

## CASO PRÁCTICO

# Una aproximación volumétrica a la desagregación espacial de la población combinando cartografía temática y datos LIDAR

Goerlich, F.J.

*Departamento de Análisis Económico, Universidad de Valencia, Campus de Tarongers, Valencia, España.*

**Resumen:** La disponibilidad de información sobre la distribución de la población a gran escala, e independiente de las unidades administrativas de recogida de la información, es una necesidad en muchos ámbitos: evaluación de riesgos de terremotos, inundaciones o incendios, integración de estadísticas demográficas con información medioambiental en formatos diversos, diseño de políticas de prestación de servicios públicos o estudios de accesibilidad y movilidad en las grandes ciudades. Por estas razones, la literatura ha explorado numerosos métodos de desagregación de la población, recogida a nivel municipal o de sección censal, en polígonos urbanos procedentes de los sistemas de información de coberturas y usos del suelo o celdas de una malla regular, permitiendo de esta forma la generación de capas *raster* de población, fáciles de combinar con información medioambiental y geográfica. Una característica común a todos estos métodos de desagregación es que la información auxiliar utilizada no incorpora la altura de las edificaciones. De esta forma los algoritmos de desagregación no distinguen entre asentamientos extensos, basados en edificaciones con pocas alturas, frente a asentamientos compactos, con predominio de edificios con muchas alturas. Este trabajo examina la reducción del error que se comente al desagregar la población por secciones censales a una malla regular, de 1×1 km y de 1 ha, cuando incorporamos la tercera dimensión, la altura de los edificios a partir de datos LIDAR, sobre una capa *raster* de densidad de edificación, y utilizamos algoritmos de desagregación muy simples, basados en distribuciones proporcionales por área o por volumen, con y sin información auxiliar de coberturas y usos del suelo. Los resultados indican que las mejoras, incluso con estos algoritmos sencillos, son notables; comparables a las que se producen al incorporar información sobre usos del suelo a métodos dasimétricos binarios. En consecuencia incorporar la tercera dimensión a los procesos de re-escalado espacial de la población resulta totalmente esencial.

**Palabras clave:** rejillas de población, mapas dasimétricos, desagregación espacial, coberturas y usos del suelo, SIOSE, datos LIDAR.

## **A volumetric approach to spatial population disaggregation using a raster build-up layer, land use/land cover databases (SIOSE) and LIDAR remote sensing data**

**Abstract:** Availability of high resolution population distribution data, independent of the administrative units in which demographic statistics are collected, is a real necessity in many fields: risk evaluation due to earthquakes, flooding or fires, to name just a few, integration between socio-demographic and environmental or geographical information collected in different formats, policy design for the provision public services, such as health, education or public transport, or mobility studies in urban areas or metropolitan regions. Because of this, the literature has explored various methods of population downscaling, collected at communality or census tract level, into smaller areas; typically urban polygons from high resolution topographic maps or land use/land cover databases, or grid cells, allowing the elaboration of raster population layers. A common feature of all these methods is that they don't incorporate building height.

\* Autor para la correspondencia: [Francisco.J.Goerlich@uv.es](mailto:Francisco.J.Goerlich@uv.es)

In this way, downscaling methods don't distinguish between the urban sprawl type of settlement, where most of the houses are detached or semi-detached, and compact cities with high buildings. This paper examines error reduction in downscaling census tract population into 1×1 km and 1 ha grids, when we add the third dimension, building height from LIDAR remote sensing data. Algorithms used are simple, and based on areal weighting with or without auxiliary land use/land cover information, since our focus is not in fine turning algorithms, but in measuring improvements due to the missing dimension: building height. Our results indicate that improvements are noticeable. They are comparable to the ones obtained when we move from binary dasymetric methods to more general models combining densities for different land use/land cover types. Hence, adding the third dimension to population downscaling algorithms seems worth pursuing.

**Palabras clave:** population grids, dasymetric mapping, spatial disaggregation, land use/land cover, SIOSE, LIDAR data.

## 1. Introducción

La estadística demográfica se recoge a partir de la división administrativa de un país, regiones o municipios, de forma que es imposible conocer cuál es la distribución real de la misma sobre el territorio por debajo del área de recogida de la información. El supuesto de una distribución uniforme, implícita en los mapas de coropletas al dibujar densidades, es tan irreal como inútil desde el punto de vista del análisis espacial. Para muchos problemas prácticos es necesario conocer la distribución de la población sobre el territorio a una escala relativamente grande, por ejemplo en la determinación de población potencialmente afectada por áreas inundables, en el análisis de los flujos de movilidad en grandes ciudades o para el estudio de la accesibilidad de la población a determinados servicios públicos fundamentales, por señalar sólo algunos ejemplos. Sin embargo en la práctica no conocemos la distribución de la población por debajo del nivel de sección censal. La disponibilidad de estadísticas de población georeferencia a nivel de coordenada puntual está todavía lejos de ser una realidad, a pesar de los esfuerzos realizados por el Instituto Nacional de Estadística (INE) con ocasión del último censo de 2011.

Por esta razón, la literatura cartográfica ha explorado, desde hace tiempo, una serie de técnicas con las que ir acotando la localización de la población sobre el territorio, de forma que fuera posible situarla, de forma aproximada, sobre zonas relativamente homogéneas o dasimétricas generadas a partir de información auxiliar (Wright, 1936; Tobler, 1979; Langford y Unwin, 1994; Fisher y Langford, 1995; Eicher y Brewer, 2001; Mennis, 2003, 2009; Wu y Murray, 2005, Mennis y Hultgren, 2006; Langford, 2006, 2007). Con el

desarrollo de la teledetección, y la consiguiente generación de bases de datos de ocupación y usos del suelo, el énfasis se dirigió hacia el aumento en la calidad y resolución de la información auxiliar (Yuan *et al.*, 1997; Harvey, 2002; Wu *et al.*, 2005; Liu y Herold, 2006; Lu *et al.*, 2010; Kose, 2015), ya que es bien conocido que son estas características las más relevantes en la generación de resultados precisos, muy por encima de la elección de un algoritmo particular de desagregación (Martin *et al.*, 2000).

La irrupción de la teledetección en los procesos de desagregación espacial de la población también puso de manifiesto la utilidad de disponer de información demográfica en formato *raster* o de malla geográfica regular, en lugar de desagregarla sobre las unidades administrativas de partida. Por una parte, este formato permitía combinar la información demográfica con otra de carácter medioambiental, habitualmente recogida en este formato. Por otra parte, ello permitía la manipulación de ingentes cantidades de información, generándose incluso distribuciones mundiales de la población en formato malla geográfica de alta resolución (Tobler *et al.*, 1997; Dobson *et al.*, 2000; Balk y Yetman, 2004; Bhaduri *et al.*, 2007).

En el contexto europeo, la Agencia Europea del Medio Ambiente (AEMA) promovió la desagregación de la población a formato de malla regular con celdas de 1×1 km para el conjunto de Europa con la finalidad de combinar adecuadamente los procesos de interacción entre el hombre y el medio ambiente.<sup>1</sup> Se ensayaron diversos métodos de desagregación a partir de los datos de población

<sup>1</sup> El sistema de cuadrículas geográficas, como parte del conjunto de unidades estadísticas de representación de información, está normalizado a nivel Europeo (INSPIRE 2010).

municipal del censo de 2001 y *Corine Land Cover (CLC)* como información auxiliar (Gallego, 2010; Gallego, *et al.*, 2011). El resultado fue una malla de ámbito europeo que, aunque suponía una mejora respecto a la información demográfica del censo, era de calidad heterogénea. Sobreestimaba la población en ámbitos rurales y la subestimaba en ámbitos urbanos. La razón había que buscarla en la escasa resolución tanto de la información demográfica de partida como de la información auxiliar. Por esta razón los esfuerzos siguientes se dirigieron hacia la mejora de la resolución, tanto de la información auxiliar (Batista e Silva *et al.*, 2013b), como de la información demográfica (Goerlich y Cantarino, 2012). Ello permitió obtener una nueva malla europea, con fecha de referencia 2006 (Batista e Silva *et al.*, 2013a), que ya ha sido utilizada para el diseño de políticas en el contexto europeo, por ejemplo la delimitación de ciudades del proyecto *Urban Audit* (Dijkstra y Poelman, 2012). En esta línea, la oficina de estadística de la Unión Europea, *Eurostat*, se propuso la elaboración de una malla derivada del censo de 2011 que, en la medida de lo posible, debería construirse a partir de población geo-referenciada a nivel de entidades puntuales, y por tanto representar la verdadera distribución de la población a la escala de la cuadrícula representada (Goerlich y Cantarino, 2016). La inexistencia de población georeferenciada para muchos países europeos ha hecho, sin embargo, que los procedimientos de desagregación espacial, utilizando información auxiliar, no se hayan podido evitar en muchos casos. Esta información ha procedido, en este caso, de fuentes muy diversas: una versión refinada de *CLC*, capas *raster* de alta resolución de sellado del suelo e incluso información topográfica digital de libre acceso como *Open Street Maps*.<sup>2</sup>

Una característica importante de estos procesos de desagregación es que la totalidad de la información auxiliar utilizada es bidimensional: *CLC*, sellado del suelo, información topográfica, proporción de superficie edificada. Tan sólo el

ejercicio de Goerlich y Cantarino (2012), que utiliza el Sistema de Información del Suelo en España (SIOSE) como información auxiliar sobre coberturas, emplea el atributo de edificación, permitiendo distinguir de esta forma viviendas unifamiliares y edificios, y recoger de forma burda aspectos relacionados con la volumetría de las construcciones donde se asigna la población.

La ausencia de información sobre la altura de las áreas construidas ha sido señalada como una debilidad de los métodos de desagregación espacial de la población (Batista e Silva, 2013c), ya que la relación entre volúmenes edificados y volúmenes de población, además de intuitiva, ha sido reiteradamente mostrada por la literatura (Ramesh, 2009; Lwin y Murayama, 2009, 2011; Qiu *et al.*, 2010; Yu *et al.*, 2010; Lu *et al.*, 2011; Hermosilla *et al.*, 2013). Si la información sobre alturas no ha sido considerada con generalidad en los algoritmos de desagregación espacial ha sido debido a la falta de información. La disponibilidad reciente de datos LIDAR (*Light Detection and Ranging*) para áreas extensas puede cambiar esta situación en un futuro próximo, de forma que los algoritmos basados en interpolación por área puedan ser adaptados de forma conveniente a algoritmos de interpolación por volúmenes, lo que no excluye la utilización de otra información auxiliar sobre coberturas y usos de los edificios (Rottensteiner y Jansa, 2002; Rottensteiner, 2003; Kaufholz, 2004; Kim *et al.*, 2006).

El objetivo de este trabajo es evaluar, en el contexto del esfuerzo europeo por desagregar la población a un formato de malla estándar, las mejoras que se obtienen al introducir información sobre altura de los edificios obtenidos a partir de datos LIDAR. Para conseguir dicho objetivo se seleccionará un área relativamente grande, que incluya zonas altamente pobladas con otras deshabitadas, y con diferente estructura en las alturas de las edificaciones. Dichas alturas se asignarán sobre una capa *raster* de porcentaje de superficie edificada, desagregándose la población por métodos sencillos, lo que permitirá examinar la reducción de errores cuando se considera la altura de los edificios, frente a cuando no se considera

La estructura del trabajo se expone a continuación. El apartado siguiente describe la zona de estudio, el material y los métodos empleados; así como los datos de validación utilizados. A continuación se

<sup>2</sup> Las cuadrículas poblacionales europeas de 2006 y 2011 están disponibles en la web de *Eurostat* (<http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/population-distribution-demography>). La cuadrícula de 2001 está disponible en la web de la AEMA (<http://www.eea.europa.eu/data-and-maps/data/population-density-disaggregated-with-corine-land-cover-2000-2>).

describen los resultados obtenidos y un apartado final concluye brevemente.

## 2. Zona de estudio, material y métodos

### 2.1. Zona de estudio

Como zona de estudio consideramos la Comunidad Autónoma de Madrid, unos 8000 km<sup>2</sup>, con una densidad media de algo más de 800 hab/km<sup>2</sup>, muy por encima del promedio nacional que no alcanza los 100 hab/km<sup>2</sup>, al tratarse de un área altamente urbanizada. Está constituida por 179 municipios, con poblaciones que oscilan entre los 53 y los más de 3 millones de residentes, y 4275 secciones censales, la mitad de las cuales, 2409, pertenecen al municipio de Madrid.

La población a desagregar es la del Padrón, con fecha de referencia 1 de enero de 2012, unos 6,5 millones de habitantes, la mitad de los cuales residen en el municipio de Madrid. Se trata por tanto de la población residente, o lo que la literatura conoce como ‘población de noche’, por contraposición a la denominada ‘población de día’, o por lugar de trabajo. El sistema zonal de origen está constituido por las secciones censales, que es la unidad más pequeña para la que se recoge la información demográfica en España.<sup>3</sup> Dicha información está disponible públicamente en la *web* del INE. Los contornos de dichas secciones censales se obtuvieron, en formato *shape*, de la *web* del Instituto de Estadística de Madrid (IEM).

### 2.2. Material

El objetivo final es desagregar dicha población por secciones censales a la cuadrícula estándar Europea, con resoluciones de 1 ha y de 1×1 km, en sistema geodésico de referencia ETRS89 y proyección *Lambert Azimutal Equal Area (LAEA)*. Así pues las secciones censales constituyen el sistema zonal de origen, mientras que las mallas geográficas constituyen el sistema zonal de destino. El sistema zonal de origen está constituido

por 4.275 unidades espaciales, mientras que los sistemas zonales de destino están constituidos por prácticamente el doble de unidades en el caso de la malla de 1×1 km, y por casi un millón de unidades en el caso de la de 1 ha de resolución. Muchas de estas unidades están deshabitadas. Como sistema zonal intermedio operativo se utilizará un pixel de 10 m, ya que esta es la resolución de una capa de información auxiliar esencial en el proceso de desagregación. Las cuadrículas que constituyen el sistema zonal de destino, 1 ha y 1×1 km, se obtendrán por agregación a partir de esta resolución.

Como información auxiliar en el proceso de desagregación consideramos tres tipos de datos. En primer lugar, una capa *raster* de densidad de edificación (*built-up area*) elaborada por el *Joint Research Centre (JRC)* de la Unión Europea (Pesaresi *et al.*, 2013; Ferri *et al.*, 2014; Florczyk *et al.*, 2015) en el contexto de un proyecto global de determinación de los asentamientos urbanos a nivel mundial, y que se conoce como la *Global Human Settlement Layer (GHSL)*. A escala europea este producto se distribuye como EU-GHSL o *European Settlement Map (ESM)*, está disponible con una resolución de 10 m pixel, y sistema de referencia ETRS89–LAEA.<sup>4</sup> Su origen procede de la agregación en promedio de capas *raster* de 2,5 m de resolución obtenidas mediante procesos de extracción automática de información de las imágenes de los satélites SPOT5 y SPOT6 a dicha resolución. A partir de aquí el método *GHSL* utiliza técnicas de aprendizaje automático (*machine learning*) para entender las relaciones sistemáticas entre las características morfológicas y de textura, extraídas de la bandas multispectral y pancromática (si está disponible), que describen los asentamientos humanos. Los conjuntos de datos utilizados en el aprendizaje automático incluyen *CLC* e información sobre el sellado del suelo. Adicionalmente, información vectorial sobre presencia de edificaciones, por ejemplo información catastral, fue incorporada mediante técnicas de fusión de información cuando ello fue

<sup>3</sup> Por debajo de municipio se dispone, además de las secciones censales, de información demográfica para las unidades poblacionales del nomenclátor, que sin embargo no disponen de territorio asociado y no están georreferenciadas con generalidad.

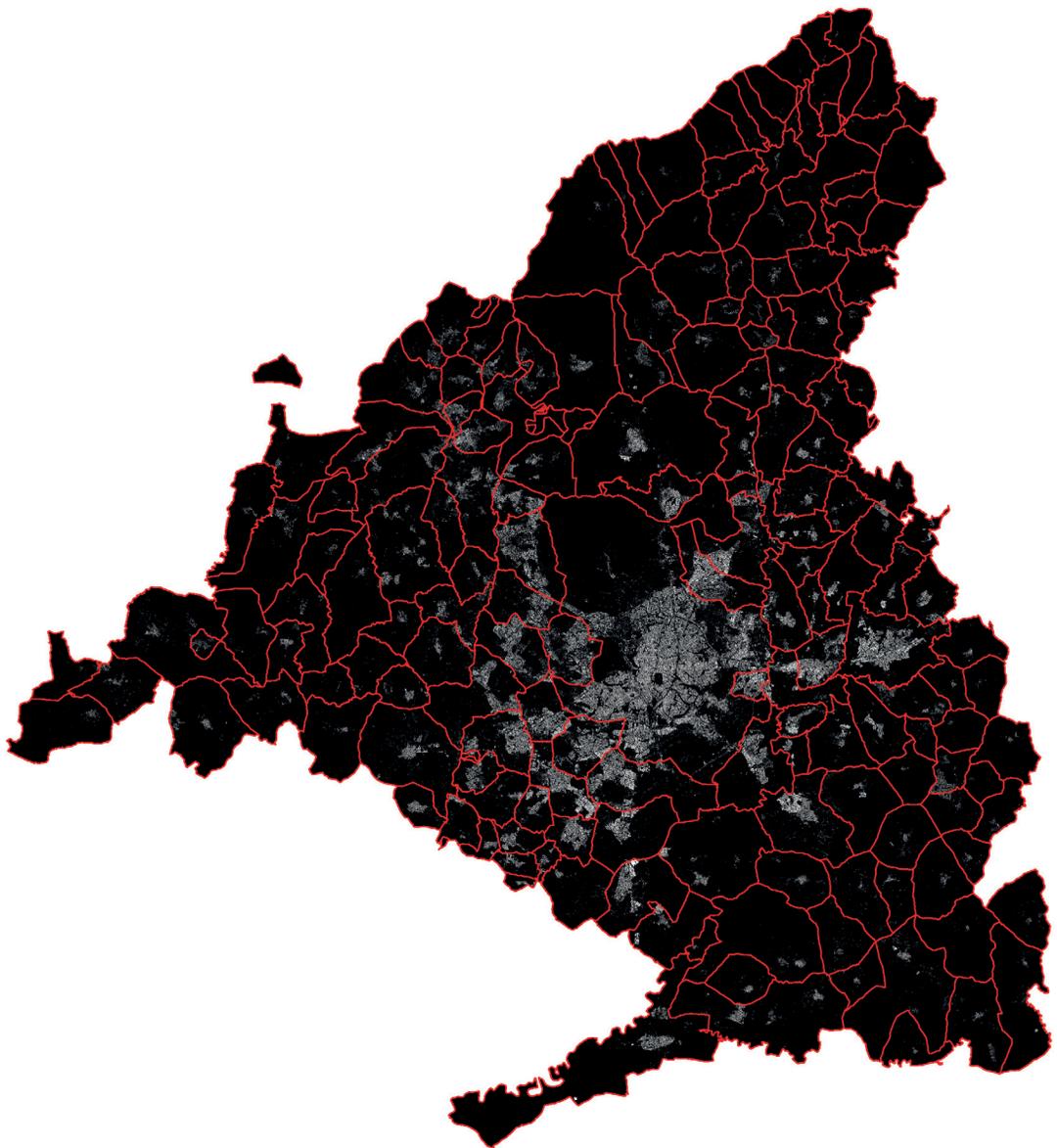
<sup>4</sup> La última versión de esta información, *ESM2016*, es actualmente pública a resolución local, 10 m pixel, en <http://land.copernicus.eu/pan-european/GHSL/view>. La capa utilizada en este trabajo, con la misma resolución, fue la versión anterior, *ESM2014*, cuya fecha de referencia es la misma, y que fue amablemente suministrada por Matina Halkia del *JRC*, para cuestiones de investigación.

posible. Una validación sistemática del producto final, frente a la información contenida en LUCAS 2012 (*Land Use and Cover Area frame Survey*), arrojó una precisión superior al 95% para la clase área urbana.

Cada píxel de dicho *raster* representa el porcentaje de edificación contenido en el mismo, lo que dado el tamaño del píxel son los metros cuadrados edificados. Esta capa representa el soporte potencial

de la población, en el sentido de que en píxeles con valor cero no habrá población, aunque en celdas con edificios si puede estimarse población nula. La Figura 1 ofrece una imagen de la zona de estudio, con los contornos administrativos a nivel municipal y la capa *raster* del *ESM* como fondo.

La Figura 2 ofrece un detalle urbano de parte de la zona de estudio, el Parque del Retiro, la M-30 y la estación Puerta de Atocha, del municipio de



**Figura 1.** Zona de estudio – Madrid. Límites municipales y capa *raster* de densidad de edificación, *European Settlement Map*, con resolución 10 m píxel.



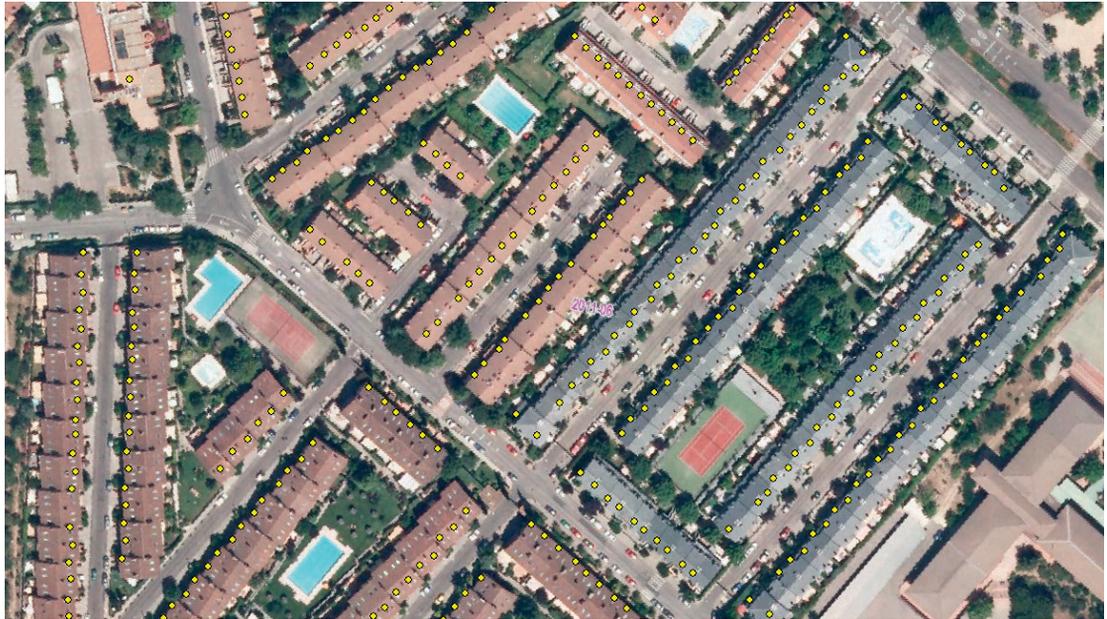
**Figura 2.** Detalle de la zona de estudio – Madrid – Zona del Parque del Retiro. Capa *raster* de densidad de edificación, *European Settlement Map*, con resolución 10 m pixel, y ortofoto del PNOA como fondo.

Madrid. En ella se aprecia, a escala muy detallada, con un fondo de ortofoto, como la información de la *ESM* filtra adecuadamente los viales y otras superficies de suelo sellado que no son edificaciones.

En segundo lugar, consideramos como información adicional sobre coberturas y usos del suelo SIOSE 2011. La principal característica de SIOSE es que se trata de un modelo de datos orientado a objetos, no es por tanto un modelo jerárquico en el que cada polígono es clasificado dentro de un conjunto de clases limitado y perfectamente definido a priori, sino que para cada polígono se dispone de una leyenda que describe sus contenidos de acuerdo con una reglas pre-determinadas, y un conjunto inicial de 40 coberturas simples. Estas clases pueden combinarse con sus diferentes porcentajes de ocupación y atributos, formando coberturas compuestas, ya sea de forma arbitraria o pre-definidas dentro del modelo de datos SIOSE (SIOSE 2011). SIOSE se distribuye en formato vectorial por el Instituto Geográfico Nacional (IGN) a través del centro de descargas del Centro Nacional de Información Geográfica (CNIG). SIOSE será utilizado de forma sencilla, simplemente como

máscara para filtrar sobre *ESM* celdas cuya clase sea no residencial, a los cuales no hay que asignarles población. Otras aplicaciones más elaboradas son posibles (Goerlich y Cantarino, 2012), en particular una extensión natural en este contexto es introducir ponderaciones diferenciadas por tipo de cobertura, que reflejan densidades de población diferentes según uso del suelo (Mennis, 2003; Mennis y Hulgren, 2006; Freire *et al.*, 2015), sin embargo, dado que el objeto del ejercicio es aislar los beneficios potenciales de incorporar la altura de los edificios en la desagregación espacial de la población nos mantendremos en la utilización de métodos sencillos.

En tercer lugar, la información sobre alturas de los edificios procede del tratamiento de los datos LIDAR del Plan Nacional de Ortofotografía Aérea (PNOA). Dicha información es obtenida mediante sensores LIDAR aerotransportados, con una densidad nominal de 0,5 puntos/m<sup>2</sup>, y un valor de RMSE en la coordenada Z inferior a 20 cm. Siendo las coordenadas Z altitudes ortométricas. La nube de puntos ha sido clasificada, tras su captura, mediante procedimientos automáticos. Los datos se



**Figura 3.** Área de Las Rozas – Madrid. Muestra del Padrón 2012 georreferenciado a nivel de entidades puntuales (Aproximación Postal Principal).

distribuyen a través del CNIG en ficheros digitales de  $2 \times 2$  km de extensión. El formato de descarga es LAZ, versión comprimida del formato estándar LAS. Dicho formato incluye la clasificación definida por la Sociedad Americana de Fotogrametría y Teledetección (ASPRS) e incluyen las dos clases que nos son de interés: suelo o terreno y edificios. Estas dos clases serán las únicas utilizadas en este ejercicio, tal y como se describe en la sección siguiente. El IGN ofrece una herramienta de descompresión y visualización, pero no de tratamiento de la información. El sistema geodésico de referencia es ETRS89 y proyección UTM en el huso 30N.

### 2.3. Validación

Como conjunto de datos de validación utilizamos el Padrón de la Comunidad de Madrid georreferenciado a nivel de coordenada puntual en un 99,4% de su población. La Figura 3 ofrece una muestra de dicho conjunto de datos, donde los puntos amarillos representan las coordenadas de la población residente. Dicha información no es pública, y fue suministrada por el Instituto de Estadística de la Comunidad de Madrid (IEM) a efectos de validación del ejercicio de desagregación. El origen de

la misma es la georreferenciación de los domicilios postales en las inscripciones del padrón, y es correcta para toda la población a la que puede atribuírsele un domicilio normalizado. De esta forma, para la zona de estudio conocemos la verdadera distribución de la población al nivel de agregación que deseemos, que no tiene por qué coincidir con los contornos administrativos de recogida de la información demográfica. Una consulta espacial punto a polígono determinó que el número de celdas habitadas en la malla de 1 ha era de 56.920; mientras que el número de celdas habitadas en la malla de  $1 \times 1$  km era de 2.449. Estas cuadrículas, construidas de abajo arriba (*bottom-up*), representan la verdadera distribución de la población respecto a la que mediremos la bondad de nuestros métodos de desagregación espacial, que construyen cuadrículas de arriba abajo (*top-down*).

La Tabla 1 muestra un resumen de los diferentes conjuntos de datos utilizados en el trabajo, el papel desarrollado por cada uno de ellos, así como las fechas de referencia.

**Tabla 1.** Datos utilizados, descripción y fechas de referencia.

Categoría	Descripción	Año de referencia
Sistema zonal de origen	Registro de población: Padrón	
	Población residente	Secciones Censales
Sistema zonal de destino	Grids Europeas	1 ha 1 km
Sistema zonal intermedio operativo	ESM (EUGHSL) a escala local: píxel 10 m	2011
Información auxiliar	SIOSE: Sistema de Información de Ocupación del Suelo en España	
	Capa de altura de edificios con la resolución de ESM elaborada a partir de datos LIDAR	2010
Datos de validación	Padrón para la Comunidad de Madrid georreferenciado a nivel de coordenada puntual	2012, 1 de enero

## 2.4. Métodos

Los métodos son variaciones simples de una distribución de la población mediante interpolación según el área edificada (*areal weighting*). El primer paso consiste en hacer compatibles los contornos vectoriales de las secciones censales con la capa *raster* de densidad de edificaciones (*ESM*). Para ello, y dada la resolución del *ESM*, 10 m píxel, se transformó a *raster* el fichero vectorial de secciones censales ajustándolo al *ESM*. Ello permite una aplicación muy sencilla de los métodos que se exponen a continuación.

Sea  $B_{i,c}$  el área edificada (*built-up area*) en el píxel  $i$  de la sección censal  $c$ , que viene dada por la capa *raster* de densidad de edificación, *ESM*, la población asignada a dicho píxel se obtiene como

$$P_{i,c} = P_c \times \frac{B_{i,c}}{B_c} \quad (1)$$

donde  $P_c$  es la población de la sección censal  $c$ , y  $B_c$  es el área edificada en dicha sección censal,  $B_c = \sum_{i=1}^n B_{i,c}$ . Por construcción, (1) verifica la restricción de agregación de la población a nivel de sección censal (Tobler, 1979), ya que es esta la resolución de los datos de partida.

La desagregación inicial se realiza a nivel de píxel de 10 m, que es el sistema zonal intermedio

operativo, para luego agregar los resultados a nivel de celda de 1 ha y de 1×1 km.

El flujo general de trabajo se ofrece en la Figura 4, donde los diferentes métodos difieren en la cantidad de información utilizada, pero todos son variación sobre (1). Los cálculos se implementaron en ArcGIS 10.2 (ESRI 2013).

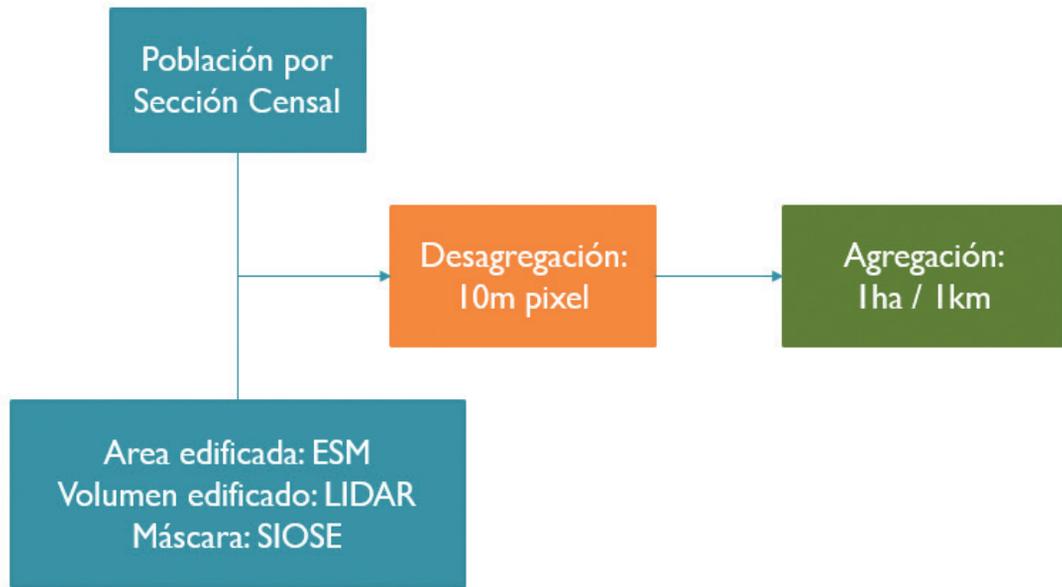
Los métodos que utilizan información de SIOSE simplemente filtran determinados píxeles donde se asume que, aunque hay edificios estos no albergan población. Por otra parte, la incorporación de la altura de los edificios permite sustituir  $B_{i,c}$  en (1) por un volumen,  $V_{i,c}$ , determinado a partir de *ESM* y la información LIDAR, constituyendo de esta forma un reparto por volúmenes (*built-up volume*), y no por áreas. Así pues, en este caso la fórmula (1) se transforma en

$$P_{i,c} = P_c \times \frac{V_{i,c}}{V_c} \quad (2)$$

Donde  $V_{i,c}$  es el volumen edificado en el píxel  $i$  de la sección censal  $c$  y  $V_c$  es el volumen edificado en dicha sección censal,  $V_c = \sum_{i=1}^n V_{i,c}$ .

En concreto los 4 métodos ensayados son los siguientes:

- Distribución por área edificada a partir de *ESM* únicamente como información auxiliar.



**Figura 4.** Esquema genérico del proceso de desagregación de la población.

- ii) Selección de coberturas de SIOSE con edificios residenciales, filtrado de *ESM* por estas coberturas y distribución por área edificada.

Los otros dos métodos utilizan la desagregación por volúmenes, fórmula (2), en lugar de la desagregación por áreas.

- iii) Distribución por volumen edificada a partir de *ESM* como soporte edificada y una estimación de la altura media por pixel a partir de datos LIDAR.
- iv) Filtrado de *ESM* con las coberturas de SIOSE seleccionadas en (ii), y distribución por volumen edificada.

Otros métodos más sofisticados son posibles (Mennis, 2003; Wu *et al.*, 2005; Mennis y Hulten, 2006; Liu, y Herold, 2006; Goerlich y Cantarino, 2012), pero el énfasis del trabajo consiste en determinar la reducción del error en la desagregación asociado a la sustitución de  $B_{i,c}$  en (1) por  $V_{i,c}$  en (2), esto es, las ganancias asociadas a la incorporación en el proceso de las alturas de edificios.

Una vez realizada la distribución de la población por cada uno de los métodos anteriores calculamos el error relativo respecto a la malla de referencia obtenida para el conjunto de validación, normalizado al intervalo 0-1, y expresado en tanto por cien:

$$\delta = \frac{\sum_{j=1}^J |P_j - P_j^{ref}|}{2 \times \sum_{j=1}^J P_j} \quad (3)$$

donde  $P_j$  es la población estimada para la celda  $j$ , y  $P_j^{ref}$  es la población en la misma celda para la malla de referencia. De esta forma  $\delta$  puede interpretarse como el porcentaje de población que es localizada de forma incorrecta por el método de desagregación, dada la resolución de la cuadrícula. Al margen de este estadístico, como medida del error total en el proceso de desagregación, también es útil examinar el número de falsos positivos: celdas donde el algoritmo sitúa población pero en realidad no la hay,  $P_j > 0$  y  $P_j^{ref} = 0$ , y falsos negativos: celdas donde el algoritmo no sitúa población pero en realidad si la hay,  $P_j = 0$  y  $P_j^{ref} > 0$ .

El sistema de referencia geodésico y la proyección de trabajo son ETRS89-LAEA, de forma que toda la información que no estaba inicialmente en dicho sistema se re-proyectó al mismo. El formato de trabajo es *raster*, lo que viene determinado por la resolución del *ESM*, así como por el hecho de que el sistema zonal de destino

sea en formato de malla geográfica regular.<sup>5</sup> Por esta razón los contornos de las secciones censales se convirtieron a formato *raster* con resolución de 10 m, y ajustados al *ESM*, lo que permite una aplicación directa del método (i), desagregación por área edificada, a partir de (1).

Para el filtrado de *ESM* según coberturas residenciales se extrajeron de SIOSE todos aquellos polígonos con edificaciones residenciales, es decir, que contuvieran la cobertura simple *edificio* con atributos de edificio aislado, entre medianeras, o viviendas unifamiliares, ya sean aisladas o adosadas; y que además estuvieran dentro de coberturas compuestas predefinidas de tipo urbano mixto –casco, ensanche o discontinuo–, asentamiento agrícola residencial o huerta familiar, así como algunas coberturas compuestas predefinidas de tipo terciario o de equipamiento y dotacional. Se excluye así a los polígonos de infraestructuras de la posibilidad de albergar

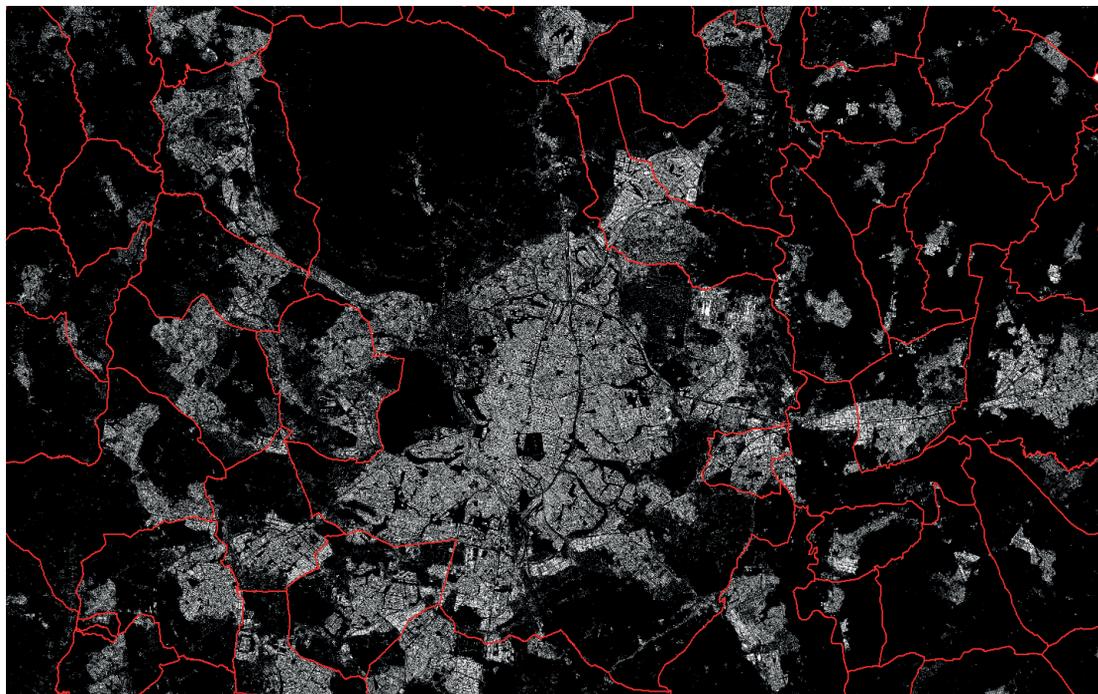
población residente, aun cuando dispongan de edificios residenciales.<sup>6</sup>

La elección de las coberturas susceptibles de albergar población es una de las fuentes de error en todo proceso de desagregación espacial que utilice este tipo de información. Su elección implica siempre un equilibrio entre dispersar la población en exceso, falsos positivos, o no asignarla a algunas coberturas donde existe población residente, falsos negativos (Goerlich y Cantarino, 2016).

La aproximación a la selección de dichas coberturas fue en gran parte empírica. Por una parte, manteniendo siempre los polígonos urbanos se probaron diversas coberturas artificiales no urbanas para a continuación examinar el error en la desagregación. Por otra parte, el análisis del Padrón georreferenciado nos permite saber en qué coberturas existe población empadronada. Ello nos proporcionó una indicación sobre la

<sup>5</sup> En el contexto de INSPIRE (2010) las mallas de población se implementan, sin embargo, en formato vectorial.

<sup>6</sup> En concreto, entre las coberturas compuestas susceptibles de soportar población residente se incluyen las de tipo terciario: comercial y oficinas, complejo hotelero y camping; y las de equipamiento y dotacional: administrativo institucional, sanitario, educativo, penitenciario y religioso.



**Figura 5.** Área edificada: Municipio de Madrid y alrededores.

conveniencia de incluir las coberturas de tipo terciario y equipamiento, por ejemplo existe población empadronada en complejos religiosos o penitenciarios, al tiempo que debíamos excluir las de infraestructuras.

Una vez seleccionados los polígonos de SIOSE susceptibles de albergar población residente, estos se convirtieron a *raster* con resolución de 10 m, ajustados al *ESM*, y se utilizaron como máscara sobre *ESM*. Esta operación reduce los píxeles sobre los que asignamos población en un 48%, y permite una aplicación directa del método (ii). La población se distribuye de nuevo utilizando (1) sobre esta nueva capa, que simplemente excluye ciertos píxeles del cálculo.

Para incorporar la altura media de los edificios a los métodos anteriores utilizamos los datos LIDAR para la zona de estudio. El proceso utiliza la clasificación de puntos incluida en los datos LIDAR del PNOA, y emplea solo las dos clases mencionadas anteriormente: suelo, código 2, y edificaciones, código 6. A partir de estos puntos se construye un modelo digital de superficies normalizado, *nDSM*, en 3 pasos, utilizando para ello las herramientas de geoprocesamiento LIDAR

contenidas en ArcGIS (ESRI 2013): (1) A partir de los puntos de suelo en los datos LIDAR del PNOA se generó un modelo digital del terreno, DTM. (2) A partir de los puntos clasificados como edificación en dichos datos LIDAR se generó un modelo digital de superficies, DSM. En ambos casos, el *raster* generado se ajustó al *ESM* y su resolución, 10 m pixel, tomando las cotas mínimas para el DTM y máximas para para el DSM. (3) El *nDSM* se generó por diferencia ( $nDSM = DSM - DTM$ ) (Figuras 5 y 6).

Al igual que sucede con el algoritmo de desagregación, o la selección de coberturas SIOSE capaces de soportar población residente, la incorporación de volúmenes edificados a través de información LIDAR es susceptible de mejora (Hermosilla *et al.*, 2011), lo que sin duda fortalecerá las conclusiones del trabajo en términos de reducción del error en los procesos de desagregación espacial de la población. El ajuste de la información LIDAR para mejorar estos procesos es un área activa de investigación actual en diferentes contextos.

Una vez calculada la altura media de los edificios por pixel del *ESM* se calculó una capa *raster* de volumen edificado por pixel,  $V_{i,c}$ , que sustituye a

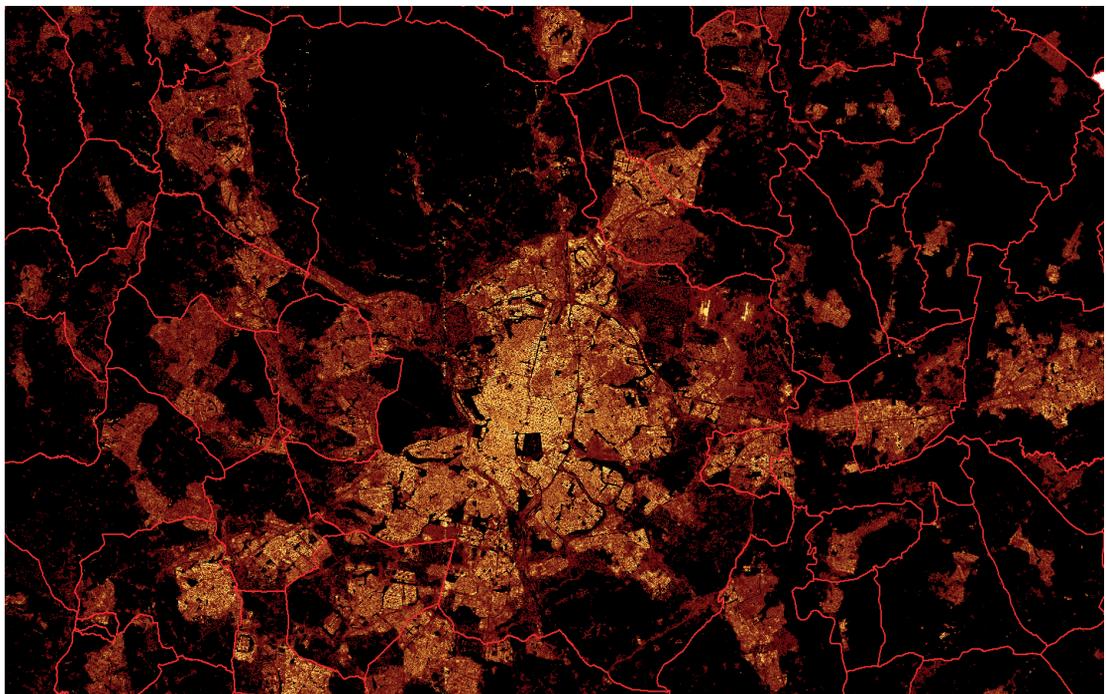


Figura 6. Volumen edificado: Municipio de Madrid y alrededores.

$B_{i,c}$  en (1) en el proceso de desagregación, lo que genera la fórmula (2). Esto constituye el método (iii), que desagrega la población sobre los mismos píxeles que el método (i), y en consecuencia cualquier mejora se debe a la estimación de la altura de los edificios.

En el método (iv) las coberturas de SIOSE susceptibles de albergar población actúan como máscara sobre la capa de volúmenes edificadas, y la distribución se hace sobre la capa resultante. Este método distribuye la población sobre los mismos píxeles que el (ii), y en consecuencia cualquier mejora es atribuible a la estimación de la altura de los edificios.

### 3. Resultados

La Tabla 2 ofrece los resultados de aplicar los 4 métodos de desagregación de la población a nivel de sección censal, expuestos en la sección anterior para dos resoluciones diferentes: la malla de 1 ha, y la de 1×1 km, que constituye el estándar europeo. Las dos primeras columnas ofrecen el error relativo, (3), mientras que las otras tres indican el número de celdas habitadas estimadas y su distribución entre aciertos y falsos positivos.

En todos los casos se produce una dispersión excesiva de la población, si comparamos las celdas que se estiman como habitadas y las que se derivan del Padrón geo-referenciado. Los resultados cualitativos son similares para ambas resoluciones. Tanto

la utilización de información sobre coberturas de SIOSE, como la derivada de la estimación de los volúmenes edificadas a partir de datos LIDAR mejoran la distribución en un orden de magnitud relativo similar. Sin embargo, los resultados cuantitativos son diferentes y merecen algunos comentarios.

Para la resolución de 1 ha el error relativo es excesivo en todos los casos. Para el método más sencillo, distribución por áreas a partir de *ESM*, (i), el error supera ligeramente el 30%, y las celdas estimadas con población triplican las reales. En consecuencia, el número de falsos positivos es muy elevado, casi un 70%. Sustituyendo los volúmenes por las áreas proporciona la misma dispersión excesiva de la población, método (iii), pero reduce el error en alrededor de 4 puntos porcentuales. Los edificios residenciales tienen, en promedio, más altura que los industriales y de infraestructuras, de forma que este ajuste mejora notablemente la localización de la población. Las ganancias no deben ser uniformes a lo largo del territorio, sino que por el contrario deben ser mayores en zonas donde la población reside en edificios elevados, ofreciendo cierto contraste con los edificios no residenciales.

Para evitar situar a la población en este tipo de edificación no residencial es necesario acudir a la información sobre coberturas del suelo, SIOSE en nuestro caso. Ello consigue reducir el número de celdas habitadas a menos de la mitad y, en

**Tabla 2.** Evaluación de los diferentes métodos de desagregación según la resolución de la malla.

	Error relativo ( $\delta$ )		Celdas habitadas		
	Distribución por área	Distribución por volúmenes	Totales	Aciertos	Falsos positivos
Resolución de la malla: 1 ha					
Sin información sobre coberturas	31,1%	27,7%	178.947	31,1%	68,9%
Con información sobre coberturas	26,0%	22,9%	83.933	64,2%	35,8%
Celdas habitadas procedentes del Padrón georeferenciado				56920	
Resolución de la malla: 1 km					
Sin información sobre coberturas	7,5%	6,7%	5.734	41,7%	58,3%
Con información sobre coberturas	4,8%	4,1%	2.782	77,3%	22,7%
Celdas habitadas procedentes del Padrón georeferenciado				2.449	

consecuencia, el número de falsos positivos disminuyen hasta alrededor de 1/3. La distribución por áreas a partir de *ESM*, utilizando como máscara las coberturas seleccionadas de SIOSE, –método (ii)– reducen el error en alrededor de 5 puntos porcentuales, y la consideración de volúmenes –método (iv)– en otros 3 puntos adicionales.

En conjunto la información sobre coberturas del suelo, junto con la información de alturas edificadas es capaz de reducir el error en la desagregación desde un 31,1% hasta un 22,9%. Una reducción significativa, aunque todavía un margen de error inaceptablemente alto.<sup>7</sup> Resulta interesante comprobar como la información sobre coberturas y la de alturas de los edificios es complementaria, y su contribución a la reducción del error es aproximadamente aditiva.

Para la resolución de 1×1 km el error relativo es sustancialmente menor, tan solo de un 7,5% para la distribución por áreas, que al igual que en el caso anterior tiende a dispersar la población en exceso. El número de celdas habitadas estimadas es más del doble de las obtenidas a partir del Padrón georreferenciado, y el número de falsos positivos de casi un 60%. La incorporación de la información por volúmenes reduce el error relativo en aproximadamente 1 punto porcentual.

Para acotar las celdas donde localizamos población, la incorporación de la información sobre coberturas del suelo es esencial. En este caso el número de celdas habitadas estimadas es solo ligeramente superior a las reales, y el número de falsos positivos se reducen a algo más de un 20%. La incorporación de los volúmenes edificadas reduce el error relativo en casi otro punto porcentual adicional. A este nivel de resolución, 1×1 km, el error relativo es comparable, o incluso ligeramente inferior en el caso del método (iv), al obtenido con métodos de desagregación mucho más sofisticados (Goerlich y Cantarino, 2012). Un error muy reducido dada la simplicidad de los

métodos, lo que indica que todavía existe margen para la mejora considerando modelizaciones más complejas.

#### 4. Conclusiones

Los resultados de este trabajo destacan la importancia de incorporar los volúmenes edificados como información auxiliar en los procesos de desagregación espacial de la población, así como la posibilidad práctica de hacerlo para grandes áreas a partir de la información LIDAR, actualmente disponible para todo el territorio nacional.

Los métodos ensayados confirman una importante reducción del error relativo en algoritmos de distribución relativamente simples. Si la información sobre coberturas es útil en acotar el dominio sobre el que se distribuye la población, la altura de las edificaciones proporciona heterogeneidad, mejorando de forma aproximadamente aditiva los resultados. En este sentido ambos tipos de información son claramente complementarias.

Los errores son muy reducidos para la resolución de 1×1 km de tamaño de celda, comparables a los de otros métodos más sofisticados, pero todavía excesivamente elevados para la malla de 1 ha. Sin embargo, dada la simplicidad de nuestros métodos, es de esperar que sean posibles mejoras relevantes mediante la utilización de algoritmos más sofisticados. Por ejemplo, la introducción de ponderaciones diferentes según el tipo de cobertura (Mennis, 2003; Mennis y Hultgren, 2006) debería mejorar de forma sustancial la bondad de las estimaciones, especialmente para la malla de 1 ha.

Por la misma razón, un tratamiento más preciso de los datos LIDAR, para incorporar información sobre alturas de los edificios, debería mejorar los resultados sobre la generación de cuadrículas de población a nivel europeo, cuyo objetivo es actualmente 1 ha de tamaño de celda. (Halkia, 2014; Halkia y Freire, 2015).

Aunque los métodos aplicados en este trabajo han utilizado una malla como sistema zonal de destino, nuestros resultados son extensibles a otros sistemas zonales, como los polígonos de *Urban Atlas* (Batista e Silva, 2013c), los núcleos urbanos de los mapas topográficos, o las manzanas y parcelas de la cartografía catastral (Santos Preciado,

<sup>7</sup> El margen de error para el método (iv) está, sin embargo, en muchos casos por debajo del reportado por Gallego (2010) para la desagregación de la población municipal del censo de 2001 a una resolución de 1×1 km de celda, y utilizando solamente *CLC* como información auxiliar. El hecho de que consigamos errores similares para una resolución de 1 ha de tamaño de celda se debe, en gran medida, a la mayor resolución de la información de partida.

2015).<sup>8</sup> El argumento esencial es que la altura de los edificios es una variable relevante en los procesos de desagregación, que reduce el error de forma significativa más allá de la información sobre coberturas del suelo, y esto es independiente tanto del sistema zonal de destino, así como del concepto de población manejado. Aunque en nuestro ejercicio hemos considerado a la población residente, mejoras similares deberían esperarse en la desagregación de la población por lugar de trabajo (Batista e Silva, 2015; Martin, 2015).

## Agradecimientos

Este trabajo ha evolucionado a partir de la presentación *A volumetric approach to population disaggregation using LIDAR remote sensing data*, en el *European Forum for Geography and Statistics (EFGS) Conference 2015*, 10-12 November, Vienna, Austria. El autor agradece los comentarios de los participantes en dicho foro, y en particular los Matina Halkia (JRC), que facilitó la capa *raster* del *European Settlement Map (ESM)* a escala local (10 m pixel). Se agradece también la disponibilidad a facilitar información para la Comunidad de Madrid a Ángel Sanchez (IEM), Dolores Nuñez (IEM) y Rosario Arenas (IEM). Una parte importante de este trabajo no hubiera sido posible sin los datos del *ESM* y de la Comunidad de Madrid que amablemente pusieron a nuestra disposición. Comentarios de dos revisores anónimos contribuyeron a una importante mejora sobre la redacción inicial del trabajo.

El autor agradece la ayuda del programa de Investigación FBBVA-Ivie, así como la del proyecto ECO2015-70632-R. Resultados mencionados en el texto pero no ofrecidos están disponibles si se solicitan al autor.

## Referencias

Bhaduri, B., Bright, E., Coleman, P., Urban, M. L. 2007. LandScan USA: A high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal*, 69(1-2), 103-117. <http://dx.doi.org/10.1007/s10708-007-9105-9>

Balk, D., Yetman, G. 2004. The global distribution of population: Evaluating the gains in resolution refinement. *Center for International Earth Science Information Network (CIESIN), Columbia University*. New York, USA. Accesible en: [http://sedac.ciesin.columbia.edu/downloads/docs/gpw-v3/gpw3\\_documentation\\_final.pdf](http://sedac.ciesin.columbia.edu/downloads/docs/gpw-v3/gpw3_documentation_final.pdf) (Acceso: junio/2016).

Batista e Silva, F. 2015. The ENACT project: towards spatiotemporal activity and population mapping in Europe. En: *European Forum for Geography and Statistics*. Vienna: 10-12 noviembre. Accesible en: [http://cpanel53.proisp.no/~efgsiowd/wp-content/uploads/conferences/efgs/2015/Conference\\_EFGS2015\\_1110\\_7\\_7\\_ENACT-Filipe-BatistaPresentation.pptx](http://cpanel53.proisp.no/~efgsiowd/wp-content/uploads/conferences/efgs/2015/Conference_EFGS2015_1110_7_7_ENACT-Filipe-BatistaPresentation.pptx) (Acceso: junio/2016).

Batista e Silva, F., Gallego, J., Lavalle, C. 2013a. A high-resolution population grid map for Europe. *Journal of Maps*, 9(1), 16-28. <http://dx.doi.org/10.1080/17445647.2013.764830>

Batista e Silva, F., Lavalle, C., Koomen, E. 2013b. A procedure to obtain a refined European land use/cover map. *Journal of Land Use Science*, 8(3), 255-283. <http://dx.doi.org/10.1080/1747423X.2012.667450>

Batista e Silva, F., Poelman, H., Martens, V., Lavalle, C. 2013c. Population Estimation for the Urban Atlas Polygons. *Joint Research Centre*. Technical Report 24437.

Dijkstra, L., Poelman, H. 2012. Cities in Europe. The new OECD-EC definition, *Regional Focus*, 1/2012, Bruselas, EC-DG REGIO.

Dobson, J. E., Bright, E. A., Coleman, P. R., Durfee, R. C., Worley, B. A. 2000. LandScan: A global population database for estimating population at risk. *Photogrammetric Engineering and Remote Sensing*, 66(7), 849-857.

Eicher, C., Brewer, C. 2001. Dasymetric mapping and areal interpolation: implementation and evaluation. *Cartography and Geographic Information Science*, 28(2), 125-138. <http://dx.doi.org/10.1559/152304001782173727>

ESRI. 2013. ArcGIS Desktop. Release Version 10.2.0. Redlands, CA: Environmental Systems Research Institute.

Ferri, S., Syrris, V., Florczyk, A., Scavazzon, M., Halkia, M., Pesaresi, M. 2014. A new map of the European settlements by automatic classification of 2.5m resolution SPOT data. In *IEEE International Geoscience and Remote Sensing Symposium* (pp. 1160-1163). IGARSS. <http://dx.doi.org/10.1109/IGARSS.2014.6946636>

<sup>8</sup> De donde también es posible obtener información sobre alturas de los edificios.

- Fisher, P. F., Langford, M. 1995. Modelling the errors in areal interpolation between zonal systems by Monte Carlo simulation. *Environment and Planning A*, 27, 211-224. <http://dx.doi.org/10.1068/a270211>
- Florczyk, A. J., Ferri, S., Syrris, V., Kemper, T., Halkia, M., Soille, P., Pesaresi, M. 2015. A New European Settlement Map From Optical Remotely Sensed Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(5), 1978-1992. <http://dx.doi.org/10.1109/JSTARS.2015.2485662>
- Freire, S., Halkia, M., Pesaresi, M., Ehrlich, D. 2015. Urban and Regional Built-up Analysis (URBA). Production of a population grid in Europe (1 km of resolution) *JRC Technical Reports*. Report EUR 27482. <http://dx.doi.org/10.2788/954596>
- Gallego, F. J. 2010. A population density grid of the European Union. *Population & Environment*, 31(6), 460-473. <http://dx.doi.org/10.1007/s11111-010-0108-y>
- Gallego, F. J., Batista, F., Rocha, C., Mubareka, S. 2011. Disaggregating population density of the European Union with CORINE land cover. *International Journal of Geographical Information Science*, 25(12), 2051-2069. <http://dx.doi.org/10.1080/13658816.2011.583653>
- Goerlich, F. J., Cantarino, I. 2012. A population density grid for Spain. *International Journal of Geographical Information Science*, 27(12), 2247-2263. <http://dx.doi.org/10.1080/13658816.2013.799283>
- Goerlich, F. J., Cantarino, I. 2016. Grid poblacional 2011 para España. Evaluación metodológica de diversas posibilidades de elaboración. *Estudios Geográficos*. Pendiente de publicación.
- Halkia, M. 2014. GHSL application in Europe: Towards new population grids. En: *European Forum for Geography and Statistics*. Krakov: 22-24 octubre. [https://geo.stat.gov.pl/documents/20182/25353/2\\_EFGS2014+Halkiav.pdf/a4fb7e4c-9013-4d0f-8f44-198ee77a56ec](https://geo.stat.gov.pl/documents/20182/25353/2_EFGS2014+Halkiav.pdf/a4fb7e4c-9013-4d0f-8f44-198ee77a56ec) (Acceso: junio/2016).
- Halkia, M., Freire, S. 2015. Towards a 100-m GHSL-based population grid in Europe. En: *European Forum for Geography and Statistics*. Vienna: 10-12 noviembre. Accesible en: [http://cpanel53.proisp.no/~efgsiowd/wp-content/uploads/conferences/efgs/2015/Conference\\_EFGS2015\\_1110\\_7\\_6\\_EC\\_JRCPresentation.pptx](http://cpanel53.proisp.no/~efgsiowd/wp-content/uploads/conferences/efgs/2015/Conference_EFGS2015_1110_7_6_EC_JRCPresentation.pptx) (Acceso: junio/2016).
- Harvey, J. T. 2002. Population Estimation Models Based on Individual TM Pixels. *Photogrammetric Engineering & Remote Sensing*, 68(11), 1181-1192.
- Hermosilla, T., Ruiz, L. A., Recio, J. A., Estornell, J. 2011. Evaluation of Automatic Building Detection Approaches Combining High Resolution Images and LiDAR Data. *Remote Sensing*, 3(6), 1188-1210. <http://dx.doi.org/10.3390/rs3061188>
- Hermosilla, T., Ruiz, L. A., Gil-Yepes, J. L., Recio, J. A., Pardo, J. E. 2013. Multi-Level Object-Based Urban Mapping: From Remote Sensing and GIS Data. En *GIS Ostrava - Geoinformatics for City Transformation*. Ostrava: 21-23 enero. pp. 1-10.
- INSPIRE. 2010. D2.8.I.2 INSPIRE specification on geographical grid systems - guidelines. Version 3.0.1. Brussels: INSPIRE Thematic Working Group Coordinate Reference Systems and Geographical Grid Systems (26 April 2010). Accesible en <http://inspire.jrc.ec.europa.eu/index.cfm/pageid/2>. (Acceso: junio/2016).
- Kaufholz, R. P. 2004. *Building Densities. A Physical Volume Approach*. MSc dissertation, MSc thesis Faculteit der Aarden levenswetenschappen Vrije Universiteit Amsterdam.
- Kim, T., Lee, T.-Y., Kim, K.-O. 2006. Semiautomatic Building Line Extraction from Ikonos Images Through Monoscopic Line Analysis. *Photogrammetric Engineering & Remote Sensing*, 72(5), 541-549. <http://dx.doi.org/10.14358/PERS.72.5.541>
- Kose, M. 2015. *Improving Population Estimation Models Using Remotely Sensed and Ordnance Survey Datasets*. Doctoral dissertation, Department of Geography. University of Leicester.
- Langford, M. 2006. Obtaining population estimates in non-census reporting zones: An evaluation of the 3-class dasymetric method. *Computers, Environment and Urban Systems*, 30(2), 161-180. <http://dx.doi.org/10.1016/j.compenvurbsys.2004.07.001>
- Langford, M. 2007. Rapid facilitation of dasymetric-based population interpolation by means of raster pixel maps. *Computers, Environment and Urban Systems*, 31(1), 19-32. <http://dx.doi.org/10.1016/j.compenvurbsys.2005.07.005>
- Langford, M., Unwin, D. J. 1994. Generating and mapping population density surfaces within a geographical information system. *Cartographic Journal*, 31(1), 21-26. <http://dx.doi.org/10.1179/caj.1994.31.1.21>
- Liu, X., Herold, M. 2006. Population Estimation and Interpolation Using Remote Sensing. En Q. Weng y D. A. Quattrochi (Eds.), *Urban Remote Sensing*, Taylor & Francis Group: pp. 269-290. <http://dx.doi.org/10.1201/b15917-18>
- Lu, Z., Im, J., Quackenbush, L. 2011. A volumetric approach to population estimation using LIDAR remote sensing. *Photogrammetric Engineering and Remote Sensing*, 77(11), 1145-1156. <http://dx.doi.org/10.14358/PERS.77.11.1145>
- Lu, Z., Im, J., Quackenbush, L., Halligan, K. 2010. Population estimation based on multi-sensor data fusion. *International Journal of Remote Sensing*, 31(21), 5587-5604. <http://dx.doi.org/10.1080/01431161.2010.496801>

- Lwin, K. K., Murayama, Y. 2009. A GIS approach to estimation of building population for micro-spatial analysis. *Transactions in GIS*, 13(4), 401-414. <http://dx.doi.org/10.1111/j.1467-9671.2009.01171.x>
- Lwin, K. K., Murayama, Y. 2011. Estimation of Building Population from LIDAR Derived Digital Volume Model. In Y. Murayama & R. B. Thapa (Eds.), *Spatial Analysis and Modeling in Geographical Transformation Process: GIS-based Applications* (First, pp. 87-98). New York, USA. [http://dx.doi.org/10.1007/978-94-007-0671-2\\_6](http://dx.doi.org/10.1007/978-94-007-0671-2_6)
- Martin, D., Cockings, S., Smith A. 2015. Exploiting census workplace data to build a daytime grid map of England and Wales. En: *European Forum for Geography and Statistics*. Vienna: 10-12 noviembre. Accesible en: [http://cpanel53.proisp.no/~efgsiowd/wp-content/uploads/conferences/efgs/2015/Conference\\_EFGS2015\\_1110\\_7\\_4\\_DaveJMartinPresentation.pptx](http://cpanel53.proisp.no/~efgsiowd/wp-content/uploads/conferences/efgs/2015/Conference_EFGS2015_1110_7_4_DaveJMartinPresentation.pptx) (Acceso: junio/2016).
- Martin, D., Tate, N. J., Langford, M. 2000. Refining population surface models: experiments with Northern Ireland census data. *Transactions in GIS*, 4(4), 343-360. <http://dx.doi.org/10.1111/1467-9671.00060>
- Mennis, J. 2003. Generating surface models of population using dasymetric mapping. *Professional Geographer*, 55(1), 31-42. <http://onlinelibrary.wiley.com/doi/10.1111/0033-0124.10042/abstract> (Acceso: junio/2016).
- Mennis, J. 2009. Dasymetric mapping for estimating population in small areas. *Geography Compass*, 3(2), 727-745. <http://dx.doi.org/10.1111/j.1749-8198.2009.00220.x>
- Mennis, J., Hultgren, T. 2006. Intelligent Dasymetric Mapping and Its Application to Areal Interpolation. *Cartography and Geographic Information Science*, 33(3), 179-194. <http://dx.doi.org/10.1559/152304006779077309>
- Qiu, F., Sridharan, H., Chun, Y. 2010. Spatial Autoregressive Model for Population Estimation at the Census Block Level Using LIDAR-derived Building Volume Information. *Cartography and Geographic Information Science*, 37(3), 239-257. <http://dx.doi.org/10.1559/152304010792194949>
- Pesaresi, M., Huadong, G., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., ... & Marin-Herrera, M. A., Ouzounis, G.K., Scavazon, M., Soille, P., Syrris, V., Zanchetta, L. 2013. A Global Human Settlement Layer from optical HR/VHR RS data: Concept and first results. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(5), 2102-2131. <http://dx.doi.org/10.1109/JSTARS.2013.2271445>
- Ramesh, S. 2009. *High resolution satellite images and LiDAR data for small-area building extraction and population estimation*. MSc dissertation, University of North Texas. Accesible en: [http://digital.library.unt.edu/ark:/67531/metadc12188/m2/1/high\\_res\\_d/thesis.pdf](http://digital.library.unt.edu/ark:/67531/metadc12188/m2/1/high_res_d/thesis.pdf) (Acceso: junio/2016).
- Rottensteiner, F. 2003. Automatic Generation of High-Quality Building Models from LIDAR Data. *IEEE Computer Graphics and Applications*, 23(6), 42-50. <http://dx.doi.org/10.1109/MCG.2003.1242381>
- Rottensteiner, F., Jansa, J. 2002. Automatic extraction of buildings from LIDAR data and aerial images. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(4), 569-574. Accesible en: [http://testcis.cis.rit.edu/~cnspci/references/dip/urban\\_extraction/rottensteiner2002.pdf](http://testcis.cis.rit.edu/~cnspci/references/dip/urban_extraction/rottensteiner2002.pdf) (Acceso: junio/2016).
- Santos Preciado J. M. 2015. La cartografía catastral y su utilización en la desagregación de la población. Aplicación al análisis de la distribución espacial de la población en el municipio de Leganés (Madrid). *Estudios Geográficos*, 76(278), 309-333. <http://dx.doi.org/10.3989/estgeogr.201511>
- SIOSE. 2011. Descripción del Modelo de Datos y Rótulo SIOSE2005. Versión 2.2. Sistema de Información de Ocupación del Suelo de España. *Equipo Técnico Nacional SIOSE*. Instituto Geográfico Nacional. Accesible en: [http://www.siose.es/SIOSEtheme-theme/documentos/pdf/ModeloDatos\\_Rotulo\\_SIOSE\\_v2.1.pdf](http://www.siose.es/SIOSEtheme-theme/documentos/pdf/ModeloDatos_Rotulo_SIOSE_v2.1.pdf) (Acceso: junio/2016).
- Tobler, W. R. 1979. Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association*, 74(367), 519-530. <http://dx.doi.org/10.1080/01621459.1979.10481647>
- Tobler, W. R., Deichmann, U., Gottsegen, J., Maloy, K. 1997. World Population in a Grid of Spherical Quadrilaterals. *International Journal of Population Geography*, 3(3), 203-225. [http://dx.doi.org/10.1002/\(SICI\)1099-1220\(199709\)3:3%3C203::AID-IJPG68%3E3.0.CO;2-C](http://dx.doi.org/10.1002/(SICI)1099-1220(199709)3:3%3C203::AID-IJPG68%3E3.0.CO;2-C)
- Wright, J. K. 1936. A method of mapping densities of population with Cape Cod as an example. *The Geographical Review*, 26, 103-110. <http://dx.doi.org/10.2307/209467>
- Wu, C., Murray, A. T. 2005. A cokriging method for estimating population density in urban areas. *Computers, Environment and Urban Systems*, 29(5), 558-579. <http://dx.doi.org/10.1016/j.compenvurbysys.2005.01.006>
- Wu, S., Qiu, X., Wang, L. 2005. Population Estimation Methods in GIS and Remote Sensing: A Review. *GIScience Remote Sensing*, 42(1), 80-96. <http://www.tandfonline.com/doi/abs/10.2747/1548-1603.42.1.80> (Acceso: junio/2016).

- Yu, B., Liu, H., Wu, J., Hu, Y., Zhang, L. 2010. Automated derivation of urban building density information using airborne LIDAR data and object-based method. *Landscape and Urban Planning*, 98(3-4), 210-219. <http://dx.doi.org/10.1016/j.landurbplan.2010.08.004>
- Yuan, Y., Smith, R. M., Limp, F. W. 1997. Remodeling Census Population with Spatial Information from Landsat TM Imagery. *Computers, Environment and Urban Systems*, 21(3-4), 245-258. [http://dx.doi.org/10.1016/S0198-9715\(97\)01003-X](http://dx.doi.org/10.1016/S0198-9715(97)01003-X)