire Graphique* (RPG). This database is available with an open license [1] and is updated yearly (in general with a delay of 2 years) on the basis of the farmer's Common Agricultural Policy (CAP) (Barbottin et al., 2018). For comparison purposes, and to avoid inherent errors in the LPIS reported crop type, each parcel of the customer database was matched with a corresponding LPIS parcel. Some parcels were not defined in the LPIS file, which explains why the number of parcels available for the LPIS analysis is slightly smaller than the number of parcels obtained when using the customer database.

Finally, to avoid problems in parcel frontiers, a buffer of 10 m was applied and too small parcels (area less than 0.5 ha) were discarded for both databases.

2.3. Remote sensing data

The multispectral S2 optical and SAR S1 images used in this study have been selected and downloaded from the PEPS platform of the French National Centre for Space Studies (Centre National d'Études Spatiales, CNES)[2]. Each S2 satellite has a theoretical revisit time of 10 days. Multispectral S2 optical images have 13 spectral bands. The blue, green, red, and near infrared channels (respectively Band 2 (490 nm), Band 3 (560 nm), Band 4 (665 nm) and Band 8 (842 nm)) have a spatial resolution of 10 m. The other channels have a spatial resolution of 20 m (Band 5 (705 nm), Band 6 (740 nm), Band 7 (783 nm), Band 8a (865 nm), Band 11 (1610 nm) and Band 12 (2190 nm)) or 60 m (Band 1 (443 nm), Band 9 (940 nm) and Band 10 (1375 nm)).

Radar S1 images were acquired by a C-band synthetic aperture radar (SAR) operating at a center frequency of 5.405 GHz. Each S1 satellite has a revisit time of 12 days. Ground Range Detected (GRD) products were used in the Interferometric Wide (IW) swath mode, allowing us to work with a 5 m by 20 m spatial resolution. GRD products provide focused SAR data that have been detected, multi-looked and projected to ground range using an Earth ellipsoid model. Phase information is lost but images are less sensitive to noise. Working with GRD images allowed the amount of images to be reduced for the analysis in order to keep a reasonable computational complexity of the learning step. Finally, the SAR images are available in dual polarization (VH+VV).

The acquisition dates of SAR (S1) and optical (S2) images are depicted in Fig. 2 for the 2016/2017

[1] `https://www.data.gouv.fr/fr/datasets/58d8d8a0c751df17537c66be/`, online accessed 10 March 2020
[2] `peps.cnes.fr/`, online accessed 10 March 2020

and 2017/2018 growing seasons. It was decided to select S2 images with a low cloud coverage (cloud coverage lower than 20%). The strategy considered to deal with remaining clouds will be detailed in Section 3.1. Finally, 41 S1 images and 10 S2 images were selected in 2016/2017 whereas 39 S1 images and 13 S2 images were selected in 2017/2018. Due to different cloud coverage for the two considered growing seasons, a lot of images were available in June 2018 whereas only one image was available in June 2017. Also, no image was selected between December 2017 and January 2018 whereas 3 S2 images were selected in 2016/2017 during the same period.

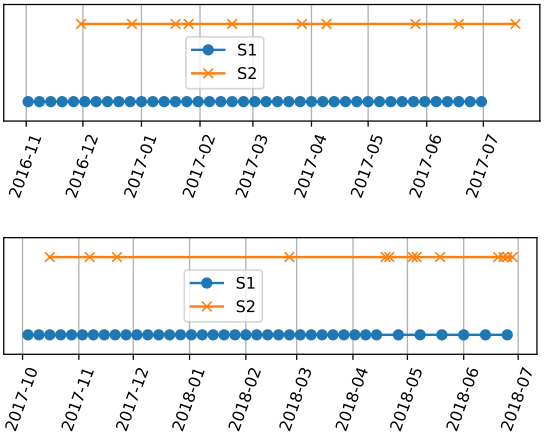


Fig. 2: Acquisition dates of the used images for 2016/2017 and 2017/2018 seasons. Each marker correspond to an available image. Too cloudy images were not used and are not represented.

3. Proposed strategy

The procedure proposed in this paper consists of four sequential steps depicted in Fig. 3 and discussed in detail in what follows.

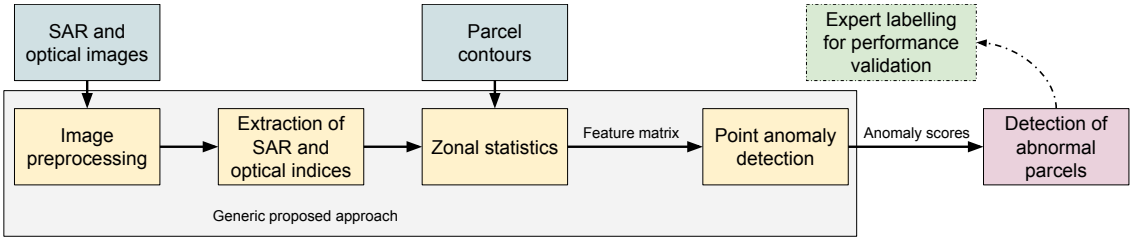


Fig. 3: Diagram summarizing the main steps of the proposed approach.

3.1. Image preprocessing

S2 images were preprocessed using the online MAJA processing chain (Hagolle et al., 2015) available on the PEPS platform of the Centre National d'Études Spatiales (CNES). This preprocessing step provides bottom-of-the atmosphere level-2A ortho-rectified products expressed in surface reflectance. In addition to atmospheric correction, Level-2A images are available with a cloud mask that was used to discard cloud pixels in the images. A resampling strategy was adopted to obtain a spatial resolution of 10 m for the channels with a lower spatial resolution (for instance each 20 m by 20 m pixel was divided into four 10 m by 10 m pixels with the same value). Parcels fully covered by clouds during at least one time instant were discarded from the database. Conversely, parcels partially covered by cloud were analyzed using pixels not covered by the cloud mask. As previous work tackled the missing data problem (Albughdadi et al., 2017), the work on this paper concentrates on parcels with no missing data.

The PEPS platform was also used to build the database of S1 images. Its online ortho-rectification of S1 images on the S2 grid was used to have a correspondence in the pixels analyzed in S1 and S2 data. Ortho-rectification leads to time series associated with calibrated to γ^0, ortho-rectified and filtered Sentinel-1 images. Calibration to γ^0 produces a backscattering coefficient normalized by the cosine of the incidence angle (Holecz et al., 1994). The ortho-rectification used in this work is based on the Orfeo Toolbox (Inglada and Christophe, 2009), with a detailed version of the code available online[3]. After these steps, SAR images were projected on the S2 grid yielding pixels having a 10 m by 10 m resolution. Finally, a multitemporal speckle filtering was applied to denoise SAR images (Quegan and Jiong Jiong Yu, 2001). As pointed out in (Quegan and Jiong Jiong Yu, 2001), multitemporal speckle filtering is much more efficient than spatial filtering for denoising SAR images. However, it requires a heavier computational power, which can be an issue for practical applications. As a consequence, both filtered and non-filtered S1 images were considered in this study.

3.2. Extraction of SAR and optical indicators

Using raw pixels from remote sensing images is rarely adapted to study the vegetation status. Generally, a spectral transformation of two or more bands is computed for each pixel in order to enhance the contribution of vegetation properties (Bannari et al., 1995). Since the proposed approach is unsupervised, feature selection is more complicated. In practice, an unsupervised algorithm is likely to

[3] `https://github.com/CNES/S1Tiling`, online accessed 10 March 2020

use all the features available for the analysis. Choosing irrelevant feature generally leads to irrelevant results, which explains why the indicators used in the study have to be selected carefully. The following section describes the indicators derived from optical and SAR images considered in this work and their importance for monitoring crop growth.

3.2.1. Optical vegetation indices

Many optical Vegetation Indices (VIs) have been proposed in the literature *e.g.*, (Bannari et al., 1995; Wu et al., 2008). A VI relates the acquired spectral information to the observed vegetation, and thus allows better quantitative and qualitative evaluations of the vegetation covers. It is also important to choose indicators whose interpretation is easy, in order to understand why an anomaly has been detected. Our objective is not to analyze the various VIs available in the literature for crop monitoring but rather to propose a generic approach for detecting anomalies in different crops. This paper considers five benchmark optical VIs reported in Table 1.

The Normalized Difference Vegetation Index (NDVI) is a benchmark indicator for agronomic analysis that is mainly related to the plant vigor (Rouse et al., 1974; Bannari et al., 1995). The Normal Difference Water Index (NDWI) actually refers to two different benchmark indicators. The first version uses Near infrared (NIR) and Short Wave infrared (SWIR) to monitor changes in the water content of leaves (Gao, 1996). The second version uses the green band and NIR to monitor changes related to content in water bodies (McFeeters, 1996). Both formulas are similar to NDVI with different bands involved. The SWIR version of NDWI seems to be more appropriate for crop analysis. However, the GREEN version of NDWI can also provide relevant information, *e.g.*, for flooded parcels. The Modified Chlorophyll Absorption Ratio Index (MCARI) was designed to extract information from the chlorophyll content in plants with a resistance to the variation of the Leaf Area Index (LAI). A variant called MCARI/OSAVI is using the Optimized Soil Adjusted Vegetation Index (OSAVI) to minimize the contribution of background reflectance (Wu et al., 2008; Daughtry et al., 2000). The Green Red Vegetation Index (GRVI) is similar to NDVI but uses the red and green bands. According to Motohka et al. (2010), GRVI *can be a site-independent single threshold for detection of the early phase of leaf green-up and the middle phase of autumn coloring*.

3.2.2. SAR indicators

Many works have been performed to establish a relationship between SAR images and vegetation. One can mention for instance two recent reviews reporting the advances in this area of investigation (McNairn and Shang, 2016; Liu et al., 2019). The relationship between SAR and vegetation is not straightforward as other factors such as soil moisture or soil structure also impact SAR signals. A Radar Vegetation Index (RVI) was developed in (Kumar et al., 2013) but it requires polarizations that are not available with S1 images. The backscattering coefficients (denoted as γ^0_{VH} and γ^0_{VV}) have been used intensively in the literature as for example in (Whelen and Siqueira, 2018; Khabbazan et al., 2019). Their ratio $\gamma^0_{\mathrm{VH}}/\gamma^0_{\mathrm{VV}}$ has also been used for vegetation analysis (Denize et al., 2018; Abdikan et al., 2016). It has the advantage of counterbalancing the effect of soil moisture (Veloso et al., 2017; Vreugdenhil et al., 2018). In addition to optical VIs, this paper considers the three SAR indicators summarized in Table 1.

Table 1
Vegetation indices computed from S2 images and SAR indicators computed from S1 images used in this paper. For S2, The near infrared (band 8), red edge (band 5), short wave infrared (band 11), green (band 3) and red (band 4) channels are denoted as NIR, RE, SWIR, GREEN and RED, respectively.

Sensor type	Indicator	Formula
Optical	NDVI	$\frac{\mathrm{NIR}-\mathrm{RED}}{\mathrm{NIR}+\mathrm{RED}}$
	$\mathrm{NDWI}_{\mathrm{SWIR}}$	$\frac{\mathrm{NIR}-\mathrm{SWIR}}{\mathrm{NIR}+\mathrm{SWIR}}$
	$\mathrm{NDWI}_{\mathrm{GREEN}}$	$\frac{\mathrm{GREEN}-\mathrm{NIR}}{\mathrm{GREEN}+\mathrm{NIR}}$
	$\frac{\mathrm{MCARI}}{\mathrm{OSAVI}}$	$\frac{(\mathrm{RE}-\mathrm{IR})-0.2(\mathrm{RE}-\mathrm{RED})}{(1+0.16)\frac{\mathrm{NIR}-\mathrm{RED}}{\mathrm{NIR}+\mathrm{RED}+0.16}}$
	GRVI	$\frac{\mathrm{GREEN}-\mathrm{RED}}{\mathrm{GREEN}+\mathrm{RED}}$
SAR	Cross-polarized backscattering coefficient VH	γ^0_{VH}
	Co-polarized backscattering coefficient VV	γ^0_{VV}
	Polarization ratio	$\frac{\gamma^0_{\mathrm{VH}}}{\gamma^0_{\mathrm{VV}}}$

3.3. Input data for the anomaly detection algorithms

3.3.1. Zonal statistics

The proposed approach detects crop anomalies at the parcel-level, through appropriate statistics referred to as "zonal statistics" (*i.e.*, spatial statistics). Zonal statistics are computed from S1 and S2 indicators and summarize the information contained in each parcel (that are determined using the parcel delineation). Multiple statistics were considered such as the median, interquartile range (IQR), skewness and kurtosis. The median allows the mean behavior of a parcel to be captured (with more robustness than the classical mean) and can thus detect global anomalies such as anomalies in crop vigor. On the other hand, IQR contains important information related to heterogeneity. IQR is equal to the difference between the 75th and 25th percentiles. These two statistics are known to be robust to outliers that can affect the extracted features. Skewness aims at measuring the asymmetry of a distribution whereas kurtosis can be used to characterize the tail of a distribution. Both measures are complementary to median and IQR to analyze the distribution of the pixel within a crop parcel. These statistics were extracted for each indicator at each date for each parcel using the python libraries SciPy (Virtanen et al., 2020) and rasterstats[4]. As mentioned in Section 3.1, cloud pixels were discarded meaning that these statistics were computed from pixels not filtered by the cloud mask.

3.3.2. Feature matrix

From the zonal statistics of the different indicators computed at different dates, a feature matrix is constructed and used as the input of point anomaly detection algorithms. Each parcel is represented by a vector concatenating all the statistics computed for the different indicators at different dates. The number of features for a given parcel is $N_f = N_s \left(N_{1,i} N_{1,f} + N_{2,i} N_{2,f} \right)$, where $N_{1,i}$ is the number of S1 images, $N_{1,f}$ is the number of features for each S1 image, N_s is the number of statistics and similar definitions apply to $N_{2,i}$ and $N_{2,f}$ for S2 images. For instance, in 2017/2018, if all S1 and S2 images are used and two statistics (median/IQR) are computed for all the indicators (5 optical indicators and 4 SAR indicators), the number of features is equal to $N_f = 442$. Various sets of features were tested and analyzed as presented in the next section. The construction of the feature matrix is illustrated in Table 2 when using a unique indicator NDVI with 2 statistics. As each column corresponds to a unique combination statistics/indicator/time, it is possible to compare each parcel columnwise.

As illustrated in Table 2, there is a high number of features, which can induce a high computa-

[4] https://pythonhosted.org/rasterstats/, online accessed 10 March 2020

Table 2
Simplified version of the feature matrix using NDVI only and two statistics (median/IQR) for n dates and M parcels. NDVI_{t_n} means NDVI computed for image #n whereas median_{P_M} means median of the indicator computed inside the parcel #M

Parcel #	Feature 1	Feature 2	.	Feature L-1	Feature L
P_1	$\text{median}_{P_1}(\text{NDVI}_{t_0})$	$\text{IQR}_{P_1}(\text{NDVI}_{t_0})$	.	$\text{median}_{P_1}(\text{NDVI}_{t_n})$	$\text{IQR}_{P_1}(\text{NDVI}_{t_n})$
P_2	$\text{median}_{P_2}(\text{NDVI}_{t_0})$	$\text{IQR}_{P_2}(\text{NDVI}_{t_0})$	.	$\text{median}_{P_2}(\text{NDVI}_{t_n})$	$\text{IQR}_{P_2}(\text{NDVI}_{t_n})$
...	...	...	.	...	...
P_M	$\text{median}_{P_M}(\text{NDVI}_{t_0})$	$\text{IQR}_{P_M}(\text{NDVI}_{t_0})$	.	$\text{median}_{P_M}(\text{NDVI}_{t_n})$	$\text{IQR}_{P_M}(\text{NDVI}_{t_n})$

tional complexity. In order to limit the data dimensionality, various approaches were tested. One can restrict the analysis to the most informative statistics and indicators (*e.g*, using only median statistics for S1 data) or by choosing specific dates. Another approach can be to use dimensionality reduction techniques in order to summarize all (or a part) of the features into a reduced set of interesting features. A benchmark method is the Principal Component Analysis (PCA), which has received a considerable attention in the literature (Jolliffe, 1986). Multidimensional Scaling (MDS) (Borg and Groenen, 1997) was also investigated in this study and provided very similar results to PCA.

The dataset considered in this work is also highly unbalanced with a majority of S1 data. Since SAR images are available regardless of the day time and cloud coverage, the number of corresponding features can be very important compared to the features associated with optical images. In practice, because important information regarding vegetation is contained in optical images and because redundant information is contained in SAR images acquired at close acquisition dates, it was observed that using a majority of S1 data leads to poor results (more details will be provided in Section 4). One way of reducing the weight of S1 images (compared to S2 images) is to use monthly averaged S1 images (which are simply obtained by averaging the values of each month). Using monthly averaged S1 images provided interesting results that will be analyzed. Other techniques summarizing the information contained in the feature matrix were tested (moving averages, polynomial curve fitting, applying PCA to S1 data only) but did not improve the analysis significantly. Various combinations of features with and without dimensionality reduction techniques have been considered and will be described in Section 3.6, devoted to experimental results.

3.4. Point anomaly detection algorithms

Four benchmark anomaly detection algorithms are tested in this paper: One-Class Support Vector Machine (OC-SVM) (Schölkopf et al., 1999), Local Outlier Factor (LOF) (Breunig et al., 2000),

its probabilistic extension Local Outlier Probabilities (LoOP) (Kriegel et al., 2009) and Isolation Forest (IF) (Liu et al., 2012). To run the experiments conducted in this study, the Python Scikit-learn implementations of OC-SVM, LOF and IF were used (Pedregosa et al., 2011). The Python library PyNomaly was used for the implementation of the LoOP algorithm (Constantinou, 2018). Details for each detection methods are given below, highlighting what makes them interesting for detecting anomalies in crops.

The LOF algorithm is an unsupervised anomaly detection method based on the nearest neighbors of the observed samples. The main idea behind LOF is that normal data instances occur in dense neighborhoods and anomalies occur far from their closest neighbors (Chandola et al., 2009). This algorithm compares the local density of each instance with that of its k-nearest neighbors (*knn*). More precisely, the LOF score is defined as the ratio between the average local density of the *knn* and the local density of the data instance itself. A LOF score higher than 1 means that the analyzed instance has a local density lower than its neighbors and is very likely an outlier. A LOF score close to 1 indicates that the data instance under test has a normal behavior. One major issue in nearest neighbors techniques is that they rely on the distance measure and the number of neighbors used in the analysis. Choosing a relevant distance is often challenging, and an irrelevant distance measure is not able to separate inliers and outliers. Another drawback of the LOF score is that the outlier factor is difficult to interpret and is not consistent over the complete dataset (*i.e.*,a given LOF score can correspond to different degrees of outlierness depending on the region where it is computed) (Kriegel et al., 2009). This problem is addressed by the LoOP algorithm which is also evaluated here. The LoOP algorithm uses local statistics that are less sensitive to the choice of *knn* and provide an anomaly score belonging to the interval $[0, 1]$. A detailed formulation of LOF is provided below to understand the idea behind anomaly detection based on nearest neighbors methods. The reader is invited to consult (Kriegel et al., 2009) for more details about the LoOP algorithm and its differences respect to LOF. Let knn(P) be the set of k nearest neighbors of an object P.[5] Let k-dist(P) be the distance between the object P and its *k-th* nearest neighbor. A mathematical formulation of the LOF score can be obtained in 3 steps. First, the reachability distance $\mathrm{rd}_k(P, O)$ of P with respect to O is defined as follows:

$$\mathrm{rd}_k(P, O) = \max\{d(P, O), \text{k-dist}(O)\}. \tag{1}$$

[5]The number of k-nearest neighbors of the object P denoted as |knn(P)| can be greater than k in the case of multiple instances located at the same distance.

For an object P far from O, the reachability distance is the actual distance between P and O. However, for closer objects, the true distance is replaced by k-dist(O) for stability reasons (Breunig et al., 2000). This smoothing effect can be controlled by the parameter k. Using the notion of reachability distance, a local reachability distance of an object P denoted as lrd(P) can be computed as follows:

$$\mathrm{lrd}_k(P) = \left[\frac{\sum_{O \in \mathrm{knn}(P)} \mathrm{rd}(P, O)}{|\mathrm{knn}(P)|} \right]^{-1}. \tag{2}$$

The local reachability distance $\mathrm{lrd}_k(P)$ is the inverse of the average reachability distance of the object P from its neighbors, *i.e.*, the average distance at which P can be reached from its neighbors. Finally, the LOF score of an object P is the average local reachability of its neighbors divided by the local reachability density of this object, i.e.,

$$\mathrm{LOF}_k(P) = \frac{\sum_{O \in \mathrm{knn}(P)} \left(\frac{\mathrm{lrd}(O)}{\mathrm{lrd}(P)} \right)}{|\mathrm{knn}(P)|}. \tag{3}$$

For both LOF and LoOP, k is a hyperparameter that has to be fixed. Note that LoOP require to adjust an additional parameter known as the "extent" parameter and denoted as λ. This parameter defines the statistical notion of an outlier as an object deviating more than a given λ times the standard deviation from the mean (Constantinou, 2018).

OC-SVM (Schölkopf et al., 1999; Tax and Duin, 2004) is a model-based technique, which assumes that all training instances are part of the same class delimited by a separating boundary and that the instances that are not inside the learned boundary are anomalies (Chandola et al., 2009). One important difference between the OC-SVM algorithm and the three other algorithms considered in this work is that OC-SVM was initially developed for novelty detection, which is slightly different from anomaly detection (or outlier detection). In novelty detection, a model is learned using normal data and the aim is to detect if new unseen data points are novel (*i.e.*, abnormal compared to the training data). However, during the training phase, the estimated boundary frontier does not necessarily include all the training instances for robustness reasons (*i.e.*, to avoid overfitting). Training instances located outside the estimated frontier are then considered as anomalies. An upper bound for the fraction of training samples located outside the frontier has to be fixed by the user. The choice of this bound can largely influence the behavior of the classifier (Schölkopf et al., 1999; Pimentel et al., 2014).

The linear OC-SVM algorithm, as defined in Schölkopf et al. (1999) looks for the maximal margin hyperplane between the data points and the origin (Lampert, 2009). For a set of instances $x_i \in \mathcal{X}$ (where $\mathcal{X} = \{x_1, ..., x_n\}$ is the training set) with a hyperplane defined by $w^T x + b = 0$, the OC-SVM algorithm requires to solve the following problem:

$$\min_{w,\xi,\rho} \frac{1}{2}||w||^2 + \frac{1}{\nu n}\sum_{i=1}^{n}\xi_i - \rho. \tag{4}$$

subject to

$$w^T x_i \geq \rho - \xi_i \quad \text{and} \quad \xi_i \geq 0, \forall i = 1, ..., n. \tag{5}$$

The first parameter ν usually approximates the fraction of training errors and support vectors (it is an upper bound for the fraction of training errors and a lower bound for the fraction of support vectors). As mentioned before for outlier detection, ν may be difficult to choose if there is no prior information about the percentage of outliers. The slack variables ξ_i allow the classifier to create a soft margin in order to prevent from over-fitting. Indeed, the slack variable ξ_i allows the vector x_i to lie within the margin.

A kernel can be used to transforms the OC-SVM linear model into a non-linear model as for the classical SVM method. A Radial Basis Function (RBF) kernel was investigated in this paper defined for two vectors x and x' as

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right). \tag{6}$$

The choice of this kernel was mainly motivated by its effectiveness that has been observed in many different practical applications (Schölkopf et al., 2004; Kuo et al., 2014). The RBF kernel has a single hyperparameter referred to as kernel bandwidth and denoted as σ which has to be adjusted for each dataset. In practice, an efficient heuristic (Jaakkola et al., 1999; Aggarwal, 2017) consists of estimating the parameter σ as the median of the pairwise Euclidean distances between vectors from the learning set $\mathcal{X}$, denoted as median(dist($\mathcal{X}$)). An important difference exists between OC-SVM and the other algorithms regarding the anomaly score given at each instance. As explained before, in order to construct the OC-SVM boundary, the parameter ν has to be fixed by the user. If the percentage of anomaly

to be detected is changing with time, a new separating frontier has to be learned. Consequently, in the OC-SVM algorithm, anomaly scores can change according to the amount of anomalies to be detected. Conversely, for LOF, LoOP and IF, an anomaly score is given for each instance analyzed independently from the amount of abnormal parcels to be detected. Thus it is possible to select the percentage of anomaly to be detected by sorting the instances using their anomaly scores.

The IF algorithm uses a new concept called isolation that aims at detecting anomalies without using any distance or density measure, assuming that outliers can be isolated more easily than other instances. Using binary isolation trees to separate instances, outliers are more likely to be isolated at the root of the trees whereas inliers tend to be isolated at deeper parts of the trees. The IF algorithm constructs multiple random isolation trees defining a so-called forest of iTrees. The construction of an iTree is detailed in Algorithm 1. It can be observed that it is mostly random: at each node, a random feature is chosen with a random split value. When using random splits with random features, outliers are more likely to be isolated first. The number of splittings required to isolate an instance is called the path length. The anomaly score of a given instance can be defined from the averaged path length in the forest. Outliers will have a short average path length whereas inliers will be isolated with a large number of splittings. The IF algorithm is known to be very fast compared to other algorithms (especially for large datasets) since it does not compute a pairwise distance matrix.

Algorithm 1 iTree($\mathcal{X}, l, e$) (Liu et al., 2008)

Input: $\mathcal{X}$ - input data, e - current tree height, l - height limit
Output: an iTree
if $e \geq l$ or $|\mathcal{X}| \leq 1$ **then**
  return external-node{$Size \leftarrow |\mathcal{X}'|$}
else
  Let Q be a list of attributes (i.e., features) of $\mathcal{X}$
  1) randomly select an attribute $q \in Q$
  2) randomly select a split point $p \in [\min(q), \max(q)]$
  3) $\mathcal{X}_l \leftarrow$ filter($\mathcal{X}', q < p$), {all instances with attribute $q < p$}
  4) $\mathcal{X}_r \leftarrow$ filter($\mathcal{X}', q \geq p$)
  return internal-node{Left $\leftarrow$ iTree({$\mathcal{X}_l, e+1, l$),
                Right $\leftarrow$ iTree($\mathcal{X}_r, e+1, l$)
                splitFeature $\leftarrow$ q,
                splitValue $\leftarrow$ p}
end if

3.5. Performance evaluation

3.5.1. Labeling process

In order to validate the performance of the proposed strategy, having access to labeled parcels is mandatory. A partial labeling of the dataset was made by analyzing the parcels detected as abnormal by the different algorithms. In other words, each time a new set of features or a new algorithm tuning was tested, the parcels declared as anomalies were counterchecked by experts, confirming the anomaly or not, and determining the type of anomaly (see details later). Three main reasons motivated a partial labeling of the dataset: 1) labeling exhaustively a complete dataset would have been time consuming and not efficient because there is by definition a majority of normal parcels inside each dataset, 2) because the proposed approach is unsupervised, a complete labeling is not necessary to run an anomaly detection algorithm and 3) a relatively high amount of parcels was analyzed in each scenario in order to have a selection of the biggest anomalies present in the dataset. As a consequence, the partial labeling made in this work ensures a good balance between feasibility of the labeling and interpretation of the results.

The labeling was conducted by photo-interpretation with the help of agronomic experts from the TerraNIS company located in Toulouse, France. In addition to photo interpretation, the different time series were visualized for each indicator/statistic to compare an analyzed parcel to the rest of the data. Using time series indicators often allowed more precise interpretation of the anomalies. More details and examples regarding the labeling process can be found in the supplementary material (see Appendix A). Using photo-interpretation provided enough information to label with confidence a significant part of the parcels. The different anomalies detected throughout the study can be decomposed into 4 main categories: heterogeneity problems, growth anomalies, boundary errors and non-agronomic anomalies. A description of each category is presented in Table 3 and more details are provided below. The term "subtle" will be used to characterize transient anomalies that do not have a strong and global impact on the parcels and the corresponding features. Generally, only a limited time period of the full season is affected by a subtle anomaly.

- *Heterogeneity* corresponds to parcels presenting a clear heterogeneous development. The most common case of heterogeneity occurs during the majority of the season dates and can be related for instance to soil problems, presence of weed or diseases. More subtle cases of heterogeneity were observed at the beginning (*early heterogeneity*) or at the end of the season (*heterogeneity*

Table 3
Description of the different categories of detected anomalies during the labeling process. Subcategories were added to have a more precise description.

Category	Subcategory	Description
Heterogeneity	Heterogeneity	Occurs during most of the season
	Heterogeneity (2 different parts)	The parcel is separated into two homogeneous different parts
	Heterogeneity after senescence	Occurs during senescence
	Early heterogeneity	Occurs during early season
Growth	Late growth	A late development is observed (non-vigorous crop)
	Vigorous crop	A vigorous development is observed
	Early flowers	Early flowering phase
	Early senescence	Early senescence phase
	Late senescence	Late senescence phase
Boundary file error	Wrong type	A wrong type of crops was declared.
	Wrong shape	Wrong parcels boundaries were declared.
Non-agronomic	Too small	The parcel seems too small to be labeled accurately
	Not sure / unknown reason	The parcel is probably normal (no strong anomaly visible)
	SAR anomaly	Only SAR indicators seem abnormal
	Red channel problem	The optical red band has value equal to zero (only 1 image affected) causing abnormal indicators.
	Shadow perturbation (cloud or forest)	Shadows seem to cause abnormality in the optical indicators.

after senescence) and can be for instance related to differences in soil water content. A particular (and rare) case of heterogeneity (called *Heterogeneity (2 different parts)*) was observed when a parcel presents two homogeneous parts of the same crop separated by a clear frontier, one of the two parts being affected by a delay in its development. This anomaly can be explained by different crop varieties, different seeding dates or an error in the parcel boundaries.

- *Growth anomalies* are related to an abnormal development of the parcel. It can be beneficial (*i.e.*, *vigorous crops*) or undesirable (*i.e.*, *late growth*). It was observed that some highly vigorous crops are also highly homogeneous. Also, a difference between rapeseed and wheat crops was observed regarding late growth anomalies. For rapeseed crops, a late growth has generally no influence on the end of the growing season (*i.e.*, the flowering and senescence cycle is generally normal). On the other hand for wheat crops, a late growth problem generally affects all the growing season, leading to a global low vegetation development.
- *Boundary* issues are considered as relevant anomalies to be detected. This kind of anomaly can be considered as an agronomic anomaly because a wrong crop type is analyzed (the wrong parcel shape is due to the presence of multiple crop types inside the parcel). This category of anomalies presents in general a strong sign of abnormality and is easy to label and detect. In

practice, detecting such anomaly could be interesting to detect errors in the list of boundaries provided by farmers.

- The *"Not sure"* label concerns a majority of non-agronomic anomalies. This label was given to parcels that seem to be normal when inspecting the indicators and images. In some cases, some few extreme values have been observed but not sufficiently to be labeled as an anomaly. Because normality is difficult to be confirmed, the label *"not sure"* was used instead of "normal". The label *"Unknown reason"* concerns only few parcels where no apparent reason explained why the parcel has been detected as an anomaly. In any case, all these parcels should have a lower anomaly score than parcels affected by a real agronomic anomalies (*e.g.*, heterogeneity or growth anomaly).
- Other non-agronomic anomalies affect a small part of the analyzed parcels. Some very small size parcels were still present in the dataset and are labeled as *"too small"* (it is sometimes difficult to clean efficiently too small parcels that are long and narrow). Analyzing this type of parcel is not possible due to the resolution limit of Sentinel data. This problem has generally a low impact with a good selection of features and an efficient anomaly detection algorithm. *Shadow* from forests generally occurs at multiple dates and is detected when shadows cover a large part of the parcel. Shadows caused by clouds are generally due to an error in the cloud mask. Finally, a particular issue was observed in the S2 image acquired in March 27, 2016 (wheat crop analysis) where some parcels have a majority of pixels from the red channel equal to zero. More details regarding this problem are provided in the Supplementary Material (see Appendix A).
- A subcategory of non-agronomic anomalies are *SAR anomalies*. These anomalies are parcels where SAR indicators have an abnormal time evolution, whereas optical images and their indicators were counterchecked as normal. These types of anomalies are generally as non-agronomic because SAR data can be affected by other factors such as soil moisture or soil structure.

Real world examples of parcels and indicators associated with each category can be found in the supplementary material (see Appendix A).

3.5.2. Performance measures

The labeled dataset introduced in Section 3.5 allows the performance of the different anomaly detection algorithms for various sets of features to be compared. One metrics that can be used to

evaluate the quality of the results provided by an anomaly detection method is the precision defined as

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{7}$$

where TP and FP are the numbers of true positives and false positives, respectively. The precision expresses the percentage of detected parcels that are true positives (here, agronomic anomalies). Ideally, the precision should be as high as possible because it means that only interesting anomalies are detected. A major limitation of using precision alone is that it does not provide information regarding the non-detected parcels. Indeed, it is possible to have a good precision even if a lot of actual anomalies are not detected.

Another efficient way to compare various sets of features and different anomaly detection algorithms is to plot the precision against the outlier ratio (the percentage of anomalies to be detected). This paper uses the term "outlier ratio" instead of "contamination ratio", which is used in scikit-learn. A good algorithm (or feature choice) will globally have detection results with a higher precision for a given outlier ratio. In our case, because we do not know the percentage of anomalies to be detected, these curves are interesting for comparison purposes. These curves are similar to the Receiver Operating Characteristics (ROC). However, they have the advantage to be more adapted to crop monitoring (*i.e.*, a maximum outlier ratio can be fixed in order to keep realistic conditions). In any case, precision against outlier ratio and ROC curves were tested and compared. We observed that both curves lead to the same conclusions, with an easier interpretation for precision against outlier ratio curves. One limitation is that this representation does not give information regarding the distribution of the different detected categories. Two algorithms can have the same precision without detecting the same parcels (*e.g*, one algorithm can detect more heterogeneous parcels whereas another one detects more late growth anomalies). Using the distribution of the different types of anomalies detected for a given outlier ratio is a complementary way to address this limitation.

3.6. Conducted experiments

In this study, various experiments are conducted to evaluate the proposed procedure and its relevance for crop monitoring. The aim of these experiments is also to understand which factors have an influence on the detected anomalies. To that extent, different sets of features and algorithms are

compared in different initial configurations, which are reported in Table 4. Various tests are made for analyzing a complete season (*i.e.*, using all the images available for a season) to detect anomalies in the rapeseed and wheat parcels. Evaluating the data from a complete season is interesting to evaluate the capacity of the proposed approach to detect anomalies occurring at different periods of the crop growth, and to determine whether differences between the detected parcels are observable. In addition, some other tests are made with a lower amount of data (and in particular for a mid season analysis between October and February). Early detection can be of interest for warning purposes at the beginning of the growth cycle. Moreover, it gives more details on the effect of having only few data available for the analysis. The influence of the amount of parcels to be detected (called outlier ratio) is also tested to analyze the relevance of the anomaly score given to each parcel. A total number of 203 different tests were made for rapeseed crops. Fewer tests were made with the wheat dataset because the main idea was to determine wether the proposed algorithm can be applied with minor modifications to other kinds of crops or not.

Table 4
Summary of the different experiments conducted in the study to evaluate the proposed approach.

Evaluated crop type	Time interval	Section	Evaluated factors
Rapeseed	Complete season	4.3.1	Point anomaly detection algorithms
		4.3.2	Feature sets
		4.3.3	Outlier ratio
		4.3.4	Zonal statistics
		4.3.5	S1 features alone
		4.3.6	Missing S2 images
		4.3.7	External factor: sowing date
	Mid season	4.3.8	Feature sets, S1 alone, outlier ratio, algorithms
Wheat	Complete season	4.4	Feature sets, S1 alone, sowing date (not displayed: algorithm, outlier ratio)
	Mid season	-	(not displayed: feature sets, S1 alone)
Rapeseed	Complete season	4.5	Robustness to changes in parcel boundaries

The different feature combinations tested in this study are identified using abbreviations that are defined in Table 5 with representative examples. Each set of features is labeled using the indicators and statistics used for both S1 and S2 sensors. The potential use of PCA is also indicated, with the number of corresponding discriminant axes (e.g., PCA3 means PCA with 3 discriminant axes). Finally, two types of S1 images were tested, namely monthly averaged images in dB scale referred to as "monthly dB" and speckle filtered images (without monthly averaging) referred to as "filtered".

Table 5
Examples of abbreviations with their corresponding sets of features used for anomaly detection.

Abbreviated name	Features used
S2: all (median / IQR)	Median and IQR of the 5 S2 indicators listed in Section 3.2
S2: NDVI (median / IQR)	Median and IQR of NDVI
S2: all (median / IQR), S1: all monthly db (median)	Median and IQR of the 5 S2 indicators and median of the 3 S1 indicators listed in Section 3.2. S1 images used are monthly averaged in dB scale.
S2: all (median / IQR), S1: all monthly db (median), PCA3	3 dimensional PCA applied to median and IQR of all S2 indicators and median of all S1 indicators. S1 images used are monthly averaged in dB scale.

4. Results and discussion

4.1. Distribution of labeled parcels

Fig. 4 summarizes the distribution of the anomaly categories for both wheat and rapeseed crops, that will be used for performance evaluation. This figure shows that a majority of agronomic anomalies were detected for both wheat and rapeseed crops during the multiple tests conducted to evaluate the proposed approach. In particular, heterogeneity and growth problems are the most numerous detected anomalies for both types of crops, which is interesting from an agronomic point of view.

Approximately 40% of the rapeseed dataset was counterchecked by the photo-interpreter and at least 30% of this dataset can be considered as abnormal in an agronomic sense. The large percentage of agronomic anomalies encountered present in this dataset can be explained by the fact that there are subtle anomalies affecting a limited time interval of the growing season. These transient anomalies are for instance delays during the senescence phase. As explained before, fewer tests were made with the wheat dataset but almost 25% of the parcels was counterchecked.

4.2. Hyperparameter tuning

This section summarizes the different hyperparameter tunings that have been considered for each algorithm. The resulting values have been reported in Table 6.

For the LOF and LoOP algorithm, the number of *knn* was fixed by grid search to $k = 701$. This value provided detection results of higher precision compared to the other values tested (note that small changes in the value of k do not significantly affect the result). It was found that choosing a too small number of neighbors (*e.g.*, choosing an odd-valued integer close to the square root of the number of

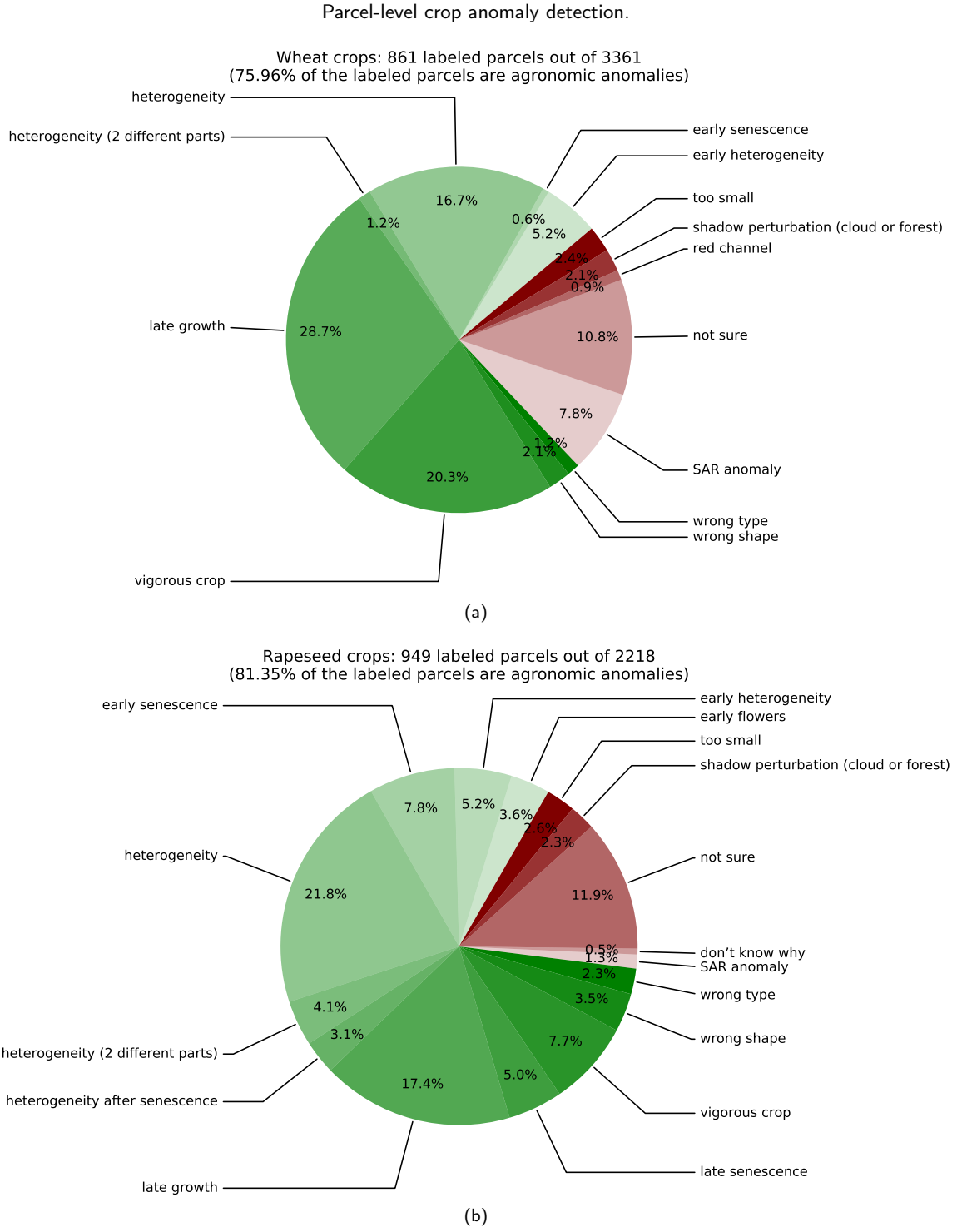


Fig. 4: Pie charts representing the distribution of the labeled parcels for wheat crops (a) and rapeseed crops (b). Green categories correspond to relevant agronomic anomalies considered as true positive. Red categories correspond to non-agronomic anomalies considered as not relevant.

observations in the data, for instance $k = 51$) leads to detect too subtle anomalies that are not related to agronomic issues. Choosing a relative large value of *knn* provides better results. The extent parameter of LoOP was fixed to $\lambda = 2$ (as recommended in (Constantinou, 2018)) and provided good results in all the analyses (no significant diffrences were observed when the value of λ was changed in a neighborhood of $\lambda = 2$).

For the OC-SVM algorithm, the estimator of σ presented in Section 3.4 provided good results without a need to a manual tuning for each new dataset.

The IF algorithm was used with a number of iTrees equal to $n_{trees} = 1000$ and a subsampling fixed to $n_{samples} = 256$ as in the original paper (Liu et al., 2008). Changing these two parameters did not have a significant effect on the results, which is a crucial advantage compared to LOF and OC-SVM. Indeed, it seems that the IF algorithm does not require a fine hyperparameter tuning. It is important because in real conditions and with an unsupervised detection strategy, a bad tuning might lead to poor results. Because they are using distances, the OC-SVM, LoOP and LOF algorithms also require a normalization in order to have input features in the interval $[0, 1]$, while this step is not mandatory when using the IF algorithm.

Table 6
Hyperparameters used in the different algorithms

Algorithm	Hyperparameter	Value
IF	n_{trees}	1000
IF	$n_{samples}$	256
LOF and LoOP	k	701
LOF and LoOP	λ	2
OC-SVM	σ	median(dist($\mathcal{X}$))

4.3. Anomaly detection results for rapeseed crops (2017/2018)

This section presents detailed results about anomaly detection for rapeseed parcels. Note that two different sets of features or algorithms sometimes provide similar results. Having similar results with different features or algorithms means that the global process is robust to changes and is not impacted by small differences in the initial configuration. In that case, simplest configurations should obviously

be preferred.

4.3.1. Effect of the algorithm used for crop anomaly detection

The performance of the different anomaly detection algorithms (LOF, LoOP, OC-SVM and IF) was tested for a complete season analysis. A first experiment was made by computing precision vs. outlier ratio curves for the different algorithms. As explained in Section 3.4, for the OC-SCVM algorithm the parameter ν sets the maximum amount of parcels that can be outside the learned frontier. In practice, a smaller amount of outliers can be detected (even if it is in general very close to ν). For the others algorithms, an anomaly score was attributed to each parcel. It was then possible to choose the outlier ratio by sorting the parcels according to their anomaly score (in order to select an appropriate percentage of parcels to be detected). Fig. 5 shows that all the anomaly detection algorithms provide similar precision for a given outlier ratio, confirming that multiple methods can lead to similar accuracy. Another observation is that a majority of agronomic anomalies is detected even with high outlier ratios (*e.g.*, 30%).

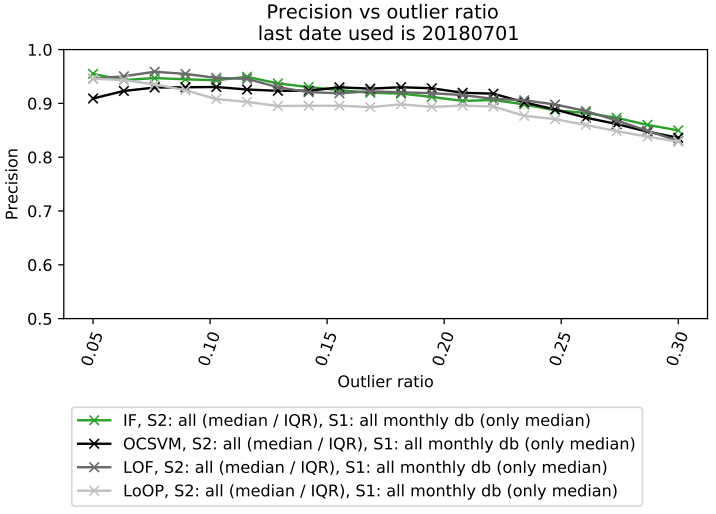


Fig. 5: Precision vs. outlier ratio for various algorithms using the same set of features for the rapeseed dataset. The nomenclature provided in Section 3.6 is used for each name of feature set.

To analyze potential differences within the categories detected, the 4 algorithms were run with an outlier ratio fixed to 20%. The numbers of detected anomalies within each category for LOF, LoOP, OC-SVM and IF are displayed in Fig. 6. The histograms obtained with LOF, LoOP and IF are very similar, whereas OC-SVM seems to detect less heterogeneity anomalies in favor of anomalies related

to delay in growth (late growth, vigorous crop, senescence anomalies). As discussed in Section 4.3.2, it is more logical to detect heterogeneity issues having a global impact on the crop season before senescence problems, which occurred only at the end of the season. To that extent, results obtained with the LoOP, LOF and IF algorithms are more coherent. In what follows, the IF algorithm will be used since its hyperparameters are easy to adjust. Another comparison between these algorithms is made in Section 4.3.8 for a mid season analysis.

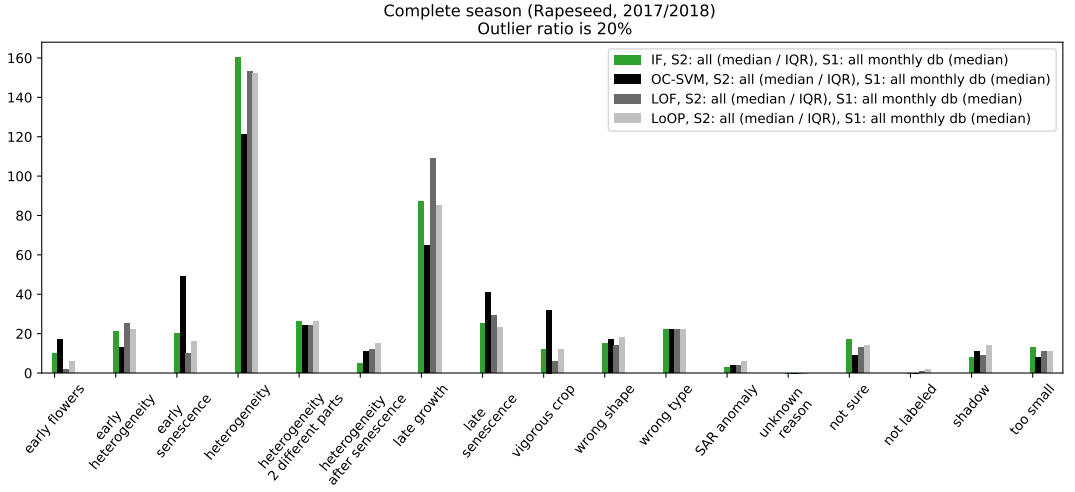


Fig. 6: Numbers of detected anomalies within each category for rapeseed parcels using various anomaly detection algorithm for a complete season analysis. All the categories at the right side of the category "wrong type" are considered as non-agronomic anomalies (false positives) as explained in Section 3.5. The nomenclature provided in Section 3.6 is used for each name of feature set.

4.3.2. Effect of the feature set used for crop anomaly detection

Fig. 7 shows the precision against outlier ratio for a selection of the best feature sets resulting from S1 and S2 data using the IF algorithm. One can notice that the different feature sets lead to similar results in term of precision. Additionally, even with relatively high outlier ratios it is possible to detect a majority of true positives (*i.e.*, more than 90% of relevant agronomic anomalies). One important result is that using only S2 indicators or even NDVI alone (with median and IQR statistics) was sufficient to obtain accurate results in that case. Note also that the best results were obtained using median and IQR statistics. More details about the test statistics will be presented in Section 4.3.4. Using jointly S2 and S1 indicators (*i.e.*, median and IQR of all indicators) leads to poor results with low precision, which are not displayed here. One explication is that most of the information is contained in S2 data and using a

majority of S1 data implies a decrease in the quality of the results. Using monthly averaged S1 images leads to results with good precision. It was observed that when there is a sufficient amount of S2 data, the use of S1 features does not improve detection results. Other cases where only few amount of S2 data is available will be discussed later. At this point, it is interesting to mention that using a global PCA (*i.e.*, apply PCA to the global feature matrix) provides good detection results for this example. However, further investigations showed that using global PCA was not adapted to datasets containing few images, which makes sense. This case will be discussed in the next sections.

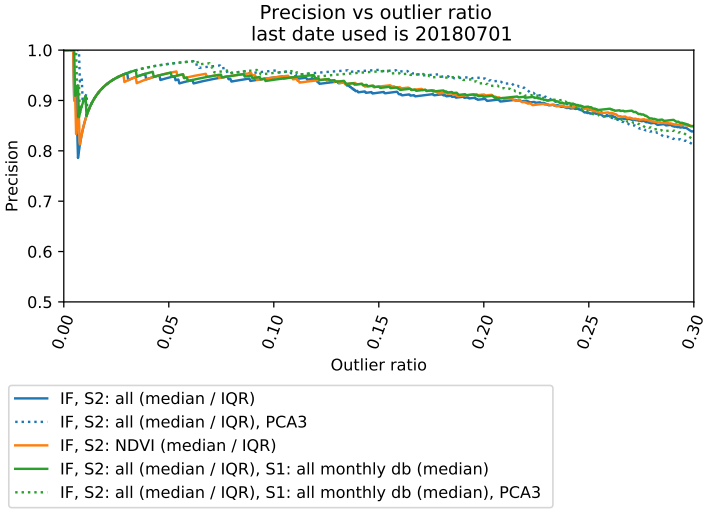


Fig. 7: Precision vs. outlier ratio for various feature sets using the IF algorithm for the rapeseed dataset.

In order to complement the results of Fig. 7, the numbers of detected anomalies within each category for an outlier ratio sets to 20% are depicted in Fig. 8. The different sets of features provide similar results, even if more growth anomalies (late/vigorous growth) and less heterogeneity are detected when using a 3 dimensional PCA for the whole dataset. A majority of the detected anomalies are heterogeneous or late growth parcels, which were identified as one of the most relevant agronomic anomalies to be detected in Section 3.5. More subtle heterogeneity (early/late) can be detected with a lower proportion.

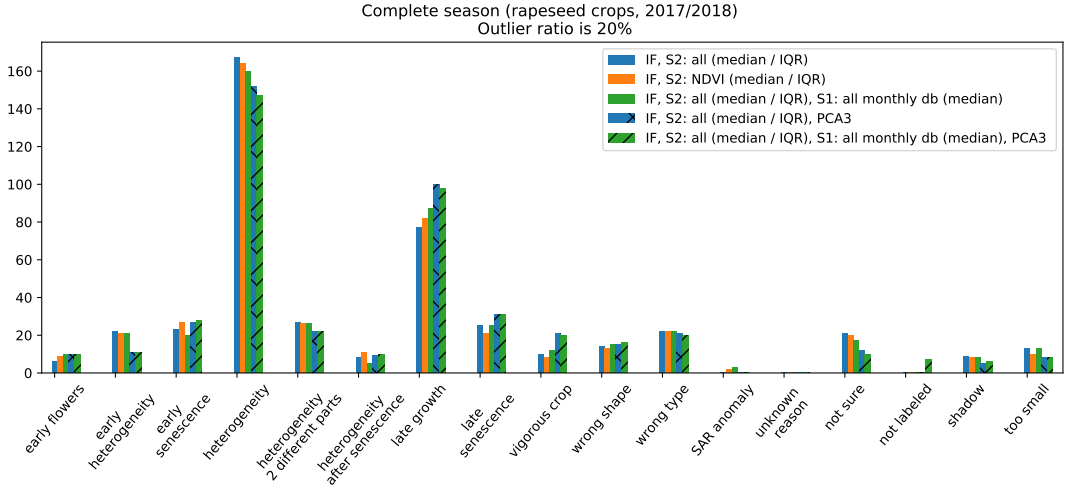


Fig. 8: Numbers of detected anomalies within each category for rapeseed parcels using various feature sets and IF for a complete season analysis.

4.3.3. Effect of the outlier ratio

Three experiments were run using the median and IQR statistics derived from S2 images and the IF algorithm, varying the outlier ratio in $\{0.1, 0.2, 0.3\}$. Fig. 9 shows the distribution of abnormal parcels in the different anomaly categories for each of these experiments. Interestingly, for a low outlier ratio (10%), the detected anomalies are mostly concentrated in wrong types, late growth and global heterogeneity. Increasing the outlier ratio leads to detect anomalies affecting smaller time periods of the season, such as early flowering and senescence problems. Finally, for an outlier ratio of 30%, much more false positives are observed (labeled as "not sure" or unlabeled). Even if more than 200 tests were made with various feature sets and algorithms, these unlabeled parcels can be considered as normal or are affected by a subtle anomaly. These results are coherent with the observations made in Section 3.5 during the labeling process. Indeed, during this labeling, a separation was made for instance between global heterogeneity and early or late heterogeneity, which was considered as more subtle. The same observations were made for senescence anomalies. These results show that the IF algorithm provides a relevant anomaly score. Particularly, more severe anomalies are detected with smaller outlier ratios (*i.e.*, they have a highest anomaly score). It is interesting because in practice estimating an optimal outlier ratio would be difficult as it depends on the farmer's needs. Moreover, because the score given by IF is computed only once, there is no need to run the algorithm several times when changing the outlier ratio.

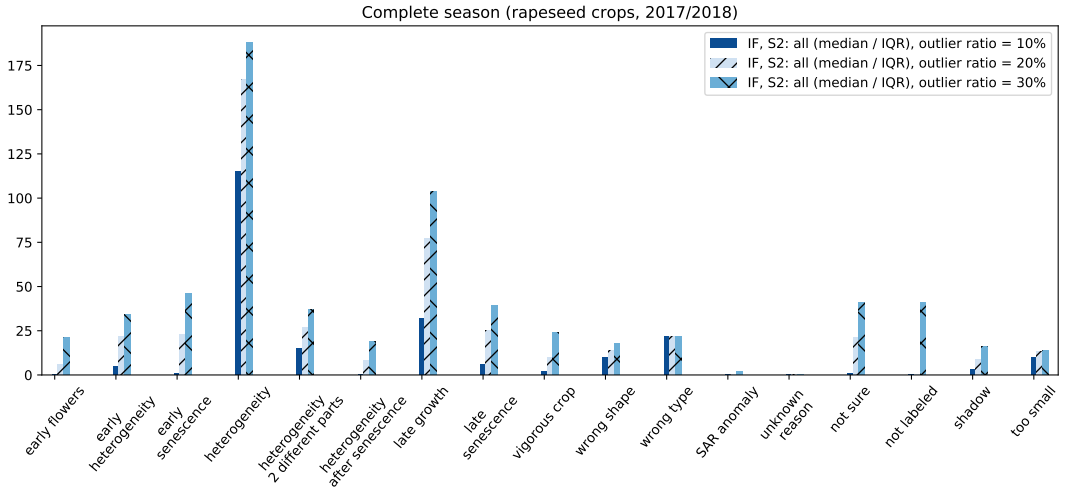


Fig. 9: Histograms of anomaly categories detected in the rapeseed parcels for a complete season analysis and the same set of features for various outlier ratios.

4.3.4. Effect of adding new statistics

All the previous experiments were conducted using the median and IQR as statistics (computed at the parcel-level). This section investigates two new statistics, namely the skewness and kurtosis (*i.e.*, the normalized third and fourth order moments of the indicators). Fig. 10 shows the Precision vs. outlier ratio when using the IF algorithm and these two additional statistics computed from S2 images to detect anomalies in rapeseed parcels. One issue encountered when using these new statistics is the detection of too subtle anomalies that are not always related to agronomic anomalies. Moreover, the precision when using skewness and kurtosis is not improved. Using only median is tested too, showing that IQR is an important statistics for anomaly detection.

4.3.5. Can we use S1 data alone?

For a complete season analysis, as shown in Section 4.3.2, no improvement is observed when using S1 and S2 data jointly. However, using S1 data alone could be needed when S2 images are not available (*e.g*, due to presence of clouds). Fig. 11 illustrates various precision vs. outlier ratio curves obtained using S1 images alone. The main conclusion is that for a relatively small outlier ratio (10%), using S1 data alone is sufficient to detect interesting agronomic anomalies. This confirms that reducing the outlier ratio leads to more precise detection and that the anomaly scores are relevant. Moreover, for a high outlier ratio (20%), an acceptable precision (close to 80%) can be reached depending on the use

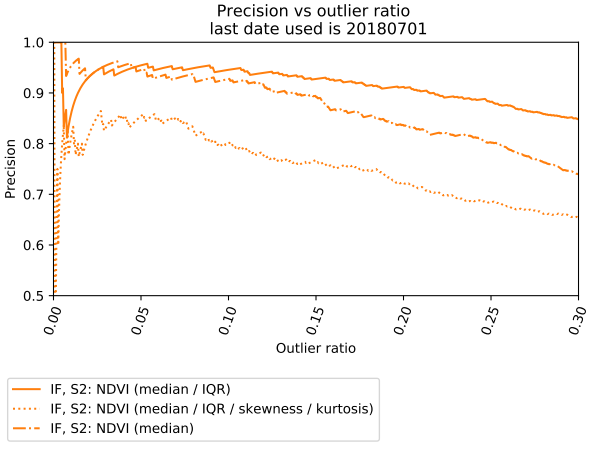


Fig. 10: Precision vs. outlier ratio for rapeseed dataset. Note that all the parcels are labeled for outlier ratios smaller than 10% and unlabeled parcels were considered as false positives.

case. Using monthly averaged images leads to slightly better results than those obtained with speckle filtered images. In addition, monthly averaged images are processed much faster and can be used easily with S2 images. When using S1 data only, the median statistics provides better results than the pair (median, IQR). Another interesting remark is that using the ratio VH/VV alone provides poor detection results. It means that VV and VH polarizations contain relevant additional information. Finally, PCA does not help when using S1 data alone, contrary to S1 and S2 images used jointly.

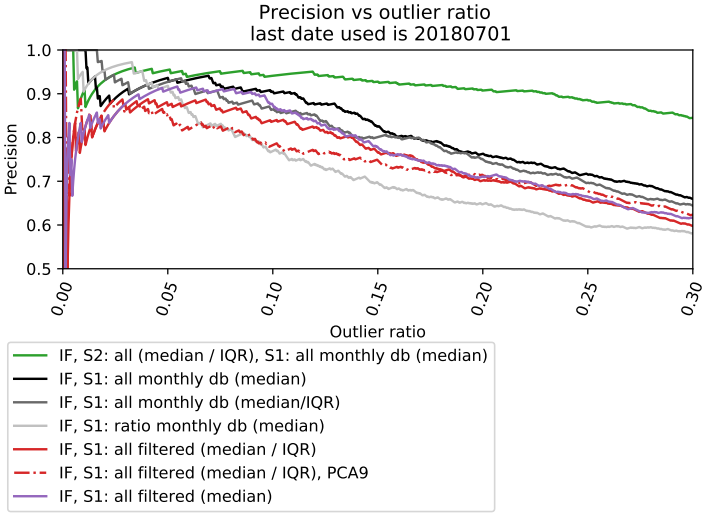


Fig. 11: Precision vs. outlier ratio for the rapeseed dataset. Note that all the parcels are labeled for outlier ratios smaller than 10% and unlabeled parcels are considered as false positives.

The distribution of the detected anomalies in each category is depicted in Fig. 12. For this experiment, the outlier ratio was fixed to 10% because higher values yield results with a lower precision, as shown in Fig. 11. In Fig. 12 it can also been observed that S1 data are less adapted to detect heterogeneity problems, which explain why it was difficult to detect a large amount of relevant anomalies using S1 data only.

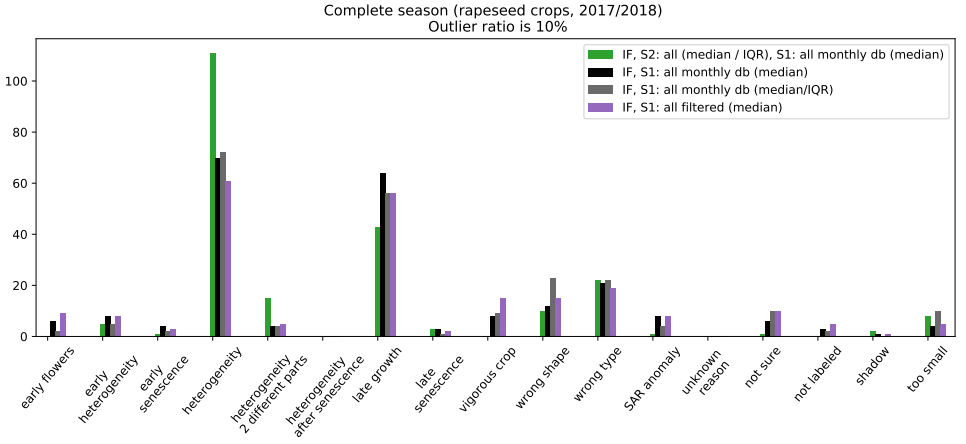


Fig. 12: Numbers of detected anomalies within each category using a complete season of rapeseed crops with various feature sets and the IF algorithm for a outlier ratio fixed to 10%.

4.3.6. Effect of missing S2 images

In order to evaluate the effect of missing S2 images, anomaly detection was investigated using 6 S2 images instead of 13. Only 1 image out of 2 was considered for the detection (the first S2 image was not used, the second S2 image was used and so on). Precision vs. outlier ratio curves are presented in Fig. 13, where it can be observed that the proposed strategy is robust to missing S2 images. This result shows that a lot of information can be extracted by using few S2 images.

Another experiment is conducted to evaluate the impact of using S1 data to complete sparse S2 images. Precisely, we consider 5 dates of S2 data between May and June that are used jointly with all S1 images. The numbers of detected anomalies with and without using S1 data are shown in Fig. 14. In the extreme case where no S2 image is available at the beginning of the season, using S1 images help to detect almost the same amount of late growth crops when compared to using a complete season of S2 images. As a consequence, S1 data are well suited to detect growth anomalies and can be of interest when there is a lack of S2 images.

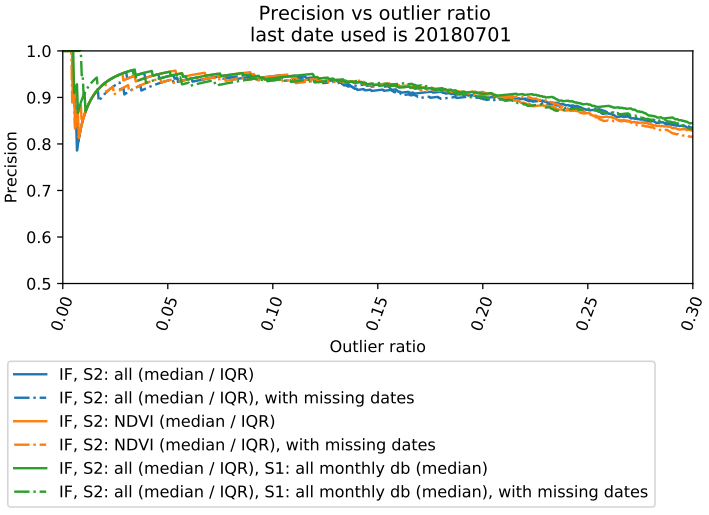


Fig. 13: Precision vs. outlier ratio, rapeseed dataset. Missing dates means that only 1 images out of 2 was taken (6 S2 images instead of 13).

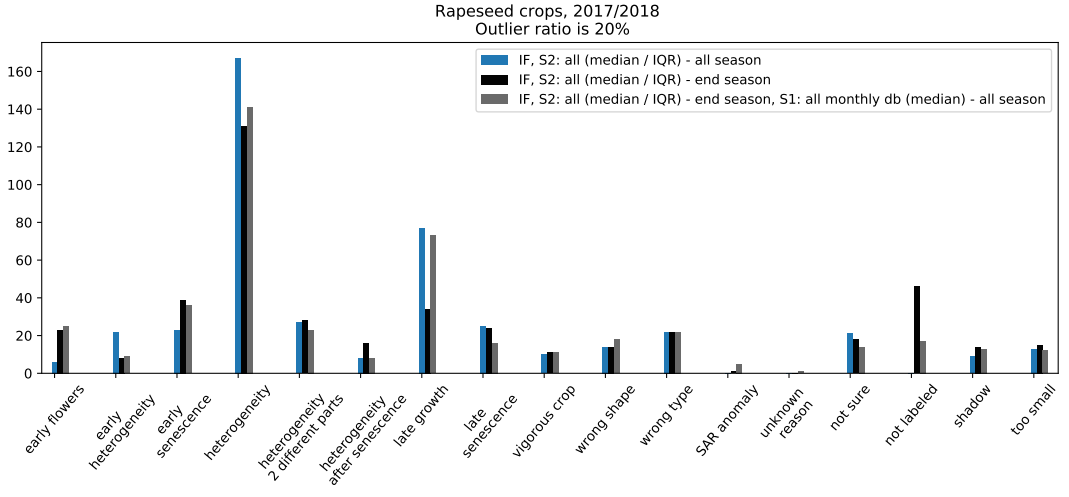


Fig. 14: Histograms of anomalies detected within each category for the rapeseed dataset. End season means that only 5 dates between May and June are used. Dates 23/06 and 28/06 were not used for this experiment to limit the impact of senescence anomalies.

4.3.7. Effect of external factors: sowing date

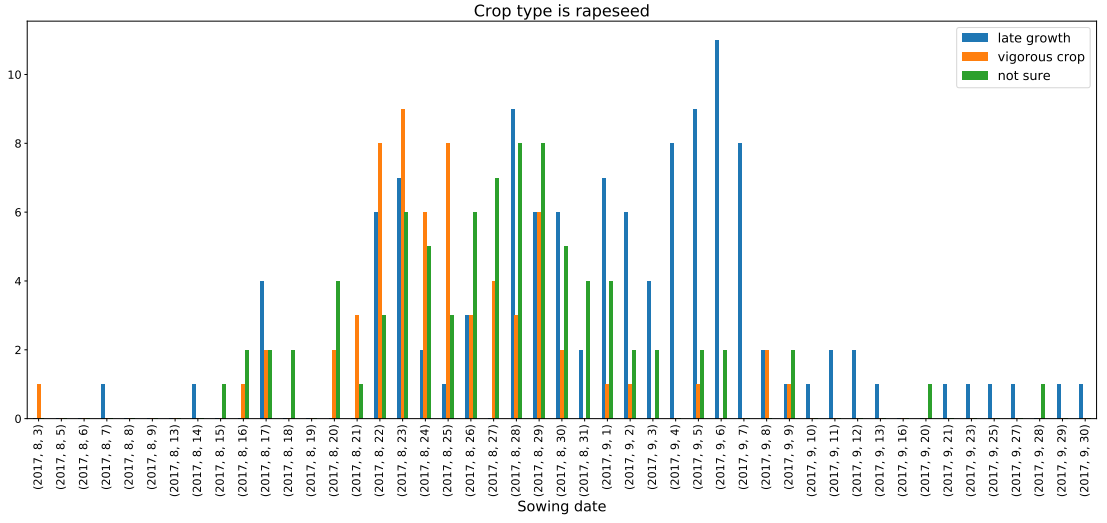
A large part of the parcels analyzed in this study had complementary information regarding for instance crop variety, sowing density and sowing date. It is interesting to verify if the detected anomalies can be explained by information already known by the farmers. The numbers of anomalies detected for a given sowing date are displayed for different categories in Fig. 15.a. Parcels with late seeding

dates are clearly more affected by late growth problems whereas vigorous crops are more likely to have an early seeding date. However, it also appears that a majority of parcels with the same sowing dates are affected by both anomaly types. Moreover, Fig. 15.b puts these results in perspective with the whole dataset. A majority of never detected parcels (which can probably be considered as normal) have the same sowing date as late growth or vigorous parcels (even if for extreme dates the proportion of observed anomalies is higher). Therefore, it can be concluded that the sowing date is a factor of anomaly but is not sufficient to explain all the anomalies encountered by the proposed approach. It is particularly the case for rapeseed crops, which have a relative short sowing period (the sowing period for wheat is much larger and thus its effect is stronger). Finally, we were not able to confirm any link between other external factors (such as crop variety or sowing density) and the detected anomalies. The conclusions of this section show that the proposed anomaly detection method provides relevant information that could be of interest for farmers.

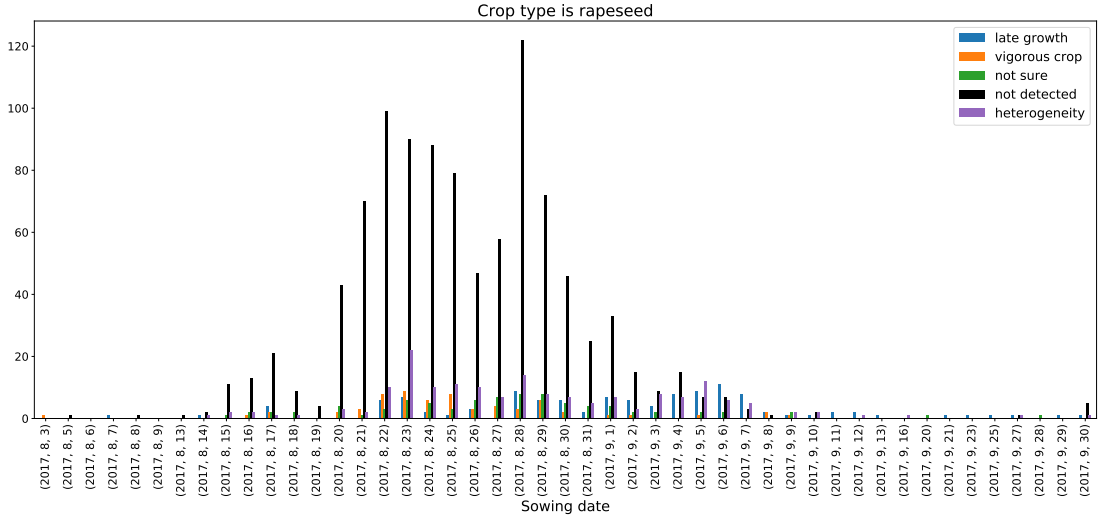
4.3.8. Mid-season analysis

A mid-season analysis (using only dates before February) was conducted for multiple reasons detailed in Section 3.6. A first experiment was made with the best sets of features selected in Section 4.3.2 for a complete season analysis using rapeseed parcels. Results displayed in Fig. 16 show that even with a small number of images, many agronomic anomalies are detected. This confirms the previous results found in the case of missing S2 images. Fig. 16 also shows that using S1 and S2 images jointly allows a slightly larger amount of actual anomalies to be detected compared to using S2 images only, especially for high outlier ratios (30%). Using NDVI alone provides results of lower precision for outlier ratios higher than 20%. Thus, using additional S2 indicators provides limited but relevant information about rapeseed crops. It was not possible to confirm anomalies detected only by SAR data because no optical image was available between December and February, as explained in Section 3.5. Finally, the use of PCA for this use case is not recommended since less anomalies are detected.

Another experiment was conducted to analyze the impact of a mid-season analysis regarding the different categories of anomalies detected. Results are depicted in Fig. 17. The mid-season analysis does not allow senescence problems to be detected, which is normal. Even with only 3 S2 images acquired between October and December, most other agronomic anomalies are detected by the algorithm. A mid-season analysis provides more late growth anomalies and fewer heterogeneous parcels because late growth is impacting mostly the beginning of the season (especially for rapeseed crops).



(a)



(b)

Fig. 15: Histograms of the sowing dates for various anomaly categories in rapeseed crops. a): only late growth and vigorous crops anomalies. b): various categories of anomalies. "Not detected" refers to a parcel that has never been detected during the study and that was not labeled.

Using S2 and S1 images jointly has a reduced impact on the detected anomaly categories when the outlier ratio is set to 20%.

The IF algorithm was run for different outlier ratios in the case of a mid season analysis, yielding the results presented in Fig. 18. As for a complete season analysis, increasing the outlier ratio logically leads to detect more subtle anomalies. Few early heterogeneity and vigorous crops are detected with

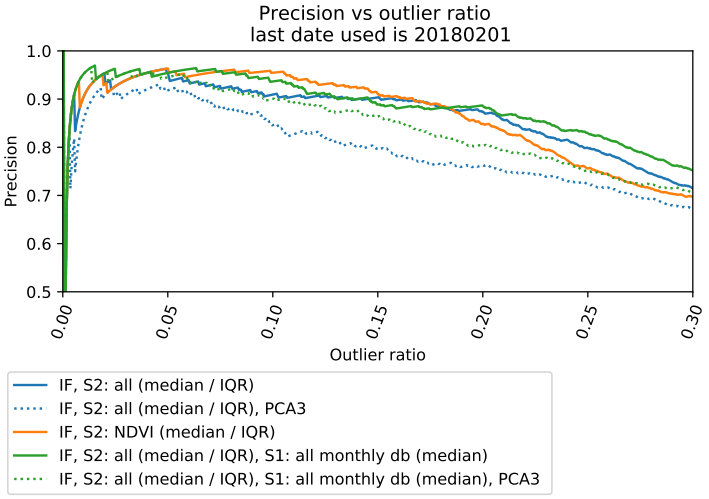


Fig. 16: Precision vs. outlier ratio for a mid-season analysis of rapeseed parcels (all images available before February) for various sets of features and the IF algorithm.

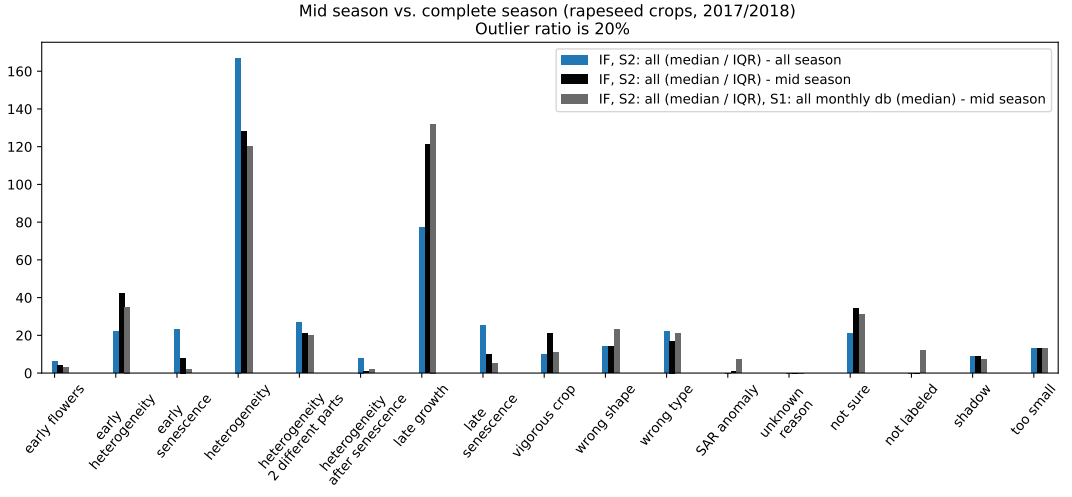


Fig. 17: Numbers of detected anomalies within each category in the rapeseed dataset for a mid season analysis and a complete season analysis.

an outlier ratio of 10%. Early heterogeneity is a more subtle anomaly than global heterogeneity, which confirms the relevant separation between these two categories. For vigorous crops, it was also observed during the labeling that this type of anomaly was not very strong when occurring with rapeseed crops.

Using S1 data alone was also tested for a mid season analysis. The detection results obtained for an outlier ratio of 10% are still accurate, as for the complete season analysis and are thus not presented here. These results confirm that S1 images are adapted to an early season analysis.

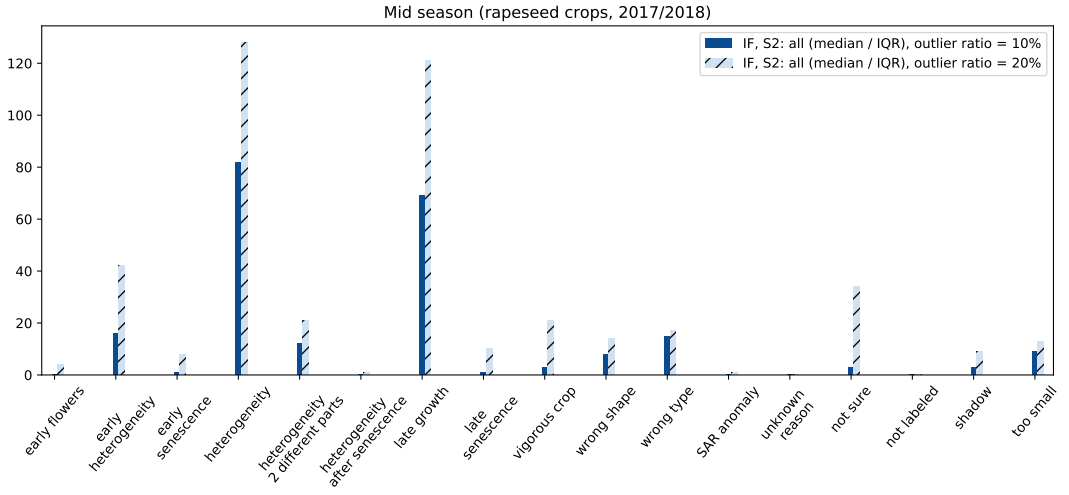


Fig. 18: Numbers of detected anomalies within each category using the IF algorithm and different outlier ratios for a mid-season analysis (images acquired before February).

The effect of changing the anomaly detection algorithm was also tested for a mid-season analysis, as shown in Fig. 19 (obtained for an outlier ratio of 10%). More unlabeled and normal (*i.e.*, labeled "not sure") parcels are detected with LOF and OC-SVM algorithms. This analysis shows that even with a low outlier ratio, the IF algorithm can detect many actual agronomic anomalies and is more robust to changes. As for a complete season analysis, the OC-SVM algorithm detects more vigorous crops than the other algorithms and is more affected by delay in growth anomalies. These differences in performance can be explained by the fact that OC-SVM, LoOP and LOF algorithms are more difficult to tune compared to the IF algorithm.

4.4. Extension to wheat crops (2016/2017)

A complementary analysis was made to measure the robustness of the proposed method to a change in the crop type. A first experiment is presented with the selection of the best sets of features used for rapeseed crops analysis. The IF algorithm was used to detect abnormal wheat parcels for a complete growing season with an outlier ratio of 10%. The histogram of anomaly categories is presented in Fig. 20. The percentage of (actual) detected anomalies is high (near 90%) and all the different sets of features lead to similar results. Using S2 images only provides accurate results. In that case, NDVI provides slightly better results than the other feature sets, mainly because it is not affected by the red channel problems (NDVI is saturated in that case). Finally, using S2 and monthly averaged S1 images

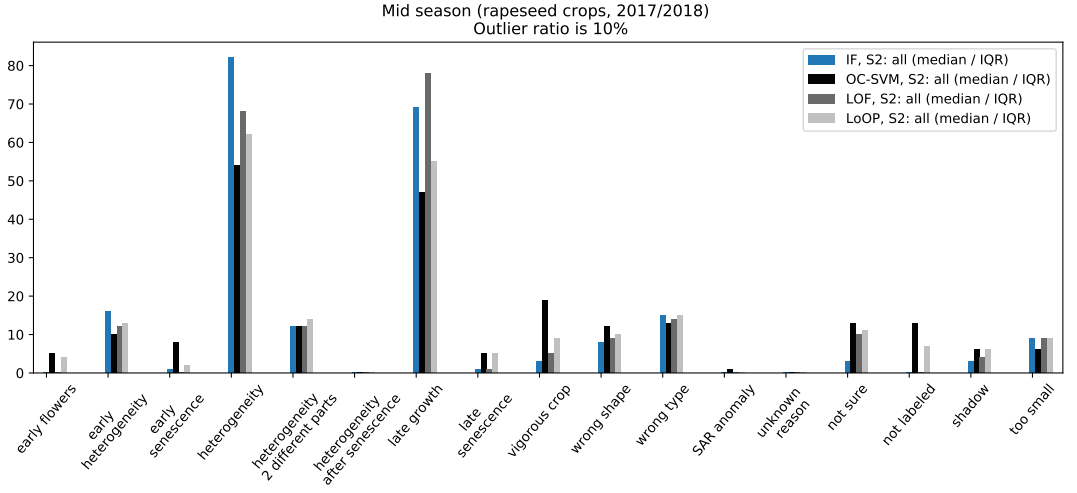


Fig. 19: Numbers of detected anomalies within each category using different detection algorithms for a mid-season analysis (images acquired before February).

jointly leads to detect less heterogeneous parcels and more parcels affected by growth anomalies. These results confirm the interest of the proposed approach and its robustness to changes in crop type.

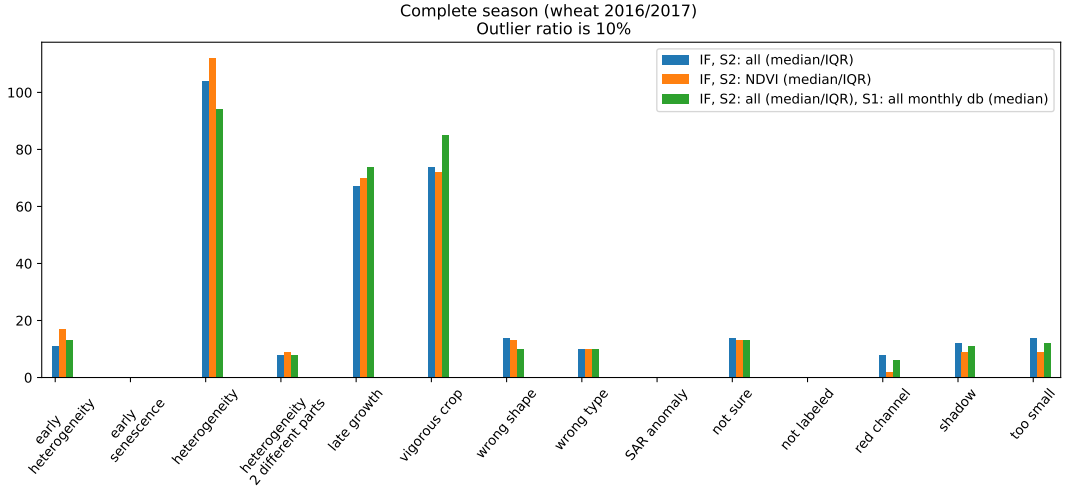


Fig. 20: Numbers of detected anomalies for wheat crops during a complete season.

The next experiment consists of detecting anomalies in wheat crops using S1 and S2 data considered jointly or separately. Results are depicted in Fig. 21. Heterogeneity problems are difficult to detect using S1 images only, whereas late growth anomalies are well detected. These results are in agreement with those obtained with rapeseed parcels.

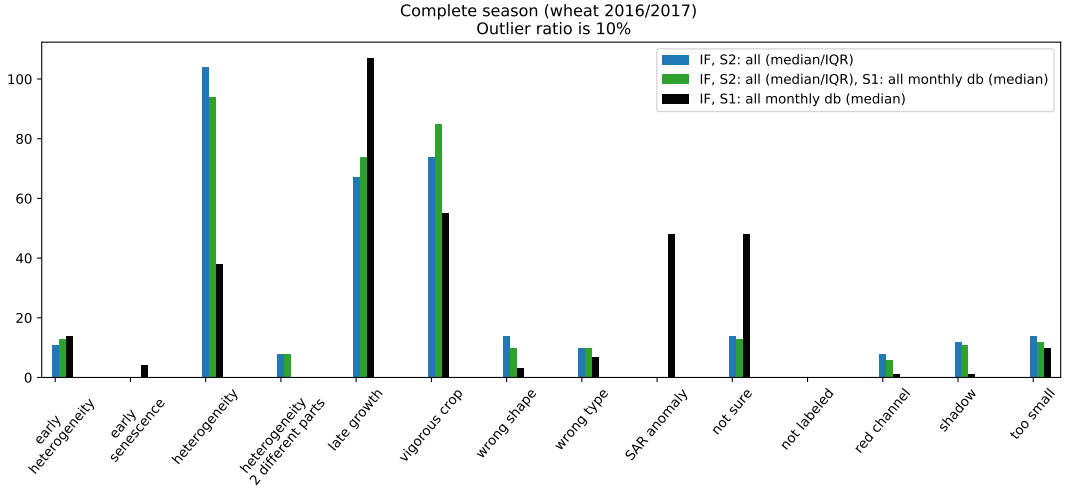
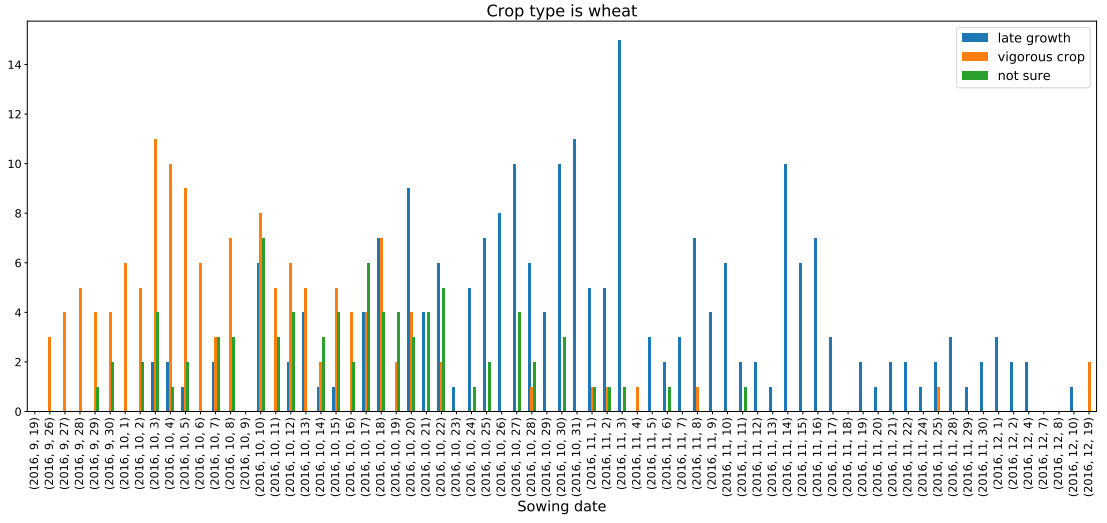
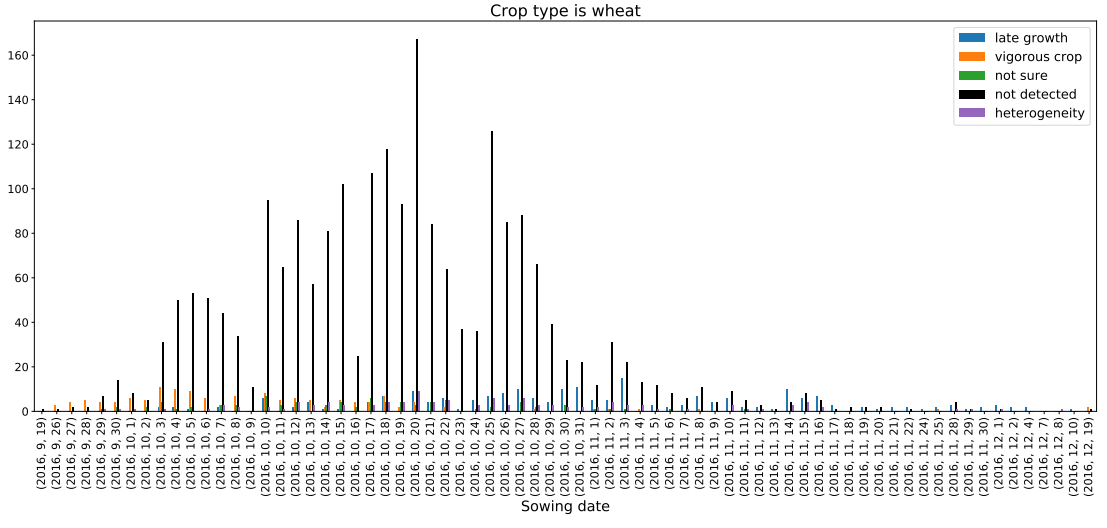


Fig. 21: Histograms of anomaly categories detected for wheat crop parcels during a complete season analysis.

Finally, the effect of sowing dates was analyzed in the case of wheat crops and is displayed in Fig. 22. A correlation between the sowing dates and late growths or vigorous crops can be observed. It may be due to a larger sowing period for wheat crops than for rapeseed crops. However, as for rapeseed crops, having the sowing dates only is not sufficient to detect and characterize an anomaly. Moreover, no correlation between the detected anomalies and other external factor was found in the case of wheat crops (confirming the results obtained for rapeseed crops). This analysis shows that the proposed strategy provides complementary information with respect to external factors such as sowing dates for the two analyzed crop types.



(a)



(b)

Fig. 22: Histograms of the sowing dates for different categories of anomalies detected in wheat parcels. a): various categories of anomaly ("Not detected" parcels means that these parcels were never detected during the study and were not labeled by photo-interpreters. b): late growth and vigorous crop anomalies.

4.5. Robustness to changes in parcel boundaries

As explained in Section 2, another parcel delineation coming from the French LPIS was tested to measure the consistency of the proposed approach to changes in parcel boundaries. More details about these two databases are provided in Section 2. Examples of parcel delinations obtained with LPIS and the Terranis parcellation system are depicted in Fig. 23. The parcel frontiers obtained using LPIS are

generally less accurate than those resulting from the Terranis systemn motivating the use of a buffer around the different parcels and robust zonal statistics.

Fig. 23: Example of parcel boundaries (rapeseed crop, growing season 2017/2018). In orange: custommer database, in green: LPIS database.

Anomaly detection was run with an outlier ratio of 20% using these two different databases. The numbers of detected anomalies for each category are depicted in Fig. 24. No significant difference can be observed when using the customer and LPIS parcels, showing that the proposed detection method is robust to this type of changes (probably because robust zonal statistics are used for anomaly detection).

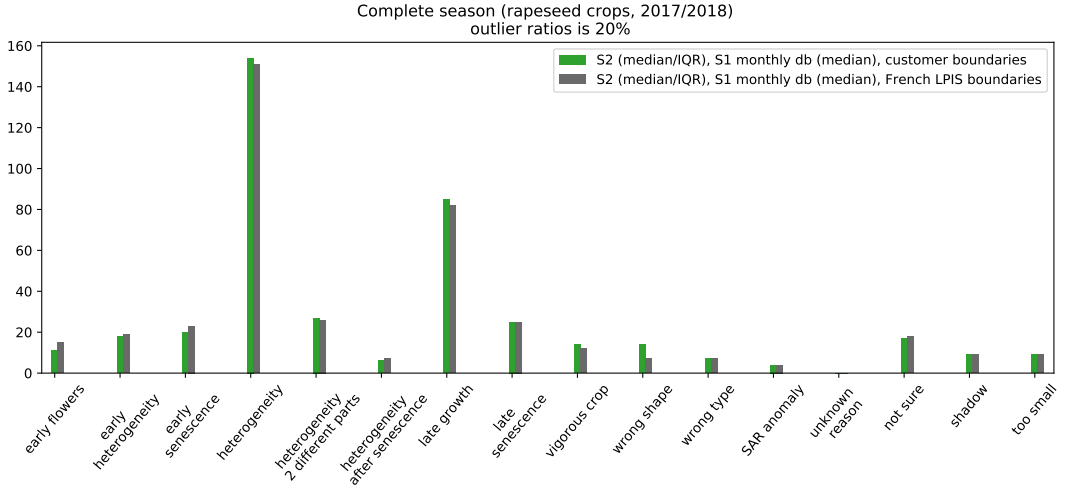


Fig. 24: Histograms of the detected anomalies within each category for parcels computed using LPIS and proprietary parcellation databases.

5. Conclusion

This paper studied a new anomaly detection method for crop monitoring. This method first computes multi-temporal vegetation indices from optical and synthetic aperture radar images. Zonal statistics are then computed at the parcel-level for each vegetation index at each date. A point anomaly detection is finally investigated to identify abnormal parcels from multi-temporal zonal statistics computed at a parcel-level. The proposed approach is unsupervised and can be used without a large amount of historical data. It can be applied to different kinds of crops such as rapeseed or wheat and is able to detect a majority of parcels that are abnormal in an agronomic sense. Moreover, a relevant anomaly score can be defined for each parcel allowing the importance of this anomaly to be quantified.

This study shows that multi-temporal vegetation indices computed at the parcel-level from optical images such as Sentinel-2 images provide relevant results for detecting anomalies in crops. These indices include the Normalized Difference Vegetation Index (NDVI), which can detect a majority of agronomic anomalies. Adding other indicators such as the Green-Red Vegetation Index (GRVI), the Normalized Difference Water Index (NDWI) and a variant of the Modified Chlorophyll Absorption Ratio Index (MCARI/OSAVI) allows additional and more subtle anomalies to be detected. Using indicators extracted from SAR images (such as Sentinel-1 images) jointly with optical indicators can help to detect subtle anomalies. Moreover, when an important part of the season misses optical images (such as the growing phase at the beginning of the season), SAR data are complementary to optical images and help to detect growth anomalies. In order to use SAR and optical images jointly, it is necessary to reduce the number of SAR indicators to avoid unbalanced datasets. This dimensionality reduction can be performed by averaging monthly SAR images or by using a principal component analysis. Using SAR images only can be suited for cloudy region with few optical images allowing growth anomalies (vigorous or late growth) to be detected. However, it was observed that heterogeneity problems are more difficult to be detected with SAR images only.

Spatial statistics computed at a parcel-level reduce the computational complexity of the algorithm and summarize efficiently the information regarding the analyzed parcels. More precisely, median and interquartile range (IQR) of vegetation indices (capturing the mean and dispersion of the indicators) computed at a parcel-level can be used efficiently for crop anomaly detection. An interesting property of the proposed method is its robustness to parcellation errors, which confirms that a parcel-level approach with robust statistics is adapted to crop monitoring.

An important conclusion is that point anomaly detection techniques are well suited to detect agronomic anomalies in crops, in an unsupervised way without a need for a large amount of historical data. Among all the tested anomaly detection methods, the isolation forest algorithm provided more consistent results with a simple hyperparameter tuning.

Further investigation should be conducted to determine whether other indicators (*e.g.*, using biophysical indicators such as Leaf Area Index (LAI) or fraction of green vegetation cover (fCover) (Poilvé, 2010; Berger et al., 2018)) can improve anomaly detection results. Another line of research is to take into account the temporal structure of vegetation indices to potentially improve anomaly detection and estimate the dates where the detected anomalies have appeared. For instance, contextual anomaly detection might be interesting for disturbance or inter-annual anomaly detection. Including a contextual anomaly detection in the proposed strategy might provide complementary information to detect both inter-annual and intra-annual anomalies. Coupling the proposed anomaly detection method with a supervised or unsupervised classification algorithm is another prospect, in order to assign to each detection an anomaly type that could be helpful for farmers, for example to identify heterogeneity or growth problems. Finally, online anomaly detection might be useful for practical implementation and near real-time (NRT) applications.

6. Acknowledgements

The authors would like to thanks Alexandre Bouvet and his colleagues from CESBIO (Centre d'Etudes Spatiales de la BIOsphère) for the help regarding the processing of S1 images and particularly for multi-temporal speckle filtering.

Appendix A Supplementary material

Supplementary data to this article can be found in the document provided in complement of the manuscript.

References

Abdikan, S., Balik Sanli, F., Üstüner, M., Calò, F., 2016. Land cover mapping using Sentinel-1 SAR data, in: Proc. ISPRS, Prague, Czech Republic. pp. 757–761. doi:10.5194/isprs-archives-XLI-B7-757-2016.

Aggarwal, C.C., 2017. Linear Models for Outlier Detection. Springer International Publishing, Cham. chapter 3. p. 93. URL: https://doi.org/10.1007/978-3-319-47578-3_3, doi:10.1007/978-3-319-47578-3_3.

Albughdadi, M., Kouamé, D., Rieu, G., Tourneret, J.Y., 2017. Missing data reconstruction and anomaly detection in crop development using agronomic indicators derived from multispectral satellite images, in: Proc. IEEE IGARSS, Fort Worth, TX, USA. pp. 5081–5084. doi:10.1109/IGARSS.2017.8128145.

Atzberger, C., Eilers, P.H.C., 2011a. Evaluating the effectiveness of smoothing algorithms in the absence of ground reference measurements. Int. J. Remote Sens. 32, 3689–3709.

Atzberger, C., Eilers, P.H.C., 2011b. A time series for monitoring vegetation activity and phenology at 10-daily time steps covering large parts of South America. Int. J. Digit. Earth 4, 365–386.

Bannari, A., Morin, D., Bonn, F., Huete, A.R., 1995. A review of vegetation indices. Remote Sensing Reviews 13, 95–120. URL: https://doi.org/10.1080/02757259509532298, doi:10.1080/02757259509532298, arXiv:https://doi.org/10.1080/02757259509532298.

Barbottin, A., Bouty, C., Martin, P., 2018. Using the French LPIS database to highlight farm area dynamics: The case study of the niort plain. Land Use Policy 73, 281 – 289. URL: http://www.sciencedirect.com/science/article/pii/S0264837717302909, doi:https://doi.org/10.1016/j.landusepol.2018.02.012.

Beck, P.S., Atzberger, C., Høgda, K.A., Johansen, B., Skidmore, A.K., 2006. Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. Remote Sens. Environ. 100, 321–334.

Berger, K., Atzberger, C., Danner, M., D'Urso, G., Mauser, W., Vuolo, F., Hank, T., 2018. Evaluation of the PROSAIL model capabilities for future hyperspectral model environments: A review study. Remote Sens. 10, 85. doi:10.3390/rs10010085.

Betbeder, J., Rémy, F., Philippets, Y., Ferro-Famil, L., Baup, F., 2016. Contribution of multitemporal polarimetric synthetic aperture radar data for monitoring winter wheat and rapeseed crops. J. Appl. Remote Sens. 10, 026020. doi:10.1117/1.JRS.10.026020.

Borg, I., Groenen, P., 1997. Modern Multidimensional Scaling. Springer-Verlag New York.

Breunig, M., Kriegel, H.P., Ng, R.T., Sander, J., 2000. LOF: Identifying density-based local outliers, in: Proc. ACM SIGMOD, Dallas, TX, USA. pp. 93–104.

Chandola, V., Banerjee, A., Kumar, V., 2009. Survey of anomaly detection. ACM Comput. Surveys 41, 15:1–15:58.

Constantinou, V., 2018. PyNomaly: Anomaly detection using local outlier probabilities (LoOP). J. Open Source Softw. 3, 845. URL: https://doi.org/10.21105/joss.00845, doi:10.21105/joss.00845.

Daughtry, C., Walthall, C., Kim, M., de Colstoun, E., McMurtrey, J., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sens. Environ. 74, 229 – 239. doi:https://doi.org/10.1016/S0034-4257(00)00113-9.

Denize, J., Hubert-Moy, L., Betbeder, J., Corgne, S., Baudry, J., Pottier, E., 2018. Evaluation of using Sentinel-1 and Sentinel-2 time-series to identify winter land use in agricultural landscapes. Remote Sens. 11. URL: https://www.mdpi.com/2072-4292/11/1/37.

Gao, B., 1996. NDWI — A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257 – 266. doi:https://doi.org/10.1016/S0034-4257(96)00067-3.

Gómez, C., White, J.C., Wulder, M.A., 2016. Optical remotely sensed time series data for land cover classification: A review. ISPRS J. Photogramm. Remote Sens. 116, 55 – 72. URL: http://www.sciencedirect.com/science/article/pii/

S0924271616000769, doi:https://doi.org/10.1016/j.isprsjprs.2016.03.008.

Hagolle, O., Huc, M., Villa Pascual, D., Dedieu, G., 2015. 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. 7, 2668–2691. URL: https://www.mdpi.com/2072-4292/7/3/2668, doi:10.3390/rs70302668.

Hedayati, P., Bargiel, D., 2018. Fusion of Sentinel-1 and Sentinel-2 images for classification of agricultural areas using a novel classification approach, in: Proc. IEEE IGARSS, Valencia, Spain. pp. 6643–6646. doi:10.1109/IGARSS.2018.8518327.

Hird, J.N., McDermid, G.J., 2009. Noise reduction of NDVI time series: An empirical comparison of selected techniques. Remote Sens. Environ. 113, 248–258.

Holecz, F., Meier, E., Piesbergen, J., Wegmuller, U., Nuesch, D., 1994. Radiometric calibration of airborne SAR imagery, in: 1994 IEEE IGARSS, Pasadena, CA, USA. pp. 1096–1098 vol.2. doi:10.1109/IGARSS.1994.399354.

Inglada, J., Christophe, E., 2009. The Orfeo Toolbox remote sensing image processing software, in: 2009 IEEE IGARSS, Cape Town, South Africa. pp. IV–733–IV–736. doi:10.1109/IGARSS.2009.5417481.

J. Verbesselt, A. Zeileis, M.H., 2012. Near real-time disturbance detection using satellite image time series. Remote Sens. Environ. 123, 98 – 108.

Jaakkola, T., Diekhans, M., Haussler, D., 1999. Using the Fisher kernel method to detect remote protein homologies, in: ISMB, AAAI Press, Menio Park, CA, USA. pp. 149–158.

Jolliffe, I.T., 1986. Principal Component Analysis. Springer New York, New York, NY. URL: https://doi.org/10.1007/978-1-4757-1904-8_8, doi:10.1007/978-1-4757-1904-8_8.

Khabbazan, S., Vermunt, P., Steele-Dunne, S., Ratering Arntz, L., Marinetti, C., van der Valk, D., Iannini, L., Molijn, R., Westerdijk, K., van der Sande, C., 2019. Crop monitoring using Sentinel-1 data: A case study from the Netherlands. Remote Sens. 11. URL: https://www.mdpi.com/2072-4292/11/16/1887, doi:10.3390/rs11161887.

Klisch, A., Atzberger, C., 2016. Operational drought monitoring in Kenya using MODIS NDVI time series. Remote Sens. 8, 267.

Kriegel, H.P., Kröger, P., Schubert, E., Zimek, A., 2009. LoOP: Local outlier probabilities, in: Proc. CIKM, Honk Kong, China. pp. 1649–1652. doi:10.1145/1645953.1646195.

Kumar, D., Rao, S., Sharma, J., 2013. Radar vegetation index as an alternative to NDVI for monitoring of soyabean and cotton, in: Proc. INCA, Jodhpur, India. pp. 91–96.

Kuo, B., Ho, H., Li, C., Hung, C., Taur, J., 2014. A kernel-based feature selection method for SVM with RBF kernel for hyperspectral image classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7, 317–326. doi:10.1109/JSTARS.2013.2262926.

Lampert, C.H., 2009. Kernel methods in computer vision. Found. Trends. Comput. Graph. Vis. 4, 193–285.

Liu, C., Chen, Z., Shao, Y., Chen, J., Hasi, T., Pan, H., 2019. Research advances of SAR remote sensing for agriculture applications: A review. J. Integr. Agric. 18, 506 – 525. URL: http://www.sciencedirect.com/science/article/pii/S2095311918620167, doi:https://doi.org/10.1016/S2095-3119(18)62016-7.

Liu, F.T., Ting, K.M., Zhou, Z., 2008. Isolation forest, in: IEEE ICDM, Pisa, Italy. pp. 413–422. doi:10.1109/ICDM.2008.17.

Liu, F.T., Ting, K.M., Zhou, Z.H., 2012. Isolation-based anomaly detection. ACM Trans. Knowl. Discov. Data 6. URL:

https://doi.org/10.1145/2133360.2133363, doi:10.1145/2133360.2133363.

McFeeters, S.K., 1996. The use of the normalized difference water index (NDWI) in the delineation of open water features. Int. J. Remote Sens. 17, 1425–1432. doi:10.1080/01431169608948714.

McNairn, H., Shang, J., 2016. A review of multitemporal synthetic aperture radar (SAR) for crop monitoring, in: Ban, Y. (Ed.), Multitemporal Remote Sensing: Methods and Applications. Springer International Publishing, Cham, Switzerland. chapter 15, pp. 317–340. URL: https://doi.org/10.1007/978-3-319-47037-5_15, doi:10.1007/978-3-319-47037-5_15.

Meroni, M., Fasbender, D., Rembold, F., Atzberger, C., Klisch, A., 2019. Near real-time vegetation anomaly detection with MODIS NDVI: Timeliness vs. accuracy and effect of anomaly computation options. Remote Sens. Environ. 221, 508–521.

Moran, M.S., Hymer, D.C., Qi, J., Sano, E.E., 2000. Soil moisture evaluation using multi-temporal synthetic aperture radar (SAR) in semiarid rangeland. Agric. For. Meteorol. 105, 69 – 80. URL: http://www.sciencedirect.com/science/article/pii/S0168192300001891, doi:https://doi.org/10.1016/S0168-1923(00)00189-1.

Motohka, T., Nasahara, K.N., Oguma, H., Tsuchida, S., 2010. Applicability of green-red vegetation index for remote sensing of vegetation phenology. Remote Sens. 2, 2369–2387. URL: https://www.mdpi.com/2072-4292/2/10/2369, doi:10.3390/rs2102369.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Pimentel, M., Clifton, D., Clifton, L., Tarassenko, L., 2014. A review of novelty detection. Signal Process. 99, 215–249.

Poilvé, H., 2010. Towards an operational GMES land monitoring core service-BIOPAR methods compendium — meris fr biophysical. Geoland2: Immenstaad, Germany , 45.

Prendes, J., Chabert, M., Pascal, F., Giros, A., Tourneret, J.Y., 2015a. Change detection for optical and radar images using a Bayesian nonparametric model coupled with a Markov random field, in: Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc., Brisbane, Australia. pp. 1513–1517.

Prendes, J., Chabert, M., Pascal, F., Giros, A., Tourneret, J.Y., 2015b. A new multivariate statistical model for change detection in images acquired by homogeneous and heterogeneous sensors. IEEE Trans. Image Process. 24, 799–812.

Prendes, J., Chabert, M., Pascal, F., Giros, A., Tourneret, J.Y., 2015c. Performance assessment of a recent change detection method for homogeneous and heterogeneous images. Revue Française de Photogrammétrie et de Télédétection 209, 23–29.

Quegan, S., Jiong Jiong Yu, 2001. Filtering of multichannel SAR images. IEEE Trans. Geosci. Remote Sens. 39, 2373–2379. doi:10.1109/36.964973.

Rouse, J., Haas, R., Schell, J., Deering, D., 1974. Monitoring vegetation systems in the great plains with ERTS. NASA special publication 351, 309.

Schölkopf, B., Tsuda, K., Vert, J.P., 2004. Kernel methods in computational biology. MIT Press, Cambridge, Mass.

Schölkopf, B., Williamson, R., Smola, A., Shawe-Taylor, J., Platt, J., 1999. Support vector method for novelty detection, in: Proc. NIPS, Denver, CO, USA. pp. 582–588.

Sedano, F., Kempeneers, P., Hurtt, G., 2015. A Kalman filter-based method to generate continuous time series of medium-

resolution NDVI images. Remote Sens. 6, 12381–12408.

Tax, D.M.J., Duin, R.P.W., 2004. Support vector data description. Mach. Learn. 54, 45–66. URL: https://doi.org/10.1023/B:MACH.0000008084.60811.49, doi:10.1023/B:MACH.0000008084.60811.49.

Veloso, A., Mermoz, S., Bouvet, A., Toan, T.L., Planells, M., Dejoux, J.F., Ceschia, E., 2017. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. Remote Sensing of Environment 199, 415 – 426. URL: http://www.sciencedirect.com/science/article/pii/S0034425717303309, doi:https://doi.org/10.1016/j.rse.2017.07.015.

Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J.P., Veroustraete, F., Clevers, J.G., Moreno, J., 2015. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – a review. ISPRS J. Photogramm. Remote Sens. 108, 273 – 290. URL: http://www.sciencedirect.com/science/article/pii/S0924271615001422, doi:https://doi.org/10.1016/j.isprsjprs.2015.05.005.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Jarrod Millman, K., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C., Polat, İ., Feng, Y., Moore, E.W., Vand erPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., Contributors, S..., 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods doi:https://doi.org/10.1038/s41592-019-0686-2.

Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., Strauss, P., 2018. Sensitivity of Sentinel-1 backscatter to vegetation dynamics: An Austrian case study. Remote Sens. 10, 1396. doi:10.3390/rs10091396.

Wegmuller, U., Werner, C., 1997. Retrieval of vegetation parameters with SAR interferometry. IEEE Trans. Geosci. Remote Sens. 35, 18–24. doi:10.1109/36.551930.

Weiss, M., Jacob, F., Duveiller, G., 2020. Remote sensing for agricultural applications: A meta-review. Remote Sens. Environ. 236, 111402. URL: http://www.sciencedirect.com/science/article/pii/S0034425719304213, doi:https://doi.org/10.1016/j.rse.2019.111402.

Whelen, T., Siqueira, P., 2018. Time-series classification of Sentinel-1 agricultural data over North Dakota. Remote Sens. Lett. 9, 411–420. doi:10.1080/2150704X.2018.1430393.

Wu, C., Niu, Z., Tang, Q., Huang, W., 2008. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. Agric. For. Meteorol. 148, 1230 – 1241. doi:https://doi.org/10.1016/j.agrformet.2008.03.005.

Zhou, Z.G., Tang, P., 2016. Continuous anomaly detection in satellite image time series based on z-scores of season-trend model residuals, in: Proc. IEEE IGARSS, Beijing, China. pp. 3410–3413.

Zhou, Z.G., Tang, P., Zhou, M., 2016. Detecting anomaly regions in satellite image time series based on sesaonal autocorrelation analysis, in: Proc. ISPRS, Prague, Czech Republic. pp. 303–310. URL: https://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/III-3/303/2016/, doi:10.5194/isprs-annals-III-3-303-2016.