

Supplementary material: Parcel-level crop anomaly detection using optical and SAR images: application to wheat and rapeseed crops

Florian Mouret^{1,2}, Mohanad Albughdadi¹, Sylvie Duthoit¹, Denis Kouamé³,
Hervé Poilvé⁴, Guillaume Rieu¹, and Jean-Yves Tourneret²

¹TerraNIS, 12 Avenue de l'Europe, 31520 Ramonville-Saint-Agne, France

²University of Toulouse / IRIT-INP-ENSEEIH / TéSA, 2 Rue Charles Camichel, 31000 Toulouse, France

³University of Toulouse / IRIT-UPS, 118 Route de Narbonne, 31062 Toulouse Cedex 9, France

⁴AIRBUS Defence & Space, 5 rue des satellites, 31400 Toulouse, France

This supplementary material shows some anomalies identified by photo-interpreters, provided to illustrate the labeling process of this study. As already mentioned in this paper, a labeling with photo-interpretation has the advantage of being fast and usable at large-scale. Regarding the categories used for labeling, the right balance had to be found between details and interpretability. Too detailed categories would lead to difficulties in the analysis of the results because only few elements would have been detected in each of these categories. On the other hand, too coarse categories lead to possible misinterpretations or inaccurate analysis. A separation between early heterogeneity and global heterogeneity is a good example of an interesting separation for a more precise analysis. Within a specific category (*e.g.*, heterogeneous parcels), giving a score of abnormality would have been too subjective and subject to misinterpretation. It explains why only binary labels were given to each parcel. In what follows, NDVI time series are represented because their interpretation is generally easier and straightforward. However, the labeling process was based in all the indicators and all the RGB images were used. Also note that in general, the 4 other indicators were correlated to NDVI. In the rare case where only one indicators was abnormal, normality was preferred to abnormality. Of course, a *in situ* verification could be interesting confirm the presence of the detected anomalies (as for SAR anomalies).

1 Heterogeneity

Fig. 1 illustrates a case of heterogeneity occurring during a complete season. In the images, various parcels present signs of heterogeneity. It is sometimes difficult to characterize more precisely this kind of anomaly apart from early/late heterogeneity. **Fig. 2** depicts an example of heterogeneity where two different parts are clearly visible. This example shows that it is difficult to decide if wrong boundaries were provided, if two different varieties of rapeseed were sowed or if soil differences led to heterogeneity. **Fig. 3** shows an example of heterogeneity after senescence, where the parcel is perfectly homogeneous in May whereas it presents some heterogeneity in June. As explained in the paper, this phenomenon is likely to be related to differences in the soil water content and could be of interest for inter season analysis. **Fig. 4** shows the IQR NDVI time series of the same parcel (in yellow) and shows that late heterogeneity can be visible and that its IQR value is greater than 90% of the rest of the data. The same observations can be made for early heterogeneity. In order to compare the analyzed parcel to the rest of the data, the median, the 10th percentile and the 90th percentile of the whole dataset were displayed (similarly to a boxplot visualization). This representation allows the agronomic expert to know if the observed parcel has indicator values higher (or lower) than 90% of the data.

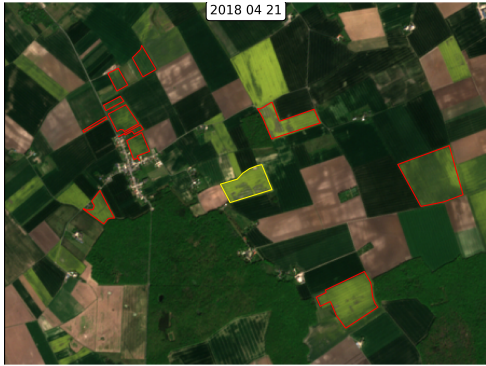


Fig. 1: Example of a parcel of rapeseed crop (yellow boundaries) where heterogeneity occurs almost during the complete season. Other parcels presents signs of heterogeneity too. Left: optical image acquired in May 2018, right: optical image acquired in June 2018.



Fig. 2: Example of a parcel of rapeseed crop (yellow boundaries) with a two-part heterogeneity. Left image was acquired in February 2018 and right image in April 2018.



Fig. 3: Example of a parcel of rapeseed crop (yellow boundaries) where heterogeneity after senescence is observed. Left image was acquired in May and right image in June.

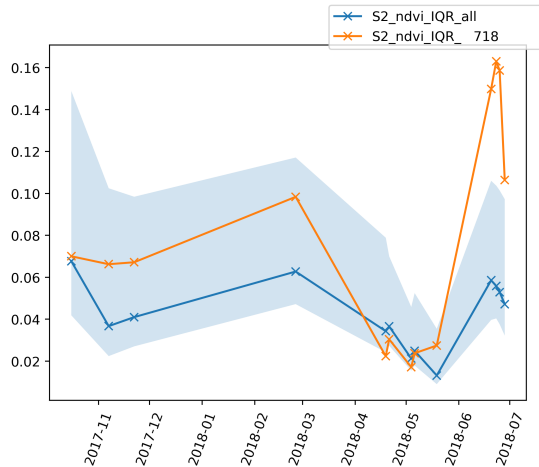


Fig. 4: Time series of IQR NDVI showing sign of heterogeneity after senescence. The blue line is the median value of the whole dataset. The blue area is filled between the 10th and 90th percentile. The orange line is a specific parcel, which is displayed in Fig. 3.

2 Growth anomalies

Growth anomalies are a bit more difficult to analyze when looking at RGB images only (or composite multispectral images) because growth anomalies occur at multiple dates. Visualizing some indicator time series is a good way to analyze growth anomalies. Fig. 5.a depicts an example of late growth problem for a rapeseed parcel. Even if important, this delay in growth has not affected the end of the season whereas for wheat it was observed to be often problematic as illustrated in Fig. 5.b.

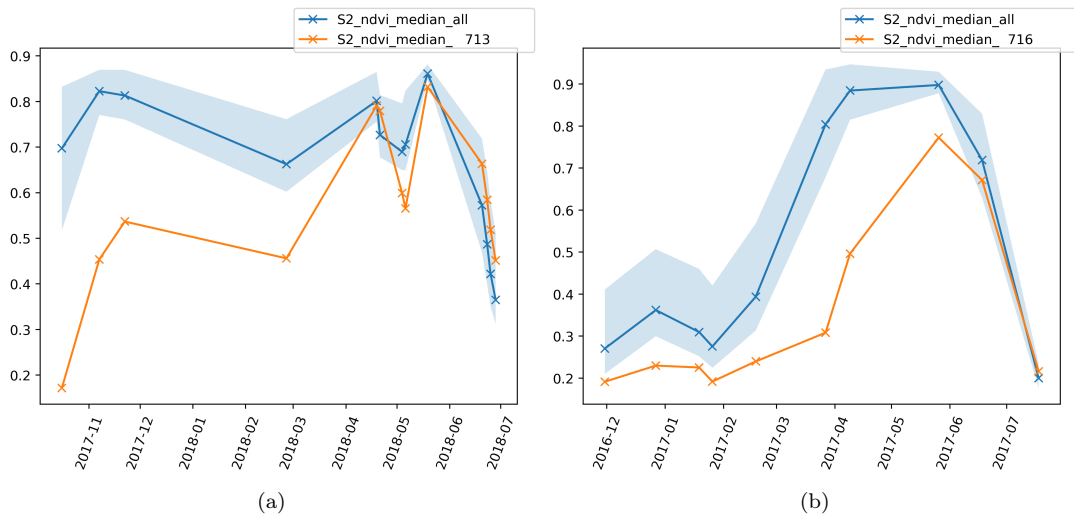


Fig. 5: Time series of median NDVI showing sign late growth for (a) a rapeseed parcel and (b) a wheat parcel. The blue line is the median value of the whole dataset. The blue area is filled between the 10th and 90th percentile. The orange line is a specific parcel analyzed.

Fig. 6 is an example of vigorous parcel. As explained in the paper, the sowing data can be a factor explaining this vigor. But in this particular example, seeding was done the 3 of October which is an early but not extreme date whereas NDVI median values was extremely high.

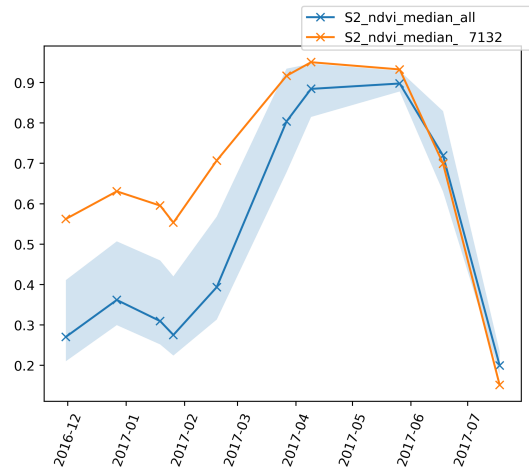


Fig. 6: Time series of median NDVI for a wheat parcel. The blue line is the median value of the whole dataset. The blue area is filled between the 10th and 90th percentile. The orange line is a specific parcel analyzed.

Fig. 7.a displays an example of early senescence. As mentioned in the paper this type of anomaly is more subtle than a global heterogeneity or a late growth problem because it occurs only at the end of the time series. In this particular example, a low median NDVI occurs in February. It could be a sign of problem during winter but it is challenging to confirm abnormality with only one NDVI value. Fig. 7.b illustrates the case of early flowering. This case is also subtle and was observed few times. In general, the early flowering category is also affected by early senescence.

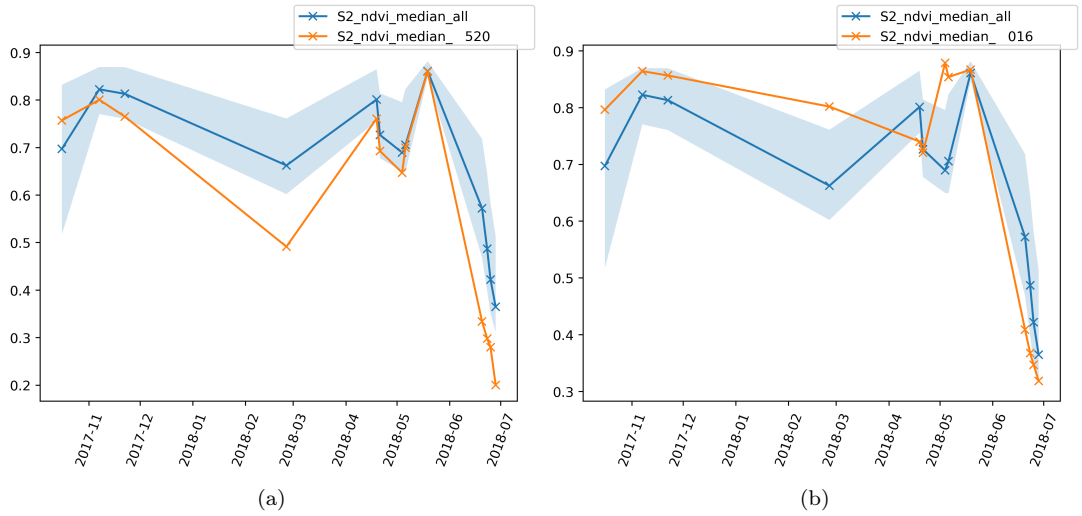


Fig. 7: Time series of median NDVI for a rapeseed parcel presenting sign of (a) early senescence and (b) early flowering. The blue line is the median value of the whole dataset. The blue area is filled between the 10th and 90th percentile. The orange line is a specific parcel analyzed.

3 Boundaries issue

Fig. 8 displays two type of anomalies. The yellow parcel has clearly wrong boundaries whereas the small parcel at its right is clearly a wrong type of crop. These two problems are important and easy to detect. For a wrong shape issue, it is often difficult determine whether the problem is related to an error in boundaries or a growth problem for part of the parcel.



Fig. 8: Example of a parcel of rapeseed crops where a wrong boundary issue and a wrong type issue are visible.

4 Non-agronomic anomalies

Fig. 9 illustrates a problem observed in March 27 2017. A majority of pixels acquired at this date have a red channel value equal to zero, explaining extreme values of the indicators. This issue was caused by the MAJA processing¹ and could be easily fixed using another processing chain like Sen2core or with a threshold for the red band. As our goal is to show all the issues encountered during the labeling process, it was decided to keep these anomalies inside the dataset to determine if the proposed method is resilient to this problem (which is generally the case). Note that NDVI saturates whereas MCARI/OSAVI is extremely affected. Extreme values are less problematic for the IF algorithm this algorithm is not directly based on a distance measure, contrary to the other algorithms tested in the study. This red channel problem was mainly observed for vigorous parcels (as for this example). Fig. 10 shows the optical image of this parcel, which is darker than its neighbors.

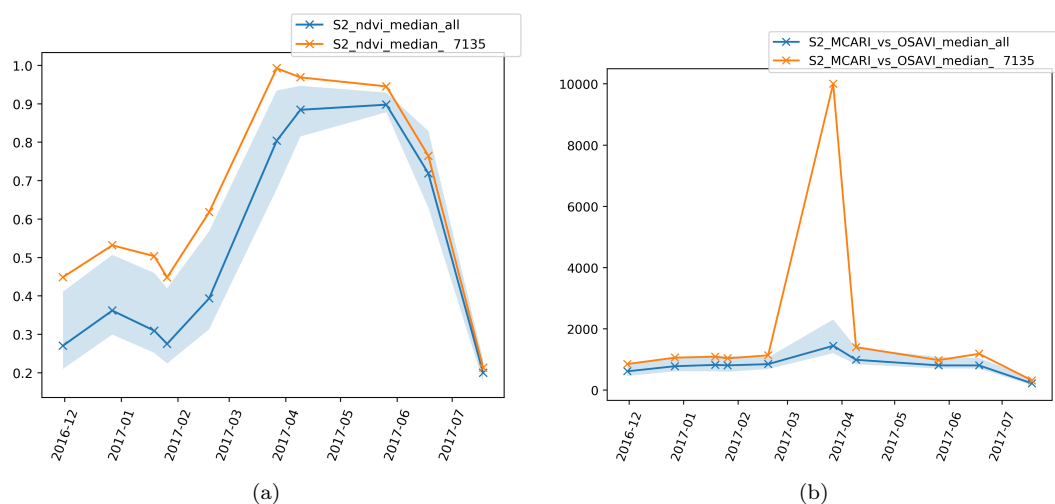


Fig. 9: Time series of median NDVI (a) and median MCARI/OSAVI (b) for a wheat parcel presenting a red channel problem in March. The blue line is the median value of the whole dataset. The blue area is filled between the 10th and 90th percentile. The orange line is a specific parcel analyzed.

¹<https://labo.obs-mip.fr/multitemp/using-ndvi-with-atmospherically-corrected-data/>, online accessed 10 March 2020



Fig. 10: Example of a parcel of rapeseed crops with a red channel problem (yellow boundaries). A late growth parcel can be observed at the bottom of the image (triangle red boundaries).

Fig. 11 displays examples of two different types of anomalies. The yellow parcel is an example of parcel affected by shadows. It can be observed that NDVI is not strongly affected by this issue whereas the other indicators such as GRVI or MCARI/OSAVI have extreme values. A case of a too small parcel is also visible in the image, the particular form of the boundaries causes difficulties to filter the parcel only using area.



Fig. 11: Example of parcels of rapeseed. The parcel with yellow boundaries is affected by shadow caused by the trees next to the parcel. Also, at the bottom a too small parcel is visible.

Fig. 12 shows an example of parcel, which was difficult to analyze by the photo-interpreter. Both MCARI/OSAVI and NDIV have low values, but in a limited time interval. This example illustrates that the label true positive / false positive was very strict and was attributed only in cases without any doubt. The other uncertain cases were labeled as “not sure”.

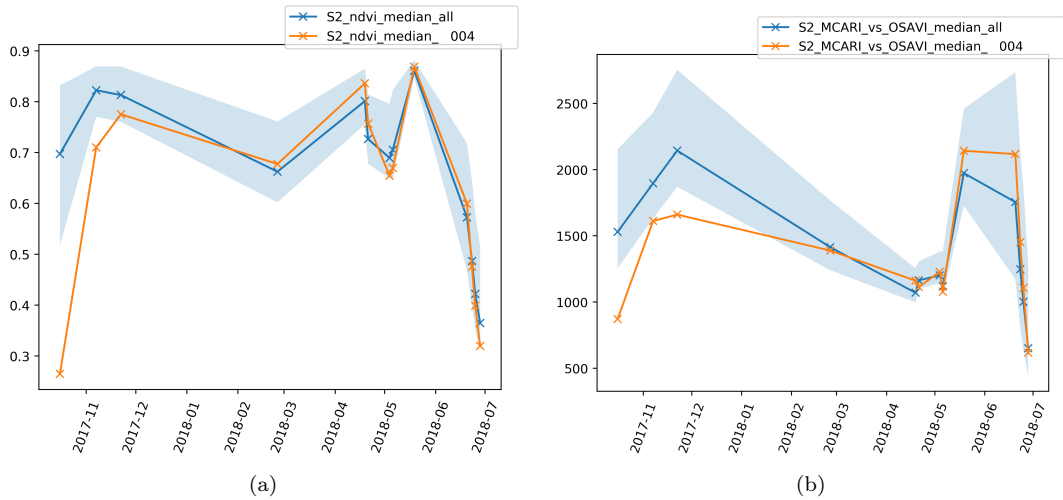


Fig. 12: Time series of median NDVI (a) and median MCARI/OSAVI (b) for a rapeseed parcel. The blue line is the median value of the whole dataset. The blue area is filled between the 10th and 90th percentile. The orange line is a specific parcel analyzed.

5 Complementary information about SAR images and their anomalies

A strong correlation between SAR and plant vigor (late growth / vigorous crop) was observed in this study. Fig. 13 illustrates the effect of late growth on S1 indicators. In this example, the time evolution of the VH, VV and VH/VV time series seem to be correlated with NDVI time series.

Fig. 14 illustrates a main issue observed when using the SAR images: heterogeneity is not always visible with SAR images which can be understood because SAR images are mainly affected by the structure of the vegetation.

Finally, Fig. 15 illustrates a case of SAR anomaly. At the end of the NDVI time series, a case of early flowering/late senescence was declared by the photo-interpreter. However, in the early season, the median NDVI time series is perfectly normal whereas the median VV is extremely high. A possible explanation is an interaction between the soil structure and the SAR signal (or an abnormal soil moisture).

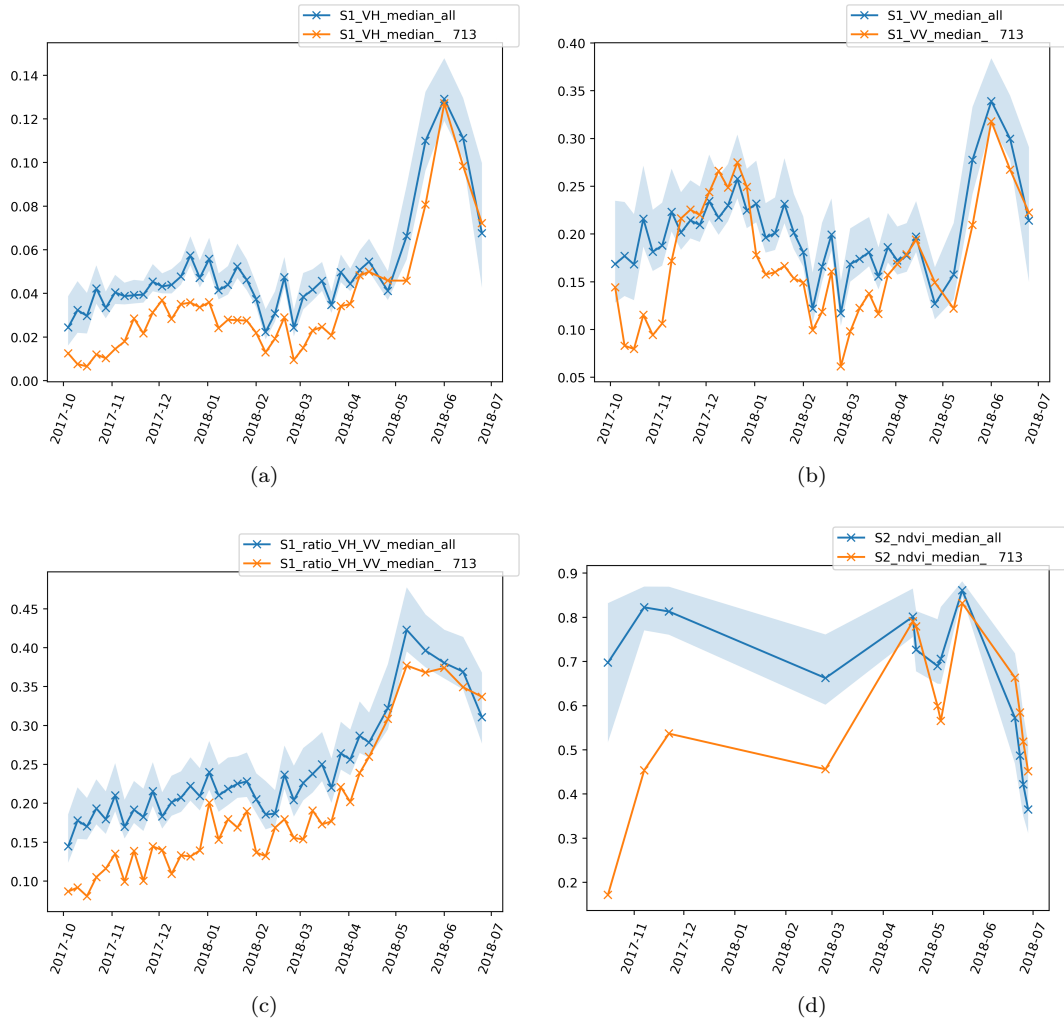


Fig. 13: Time series of median SAR indicators (VV, VH, VH/VV) and median NDVI for a rapeseed parcel. The blue line is the median value of the whole dataset. The blue area is filled between the 10th and 90th percentile. The orange line is a specific parcel analyzed. (a): median VH, (b): median VV, (c): median ratio VH/VV, (d): median NDVI



Fig. 14: Example of a parcel of rapeseed crop (yellow boundaries) where heterogeneity occurs almost during the complete season. Some other parcels show some signs of heterogeneity too. Left: optical image acquired in May, right: composite SAR image (green channel is VV polarization and red channel is VH polarization) acquired in May with multi-temporal speckle filtering.

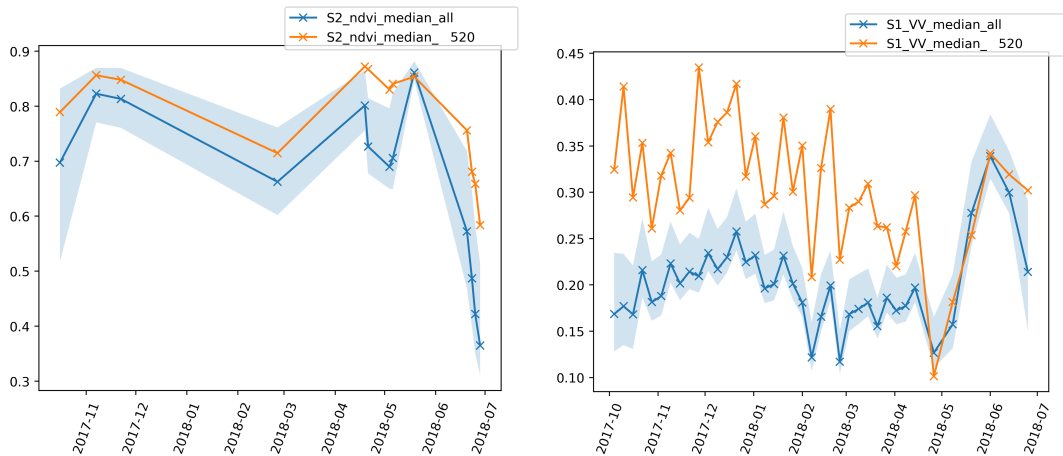


Fig. 15: Time series of median VV polarization and median NDVI for a rapeseed parcel. The blue line is the median value of the whole dataset. The blue area is filled between the 10th and 90th percentile. The orange line is a specific parcel analyzed. Left: median NDVI, right: median VV.