PAPER

Title: Degree of urbanization in Catalonia: administrative divisions versus high-resolution grids

Authors and e-mails of them: Clara Hormigos Feliu, chormigos@idescat.cat; Eduard Suñé, esl@idescat.cat; Daniel Ibáñez, dibanez@idescat.cat; Mireia Farré, mfarre@idescat.cat

Department: Department of Population and Territory, General Sub-Directorate of Production and Coordination

University: (Institution) Statistical Institute of Catalonia (Idescat)

Subject area: Methods for territorial analysis.

Abstract:

This paper studies the determination of the degree of urbanization based on the distribution of population density throughout the territory of Catalonia. Using grids with sizes of 1km, 500m and 250m, derived from the European standard 1km grid, we classify small administrative units as densely populated areas (cities), areas of intermediate density or thinly populated areas (rural areas). We find that both at the level of grid cell classification and when assigning a degree of urbanization to the municipalities of Catalonia grid size plays a major role, leading to different results as we vary grid size. We discuss as well the density cuts used in the calculation of the degree of urbanization, finding no evidence that they could be derived from data.

Keywords: Degree of urbanization; high-resolution grid; urban-rural continuum

JEL codes: O18; R14; R52; Y10

1. Introduction

The collection of statistics for urban and rural areas can represent an important tool to inform policy decisions. For instance, UN guidelines indicate that several indicators for sustainable development goals should be computed for rural and urban areas (UN, 2015), and urban-rural distinctions have already been used in studies carried out by Idescat and other entities in Catalonia (Generalitat de Catalunya, 2016). In order to improve the quality of such statistics, accurate and updated techniques to describe the urban-rural distribution of a territory constitute an important asset. To this end, the European Commission together with five other international organizations have put forward a methodology to compute the *degree of urbanization* (DEGURBA) of any given territory (Eurostat, 2021). Taking as a starting point a 1km population grid, the methodology allows to classify areas within a territory in one of three distinct classes (cities, towns and semi-dense areas and rural areas), which can then be used to produce harmonized statistics with international comparability. Using 2011 census data, Eurostat has used this methodology to determine the degree of urbanization of Local Administrative Units (LAUs) across Europe, and a major update is expected with the 2020 census (Eurostat, 2018; Eurostat, n.d.). In the case of Catalonia, the LAUs studied by Eurostat correspond to municipalities. In this paper, we implement the calculation of the degree of urbanization of the municipalities of Catalonia with a twofold motivation: firstly, performing the calculation independently at Idescat would allow us to periodically (for instance, annually) update the urban-rural classification of the territory. Additionally, it would allow us to determine the degree of urbanization of territorial units below the municipal level, such as census tracts, which is currently not provided by Eurostat.

When determining the degree of urbanization of small areas, grids with a resolution higher than the 1km European standard might improve the description of the urban-rural spectrum. Furthermore, higher-resolution and multi-resolution grids can also be used to provide local statistics with high spatial accuracy (Idescat, 2018). For these reasons, we are especially interested in studying the effects of grid size in computing the degree of urbanization, and whether the methodology put forward by Eurostat can be directly applied to higher-resolution grids. To this end, we implement the computation with grids of three different sizes (1km, 500m and 250m) and study the effects of grid size in the classification of the municipalities of Catalonia. Moreover, we want to explore as well whether the density cuts used in identifying urban, intermediate and rural areas can be derived from data. These density thresholds are to date input parameters of the classification methods and can vary significantly depending on the employed criteria. For instance, the OCDE (2011) defines rural areas to have densities below 150 inhab/km^2, while Eurostat (2021) sets this bound at 300 inhab/km^2. If population density distributions naturally contain density clusters, these density thresholds could be set by data.

Our paper is structured as follows: in Sec. 2 we describe the methodology used to compute the degree of urbanization of the municipalities of Catalonia using the standard European 1km grid, as well as our approach to apply the methodology to higher-resolution grids. Results are presented in Sec. 3, where we show the degree of urbanization at the stage of cell and municipality classification for the population of Catalonia and for three grid sizes: 1km, 500m and 250m. In Sec. 4 we examine how natural cuts in cell density can be derived from our data and compare them to the thresholds used in the determination of the degree of urbanization, and our conclusions and outlook are discussed in Sec. 5.

2. Computing the degree of urbanization

In order to determine the degree of urbanization of the municipalities of Catalonia we employ the methodology described by Eurostat (2021). It aims to provide a standard procedure to establish the degree of urbanization of small spatial units, allowing international comparisons. The methodology follows two steps: first, a 1km population grid covering the target areas is set up, and each cell is classified in one of three categories (urban center, urban cluster or rural area) depending on its population density and its location. After grid cells have been classified, each small spatial unit is assigned one of three possible degrees of urbanization (cities or densely populated areas, towns and semi-dense areas or intermediate density areas, and rural or thinly populated areas), depending on how its population is distributed and the cell types that cover its area. This classification of small spatial units corresponds to the level 1 classification described by Eurostat (2021), and subsequent classification levels can also be computed. In this paper, we focus on the level 1 classification of the municipalities in Catalonia. In what follows, we summarize the main steps needed to carry out the classification.

2.1 Grid cell classification

Grid cell classification hinges on the identification of urban centers and urban clusters. These consist of groups of contiguous grid cells where each cell falls within a given range of population density, and with the total cluster population being above certain thresholds. Each cell is assigned one category only, according to the following criteria:

- **Urban centers**: groups of contiguous cells (excluding diagonal contiguity) with cell population density higher than 1,500 inhab/km^2, and with a total aggregate cluster population of 50,000 inhabitants or more. Once an urban center is identified, its borders are smoothed according to the *iterative majority rule*, explained below.

- **Urban clusters**: groups of contiguous cells (including diagonal contiguity) with cell population density higher than 300 inhab/km^2, and with a total aggregate cluster population of 5,000 inhabitants or more. After urban clusters are identified, if cells from an urban cluster were previously found to belong to an urban center they remain urban center cells.

- **Rural areas**: grid cells which do not belong to urban centers or urban clusters. Note that these may have any population density, although it is usually below 300 inhab/km^2.

After identifying all urban centers the iterative majority rule is used to smooth their borders. The rule assigns urban center status to cells which are contiguous, including diagonal contiguity, with at least 5 cells that belong to the same urban center. The process of identifying such cells is repeated until no more cells are added to an urban center (see Eurostat, 2021, for details).

The delimitation of urban centers may be further adjusted by including cells that, even if they do not reach a population density of 1,500 inhab/km^2, are at least 50% built up. This allows for the addition to urban centers of areas that have low population density but are intensely used during daytime, such as railways, commercial or industrial areas. The inclusion of built-up areas is not considered in this paper and will be addressed in the future, as heavily industrial areas in Catalonia may be included in urban centers as a result.

However, we do not expect these corrections to impact our classification of municipalities, which hinges on the distribution of residential population.

2.2 Classification of small spatial units

After having classified each grid cell, small spatial units may be assigned one of three categories describing their level 1 degree of urbanization. Following Eurostat (2021) again, these are:

- **Cities or densely populated areas**: small spatial units where at least 50% of the population resides in urban center cells.

- **Towns and semi-dense areas or intermediate density areas**: small spatial units where less than 50% of the population resides in urban center cells and less than 50% of the population resides in rural cells.

- **Rural or thinly populated areas**: small spatial units where at least 50% of the population resides in rural cells.

Note that, in order to classify small spatial units, aggregate grid cell population and density are not sufficient, but geolocalized microdata is necessary in order to identify the percentages of population which reside in each cell type. The classification of all small spatial units captures the urban-rural continuum of the studied area.

2.3 Grid size effects

The standard determination of the degree of urbanization presented by Eurostat (2021) and summarized in the previous sections contains several parameters which must be fixed a priori. These comprise grid size (set at 1km), urban center cell density and total population (1,500 inhab/km^2 and 50,000 inhabitants, respectively), and urban cluster cell density and total population (300 inhab/km^2 and 5,000 inhabitants, respectively). In order to study the effects of changing grid sizes on the classification of small spatial units we compute, in addition to the standard European 1km grid, population grids with cell size 500m and 250m for the 2018 population of Catalonia. We then perform the classification of grid cells as described in Sec. 2a for each of the three cell sizes, and then compute the classification of municipalities based on the different outputs according to Sec. 2b. The results are discussed in Sec. 3, where we compare the outcomes obtained for each grid size together with Eurostat's DEGURBA determination based on the 2011 census (Eurostat, n.d.).

3. Degree of urbanization in Catalonia: comparing grid sizes

Using the methodology described in Sec. 2, we now determine the degree of urbanization of the municipalities of Catalonia. We use 1km, 500m and 250m population grids compatible with the European standard grid (European Environmental Agency, 2017), together with the 2018 population of Catalonia from the Statistical Register of the Population (*Registre Estadístic de Població* or REP) and the Statistical Register of the Territory (*Registre Estadístic de Territori* or RET) maintained by Idescat (Suñé et al., 2019, 2020), which contain geolocalized entries. We first classify grid cells and use these results to assign a degree of urbanization to the municipalities, using 2021 geometries provided by Institut Cartogràfic i Geològic de Catalunya.

3.1 Grid cell classification in Catalonia

The results for grid cell classification for 1km, 500m and 250m grids are shown in Figure 1 - Figure 3 for the Barcelona metropolitan area and its surroundings. In general, we find that our implementation of the cell classification correctly identifies urban centers and clusters in Catalonia.

Already at the level of cell classification we observe significant differences between the results obtained with each grid size, particularly in the identification of urban centers and urban clusters. Table 1 shows the number of urban centers and urban clusters that we identify in each case. We observe that the number of urban centers found is very similar for all grid sizes (18-19 urban centers). However, the identified urban centers are not the same for the 1km, 500m and 250m grids. In fact, we find that as grid resolution increases, some urban centers are fragmented or become smaller, and others are no longer identified. For instance, in the Maresme area (norhteast of Barcelona, between Barcelona and Mataró) we see that, using the 1km grid, the Barcelona urban center stretches norhteast until El Masnou, and another urban cluster is found containing the towns of Premià de Mar, Vilassar de Mar, Premià de Dalt, Vilassar de Dalt and Cabrils. With both higher-resolution grids, the Barcelona cluster extends northeast only until Badalona, while no more urban clusters are found between Barcelona and Mataró. Furthermore, we find that the number of urban clusters in Catalonia increases with grid resolution, as can be seen in Table 1. This is partly due to the fact that the area covered by urban centers decreases with grid size, and populated areas which were identified as urban centers with the 1km grid change their classification to urban clusters when using a higher resolution. Moreover, urban clusters are also fragmented as grid size decreases, which increments their number.

Altogether, we find that grid size has a significant effect in the classification of the urban-rural continuum at the cell level. In the next section, we explore how these diffrences affect the DEGURBA classification of small spatial units.

	1km grid	**500m grid**	**250m grid**
Urban centers	19	18	19
Urban clusters	100	132	154

Table 1. Number of urban centers and urban clusters found in Catalonia using 1km, 500m and 250m grids.

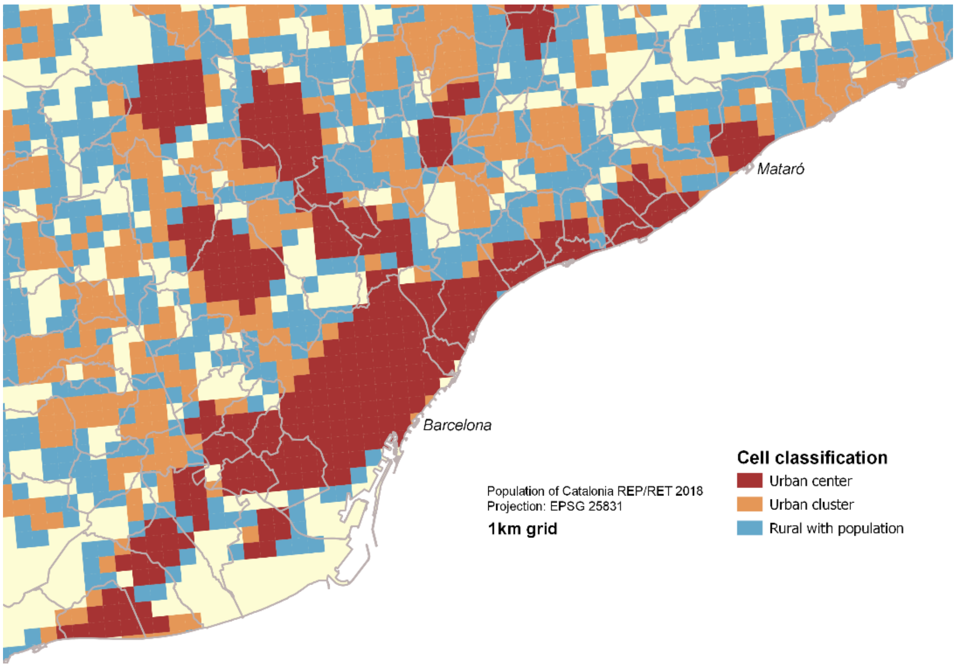

Figure 1. 1km grid cell classification in the Barcelona area. Yellow areas are unpopulated and have rural cell classification. Municipality boundaries are also shown. The Barcelona urban center extends northeast until El Masnou, and another urban cluster is found between Barcelona and Mataró.

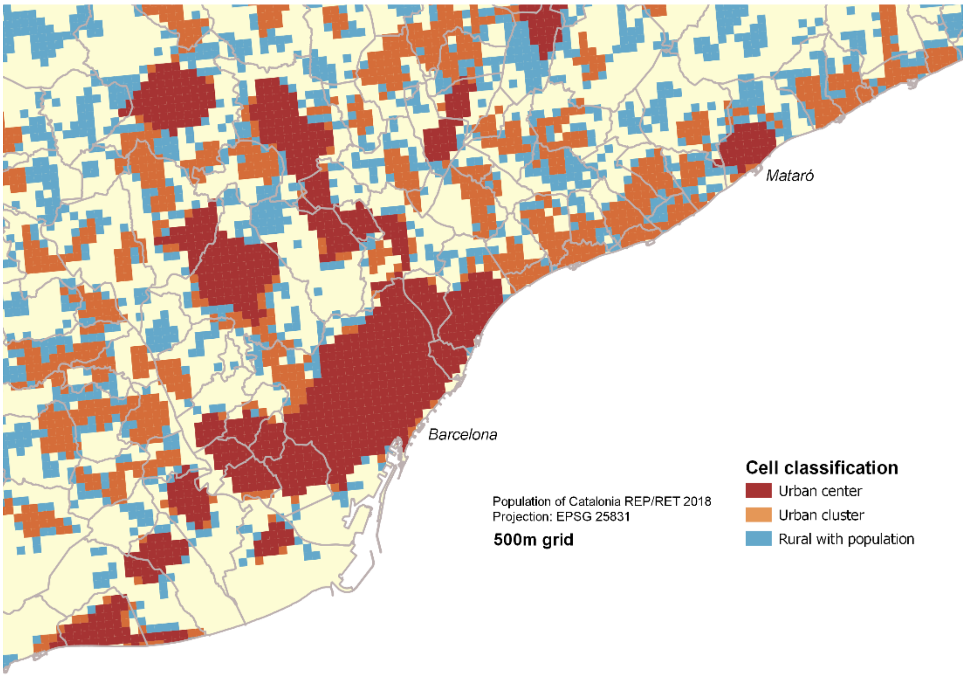

Figure 2. 500m grid cell classification in the Barcelona area. Yellow areas are unpopulated and have rural cell classification. Municipality boundaries are also shown. The Barcelona urban center extends northeast until Badalona, and no other urban centers are found along the coastline until Mataró.

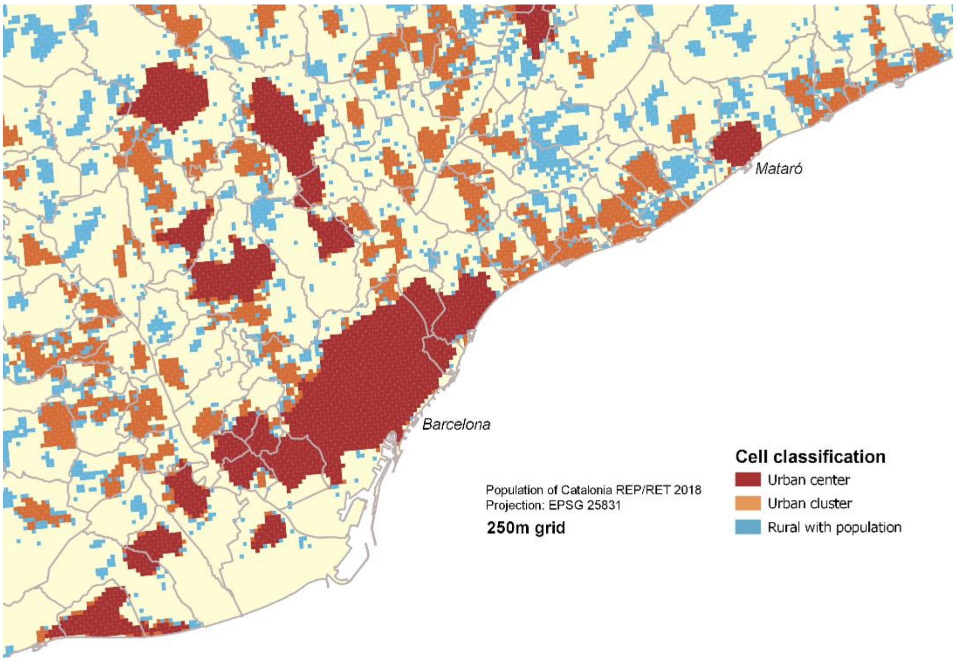

Figure 3. 250m grid cell classification in the Barcelona area. Yellow areas are unpopulated and have rural cell classification. Municipality boundaries are also shown. Comments from Figure 2 on the distribution of urban centers apply.

3.2 Classification of municipalities

After having classified the totality of the territory of Catalonia at grid cell level using three different grid sizes, we now determine the degree of urbanization of its municipalities. The results are shown in Figure 4, 5 and 6 for results based on the 1km, 500m and 250m grids, respectively. As expected from our results of Sec. 3a, we observe differences in the classification of municipalities for each base grid. These can be seen in Table 2, where we show the number of high-, intermediate- and low-density municipalities we identify with each grid size together with Eurostat's classification, which is based on the 2011 census and 2016 municipality geometries.

Firstly, we observe that using the standard European 1km grid our results already differ from Eurostat's, with our output showing more municipalities that are classified as densely populated or intermediate-density areas. These differences can be attributed to the increase of population between 2011 and 2018, and to a lower degree to the different statistical sources used to construct the population grids (census vs. REP-RET registers).

Moreover, we find that our 1km grid results correctly classify all high-density municipalities which were previously identified by Eurostat.

When comparing results for different base grid sizes, we find a clear tendency showing a decrease in the number of municipalities classified as densely populated and intermediate-density areas with decreasing grid size, as can be seen in Table 2. Taking again the example of the Maresme region, which showed already large variability at the cell classification stage, we see that the number of high-density municipalities in the region decreases dramatically from 11 to 1 as we move from the 1km grid to the higher-resolution ones, as seen in Table 3. Therefore, we see a clear impact of grid size not only at the stage of cell classification but also when assigning a degree of urbanization to small spatial units.

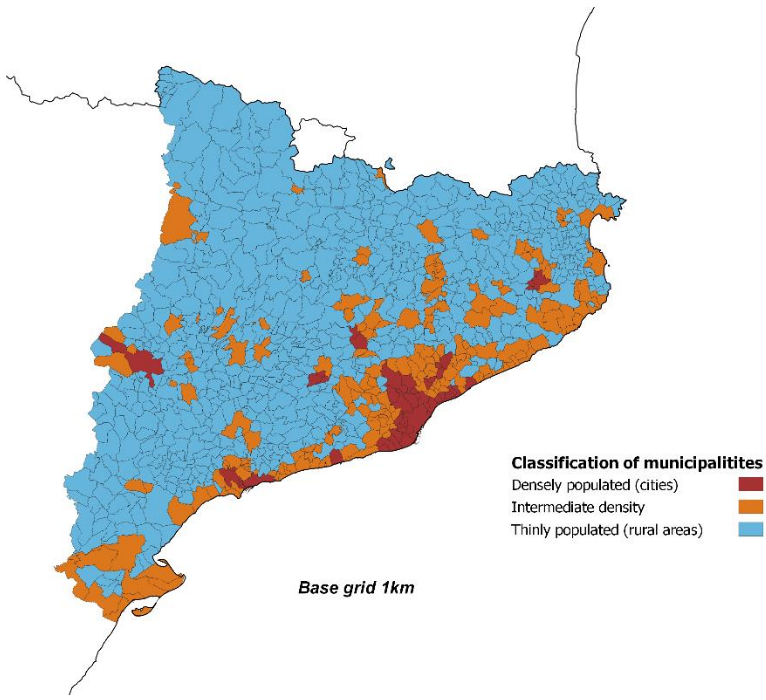

Figure 4. Classification of municipalities in Catalonia based on the 1km grid cell classification. Projection: EPSG 25831

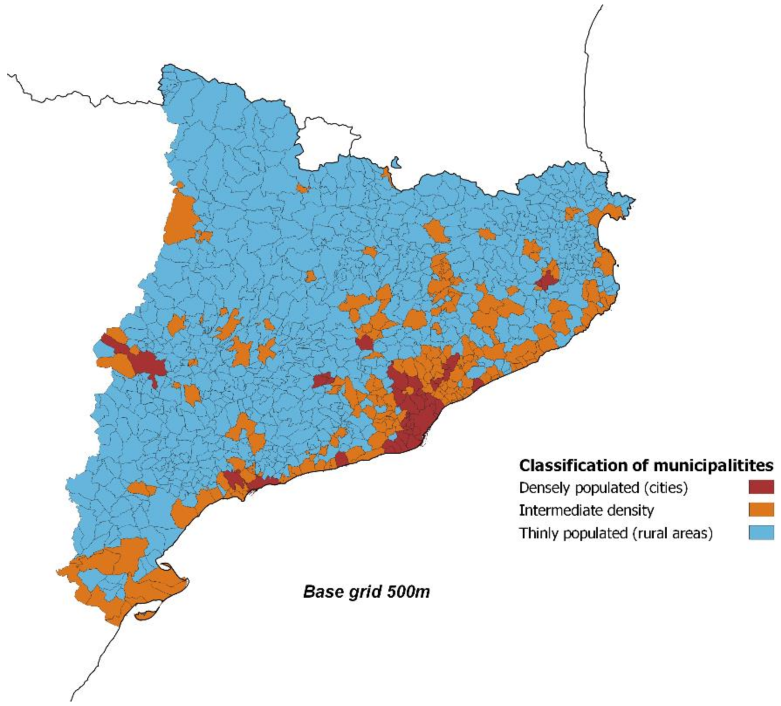

Figure 5. Classification of municipalities in Catalonia based on the 500m grid cell classification. Projection: EPSG 25831.

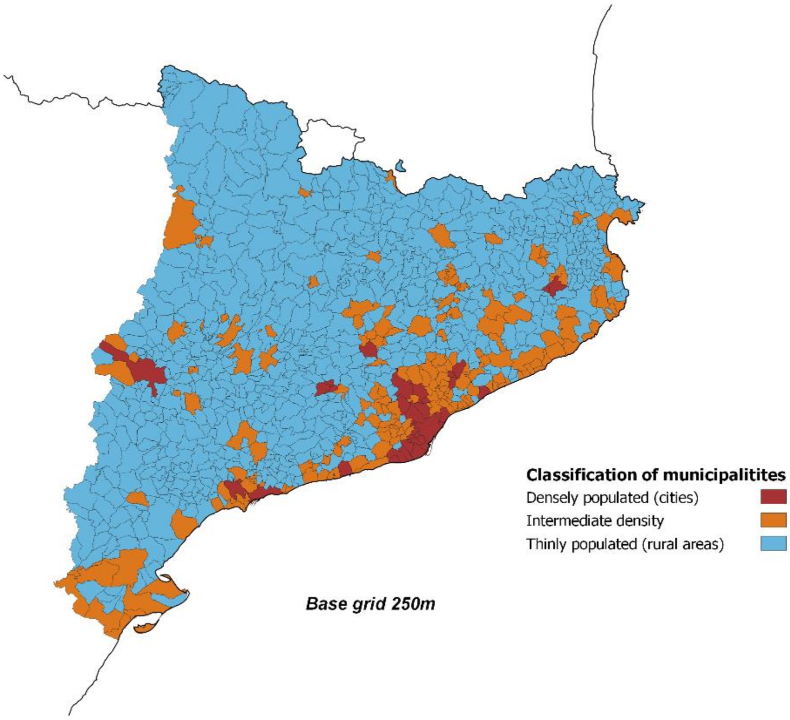

Figure 6. Classification of municipalities in Catalonia based on the 250m grid cell classification. Projection: EPSG 25831

DEGURBA	Eurostat	1km grid	500m grid	250m grid
Densely populated (cities)	42	55	40	35
Intermediate density	198	201	196	183
Thinly populated (rural areas)	707	691	711	729

Table 2. Number of municipalities in Catalonia classified as densely populated, intermediate-density and thinly populated areas, according to Eurostat (based on a 1km grid and the 2011 census) and our classification, based on 1km, 500m and 250m grids and the 2018 population of the REP and RET registers maintained by IDESCAT.

DEGURBA Maresme	Eurostat	1km grid	500m grid	250m grid
Densely populated (cities)	6	11	1	1
Intermediate density	20	15	24	23
Thinly populated (rural areas)	4	4	5	6

Table 3. Number of municipalities in Maresme (Catalonia) classified as densely populated, intermediate-density and thinly populated areas, according to Eurostat and to our computation for different grid sizes (see Table 2 for details). The classification of municipalities in the Maresme region shows a strong dependence on grid size.

4. Density thresholds

After studying grid size effects in the determination of the degree of urbanization we now explore whether the density cuts used in the cell classification algorithm can be derived from data. We use the 1km grid cell densities of Catalonia to investigate whether natural cuts in the density can be found, using both the Jenks-Fisher algorithm and 1-dimensional K-means clustering. We use the full dataset of cell densities, which span the values 0 – 54,911.5 inhab/ km^2, and apply the algorithms for the case of 3 and 4 density groups or clusters. We find that both methods yield the same results, as shown in Table 4. For 3 density groups, the algorithms yield density thresholds around 6,000 and 23,000 inhab/km^2, while for 4 groups we find thresholds around 2,600, 11,000 and 27,000 inhab/km^2. These do not match the density cuts used for cell classification in the determination of the degree of urbanization (300 and 1,500 inhab/km^2, see Sec. 2a for details). On the other hand, we find that when increasing the number of density groups the thresholds do approach the values set by Eurostat (2021) for cell classification. For instance, when applying the K-means algorithm for 12 groups the first two density thresholds are found around 400 and 1,500 inhab/km^2. However, the K-means Elbow test shows that the optimal number of density clusters is indeed 3, and there seems to be little justification to employ 12 density groups. Therefore, we do not find evidence that the cuts used in Eurostat's calculation of the degree of urbanization stem naturally from density data in the case of Catalonia.

Cuts	**Eurostat**	**Data (3 groups)**	**Data (4 groups)**
Cut 1	300	5,929	2,630
Cut 2	1,500	23,193	10,781
Cut 3			26,715

Table 4. Density thresholds in inhab/km^2 set by Eurostat (2021) in the calculation of the degree of urbanization, and those obtained by applying Jenks-Fischer and Kmeans algorithms to the 1km cell density data of Catalonia. We show thresholds for 3 and 4 density groups, with both algorithms yielding the same results.

5. Conclusions and outlook

The determination of the degree of urbanization of small spatial units allows to describe the urban-rural continuum of a territory. Based on the methods of Eurostat (2021), we have implemented the calculation of the degree of urbanization for the municipalities of Catalonia, using the 2018 population of the Statistical Population and Territory Registers maintained by Idescat and the 2021 geometries of the municipalities. We have computed the degree of urbanization using the standard European 1km grid as well as two higher resolution grids, with cell size 500m and 250m. This allowed us to study the effect of cell size in the determination of the degree of urbanization.

The degree of urbanization is computed in two steps: first, grid cells are classified in 3 categories (urban centers, urban clusters and rural areas), and then these results are used to classify small spatial units in densely populated, intermediate density and thinly populated areas. Our results for cell classification show a large dependence on grid size. As seen in Figures 1 – 3 for the Barcelona area and its surroundings, as grid size decreases we see urban centers reducing their size, breaking up or not being identified altogether, while urban clusters break up as well and their total number increases. Using the results for cell classification with three different grid sizes, we then computed the degree of urbanization of the municipalities of Catalonia, as seen in Figures 4 – 6. We found that our classification based on a 1km grid (and 2018 population) varies from Eurostat's DEGURBA determination based on the 2011 census, a variation we attribute mainly to population increase. Furthermore, we observed that the differences we noted in the classification of different-sized grid cells lead to municipality classifications depending heavily on the base grid used. For instance, for the Maresme area, the number of municipalities classified as densely populated areas decrease from 11 to 1 when reducing grid size from 1km to 500m. Therefore, we conclude that the assignation of the degree of urbanization has a strong dependence on grid size.

On the other hand, we explored as well whether the density cuts used in the step of cell classification could be derived from data. In order to do this, we used the Jenks-Fisher and K-means algorithms to find 3 or 4 natural breaks in the density values of our 1km grid, and found that these do no not resemble the density cuts set by Eurostat (2021). Therefore, we find no evidence that the density thresholds used in the determination of the degree of urbanization can be backed up by data but are rather input parameters which could in principle be changed. Altogether, we find that the development of more robust, parameter-independent methods would be desirable in order to reach a stable definition of the degree of urbanization.

In order to further validate our results, we aim, in the future, to determine the degree of urbanization using the GHS software together with our population dataset (see Eurostat, 2021, and Florczyk, 2019, for details). Moreover, we plan to include a built-up layer to refine our determination of urban areas. This would allow us, for instance, to classify heavily industrial areas in Catalonia which have a small residential population as part of urban centers or clusters. Although we expect these corrections to have little impact on the classification of municipalities, they could be important in determining the degree of urbanization of infra-municipal spatial units such as census tracts, which we also intend to compute in the future, as it would allow the production of highly localized statistics in terms of the degree of urbanization. Furthermore, the implementation of the level 2 degree of the urbanization would also be an asset, as it would provide us with a finer distinction within each level 1 classification.

References

European Environment Agency (November 2017). *EEA reference grid.* https://www.eea.europa.eu/data-and-maps/data/eea-reference-grids-2

Eurostat, (2021). *Applying the degree of urbanisation : a methodological manual to define cities, towns and rural areas for international comparisons : 2021 edition*, Publications Office. https://data.europa.eu/doi/10.2785/706535

Eurostat (2018). *Degree of urbanisation.* https://ec.europa.eu/eurostat/web/degree-of-urbanisation/background

Eurostat (n.d.). *NUTS – Nomenclature of territorial units for statistics. Local administrative units.* https://ec.europa.eu/eurostat/web/nuts/local-administrative-units

Florczyk, A., Corbane, C., Ehrlich, D., Freire, S., Kempe, T., Maffenini, L., Melchiorri, M., Pesaresi, M., Politis, P., Schiavina, M., Sabo, F. & Zanchetta, L. (2019). *GHSL data package 2019 : public release GHS P2019*, Publications Office. https://data.europa.eu/doi/10.2760/0726

Generalitat de Catalunya, Departament d'Agricultura, Ramaderia, Pesca i Alimentació (2016). *Programa de Dones del Món Rural i Marítim 2016–2020.* http://agricultura.gencat.cat/web/.content/01-departament/politiques-dones/enllacos-documents/fitxers-binaris/Programa-de-dones-2016-2020.pdf

Idescat (2018). *Població de Catalunya georeferenciada a 1 de gener de 2016.* https://biblio.idescat.cat/publicacions/Record/21104

Institut Cartogràfic i Geològic de Catalunya (2021). *Divisions administratives*. https://www.icgc.cat/Descarregues/Cartografia-vectorial/Divisions-administratives

OECD (2011). *OECD Regional typology.* https://www.oecd.org/cfe/regionaldevelopment/OECD_regional_typology_Nov2012.pdf

Reig Martínez, E., Goerlich Gisbert, F. J., & Martí Cantarino, I. (2016). *Delimitación de áreas rurales y urbanas a nivel local: Demografía, coberturas del suelo y accesibilidad*, Fundación BBVA. https://www.fbbva.es/en/publicaciones/delimitacion-de-areas-rurales-y-urbanas-a-nivel-local-demografia-coberturas-del-suelo-y-accesibilidad-2/

Suñé, E., Bernaus, A., Ibáñez, D., Condal, R., Farré, M., & Roy, C. (2019, March 12 – 14). *Implementation of the statistical territory register* [Conference presentation]. NTTS 2019, Brussels, Belgium. https://ec.europa.eu/eurostat/cros/NTTS2019_en

Suñé, E., Ibáñez, D., Bernaus, A., & Farré, M. (2020). *El registro estadístico de territorio: finalidades y calidad de sus fuentes*, Indice 76, 9 – 11. http://www.revistaindice.com/numero76/p9.pdf

UN (2015). *Transforming our World: the 2030 Agenda for Sustainable Development A/RES/70/1*. United Nations. https://sdgs.un.org/sites/default/files/publications/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf