

MULTISCALE ADMINISTRATIVE REPRESENTATION OF POPULATION DENSITY AT SUB-MUNICIPAL LEVEL IN SPAIN, THE GALICIAN CASE

ALEXANDRE NOBAJAS¹ 

ABSTRACT – The disparity between total population values and how that population is spread across an area can generate issues of perception, as though choropleth maps are usually used to represent population, its density may not be equally spread across the territory. A prime example of this situation is Galicia in NW Spain, as it has the highest number of population settlements of the country with almost 50% of the total, and yet at the same time it has only around 4% of municipalities, which means that each local authority has on average dozens of borderless settlements within its limits. However, population maps of the region are normally produced at the municipal level, providing a poor representation of the demographic complexity of the area. Consequently, this paper provides an alternative way of displaying population density by using three sub-municipal division levels obtained from official sources: census tracts, parishes and settlements. While census tracts and parishes have defined limits, settlements do not, so they are normally mapped by using punctual symbology. It is therefore necessary to generate some borders for those population entities which do not have them, as it can help understand how the population is distributed in more detail. In order to provide a suitable means of comparison with the other administrative divisions this paper presents a method to convert settlement data into polygons that includes the use of Voronoi tessellation, producing for the first time the most detailed population density choropleth map of the region at settlement level. Furthermore, the method can be applied anywhere in Spain and with non-census data.

Keywords: Borderless settlements; population density; Galicia (Spain); Voronoi tessellation; modifiable areal unit problem.

RESUMO – REPRESENTAÇÃO MULTIESCALA DA DENSIDADE DA POPULAÇÃO AO NÍVEL SUBMUNICIPAL EM ESPANHA, O CASO DA GALIZA. A disparidade

Recebido: 06/04/2022. Aceite: 30/11/2022. Publicado: 30/12/2022.

¹ School of Geography, Geology and the Environment, Faculty of Natural Sciences, William Smith Building, Keele University, Staffordshire ST5 5BG, United Kingdom. E-mail: a.nobajas@keele.ac.uk

entre os valores populacionais totais e como essa população está distribuída numa área pode gerar problemas de percepção uma vez que embora os mapas coropletos sejam geralmente usados para representar a população, a sua densidade pode não estar igualmente distribuída por todo o território. Um bom exemplo desta situação é na Galiza, no noroeste da Espanha, por ter o maior número de assentamentos populacionais do país, com quase 50% do total, e ao mesmo tempo ter apenas cerca de 4% de municípios, o que significa que cada município possui, em média, dezenas de aglomerados sem fronteiras definidas dentro de seus limites. No entanto, os mapas populacionais da região são normalmente produzidos ao nível municipal, o que fornece uma representação pobre da complexidade demográfica da área. Consequentemente, este artigo fornece uma forma alternativa de representar a densidade populacional usando três níveis de divisão submunicipal obtidos de fontes oficiais: secções censitárias, freguesias e aglomerados. Embora as secções censitárias e freguesias tenham limites definidos, os aglomerados não os têm, por isso são geralmente mapeados por meio de simbologia de pontos. Portanto, é necessário gerar limites para os aglomerados populacionais que não os possuem, pois pode ajudar a entender mais detalhadamente como a população está distribuída. A fim de fornecer um meio adequado de comparação com as outras divisões administrativas, este artigo apresenta um método para conversão de pontos em polígonos que inclui o uso de Diagrama de Voronoi, produzindo pela primeira vez, para esta região, o mapa coropleto de densidade populacional mais detalhado, à escala do aglomerado populacional. Da mesma forma, este método é aplicável ao resto da Espanha e em dados não censitários.

Palavras-chave: Aglomerados populacionais sem limites definidos; densidade populacional; Galiza (Espanha); Diagrama de Voronoi; problema de agregação espacial.

RÉSUMÉ – REPRÉSENTATION MULTI-ÉCHELLE DE LA DENSITÉ DE LA POPULATION AU NIVEAU SOUSMUNICIPAL EN ESPAGNE, LE CAS DE LA GALICE. La disparité entre les valeurs de la population totale et la distribution de cette population dans un territoire peut générer des problèmes de perception, car bien que les cartes choroplèthes soient généralement utilisées pour représenter la population, leur densité peut ne pas être également répartie sur tout le territoire. Un bon exemple de cette situation est la Galice, dans le nord-ouest de l'Espagne, car elle a le plus grand nombre d'agglomérations humaines du pays avec près de 50% du total, et en même temps, elle ne compte qu'environ 4% de municipalités, ce qui signifie que chaque autorité locale compte en moyenne des dizaines de localités sans frontières définies à l'intérieur de ses frontières. Cependant, les cartes de la population de la région sont généralement produites au niveau municipal, ce qui fournit une mauvaise représentation de la complexité démographique de la région. Par conséquent, cet article propose une autre façon de représenter la densité de population en utilisant trois niveaux de division infra-municipale obtenus de sources officielles : les sections de recensement, les paroisses et les agglomérations urbaines. Bien que les sections du recensement et les paroisses aient des limites définies, les agglomérations ne l'ont pas, de sorte qu'elles sont généralement cartographiées à l'aide d'une symbologie ponctuelle. Par conséquent, il est nécessaire de générer des limites pour les agglomérations de population qui ne les ont pas, car cela peut aider à comprendre plus en détail comment la population est répartie. Afin de fournir un moyen approprié de comparaison avec les autres divisions administratives, ce document présente une méthode de conversion des données de points en polygones qui comprend

l'utilisation de la Diagramme de Voronoï, produisant pour la première fois la carte de densité choroplèthe de population plus détaillée de la région au niveau de l'agglomération. De même, cette méthode est applicable au reste de l'Espagne et pour des données non censitaires.

Mot clés: Entités de population sans limites définies; densité de la population; Galice (Espagne); Diagramme de Voronoï; problème d'agrégation spatiale.

RESUMEN – REPRESENTACIÓN MULTIESCALA DE LA DENSIDAD DE POBLACIÓN A NIVEL SUBMUNICIPAL EN ESPAÑA, EL CASO DE GALICIA. La disparidad entre los valores totales de población y cómo ésta se distribuye sobre un área puede generar problemas de percepción, ya que aunque los mapas de coropletas se utilizan generalmente para representar la población, su densidad puede no estar igualmente distribuida en todo el territorio. Un buen ejemplo de esta situación es Galicia en el noroeste de España, que tiene el mayor número de asentamientos de población del país con casi el 50% del total, y al mismo tiempo tiene solo alrededor del 4% de municipios, lo que significa que cada autoridad local tiene en promedio decenas de asentamientos sin fronteras definidas dentro de sus límites. Sin embargo, los mapas de población de la región se producen normalmente a nivel municipal, lo que proporciona una representación deficiente de la complejidad demográfica del área. En consecuencia, este artículo proporciona una forma alternativa de representar la densidad de población mediante el uso de tres niveles de división submunicipal obtenidos de fuentes oficiales: secciones censales, parroquias y asentamientos (entidades de población). Si bien las secciones censales y las parroquias tienen límites definidos, los asentamientos no los tienen, por lo que normalmente se cartografían utilizando simbología puntual. Por tanto, es necesario generar unos límites para aquellas entidades de población que no las tengan, ya que puede ayudar a entender con más detalle cómo se distribuye la población. Con el fin de proporcionar un medio adecuado de comparación con las otras divisiones administrativas, este documento presenta un método para convertir los datos de asentamiento en polígonos que incluye el uso del Diagrama de Voronoï, produciendo por primera vez el mapa de coropletas de densidad de población más detallado de la región a nivel entidad de población. Asimismo, este método es aplicable al resto de España y en datos no censuales.

Palabras clave: Entidades de población sin límites definidos; densidad de población; Galicia (España); Diagrama de Voronoï; Problema de agregación espacial.

I. INTRODUCTION AND BACKGROUND

Population maps, which represent how the inhabitants of a given territory are distributed, are one of the most commonly produced thematic maps (Wiegand, 2003). Traditionally, geographic population distribution is displayed using a choropleth map, which uses shaded colours or patterns to represent the value of a statistical variable and helps with the geographical interpretation of otherwise complex datasets (Raisz, 1962). However, choropleth maps can be misleading as the value attributed to each polygon is across the entire area of each mapping unit, which is seldom the case, and changes only occur at polygon boundaries in an abrupt manner (Coulson, 1987). Consequently, results are

highly dependent on how polygons are designed and on the selected symbolism, something which can lead to deceptive impressions of how the population is distributed (Openshaw & Taylor, 1981).

In order to minimise this issue, cartographers have come up with several solutions. For example, one way of fine tuning where population is located is the use of population density instead of the absolute number of inhabitants, as by doing this the reader can get a better grasp of how the population is distributed across the polygon (Dixon, 1972). Additionally, a variant of population density maps is the use of dot maps, where dots representing a certain amount of people are placed at random within each polygon providing an easy way of visualising population concentrations (Lavin, 1986). Other possible alternatives are dasymetric maps, where instead of representing population across an area regardless of where housing is located, land use is taken into account in order to represent where people actually live, avoiding in this way fictitious situations such as people living on lakes or glaciers (Eicher & Brewer, 2001; Holt & Lu, 2011; Martins *et al.*, 2012; Petrov, 2012).

Another way of representing population distribution that has gained popularity in the last decade is by creating population grids (Huang *et al.*, 2019). Population grids do not represent population using a vectorial model of the world constrained by administrative limits, but rather use a continuous or raster model where population changes through space in a smoother manner (Goerlich & Cantarino, 2013). In order to create them, land use data extracted from satellite or aerial imagery is combined with population data, and this way the modifiable aerial unit problem (Openshaw & Taylor, 1981) is avoided (Batista & Silva *et al.*, 2013). It is therefore not surprising that some multinational institutions such as the European Union have embraced gridded population maps, as they are a way of smoothing out administrative differences between regions and countries (Eurostat, 2021). The Instituto Nacional de Estadística (INE), also embraced the gridded approach with the 2011 census, as they generated a 1km² grid for Spain which includes data for all cells where at least one housing unit is present. This gridded dataset was used by Eurostat to create the aforementioned European gridded population map (Instituto Nacional de Estadística [INE], 2011).

However, perhaps the most straight forward way of improving the resolution of population maps is using a standard choropleth map but augmenting the level of detail – that is reducing the polygon size. Using the USA as an example, one can represent population at the national, state, county or local level, which will provide an increasing level of detail while changing how the map is interpreted by the reader. Nevertheless, obtaining ever-increasing detail when it comes to population figures is not equally possible across the world, as while some countries produce population counts at different resolutions down to neighbourhood level (Campanera *et al.*, 2014), others do not offer such detailed information, making the production of detailed population maps a challenging task.

In the case of Spain, detailed population figures have traditionally been given at the municipal – local – level (Pampalon *et al.*, 2009), which for many uses and applications is too coarse, as municipalities can be larger than 1500km² (Gajardo *et al.*, 2014) and can include hundreds of unevenly sized settlements. This means that in the picture given

when using a municipal level of scale less populated areas may be masked by larger settlements within the same local authority, conveying misleading information. There are, however, alternatives which enable the representation of the population, whether absolute or density, at sub-municipal level. This article uses the case of Galicia to present a comparison between four different levels of detail which enable the representation of population density at local authority level, census tract and traditional parish level. It starts by explaining how population data is provided in Spain and Galicia, it continues by developing a new method to represent population density at settlement level as a choropleth map, and represents and analyses the generated cartographic outputs, the most detailed population density map available for Galicia.

II. POPULATION DATA IN SPAIN AND GALICIA

The reason why the most detailed population maps of Spain have traditionally been at the municipal level is because population counts are carried out by local authorities – using a register known as *padrón* – and once a year they send their population data to the Instituto Nacional de Estadística (INE) (Jurado, 2004). Once the INE has received data from all municipalities, it collates it and publishes the latest population figures for Spain (Goerlich, 2007). Consequently, the only annual population count available in Spain is at municipal level and not at greater levels of detail. In the year 2015 Spain was divided into 8125 municipalities (Cabezuelo-Lorenzo *et al.*, 2016), which means that the average local authority size was of around 62km² and had an average population of just over 11 000 people. However, due to historical, political and demographic reasons there is a very large variability in terms of size, population and their distribution across Spain (Rivero & Lladós, 2014). This disparity means that the cartographic representation of population distribution can often be misleading, as very different types of municipalities are all mapped using the same method.

The issue is the same across the country, but the case of Galicia, a historic nationality which currently is one of the seventeen Autonomous Communities of Spain (Rodríguez, 2016), is particularly relevant. It has the highest density of inhabited settlements in Spain (Barbosa-Brandão *et al.*, 2015) with almost half of Spain's total (Goerlich & Cantarino, 2011) in an area which represents less than 6% of the country's surface and population. The Galician population is therefore very dispersed when compared to other areas of Spain, but administratively in the year 2015 it only had 314 municipalities – locally known as *concellos* – (table I) which represent less than 4% of Spain's total. Therefore, when mapping population in Galicia at a municipal level the picture given can be misleading, as less populated areas may be masked by large settlements within the same local authority.

Up until recently it was difficult to easily obtain more detailed population data, but with the progressive popularisation of GIS software and the drive towards opening and facilitating official data to the public, sub-municipal population data has become avail-

able, allowing alternatives to the traditional municipal level of detail. In the Galician case, three sub-municipal levels of detail can be used to represent population: census tracts, traditional parishes and settlements (fig. 1).

Table I – Descriptive statistics of the different administrative divisions of Galicia by province and totals.
Quadro I – Estatísticas descritivas das diferentes divisões administrativas da Galiza por provincias e totais.

Administrative division	Variable	Province				Galicia
		A Coruña	Lugo	Ourense	Pontevedra	
Province	Population	1 128 694	338 921	318 739	948 302	2 734 656
	Surface (km ²)	7958	9860	7274	4496	29 588
	Population density (hab/km ²)	142	34	44	211	92
Local authority (LA)	Count	93	67	92	62	314
	People per LA	12 136	5 059	3 465	15 295	8 709
	Average surface (km ²)	85.6	147.2	79.1	72.5	94.2
Census tract (CT)	Count	920	298	283	748	2249
	People per CT	1227	1137	1126	1268	1216
	Average surface (km ²)	8.6	33.1	25.7	6.0	13.2
Parish	Number per LA	10	4	3	12	7
	Count	939	1271	917	664	3791
	People per parish	1202	267	348	1428	721
Settlement	Average surface (km ²)	8.5	7.8	7.9	6.8	7.8
	Number per LA	10	19	10	11	12
	Count	9779	8989	3531	6044	28 343
Settlement	People per settlement	116	38	90	157	96
	Average surface (km ²)	0.8	1.1	2.1	0.7	1.0
	Number per LA	105	134	38	97	90

Source: IGE (2002) and INE (2017a, 2017b, 2017c and 2017d)

1. Census tracts

Census tracts are used by the INE in order to carry out the census data collection every ten years and they are designed to facilitate the process and to be as homogeneous as possible population wise, so they are roughly similarly sized in terms of population with a recommended maximum of 2500 people per tract, while keeping municipal borders unaltered (Instituto Nacional de Estadística, 2015). In the Galician case, in the year 2011 there were 2249 census tracts (table I). In practical terms this means that most municipalities which are less than 2500 people will consist of one census tract only, which will exactly coincide with the municipal limits, while in more populated local authorities the total area is split into smaller units. The consequence of this is that while in rural and less populated local authorities' census tracts tend to be the same size and shape as municipal boundaries, in densely populated municipalities such as cities there can be hundreds of them, allowing for a more detailed population representation. This disparity has a cartographic consequence when displaying population related data, as there is a clear

concentration in urban areas. Therefore, when the whole of Galicia is represented using graduated colour symbols on a choropleth map, most areas appear as sparsely populated due to the very high density of census tracts within cities.

2. Parishes

Parishes – known as *parroquias* in Spanish and Galician – are the traditional sub-municipal divisions into which Galician local authorities are divided and, although they are not fully officially recognised, inhabitants tend to have a strong sense of belonging (Tolosana, 2004). Contrary to census tracts, parishes in Galicia can be used to represent the rural population in a more detailed way, while urban areas are more coarsely displayed. In the year 2015 there were 3791 official parishes in Galicia (table I), which means that on average each municipality had around 12 parishes, but values vary dramatically from town to town. Therefore, a parish population map will provide a level of detail for rural areas which is not achievable by using census tracts.

3. Settlements

Settlements are officially known as Singular Population Entities – *Entidad Singular de Población in Spanish* – and are defined by the INE as “any habitable area of the municipal terminality, inhabited or exceptionally inhabited, clearly differentiated within the same and which is known by a specific denomination that identifies it without possibility of confusion” (INE, 2017a). Therefore, any group of buildings which can be potentially inhabitable and is separated from other groups of buildings is considered an independent settlement, even if nobody currently lives there permanently. The number of settlements varies yearly as new dwellings are built and demolished, but in the year 2015 Galicia had over 28 000 inhabited settlements, which given its size, it is an astonishing number as it roughly equals to one settlement per square kilometre (table I), a value eight times higher than the Spanish average. Therefore, settlements are the minimal population unit with which Galicia – and Spain – can be mapped using official data, but due to their small size, variable number, and unofficial status, there are no boundaries linked to each settlement, just latitude and longitude coordinates associated to their centre are provided by official sources (Centro Nacional de Información Geográfica [CNIG], 2015). Other ways the INE refers to settlements is as *Nomenclátor* or *Unidad poblacional*, with information about them provided directly by each local authority.

III. METHODS

Four population density maps for each administrative division – municipal, census tracts, parishes and settlements – plus a standard guide map with Galicia’s main cities were produced to illustrate how each different division influences the way in which the population is represented.

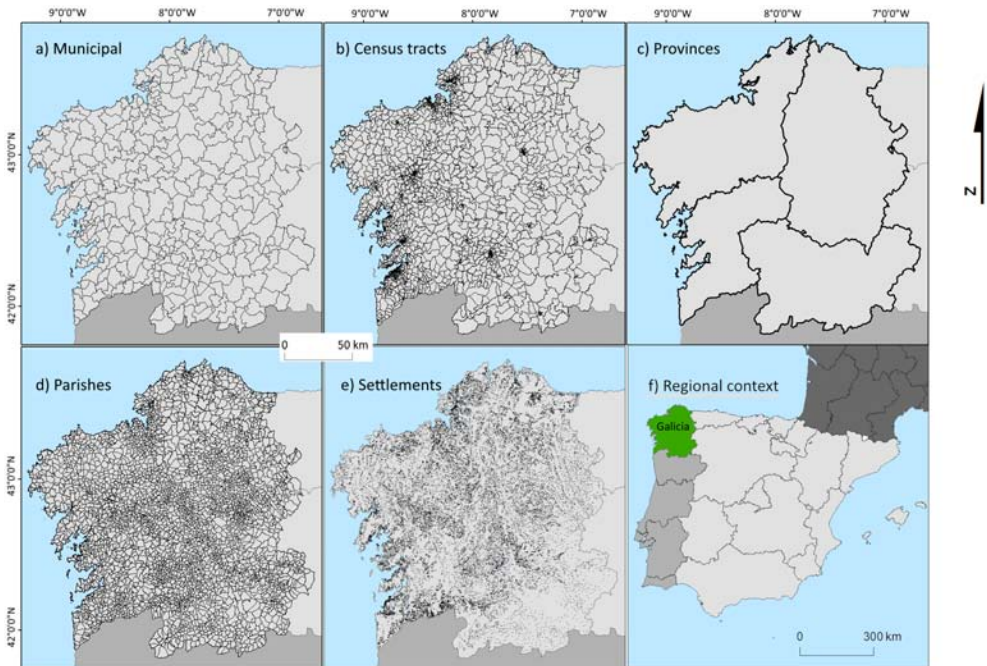


Fig. 1 – Administrative limits of Galicia at different levels of scale. Colour figure available online.

Fig. 1 – *Limites administrativos da Galiza em diferentes niveis de escala. Figura a cores disponivel online.*

Source: IET (2011) and INE (2017a, 2017b, 2017c and 2017d)

1. Municipal and census tracts

Municipal divisions were obtained for the year 2015 from the Instituto Geográfico Nacional (IGN) official cartography (CNIG, 2017) and population data from the INE's official population figures for January 2015 (INE, 2017b). Since both are official sources, each polygon representing a local authority has a statistical code which matches with the INE's population dataset, so a simple join between the shapefile's attribute table and the population dataset is enough to create the population map.

Census tract data was produced as part of the 2011 census process, so contrary to the other datasets used in this research, all cartographic (INE, 2014) and demographic (INE, 2017c) information refers to that year, and not 2015. However, it was decided that even if four years older, census tract data as provided by the INE (INE, 2017c) was comparable to the other datasets used as part of this paper. Even if not ideal, this does not represent an insurmountable issue, as during the 2011-2015 period Galician population has remained relatively stable with an overall population change of -1.4% (INE, 2017d). Just as with the municipal dataset, both cartographic and demographic data had a field in common with a numerical value for each census tract, so a simple join by attributes was performed to create the population map.

2. Parishes

Contrary to the other administrative divisions, parishes do not have a fully official status, so instead of using national data providers it was necessary to obtain the data from regional institutions. The Instituto de Estudios do Territorio (IET), part of the regional government, provides data at parish level, but it is quite dated as both demographic figures (Instituto Galego de Estatística [IGE], 2002) and the cartography associated to it (IET, 2011) are from the year 2001. Although there is population data available from more recent years (IGE, 2017), this information is not as accessible and, since 14 years have passed between 2001 and 2015, parish limits have changed in places, increasing the difficulty of linking population data with its corresponding parish. It was therefore decided that the best compromise was to use the available 2001 parish limits and the settlement population dataset, which is delivered as points with a population value attributed to them. By performing a spatial join based on location and adding all the population of the points which fell within each parish limits the total population for each parish in the year 2015 was established and mapped. On average each parish contained 7.5 settlements; with eight of them having 50 or more contained within their borders.

3. Settlements

While information about settlements has been traditionally provided as part of the national gazetteer – *nomenclátor* in Spanish (IGN, 2017), having population data associated to each settlement is more recent. The dataset is provided by the IGN by combining information from, among others, the Local Entities Register – collated by the Ministry of the Finance and Public Administrations, the INE and the cartographical data base of the IGN (CNIG, 2015). The dataset is available to download as a Microsoft Access Database or as an OpenDocument Database and includes a table named “entidades” where the information used in this paper is held. The “entidades” table contains a list of 153 125 settlements with information including a statistical code, name, province, population and X-Y coordinates which use the ETRS89 system. However, the given number of settlements does not match the number of settlements provided by other sources, which give a figure of around 60 000 (Comins & Moreno, 2003). Although this source has been superseded by an increase in the quality of the data provided by the INE in recent years, the dataset provided by the INE still has an excess number of settlements. The issue with the original dataset is that some settlements are introduced multiple times because they are classified using different denominations, such as municipality, local authority seat and settlement, multiplying the population and number of settlements. In order to eliminate said duplicities, all cases with a statistical code ending in zero – while being careful to not include those which ended in 10, 20, 30 or 40 – were eliminated, as they are duplications. To check that the process was correct, the total population per local authority was added and compared with official population data from the INE, with positive results for the Galician case and a total population discrep-

ancy of 0.06% for the Spanish case. Finally, all those settlements with a population of 0 were deleted (Comins & Moreno, 2003) as they could be abandoned houses, holiday accommodation or new settlements with no inhabitants yet. This left a total of 64 224 populated settlements in Spain with known coordinates, as those without given coordinates could not be mapped, of which 28 307 pertained to Galicia. It is important to note that all population types – cities, towns, villages, hamlets, etc. – and regardless of their size are considered a settlement.

The biggest issue when working with settlements, however, is mapping them as a choropleth map in order to allow a comparison with all other administrative divisions, as since they have no official limits no cartography exists which defines them, so they are just points without surface. It is therefore necessary to create a bespoke cartography which represents some 2D limits, even if not official but purely geometrical. The first step to achieve this is to convert the X-Y coordinates into a point layer and to split it by the town they belong to. If a municipality only has one settlement, the settlement limits are those of the town, so no more processing is required, but only two cases were found in Galicia in which this was the case. Those local authorities which have three or more settlements within their limits – in the Galician case, there are no local authorities with only two settlements – need a more complex method in order to convert punctual information into a choropleth map. While in the past interpolation methods have been used to convert settlement data values into a population raster (García González & Cebrián Abellán, 2006), this produces results which are difficult to compare with the other sub-municipal divisions, so it was decided to use a fully vector based methodology in order to provide results in the form of a choropleth map (Nobajas & Nadal, 2015). The method consists in splitting the settlements shapefile into 312 new shapefiles, one per municipality, and also splitting a municipal shapefile into the same number of new shapefiles where only the border of one local authority is stored. The next step is to create Voronoi – also known as Thiessen – polygons (Albrecht, 1998) for each of the files containing settlement data with the purpose of converting punctual data into polygon data. Afterwards each of the new polygon shapefiles composed by Voronoi polygons is clipped using its corresponding municipal boundaries and finally all the resulting polygons are merged together to generate a continuous map for all of Galicia.

This method, while being computationally demanding and more complex to perform, provides a final result which is directly comparable to the other administrative divisions, as they all use polygons. Other tested forms of representation, such as proportional symbols, did not deliver such good results due to the uneven density of settlement locations across the study area, but also to the sheer quantity of them. Punctual methods of representation produced very high levels of cartographic noise which impeded reading the maps clearly, while by using the method explained here the final result is much clearer and spatial trends are easily observable. It should be also stressed that the resulting map is just a mean of representing population data, not an administrative map, as settlements are provided as points by the INE because there are no administrative divisions associated to them.

4. Visualisation methods

In order to compare the four different maps, it was deemed best to represent population density rather than absolute population figures as it allows for an easy comparison between population datasets which have dissimilar ranges of maximum and minimum values. In Galicia the average population for municipalities is almost 9000, while for settlements it is less than 100, so comparing both administrative divisions directly could be misleading. Population density was measured by dividing the population contained in each polygon by its surface area, but while official surface figures exist for municipalities, other sub-municipal units do not have them and have to be calculated using GIS, so for the sake of consistency it was decided to use calculated figures in all the maps, even when official figures existed.

Once all population density figures for all maps were obtained, it was observed that, as expected, population density ranges between all the maps were highly different. For example, census tracts tend to have similar populations and a very varied surface due to the way they are designed, while municipalities and settlements can have very different populations but similar areas, making population density figures very different and comparisons difficult. In order to minimise this and enhance the comparison between maps, it was decided to rank population density and represent it using quintiles (Smith, 1986). By displaying information in this way polygons are divided into five equal groups each representing 20% of the administrative divisions ordered by population density (Pamalon *et al.*, 2009). The result of using this method is that maps are then directly comparable and the different characteristics of the different divisions are minimised since a common scale can be applied to all maps.

IV. RESULTS AND DISCUSSION

Thanks to the proposed method it is possible to compare the population density at different levels of detail at a glance and in an easily comparable way (fig. 2). Although total population numbers are the same in all four cases, changing the way administrative limits are defined provides very different visualisations, not only due to the different levels of granularity, but also due to the modifiable areal unit problem (Openshaw & Taylor, 1981). Representing population at municipal level, although being the most common manner of doing so, is too coarse to illustrate with the necessary level of detail how population is distributed, as the areas are usually too large and may encompass a variety of settlement types, from cities to tiny hamlets, all within the same administrative limits.

The influence that the modifiable areal unit problem has in how population density is represented is quite clear if the population is represented by using census tracts or parishes, as they have a relatively similar number of individual units – 2249 and 3791 respectively – yet they produce very different results. Due to their nature, census tracts are much denser in cities and urban areas as they are designed not to go above certain popu-

lation numbers, while in rural areas they often share boundaries with municipal limits. The consequence of this is that when using quintiles, the region looks quite depopulated, as higher densities are all found in urban centres, while rural areas have the lowest values. On the other hand, parishes are more similar in size between them, so population density is represented in a very different way, one which is perhaps more similar to the municipal distribution but with a greater level of detail.

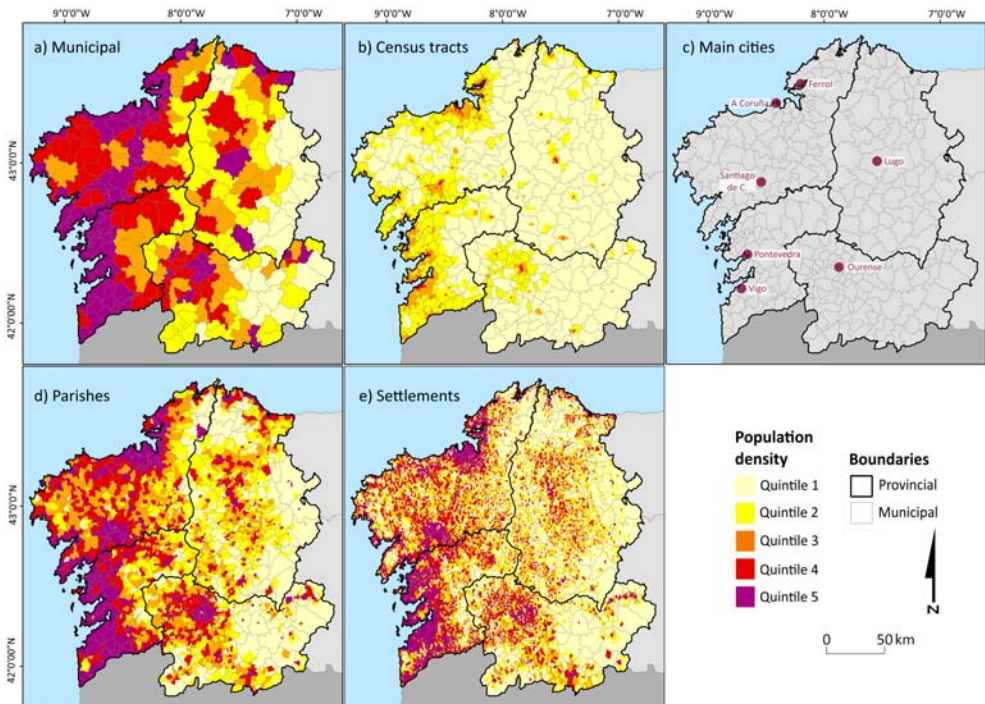


Fig. 2 – Population density of Galicia using quintiles and different administrative divisions. A map highlighting the location of the main cities was included for reference. Colour figure available online.

Fig. 2 – Densidade populacional da Galiza por quintis e diferentes divisões administrativas. Um mapa da localização das principais cidades foi incluído para referência. Figura a cores disponível online.

Source: INE (2017a, 2017b, 2017c and 2017d), IGE (2002) and IET (2011)

It is however by using settlements that the most detailed population distribution can be observed for the first time. Since the population is represented using over 28 000 individual polygons, each corresponding to a settled unit of population, the modifiable area unit problem is greatly minimised, as they represent the smallest unit for which population data can be obtained. This allows the observation of population patterns previously hidden by using other administrative divisions, while allowing a clearer understanding of the hidden complexity of population distribution. As figure 2 shows, all four population density maps show very different patterns which will convey a different type of information to different readers.

The different ways in which population representation is perceived also has a scalar dimension, as using different scales provides new interpretation possibilities. For instance, Lugo, one of the four provincial capitals and one of Galicia's largest cities with a population of around 100 000 people covers an area of over 330km² which can be divided into either 71 census tracts, 54 parishes or 302 settlements. Due to its large population it is on the top quintile for population density at the municipal level, but its internal population density is not equally spread across its area, as there are important variations within its borders. As figure 3 shows, its population distribution is very complex with a mosaic of different population densities spanning across all quintiles, something important when making planning choices or providing services, hence sub-municipal divisions allow us to understand the population distribution much more accurately.

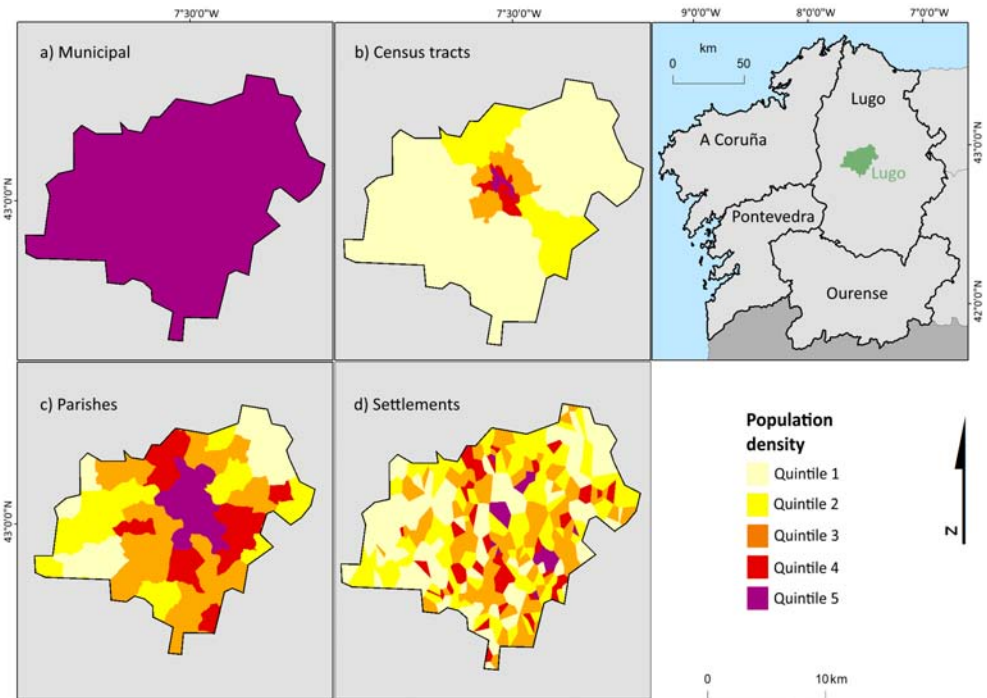


Fig. 3 – Lugo’s population density using different administrative boundaries.
 Colour figure available online.

*Fig. 3 – Densidade populacional de Lugo presentada segundo diferentes limites administrativos.
 Figura a cores disponível online.*

Source: IET (2011), IGE (2002) and INE (2017a, 2017b, 2017c and 2017d)

Sub-municipal complexities are not limited to cities, as smaller townships also present big population differences within their borders. As an example, the town of Guitiriz has a population of almost 5500 people and a surface of almost 300km², so consequently it has a lower population density which places it in Galicia's second quintile. It can also be

divided into six census tracts; of which all, bar one, are at the lowest population density quintile. If the population is divided using its 18 parishes a clearer trend appears due to the towns' rurality, but it is only when the around 300 settlements are mapped that a precise representation of how the population is distributed emerges (fig. 4). Thanks to the use of settlement level data, it is possible to easily observe where the population concentrates within the municipal boundaries.

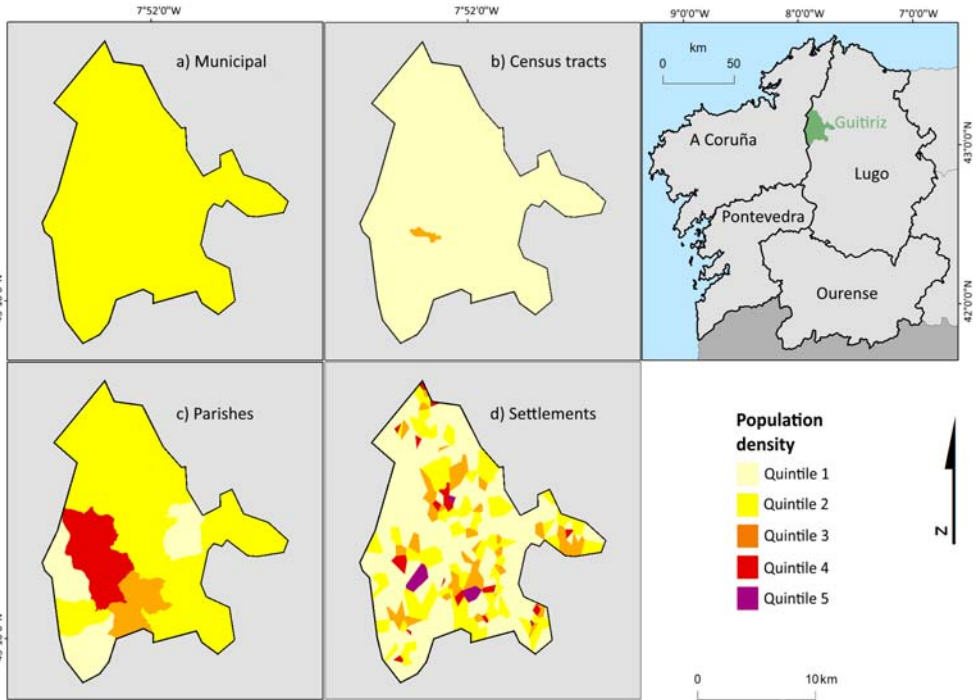


Fig. 4 – Guitiriz's population density using different administrative boundaries.
Colour figure available online.

*Fig. 4 – Densidade populacional de Guitiriz representada usando diferentes limites administrativos.
Figura a cores disponível online.*

Source: IET (2011), IGE (2002) and INE (2017a, 2017b, 2017c and 2017d)

While other more straightforward methods such as interpolation (García González & Cebrián Abellán, 2006) or a direct creation of a Voronoi diagram could be used, the results would differ. For example, as figure 5 shows, if Voronoi polygons are created without considering the municipal limits, polygons from different local authorities can become mixed, providing less accurate information, which at the level of scale achieved by working with settlements can become an issue. This makes a big difference, as it constrains the settlement limits to the municipal limits, adding an additional level of rigour.

Other methods used to avoid or minimise the modifiable area unit problem, such as dasymetric mapping or population grids, can be used in conjunction or as an alternative

to the proposed method, although the advantage of the one presented here is that it does not rely on further datasets, it is self-contained to official data sources. Both dasymetric maps and population grids require combining land use data with population data (Goerlich & Cantarino, 2012), while here a similar result can be achieved by exclusively processing the initial datasets. However, the application of the proposed method is limited to those cases where there is a need to represent population entities that have no defined geographical limits. Due to its historic and geographic characteristics, the Galician case is a prime example of this, as it would be the whole of Spain due to the way official population data is distributed, but it may not be applicable to other regions or countries unless they offer similar datasets.

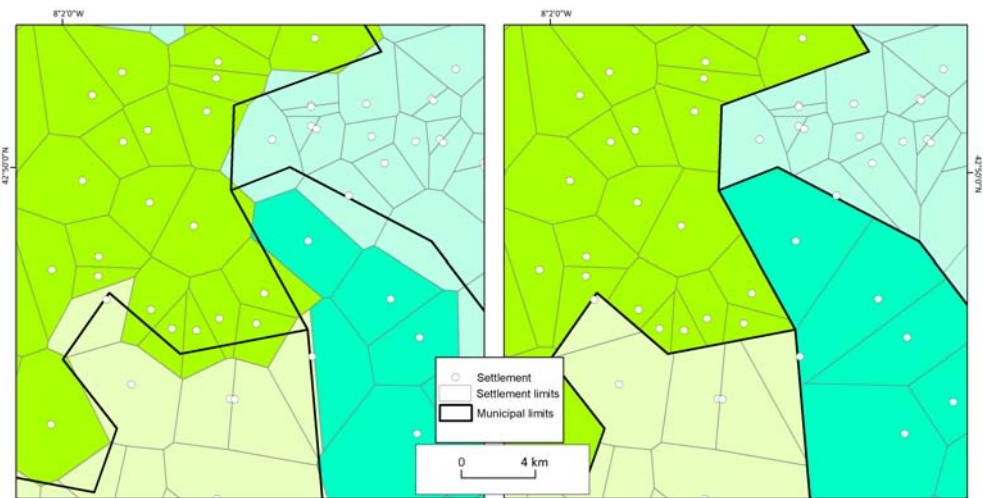


Fig. 5 – Creating Voronoi polygons directly (left) or using the proposed method (right).
Colour figure available online.

*Fig. 5 – Criação direta de Voronoi (à esquerda) ou usando o método proposto (à direita).
Figura a cores disponível online.*

Data from IET (2011), IGE (2002) and INE (2007)

V. CONCLUSIONS

Due to the way in which population data are gathered, working at sub-municipal population level in Spain and other locations can be challenging, but in certain areas it is of paramount importance as data aggregated at the municipal level is not always fit for the purpose. Even when sub-municipal population data are available, such as census tracts or parishes, they can provide unsuitable results due to their nature, which can put too much weight on urban or rural areas. Therefore, the developed method, which allows converting very detailed settlement data into a choropleth map, has enabled the mapping of Galicia's population with an unprecedented level of detail.

The process, however, is not without its challenges, especially if it is to be expanded across all of Spain or other countries. Due to the ever-changing nature of municipal and sub-municipal limits, new local authorities can be created while others eliminated, which exacerbates the difficulties working with data from different sources or years. Furthermore, in this case, settlement data is not provided ready to use, so significant amounts of tinkering are required in order to map it. At times it is not possible to fix all of these issues, which may lead to errors in the data, but fortunately they represent a very small percentage of settlements and/or population. Finally, the provided coordinates for settlements are at times either not correct or not sufficiently precise, which means that some settlements may end up in a different municipality when performing a spatial join with a municipal polygon vector file. Again, fortunately this only occurs in a very small minority of cases which can often be manually fixed.

Thus, even if challenging at times, the presented method allows the visualisation of how population is distributed using four different administrative divisions allowing the direct comparison between them, something which had not been achieved before. This should start a debate about which is the most suitable way of representing the population of an area with such a large amount of settlements, as no single representation is suitable for all purposes. Additionally, the creation of fictional limits can be controversial, as information which is mapped has a tendency to be considered true (Monmonier, 2018), which could lead to territorial conflicts, so it is important to make it clear that they are theoretical borders which have no legal or historical standing.

ORCID ID

Alexandre Nobajas  <https://orcid.org/0000-0002-9670-0188>

REFERENCES

- Albrecht, J. (1998). Universal analytical GIS operations: A task-oriented systematization of data structure-independent GIS functionality. In M. Craglia & H. Onsrud (Eds.), *Geographic Information Research: Transatlantic Perspectives* (pp. 577-591). CRC Press. <https://doi.org/10.1201/9780367804756>
- Barbosa-Brandão, V., Santé Riveira, I., & Crecente Maseda, R. (2015). Evolution of Legal Criteria for the Identification and Zoning of Rural Settlements in Galicia, NW Spain. *European Planning Studies*, 23(2), 398-429. <https://doi.org/10.1080/09654313.2013.867317>
- Batista e Silva, F., Gallego, J., & Lavalle, C. (2013). A high-resolution population grid map for Europe. *Journal of Maps*, 9(1), 16-28. <https://doi.org/10.1080/17445647.2013.764830>
- Cabezuelo-Lorenzo, F., Rey-García, P., & Tapia-Frade, A. (2016). Analysis of tools for citizens' monitoring of public authorities: Municipal transparency in Castile and León. *Revista Latina De Comunicación Social*, (71), 1261. <https://doi.org/10.4185/RLCS-2016-1145en>
- Campanera, J. M., Nobajas, A., & Higgins, P. (2014). The Relationship Between Residential Quality of Life and Socioeconomic Status in England. *Urban Affairs Review*, 50(1), 100-133. <https://doi.org/10.1177/1078087413485204>
- Centro Nacional de Información Geográfica. (2017). *Cartography of SIANE*. <http://centrodedescargas.cnig.es/CentroDescargas/buscadorCatalogo.do?codFamilia=CAANE>
- Centro Nacional de Información Geográfica. (2015). *Towns and Population Entities Geographical*

- Gazetter*. http://centrodedescargas.cnig.es/CentroDescargas/equipamiento/BD_Municipios-Entidades.zip
- Comíns, J. S., & Moreno, D. R. (2003). *Población y Poblamiento Rural de España: Un Primer Análisis a la luz del Censo 2001* [Population and Rural Settlement of Spain: A First Analysis in light of the 2001 Census]. Universidad de Alcalá.
- Coulson, M. R. (1987). In the matter of class intervals for choropleth maps: With particular reference to the work of George F Jenks. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 24(2), 16-39. <https://doi.org/10.3138/U7X0-1836-5715-3546>
- Dixon, O. M. (1972). Methods and progress in choropleth mapping of population density. *The Cartographic Journal*, 9(1), 19-29. <https://doi.org/10.1179/caj.1972.9.1.19>
- Eicher, C. L., & Brewer, C. A. (2001). Dasymeric mapping and areal interpolation: Implementation and evaluation. *Cartography and Geographic Information Science*, 28(2), 125-138. <https://doi.org/10.1559/152304001782173727>
- Eurostat. (2021). *Population grids*. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_grids
- Gajardo, C. P., Morillas J. B., Escobar, V. G., Vílchez-Gómez, R., & Gozalo, G. R. (2014). Effects of singular noisy events on long-term environmental noise measurements. *Polish Journal of Environmental Studies*, 23(6), 2007-2017.
- García González, J. A., & Cebrián Abellán, F. (2006). *La interpolación como método de representación cartográfica para la distribución de la población: Aplicación a la provincia de Albacete* [Interpolation as a cartographic representation method for population distribution: Application to the province of Albacete]. In M. T. Camacho Olmedo, J. A. Canete Pérez, & J. J. Lara Valle (Eds.). XII Congreso Nacional de Tecnologías de la Información Geográfica: El acceso a la información espacial y las nuevas tecnologías geográficas (pp. 165-178). Universidad de Granada.
- Goerlich, G. F. J. (2007). ¿Cuántos somos? Una excursión por las estadísticas demográficas del Instituto Nacional de Estadística (INE) [How many are we? The Demographic Statistics of the Spanish national Statistics Institute (INE)]. *Boletín de la Asociación de Geógrafos Españoles*, 45, 123-156. <https://dialnet.unirioja.es/servlet/articulo?codigo=2519804>
- Goerlich, F. J., & Cantarino, I. (2013). A population density grid for Spain. *International Journal of Geographical Information Science*, 27(12), 2247-2263. <https://doi.org/10.1080/13658816.2013.799283>
- Goerlich, G. F. J., & Cantarino, M. I. (2012). *Una grid de densidad de población para España* [A population density grid for Spain]. Fundación BBVA.
- Goerlich, G. F. J., & Cantarino, M. I. (2011). Cartografía y demografía: una grid de población para la Comunitat Valenciana. [Mapping and demography: a population grid for the Valencian Community] *Estadística Española*, 53(176), 127-161. <https://dialnet.unirioja.es/servlet/articulo?codigo=3786661>
- Holt, J. B., & Lu, H. (2011). Dasymeric mapping for population and sociodemographic data redistribution. *Urban Remote Sensing: Monitoring, Synthesis and Modeling in the Urban Environment*, (250), 195-210. <https://doi.org/10.1002/9780470979563.ch14>
- Huang, X., Wang, C., & Li, Z. (2019, November). High-resolution population grid in the CONUS using Microsoft building footprints: a feasibility study. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Geospatial Humanities* (pp. 1-9). <https://doi.org/10.1145/3356991.3365469>
- Instituto de Estudios do Territorio. (2011). *Mapa de Parroquias ee Galicia 1/200.000* [Map of Galician Parishes 1/200.000]. IET. http://visorgis.cmati.xunta.es/cdix/descargas/visor_basico/Parroquias.zip
- Instituto Galego de Estatística. (2017). *Nomenclátor de Galicia* [Galician Gazetteer]. IGE. <https://www.ige.eu/igebdt/esq.jsp?paxina=002001&c=-1&ruta=nomenclator/nomenclator.jsp>
- Instituto Galego de Estatística. (2002). *Censo de Poboación e Vivendas 2001. Información Relativa ás Entidades Colectivas de Poboación (Parroquias)* [Population and Housing Census 2001. Information Related to Collective Population Entities (Parishes)]. IGE. <http://www.ige.eu/igebdt/esq.jsp?paxina=002001&ruta=parroquias/index.jsp>
- Instituto Geográfico Nacional. (2021). *Nomenclátor: Población del Padrón Continuo por unidad poblacional. Resultados* [Gazetteer: Population of the Continuous Register by population unit. Results]. IGN. https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177010&menu=resultados&cidp=1254734710990#!tabs-1254736195532

- Instituto Geográfico Nacional. (2017). *Registro Central de Cartografía. Nombres Geográficos* [Central Register of Cartography. Geographic names]. IGN. <https://www.ign.es/ign/layoutIn/regcartoNombresGeograficos.do>
- Instituto Nacional de Estadística. (2017a). *Revision of Municipal Register to 1 January of each Year*. INE. <http://www.ine.es/nomen2/ Metodologia.do?c=2&l1=/nomen2/ Metodologia.do?L=1&i1=en&t1=English&a1=English&l2=/nomen2/ Metodologia.do?L=0&i2=es&t2=Espa%C3%83%C2%B1ol&a2=Espa%C3%83%C2%B1ol&L=1>
- Instituto Nacional de Estadística. (2017b). *Official Population Figures Referring to Revision of Municipal Register 1 January*. INE. <https://www.ine.es/dynt3/inebase/en/index.htm?padre=517>
- Instituto Nacional de Estadística. (2017c). *Data by Census Sections at 1 January 2011*. INE. <http://www.ine.es/dynt3/inebase/en/index.htm?type=pcaxis&file=pcaxis&path=%2Ft20%2Fe245%2Fp07%2F%2Fa2011>
- Instituto Nacional de Estadística. (2017d). *Serie Detallada desde 2002* [Detailed Series since 2002]. INE. <http://www.ine.es/jaxiT3/Tabla.htm?t=9681&L=0>
- Instituto Nacional de Estadística. (2015). *Secciones Censales de la Comunidad de Madrid* [Population Census of Madrid]. <http://www.madrid.org/ies-tadis/fijas/clasificaciones/seccioncensal.htm>
- Instituto Nacional de Estadística. (2014). *Contorno de las Secciones Censales a 1 de Noviembre de 2011 en Formato SHP* [Outline of Census Tracts as of November 1, 2011, in SHP Data Format]. INE. http://www.ine.es/censos2011_datos/cartografia_censo2011_nacional.zip
- Instituto Nacional de Estadística. (2011). *¿Cómo es España? 25 mapas para descubrirla km² a km²* [How is Spain? 25 maps to discover it km² by km²]. INE.
- Jurado, A. (2004). Diferencias entre censo de población y padrón municipal [Differences between population census and municipal register]. *Revista de Estadística y Sociedad*, 3, 12-13.
- Lavin, S. (1986). Mapping continuous geographical distributions using dot-density shading. *The American Cartographer*, 13(2), 140-150. <https://doi.org/10.1559/152304086783900068>
- Monmonier, M. (2018). *How to Lie with Maps*. University of Chicago Press.
- Nobajas, A., & Nadal, F. (2015). From historical map to online 3D recreation: the 1861 cadastral map of Horta (Barcelona). *Cartography and Geographic Information Science*, 42(3), 211-223. <https://doi.org/10.1080/15230406.2014.998285>
- Openshaw, S., & Taylor, P. J. (1981). The modifiable areal unit problem. In N. Wrigley & R. Bennett (Eds.), *Quantitative geography: a British view* (pp. 60-69). Routledge & Kegan Paul.
- Pampalon, R., Hamel, D., & Gamache, P. (2009). A comparison of individual and area-based socio-economic data for monitoring social inequalities in health. *Health Reports*, 20(4), 85-94. <https://pubmed.ncbi.nlm.nih.gov/20108609/>
- Petrov, A. (2012). One hundred years of dasymetric mapping: back to the origin. *The Cartographic Journal*, 49(3), 256-264. <https://doi.org/10.1179/1743277412y.0000000001>
- Raisz, E. (1962). *Principles of Cartography*. McGraw-Hill.
- Rivero, J. B., & Lladós, M. G. (2014). El mapa municipal de España. Una caracterización geográfica [The municipal map of Spain. A geographic characterization]. *Boletín De La Asociación De Geógrafos Españoles*, (64), 11-36.
- Rodríguez, J. J. F. (2016). *The Parliament of Galicia* [Paper presentation]. In G. Laforest & A. Lecours (Eds.), *The Parliaments of Autonomous Nations* (pp. 72-96). McGill-Queen's University Press.
- Tolosana, C. L. (2004). *Antropología Cultural de Galicia* [Galician Cultural Anthropology]. Ediciones Akal.
- Smith, R. M. (1986). Comparing traditional methods for selecting class intervals on choropleth maps. *The Professional Geographer*, 38(1), 62-67. <https://doi.org/10.1111/j.0033-0124.1986.00062.x>
- Wiegand, P. (2003). School students' understanding of choropleth maps: Evidence from collaborative mapmaking using GIS. *Journal of Geography*, 102(6), 234-242. <https://doi.org/10.1080/00221340308978555>