INTRODUCING AUTOMATIC SATELLITE IMAGE PROCESSING INTO LAND COVER MAPPING BY PHOTO-INTERPRETATION OF AIRBORNE DATA

H. Costa [1,2], P. Benevides [1], F. Marcelino [1], M. Caetano [1,2] *

[1] Direção-Geral do Território, 1099-052 Lisbon - (hcosta, pbenevides, fmarcelino, mario.caetano)@dgterritorio.pt
[2] NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Campus de Campolide, 1070-312 Lisbon, Portugal

KEY WORDS: Time series, Sentinel, Landsat, NDVI, Change detection, classification, crops, COS

ABSTRACT:

A series of five land cover maps, widely known as COS (*Carta de Uso e Ocupação do Solo*), have been produced since 1990 for mainland Portugal. Previous to 2015, all maps were produced through photo-interpretation of orthophotos. Land cover and land use changes were detected through comparison of previous and recent orthophotos, which were used for map updating, thereby producing a new map. The remaining areas of no change were preserved across the maps for consistency. Despite the value of the maps produced, the method is very time-consuming and limited to the single-date reference of the orthophotos. From 2015 onwards, a new approach was adopted for map production. Photo-interpretation of orthophoto maps is still the basis of mapping, but assisted by products derived from satellite data. The goals are three-fold: (i) cut time production, (ii) increase map accuracy, and (iii) further detail the nomenclature. The last map published (COS 2015) benefited from change detection and classification analyses of Landsat data, namely for guiding the photo-interpretation in forest, shrublands, and mapping annual agriculture. Time production and map error have been reduced comparing to previous maps. The new 2018 map, currently in production, further explores this approach. Landsat 8 time series of 2015-2018 are used for change detection in vegetation based on NDVI differencing, thresholding and clustering. Sentinel-2 time series of 2017-2018 are used to classify Autumn/Winter crops and Spring/Summer crops based on NDVI temporal profiles and classification rules. Benefits and pitfalls of the new mapping approach are presented and discussed.

1. INTRODUCTION

The importance of land cover and land use (LCLU) has long been recognized, such as in environmental sciences and policy, and LCLU mapping is nowadays a well-established activity all over the world at a diversity of scales. In Portugal, the national mapping agency, *Direção-Geral do Território* (DGT), produces and publishes a LCLU map for the continental territory, widely known as COS (*Carta de Uso e Ocupação do Solo*) (Figure 1). This map was first produced for the year of 1990 and updated for 1995, 2007, 2010, and 2015. Currently, COS for 2018 is under production and its publication is expected by the end of 2019.

COS is used nationally and internationally by a large spectrum of users for numerous purposes, including landscape planning, decision-making, reporting obligations, research, education, and business. The uptake of COS is related to the fact that it is made available freely through an open data policy, and is the most detailed product of mainland Portugal in terms of thematic content and spatial representation.

The minimum mapping unit (MMU) of COS is 1 hectare (Figure 1), and the LCLU nomenclature is based on that of CORINE Land Cover (CLC), which constitutes a reference LCLU product at European level (Büttner, 2014). The nomenclature of COS 2015 includes 48 classes aggregated to 9 main classes: Artificial land, Agriculture, Pastures, Agro-forestry, Forest, Shrublands, Bare soil and sparse vegetation, Wetlands, and Water.

Figure 1. Subset of COS 2015 over the city of Coruche (38°57'30,05"N; 8°31'36,013"W). Background: false colour orthophoto of 2015.

COS has been produced through photo-interpretation of orthophotos (spatial resolution of 50 cm and better). LCLU changes are normally detected through comparison of previous and recent orthophotos. The changes detected are mapped and superimposed on the outdated map, thereby producing a new map. The remaining areas of no change are preserved across the maps for consistency.

The traditional methods used in map production holds several problems. Mostly, map production is very time-consuming, expensive, and orthophotos provide information only related to the specific date of the flight campaigns (while multi-temporal data are valuable for accurate photo interpretation of some classes, such as annual crops).

---

* Corresponding author

From 2015 onwards, map production is evolving to include (semi)-automatic methods of image analysis applied to Landsat and Sentinel-2 data. Therefore, while photo-interpretation of orthophotos is still the basis of mapping, additional products derived from multi-temporal satellite data are used with three main objectives: (i) cut time production, (ii) further detail the nomenclature, and (iii) increase map accuracy.

COS 2015 already benefited from change detection analysis of Landsat multi-temporal data, namely inter-annual differencing of Normalized Difference Vegetation Index (NDVI). This was used for guiding the photo-interpretation on detection of changes between 2010 and 2015, and classifying Autumn/Winter crops and Spring/Summer crops. Here, however, the production of the new 2018 map is presented, namely the (semi)-automatic methods used to analyze satellite data.

First, a change detection method based on inter-annual time series of Landsat 8 data is described. This method applies thresholding and k-means clustering on image differencing of NDVI data of 2015-2018 for detecting changes in forest and shrublands. Then, a classification method based on intra-annual Sentinel-2 data is presented for annual croplands. This method uses expert knowledge and statistics extracted from the intra-annual time series to distinguish between Autumn/Winter crops and Spring/Summer crops.

2. COS PRODUCTION

COS 2018 is under production and its publications is expected by the end of 2019. It will be a map of polygons released in vector format with 1 hectare MMU. The technical specifications of COS 2018 (Table 1) are similar to the precedent map of 2015, except the number and name of some classes. The maps are, however, compatible as all classes have direct correspondence among maps. The differences between nomenclatures are not discussed here, but they result from collaborative work between DGT and other Portuguese public institutions, so COS can serve them better.

Reference year	1990, 1995, 2007, 2010, 2015 (and 2018)
Format	Vector (polygons)
MMU (ha)	1 ha
Minimum distance between lines (m)	20 m
Base data	Orthorectified digital aerial images
Nomenclature	Hierarchical classification system
Production method	Visual interpretation
Geometric accuracy	$\geq 5,5m$
Thematic Accuracy target	$\geq 85\%$
Coordinate reference system	ETRS89/PT-TM06

Table 1. Technical specifications of COS

The production of COS 2018 is founded on visual interpretation of orthophotos and assisted and complemented with novel methods of image analysis. First, photo-interpreters are assisted with auxiliary information produced with Landsat 8 data to find changes in forest and shrublands. The auxiliary information indicates where vegetation loss occurred between 2015 and 2018, which is indicative of changes, such as new constructions or agricultural fields. Photo-interpreters use these auxiliary data as alerts deserving careful inspection. The benefit of these alerts

is that photo-interpreters focus their attention on sites of probable change, hence reducing time wasted in inspecting areas where vegetation change was not observed from space. Because forest and shrubland cover 51% (2015) of mainland Portugal, the time needed to detect and map changes across a large area is reduced considerably.

Second, photo-interpretation is complemented with analysis of Sentinel-2 data to distinguish two main types of annual crops: Autumn/Winter crops and Spring/Summer crops. These two classes have a strong seasonal variation in terms of vegetation cover and status, which a single-date orthophoto misses to capture. On the contrary, the temporal resolution of Sentinel-2 is large and tracks the vegetation cycles typical of annual agriculture. Therefore, the photo-interpreters are responsible for delimiting the borders of annual agricultural fields without making a distinction on the type of annual crops. Later, Sentinel-2 data is classified on either Autumn/Winter crops or Spring/Summer crops across the agricultural fields identified previously. The relative abundance of each of the crop type found in the polygons classified as annual crops is used as an attribute. The satellite data overcomes the temporal limitation of the orthophotos, making it possible to map two classes with otherwise insufficient accuracy.

3. METHODS

3.1 Detection of vegetation loss

Comparison of remotely sensed data acquired on different dates is a simple but effective approach for change detection. Some techniques based on this approach, often called differencing or layer arithmetic (Tewkesbury et al., 2015; Zhu, 2017), calculate descriptive statistics for the difference of vegetation indices between two points in time. Then, the descriptive statistics are analysed to discriminate change from no-change. This approach was used here to analyse differences of NDVI (Tucker, 1979) across 2015-2018.

The analysis of NDVI differencing is summarized in Figure 2. All available images (Surface Reflectance Level-2 Data Products) with cloud cover < 50% and covering mainland Portugal (Figure 3a) during summer of 2015-2018 were downloaded from EarthExplorer. Phenological differences between the years are expected to be small as only images of summer are used. All images were processed according to the specifications described in the Landsat 8 Land Surface Reflectance Product Guide. Specifically, pixels affected by clouds, shadow and with invalid values were reclassified as "no data".
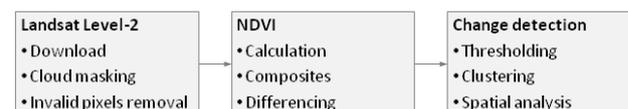
Figure 2. Methods used detecting vegetation loss.

The NDVI was calculated from each processed image. Then, images were combined to produce composites without gaps ("no data"). Therefore, images of the same year (e.g. 2015) were combined to produce one composite per year. Each pixel was given the minimum NDVI value among all images of the same year.
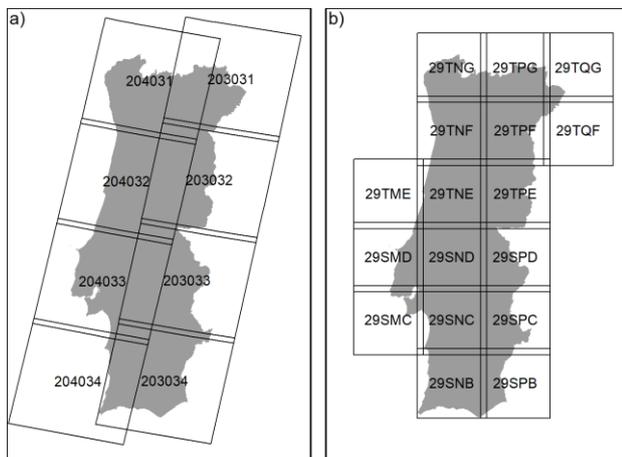
Figure 3. Landsat pathrows (a) and Sentinel-2 tiles (b) over mainland Portugal.

The minimum operator was selected for the composites production because it selects the NDVI associated with potential vegetation loss. For example, if three images are available from June, July and August, and a clear cut of forest occurs in August, the composite will select the NDVI associated with change (the smallest NDVI of August). Other operators such as the mean would select a value unrepresentative of change. There is one exception thought. For the first year of the analysis (2015), the maximum NDVI was selected for the composite because 2015 is the baseline of comparison. If the minimum NDVI was selected, any change occurring in 2015 would go unnoticed while comparing it with 2016, in which the NDVI will certainly remain small. Using the maximum NDVI of 2015 ensures that the removals of vegetation of 2015 are detected as compared to 2016.

Differences of NDVI between the years were then calculated and analysed based on the assumption that spectral changes among years are due to many reasons (e.g., vegetation health), but only substantial changes reflect land cover changes. Therefore, the difference of NDVI between two years should follow a normal distribution as most of the differences observed are small (close to zero) and only few marked differences are caused by land cover change (Jin et al., 2013; Lu et al., 2004; Pu et al., 2008).

The NDVI differences were submitted to two analyses: thresholding and clustering. The former relies on a threshold of difference. That is, if the difference of NDVI exceeds the threshold, a land cover change is expected to have occurred in the pixel. Here, a threshold was defined empirically and corresponded to -1.5 standard deviations of the average NDVI difference (Pu et al., 2008). The second analysis was k-means clustering (Hartigan and Wong, 1979). Specifically, the NDVI difference was classified in 10 clusters, and those characterized by a median difference of < -0.25 were flagged as change. Because the analyses are based on relative comparisons between pixels, they are sensitive to the range of NDVI as a function of the land cover class. Therefore, calculations were done independently by strata, which were eucalyptus, pines, other forest species, and shrubs, all of them as mapped in 2015.

All pixels associated with change, irrespective of type of analysis (thresholding or clustering) and strata, were merged together in a single layer to represent the cumulative changes identified across the whole period on analysis in forest and shrubland. Then changes were submitted to spatial analysis.

Patches of change were enlarged and contracted by one pixel, and only patches >= 5 pixels (4.5 hectares) were retained. This step removed noise and extremely small changes. Finally, the changes were converted to vector format.

The accuracy of the method was assessed in two different ways. First, a set of polygons produced in advanced for COS by the photo-interpreters without the help of satellite data were taken as reference data. These polygons were labelled as either "change" or "no-change". The first case corresponds to changes to be mapped in COS 2018, and the latter corresponds to area of no change between 2015 and 2018. Commission and omission error were estimated simply by counting the relative number of reference polygons that spatially intercepted the alerts.

Second, the accuracy of the method was assessed based on a sample of alerts collected across the Landsat pathrow 204032 (Figure 3a), which were labelled as either "change" or "no-change" by visual interpretation of orthophotos and satellite data. In this case, "change" was interpreted as all obvious spectral changes visible in the satellite data associated with vegetation loss, regardless if they should correspond to a thematic change in COS. Commission error was estimated simply by counting the relative number of alerts associated with no vegetation loss. No omission error was calculated.

3.2 Classification of crops

Autumn/Winter crops and Spring/Summer crops normally follow a different annual cycle, which satellites can track from space. The analysis of temporal profiles of vegetation indices can help detect phenological traits of crops such as vegetation growth and time of maturation. Temporal analysis of vegetation indices are increasingly used for crop monitoring (Belgiu and Csillik, 2018; Defourny et al., 2019).

The analysis performed is summarized in Figure 4. All Sentinel-2 level-2 (L2A) data available with cloud cover < 50%, covering mainland Portugal (Figure 3b), and acquired between the 1st October 2017 to 30th September 2018 were downloaded from the French Theia Land Data Centre (THEIA). The period analysed corresponds to the agricultural year relevant for 2018. The level 2A data distributed by THEIA corrects the level-1 data for atmospheric and slope effects using MAJA (Baetens et al., 2019). Clouds and their shadows were converted to "no data". All bands were saved with 10 m pixel size without resampling methods applied.
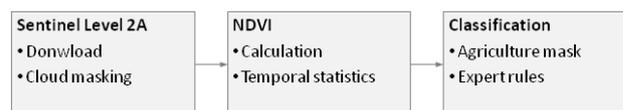
Figure 4. Methods used for crop classification.

The NDVI was calculated from each processed image, and the entire time series was analysed with the estimation of a set of statistical indicators. For example, extracting the date of the maximum NDVI can be important to identify different types of crops. Some statistics take into account the entire agricultural year, while other statistics were extracted independently for each quarter of the agricultural year.

The statistics extracted were then used together with rules predefined with expert knowledge to make the particular distinction between Autumn/Winter crops and Spring/Summer crops. A pixel was classified as Autumn/Winter crop if the

maximum annual amplitude of NDVI was >0.4, the mean annual value of NDVI was > 0.2 and the mean NDVI observed in the third year quarter (April to June) was greater than that of the fourth year quarter (July to September). Spring/Summer crops were classified if the difference between the maximum NDVI of the fourth year quarter and the minimum NDVI of the third quarter was > 0.4, combined with the mean annual NDVI value being > 0.3.

The rules were assessed by inspecting a large sample of NDVI temporal profiles grouped by Autumn/Winter and Spring/Summer crops, such as the ones presented in Figure 5. Here, we can visualize that the maximum NDVI for Autumn/Winter crops is observed at the end of the winter and the respective maximum for the Spring/Summer crops at the end of Summer. Large amplitude values between the minimum and the maximum value throughout the year are also noticeable for both types of crop.
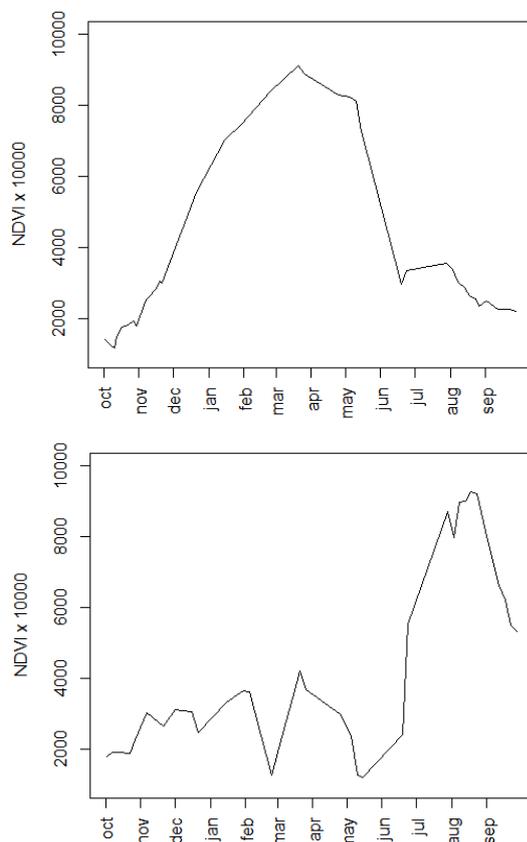
Figure 5. Examples of typical NDVI profiles observed for Autumn/Winter crops (top) and Spring/Summer crops (bottom).

The classification was applied only to pixels that fall inside polygons identified as annual crops in the visual interpretation of orthophotos performed in COS production. Therefore, the rules used are under revision while the geometry on COS is not closed yet for the entire country. Finally, the polygons will be enriched with the relative abundance of each of the crop type as an attribute.

The accuracy of the method was assessed by comparing the 2018 crop map produced to the Land Parcel Identification System (LPIS) of the *Instituto de Financiamento da Agricultura e Pescas* (IFAP) in tile T29SND (Figure 3b). LPIS is a geographical data set used in the framework of the Common Agricultural Policy (CAP) of the European Union for the administration and control of payments to farmers. The agreement between the classification and LPIS was calculated, corresponding to the area of the LPIS's polygons intersected by the crop map of the same class. Commission error was also calculated, corresponding to the area of the LPIS's polygons intersected by the crop map of a different class.

4. RESULTS

4.1 Change detection

The change detection method produced 148793 polygons used as alerts for vegetation loss (Figure 6). When considering the polygons produced in advanced for COS as reference (first accuracy assessment), around 77% of the "change" polygons and 49% of the "no-change" polygons overlap an "alert". This may be interpreted as a crude estimation of 33% omission error and 51% commission error. The second accuracy assessment undertaken only with alerts in pathrow 204032 found 85% of the cases associated with vegetation loss, which corresponds to 15% commission error.

Commission and omission error were observed to occur mostly in small areas of land cover change. Therefore, when using the 148793 alerts produced operationally in the production of COS, they were divided by size and given to the photo-interpreters as different layers. Alerts >2 ha were first inspected, than alerts from 1 to 2 ha, and finally <1 ha.
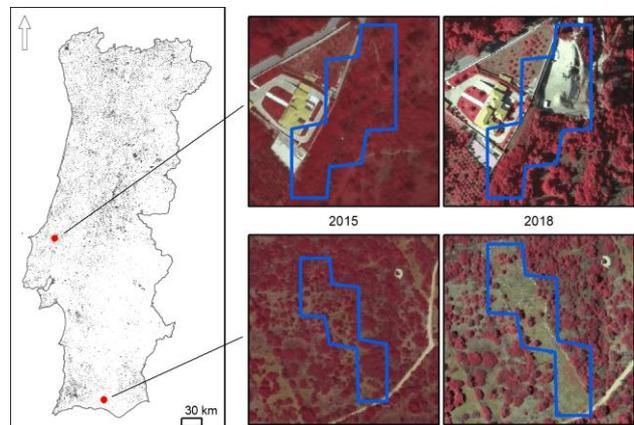
Figure 6. Alerts produced for vegetation loss in forest and shrubland. Two examples of change between 2015 and 2018 are shown for central and south Portugal.

4.2 Classification

The intra-annual statistics extracted from Sentinel-2 dataset combined with the set of predefined were able to produce two classification maps in raster format, one of each for Autumn/Winter and Spring/Summer crops (Figure 7).

The classification of Autumn/Winter crops covers 91% of the area represented by the polygons of the LPIS of the same class (9% omission error), and 39% of the area identified as Spring/Summer crop (commission error). On the other hand, the classification of Spring/Summer crops reaches 94% of the area covered by the same class in LPIS (6% omission error) with 24% commission error.
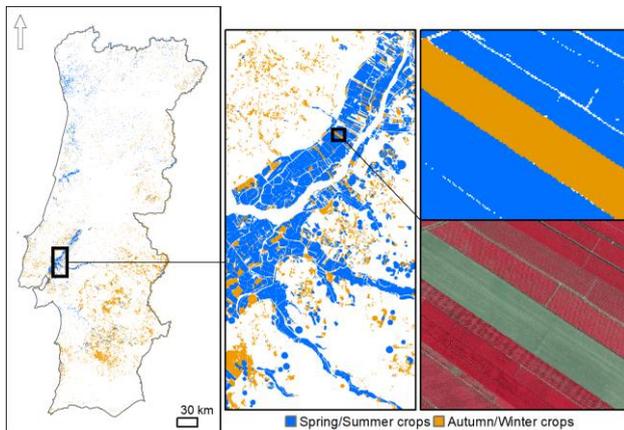
Figure 7. Classification of annual crops into Autumn/Winter crops and Spring/Summer crops.

5. DISCUSSION

The change detection method produced numerous alerts for false changes in vegetation cover, which are expressed by large commission error (51%) in the first accuracy assessment based on COS polygons. However, this was expected because the method was implemented to minimise omission error. Commission error is preferred over omission error as the former has no negative impact on the accuracy of COS. For example, forest clear-cuts occurring as part of the normal cycle of forest management are detected as the NDVI drops sharply, but no change is mapped as land use remains forest, and a new forest patch is expected to follow. The only inconvenient of commission error is that photo-interpreters inspect areas of no thematic change, wasting some time and hence diminishing the effectiveness of the change detection method.

However, it is still important to assess the accuracy of the alerts produced. The second accuracy assessment performed in pathrow 204032 shows a larger accuracy as commission error was only around 15%. The difference between the two assessments is that the later considered spectral changes such as forest clear-cuts as change in vegetation cover (even if that change should not be mapped).

Omission error, on the contrary, impact negatively on mapping as real changes may go unnoticed and hence not mapped. Omission error was larger than desired (33%) and should be reduced. This may be achieved, for example, by improving the implementation of thresholding and clustering analyses, which were conducted sometimes based on empirical decisions on parameterization (e.g., number of clusters in k-means). Exploring change detection methods other than these based on NDVI differencing is also recommended (e.g. Verbesselt et al., 2010).

Despite the large omission error of the alerts, the real omission error of COS 2018 across forest and shrublands will be certainly smaller than 33%. This is because the first accuracy assessment performed was a simple spatial overlap operation between the alerts and the reference data, which maps thematic changes only. Furthermore, the photo-interpreters see beyond the spatial extent of the alerts, and hence are able to spot changes not flagged automatically.

With regard to the classification of crop types, classification problems are expressed mainly by large commission errors. This

is caused by the fact that the same polygon of LPIS could be considered as both Autumn/Winter crop and Spring/Summer crop in the accuracy assessment, as long as intersecting the crop map of these classes. This was regarded as appropriate to simulate the cases in which COS includes different crop types in the same polygon, such as in complex cultivation patterns. Therefore, a single polygon can include Autumn/Winter and Spring/Summer crops, whose abundance is to be included in the polygons attributes.

Note that these classification rules are a preliminary approach to the technique and will be improved. The rules were initially derived analysing the typical temporal profile signature observed for these types of crops (Figure 5), and thus they disregard less common cases such as the occurrence of multiple crops during the crop year. Therefore, alternative rules are likely to be implemented to address specific cases. In addition, semi-automatic approaches based on supervised classification should be tested and adopted.

6. WAY AHEAD

The methods discussed will be further developed beyond 2018 to satisfy the production of the future mapping model of COS. The latter will be a map of polygons similar to previous maps, but enriched with attributes that characterize the polygons. While the geometry of the map should continue to be defined by visual interpretation of orthophotos, the attributes will be produced from raster maps through automatic change detection and classification of satellite time series data.

ACKNOWLEDGEMENTS

REFERENCES

Baetens, L., Desjardins, C., Hagolle, O., 2019. Validation of copernicus Sentinel-2 cloud masks obtained from MAJA, Sen2Cor, and FMask processors using reference cloud masks generated with a supervised active learning procedure. *Remote Sens.* 11, 1–25. doi:10.3390/rs11040433

Belgiu, M., Csillik, O., 2018. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sens. Environ.* 204, 509–523. doi:10.1016/j.rse.2017.10.005

Büttner, G., 2014. CORINE Land Cover and Land Cover Change Products, in: Manakos, I., Braun, M. (Eds.), *Land Use and Land Cover Mapping in Europe: Practices & Trends.* Springer Netherlands, Dordrecht, pp. 55–74. doi:10.1007/978-94-007-7969-3_5

Defourny, P., Bontemps, S., Bellemans, N., Cara, C., Dedieu, G., Guzzonato, E., Hagolle, O., Inglada, J., Nicola, L., Rabaute, T., Savinaud, M., Udroiu, C., Valero, S., Bégué, A., Dejoux, J.F., El Harti, A., Ezzahar, J., Kussul, N., Labbassi, K., Lebourgeois, V., Miao, Z., Newby, T., Nyamugama, A., Salh, N., Shelestov, A., Simonneaux, V., Traore, P.S., Traore, S.S., Koetz, B., 2019. Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems

around the world. *Remote Sens. Environ.* 221, 551–568. doi:10.1016/j.rse.2018.11.007

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. *Appl. Stat.* 28, 100–108.

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., Xian, G., 2013. A comprehensive change detection method for updating the National Land Cover Database to circa 2011. *Remote Sens. Environ.* 132, 159–175. doi:http://dx.doi.org/10.1016/j.rse.2013.01.012

Lu, D., Mausel, P., Brondízio, E., Moran, E.F., 2004. Change detection techniques. *Int. J. Remote Sens.* 25, 2365–2401. doi:10.1080/0143116031000139863

Pu, R., Gong, P., Tian, Y., Miao, X., Carruthers, R.I., Anderson, G.L., 2008. Using classification and NDVI differencing methods for monitoring sparse vegetation coverage: a case study of saltcedar in Nevada, USA. *Int. J. Remote Sens.* 29, 3987–4011. doi:10.1080/01431160801908095

Tewkesbury, A.P., Comber, A.J., Tate, N.J., Lamb, A., Fisher, P.F., 2015. A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sens. Environ.* 160, 1–14. doi:10.1016/j.rse.2015.01.006

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation, *Remote Sens. Environ.* doi:10.1016/0034-4257(79)90013-0

Verbesselt, J., Hyndman, R., Zeileis, A., Culvenor, D., 2010. Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. *Remote Sens. Environ.* 114, 2970–2980. doi:10.1016/j.rse.2010.08.003

Zhu, Z., 2017. Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications. *ISPRS J. Photogramm. Remote Sens.* 130, 370–384. doi:10.1016/j.isprsjprs.2017.06.013