

Distribución espacial de la población en Andalucía. Año 2013

IRIA ENRIQUE | JOSÉ E. MOLINA | SERAFÍN OJEDA | MARÍA ESCUDERO | GERMÁN PÉREZ

Instituto de Estadística y Cartografía de Andalucía

1. Introducción

El Instituto de Estadística y Cartografía de Andalucía (IECA) ha acometido un proyecto de investigación orientado a la georreferenciación de la población, cuyos resultados han dado lugar a la presentación del producto estadístico-cartográfico denominado “La distribución espacial de la población en Andalucía. Año 2013”. En este proyecto, nuestra región ha sido dividida en una malla de celdas de 250 m de lado. Como fuente principal de información para la localización de la población se han usado datos de carácter administrativo recogidos en el Registro de Población de Andalucía (RPA), actualizados a fecha de 1 de enero de 2013. En este registro figuran todas las personas empadronadas en algún municipio de Andalucía, añadiéndose información relativa a la dirección postal, así como ciertas características demográficas relativas al sexo, la edad y la nacionalidad de cada una de las personas registradas.

Esta información ha sido georreferenciada asignando las coordenadas geográficas correspondientes a las personas residentes en cualquier vivienda situada en Andalucía. La información georreferenciada proveniente del RPA se ha agregado a la malla regular. Esta aproximación *bottom-up* ha permitido la georreferenciación del 87,4% de la población registrada en el RPA. El 12,6% de la población restante ha sido asignado a las celdas correspondientes usando técnicas estadísticas. Para ello se utiliza información auxiliar proveniente del RPA y del Catastro (aproximación *top-down*). Para la difusión de los resultados en internet se ha desarrollado una capa de información geográfica descargable, un servicio WMS y un visualizador cartográfico.

2. Sistema de codificación y fuentes de información

Se ha generado una capa de celdas regulares. Se ha utilizado para ello la herramienta de generación de mallas desarrollada por Eurostat¹ incorporada a ArcGIS 10. Esta malla regular se ha configurado siguiendo las recomendaciones del proyecto ESSnet GEOSTAT 1A², y se ha creado inicialmente en el sistema de referencia ETRS89-LAEA, adaptado posteriormente al ETRS89-UTM Zone 30N, que es el sistema de referencia estándar para la cartografía en España y Andalucía. Se obtiene una capa en formato shape con un total de 1.416.093 celdas (250 x 250 m cada una de ellas).

Las fuentes de información principales que se han utilizado son: Registro de Población de Andalucía, Callejero Digital de Andalucía y Catastro de Urbana.

1. <http://www.efgs.info/data/eurogrid/eurostat-grid-generation-tool-for-arcgis/view>

2. <http://www.efgs.info/geostat/1A>

3. Georreferenciación de la población. Enfoque *bottom-up*

El Registro de Población de Andalucía (RPA) proporciona una dirección postal para cada individuo y grupo de individuos (hogares). En el Callejero Digital de Andalucía (CDA) cada uno de los portales tiene su dirección postal y sus coordenadas geográficas correspondientes. A través de complejas técnicas de enlace, las direcciones postales contenidas en el RPA se vinculan con el portal correspondiente del CDA, asignándole a cada una de ellas una coordenada geográfica.

La georreferenciación de los portales y de la población que vive en ellos es posible a través de Tecnología de Sistemas de Información Geográfica (SIG) una vez que se han asignado las coordenadas. El proceso de enlace que se ha ejecutado ofrece también información del grado de coincidencia entre la dirección postal de RPA y la de CDA (fiabilidad en el emparejamiento y por lo tanto de la coordenada asignada).

En esta fase se genera una capa de puntos. Cada punto representa un portal al que se le asocia el número de habitantes que residen en él y las características demográficas de éstos. El paso siguiente consiste en la agregación de esa información a cada celda, añadiendo los datos de todos los puntos (portal y su población) localizados en cada celda. Esto se realiza mediante una operación de unión espacial. Cada celda contiene el número total de habitantes pertenecientes a cada celda según la localización de los puntos (portales). Con este procedimiento se han localizado 7.453.107 de habitantes (87,4% de la población registrada en el RPA).

4. Georreferenciación de la población. Enfoque *top-down*

El 12,6% de la población residente en hogares no ha podido ser georreferenciada a través de la aproximación *bottom-up*. Para solventar esta circunstancia se ha diseñado un procedimiento *top-down* de asignación de cada uno de estos hogares no georreferenciados a una celda. La sección censal es el nivel máximo de desagregación territorial disponible para estos hogares. De este modo se dispone del número de hogares y de personas no georreferenciados para cada una de las secciones censales. A partir de técnicas estadísticas y con la información auxiliar, las operaciones *top-down* asignan cada uno de esos hogares a una de las celdas que componen esa sección censal.

Antes de ejecutar el modelo de localización de hogares se llevó a cabo un importante trabajo de procesamiento de la información geográfica. Éste se realiza fundamentalmente con la información proveniente de dos capas muy útiles para la construcción del modelo de localización: la capa de secciones censales elaborada por el Instituto Nacional de Estadística (INE) y la capa del parcelario del catastro de urbana, elaborada por la Dirección General de Catastro.

4.1. Información de secciones censales

La localización de los hogares no georreferenciados está disponible a nivel de sección censal. El objetivo principal de esta fase del proyecto es transferir esos datos (hogares no georreferenciados) desde la sección censal a la celda que le corresponde dentro de esa sección censal.

Con este fin se genera una capa nueva por la intersección de la capa de secciones censales con la de la malla de celdas regulares. Esta capa nueva establece la división en entidades de secciones y celdas, mostrando qué celdas quedan total o parcialmente dentro de cada una de las secciones censales en Andalucía. Las entidades espaciales generadas en esta capa de intersección no son

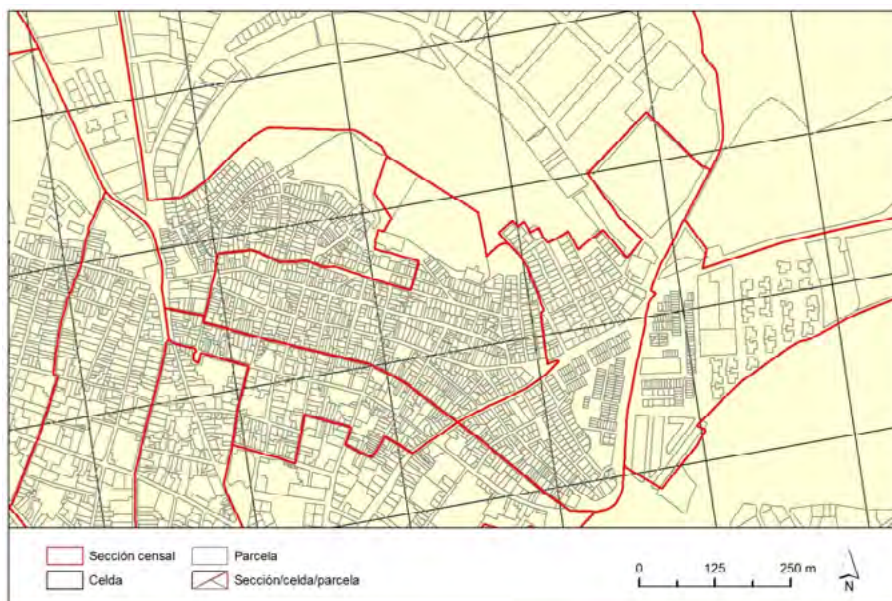
ni secciones ni celdas, sino que es la intersección de ambas, la *sección/celda*. Ésta es la unidad de referencia para la realización del análisis, de tal modo que el modelo de localización asigna población desde la sección censal a la entidad sección/celda y las celdas se construyen posteriormente por la suma de todas las secciones/celdas que componen una celda.

4.2. Información catastral

El catastro de urbana proporciona información correspondiente a la ocupación del territorio identificando y localizando, entre otros usos, parcelas con uso residencial. Las parcelas con uso residencial son la clave del proyecto, de ahí que la información de la que se dispone en catastro relativa a sus atributos y sus características ha sido analizada en detalle.

La unidad de delimitación territorial del Catastro es la parcela, mientras que la unidad de análisis en el proyecto es la sección/celda. Por ello la intersección entre la capa de parcelas con la de sección/celda establece la división de cada parcela por sección/celda. Así, se genera una capa nueva con las parcelas que contiene cada una de las entidades sección/celda. Una parcela puede estar total o parcialmente dentro de una sección/celda. La unidad espacial de información en esta capa de intersección es la *sección/celda/parcela* (Mapa 1).

Mapa 1. Unidades espaciales: secciones censales, celdas, parcelas y sección/celda/parcelas



Fuente: elaboración propia.

El análisis del catastro se centra en las capas de parcelas urbanas que contienen información de uso, tanto residencial como de otros usos. Este análisis establece las áreas que podrían potencialmente acoger población (áreas residenciales, comerciales y recreativas, frente a las áreas industriales o en construcción).

La información catastral aporta también información de los bienes inmuebles relativos a la edad de éstos, tipo de vía donde se localiza el bien inmueble, etc. Esta información es relevante porque la no georreferenciación es fundamentalmente el resultado de la no posibilidad de enlazar los registros correspondientes a través del algoritmo de enlace desarrollado. De este modo, la información referida al grado de consolidación urbana (áreas urbanas de construcción reciente,

áreas con alto porcentaje de tipos de vías poco habituales, áreas con alto porcentaje de edificios de almacén, etc.) podría arrojar luz sobre los factores que dificultaron el proceso de vinculación entre registros de las dos bases de datos.

Como se ha mencionado anteriormente, la unidad de análisis espacial de este proyecto es la sección/celda, de tal modo que los atributos de cada parcela han sido transferidos a la unidad sección/celda/parcela. La sección/celda/parcela es la unidad de transferencia. Los atributos de las parcelas han sido distribuidos entre las sección/celda/parcelas que la componen y entonces sus atributos se agregan por cada sección/celda para obtener el número total de bienes inmuebles, media de año de construcción, porcentaje de bienes inmuebles registrados en avenidas, etc., por sección/celda.

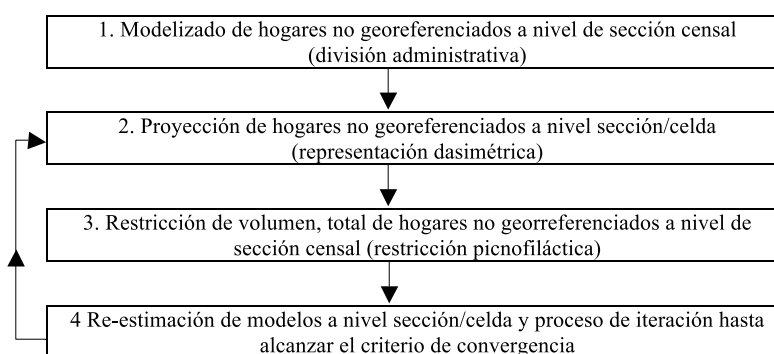
4.3. Modelos de desagregación espacial

Una vez que toda la información alfanumérica y espacial disponible del RPA y del catastro ha sido procesada, se desarrolla un proceso de análisis descriptivo. La población no georreferenciada no se distribuye uniformemente a través del territorio por lo que hay que llevar a cabo un análisis de la información disponible de RPA y catastro a nivel de sección censal. La población no georreferenciada registrada por secciones censales presenta ciertos patrones de carácter sociodemográficos, aunque la no georreferenciación podría no estar asociado exclusivamente a las características de la población, como se ha mencionado anteriormente.

Basándose en este análisis descriptivo y en las correlaciones observadas, se han definido cuatro modelos de hogares no georreferenciados, dependiendo de la incidencia de la no georreferenciación en cada sección censal. Todos los modelos aportaron resultados satisfactorios de regresión de los hogares no georreferenciados. Sobre los modelos estimados a nivel de sección censal se ha obtenido una primera proyección de hogares no georreferenciados al nivel de sección/celda, iniciando un proceso iterativo de desagregación espacial, tal y como se muestra en la Imagen 1.

El proceso se detiene una vez que se alcanza el criterio de convergencia, cuando la discrepancia entre los parámetros estimados en las dos iteraciones es más bajo de 10^3 . Los modelos estimados distribuyen los hogares no georreferenciados de cada sección censal entre las secciones/celdas que la componen. Los hogares e individuos localizados en el proceso *top-down* se suman a los datos de hogares e individuos georreferenciados (enfoque *bottom-up*) para obtener el número total de hogares e individuos que hay en cada celda, completando un método híbrido para desarrollar la malla regular de población.

Imagen 1. Proceso iterativo de desagregación espacial



Fuente: elaboración propia

5. Salvaguarda del secreto estadístico y la confidencialidad

Antes de llevar a cabo la difusión de los datos obtenidos con el enfoque híbrido, éstos han sido sometidos a un proceso de Control de Divulgación Estadística (SDC, en su acrónimo inglés). Los métodos de SDC minimizan los riesgos en la divulgación de datos sensibles en un nivel aceptable mientras que permite difundir toda la información posible. Se ha tomado la decisión de divulgar a partir de un determinado umbral. Así, los datos son protegidos para evitar la difusión de datos reales por debajo de un cierto umbral, así como la deducción exacta de los datos por debajo de determinado umbral.

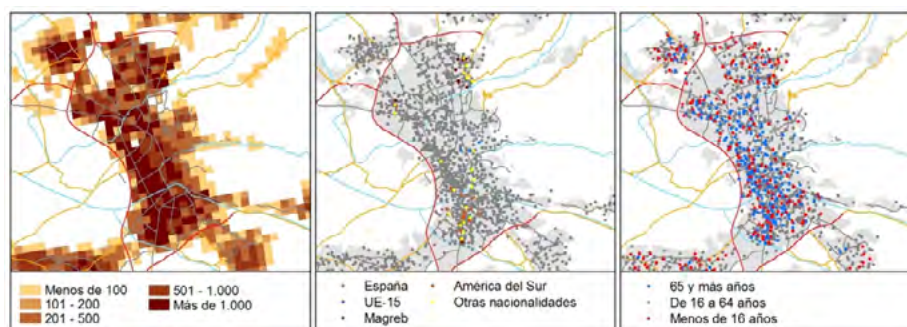
6. Difusión de los datos

Dentro del grupo de capas G07 Sistema Urbano³ de la colección de datos espaciales DERA (Datos Espaciales de Referencia de Andalucía), se encuentran accesibles para su descarga en formato de archivo de formas los datos del proyecto resultantes tras el proceso de protección SDC, su09_grid_poblacion_250.shp.

Igualmente, se ha incorporado al proyecto la estrategia de difusión, acceso y reutilización de la información del Plan Estadístico y Cartográfico de Andalucía 2013-2017, al ofrecer en la infraestructura de datos espaciales IDE Andalucía las capas de información espacial generadas como servicios interoperables. Ello permitirá a los usuarios consumir servicios de visualización WMS⁴, con lo que se facilitará su uso en combinación con otras capas de información espacial, y se contribuirá así al desarrollo de procesos de generación de valor añadido basados en la reutilización de la información.

Por otro lado, la página web del IECA⁵ ofrece también un visualizador cartográfico, que permite realizar consultas interactivas. Este visualizador ha sido desarrollado para mostrar la información en internet y hacer posible la consulta de este tipo de datos. Los datos son representados en tres mapas de celdas: un mapa de coropletas, “Población total”, y dos mapas de densidad de puntos, “Población por nacionalidad” y “Población por grupos de edad, como se muestran en los Mapas 2, 3 y 4.

Mapas 2, 3 y 4. Mapas de celdas interactivos: población total, población por nacionalidad y población por grupos de edad



Fuente: elaboración propia

3. http://www.juntadeandalucia.es/institutodeestadisticaycartografia/DERA/ficheros/G07_Sistema_Urbano.zip

4. <http://www.juntadeandalucia.es/institutodeestadisticaycartografia/geoserver-ieca/grid/wms?>

5. Se puede acceder al visualizador a través de la siguiente dirección: <http://www.juntadeandalucia.es/institutodeestadisticaycartografia/distribucionpob/index.htm>