PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Detection and classification of changes in agriculture, forest, and shrublands for land cover map updating in Portugal

Hugo Costa[1,2], Pedro Benevides[1], Francisco D. Moreira[1], Mário Caetano[1,2]

[1]*Direção-Geral do Território (Portugal)*

[2]*Univ. Nova de Lisboa (Portugal*)

Costa, H., Benevides, P., Moreira, F. D., & Caetano, M. (2022). Detection and classification of changes in agriculture, forest, and shrublands for land cover map updating in Portugal. In C. M. U. Neale, & A. Maltese (Eds.), Proceedings of SPIE. Sensing for Agriculture, Ecosystems, and Hydrology XXIV (Vol. 12262, pp. 19). SPIE. Society of Photo-Optical Instrumentation Engineers. https://doi.org/10.1117/12.2636127.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Detection and classification of changes in agriculture, forest and shrublands for land cover map updating in Portugal

Hugo Costa*[ab], Pedro Benevides[a], Francisco D. Moreira[a], Mário Caetano[ab]
[a]Direção-Geral do Território, 1099-052 Lisboa, Portugal; [b]NOVA Information Management School (NOVA IMS), Universidade Nova Lisboa, Campus de Campolide, 1070-312 Lisbon, Portugal

ABSTRACT

Portugal produced a land cover map for 2018 based on Sentinel-2 data and represents 13 classes, including agriculture, six tree forest species, and shrubland. The map was updated for 2020. The strategy focused on three strata where annual changes occur: S1 (agriculture) due to crop rotation, S2 (forest and shrubland) due to wildfires and clear-cuts, and S3 (fire scars and clear-cuts of previous years) where vegetation regeneration occurs. The methodology included i) change detection, ii) classification, and iii) knowledge-based rules. Stratum S1 was classified with images of the entire 2020 crop year and a training dataset extracted from the national Land Parcel Identification Systems (LPIS) of 2020. The land cover nomenclature was expanded and class agriculture was split in three distinct classes, hence resulting a map with 15 classes in total. Change detection, implemented in stratum S2, analyzed the profile of NDVI since 2018 to find potential loss of vegetation. S2 and S3 were classified through two stages. First, images of the entire 2020 crop year were used and then data of October 2020 (end of crop year) to capture late changes. The training points of the 2018 land cover map were used, but only if not associated with NDVI change. For all the three strata, knowledge-based rules corrected misclassifications and ensured consistency between the maps. A comparison between 2018 and 2020 reveal important land cover dynamics related to vegetation loss and regeneration on ~5% of the country.

Keywords: Sentinel-2, COSsim, COSc, SMOS, NDVI, time series, operational programme, LCLU monitoring

1. INTRODUCTION

Land cover mapping is fundamental for a wide range of applications spanning multiple domains such as policy and research[1,2]. Thus, new initiatives from continental to global scales have started to explore the wealth of data provided by the open archives of missions Landsat and Sentinel to narrow the gap between exploratory and operational applications[3–5]. However, challenges remain at the national scales, which often require frequent updating of land cover maps[6,7].

Portugal has a recent national land cover monitoring system called SMOS (*Sistema de Monitorização da Ocupação do Solo*) developed to produce land cover products to assist decision-making, assessment of policies, business, and so forth. One such product is an annual land cover map in raster format with pixels of 10 m derived from Sentinel-2 multispectral and multitemporal data. The first map was produced relative to the crop year of 2018 (from October 2017 to September 2018). The methodology of this map implements a spatially stratified and multi-stage approach to automatically classify Sentinel-2 data[7].

The methodology developed for the 2018 map is suitable to produce a map from the outset, but needs adaptations for consecutive and annual updating. This is because the repetition of the methodology for a more recent year would often allocate pixels to different land cover classes simply due to casual variations on classification, for example, due to classification uncertainty. Such variations across maps of different years could be interpreted as land cover change. Therefore, additional research has been undertaken[8–10] to redesign the methodology and ensure consistency between the annual land cover maps.

The new methodology consists in change detection and classification of the changes occurred in annual agriculture, forest and shrublands. That is, the map of the previous year is regarded when producing a new map by focusing only on the changes occurred meantime and allocating them to a new suitable class. The methodology initially developed for 2018 is the basis for map production and hence a spatially stratified and multi-stage approach is still used with some modifications. This paper describes the methodology used to update the map of 2018 for 2020.

*hcosta@dgterritorio.pt; phone (+351) 21 381 96 00; www.dgterritorio.gov.pt

2. BASE MAP AND DATA SETS

The base map subject to updating is a land cover map of continental Portugal for 2018. This map includes 13 classes, namely Artificial land, Agriculture, Evergreen oaks, Eucalyptus, Other broadleaves, Maritime pine, Stone pine, Other conifers, Shrubland, Natural grassland, Bare soil, Wetland, and Water. The map was produced based on Sentinel-2 data and inherits the pixel size of 10 meters. The overall accuracy of the map was estimated at 81.3%, with a confidence interval of ± 2.1%[7].

The satellite data used to update the base map for 2020 were all Sentinel-2 images with cloud cover below 50% acquired in the crop years of 2019 and 2020 (i.e., October 2018 to September 2020). Although updating was related to the 2020 crop year, the previous crop year was needed for change detection. One additional month after the period of interest (October 2020, which belongs to the 2021 crop year), was also used for one specific analysis. The Sentinel-2 data were processed with the MAJA processor[11] as downloaded from the Theia Land Data Centre (Theia). The masks produced by MAJA removed clouds or cloud shadows from all Sentinel-2 bands, except bands 1, 9, and 10 as they were not used in classification because they are used for correction of atmospheric effects. The spectral bands acquired at 20 m spatial resolution were disaggregated to 10 m resolution.

The Normalized Difference Vegetation Index (NDVI) was calculated from each processed image and used for change detection. For classification, 12 multispectral composites were produced (2020 crop year) representing cloud-free monthly medians. One additional composite for October 2020 (2021 crop year) was produced too. Five spectral indices were calculated from the monthly composites in addition to spectral–temporal metrics to summarize the 2020 crop year. The image processing followed the methodology of the base map[7].

Auxiliary data were used for multiple purposes at the various stages of map production. They included some High-Resolution Layers (HRL) from the Copernicus Land Monitoring Service (CLMS) and national data sets, including the official land use and land cover map of 2018 (COS), official burnt area perimeters of 2020 and previous years, the Land Parcel Identification System (LPIS) of 2020, and a national digital terrain model (DTM) at 25 m spatial resolution.

3. METHODOLOGY

3.1 Mapping zones and subzones

Continental Portugal was divided in a set of strata with two hierarchical levels called mapping zones and subzones (Figure 1). Mapping zones were defined to deal with the landscape diversity of Portugal whereas mapping subzones represented different characteristics and change dynamics of the territory. Because the mapping zones relate to the Portuguese landscape, the same 14 mapping zones of the base map[7] were used for map updating. The mapping subzones, on the contrary, were adapted to focus on expected relevant change dynamics occurred since 2018.

Three mapping subzones were defined to delimit areas where most of the land cover change occurs in Portugal in short periods such as one or two years. The first subzone (S1) focused on annual agriculture because of crop rotation. The second subzone (S2) focused on forest and shrubland that may have faced disturbance due to wildfires or clear-cuts since 2018. The third subzone (S3) focused on fire scars and clear-cuts previous to 2018 where vegetation regeneration is expected, or on the contrary, new clearings may have occurred (e.g., removal of burnt vegetation for new plantations). The remaining area of the mapping zones, not included in any of the subzones (e.g., settlements), were considered stable over time, and therefore, the updated map preserved the class of the base map for consistency. However, the updated map presents a more detailed land cover nomenclature as compared to the base map. Because subzone S1 focus on annual agriculture, the new map splits the 2018's class Agriculture in three new classes, namely two types of annual crops (Autumn/Winter crops and Spring/Summer crops), and Other agricultural areas (permanent crops and managed grassland). The resulting map represents 15 classes rather than 13 classes.

The subzones were delimited based on overlapping and comparing the base map, auxiliary data, and change detection analysis. Subzones S1 (agriculture) corresponded to the pixels classified in the base map as Agriculture plus Shrubland, Natural grassland, and Bare soil that overlap with the annual agriculture class in COS (meaning potential annual agricultural land use)[10]. S2 used the official burnt area maps of 2019 and the change detection results to capture potential wildfires in 2020 and clear-cuts. This subzone included only pixels classified as forest trees and shrublands in the base map because the methodology assumes only these pixels are subject to vegetation loss. S3 gathered areas where forest and shrubland suffered disturbance in 2018 or earlier, meaning that vegetation possibly recovered between 2018 and 2020. Vegetation growth can be slow, but may require updating after some time (e.g., Natural grassland evolving to

Shrubland). S3 was delimited based on past data sets[7], namely official burnt area maps and clear-cuts detected in 2015-2018 with Landsat and during the production of the base map with Sentinel-2. Only pixels classified as Shrubland, Natural Grassland and Bare soil in the base map were used in this stratum, as these are the classes with potential recovery.

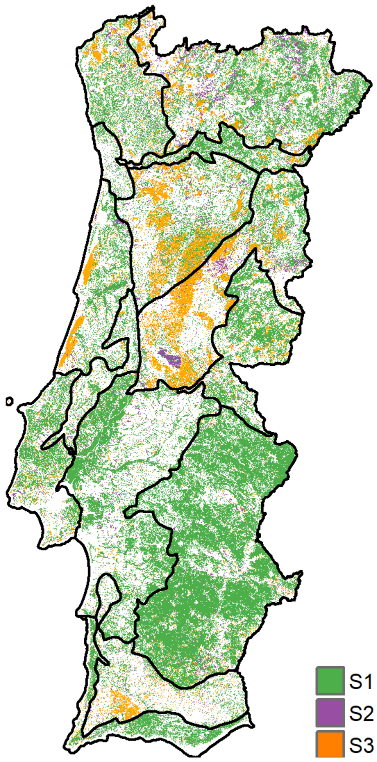

Figure 1. Mapping zones (black outline) and subzones focused on agriculture (S1), forest and shrubland that may have faced wildfires or clear-cuts since 2018 (S2), and fire scars and clear-cuts previous do 2018 (S3).

3.2 Change detection

The detection of changes followed a method used in a specific production stage of the base map[7]. This method detects abrupt decline of the NDVI by calculating variations within a moving time window of 10 observations. This window calculates the difference between the NDVI values after and before the central instance of the window and returns the largest decrease. Then, a threshold can be applied to define regions of substantial NDVI decline potentially associated with loss of vegetation.

The method analyzed the crop years of 2019 and 2020 and the largest decrease of NDVI along the time series in each pixel was retained for both years. A threshold of -0.17[7] was applied to identify pixels potentially associated with loss of vegetation. All patches smaller than 50 pixels (i.e., 0.5 ha at the resolution of Sentinel-2) were ignored to filter out spurious spectral variations.

3.3 Classification stages

Map updating followed a series of stages, namely: 1) image classification, 2) late change classification, and 3) knowledge-based rules. The first and second stages used random forest for classification, but differed on their goals. The third stage used expert-knowledge to fix errors of the classifications and improve logic and consistency between the updated map and the base map. The three stages were implemented on the mapping zones independently. The final map resulted from merging the 14 mapping zones.

The first image classification included spectral data of the whole 2020 crop year to provide the classifier with information on spectral seasonality or stability. In Agriculture (S1), the random forest was trained with classes representing annual crops (e.g., corn), Shrubland, Natural grassland, and Bare soil. Classes other than annual crops were needed to detect fallow and abandonment. The annual crops were merged into two classes: Autumn/Winter crops and

Spring/Summer crops[10]. In Subzones S2 and S3, the training classes were Shrubland, Natural grassland, Bare soil, and the several forest species. The classes were selected because they are expected as a result of either vegetation loss or vegetation recovery. However, in both S2 and S3, the classification result could be similar to the class on the base map (e.g., Shrubland in 2018 and 2020) because the subzones delimited potential change, including false change, and thus the classification stage was an opportunity to fix inaccuracies of the change detection method. The training data produced to represent the 2020 crop year corresponded to the same locations (pixels) of the 2018 training data after excluding the locations associated with potential change. The exception was the annual crops, whose training data were completely new and automatically generated based on the LPIS of 2020.

The second classification was needed in some of the pixels of subzones S2 and S3 because wildfires or clear-cuts occurring at the end of the crop year may pass unnoticed in the first classification. This happens when most of the spectral information in a pixel represents vegetation before change. However, the updated map should represent the land cover after change, even if recent. Therefore, late change classification was performed with spectral data of October 2020, which belongs to the following crop year, but enabled the updating map to capture changes by September 2020.

Knowledge-based rules combined the outputs of both classifications and corrected errors, like originally implemented in the production of the base map. For map updating, the rules were also needed to ensure consistency between the maps. For example, if a pixel of undisturbed forest was included in subzones S2 or S3 (i.e., false change), and the random forest classified that pixel with a forest species different from that of the base map, the rules reverted the new classification to maintain the original class. However, for cases considered correctly classified in 2020, the rules propagated the 2020 classification back to the base map. Therefore, the production of the 2020 map resulted in a revised version of the base map.

4. RESULTS AND DISCUSSION

The methodology produced the new map for the 2020 crop year and a revised version of the base map (Figure 2). The detail in agriculture provided by the new map has a large impact on its appearance as compared to the base map. Spring/Summer crops are evident along the largest rivers of the country, notably river Tagus, and generally scattered in the country where center pivots abound or the natural conditions for this type of annual crops are most favorable. The new map also reveals large areas of Autumn/Winter crops, but Other agriculture is the most common, which includes permanent crops and managed grassland.

The change detection method was able to delimit the various land cover changes related to wildfires and clear-cuts occurred since 2018. Loss of vegetation during 2019 is visible on the map, as well as events occurred at the end of the 2020 crop year (September). For example, Figure 3 shows the fire scar of the largest wildfires of 2019 and 2020, which occurred relatively close to each other in central Portugal. The 2020 fire extinguished close to the end of the crop year (17 September 2020) and was mapped thanks to the second image classification implemented to capture this kind of situations. Most of the fire scar was mapped as Bare soil because of fire severity and recent date while the fire scar of 2019 shows some vegetation recovery.

A comparison between the maps reveals important land cover dynamics (Table 1). Changes from Eucalyptus to Natural grassland (20 kha) or to Bare soil (12 kha), and also from Shrubland to Natural grassland (59 kha) or Bare soil (9 kha) are typical cases of vegetation loss caused by wildfires and clear-cuts. On the contrary, vegetation recovery is revealed by changes such as from Natural grassland to Eucalyptus (23 kha) and from Natural grassland to Shrubland (79 kha). Dynamics in agricultural areas are largely obscured by the base map nomenclature and thus Table 1 only shows the general class Agriculture. Nevertheless, it is notable the changes involving classes Agriculture, Shrubland, Natural grassland and Bare soil, which result from agricultural practices such as fallow.

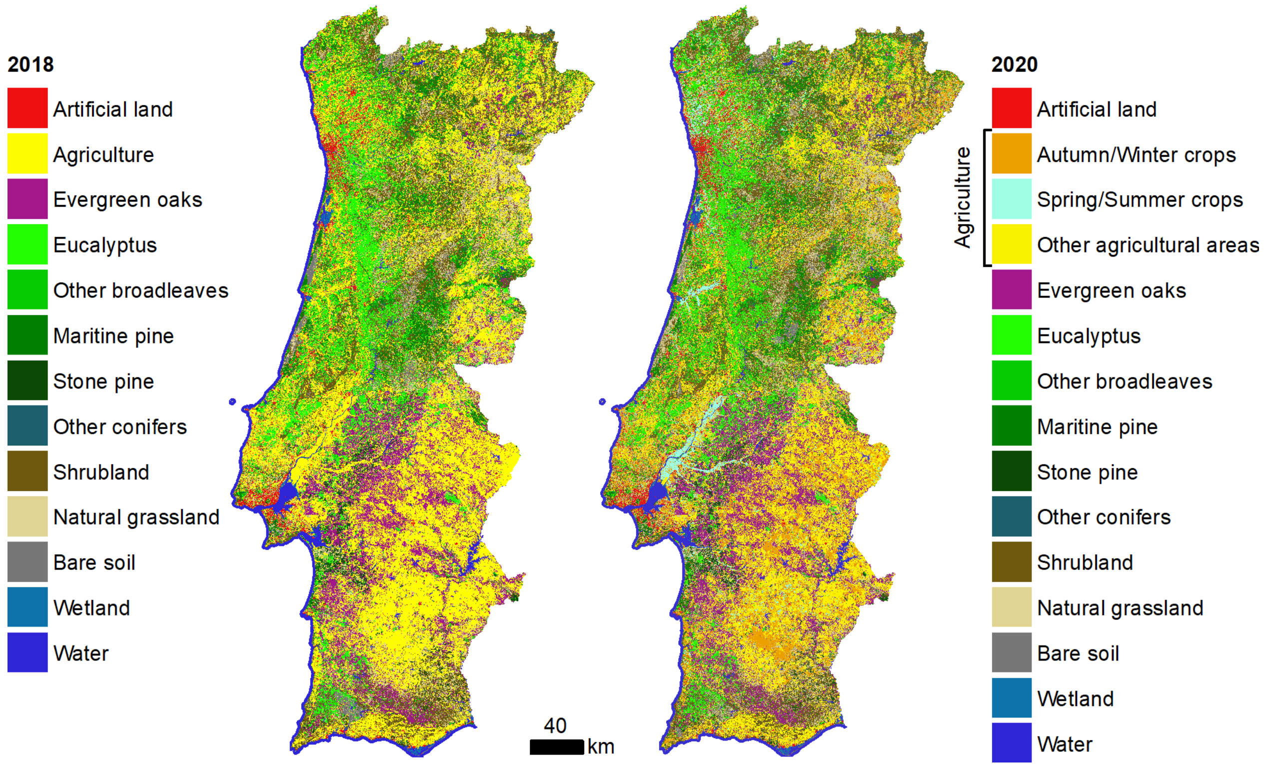

Figure 2. Land cover maps of 2018 and 2020.

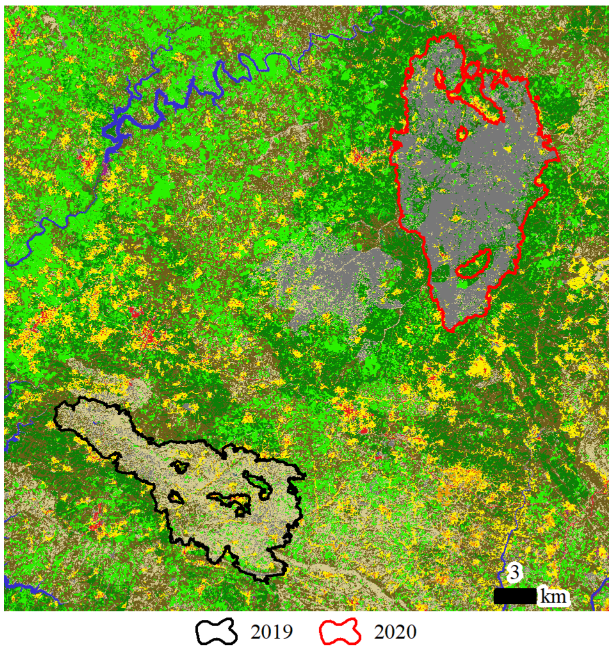

Figure 3. Largest fire scars of 2019 and 2020 (see Figure 2 for the map legend).

Table 1. Land cover change between 2018 (rows) and 2020 (columns) expressed in thousands of hectares (kha). Values <0.05 are omitted.

	AL	**Ag**	**EO**	**Eu**	**OB**	**MP**	**SP**	**OC**	**Sb**	**NG**	**BS**	**We**	**Wa**
Artificial land (**AL**)	251												
Agriculture (**Ag**)		3055							3	8	12		
Evergreen oaks (**EO**)			856								1		
Eucalyptus (**Eu**)				759					3	20	12		
Other broadleaves (**OB**)					397				2	3	1		
Maritime pine (**MP**)						414			6	8	12		
Stone pine (**SP**)							139				1		
Other conifers (**OC**)								19	0	0	0		
Shrubland (**Sb**)		1		0	0		0		1386	59	9		
Natural grassland (**NG**)		5		23					79	938	29		
Bare soil (**BS**)		1		5					29	53	154		
Wetland (**We**)												21	
Water (**Wa**)													317

The area associated with change between the maps corresponds to 5.4% of the total area. The rate of change for continental Portugal is usually smaller[12,13], but typically statistics rely on maps such as COS and CORINE Land Cover, which have a stronger emphasis on land use. On the contrary, the new maps focus more on land cover and therefore represent changes often not mapped, such as the impact of wildfires and clear-cuts, which do not represent directly land use change. Therefore, land cover changes are expected to be larger than land use changes. Future updates of the map, which can include crop rotation due to the expansion of the land cover nomenclature to 15 classes, can reveal even larger changes and help to better understand the annual practices in agriculture.

5. ACKNOWLEDGEMENTS

The work has been supported by projects foRESTER (PCIF/SSI/0102/2017), and SCAPE FIRE (PCIF/MOS/0046/2017), and by Centro de Investigação em Gestão de Informação (MagIC), all funded by the Portuguese Foundation for Science and Technology (FCT). Value-added data processed by CNES for the Theia data centre www.theia-land.fr using Copernicus products. The processing uses algorithms developed by Theia's Scientific Expertise Centres.

REFERENCES

[1] Cieślak, I., Biłozor, A. and Szuniewicz, K., "The Use of the CORINE Land Cover (CLC) Database for Analyzing Urban Sprawl," Remote Sens. **12**(2), 282 (2020).

[2] Pettorelli, N., Wegmann, M., Skidmore, A., Mücher, S., Dawson, T. P., Fernandez, M., Lucas, R., Schaepman, M. E., Wang, T., O'Connor, B., Jongman, R. H. G., Kempeneers, P., Sonnenschein, R., Leidner, A. K., Böhm, M., He, K. S., Nagendra, H., Dubois, G., Fatoyinbo, T., et al., "Framing the concept of satellite remote sensing essential biodiversity variables: challenges and future directions," Remote Sens. Ecol. Conserv. **2**(3), D. Boyd, Ed., 122–131 (2016).

[3] Brown, J. F., Tollerud, H. J., Barber, C. P., Zhou, Q., Dwyer, J. L., Vogelmann, J. E., Loveland, T. R., Woodcock, C. E., Stehman, S. V., Zhu, Z., Pengra, B. W., Smith, K., Horton, J. A., Xian, G., Auch, R. F., Sohl, T. L., Sayler, K. L., Gallant, A. L., Zelenak, D., et al., "Lessons learned implementing an operational continuous

United States national land change monitoring capability: The Land Change Monitoring, Assessment, and Projection (LCMAP) approach," Remote Sens. Environ. **238**(August 2019), 111356 (2020).
[4] Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, L., Tsendbazar, N. E., et al., "ESA WorldCover 10 m 2020 v100."
[5] Malinowski, R., Lewiński, S., Rybicki, M., Gromny, E., Jenerowicz, M., Krupiński, M., Nowakowski, A., Wojtkowski, C., Krupiński, M., Krätzschmar, E. and Schauer, P., "Automated production of a land cover/use map of europe based on sentinel-2 imagery," Remote Sens. **12**(21), 1–27 (2020).
[6] Tulbure, M. G., Hostert, P., Kuemmerle, T. and Broich, M., "Regional matters: On the usefulness of regional land-cover datasets in times of global change," Remote Sens. Ecol. Conserv., 1–12 (2021).
[7] Costa, H., Benevides, P., Moreira, F. D., Moraes, D. and Caetano, M., "Spatially Stratified and Multi-Stage Approach for National Land Cover Mapping Based on Sentinel-2 Data and Expert Knowledge," Remote Sens. **14**(8), 1865 (2022).
[8] Moraes, D., Benevides, P., Costa, H., Moreira, F. and Caetano, M., "Assessment of the introduction of spatial stratification and manual training in automatic supervised image classification," Proc. SPIE - Int. Soc. Opt. Eng., 36 (2021).
[9] Moraes, D., Benevides, P., Costa, H., Moreira, F. D. and Caetano, M., "Influence of Sample Size in Land Cover Classification Accuracy Using Random Forest and Sentinel-2 Data in Portugal," 2021 IEEE Int. Geosci. Remote Sens. Symp. IGARSS, 4232–4235, IEEE (2021).
[10] Benevides, P., Costa, H., Moreira, F. D., Moraes, D. and Caetano, M., "Annual Crop Classification Experiments in Portugal Using Sentinel-2," 2021 IEEE Int. Geosci. Remote Sens. Symp. IGARSS, 5838–5841, IEEE (2021).
[11] Baetens, L., Desjardins, C. and Hagolle, O., "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**(4), 1–25 (2019).
[12] Costa, H., Almeida, D., Vala, F., Marcelino, F. and Caetano, M., "Land cover mapping from remotely sensed and auxiliary data for harmonized official statistics," ISPRS Int. J. Geo-Information **7**(4) (2018).
[13] Meneses, B., Reis, E., Reis, R. and Vale, M., "The Effects of Land Use and Land Cover Geoinformation Raster Generalization in the Analysis of LUCC in Portugal," ISPRS Int. J. Geo-Information **7**(10), 390 (2018).