

CASO PRÁCTICO

The 2017 Land Use/Land Cover Map of Catalonia based on Sentinel-2 images and auxiliary data

González-Guerrero, O. ^{*}, Pons, X. 

Grumets research group. Geography Department. Building B. Universitat Autònoma de Barcelona. 08193 Bellaterra, Catalonia, Spain.

Abstract: This paper details the process of generating the 2017 Land Use/Land Cover Map of Catalonia (MUCSC) using automatic classification of satellite imagery and auxiliary cartographic and remote sensing data. A total of 60 images (6 dates for each of the 10 tiles covering Catalonia) captured by the Sentinel-2A and Sentinel-2B satellites were used. These images as well as texture variables, terrain models derived from lidar processing, and vegetation and wetness indices were classified using the k-Nearest Neighbor algorithm (kNN) to obtain a map with 25 categories. The categories related to urbanized areas (urban areas, urbanizations and industrial zones/commercial areas), road infrastructures and burned areas were edited using official cartographic datasets of the Catalan Government [Generalitat]. The results have an overall accuracy greater than 98 %, which was evaluated with a set of more than 8.6 million independent test pixels. This work represents an important milestone in terms of the computational effort it involves due to the territorial extension (32 000 km²), the spatial detail of between 2 and 20 m, the use of up to 58 variables, the relative completeness of the legend and the level of success achieved. The MUCSC 2017, which is part of a 30-year quinquennial series beginning in 1987, can be downloaded in different formats (also in MMZX: new ISO 19165-2) and at resolutions of 10 m and 30 m pixel side from the Ministry of Territory and Sustainability website of the Catalan Government.

Key words: Land Use/Land Cover, LULC, Sentinel-2, Catalonia.

Uso de Sentinel-2 y datos auxiliares para la generación, mediante clasificación de imágenes, del Mapa de Usos y Cubiertas del Suelo de Cataluña 2017

Resumen: En este artículo se detalla el proceso de generación del Mapa de Usos y Cubiertas del Suelo de Cataluña (MUCSC) 2017 mediante clasificación automática de imágenes de satélite y datos cartográficos y de teledetección auxiliares. Con este propósito se han utilizado un total de 60 imágenes (6 fechas sobre cada una de las 10 teselas que cubren Cataluña) captadas por los satélites 2A y 2B de la constelación Sentinel-2. Estos datos, junto con variables de textura, modelos del terreno derivados del procesado lidar para todo el territorio e índices de vegetación y humedad, han sido clasificados con el algoritmo de inteligencia artificial kNN para obtener un mapa de 25 categorías, de las cuales las referentes a zonas urbanizadas (zonas urbanas, urbanizaciones y zonas industriales y comerciales), vías de comunicación y zonas quemadas han sido editadas utilizando bases cartográficas oficiales de la Generalitat [Gobierno] de Catalunya. Los resultados muestran un acierto global superior al 98 % evaluado mediante un conjunto de más de 8,6 millones de píxeles independientes de test. Este

To cite this article: González-Guerrero, O., Pons, X. 2020. The 2017 Land Use/Land Cover Map of Catalonia based on Sentinel-2 images and auxiliary data. *Revista de Teledetección*, 55, 81-92. <https://doi.org/10.4995/raet.2020.13112>

^{*} Corresponding author: oscar.gonzalez.guerrero@uab.cat

trabajo representa un hito importante tanto por el esfuerzo de cálculo que ha supuesto (extensión territorial de 32.000 km², detalle espacial de entre 2 y 20 m y uso de hasta 58 variables), como por la relativa completitud de la leyenda y por el nivel de acierto conseguido. El MUCSC 2017, que forma parte de una serie quinquenal de 30 años desde 1987, está disponible para descarga en distintos formatos (también en MMZX: nueva ISO 19165-2) y a resoluciones de 10 m y 30 m de lado de píxel a través de la página web en el Departamento de Territorio y Sostenibilidad de la Generalitat de Catalunya.

Palabras clave: cubiertas y usos del suelo, LULC, Sentinel-2, Cataluña.

1. Introduction

One of the basic elements for monitoring the state of the environment properly is to study the land use/land cover (LULC) dynamics. This makes it possible to understand the changes in the territory, anticipate future processes and plan possible actions. It is therefore necessary to have thematic cartography datasets with high spatial and temporal resolutions so that detailed analyses can be carried out (Townshend et al., 2012).

Currently, the main data source for generating land cover maps is satellite imagery, which enables data to be obtained recurrently, at a low cost and at both local and regional scales (Elvidge et al., 2004; Hansen and Loveland, 2012). The Landsat program of the United States of America is a paradigmatic example of this type of data. It has had eight Earth observation satellites and, since 2008, has provided images for free for all users. These images are from all around the globe and have been captured between the early 1970s and the current date (Woodcock et al., 2008; Wulder et al., 2012).

However, Landsat data has certain limitations in terms of automatically classifying satellite images for obtaining land cover cartography, mainly due to its temporal resolution. This resolution has varied between platforms over time, going from the 18-day revisit period of the first satellites launched into orbit, to the 16-day period of the Landsat 4 satellites onwards (Loveland and Dwyer, 2012). This periodicity implies that in some areas of the planet it is not possible to obtain cloud-free images in key phenological periods to distinguish between certain types of vegetation cover. This means the producer has to use data from different years or to work with less detailed legends.

The launching of the Sentinel-2A and Sentinel-2B satellite constellation by the European Space Agency (ESA) has made possible, under standard

conditions, image availability of the same point on Earth every five days, which increases the probability of capturing images in the best atmospheric conditions. Furthermore, the Sentinel-2 constellation also improves the spatial and spectral resolution compared to Landsat, as Sentinel-2 has a larger number of available spectral bands and spatial resolutions of 10, 20 and 60 m (ESA, 2015).

In Catalonia there are different land use and cover maps. The series known as “Land Use/Land Cover Maps of Catalonia” (MUCSC) is the longest and most regular available to date. This series was initiated by the current Cartographic and Geological Institute of Catalonia (ICGC) thanks to the support of the Ministry of the Environment of the Catalan Government [Generalitat]. In the first phase, four “land use” maps were generated, although their legend was more the one of a land cover map, corresponding to the years 1987, 1992, 1997 and 2002. They were created using satellite data information from the Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) sensors onboard Landsat satellites. These maps, generated according to a raster data model with the same legend of 21 categories, cover the entire Catalan territory and have a pixel size of 30 m × 30 m (ICG, 1992).

In a second phase, in 2017, the Geography Department of the Universitat Autònoma de Barcelona (UAB) and the Center for Ecological Research and Forestry Applications (CREAF), with the support of the Ministry of Territory and Sustainability of the Generalitat, generated two new maps corresponding to the years 2007 and 2012. As in the previous cases, these maps were made by automatic classification of images captured by the TM and Operational Land Imager (OLI) sensors of the Landsat satellites. These maps, mainly produced for LULC dynamics studies, have the same raster grid as the maps made by ICGC; however, they show an extended legend

of 23 categories, but with backward compatibility (the extension corresponds to the subdivision of some of the previous categories).

In 2018, the same Ministry, and with the same aim of having data for accurate LULC dynamics studies, commissioned CREAM and the UAB to make the 2017 map to completely update the five-year series, expanding the legend to 25 categories [again with backward compatibility] (Figure 4) and improving the spatial resolution by using data from the MSI sensor of the Sentinel-2A and Sentinel-2B satellites.

Thus, the purpose of this article is to explain the process of generating the MUCSC 2017 from Sentinel-2 satellite images and other auxiliary cartographic and remote sensing datasets.

2. Study area

Located in the northeast of the Iberian Peninsula, Catalonia (Figure 1) is part of its territory along the Mediterranean Sea and has an area of about 32 000 km². This surface includes both mountain areas and large plains. It has an average altitude above sea level of close to 600 m, and the altitude ranges from 0 m to more than 3 000 m at some of the mountain peaks in the Pyrenees.

There is a climatic gradient going from the Pyrenean zones in the north, which have a temperate boreal climate, towards the southern zones, which have a more Mediterranean climate characterized by mild winters and dry summers. There is also a continental gradient running towards the

west, where the climate is semiarid with very warm summers and colder winters. The described geography and climate contribute to the configuration of a landscape dominated by forests, shrublands and grasslands (60 %), while crops represent approximately one third of the total area, generally occupying areas with lower slope.

3. Material and methods

Unlike previous MUCSCs, in order to create the 2017 map, it was decided to work with images from the Sentinel-2A and Sentinel-2B satellite constellation. The combined use of images from the two satellites reduces the revisit period, which increases the probability of having relatively cloud-free images within the study year, providing a product with greater temporal consistency.

MUCSC 2017 was obtained by automatic classification with the non-parametric algorithm kNN (k-Nearest Neighbors). Specifically, we used a parallelized, 64-bit implementation version of this algorithm in MiraMon (Pons 2004), which allowed us to generate the maps within a reasonable time taking advantage of the 40 processors and 128 Gbytes of RAM available in the computer. In addition, this implementation makes it possible to obtain a wide range of quality auxiliary layers giving information about uncertainty through maps of minimum and maximum distances to the final classes, maps of medians of the k distances, maps of the number of used neighbors from the training areas used and the uncertainty map itself, particularly informative.

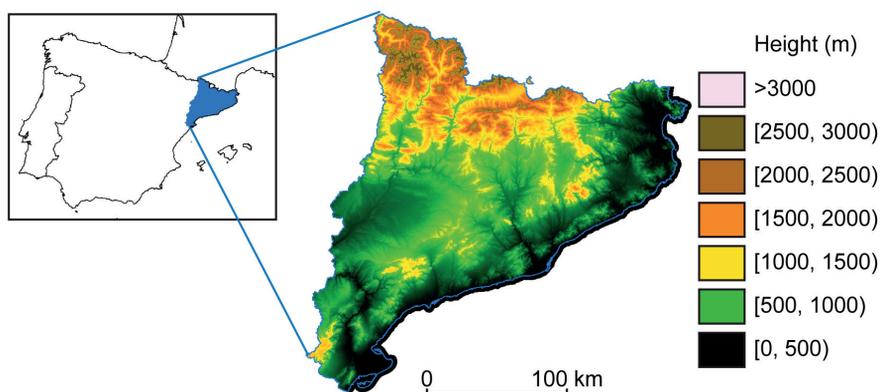


Figure 1. Map of the location and altitudinal distribution of Catalonia.

Some of the legend categories are edited *a posteriori* from different official cartographic sources (see Section 3.5).

Below we describe the image processing, the different variables included in the classifier (58 variables), the selection of training and test areas, and the editing process of the results.

3.1. Pre-processing of the Sentinel-2 images

Ten tiles of 100 km×100 km, corresponding to the orbits 008 and 051, are necessary to cover the entire Catalan territory (Figure 2). From each of the tiles, six images distributed between March and September 2017 were selected to obtain a good representation of the phenological dynamics of the natural vegetation and crop covers (Moré and Pons, 2007). The full list of images used can be found in Table 1.

We made a mosaic of all the tiles that correspond to the same image capture date, so that the areas are unified according to the acquiring orbit. Therefore, we can work with just two large areas of the territory determined by orbits 051 and 008 (Table 1).

To eliminate the areas covered by clouds and their shadows we manually edited the two series of image mosaics. This process is more accurate than

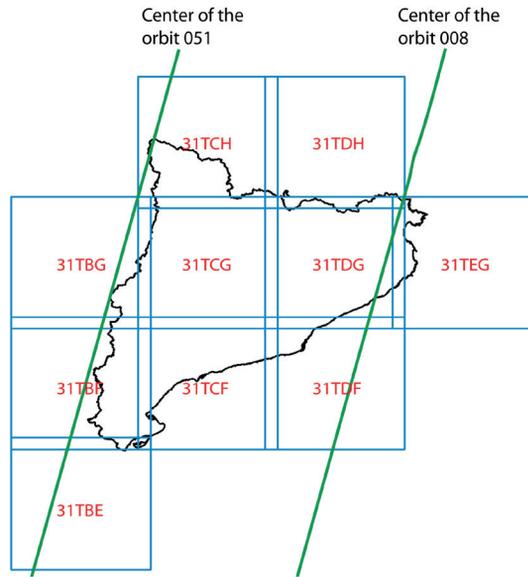


Figure 2. Distribution of Sentinel-2 tiles necessary to cover Catalonia.

automatic algorithms, which are faster but obtain lower quality results.

Next, we carried out a radiometric correction process (atmospheric and topographic) using as a reference the pseudo-invariant radiometry areas obtained from the temporal analysis of the MODIS image series, as explained in Pons et al. (2014), Padró et al. (2017) and Padró et al. (2018), with

Table 1. The “x” marks the dates of the images selected for each of the Sentinel-2 tiles that cover the Catalan territory.

		14/03/2017	17/03/2017	03/04/2017	06/04/2017	06/05/2017	16/05/2017	26/05/2017	12/06/2017	05/07/2017	12/07/2017	27/07/2017	04/08/2017	14/08/2017	21/08/2017	26/08/2017	05/09/2017	10/09/2017	13/09/2017	18/09/2017	
Orbit 051	31TBE	x		x		x			x			x									x
	31TBF		x	x		x			x			x									x
	31TBG		x		x		x			x			x								x
	31TCF		x		x	x				x			x								x
	31TCG		x		x			x		x			x								x
	31TCH		x		x			x		x			x								x
Orbit 008	31TDF		x						x		x					x					x
	31TDG		x						x			x			x		x				
	31TDH		x						x			x			x		x				
	31TEG		x						x			x			x		x				

additional improvements that make it possible to have data in areas where only diffuse radiation is received.

Once the images were radiometrically corrected, those images of the same orbit captured in a period less than or equal to 15 days were considered to have sufficiently similar capture conditions to be unified. Although the radiometric correction method used guarantees the temporal consistency of the images, it is possible to find differences in reflectance of the same cover on two different dates of up to 2%. Therefore, this second mosaic process was not performed directly, but instead the radiometric values were adjusted by linear regression on the overlap areas between images to ensure maximum consistency of the results (Table 2).

Table 2. Mosaics composed of images captured with a margin of up to 15 days, necessary for covering the entire Catalan territory at six different time points for each of the orbits.

Orbit 051	17/03/2017	Orbit 008	14/03/2017
	06/04/2017		03/04/2017
	16/05/2017		12/06/2017
	05/07/2017		27/07/2017
	14/08/2017		21/08/2017
	13/09/2017		05/09/2017

A mask generated with thresholds of the Normalized Difference Snow Index (NDSI) (Cea et al., 2007) was applied to the tiles resulting from the above described process to remove the areas covered by snow.

Finally, due to the large volume of data represented by the mosaics of the 051 orbit images and to speed up the classification process, we divided the information into three parts: one part covered by the 008 orbit (Figure 3.A), and two parts covered by the 051 orbit (Figure 3.B and C).

3.2. Selection of variables for the classification generated from Sentinel-2 images

Based on the previous classification tests with different combinations of spectral variables on the 31TDG tile, it was determined that, of the 13 available spectral bands, the ones that work best in the

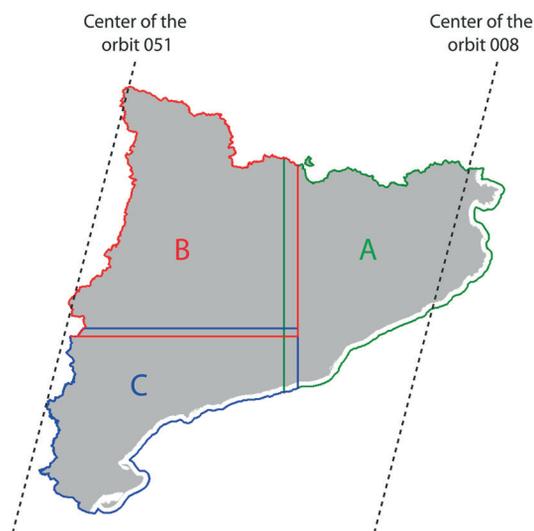


Figure 3. Division of the surface of Catalonia into the three areas to be classified, A: 1213 304.63 ha; B: 1 380 918.63 ha; C: 953 090.09 ha.

classification process for the legend that we want to obtain are bands 2, 3, 4, 8, 11 and 12 (these last two were resampled by bilinear interpolation at 10 m pixel side). This corresponds to the visible, near-infrared and short-wave infrared spectral regions.

In addition to the spectral bands, we included in the classifier the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI) (calculated with bands 11 and 8 corresponding to short wave infrared and near infrared) of the different dates. We also added a texture variable calculated with the standard deviation of band 8 in 21 pixel \times 21 pixel moving windows.

This makes a total of 54 variables derived from the images themselves.

The digital elevation model, the digital slope model and the solar radiation at the summer solstice were also used as classifier variables (Pons and Ninyerola, 2008). Finally, we also added a digital height model, DHM, with the elevation above the surface (DHM=digital surface model - digital elevation model=DSM-DEM) generated by processing some 40 000 million points of the first lidar cover of Catalonia (ICGC 2017) (minimum

density: 0.5 points/m²), the only lidar cover available in the territory at the time the MUCSC 2017 was created, and containing data captured between March 2008 and July 2012.

Thus, these four variables added to the 54 variables derived from the satellite images make up a set of 58 input variables for the classifier.

3.3. Obtaining the training and test areas

Training and test areas of the categories corresponding to cultivated areas were obtained from a plot selection of the 2017 Catalan Crop Map generated from the Declaración Única Agraria (DUN) and the SIGPAC [GIS of the Common Agricultural Policy] geographical database of the Ministry of Agriculture, Livestock and Fisheries of the Generalitat. This map provides information on the type of crop (at the species level) and at the land plot level of all the owners who have an area for agricultural production (production for personal consumption was not included). An internal buffer of 30 m was applied to this selection and all the resulting polygons with an area of less than 1 ha were eliminated to avoid possible confusion between spatially neighboring categories.

For the rest of the legend categories, the training and test areas were obtained by photointerpretation of the current 1:5 000 orthophoto for the study period (orthophotos with a capture date between 2016 and 2018) and taking as a starting point the areas common to the same category in the entire series of the five-year maps since 1987. In the case of the different forest typologies, the databank of parcels of the fourth Spanish National Forest Inventory (density: 1 point/km² of forest) was used as auxiliary data.

All of the above natural surface data were filtered depending on their response to the NDVI. The filters applied were based on the work of Padial et al. (2019) and aim to characterize each cover according to its response to multiple criteria, such as the variation range of NDVI accumulated over all the available dates, the characterization of seasonal variability, thresholds at specific phenological times, etc.

Finally, more than 250 000 training pixels and 8.6 million test pixels were obtained through a

process of random selection stratified by categories. The strict randomness between them contributes to a minimal spatial dependence.

3.4. Post-processing of the classification

The result, obtained by automatic classification of the 58 variables discussed above using the Artificial Intelligence kNN algorithm, was edited using a selective mode filter, which was applied exclusively to polygons with surfaces less than 900 m². This process was carried out to avoid the “salt and pepper” effect in highly fragmented landscape areas that do not make sense at the scale at which the map is published according to the pixel dimension and the planimetric accuracy of the images used (1:25 000), and also to decrease the risk associated with the usage a minimum mapping unit too close to the spatial resolution of the sensor. In addition, vectorization was used to review those polygons with the higher figures in the uncertainty maps provided by the MiraMon kNN classifier.

3.5. Auxiliary cartographic datasets

The categories corresponding to urbanized areas (low and high density urban areas and industrial and commercial areas) and road infrastructures were mapped from official datasets such as the Urbanistic Map of Catalonia and the road graph of the Generalitat. These layers were reviewed according to the current 1:5 000 orthophotos (2016-2018) of the ICGC.

We selected the following road infrastructures from the road graph: toll motorways, free highways and motorways, access and ring roads, national highways, autonomous motorways and highways, and the basic network, in addition to those roads that have appeared in previous MUCSC versions although they do not belong to any of the above groups or which are clearly identifiable in the classification result. Likewise, we included those railway infrastructures that the classifier could detect due to their size. All elements previously selected were rasterized in the classification result with a width of 20 m for motorways and 10 m for the other roads.

The areas affected by forest fires were rasterized based from the official dataset of burned areas of the Generalitat, also available [here](#) in MMZX format.

4. Results

The classification and later editing resulted in the 2017 Land Use/Land Cover Map of Catalonia shown in Figures 4 and 5.

The overall accuracy of the map, evaluated with more than 8.6 million independent test pixels, was found to be 98.5% (kappa 0.98), with commission errors less than 0.03% (value corresponding to deciduous forests) and omission errors below 6.8% (value corresponding to lowland grasslands) in all categories of the legend. The confusion matrix between the different categories is shown in the Annex. This table does not include the categories “Permanent snowpacks”, “Low-density urban areas” or “Burned areas” as these categories were

edited later from data external to the classification. The evaluation of these categories would not be independent and including them could make the results misleading by increasing the overall accuracy obtained by the classification (although evidently they do increase the overall accuracy of the map).

Regarding the distribution of the land covers (Table 3), the results show that the main components of the territory are forests (34.7% of the total area) and the different types of crops (30.1%), followed by the combination of shrublands and grasslands (27.3%, including wetland vegetation). It is worth highlighting the area occupied by shrublands (22.4% of the total), which is the main cover in terms of area at an individual level (more than 700 000 ha).

Forests are distributed in nearly equal proportions of the total of 1 114 454.2 ha in the provinces

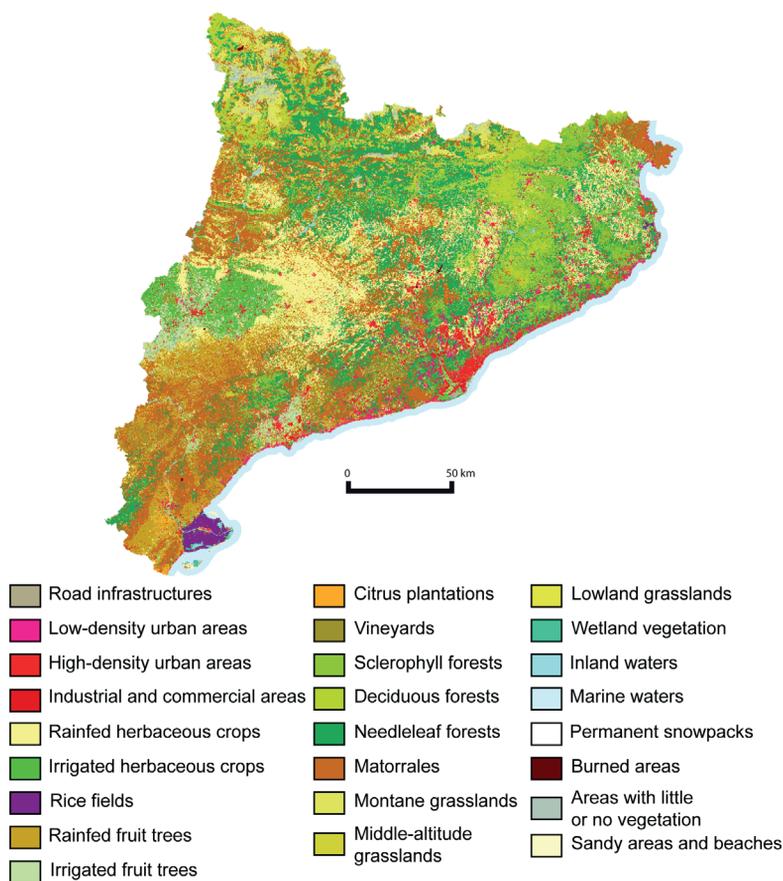


Figure 4. General view of the 2017 Land Use/Land Cover Map of Catalonia.

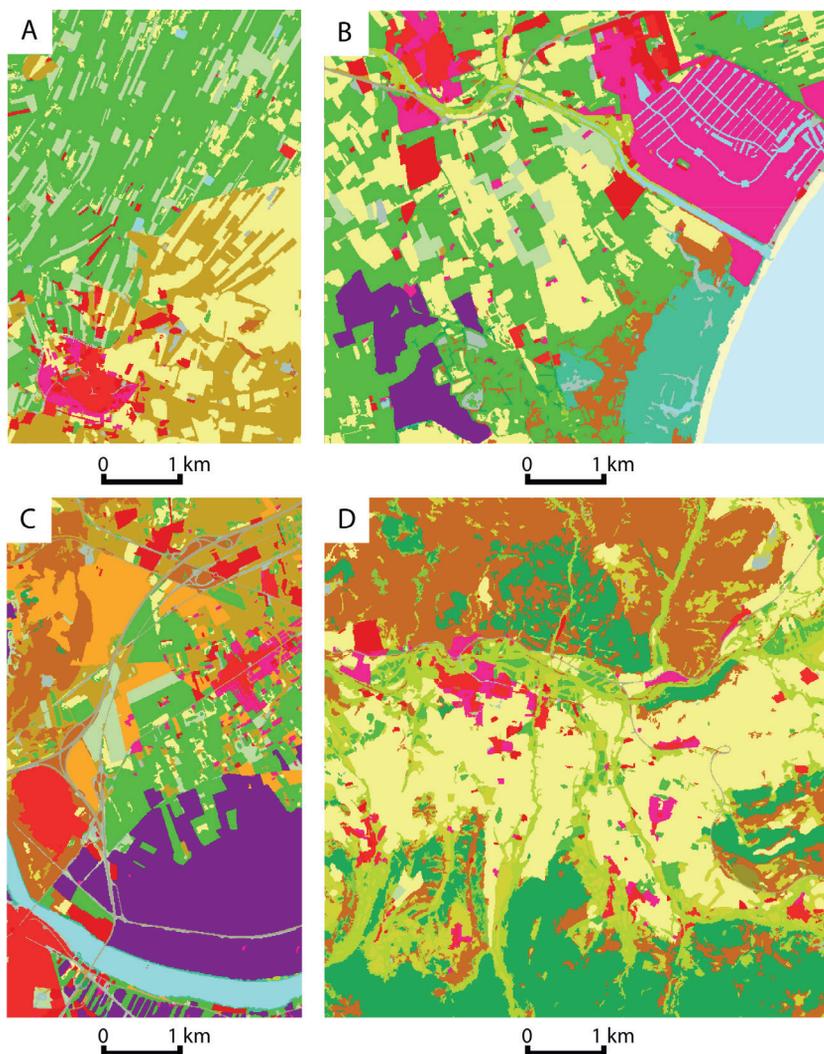


Figure 5. Details of the 2017 Land Use/Land Cover Map of Catalonia. A: Surroundings of Arbeca (Lleida); B: Empuriabrava and Aiguamolls de l'Empordà (Girona); C: Surroundings of Amposta (Tarragona); D: Surroundings of Bellver de Cerdanya (between Girona and Lleida).

of Barcelona, Girona and Lleida, leaving just over 100 000 ha for the province of Tarragona. The most abundant type of forests are needle-leaf forests, which represent 19.8% of the total. Sclerophyll forests are the less abundant in Catalonia (198 043.4 ha) and are concentrated mainly in Girona, representing 18.2% of the total area of this province.

In terms of crops, rainfed herbaceous crops stand out (13.7%) and are concentrated mainly in the provinces of Barcelona, Girona and Lleida (which represent between 13 and 17% of the surface area of the province). In Tarragona, however, the main

crop is rainfed fruit trees (17.3% of the province's surface area). Irrigated fruit trees are concentrated mainly in Lleida and Tarragona and represent 2.7% of the total surface area of Catalonia, while vineyards, which are important due to the economic activity they generate, account for 2.1% of the surface.

The urban areas and road infrastructures represent 5.4% of the total surface of Catalonia and, obviously, are concentrated in the provinces of Barcelona (more than 7 500 ha) and Tarragona (more than 4 000 ha).

Table 3. Total distribution, and by provinces, of the surface area of the different land covers.

	Area (ha)				Total	Total (%)
	Barcelona	Girona	Lleida	Tarragona		
Road infrastructures	7543.3	2571.2	3038.2	4085.9	17238.7	0.5
Low-density urban areas	34897.7	17241.7	5080.6	13862.0	71082.0	2.2
High-density urban areas	21831.6	6103.3	5302.2	6509.3	39746.5	1.2
Industrial and commercial areas	21929.6	6360.7	8694.3	8723.0	45707.6	1.4
Rainfed herbaceous crops	130511.6	76797.8	198075.3	34971.7	440356.3	13.7
Irrigated herbaceous crops	14338.2	34502.9	98342.1	6723.5	153906.7	4.8
Rice fields	-	1030.5	64.5	21267.0	22362.1	0.7
Rainfed fruit trees	4828.5	3913.3	70315.2	108941.6	187998.7	5.9
Irrigated fruit trees	1276.1	4951.5	54392.6	25359.7	85979.8	2.7
Citrus plantations	10.8	0.1	0.9	9153.4	9165.2	0.3
Vineyards	28099.5	2478.4	4404.5	32648.3	67630.7	2.1
Sclerophyll forests	48837.9	107508.3	25319.7	16377.6	198043.4	6.2
Deciduous forests	61946.2	98892.8	116578.9	3077.9	280495.7	8.7
Needleleaf forests	221774.5	117476.3	203133.9	93530.5	635915.1	19.8
Shrublands	149477.7	75841.3	273001.3	220433.3	718753.6	22.4
Montane grasslands	3281.7	14313.9	73127.4	-	90723.0	2.8
Middle-altitude grasslands	6509.6	10095.3	24706.6	3737.8	45049.4	1.4
Lowland grasslands	2010.7	1072.2	10834.8	4785.2	18702.8	0.6
Wetland vegetation	323.9	811.8	432.1	3066.8	4634.5	0.1
Inland waters	1460.7	1596.7	8383.8	4260.1	15701.3	0.5
Permanent snowpacks	-	-	203.7	-	203.7	0.0
Burned areas	707.7	8.5	793.2	177.8	1687.0	0.1
Areas with little or no vegetation	11109.0	6380.0	32328.4	6234.8	56052.2	1.7
Sandy areas and beaches	374.6	289.5	-	2652.4	3316.5	0.1
Total	773080.9	590237.9	1216554.0	630579.6	3210452.3	100.0

The MUCSC 2017, made entirely with MiraMon, is published at a 1:25 000 working scale and can be downloaded as a 10 m pixel-side raster data model in different formats (including MMZX: new ISO 19165 based on the article by Pons and Masó (2016)) along with the maps of all the five-year periods since 1987 from the webpage of the Ministry of Territory and Sustainability of the Generalitat. In addition, a version of the same map with a 30 m pixel side has also been generated, which makes a pixel-by-pixel comparison with previous maps possible. An example of an analysis of this series, in this case for agricultural studies, can be seen in González-Guerrero et al. (2019).

5. Conclusions

The launching of the two satellites that make up the Sentinel-2 constellation has made it possible to generate the Land Use/Land Cover Map of Catalonia (~32 000 km²) from images entirely from the year to be represented. This is an improvement over previous maps, in which the limitation

of images free from clouds, cloud shadows or other artifacts meant that images from various, close, years had to be used in order to cover key phenological periods so that the classifier could distinguish some of the land covers.

In addition, the higher spatial resolution compared to the sensors of the different Landsat platforms means the beginning of a new stage in the MUCSC series that opens the door to more detailed land analyses in the future. This will not, however, generate major problems for making comparisons with historical maps (after legend reclassifying) due to the care put into the backward compatibility of the new categories.

Similarly, thanks to both the spectral and spatial and radiometric resolution of the Sentinel-2 images and the available auxiliary information (it is worth noting another key source of remote sensing today: lidar data available for the entire territory), it has been possible to expand the map legend up to 25 categories, keeping the legend consistent with the previous maps and obtaining a very high level

of accuracy. Indeed, evaluation using independent test areas shows an overall accuracy of over 98%, guaranteeing high-quality use when crossed with other maps.

In terms of the geographical distribution of land covers in Catalonia, shrublands (22.4% of the surface), needleleaf forests (19.8%), and rainfed herbaceous crops (13.7%) stand out since they cover a far larger area than the rest of the land covers.

All the five-year maps of the MUCSC series, from 1987 to 2017, can be downloaded at the aforementioned website.

Acknowledgements

This study was funded by the Generalitat (Grumets SGR2014-1491) and by the Spanish MCIU through the NEWFORLAND project (RTI2018-099397-B-C21 MCIU/AEI/ERDF, EU). X. Pons is the recipient of an ICREA Academia 2016 -2020 Excellence in Research Grant.

References

Cea, C., Cristóbal, J., Pons, X. 2007. An improved methodology to map snow cover by means of Landsat and MODIS imagery. En: *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007*. Barcelona. *IEEE International*, 4217-4220. DOI <https://doi.org/10.1109/IGARSS.2007.4423781>

Elvidge, C.D., Sutton, P.C., Wagner, T.W., Ryznar, R., Goetz, S.J., Smith, A.J., Jantz, C., Seto, K., Imhoff, M.L., Vogelmann, J., wang, Y.Q., Milesi, C., Nemani, R. 2004. Urbanization. En: Gutman, G. et al. (ed.). *Land change science: Observing, monitoring, and understanding trajectories of change on the earth's surface*. Dordrecht, Países Bajos: Kluwer Academic Publishers, pp. 315-328. https://doi.org/10.1007/978-1-4020-2562-4_18

ESA. 2015. Sentinel-2 User Handbook. Recuperado de https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook Último acceso: 4 de febrero, 2020.

González-Guerrero, O., Pons-Fernández, X., Bassols-Morey, R., Camps-Fernandez, F.J. 2019. Dinàmica de les superfícies de conreu a Catalunya mitjançant Teledetecció en el període 1987-2012. *Quaderns Agraris*, 46, 59-91.

Hansen, M.C., Loveland, T.R. 2012. A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment*, 122, 66-74. <https://doi.org/10.1016/j.rse.2011.08.024>

ICC. 1992. Mapa d'usos del sòl de Catalunya. Institut Cartogràfic de Catalunya. Barcelona. 118 p.

ICGC (2017). *Datos lidar*. Institut Cartogràfic i Geològic de Catalunya. Recuperado de <https://www.icgc.cat/es/Descargas/Elevaciones/Datos-lidar> Último acceso: 1 de mayo, 2020.

Loveland, T.R., Dwyer, J.L. 2012. Landsat: Building a strong future. *Remote Sensing of Environment*, 122, 22-29. <https://doi.org/10.1016/j.rse.2011.09.022>

Moré, G., Pons, X. 2007. Influencia del número de imágenes en la calidad de la cartografía detallada de vegetación forestal. *Revista de Teledetecció*, 28, 61-68. Recuperado de http://www.aet.org.es/revistas/revista28/7-111_More_revisado.pdf Último acceso: 1 de mayo, 2020.

Padial, M., Vidal-Macua, J.J., Serra, P., Ninyerola, M., Pons, X. 2019. Aplicación de filtros multicriterio basados en NDVI para la extracción de áreas de entrenamiento desde la base de datos SIOSE. Ruiz L.A., Estornell J., Calle A., Antuña-Sánchez J.C. (eds) *Teledetecció: hacia una visión global del cambio climático*, pp. 311-314. ISBN: 978-84-1320-038-5. Libro de actas XVIII Congreso de la Asociación Española de Teledetecció, 24 - 27 Septiembre, Valladolid (Spain).

Padró, J.C., Pons, X., Aragonés, D., Díaz-Delgado, R., García, D., Bustamante, J., Pesquer, L., Domingo-Marimon, C., González-Guerrero, O., Cristóbal, J., Doktor, D., Lange, M. 2017. Radiometric Correction of Simultaneously Acquired Landsat-7/Landsat-8 and Sentinel-2A Imagery Using Pseudoinvariant Areas (PIA): Contributing to the Landsat Time Series Legacy. *Remote Sensing*, 9(12), 1319. <https://doi.org/10.3390/rs9121319>

Padró, J.C., Muñoz, F.J., Ávila, L.A., Pesquer, L., Pons, X. 2018. Radiometric Correction of Landsat-8 and Sentinel-2A Scenes Using Drone Imagery in Synergy with Field Spectroradiometry. *Remote Sensing*, 10(11), 1687. <https://doi.org/10.3390/rs10111687>

Pons, X. 2004. MiraMon. Sistema de Información Geográfica y software de Teledetecció. Centre de Recerca Ecològica i Aplicacions Forestals, CREA. Bellaterra. ISBN: 84-931323-4-9. Recuperado de http://www.mirammon.cat/Index_es.htm Último acceso: 1 de mayo, 2020.

- Pons, X., Ninyerola, M. 2008. Mapping a topographic global solar radiation model implemented in a GIS and refined with ground data. *International Journal of Climatology*, 28(13), 1821-1834. <https://doi.org/10.1002/joc.1676>
- Pons, X., Pesquer, L., Cristóbal, J., González-Guerrero, O. 2014. Automatic and improved radiometric correction of Landsat imagery using reference values from MODIS surface reflectance images. *International Journal of Applied Earth Observation and Geoinformation*, 33, 243-254. <https://doi.org/10.1016/j.jag.2014.06.002>
- Pons X., Masó J. 2016. A comprehensive open package format for preservation and distribution of geospatial data and metadata. *Computers & Geosciences*, 97, 89-97. <https://doi.org/10.1016/j.cageo.2016.09.001>
- Townshend, J.R., Masek, J.G., Huang, C., Vermote, E.F., Gao, F., Channan, S., Sexton, J.O., Feng, M., Narasimhan, R., Kim, D., Song, K., Song, D., Song, X. P., Noojipady, P., Tan, B., Hansen, M.C., Li, M., Wolfe, R.E. 2012. Global characterization and monitoring of forest cover using Landsat data: Opportunities and challenges. *International Journal of Digital Earth*, 5(5), 373-397. <https://doi.org/10.1080/17538947.2012.713190>
- Woodcock, C.E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S.N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P.S., Vermote, E.F., Vogelmann, J., Wulder, M.A., Wynne, R. 2008. Free access to Landsat imagery. *Science*, 320(5879), 1011 <https://doi.org/10.1126/science.320.5879.1011a>
- Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., Woodcock, C.E. 2012. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, 122, 2-10. <https://doi.org/10.1016/j.rse.2012.01.010>

Annex 1

		Test areas																	Total	CE	UA							
		ILW	MRW	RF	HDU	IND	RHC	IHC	RFT	IFT	VIN	MOG	SHR	IMGL	LGL	SCL	DCD	NDL	WET	BAR	BEA	RIC	CTR					
	NODATA	0	0	0	0	0	2 862	11	37	0	58	0	0	254	0	0	47	137	0	48	0	1	0	3 456				
	ILW	41 352	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	0	13	0	0	0	0	41 378	0.001	100.00		
	MRW	0	361 096	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	65	0	0	0	361 161	0.000	100.00	
	RF	0	0	43 905	68	0	0	0	0	0	0	0	0	0	0	0	7	0	0	145	0	0	0	0	44 125	0.005	100.00	
	HDU	0	0	0	104 602	0	26	1	0	0	0	0	0	0	0	0	0	0	0	252	1	0	0	0	104 881	0.008	100.00	
	IND	0	0	0	118 303	0	410	132	31	14	0	0	0	854	1 314	0	100	553	0	8	0	857	0	0	0	119 756	0.012	99.99
	RHC	0	0	0	0	0	1 153 882	17 669	0	0	0	0	0	2 132	1 314	0	121	0	0	0	22	0	0	0	0	1 176 515	0.019	99.98
	IHC	0	0	0	0	0	2 604	387 565	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	390 312	0.007	99.99
	RFT	0	0	0	20	0	0	0	494 855	9	7	0	0	0	0	0	0	0	0	0	12	0	0	0	0	494 950	0.000	100.00
	IFT	0	0	0	0	0	0	0	226 373	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	226 402	0.000	100.00
	VIN	0	0	0	0	0	0	0	40	177 945	0	0	0	0	0	0	0	0	0	29	0	0	0	0	0	178 019	0.000	100.00
	MOG	0	0	0	0	0	0	0	0	0	238 891	0	0	203	0	0	7	0	0	0	17	0	0	0	0	239 117	0.001	100.00
	SHR	0	0	0	10	0	0	0	25	5	9	1 925 287	762	1 877	209	114	3 081	27	138	0	0	0	0	0	0	1 931 548	0.008	100.00
	IMGL	0	0	0	0	0	0	0	0	0	0	1 566	114 850	0	0	0	47	4	0	56	0	0	0	0	0	116 523	0.014	99.99
	LGL	0	0	0	0	0	0	0	0	0	0	427	0	0	45 931	0	0	0	0	22	0	0	0	0	0	46 379	0.012	99.99
	SCL	0	0	0	0	0	0	0	0	0	0	1 424	51	0	514 744	0	8 778	0	0	0	0	0	0	0	0	524 997	0.020	99.98
	DCD	0	0	0	0	0	0	0	0	0	61	16 086	406	0	186	738 073	5 115	5	2	0	0	0	0	0	0	759 935	0.029	99.97
	NDL	0	0	0	0	0	0	0	0	0	0	1 139	0	0	6 459	174	1 604 208	0	0	0	0	0	0	0	0	1 611 980	0.005	100.00
	WET	0	0	0	0	0	0	0	0	0	0	0	0	0	0	67	0	12 155	0	0	0	0	0	0	12 222	0.005	99.99	
	BAR	4	0	1 462	20	2 089	0	0	0	0	0	0	0	94	6	0	6	8	0	146 080	33	0	0	0	0	149 745	0.025	99.98
	BEA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8 639	0.000	100.00	
	RIC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	58 899	0.000	100.00	
	CTR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	24 138	0.000	100.00	
	Total	41 356	361 096	45 406	104 690	120 392	1 159 784	405 378	494 948	226 442	178 010	238 960	1 946 782	118 658	49 262	521 604	738 797	1 621 897	12 206	147 680	8 735	58 901	24 138	0	8 625 072			
	OE	0.0	0.0	3.3	0.1	1.7	0.5	4.4	0.0	0.0	0.0	1.1	3.2	6.8	1.3	0.1	1.1	0.4	1.1	1.1	1.1	1.1	0.0	0.0	0.0			
	PA	100.0	100.0	96.7	99.9	98.3	99.5	95.6	100.0	100.0	100.0	98.9	96.8	93.2	98.7	99.9	98.9	99.6	98.9	98.9	98.9	100.0	100.0	100.0	100.0			

Confusion matrix between the different legend categories. The table does not include the categories “Permanent snowpacks”, “Low-density urban areas” or “Burned areas” as these categories were edited later from data external to the classification process. ILW: Inland waters; MRW: Marine waters; RIF: Road infrastructure; HDU: High-density urban areas; IND: Industrial and commercial areas; RHC: Rainfed herbaceous crops; IHC: Irrigated herbaceous crops; RFT: Rainfed fruit trees; IFT: Irrigated fruit trees; VIN: Vineyards; MOG: Montane grasslands; SHR: Shrublands; MGL: Middle-altitude grasslands; LGL: Lowland grasslands; SCL: Sclerophyll forests; DCD: Deciduous forests; NED: Needleleaf forests; WET: Wetland vegetation; BAR: Areas with little or no vegetation; BEA: Sandy areas and beaches; RIC: Rice fields; CTR: Citrus plantations.