

Fundación **BBVA**

Una *grid* de densidad de población para España

Francisco J. Goerlich Gisbert
Isidro Cantarino Martí

Informes 2012

Economía y Sociedad

**Una *grid* de densidad
de población para España**

Una *grid* de densidad de población para España

Francisco J. Goerlich Gisbert
Isidro Cantarino Martí

Primera edición, noviembre 2012

© Los autores, 2012

© Fundación BBVA, 2012
Plaza de San Nicolás, 4. 48005 Bilbao
www.fbbva.es
publicaciones@fbbva.es

Copia digital de acceso público en www.fbbva.es

Al publicar el presente informe,
la Fundación BBVA no asume responsabilidad alguna
sobre su contenido ni sobre la inclusión en el mismo
de documentos o información complementaria
facilitada por los autores.

Edición y producción: Editorial Biblioteca Nueva, S. L.

ISBN: 978-84-92937-39-4
Depósito legal: BI-1.718-2012

Impreso en España – *Printed in Spain*

Impreso por Lável Industria Gráfica, S. A.
sobre papel con un 100% de fibras recicladas
y elaborado según las más exigentes normas ambientales europeas.

Índice

AUTORES	9
AGRADECIMIENTOS	11
RESUMEN / SUMMARY	13
INTRODUCCIÓN	15
1. ESTADÍSTICAS SOCIOECONÓMICAS Y SISTEMAS ZONALES	25
1.1. Métodos	28
1.2. Transferencia de datos entre diferentes sistemas zonales	30
2. LA INFORMACIÓN DE BASE	35
2.1. Población residente: secciones censales	35
2.2. Coberturas del suelo: SIOSE	36
2.3. Otra información utilizada en el estudio	36
2.4. Un comentario sobre las fechas de referencia	37
3. DISTRIBUCIÓN ESPACIAL DE LA POBLACIÓN	41
3.1. Municipios	41
3.2. Estadísticas inframunicipales: el nomenclátor de unidades poblacionales ...	52
3.3. Estadísticas inframunicipales: secciones censales	63
4. CORINE LAND COVER Y LA DISTRIBUCIÓN DE LA POBLACIÓN	71
4.1. El proyecto <i>CORINE Land Cover</i>	72
4.2. Desagregación espacial (<i>downscaling</i>) con CLC	73
5. EL SISTEMA DE INFORMACIÓN DE OCUPACIÓN DEL SUELO EN ESPAÑA: SIOSE ..	79
6. LA DISTRIBUCIÓN DE LA POBLACIÓN A PARTIR DEL SIOSE: UNA GRID DE DENSIDAD DE POBLACIÓN PARA ESPAÑA	83
6.1. Determinación de las coberturas susceptibles de soportar población residente	84
6.2. Intersección con el fichero vectorial de secciones censales	86
6.3. Análisis de secciones censales puras	89
6.4. Modelización estadística y generalización al resto de secciones censales ..	95

6.5. Generación de la <i>grid</i>	97
6.6. Validación	100
RESUMEN Y CONCLUSIONES: USOS POTENCIALES	105
APÉNDICES	109
A.1. Coberturas de <i>CORINE Land Cover</i>	109
A.2. El modelo de datos del SIOSE: coberturas simples, compuestas y atributos ..	114
A.3. Sistemas de referencia geográfica en la UE. INSPIRE	122
BIBLIOGRAFÍA	127
ÍNDICE DE CUADROS, ESQUEMAS Y GRÁFICOS	135
ÍNDICE DE MAPAS	137



Autores

Francisco J. Goerlich Gisbert es licenciado en Ciencias Económicas por la Universidad de Valencia, máster en Económicas por la London School of Economics, doctor por la Universidad de Valencia, catedrático del Departamento de Análisis Económico en la misma universidad y profesor investigador del Ivie. Coautor de más de diez libros, ha publicado medio centenar de artículos sobre temas de macroeconomía,

econometría y economía regional en diversas revistas nacionales e internacionales, tales como *Investigaciones Económicas*, *Revista Española de Economía*, *Revista de Economía Aplicada*, *Investigaciones Regionales*, *Estadística Española*, *Review of Income and Wealth*, *Regional Studies*, *Journal of Regional Science*, *Applied Economics*, *Population*, *Economics Letters* o *Econometric Theory*.

Isidro Cantarino Martí es doctor ingeniero agrónomo por la Universidad Politécnica de Valencia y profesor titular del Departamento de Ingeniería del Terreno de la citada universidad. A lo largo de su carrera profesional, ha realizado diversos trabajos técnicos y de investigación en la determinación y gestión de recursos hidráulicos con aplicación en ámbitos

montañosos, promovidos por la Dirección General de Obras Hidráulicas. En la actualidad se encuentra colaborando con la Consellería de Cultura, Educación y Deportes para implementar un Sistema de Información Geográfica que gestione la localización y clasificación patrimonial de yacimientos paleontológicos, así como cartografías temáticas asociadas.



Agradecimientos

Este informe se ha beneficiado de la ayuda y comentarios de muchas personas. Aun corriendo el riesgo de omitir a alguien, queremos agradecer la colaboración prestada por Pilar Chorén y Héctor García, del Instituto Valenciano de Investigaciones Económicas (Ivie); Ignacio Duque, José Luis Maldonado y Josefa Barragán, del Instituto Nacional de Estadística (INE); Ángel Sánchez y Dolores Núñez, del Instituto de Estadística de la Comunidad de Madrid; M.^a Elena Caballero y Julián Delgado, del Equipo Nacional del Sistema de Información sobre Ocupación del Suelo de España (SIOSE); Juan Carlos Losana García y Angélica Castaño, del Registro Central de Cartografía del Instituto Geográfico

Nacional; Roger Milego, del European Topic Centre for Spatial Information and Analysis (ETC-SIA) y los asistentes a las Conferencias del European Forum for Geostatistics de Tallin (Estonia), 2010 y Lisboa (Portugal), 2011. Los autores agradecen el apoyo del programa de investigación de la Fundación BBVA-Ivie para la realización de este documento. Adicionalmente, Francisco J. Goerlich reconoce la ayuda del proyecto del Ministerio de Ciencia y Tecnología ECO2011-23248. Los resultados mencionados en el texto pero no ofrecidos en el mismo están disponibles si se solicitan a los autores, así como un extenso apéndice estadístico que sustenta las afirmaciones del informe.

Resumen *Summary*

Este informe elabora un mapa de la distribución de la población en España. Tradicionalmente, la distribución de la población se representa mediante mapas de coropletas en los que la población de una determinada área administrativa, ya sea una región o un municipio, se distribuye de forma uniforme sobre todo el territorio de dicha área. Sin embargo, esta forma de distribuir la población no es satisfactoria, y produce resultados engañosos que no se ajustan a la realidad.

Una forma natural de escapar de las restricciones impuestas por los límites administrativos es representar a la población en formato de rejilla o *grid*. Este sistema zonal consiste en celdas cuadradas de idéntico tamaño, en las que representamos la población residente convenientemente distribuida. Para realizar este reparto, se ha utilizado una aproximación *top-down*, en la cual los datos disponibles de población se desagregan mediante técnicas estadísticas utilizando como material auxiliar información sobre coberturas del suelo.

In this study we elaborate a dasymetric map of population distribution for Spain. Population distribution is traditionally represented by means of choropleth maps in which the population of a given administrative unit, whether a region or municipality, is distributed evenly across its territory. But this is unsatisfactory, and produces misleading results that match poorly with reality.

A natural way to elude the restrictions imposed by administrative boundaries is to represent the population in grid form. This zonal system consists of square cells of equal size in which we represent the resident population, distributed as appropriate. To effect this distribution, we use a top-down

En nuestro caso, se ha utilizado la población residente a nivel de sección censal para 2010 y la información referente a las coberturas del suelo del Sistema de Información de Ocupación del Suelo en España (SIOSE), de reciente aparición.

A partir del SIOSE seleccionamos, en primer lugar, las coberturas susceptibles de albergar población residente. El reparto de población en estas superficies se efectúa mediante un proceso iterativo en función del tipo de edificación y manteniendo la restricción de volumen, es decir, si recuperamos las secciones censales originales se recobra su población.

Los resultados obtenidos se han validado frente a un fichero de población georreferenciada para la Comunidad de Madrid. Se ha calculado un índice de discrepancia que se ha comparado con los resultados obtenidos para otros países europeos que han seguido metodologías similares. De acuerdo con esta comparativa, la *grid* obtenida en este informe ofrece unos parámetros de calidad muy superiores a la de otros países europeos.

approach in which available population data are disaggregated using statistical techniques, with secondary reference to land cover information.

In our case, we use the resident population at census tract level for 2010 and land cover data from the recently released Land Cover and Use Information System of Spain (SIOSE).

From SIOSE we select, in the first instance, the polygons susceptible to hold residential population. To establish the share-out of population across these areas, we apply an iterative process based on type of construction, maintaining the volume constraint, such that if we recover the original census tracts their population is preserved.

The results obtained were validated against a geo-referenced population file for the Madrid Region. We also calculated a discrepancy index which was then compared with the results obtained for other

European countries using similar methodologies. This comparison shows that the grid featuring in this report offers quality parameters far superior to those of other European countries.

Introducción

Para conocer un territorio, es necesario contar con instrumentos que nos permitan un acercamiento a la realidad, y, en este sentido, el primero de los instrumentos necesarios para el análisis y comprensión de una realidad cambiante y compleja es una base cartográfica.

Densidad de población de los municipios españoles
Mapas provinciales
Censos de población y viviendas 1991 (INE 1994)

El interés por representar sobre un mapa la distribución de la población se remonta al propio desarrollo de la cartografía (Reher 2006). El mapa 1 muestra el mapa de densidades más antiguo disponible en los fondos cartográficos del Instituto Geográfico Nacional (IGN). En él se representa un mapa de *coropletas* de la densidad de población de España por partidos judiciales, según los datos demográficos del censo de 1860, y se observan ya los rasgos fundamentales de la actual distribución de la población, con su elevada concentración en el litoral y la *isla* de Madrid.

Esta distribución comenzó a fraguarse en el último tercio del segundo milenio (Pérez Moreda 2004), tras un corto período de hegemonía castellana a mediados del siglo xv (Yun 2002), y que a su vez supuso una ruptura drástica respecto al comienzo del segundo milenio, cuando las ciudades de al-Andalus, con Córdoba a la cabeza, eran las más pobladas de Europa (Casado 2002).

En la actual organización administrativa del Estado español, los municipios constituyen las unidades administrativas menores en las que se divide el territorio nacional y que tienen asignados lindes precisos sobre los que se extienden sus competencias¹. Por esta razón, y también por la disponibilidad estadística, los trabajos que estudian la localización de la población para áreas geográficas amplias suelen hacerlo, en el mejor de los casos, descendiendo al nivel municipal (Reques y Rodríguez 1998; Zoido y Arroyo 2004; De Cos y Reques 2005; Goerlich *et al.* 2006; Goerlich y Mas 2008a, 2008b, 2009). Sin embargo, como han reconocido numerosos autores (Reher 1994), desde el punto de vista del estudio del asentamiento de la población sobre el territorio esta división es claramente insuficiente y debemos aumentar la escala geográfica del análisis.

Por otro lado, los municipios son la unidad básica de recogida y clasificación de la información estadística. Una gran parte de los datos demográficos está disponible a nivel municipal, si bien la mayor parte de información estadística de carácter socioeconómico solo existe a niveles superiores de agregación, ya sean provincias o comunidades autónomas. Desde el punto de vista geográfico-estadístico, es bien conocido que la división del territorio es relevante para un tema particular si las áreas que constituyen la partición son homogéneas, o cumplen una determinada función, para el problema que deseamos

¹ Es cierto, sin embargo, que por debajo del municipio existe el concepto de *Entidad Local de Ámbito Territorial Inferior al Municipio* (entidades locales menores, ELM), definido por la Ley 7/1985, de 2 de

abril, Reguladora de las Bases del Régimen Local como una unidad de gestión, administración descentralizada y representación política dentro del municipio.

investigar (Vidal, Gallego y Kayadjanian 2001). No hay una partición geográfica uniformemente mejor para todas las situaciones, sin embargo la captación de estadísticas demográficas y socioeconómicas parte de áreas administrativas predefinidas, organizadas según una estructura jerárquica, y no relacionadas directamente con los problemas que se pretenden analizar, sino más bien con la estructura de toma de decisiones política.

Este problema es más evidente cuando analizamos variables climáticas o medioambientales, en las que las zonas homogéneas o funcionales de análisis tienen claramente que ver con características geográficas o relacionadas con el paisaje: cortes altimétricos, coberturas del suelo, zonas climáticas homogéneas o cuen-

cas hidrográficas, por ejemplo; pero subsiste igualmente en el caso de variables demográficas o socioeconómicas. Así pues, si deseamos analizar la distribución de la población a partir de datos municipales deberíamos preguntarnos sobre la homogeneidad de los municipios para este fin. Los problemas en este contexto son básicamente tres:

1. Por una parte, los estudios sobre localización de la población en el territorio están excesivamente condicionados por los lindes administrativos que es necesario superar; dichos lindes no guardan relación muchas veces con la estructura real de los asentamientos; la población se recoge a nivel municipal, pero hay núcleos de población separados por una calle que pertenecen a



Mapa 1 Densidad de la población de España por partidos judiciales. Censo de 1860

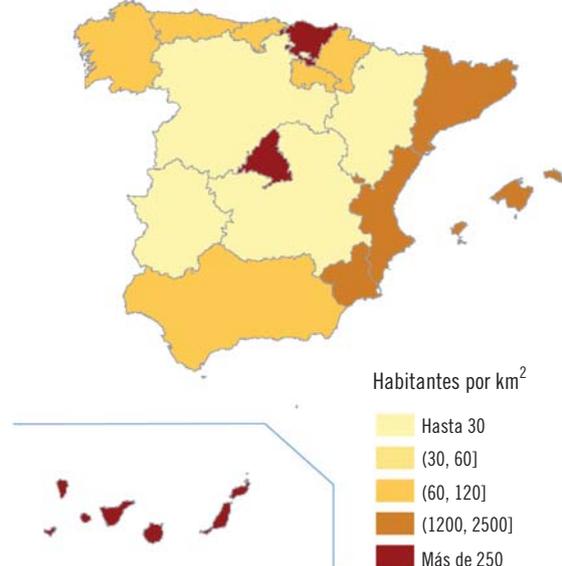
Escala 1:2.500.000. 250 kilómetros [= 9,8 cm].

Fuente: Instituto Geográfico Nacional (IGN 2012).

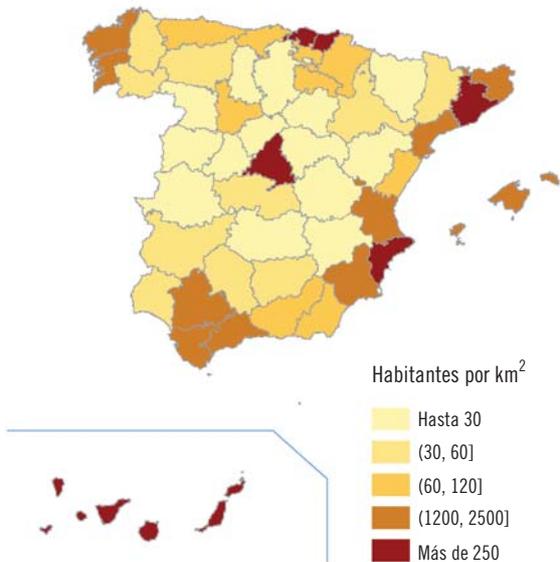
a) España



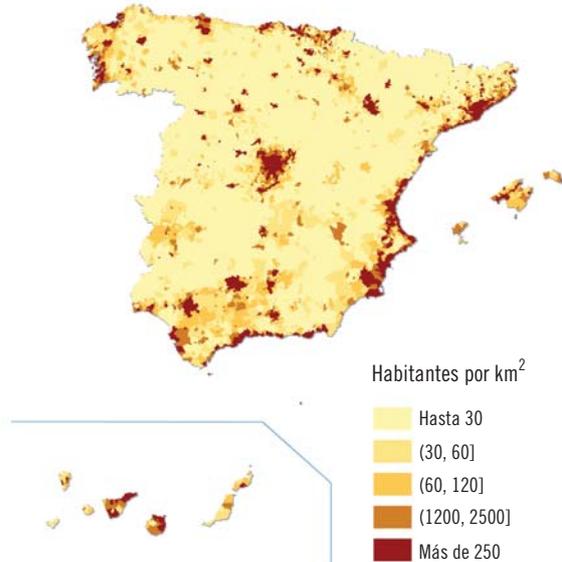
b) Comunidades autónomas



c) Provincias



d) Municipios



Mapa 3 Densidad de población. Padrón del 2010. Efecto de la escala de representación

Fuente: INE (2010) y elaboración propia.

totalmente a dicho problema, y solo podemos aspirar a minimizar sus efectos mediante procedimientos de agregación o desagregación de las unidades territoriales que tengan en cuenta el nivel de autocorrelación espacial local (Openshaw 1984; Rase 2001; Wong 2004; *European Observation Network*

for Territorial Development and Cohesion [ESPON] 2006). En definitiva, el MAUP indica que la correlación entre dos fenómenos distribuidos en el espacio puede cambiar de forma dramática de acuerdo con el tamaño y la forma de las unidades espaciales sobre las que el fenómeno es medido.

3. Finalmente, los municipios españoles tienen un grado considerable de heterogeneidad, su estructura ha sido determinada por razones históricas y sus delimitaciones no responden necesariamente a criterios relacionados con la distribución de la población sobre el territorio, sino más bien a criterios de conveniencia administrativa y voluntad política². La complejidad en lo referente a la definición y alteración de los lindes municipales es un tema recurrente en la generación y mantenimiento de una cartografía digital en España por parte del IGN (Rodríguez, Martín-Asín y Astudillo 1997; Capdevila i Subirana 2009)³. Uno de los problemas más graves en este sentido es la presencia de algunos términos municipales muy grandes, que conviven con algunos de tamaño ridículo, como tendremos ocasión de comprobar más adelante.

Así pues, un examen detallado de la distribución de la población sobre el territorio exige, por una parte, descender más allá del nivel de agregación municipal, y por otra, superar de alguna forma el rígido marco de los lindes administrativos (Muguruza y Santos 1989). La superación de los lindes administrativos se hace todavía más necesaria cuando se desea integrar

datos demográficos y socioeconómicos con datos geográficos o medioambientales, recogidos en muchas ocasiones en formato de rejilla (*grid*), o *raster* utilizando la terminología de los Sistemas de Información Geográfica (GIS)⁴.

Conscientes de esta problemática, hace unos años, algunos institutos nacionales de estadística de diversos países de la Unión Europea (UE), con el apoyo del Statistical Office of the European Communities (Eurostat), lanzaron la *Grid Club Initiative*, actualmente conocido como el *European Forum for Geostatistics* (v. sitio web del EFGS [2012]), con el objetivo de armonizar estadísticas europeas sobre la base de una *grid* de 1 × 1 kilómetros y en un sistema geodésico de referencia común, en lugar de sobre la base territorial regional actual basada en las regiones europeas (*Nomenclature of Territorial Units for Statistics*, NUTS, Eurostat 2007) o en las unidades administrativas locales (LAU, *local administrative units*). La primera variable que se identificó para tal armonización fue la población, a la que en el medio plazo deberían añadirse otros indicadores.

En esta dirección, la mejor forma de generar mapas sobre la distribución de la población sería georreferenciar todos los edificios resi-

² El primer censo que presenta el conjunto completo de los municipios que cubren el territorio español es el denominado *Censo de la matrícula catastral*, fechado en 1842. *Alteraciones de los municipios en los censos de población desde 1842*, INE (sitio web).

³ Baste señalar que hasta mayo del 2011 el Centro Nacional de Información Geográfica (CNIG) del IGN no ha puesto a disposición de los usuarios en su Centro de Descargas un fichero vectorial de las líneas de límite municipales, en formato *shape*, que estuviera libre de errores.

Aun en este fichero los municipios de Ansó y Fago, en la provincia de Huesca, aparecen englobados en un único recinto, dado que en la actualidad no existe, formalmente, el deslinde jurisdiccional entre ambos municipios. La línea que separa los polígonos asignados a Fago y Ansó que aparece en los mapas municipales (INE 1994), fue incorporada a la base de datos de límites jurisdiccionales del IGN en 1995, a partir de un documento gráfico procedente del INE por una necesidad topológica, ante la imposibilidad de atribución de datos estadísticos a nivel municipal. Por esta razón, «se procedió a señalar un polígono artificial para Fago que corresponde al 20% de su participación en la Comunidad de Villa y Tierra con Ansó. En ningún caso

esta línea supone el reconocimiento de un término municipal para el municipio de Fago».

Ante la inexistencia de documentación oficial (acta de deslinde) respecto a la delimitación de ambos términos municipales, el IGN eliminó la línea de separación entre dichos municipios de su base de datos de líneas de límite, en tanto en cuanto no se produzca el acto de deslinde entre ambos municipios. Por esta razón, a nivel cartográfico tenemos la peculiaridad de que ambos municipios están englobados en un único polígono. (Correspondencia electrónica (28/12/2010) con Juan Carlos Losana García, del Área de Registro Central de Cartografía del IGN.)

En nuestra cartografía de términos municipales, y dado que tenemos la misma necesidad topológica que el INE manifestó en su momento, recuperamos el término municipal de Fago de una base de datos más antigua del propio IGN, descargada de la Infraestructura de Datos Espaciales de España (IDEE), y la sustraemos del conjunto Ansó-Fago. Este polígono coincide con el de la sección censal del municipio de Fago en el fichero vectorial del INE.

⁴ GIS: *Geographical Information System*. Los acrónimos utilizados en el texto se corresponden a su original en inglés, por la internacionalización de su uso y para evitar equívocos.

denciales de un país, determinar la población que reside habitualmente en ellos, contarla y asignarla al píxel correspondiente. Esta aproximación, que construye el mapa desde su base, se conoce como *bottom-up approach*, y su realización es posible a partir de: 1) enormes bases de datos geográficas sobre las construcciones residenciales del país; y 2) de un censo o un registro exhaustivo de población (como el padrón), convenientemente enlazado con la base de datos de edificios. Este tipo de procedimientos solo pueden ser implementados por institutos oficiales de estadística u organismos gubernamentales y su difusión presenta ciertos inconvenientes relacionados con la confidencialidad, especialmente si la *grid* es fina o se utilizan características peculiares de la población⁵.

A pesar de ello, la primera *grid* de población que se conoce para una región de España, Madrid, se remonta a 1976, siguió esta filosofía y fue realizada por un investigador individual, Julio Vinuesa (1976). Un trabajo ingente, si pensamos que fue realizado antes del desarrollo de las técnicas de los GIS y con cartografía en papel, y que lamentablemente no tuvo continuidad.

Frente a esta aproximación de construcción de una *grid* de población mediante agregación podemos pensar en métodos de desagregación espacial. Es la aproximación conocida como *downscaling* o *top-down approach*. En este caso se parte de datos municipales o incluso de un nivel inferior (secciones censales, entidades o núcleos de población) y se debe desagregar de alguna forma hasta el nivel de resolución de la *grid*. La literatura sobre métodos de desagregación espacial en este contexto es enorme

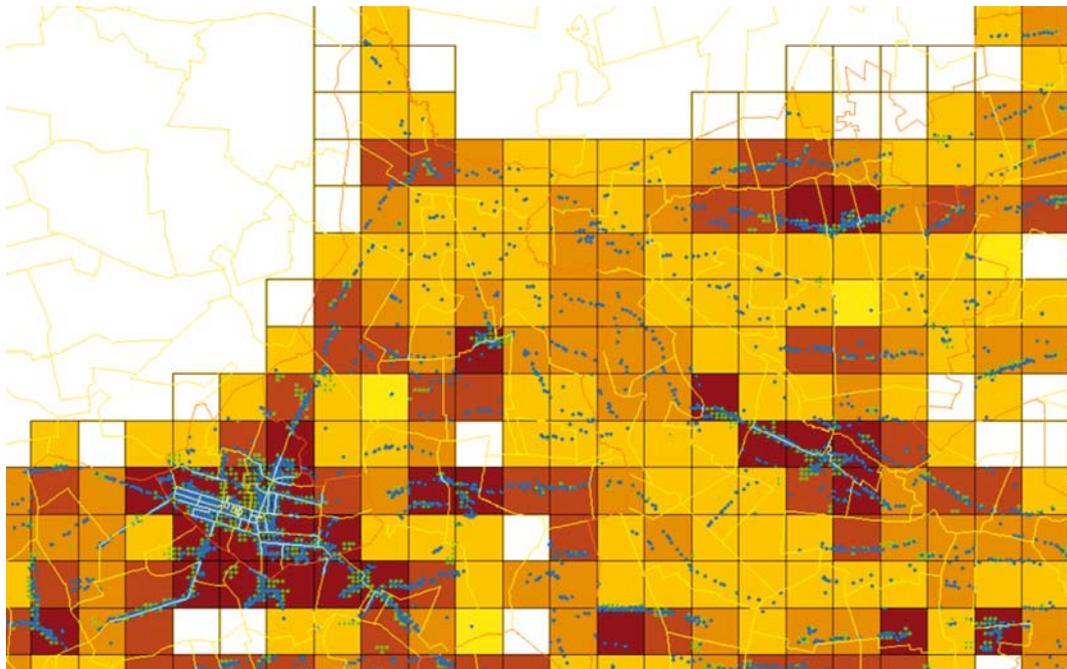
(Grasland, Mathian y Vincent 2000; Rase 2001) y forma parte de un problema más general en el análisis de datos espaciales referente a la transferencia de datos geográficos entre diferentes sistemas zonales (Flowerdew, Green y Kehris 1991).

Los mapas 4 y 5 contraponen ambos métodos. Los puntos azules representan la población georreferenciada, para cada habitante disponemos de sus coordenadas de acuerdo a un determinado criterio, residencia o lugar de trabajo por ejemplo. La superposición sobre estos puntos de una rejilla cartográfica de la resolución deseada nos genera de forma automática la *grid* que estamos buscando. Esta forma de proceder genera resultados exactos, en el sentido de que si la base de datos es completa, la localización de la población es perfecta. Obsérvese que en el mapa 4 a todas las celdas que contienen puntos azules se les asigna densidad.

Cuando la información de base no es suficiente para lograr este resultado y partimos de datos demográficos por unidades administrativas, debemos utilizar métodos de desagregación con información auxiliar. Un mapa de este tipo es conocido como *dasimétrico*. La calidad del resultado final dependerá en gran medida de la resolución y calidad de la información de partida, y en menor medida del algoritmo concreto de desagregación espacial utilizado (Martin, Tate y Langford 2000), pero al tratarse de un método estadístico es necesariamente menos perfecto que el anterior, y es inevitable etiquetar celdas como vacías, cuando en realidad contienen población. (Las celdas en blanco con puntos azules en el mapa 5.) Si mantenemos la restricción de volumen sobre la población representada, ello implica que en otras celdas tendremos más población de la que deberíamos.

⁵ El proyecto del INE sobre el censo del 2011, dirige sus esfuerzos hacia la georreferenciación de los edificios con la

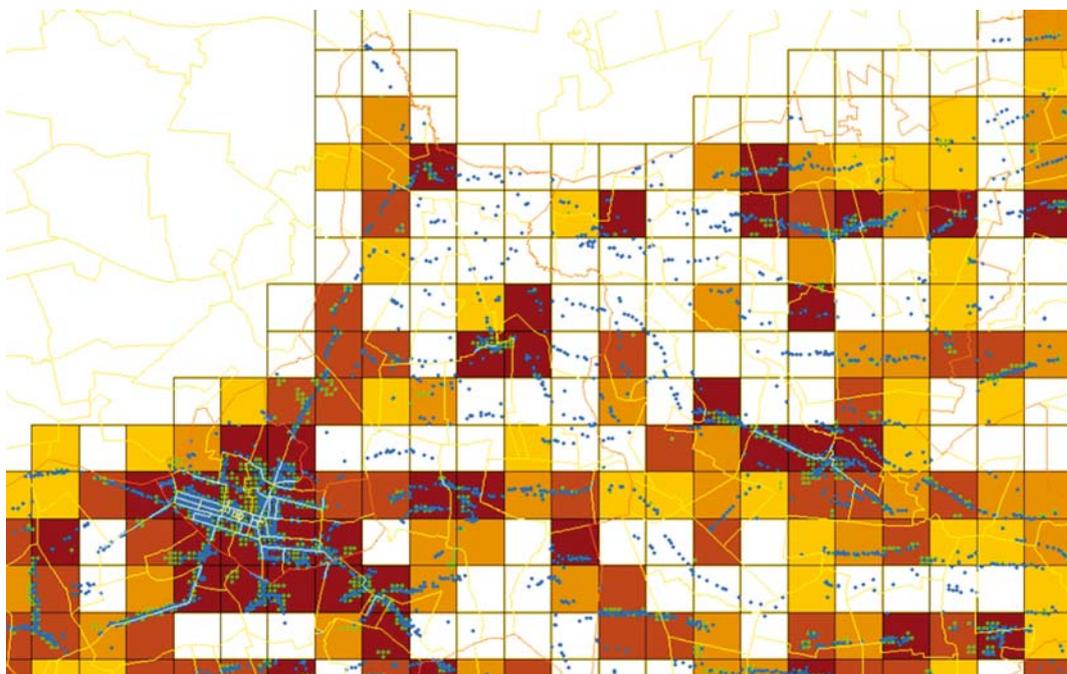
intención de localizar a la población en ellos y, menciona la posibilidad de difusión de resultados en formato de *grid*.



Mapa 4 Población georreferenciada: aproximación *bottom-up*

Nota: Para cada habitante disponemos sus coordenadas de acuerdo con el criterio de residencia. La superposición de la rejilla cartográfica deriva de forma inmediata la *grid* buscada.

Fuente: Jablonski (2011).



Mapa 5 Mapa dasimétrico: aproximación *top-down* o *downscaling*

Nota: Información auxiliar sobre usos del suelo, contornos urbanos o de edificios residenciales, es utilizada como información auxiliar en el algoritmo de desagregación espacial.

Fuente: Jablonski (2011).

1. ANTECEDENTES

Al margen de la investigación pionera de Vinuesa (1976), los intentos de elaboración de rejillas de población son escasos en España. Desde el punto de vista de las representaciones cartográficas los estudios de De Cos (2004) y García González y Cebrián (2006) han tratado esta cuestión, mediante la utilización de *kernels* (densidades), en el primer caso, o de métodos de interpolación, en el segundo. Las objeciones fundamentales a este tipo de aproximación son dos. Por una parte, el énfasis fundamental de este tipo de ejercicios descansa básicamente sobre un problema visual de representación cartográfica. Por otra parte, este tipo de desagregación no hace uso de ningún tipo de información auxiliar que determine realmente la localización de la población (o de la variable de interés) a partir del punto o del área de concentración de la misma. Sin embargo, la distribución de la población está altamente relacionada con las coberturas y usos del suelo, y esta información puede y debe ser introducida como información auxiliar para la desagregación. Adicionalmente, Martín (1996) ha puesto de manifiesto la importancia de satisfacer la restricción de volumen o picnofiláctica (Tobler 1979) para la población de partida, y que no es normalmente satisfecha por estos métodos de suavizado inicialmente utilizados en la elaboración de *grids* de población (Bracken y Martín 1989, 1995).

En el caso español, las publicaciones de Santos y García Lázaro (2003) y Suárez, Santos y Dorta (2008) han tratado de incorporar información sobre usos del suelo urbano en el proceso de distribución de la población para representarla en formato de *grid*. Se trata de estudios de carácter muy local (Leganés, en el primer caso, y la isla de Gran Canaria, en el segundo), si bien la resolución es extremadamente elevada, un área (10 × 10 metros de tamaño de celda).

Como antecedente a este informe Goerlich y Cantarino (2011) elaboraron una *grid* de pobla-

ción para la Comunitat Valenciana a partir de la información sobre el nomenclátor de unidades poblacionales del INE, junto con información puramente cartográfica, y un método dasimétrico binario (Langford y Unwin 1994; Langford 2007), que solo asigna población a polígonos urbanos convenientemente delimitados. Sin embargo, la escasa calidad de la información geográfica y su caótica organización⁶, así como la dificultad y falta de criterio en la delimitación de contornos urbanos residenciales a partir de la densidad de edificios, dificulta la extensión de sus métodos para el conjunto nacional.

En el contexto europeo, la aparición de los datos de coberturas del suelo del *Coordination of Information on the Environment (CORINE Land Cover [CLC])* ha planteado la cuestión de la integración de los mismos con los de estadísticas demográficas y socioeconómicas, fundamentalmente con el objetivo de avanzar en el conocimiento de la distribución de la población y de las relaciones entre la actividad humana y el medio ambiente. Esta integración de datos demográficos y socioeconómicos con estadísticas medioambientales solo es posible si utilizamos un sistema de unidades geográficas de referencia común, y en este sentido las *grids* parecen ser el marco de referencia más adecuado. Este sistema zonal, no solo se ajusta fácilmente a los GIS, sino que está especialmente adaptado al tratamiento de imágenes, de donde procede la mayor parte de la información primaria de carácter climático o de usos del suelo (teledetección), y además incrementa la comparabilidad entre unidades; todas tienen el mismo tamaño, independientemente de cuál sea la cuestión bajo investigación.

Por esta razón, Eurostat, a través del Centro Común de Investigación (*Joint Research Centre [JRC]*) de la UE, impulsó la elaboración de una *grid* de población a nivel europeo que desagregara la población municipal mediante la utilización de CLC (Gallego y Peedell 2001), y mostrara una distribución más realista de lo que se observa con un mapa de coropletas. Sin embargo, no se trata solamente de una cuestión visual de represen-

⁶ Piénsese que la unidad de distribución de la cartografía a escala 1:25.000 es la hoja del Mapa Topo-

gráfico Nacional (MTN25), lo que para el conjunto nacional representa manejar más de 4.000 hojas.

tación cartográfica, sino de integrar estadísticas socioeconómicas y medioambientales. Gracias a este esfuerzo existe una *grid* de población para España con resolución de 1 ha, fecha de referencia el censo del 2001, y que se puede obtener en formato *raster* del sitio web de la Agencia Europea del Medio Ambiente (AEMA 2012a; European Environment Agency [EEA]) (v. Gallego 2009). Este estudio está actualmente en revisión (Batista e Silva 2011), con fecha de referencia del 2006, ya que la escasa resolución de CLC para este fin ha hecho que la *grid* no se ajustara a la realidad como se esperaba.

Además de esta iniciativa existen datos sobre una *grid* de población para España en dos proyectos de cobertura mundial: *LandScan* (Dobson *et al.* 2000; Bhaduri *et al.* 2002) y la *Gridded Population of the World* (GPW) del Center for International Earth Science Information Network (CIESIN 2005). Sin embargo, estas investigaciones hacen referencia a un concepto de población referido como *población ambiente*, un promedio ponderado temporal de la población que existe en una determinada área, más que al concepto de población residente en dicha área, más habitual en las estadísticas demográficas, y que será el que adoptemos en este informe.

2. PLAN DEL ESTUDIO

Este trabajo realiza un ejercicio de desagregación espacial (*downscaling*) de la población residente española a una *grid* de 1 × 1 kilómetros de resolución, que mejore y actualice la disponible en el sitio web de la AEMA. Para ello, partimos de la población según la sección censal procedente de la explotación estadística del padrón municipal del 2010 del INE, y de la información del Sistema de Información de Ocupación del Suelo en España (SIOSE), como información auxiliar sobre coberturas del suelo.

El objetivo es doble. Por una parte, se trata de un trabajo metodológico, ensayamos métodos de desagregación que sean aplicables en contextos más generales, lo que nos permitirá en un futuro aplicarlos a otras variables, no solo demo-

gráficas sino también de carácter económico. Por otra parte, ofrecemos un resultado actual que tiene multitud de aplicaciones.

La estructura de este informe es la siguiente. El capítulo 1 ofrece unas pinceladas sobre el problema de la transferencia de información espacial entre diferentes sistemas zonales, con especial hincapié en las estadísticas socioeconómicas. El capítulo 2 describe la información utilizada en el estudio. El capítulo 3 ofrece unos rasgos sobre la distribución de la población a nivel nacional con datos *inframunicipales*, el nomenclátor y las secciones censales, así como un breve examen de la heterogeneidad en los lindes municipales de nuestro país. Se indican los pros y los contras de la utilización de ambas fuentes de información demográfica para la construcción de una *grid* de población mediante *downscaling*.

A partir de estos resultados, el capítulo 4 examina, para el caso de España, la idoneidad de utilizar CLC para desagregar la población municipal a un formato tipo *grid* de elevada densidad, digamos 1 ha, como la construida por el JRC y accesible en el sitio web de la AEMA (Gallego 2010), o incluso de 1 km², como la difundida por *Eurostat-Geographical Information System at the Commission* (GISCO) en 2008 en formato gráfico. El argumento básico es que, dada la resolución de CLC y la distribución espacial de la población española, altamente concentrada en pequeños núcleos de población fuera de las grandes ciudades, esta fuente de información de coberturas del suelo producirá resultados de baja calidad en una gran parte del territorio nacional. Descartado CLC como información auxiliar sobre usos del suelo en el ejercicio de desagregación, el capítulo 5 describe la base de datos utilizada: el SIOSE.

El capítulo 6 describe con detalle la labor de *downscaling*, incluyendo un ejercicio de validación de nuestros métodos. El último capítulo ofrece unas conclusiones finales y una relación de usos potenciales de la *grid* elaborada. Finalmente, se incluyen una serie de apéndices con información adicional sobre CLC, SIOSE y otros aspectos técnicos referentes a los sistemas de referencia geodésicos.

1

Estadísticas socioeconómicas y sistemas zonales

Las diversas disciplinas del conocimiento humano recogen información en la forma que les es más natural, desde los experimentos puros de laboratorio, que nada tienen que ver con el tiempo y el espacio, hasta los trabajos de campo, en los que la referencia temporal y espacial forma parte esencial de la información recopilada. Cuando la geografía es importante siempre hay un sistema zonal de referencia detrás de las estadísticas, ya sea definido por la propia área de conocimiento o venga impuesto desde fuera de la misma.

Los investigadores no reflexionan normalmente sobre las implicaciones que su sistema zonal tiene sobre el conocimiento que generan; en muchas ocasiones este viene totalmente determinado de forma exógena y no tienen ninguna influencia sobre él. Sin embargo, cuando es necesario combinar la aportación de diferentes disciplinas, cuya información ha sido generada a partir de sistemas zonales diferentes, entonces debemos ponernos de acuerdo sobre el marco geográfico común a partir del cual efectuar la integración de estadísticas con una base geográfica.

Problemas similares se plantearon hace tiempo en el ámbito de las series temporales cuando era necesario, por ejemplo, combinar información anual con trimestral, mensual o incluso diaria. La literatura estadística desarrolló una serie de herramientas para poner toda esta información sobre una escala común, que van desde los primeros métodos puramente mecánicos (Denton 1971), hasta los más sofisticados de modelización probabilística (Harvey 1989), con información auxiliar o sin ella. Todo ello

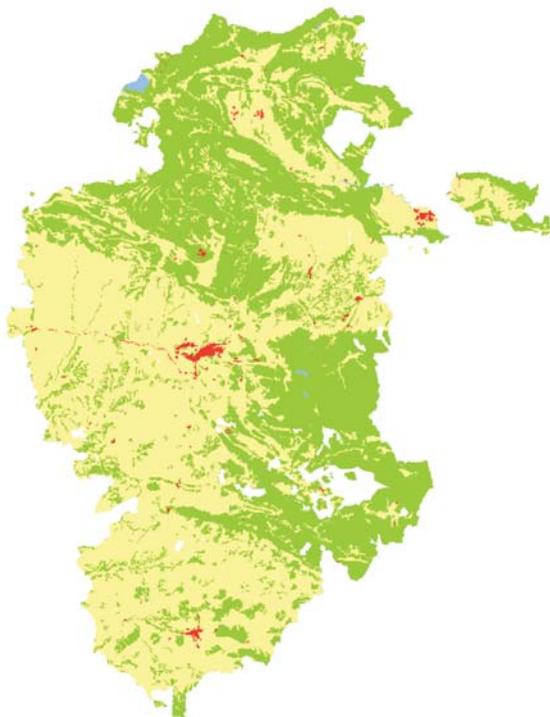
teniendo en cuenta la naturaleza de la información, en particular si las variables implicadas son flujos o *stocks*. En el ámbito espacial este problema es mucho más reciente, como consecuencia, en parte, de que hasta el desarrollo moderno del GIS la geografía no ha alcanzado el estatus de disciplina cuantitativa. Es también un problema mucho más difícil de abordar, puesto que la flecha del tiempo es unidireccional, pero en el espacio nos movemos en dos, o incluso tres direcciones.

Desde nuestro punto de vista *tres* son los sistemas zonales de interés a la hora de recopilar información con ámbito geográfico:

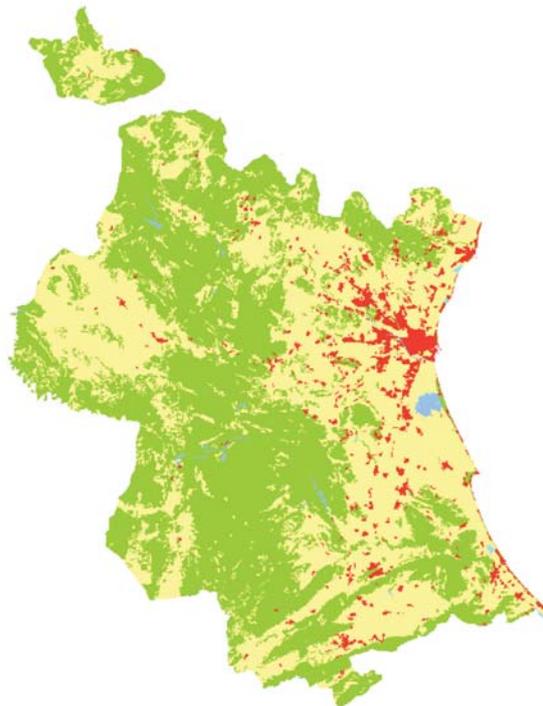
1. *Zonas funcionales* o geográficas definidas expresamente para una cuestión particular. Se construyen a partir de características geográficas relevantes, por ejemplo, las cuencas hidrográficas o zonas climáticas homogéneas, y pueden estar definidas por cuestiones meramente físicas, por ejemplo, cortes altimétricos, o incluir cierto grado de subjetividad en su definición, como las coberturas y usos del suelo, los espacios de biodiversidad (Natura 2000) o las zonas urbanas morfológicas (*urban morphological zones* [UMZ] 2006).

Su principal ventaja es que están diseñadas para un problema específico y concreto, pero con frecuencia carecen de homogeneización en sus definiciones y cuando se requiere la intervención de políticas sobre estas áreas siempre subsiste la cuestión de las competencias sobre las mismas.

a) Burgos



b) València/Valencia



Mapa 1.1 Coberturas de suelo de *CORINE Land Cover 2006*

Nota: Polígonos de municipios.

Fuente: Elaboración propia a partir de CLC2006 IGN (2010c).

El mapa 1.1 muestra las coberturas de CLC2006 al primer nivel para las provincias de Valencia y Burgos.

2. *Zonas administrativas o normativas:* municipios, regiones, países... Se trata de una división del territorio que es el resultado de un proceso histórico-político. Los límites están estrechamente vinculados a centros de poder y decisión política, lo que tiene ventajas en términos de quien tiene las competencias sobre el territorio, pero presenta los inconvenientes de que la información recopilada puede estar solo vagamente relacionada con la base territorial a partir de la cual se ha reunido. Típicamente estos sistemas zonales generan divisiones del territorio muy heterogéneas, lo que en muchos casos constituye un obstáculo para el análisis.

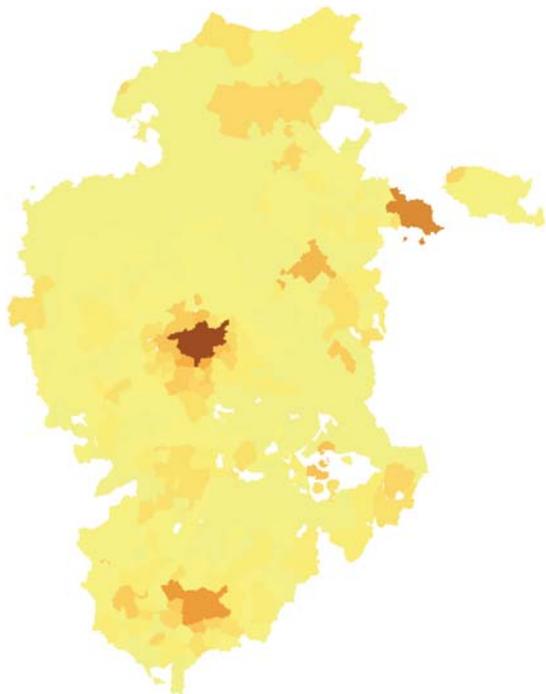
Generalmente, las estadísticas demográficas y socioeconómicas se organizan sobre la

base de este sistema zonal, fundamentalmente por ventajas en su recopilación. Normalmente los investigadores sociales toman este sistema como dado y no suelen pararse a reflexionar sobre cómo sus conclusiones podrían verse alteradas ante una modificación del sistema zonal de referencia.

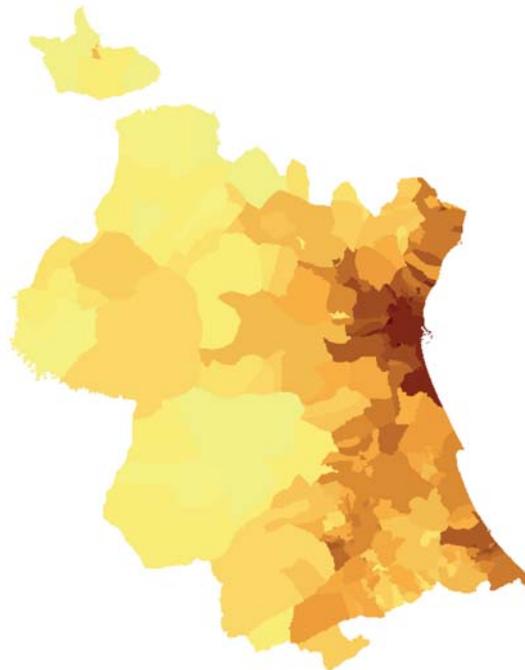
El mapa 1.2 muestra la densidad de población para el 2010 sobre la base de los términos municipales y las mismas dos provincias que aparecen en el mapa 1.1. Como tendremos ocasión de comprobar la distribución de la población es bastante diferente de lo que este mapa ilustra.

3. *Zonas geográficamente regulares o de rejilla cartográfica (grid)* en el que todas las unidades tienen el mismo tamaño. La base de este sistema zonal está en la propia cartografía, se trata de efectuar una partición del espacio con una base geo-

a) Burgos



b) València/Valencia

**Mapa 1.2** Densidad de población municipal 2010

Fuente: Elaboración propia a partir del INE (2010).

gráfica⁷, evitando cualquier hipótesis a priori sobre la distribución espacial de la información en la que estamos interesados y sin ninguna referencia ni a zonas funcionales ni a regiones administrativas. Las rejillas cartográficas existen desde la propia aparición de los mapas, ya que fueron utilizadas, en un principio, como sistema de referencia para la localización de accidentes geográficos.

Este sistema zonal ha ganado popularidad en los últimos años por varias razones: 1) la observación de la Tierra mediante satélite facilita la recopilación de información en este sistema zonal. De hecho la mayor parte

de la información de tipo medioambiental se recoge actualmente en este formato, y posteriormente se procesa en otro formato si se considera conveniente; 2) el desarrollo de las técnicas GIS ha hecho tremendamente eficiente la manipulación y procesado de grandes volúmenes de información almacenada en este tipo de formato. En este caso la *grid* es básicamente un paso intermedio para el procesamiento posterior de la información; 3) convenientemente definidas, las *grids* aumentan la comparabilidad entre las unidades de partida, al tener todas ellas el mismo tamaño. Ello es cierto independientemente de la información que contengan; 4) los últimos desarrollos muestran que las

⁷ El hecho de que todas las unidades posean el mismo tamaño no debe confundirse con que todas tengan la misma superficie. Al margen de las cuestiones relacionadas con las proyecciones (apéndice 3), podemos tener *grids* geográficas en las que el arco en longitud disminuye con la latitud, y en consecuencia las celdas son iguales desde el punto

de vista angular, pero no desde el punto de vista de la superficie. Este es el formato *grid* detrás de los modelos digitales de elevaciones (*digital elevation model* [DEM]) de cobertura mundial *Global Topographic Data* (GTOPO30) o *Shuttle Radar Topography Mission* (SRTM90) (Goerlich y Cantarino 2010a).

grids juegan un papel primordial en el transvase de información entre sistemas zonales, lo que es esencial en la integración entre estadísticas demográficas y socioeconómicas por una parte, y estadísticas medioambientales, por otra. En este caso, una *grid* puede ser el sistema zonal último en el que almacenar información, que será rápidamente comparable, o simplemente un eslabón intermedio en el proceso de manipulación de la misma.

El mapa 1.3 muestra la imposición de una *grid* sobre los sistemas zonales de los mapas 1.1 y 1.2, en los que la población está asignada al centroide del término municipal para facilitar la visualización.

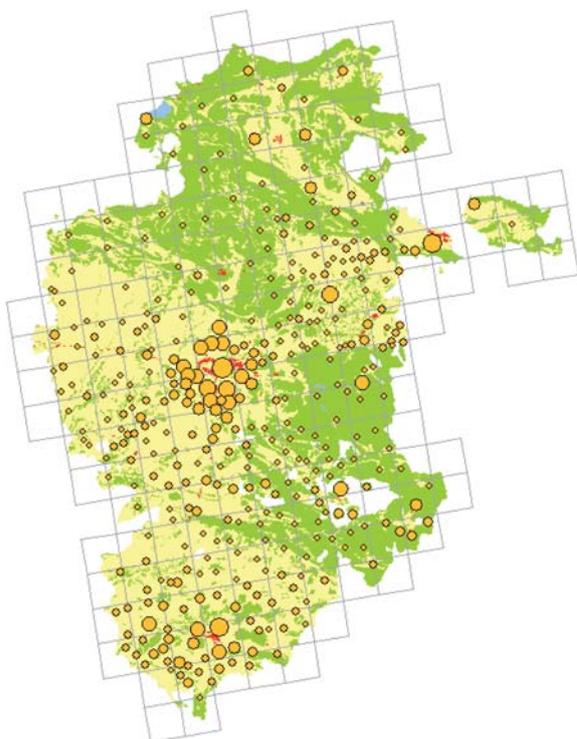
El mapa 1.4 ofrece una *grid* elemental de densidad de población, en la que a cada celda se le asigna el número de habitantes en función de determinados supuestos, información auxiliar sobre coberturas

del suelo y modelización estadística. Este es un ejemplo muy simple de una *grid* construida mediante procedimientos de desagregación espacial (*downscaling* o *top-down approach*). El ejercicio que presentamos en este informe es una versión mucho más sofisticada de esta forma de proceder.

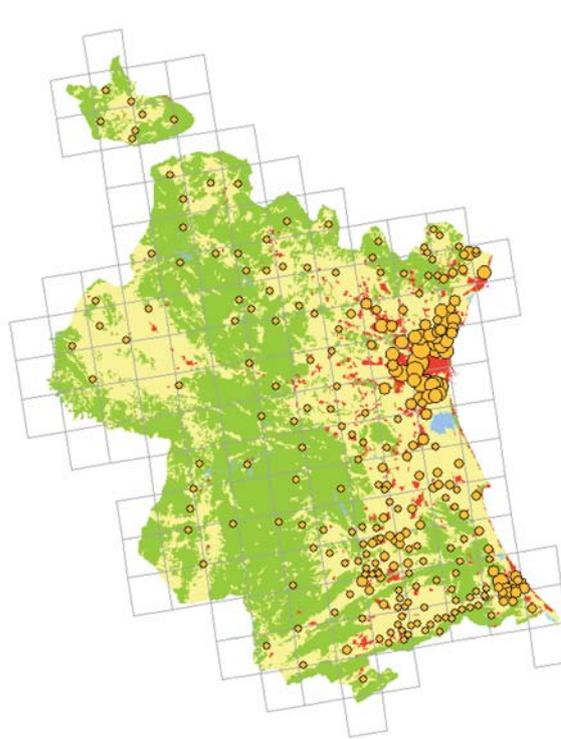
1.1. MÉTODOS

Los últimos años han visto un enorme desarrollo de los métodos estadísticos para la generalización e interpolación de superficies a partir de observaciones discretas en el espacio sobre ciertos fenómenos (triangulación, interpolación polinomial, suavizado *kernel*, *kriging*...). Estos métodos son rutinariamente aplicados a la distribución espacial de variables demográficas y socioeconómicas (Nordhaus 2002), sin tener en cuenta que la continuidad en el espacio no es, normalmente,

a) Burgos



b) València/Valencia

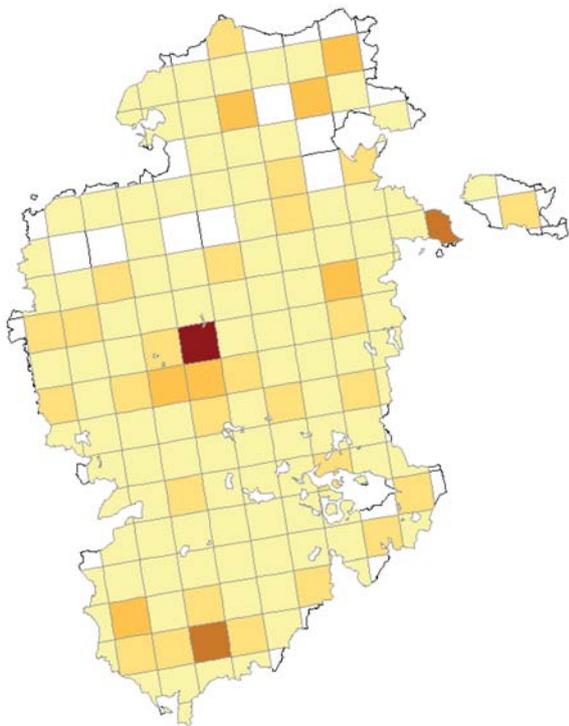


Mapa 1.3 Imposición de una *grid* sobre otros sistemas zonales

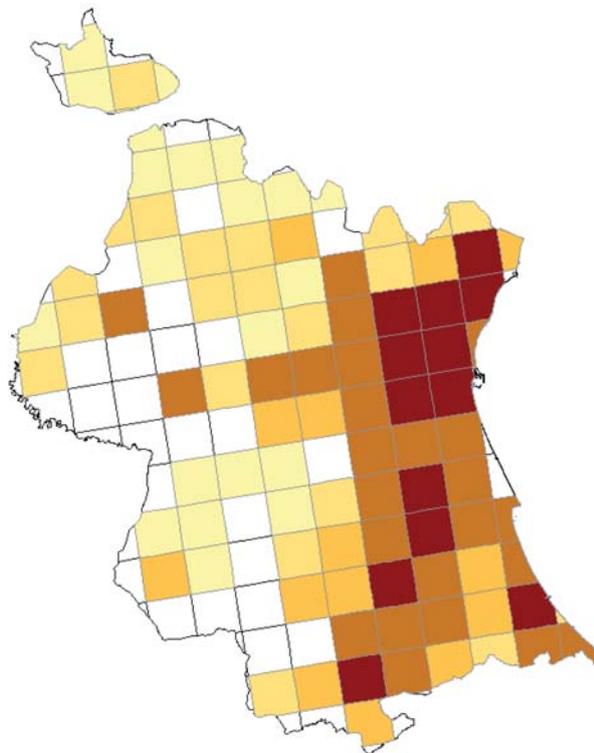
Nota: Polígonos de municipios.

Fuente: Elaboración propia a partir de CLC2006 IGN (2010c).

a) Burgos



b) València/Valencia

**Mapa 1.4** Grilla de población elemental de baja resolución

Nota: Polígonos de municipios.

Fuente: Elaboración propia a partir de CLC2006 IGN (2012c).

una característica de la distribución de este tipo de variables. No observamos población en cada punto del espacio.

Estos métodos fueron diseñados para fenómenos físicos espacialmente continuos (altitud, temperatura, presión atmosférica, humedad...), y no pueden ser aplicados de forma mecánica al análisis de distribuciones discretas en el espacio, como son típicamente aquellas relacionadas con indicadores sociales. Por ello es necesario el desarrollo de nuevas soluciones para este tipo de fenómenos. Exponemos a continuación brevemente algunos de estos métodos potenciales, haciendo especial hincapié en el tipo de técnicas utilizadas en este estudio.

Las dos operaciones básicas en la manipulación de la distribución de datos geográficos de carácter cuantitativo son la *agregación* y la *desagregación espacial*. La agregación consiste

en resumir los datos de origen en una zona de destino mayor, normalmente mediante un proceso de tipo aditivo.

La desagregación espacial, consistente en transferir datos de una zona de origen a diversas zonas de destino, cuando estas constituyen una partición de la primera. Este es un proceso más complejo cuya solución práctica depende de la naturaleza de la variable y de la información adicional de la que dispongamos. Si se trata de un total, como por ejemplo la población, podemos pensar simplemente en dividirla entre el número de zonas de destino o, lo que es más habitual, desagregarla proporcionalmente al área de dichas zonas (*areal weighting*). En este último caso estamos considerando implícitamente que la densidad de la variable en cuestión es uniforme en la zona de origen, lo que raramente es el caso en variables demográficas o socioeconómicas. El problema es similar a cuando queremos

desagregar la producción anual a la correspondiente a cada trimestre, *areal weighting* es equivalente a dividir la producción por cuatro, ignoramos de esta forma cualquier efecto de la estacionalidad o de las condiciones económicas a lo largo del año; no es difícil entender que esta forma de desagregación espacial no generará buenos resultados, excepto en casos muy excepcionales.

La forma natural de mejorar los resultados de la desagregación son dos:

1. Modelizar estadísticamente la correlación espacial de la variable en cuestión, un caso poco habitual en el contexto geográfico, al contrario de lo que sucede en el ámbito de las series temporales; o
2. incorporar información adicional. De esta forma la variable de la zona de origen se reparte sobre las zonas de destino según la distribución en las mismas de otra variable auxiliar. Esto puede hacerse a partir de una relación matemática conocida entre la variable objetivo y la variable auxiliar, o a partir de modelización estadística. Puesto que el tamaño es especialmente importante en el contexto espacial, normalmente esta modelización se realiza en términos de densidades.

Un caso particular de especial interés en el ámbito geográfico son las denominadas máscaras (*mask*), que señalan zonas que se han de excluir del análisis, es decir, áreas en las que la variable debe tomar el valor 0. Esta información también se utiliza a veces en procesos de agregación espacial.

Cuando diversas zonas implicadas en el análisis no forman una estructura jerárquica, por ejemplo, los municipios agregan a provincias, entonces es necesario efectuar una *intersección espacial* de las zonas implicadas: un producto cartesiano. Por ejemplo, si partimos de municipios y coberturas del suelo, la intersección espacial nos generará las coberturas de cada

municipio, y la proporción de su superficie cubierta. Alternativamente, también nos proporcionará los municipios que aparecen en cada una de las coberturas.

Con estas técnicas, un sencillo ejemplo nos permite observar como la información auxiliar puede mejorar la localización de variables para las que no se dispone de una referencia geográfica exacta. Supongamos que disponemos de la población de las cuatro unidades administrativas que se dibujan en el esquema 1.1, el total de población suma 100, de forma que los resultados pueden ser interpretados en términos de estructura porcentual, un mapa de densidades produciría resultados similares a los del mapa 1.2. Sin embargo, la información sobre coberturas del suelo nos acota la superficie artificial, fuera de la cual no puede existir población residente. Esta superficie está representada en color verde en el cuadrado derecho del esquema 1.1.

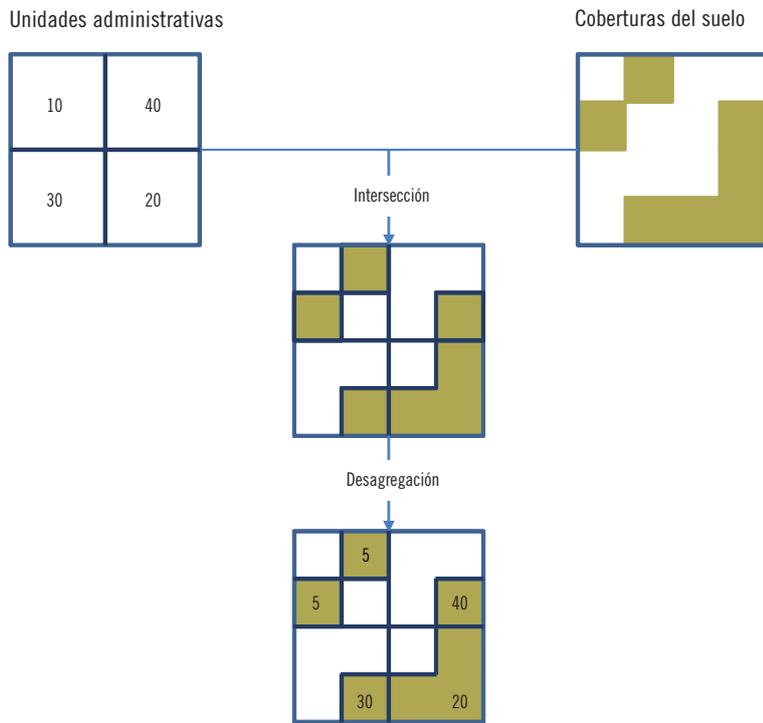
Una forma sencilla de mejorar la distribución de la población consiste en *interseccionar* ambas capas y efectuar la desagregación solamente sobre las coberturas artificiales dentro de cada unidad administrativa. Este es, esencialmente, un método dasimétrico binario, que asigna toda la población a una determinada cobertura (Langford y Unwin 1994; Langford 2007). La distribución final de la población tras este proceso podrá ser utilizada de forma más efectiva en un análisis espacial que combine otra información.

1.2. TRANSFERENCIA DE DATOS ENTRE DIFERENTES SISTEMAS ZONALES

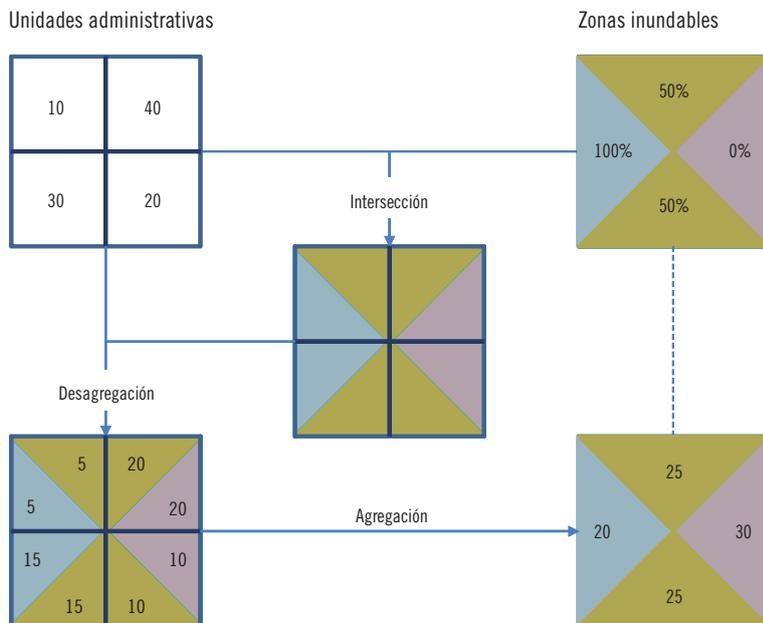
Consideremos ahora una cuestión ligeramente diferente, supongamos que deseamos saber la población en riesgo de ser afectada por una inundación. La población está disponible por unidades administrativas, pero el riesgo de inundación por zonas inundables⁸. El esquema 1.2 ilustra este ejemplo, el riesgo de inundación se

⁸ Curiosamente el *Web Map Service* (WMS), Sistema Nacional de Cartografía de Zonas Inundables del Ministerio de Medio Ambiente y Medio Rural y Marino (<http://sig.marm.es/snczi/>) contiene cuantiosa información geográfica

sobre zonas inundables, pero ninguna estimación de la población, los edificios, las infraestructuras y, más generalmente, los activos sujetos al riesgo de inundación (Gómez Andrés y Sánchez Martínez 2011).



Esquema 1.1 Aumentando la precisión de la localización espacial con información auxiliar
Fuente: Elaboración propia.



Esquema 1.2 Transferencia de datos entre sistemas zonales: *areal weighting*
Fuente: Elaboración propia.

cuantifica en tres niveles, con total certeza en una zona, con una probabilidad del 50% en otra y una tercera zona en la que no existe posibilidad de inundación bajo ningún supuesto.

Sin información adicional una forma aparentemente razonable de proceder consiste en tres pasos:

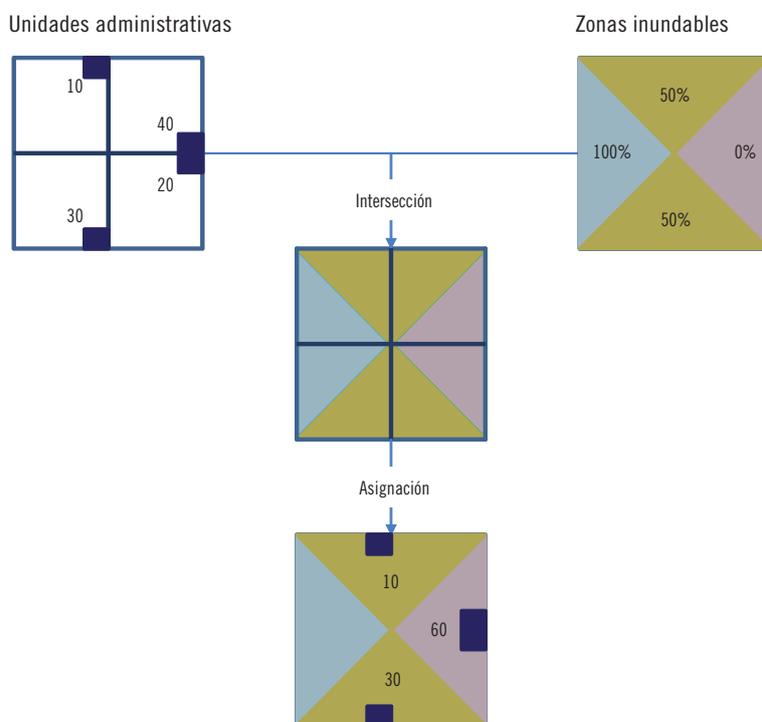
1. *Intersectamos* las unidades administrativas y las zonas inundables;
2. desagregamos la población en función de la superficie de cada área tras la intersección; y
3. finalmente, agregamos por zonas inundables.

El esquema 1.2 ilustra el proceso completo. Con los números del ejemplo, la población en riesgo ante una inundación son 45 personas. De nuestra población de 100 habitantes, 20 son afectados por la inundación con total seguridad, 50 con una probabilidad del 50% y 30 habitantes no se verán afectados bajo ninguna circunstancia.

Este es un ejemplo sencillo de transferencia de datos mediante el proceso simple de *areal weighting* que hemos introducido anteriormente (Flowerden, Green y Kehris 1991). Es conocido que suele generar resultados bastante pobres cuando la distribución de la variable dentro de las zonas de origen no es homogénea (Fisher y Langford 1995). La razón es sencilla, y puede ser ilustrada de forma natural.

Supongamos que la verdadera, pero desconocida, distribución de la población dentro de las unidades administrativas es la ilustrada por los rectángulos azul oscuro en el cuadrado izquierdo del esquema 1.3. La asignación de población a las zonas inundables es ahora inmediata. En este caso, la población en riesgo ante una inundación son solo 20 personas, pero lo que es más importante, no hay población afectada con total certidumbre, y 60 personas no se verán afectadas en ningún caso. Los resultados son, pues, bastante diferentes.

Podemos, sin embargo, mejorar el resultado final con información auxiliar procedente de coberturas



Esquema 1.3 Dificultades con la aplicación simplista de *areal weighting*

Fuente: Elaboración propia.

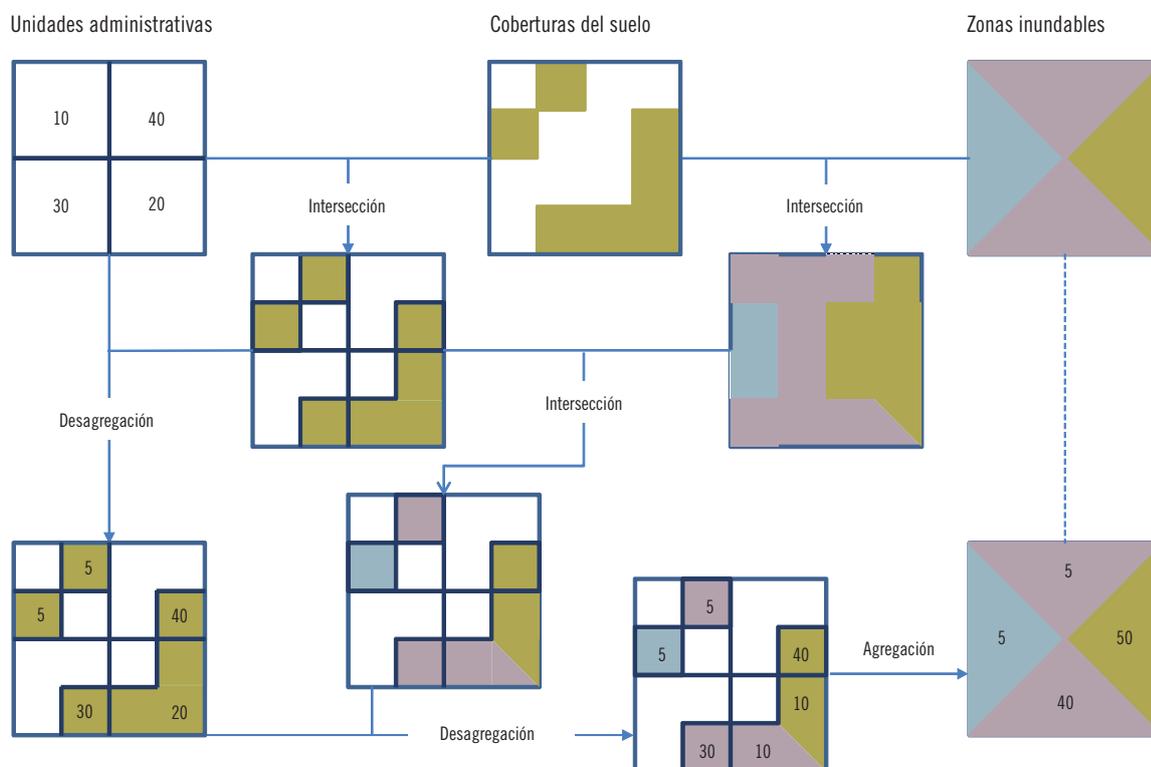
del suelo. El esquema 1.4 ilustra el proceso. Supongamos que disponemos de la capa sobre coberturas artificiales del esquema 1.1; como hemos visto, con ella mejoramos la localización de la población para disminuir el sesgo asociado a la simple distribución proporcional por superficie. Al mismo tiempo, la capa de coberturas del suelo se *intersecta* con las zonas inundables, que constituyen nuestras zonas de destino, para en una segunda desagregación asignar la población a dichas zonas. Finalmente, agregamos la población por zonas inundables. El resultado del proceso indica que, en nuestro ejemplo, la población en riesgo ante una inundación es del 27,5%, de los cuales solo el 5% lo es con certidumbre, y el 45% lo es con una probabilidad del 50%. La mitad de la población no está sujeta a este riesgo en ningún caso.

De esta forma, la aplicación en el proceso de zonas de control mejora sustancialmente los resultados finales. Esta forma de proceder no es solo

aplicable a problemas de distribución de la población sino a situaciones en la que las variables de interés son típicamente discretas en el espacio, como es el caso de muchas variables económicas (Kayadjanian y Vidal 2001).

Obsérvese que en los procesos que acabamos de describir la utilización de un sistema zonal común, por ejemplo una *grid*, tanto para la distribución de la población como para la generación de zonas inundables, lo que en terminología GIS se conoce como un modelo de datos *raster*, facilitaría enormemente las manipulaciones y el intercambio de datos de unas zonas a otras.

Este sencillo ejemplo, esquema 1.4, ilustra el enfoque que adoptaremos en el ejercicio que presentamos en este estudio, así como el tipo de datos que utilizaremos y que describimos en el capítulo siguiente. Nos instruye además sobre algunos hechos importantes y que son bien conocidos en la literatura:



Esquema 1.4 Transferencia de datos entre sistemas zonales: *areal weighting* con información auxiliar

Fuente: Elaboración propia.

1. Cuanto menos homogénea sea la distribución espacial de una variable sobre las zonas de origen, peor serán los resultados del método *areal weighting*. Dicho de otra forma, cuanto más discreta en el espacio sea la variable de interés, menos argumentos tenemos para aplicar mecánicamente los algoritmos sobre generación de superficies implementados en el *software* de GIS, que en la mayoría de los casos son utilizados como caja negra, y raramente son de aplicación a variables demográficas y socioeconómicas.
2. La fiabilidad del resultado final depende no solo de la resolución de los datos de partida respecto a la variable de interés, esto es, que los datos municipales serán preferibles a los datos regionales, sino sobre todo de la calidad y resolución de la información auxiliar. Esto es mucho más importante que la modelización estadística y el método particular de desagregación espacial utilizado (Martin, Tate y Langford 2000).
3. No es posible competir con una *grid* de densidad de población *bottom-up*. La georreferenciación exacta de los individuos permite conocer la distribución real de la población con arreglo al criterio adoptado, mientras que todos los métodos de desagregación espacial (*downscaling o top-down approach*) tienden a suavizar los resultados en cierta medida. Es importante en este caso, no solo validar el resultado final de alguna forma, sino comprobar que las discrepancias no son sistemáticas, ya que si lo fueran hay información externa potencial que podría ser incorporada al análisis.

2

La información de base

2.1. POBLACIÓN RESIDENTE: SECCIONES CENSALES

La *población* utilizada procede de la explotación estadística del padrón municipal de habitantes a nivel de *sección censal* para el año 2010. El *criterio* de población adoptado es, pues, el de *residencia*, o lo que se conoce como población nocturna, frente a la distribución de la población en los lugares de trabajo o población diurna (Hofstetter 2011)⁹. Dicha información puede ser descargada del sitio web del INE.

Descender en el nivel de agregación municipal es importante dada la distribución de la población en España y la diversa estructura de asentamientos en las diversas regiones. A nivel inframunicipal las fuentes de información demográfica son dos: las secciones censales, por una parte, y el nomenclátor de unidades poblacionales del INE del 2010, por otra. El capítulo siguiente examina brevemente la distribución de la población con arreglo a estas dos fuentes estadísticas que tienen una finalidad muy distinta, y ofrece argumentos a favor y en contra en ambos casos para el tipo de ejercicio que realizamos.

Si bien es cierto que, en principio, es la relación de unidades poblacionales del nomenclátor del

2010 la que proporciona una mayor información sobre la forma en que se asienta la población sobre el territorio, no es menos cierto que dichas unidades son de difícil tratamiento en el contexto del GIS. Por una parte, no se dispone de polígonos asociados a los núcleos de población, y por otra, la existencia de población clasificada como *diseminada* hace inviable la localización exacta de la misma dentro del municipio. Por su parte las secciones censales tienen territorio asignado que debe agregar el del término municipal, lo que facilita su manipulación en el contexto del GIS, pero por el contrario no están diseñadas para el estudio de la localización de la población, y presentan ciertas peculiaridades que es conveniente conocer. Examinamos estas cuestiones con más detalle en el capítulo 3, lo que permitirá entender con más claridad nuestra elección.

La construcción de la *grid* requiere una cartografía asociada a la población. En nuestro caso disponemos del fichero vectorial de polígonos de las secciones censales procedente del INE (en formato *shapefile*). En la elaboración de la *grid*, que se explica con detalle en el capítulo 6, este es el único fichero de contornos administrativos utilizado, de forma que los municipios y las provincias se obtienen por agregación del mismo¹⁰. Las

⁹ La población no es, sin embargo, un elemento fijo del paisaje, y por tanto sería deseable modelizar su movilidad en un contexto de *grids*, en la misma forma que se modeliza su movilidad entre unidades administrativas. Estos modelos de *grid* espacio-temporales, están actualmente en desarrollo (Bhaduri *et al.* 2007; Martin 2010; Bright y Bhaduri 2010), pero para ello es necesario tener una primera foto fija, es decir, la de la población residente. En este sentido, en España sería conveniente el desarrollo de modelos estacionales de *grids*, ya que

algunos municipios multiplican su población por un factor de 10 o más en períodos vacacionales puntuales, y además la concentración se hace sobre porciones muy reducidas del término municipal, como es el caso del litoral mediterráneo.

¹⁰ Esto es necesario por consistencia de la topología, dado que los contornos de las secciones censales, cuando son agregados a municipios, no coinciden exactamente con los contornos municipales de la base de líneas de límite del IGN.

superficies utilizadas en dicho capítulo también proceden de este fichero cuando hacen referencia a unidades administrativas, una vez proyectado adecuadamente (apéndice 3).

2.2. COBERTURAS DEL SUELO: SIOSE

La *información auxiliar* sobre *coberturas del suelo* procede de la base de datos del Sistema de Información de Ocupación del Suelo de España (SIOSE) del IGN, elaborada fundamentalmente con información del año 2005 (SIOSE2005). Dicha información puede ser descargada del Centro de Descargas del Centro Nacional de Información Geográfica (CNIG) dependiente del IGN.

El capítulo 5 se destina a dar una visión general de esta novedosa y potente fuente de información sobre coberturas del suelo en España, y el apéndice 2 ofrece una descripción más técnica.

2.3. OTRA INFORMACIÓN UTILIZADA EN EL ESTUDIO

Al margen de la información sobre población, secciones censales y su fichero de contornos correspondiente, y coberturas del suelo, SIOSE2005, a lo largo del estudio se utilizan muchas otras estadísticas.

En concreto, también se manejan las poblaciones municipales y regionales para la misma fecha de referencia, y así como el nomenclátor del 2010 sobre las entidades de población en el capítulo 3. Toda la información demográfica procede del INE.

La cartografía utilizada en los mapas que incluyen contornos administrativos procede de la base de datos de líneas de límite municipales del IGN¹¹. Cuando se requirieron contornos regionales o nacionales, estos se obtuvieron por

agregación de los municipios, con los territorios mancomunados o sin ellos, según fuera el caso. El capítulo 3 también utiliza las coordenadas de las entidades de población del nomenclátor, estas proceden del nomenclátor geográfico de municipios y entidades de población del IGN¹².

El fichero vectorial de la *grid* de referencia con la que trabajamos procede del sitio web del *European Forum for Geostatistics* (EFGS), por motivos de homogeneidad con Europa. El apéndice 3 ofrece detalles técnicos sobre sistemas de referencia geográfica que son de utilidad en nuestro contexto.

El estudio utiliza cálculos sobre superficies en diferentes partes, muchas veces implícitas en el cálculo de densidades de población. Por las razones que se exponen en el apéndice 3 este no es un tema trivial cuando tratamos con datos geográficos y requiere ser específico acerca de la fuente de información y homogeneidad en el tratamiento de la información. Las superficies finales utilizadas en la elaboración de la *grid* proceden de cálculos propios realizados con ArcMap™ de ArcGIS Desktop 9.3 (*Service Pack 1*) de ESRI (Environmental Systems Research Institute) poniendo especial cuidado en la proyección utilizada (apéndice 3). Los ficheros originales de SIOSE2005 contienen la superficie de los polígonos, que fue la que utilizamos, mientras que para las secciones censales se calculó con el *software* mencionado. El resto de superficies que aparecen en otros capítulos proceden del INE, Sección Territorio 2010; y cuando esto no es así se menciona de forma explícita.

El capítulo 4 utiliza además la base de datos de coberturas del suelo CLC2006. Dicha información procede del Centro de Descargas del CNIG, en formato vectorial, y se describe con algo de detalle en el apéndice 1. La información sobre superficies venía como atributo en el fichero descargado.

¹¹ Versión de mayo del 2011, convenientemente ajustada a nuestra fecha de referencia.

¹² Versión de junio del 2011.

2.4. UN COMENTARIO SOBRE LAS FECHAS DE REFERENCIA

Dado que el SIOSE está elaborado con información de 2005, SIOSE2005, la población más cercana a la base de datos de coberturas del suelo sería la de 2006, fechada a 1 de enero. Este sería el año natural para la elaboración de la *grid*, ya que como hemos mostrado en los ejemplos del capítulo anterior, la precisión de la información auxiliar es tremendamente relevante para la obtención de buenos resultados. Sin embargo, nuestra población de referencia está fechada 4 años después, en el 2010. Hay básicamente tres razones para esta decisión:

1. El año utilizado, 2010, es más cercano a la actualidad; lo que aumenta la novedad del trabajo. Esperamos poder comparar nuestros resultados con información del censo del 2011 cuando esté disponible, sobre todo si se distribuye información en formato de *grid* o por áreas geográficas definidas por el usuario, tal y como apunta el proyecto censal.
2. El objetivo del ejercicio es en gran medida metodológico, pretende ensayar métodos que superen la falta de resolución de CLC para este tipo de aplicaciones, tal y como se muestra en el capítulo 4. El SIOSE está en proceso de actualización a 2009, SIOSE2009; y para esta nueva versión la fecha de referencia idónea para la población sería el 2010. La investiga-

ción ha implicado una elevada carga en términos de ajustar las diferentes bases de datos y de programación de forma que, con la información adecuada ya depurada, los cálculos podrían ser rehechos en un tiempo reducido.

3. Pero la principal razón para asumir este desfase temporal entre la información sobre coberturas del suelo y la población radica en el hecho de que el fichero de polígonos de secciones censales del INE para el año 2006 contenía un elevado número de errores topológicos, tanto solapamientos entre polígonos como falta de contigüidad entre los mismos. Dada la dificultad de un tratamiento fluido de estos errores en el contexto del GIS preferimos utilizar una fecha más reciente para la población, en la que el fichero de polígonos de secciones censales careciera de errores topológicos¹³.

Conviene, sin embargo, tener una idea del orden de magnitud de los cambios en la superficie destinada a uso residencial durante este período. Como es bien sabido, el período 2006-2009 no ha sido homogéneo en cuanto a construcción residencial, el *boom* de 2006-2007 fue seguido de un fuerte período recesivo en los años siguientes que todavía no ha finalizado. Según el Ministerio de Fomento (2007, 2008, 2009, 2010) la superficie sobre el terreno ocupada por edificios residenciales durante el período 2006-2009 ascendió a 66 km²; de estos, 48 km² corresponden al período 2006-2007, y 18 km², al período 2008-2009¹⁴.

¹³ Aunque el fichero vectorial de polígonos de las secciones censales del 2010 carecía de problemas topológicos, no estaba completamente libre de errores. En concreto, presentaba algunos problemas que podemos clasificar en dos grupos:

1. Falta de consistencia en las secciones censales entre la cartografía y la explotación del padrón, probablemente por un problema de fechado de las mismas. Así, existen siete secciones censales que aparecen en la explotación del padrón pero no en el fichero de contornos; en la práctica esto supone una pérdida de población en la construcción de la *grid* de 6.197 habitantes (0,0132%). También existe el caso contrario, hay nueve secciones censales en el fichero de contornos que no aparecen en la explotación del padrón, esto no supone pérdida de población en la elaboración de la *grid*, pero genera problemas de mala localización de la misma.
2. Falta de consistencia entre los lindes municipales derivados al agragar el fichero de contornos de seccio-

nes censales y el fichero de la base de líneas de límite municipales del IGN. Muchas de estas inconsistencias carecen de relevancia por encontrarse dentro de un margen de error perfectamente tolerable. Sin embargo, hay algunos casos llamativos y que son claramente un error de asignación de superficies. Una cuantificación exacta de la población que se pierde al construir la *grid* debida a este tipo de errores no es posible por las razones que se explican en el capítulo 6.

¹⁴ Estas cifras son estimativas, ya que proceden de los datos recogidos de las licencias de obra concedidas por los ayuntamientos. En este sentido, 1) se estima un período medio de unos dieciocho meses para la terminación de la obra, y 2) además las viviendas finalmente terminadas (certificaciones de fin de obra) oscilan entre un 70 y un 95% de las iniciadas 2 años antes (Naredo, Carpintero y Marcos 2005, cap. 5, pág. 54).

El gráfico 2.1 muestra que esta tendencia es relativamente uniforme en todas las comunidades autónomas.

Esta cifra es ligeramente superior al 1% que SIOSE2005 ofrece como superficie cubierta por edificación, si bien esta clase incluye también las naves industriales (cuadro A.2.1). La edificación residencial constituye, no obstante, la mayor parte de nueva construcción, ya que la superficie ocupada por edificios no residenciales alcanzó solamente los 38 km² durante el mismo período, 2006-2009, según la misma fuente del Ministerio de Fomento (2007, 2008, 2009, 2010).

Por otra parte, es bien conocido que no todos los edificios de uso residencial albergan población

residente, lo que es particularmente cierto en muchos de los edificios de reciente construcción. Según estimaciones de Vinuesa (2008) a partir de los datos del censo del 2001, en España existían alrededor de 1,5 viviendas por hogar y sin duda esta cifra no ha disminuido durante la última década.

Por tanto, este es un problema que se debe tener en cuenta en la interpretación de resultados. Las cifras mencionadas no son despreciables, pero tampoco de una magnitud considerable. Solo la disponibilidad de un fichero de polígonos de secciones censales para 2006, en condiciones aceptables, permitiría acercar la población a la fecha de referencia de la base de datos de coberturas del suelo.

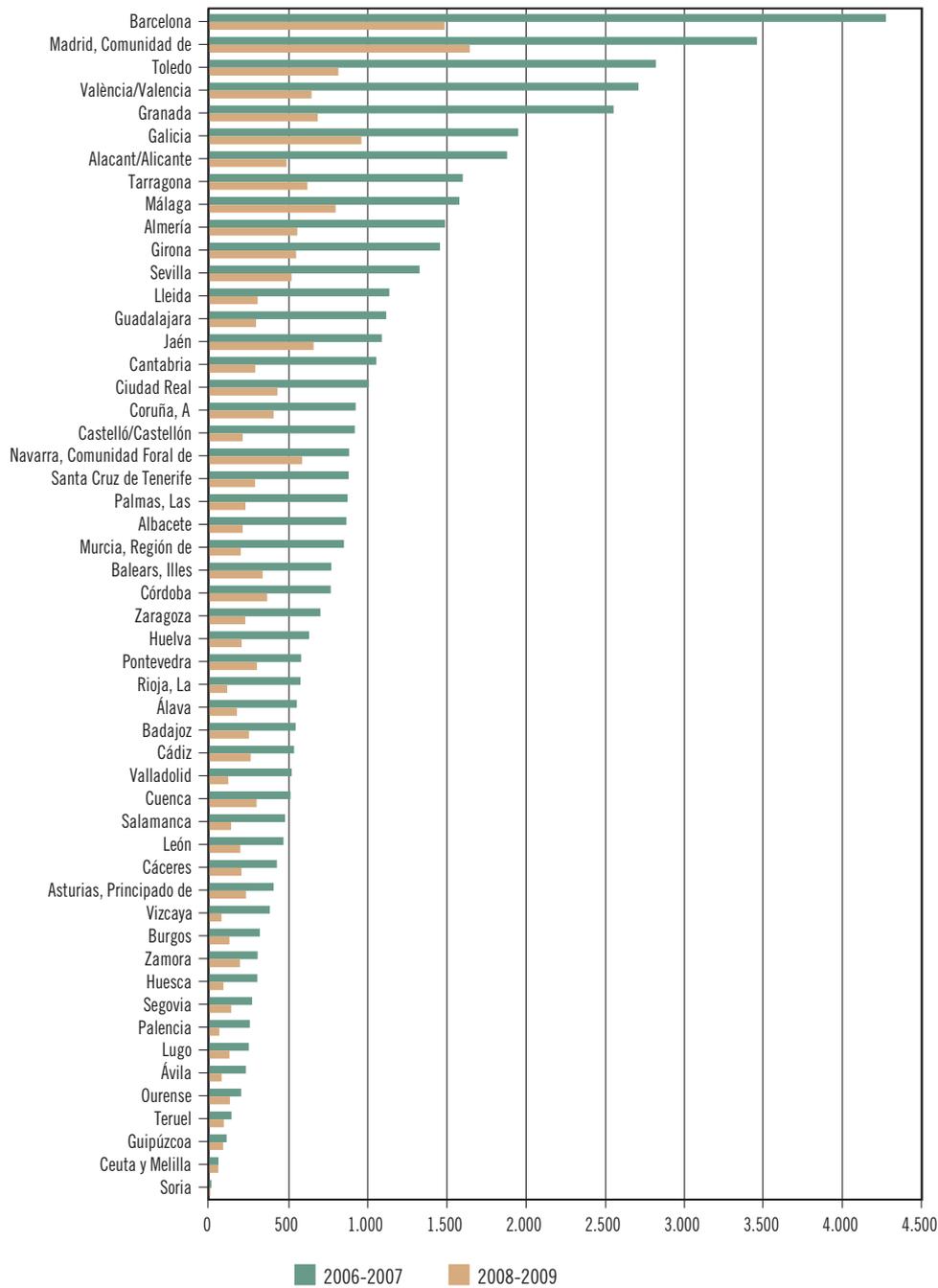


Gráfico 2.1 Superficie sobre el terreno que ocupará la edificación residencial. 2006/2007 y 2008/2009 (miles de m²)

Nota: Datos ordenados de mayor a menor según el bienio 2006-2007.

Fuente: Ministerio de Fomento (2007, 2008, 2009, 2010) y elaboración propia.

3

Distribución espacial de la población

Este capítulo examina brevemente la idoneidad de la estructura municipal para aproximarse al análisis de la distribución de la población sobre el territorio. A continuación, se exploran las posibilidades de mejora a partir de las estadísticas inframunicipales disponibles: el nomenclátor del 2010 y las secciones censales.

3.1. MUNICIPIOS

Como ya hemos señalado al principio del informe, para analizar la distribución de la población a partir de los datos municipales debemos preguntarnos por la homogeneidad de los municipios para este fin.

El gráfico 3.1 muestra un histograma de los tamaños de los municipios españoles. En él se observan dos características importantes:

1. Por una parte, una elevada heterogeneidad: el tamaño medio son 62,2 km², pero el tamaño mediano es prácticamente la mitad, 34,9 km²; por su parte, la desviación típica es 92,3 km², lo que arroja un coeficiente de variación de 1,5.
2. Por otra parte, emerge un número sustancial de municipios de muy reducidas dimensiones: el 28,1% de los municipios no supera los 20 km², y un 9,5% no rebasa los 10 km². En el extremo inferior de la

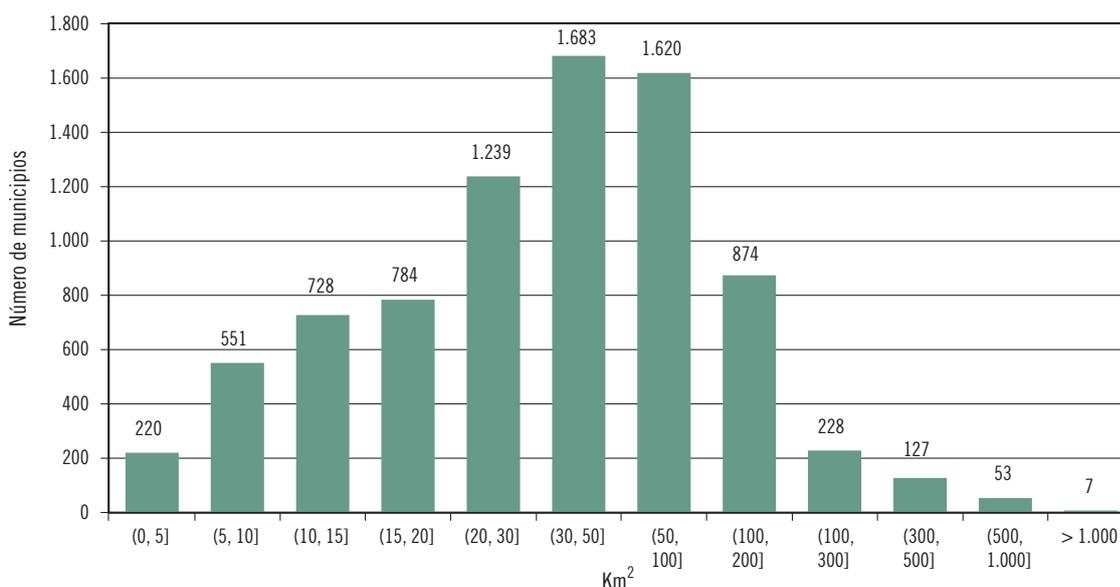


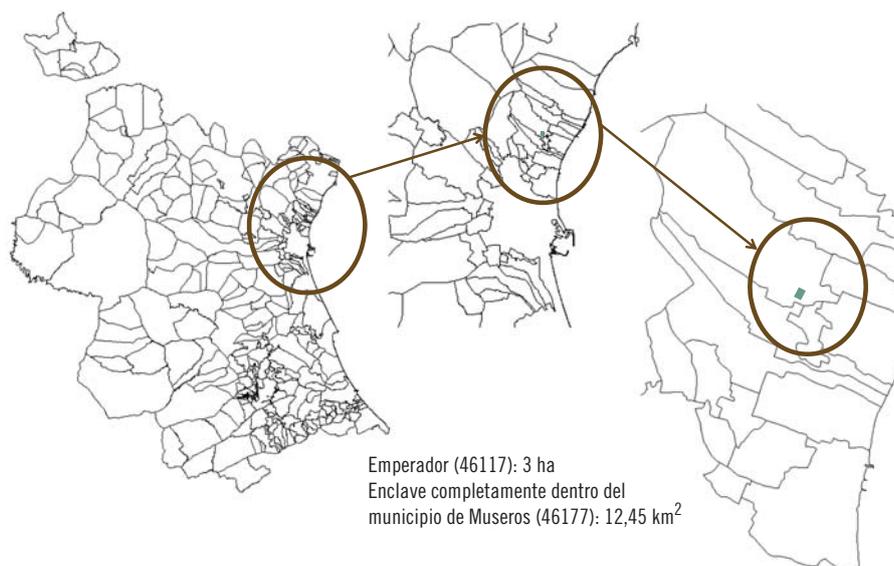
Gráfico 3.1 Distribución de municipios por tamaño (2010)

Fuente: Elaboración propia a partir del INE (2012).

distribución, encontramos 17 municipios cuyo término municipal no alcanza el km².

El rango en la distribución de tamaños es extremadamente elevado. El municipio más pequeño ocupa una superficie de solo 3 ha, Emperador en la provin-

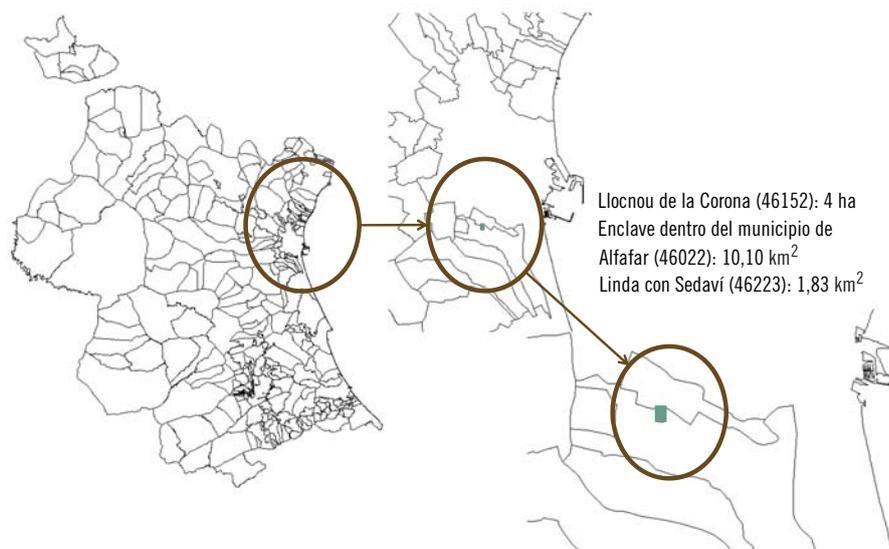
cia de Valencia; el siguiente en el *ranking* tiene solo 1 ha más, Llocnou de la Corona en la misma provincia. Su situación se muestra en los mapas 3.1 y 3.2. En ambos casos se trata de pequeños enclaves dentro de municipios de ya muy reducidas dimensiones, y en el área metropolitana de Valencia.



Mapa 3.1 Localización de Emperador (46117, Valencia)

Nota: Valencia: 266 municipios. 10.806 km².

Fuente: Elaboración propia.



Mapa 3.2 Localización de Llocnou de la Corona (46152, Valencia)

Nota: Valencia: 266 municipios. 10.806 km².

Fuente: Elaboración propia.

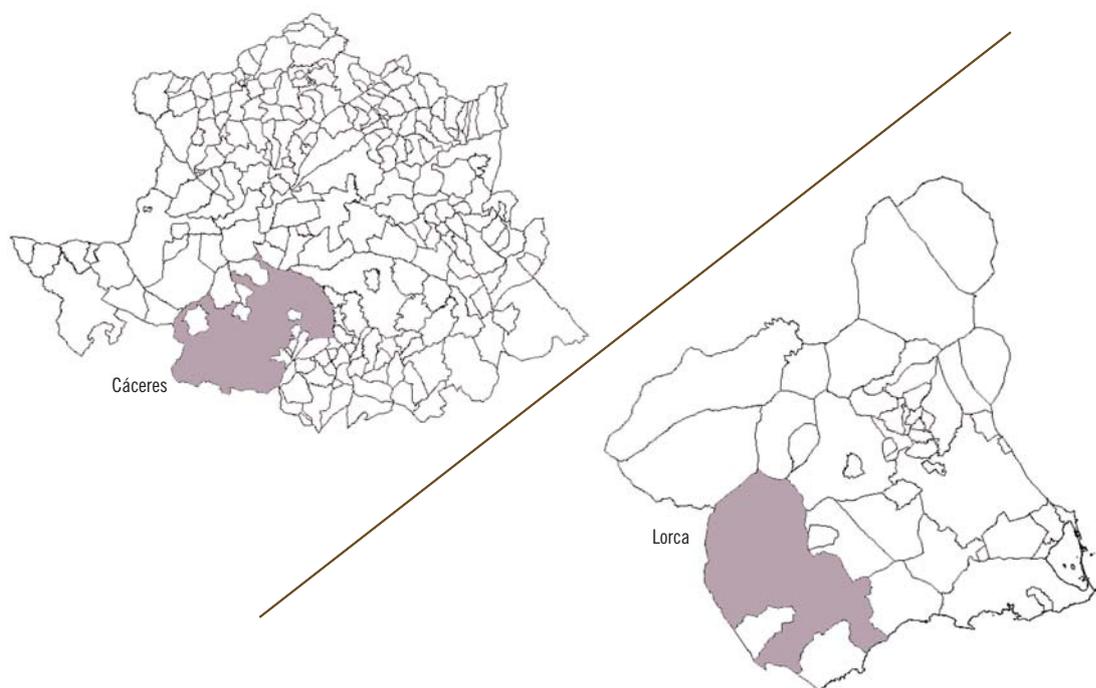
En el otro extremo de la distribución, el municipio más grande de España es una capital de provincia, Cáceres, con 1.750,33 km², seguido por Lorca, en la provincia de Murcia, con 1.675,21 km². El mapa 3.3 muestra estos dos municipios dentro de sus respectivas provincias, que en los dos casos muestran distribuciones de tamaños municipales muy dispares.

El cuadro 3.1 ofrece información sobre la superficie ocupada por los diferentes tamaños de municipios considerados, así como la población que albergan según el padrón municipal del 2010. Ya hemos observado cómo un 9,5% de los municipios no superan los 10 km², y un 28,1% no alcanzan los 20 km²; estos grupos acogen a un 5,4 y un 13,3% de la población, respectivamente. Por su parte, los 17 municipios con superficie inferior al km² albergan a 53.009 residentes, ya que entre ellos se en-

cuentran dos municipios con más de 10.000 habitantes, Badia del Vallès en Barcelona, con 13.703 habitantes, y Benetússer en Valencia, con 15.399 habitantes.

En el otro extremo, solo un 0,8% de los municipios supera los 500 km², acogiendo al 13,9% de la población. Este enorme volumen de población en tan reducido número de municipios, 60, se debe a la presencia de Madrid, la capital del Estado, que con una superficie de 605,77 km² alberga a más de 3 millones de habitantes¹⁵. De hecho la relación entre superficie y población es claramente no lineal, pero muestra un coeficiente de correlación ligeramente positivo, 0,18, que se eleva a 0,34 cuando tomamos logaritmos.

Los municipios españoles no son solo heterogéneos y de reducidas dimensiones en términos de



Mapa 3.3 Localización de Cáceres (10037) y Lorca (30024, Murcia)

Nota: Cáceres: 221 municipios. 19.868,20 km². Murcia: 45 municipios. 11.313,16 km².

Fuente: Elaboración propia.

¹⁵ Exactamente 3.273.049 residentes según el padrón

del 2010.

CUADRO 3.1: Distribución de municipios por tamaño. Superficie ocupada y población en el 2010

Tamaño (km ²)	Municipios		Superficie ocupada		Población	
	Número	Porcentaje	Km ²	Porcentaje	Habitantes	Porcentaje
(0, 5]	220	2,7	690,94	0,1	940.238	2,0
(5, 10]	551	6,8	4.240,53	0,8	1.622.147	3,4
(10, 15]	728	9,0	9.129,65	1,8	1.975.299	4,2
(15, 20]	784	9,7	13.742,51	2,7	1.732.515	3,7
(20, 30]	1.239	15,3	30.512,41	6,0	3.158.352	6,7
(30, 50]	1.683	20,7	65.348,63	12,9	7.221.216	15,4
(50, 100]	1.620	20,0	113.170,20	22,4	9.032.442	19,2
(100, 200]	874	10,8	119.228,54	23,6	8.863.388	18,8
(200, 300]	228	2,8	54.817,10	10,9	3.599.523	7,7
(300, 500]	127	1,6	47.650,69	9,4	2.336.192	5,0
(500, 1000]	53	0,7	36.431,14	7,2	5.487.777	11,7
> 1000	7	0,1	9.673,70	1,9	1.051.942	2,2
Total	8.114	100,0	504.636,04	100,0	47.021.031	100,0

Fuente: Elaboración propia a partir del INE (2010 y 2012).

la superficie que ocupan, sino también, y sobre todo, en términos del volumen de población que reside en ellos. En el 2010, el tamaño poblacional medio era de 5.795 habitantes, pero el municipio mediano tenía tan solo 582 residentes. Así pues, un 59,9% de los municipios tenía una población inferior a los 1.000 habitantes, albergando tan solo al 3,2% de la población residente. Las cifras son más llamativas todavía si nos movemos hacia los extremos de la distribución. En el 12,8% de los municipios el número de residentes no alcanzaba los 100 habitantes, y en conjunto estos municipios albergaban solamente al 0,13% de la población. Traducido a números absolutos, en 1.041 municipios residían 60.486 personas. El municipio con menos población, Illán de Vacas en la provincia de Toledo, contaba con 5 habitantes¹⁶; y el siguiente en el *ranking*, Jaramillo Quemado en la provincia de Burgos, solamente contaba con un residente más. En el extremo opuesto, los dos municipios mayores, Madrid y Barcelona, representaban el 10,4% del

total de población residente, aproximándose a los 5 millones de habitantes.

El cuadro 3.2 ofrece la distribución por provincias de los municipios por tamaños de superficie. El mensaje es claro y se percibe inspeccionando los mapas 3.2 y 3.3: la estructura es muy heterogénea a nivel provincial. En algunas provincias hay una clara concentración de municipios de muy reducidas dimensiones, como por ejemplo Valencia¹⁷, Guipúzcoa o Barcelona. En otras predominan claramente los municipios grandes, es el caso de Murcia, Albacete, Ciudad Real, Córdoba, Badajoz o Cáceres. Por su parte, el cuadro 3.3 ofrece información sobre la población a nivel provincial y estadísticos descriptivos sobre los tamaños municipales en términos de población. La abundancia de municipios de muy reducida población, concentrada en determinadas provincias, es patente.

Como consecuencia de esta elevada heterogeneidad es difícil generalizar, y resultados derivados

¹⁶ Curiosamente este municipio aparece en el censo de 1991 con nula población, tanto de hecho como de derecho, lo que lleva a preguntarse por qué no desapareció.

¹⁷ La provincia de Valencia muestra una elevada concen-

tración de municipios minúsculos, el 39,8% no superan los 10 km² y 8 de los 17 municipios de España que no alcanzan el km² de superficie se encuentran en esta provincia.

CUADRO 3.2: Distribución provincial de municipios por tamaño (km²) (2010)

(número de municipios)

Provincias	(0, 10]	(10, 20]	(20, 50]	(50, 100]	(100, 500]	> 500	Total
01 Álava	6	8	16	13	8		51
02 Albacete		2	12	24	42	7	87
03 Alacant/Alicante	33	31	41	25	11		141
04 Almería	4	11	35	30	21	1	102
05 Ávila	20	93	95	33	7		248
06 Badajoz	5	8	39	48	59	5	164
07 Balears, Illes	1	11	22	15	18		67
08 Barcelona	69	84	128	29	1		311
09 Burgos	48	124	126	41	32		371
10 Cáceres	6	20	78	60	52	5	221
11 Cádiz		4	7	12	19	2	44
12 Castelló/Castellón	10	30	49	30	16		135
13 Ciudad Real	1	2	11	23	59	6	102
14 Córdoba	1	7	11	13	37	6	75
15 Coruña, A	1	2	23	40	28		94
16 Cuenca	2	10	94	86	45	1	238
17 Girona	53	68	72	26	2		221
18 Granada	17	25	48	46	30	2	168
19 Guadalajara	8	75	142	43	20		288
20 Guipúzcoa	31	22	26	8	1		88
21 Huelva	1	5	23	15	34	1	79
22 Huesca	12	27	62	48	52	1	202
23 Jaén	1	3	22	23	46	2	97
24 León	1	14	81	65	50		211
25 Lleida	20	49	73	53	36		231
26 Rioja, La	38	56	55	21	4		174
27 Lugo	2		1	14	50		67
28 Madrid	7	24	99	38	10	1	179
29 Málaga	7	13	38	21	20	1	100
30 Murcia		8	10	5	14	8	45
31 Navarra	59	58	93	50	12		272
32 Ourense	2	1	28	38	23		92
33 Asturias	2	1	15	25	33	2	78
34 Palencia	5	40	98	35	13		191
35 Palmas, Las		1	12	5	16		34
36 Pontevedra	3	1	25	18	15		62
37 Salamanca	29	80	190	53	10		362
38 Santa Cruz de Tenerife	1	5	24	11	13		54

CUADRO 3.2 (cont.): Distribución provincial de municipios por tamaño (km²) (2010)
(número de municipios)

Provincias	(0, 10]	(10, 20]	(20, 50]	(50, 100]	(100, 500]	> 500	Total
39 Cantabria	7	15	45	28	7		102
40 Segovia	16	70	97	17	9		209
41 Sevilla	9	14	20	20	38	4	105
42 Soria	11	36	66	44	26		183
43 Tarragona	36	43	62	30	12		183
44 Teruel	1	19	113	71	32		236
45 Toledo	6	27	73	48	49	1	204
46 València /Valencia	106	38	57	35	29	1	266
47 Valladolid	4	61	116	34	10		225
48 Vizcaya	37	40	30	3	2		112
49 Zamora	6	61	114	51	16		248
50 Zaragoza	26	63	105	56	40	3	293
51 Ceuta		1					1
52 Melilla		1					1
España	771	1.512	2.922	1.620	1.229	60	8.114

Nota: Las casillas vacías tienen valor 0.

Fuente: Elaboración propia a partir del INE (2012).

CUADRO 3.3: Municipios, población y estadísticos de tamaño poblacional por provincias (2010)

Provincias	Municipios		Población		Estadísticos descriptivos de tamaño poblacional			
	N.º	Porcentaje	Habitantes	Porcentaje	Mínimo	Mediana	Media	Máximo
01 Álava	51	0,6	317.352	0,7	175	866	6.223	238.247
02 Albacete	87	1,1	401.682	0,9	58	1.217	4.617	170.475
03 Alacant /Alicante	141	1,7	1.926.285	4,1	38	2.948	13.662	334.418
04 Almería	102	1,3	695.560	1,5	84	1.069	6.819	190.013
05 Ávila	248	3,1	171.896	0,4	18	156	693	58.245
06 Badajoz	164	2,0	692.137	1,5	93	1.459	4.220	150.376
07 Balears, Illes	67	0,8	1.106.049	2,4	276	5.750	16.508	404.681
08 Barcelona	311	3,8	5.511.147	11,7	24	3.078	17.721	1.619.337
09 Burgos	371	4,6	374.826	0,8	6	128	1.010	178.574
10 Cáceres	221	2,7	415.083	0,9	67	665	1.878	94.179
11 Cádiz	44	0,5	1.236.739	2,6	479	9.505	28.108	208.896
12 Castelló/Castellón	135	1,7	604.274	1,3	16	618	4.476	180.690
13 Ciudad Real	102	1,3	529.453	1,1	103	1.239	5.191	74.345
14 Córdoba	75	0,9	805.108	1,7	374	3.622	10.735	328.547
15 Coruña, A	94	1,2	1.146.458	2,4	1.277	5.582	12.196	246.047
16 Cuenca	238	2,9	217.716	0,5	16	236	915	56.189
17 Girona	221	2,7	753.046	1,6	89	726	3.407	96.236

CUADRO 3.3 (cont.): Municipios, población y estadísticos de tamaño poblacional por provincias (2010)									
Provincias	Municipios		Población		Estadísticos descriptivos de tamaño poblacional				
	N.º	Porcentaje	Habitantes	Porcentaje	Mínimo	Mediana	Media	Máximo	
18	Granada	168	2,1	918.072	2,0	146	1.416	5.465	239.154
19	Guadalajara	288	3,5	251.563	0,5	12	97	873	83.789
20	Guipúzcoa	88	1,1	707.263	1,5	120	1.874	8.037	185.506
21	Huelva	79	1,0	518.081	1,1	50	2.194	6.558	149.310
22	Huesca	202	2,5	228.566	0,5	24	275	1.132	52.347
23	Jaén	97	1,2	670.761	1,4	414	2.940	6.915	116.790
24	León	211	2,6	499.284	1,1	56	609	2.366	134.012
25	Lleida	231	2,8	439.768	0,9	57	463	1.904	137.387
26	Rioja, La	174	2,1	322.415	0,7	9	176	1.853	152.650
27	Lugo	67	0,8	353.504	0,8	207	2.858	5.276	97.635
28	Madrid	179	2,2	6.458.684	13,7	49	3.176	36.082	3.273.049
29	Málaga	100	1,2	1.609.557	3,4	127	2.664	16.096	568.507
30	Murcia	45	0,6	1.461.979	3,1	582	15.481	32.488	441.345
31	Navarra	272	3,4	636.924	1,4	17	366	2.342	197.488
32	Ourense	92	1,1	335.219	0,7	447	1.659	3.644	108.673
33	Asturias	78	1,0	1.084.341	2,3	170	2.587	13.902	277.198
34	Palencia	191	2,4	172.510	0,4	11	164	903	82.169
35	Palmas, Las	34	0,4	1.090.605	2,3	823	16.202	32.077	383.308
36	Pontevedra	62	0,8	962.472	2,0	734	6.477	15.524	297.124
37	Salamanca	362	4,5	353.619	0,8	18	229	977	154.462
38	Santa Cruz de Tenerife	54	0,7	1.027.914	2,2	1.180	6.522	19.035	222.643
39	Cantabria	102	1,3	592.250	1,3	72	1.778	5.806	181.589
40	Segovia	209	2,6	164.268	0,3	24	171	786	55.748
41	Sevilla	105	1,3	1.917.097	4,1	350	6.516	18.258	704.198
42	Soria	183	2,3	95.258	0,2	9	89	521	39.838
43	Tarragona	183	2,3	808.420	1,7	42	925	4.418	140.184
44	Teruel	236	2,9	145.277	0,3	8	152	616	35.241
45	Toledo	204	2,5	697.959	1,5	5	1.672	3.421	88.986
46	València /Valencia	266	3,3	2.581.147	5,5	36	2.112	9.704	809.267
47	Valladolid	225	2,8	533.640	1,1	28	220	2.372	315.522
48	Vizcaya	112	1,4	1.153.724	2,5	112	1.725	10.301	353.187
49	Zamora	248	3,1	194.214	0,4	40	304	783	65.998
50	Zaragoza	293	3,6	973.252	2,1	12	260	3.322	675.121
51	Ceuta	1	0,0	80.579	0,2			80.579	
52	Melilla	1	0,0	76.034	0,2			76.034	
	España	8.114	100,0	47.021.031	100,0	5	582	5.795	3.273.049

Nota: Las casillas vacías tienen valor 0.

Fuente: Elaboración propia a partir del INE (2012).

a nivel nacional presentan importantes desviaciones cuando descendemos al nivel provincial.

No es de extrañar que esta conjunción de heterogeneidad en superficie y en población, unida a la poca relación directa entre ambas variables, genere densidades de población municipal extremadamente dispares, y en ocasiones poco razonables desde el punto de vista del análisis de la distribución de la población sobre el territorio. La densidad promedio para el conjunto del país en el 2010 es de 93 habitantes por km²; pero 93 municipios no alcanzan al habitante por km² de densidad y 1.810 no superan los 5 habitantes por km²; sin embargo, estos municipios representan la quinta parte de la superficie. Fijando el umbral en los 10 habitantes por km², el número de municipios que no superan esa cifra casi se duplica, 3.255, representando ya un 40% de la superficie nacional. En el otro extremo, 7 municipios muestran densidades de población superiores a los 15.000 habitantes por km², pero cubren una superficie de solamente 123,18 km², entre ellos está Barcelona, con 98,21 km², por lo que si la excluimos observamos que los municipios de mayor densidad son de una superficie minúscula, 2 de ellos no alcanzan el km². El municipio con mayor densidad de España, Mislata en la provincia de Valencia, alcanza los 21.258 habitantes por km², pero apenas supera los 2 km² de superficie; el siguiente en el *ranking* resulta ser el municipio más pequeño, Emperador, que con una población de 615 residentes arroja una densidad superior a los 20.000 habitantes por km². Con la excepción de Barcelona, todos estos municipios son pequeños enclaves en una gran área urbana, un ejemplo más del corsé que impone la difusión de resultados estadísticos basados en límites administrativos y que nada tienen que ver con el tema que se desea tratar.

La conclusión que se deriva de este análisis inicial es que la estructura municipal no alcanza la homogeneidad suficiente para un análisis preciso de la distribución de la población sobre el territorio. El procedimiento habitual de extender la población de forma uniforme sobre el término municipal en el que reside, con bruscos cambios al traspasar los lindes, es claramente insatisfactorio¹⁸. En cualquier caso el mapa 3.4 muestra, para futuras referencias, la densidad de población municipal calculada de esta forma, y que ya ofrecimos como parte del mapa 3, aunque ahora los intervalos considerados difieren, lo que también afecta a nuestra percepción de la realidad dibujada en el mapa (Eurostat 1999a, 2005; ESPON 2010).

Antes de pasar a examinar la distribución de la población a partir de estadísticas inframunicipales, conviene detenerse brevemente en algunas peculiaridades de los contornos de nuestros municipios, cuyo origen tiene sin duda un alto componente histórico, y sobre las que han llamado la atención los encargados de mantener la base de datos de líneas de límite (Rodríguez, Martín-Asín y Astudillo 1997). Ello explica parcialmente lo difícil que ha sido, hasta fechas muy recientes, conseguir una cartografía digital completa y libre de errores de los contornos municipales españoles.

Ya hemos señalado en la Introducción que, incluso en la actualidad, los municipios de Ansó y Fago comparten un único polígono, al no existir documentación oficial de la delimitación de ambos términos municipales por separado¹⁹. Además de esta peculiaridad, resulta curioso señalar que la partición del territorio nacional en términos municipales no es exhaustiva, es decir, existen territorios que no pertenecen a un

¹⁸ Adicionalmente, cuando la densidad se mide en habitantes por km² es razonable preguntarse por el sentido de calcular densidades de población en municipios con una superficie inferior al km², ya que entonces no estamos distribuyendo población, sino escalándola hasta que el municipio alcanza la unidad mínima de superficie considerada en nuestra escala de medida. Resulta paradójico que a Emperador, con 3 ha de superficie y una población de 615 habitantes, se le asigne una densidad de 20.500 habitantes por km², por muy correcto que sea el cálculo.

¹⁹ El INE, sin embargo, sí atribuye información estadística por separado para ambos municipios; incluso en la

sección de Territorio, atribuye una superficie a Fago de 29 ha, lo que con una población de 26 habitantes supone una densidad de 89,7 habitantes por km², dicha superficie debe corresponder solamente al casco urbano, y es notablemente más reducida que la que se deriva del polígono de la sección censal correspondiente, idéntico al que aparece en INE (1994), y que resulta ser de 28,75 km². Curiosamente la superficie ofrecida por el INE en la sección de Territorio para el municipio de Ansó, 251,83 km², se corresponde con la que el IGN certifica como superficie global de ambos municipios, Fago y Ansó.

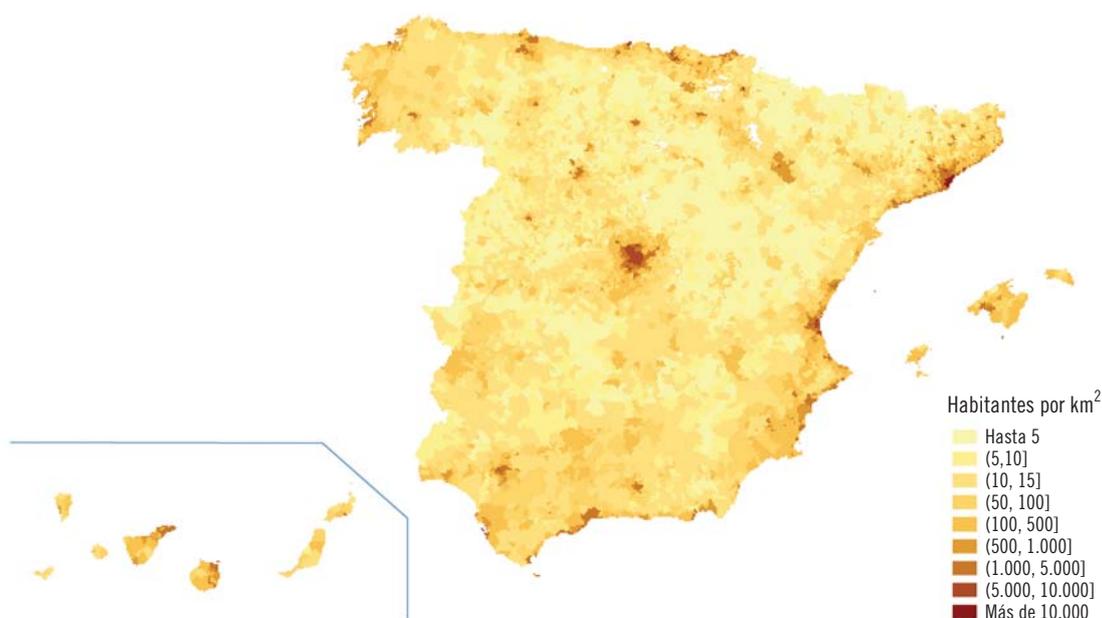
único municipio, sino que administrativamente pertenecen a varios municipios o incluso directamente al Estado, aunque sí tienen adscripción provincial. Son los llamados *condominios*, normalmente territorios de pequeña superficie, aunque las Bârdenas Reales en Navarra supera los 400 km². Se observan sobre todo en Navarra y Burgos. El mapa 3.5 ofrece la situación de los mismos.

Existen municipios con una elevada fragmentación espacial de su superficie. El mapa 3.6 muestra el municipio de Xàtiva, en la provincia de Valencia, cuyo término municipal se encuentra fragmentado, a escala 1:25.000, en 30 polígonos, algunos de ellos de dimensiones irrisorias, notablemente inferiores a la ha. Otros municipios adoptan formas caprichosas, poseen enclaves deshabitados con un único punto de tangencia al polígono principal, como es el caso de Mohedas de Granadilla, en la provincia de Cáceres (mapa 3.7); o engloban a otros municipios de reducidas dimensiones, fruto de tradiciones históricas, como es el caso del municipio de Albarracín, en la provincia de Teruel, que rodea completamente a otros 19 municipios de reducidas dimensiones, todo

ello sin perder la contigüidad de su superficie por medio de estrechas franjas y pasillos (mapa 3.8).

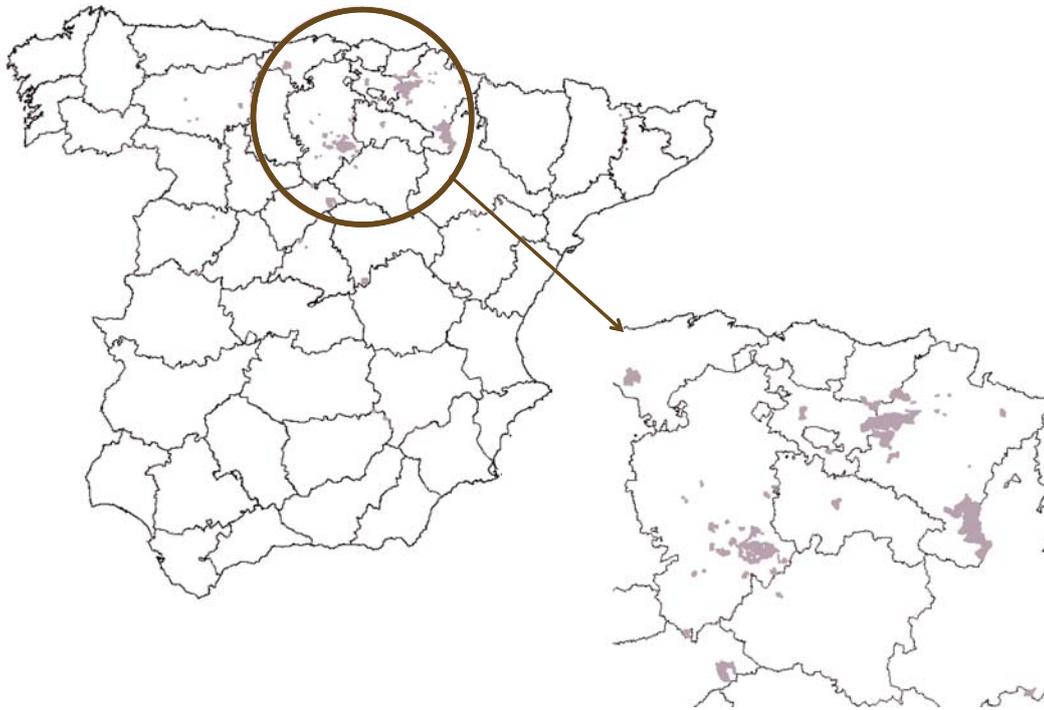
Finalmente, existen algunos municipios con una situación singular, dado que carecen de contigüidad con otros municipios de España. No se trata de Ceuta y Melilla, ambos en el continente africano, o de Formentera, en Illes Balears, y A Illa de Arousa, en la provincia de Pontevedra, ambas islas constituidas en municipios, aunque esta última esté unida a la Península por un puente, sino de dos municipios burgaleses cuyos términos municipales se sitúan totalmente dentro de condominios, Hacinas y Monasterio de la Sierra, y el municipio de Llivia, en la provincia de Girona, que se encuentra completamente rodeado por territorio francés (mapa 3.9). Por tanto, estos municipios presentan falta de continuidad topológica con el resto de municipios de la Península.

Muchas de estas peculiaridades dificultan en ocasiones el trabajo de análisis espacial en un contexto de GIS, al margen de otras singularidades relacionadas con la Administración Civil del Estado.



Mapa 3.4 Densidad de población municipal. Padrón del 2010

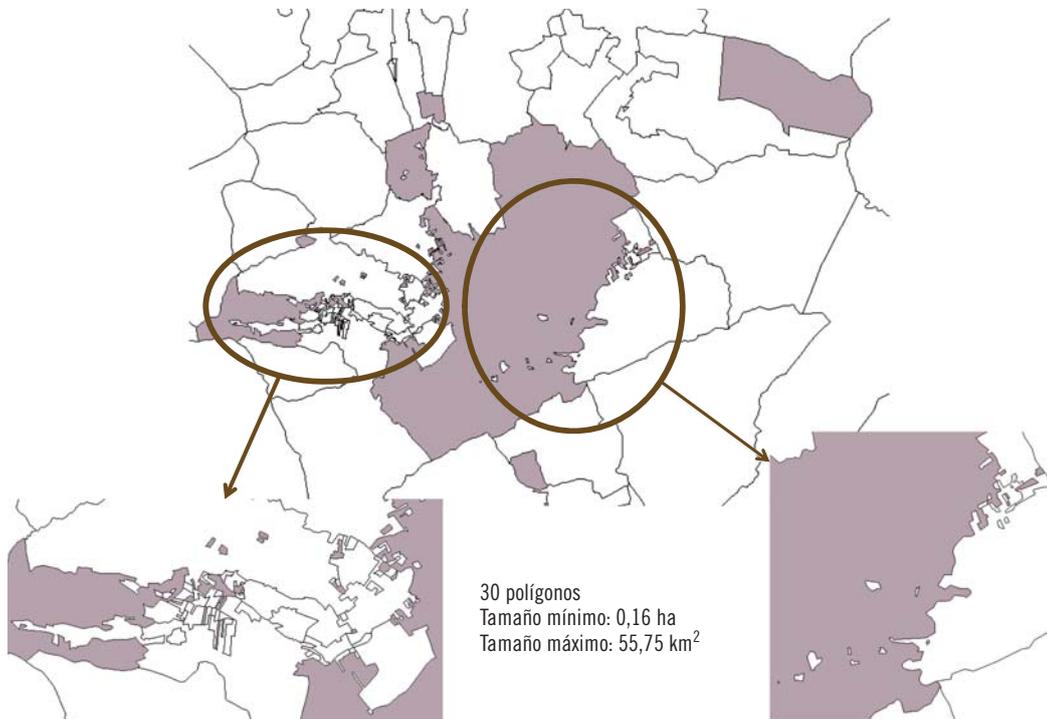
Fuente: INE (2010) y elaboración propia.



Mapa 3.5 Condominios

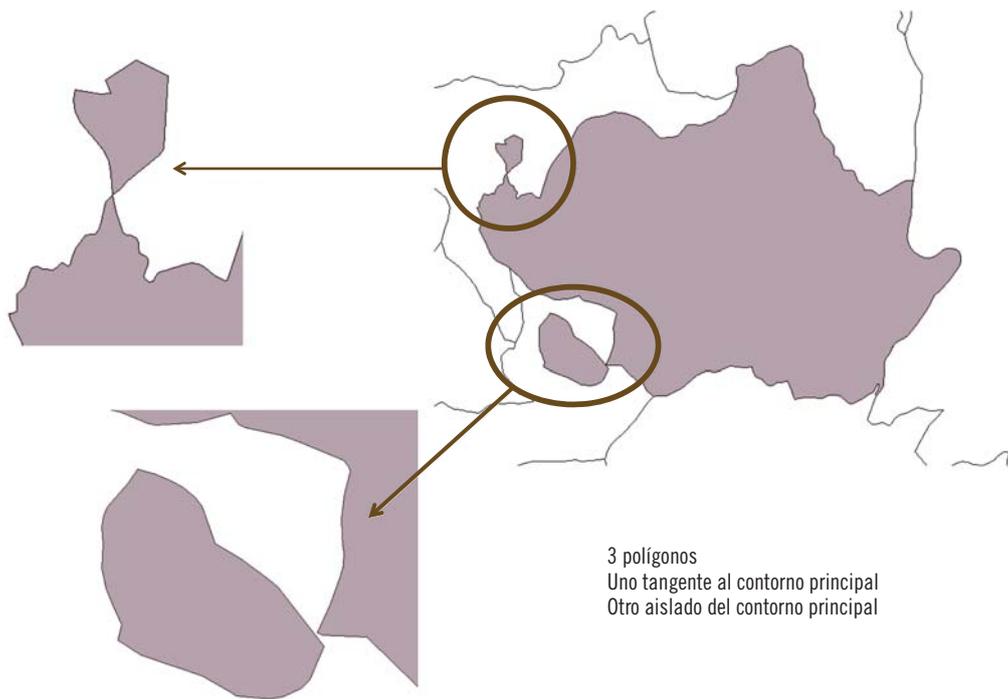
Nota: Territorios que no pertenecen a un único municipio, sino que administrativamente pertenecen a varios municipios o incluso directamente al Estado, aunque sí tienen adscripción provincial.

Fuente: Elaboración propia.



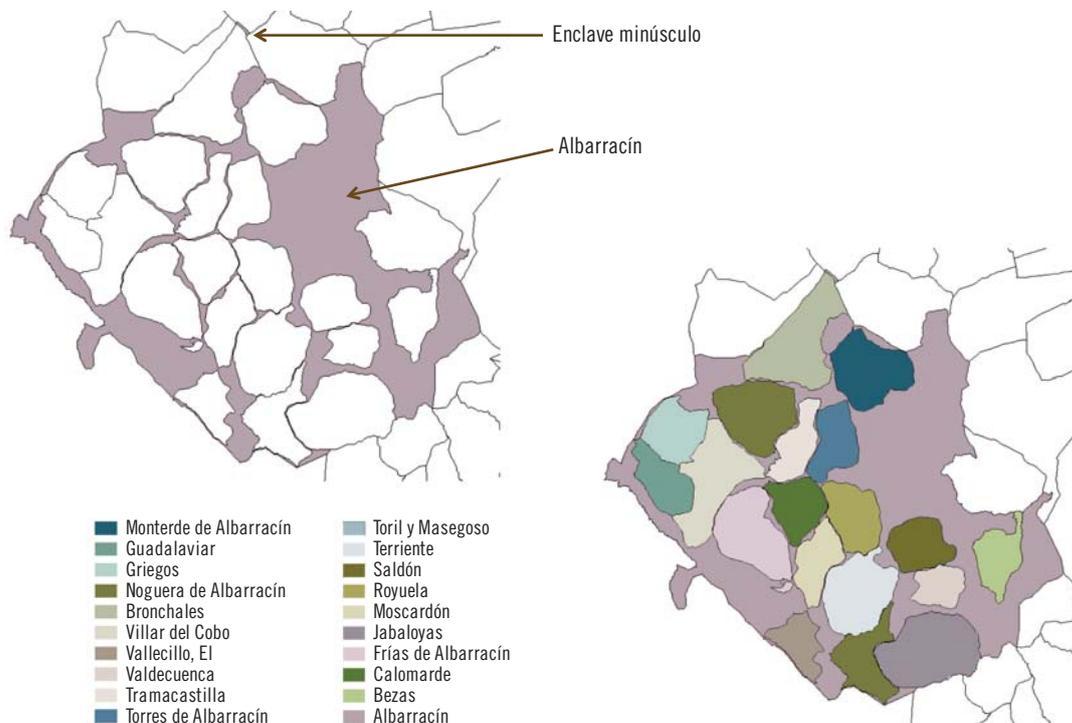
Mapa 3.6 Fragmentación de Xàtiva (46145, Valencia)

Fuente: Elaboración propia.



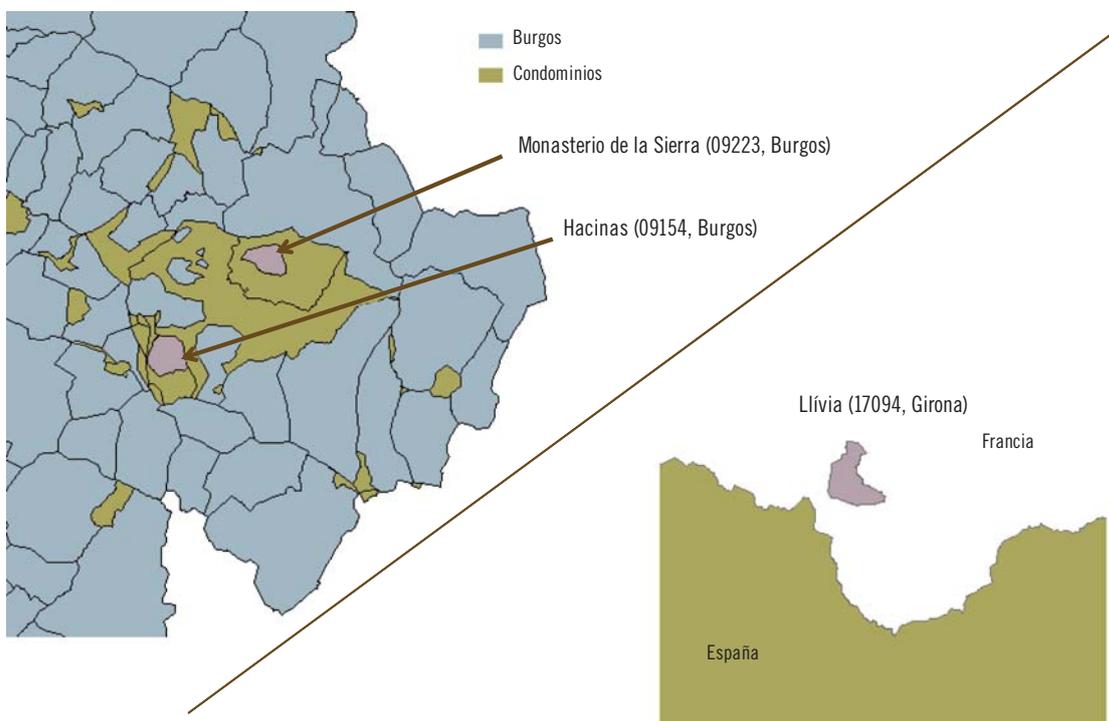
Mapa 3.7 Enclaves tangentes al contorno principal: Mohedas de Granadilla (10124, Cáceres)

Fuente: Elaboración propia.



Mapa 3.8 Polígonos *donut*: Albarracín (44009, Teruel)

Fuente: Elaboración propia.



Mapa 3.9 Municipios sin contigüidad con otros municipios españoles

Fuente: Elaboración propia.

3.2. ESTADÍSTICAS INFRAMUNICIPALES: EL NOMENCLÁTOR DE UNIDADES POBLACIONALES

Desde el punto de vista de las estadísticas demográficas el INE publica anualmente, como explotación de los ficheros padronales, el nomenclátor de unidades poblacionales del 2010, con el objetivo declarado de conocer de qué forma se asienta la población dentro de los términos municipales. Así, todos los municipios están constituidos por una o más áreas habitables, claramente diferenciadas dentro del mismo, que se las conoce bajo la denominación de *entidades singulares* de población. A su vez, las entidades singulares están consti-

tuidas por uno o varios *núcleos de población* y/o *diseminados*. Se considera núcleo de población a un conjunto de al menos diez edificaciones, que se sitúan formando calles, plazas y otras vías urbanas; si bien el número de edificaciones podrá ser inferior a 10, siempre que la población que las habita supere los 50 habitantes. Se incluyen en el núcleo aquellas edificaciones que, estando aisladas, disten menos de 200 metros de los límites exteriores del mencionado conjunto. Las edificaciones o viviendas de una entidad singular de población que no puedan ser incluidas en el núcleo se consideran en diseminado. De esta forma una entidad singular de población puede tener uno o varios núcleos, o incluso ninguno, si toda la población se encuentra diseminada²⁰. Obsérvese

²⁰ Como unidad intermedia entre la entidad singular de población y el municipio existen, en algunas provincias, agrupaciones de entidades singulares con un marcado origen histórico (parroquias, hermandades, concejos, diputaciones...) y que conforman lo que se conoce como una *entidad colectiva* de población. En el 2010 existen 4.920 entidades colectivas en 14 provincias, y su

número solo es relevante en las provincias gallegas, en Asturias y, en menor medida, en Murcia.

La presencia de entidades colectivas deshace el carácter jerárquico del nomenclátor del 2010, si bien carecen de relevancia práctica para el estudio del asentamiento de la población sobre el territorio.

que el diseminado, de difícil atribución en una aproximación que no sea *bottom-up* a la elaboración de una *grid*, se define a nivel de entidad singular, no de municipio²¹.

Así pues, frente a algo más de 8.000 municipios, el nomenclátor del 2010 contiene más de 58.000 entidades singulares y más de 71.000 núcleos y diseminados²². El cuadro 3.4 muestra la distribución provincial de estas unidades poblacionales. Se observa un gran número de ellas en Galicia, que concentra alrededor de la mitad de entidades singulares y más del 50% de los diseminados del país. La otra provincia que destaca por el gran número de unidades poblacionales es Asturias. Dado que se trata de provincias que no tienen un gran volumen de población, los tamaños medios por unidad poblacional tienden a ser muy reducidos.

El nomenclátor es, pues, una rica información sobre la distribución de la población, que puede ser analizada a dos niveles: entidades singulares o núcleos y diseminados, constituyendo estas últimas las unidades poblacionales más elementales. Además, si nos restringimos a nivel de entidad singular evitamos, inicialmente, el problema del tratamiento del diseminado.

Entre sus principales ventajas, el nomenclátor no solo ofrece una desagregación mucho mayor que la población municipal, sino que nos aporta una primera estimación de la población residente en núcleos y de aquella que reside en diseminado. El cuadro 3.5 ofrece esta distribución de la población, distinguiendo por provincias. A nivel nacional solo un 3,4% de la población, algo más

de 1,5 millones de personas, residen en diseminado; pero existe una gran heterogeneidad a nivel provincial. Destacan las provincias gallegas de Lugo y Pontevedra, que superan el 20% de su población en diseminado. Otras tres provincias superan el 8%: A Coruña, Illes Balears y Murcia, en este orden de importancia.

Esta información es, pues, el punto de partida natural para la elaboración de una *grid* de población por métodos de desagregación espacial (*top-down*), dado su elevado nivel de detalle. Sin embargo, el nomenclátor no está totalmente libre de los lindes administrativos. La relación de las unidades poblacionales del nomenclátor se realiza a partir de los municipios, de forma que un único casco urbano, como por ejemplo los núcleos principales de Elda y Petrer del mapa 2, aparecen en el nomenclátor como dos núcleos distintos, pertenecientes a dos municipios diferentes. Este tipo de situaciones son difíciles de evitar en una organización jerárquica de las estadísticas demográficas, y la única posibilidad de superación real es observar los asentamientos desde el espacio (Burchfield *et al.* 2006).

El análisis de algunas cifras del nomenclátor del 2010 arroja resultados curiosos, y muestran también una cierta influencia histórica y de práctica administrativa en la definición de unidades poblacionales. Por ejemplo, el municipio con mayor número de entidades singulares es A Estrada, en la provincia de Pontevedra, que con una población de tan solo 21.828 habitantes, contabiliza 452 entidades de población. El municipio que más núcleos registra es Vigo, también en la provincia

²¹ A nivel nacional, y a partir del nomenclátor, el 40,6% de las entidades singulares y el 33% de los municipios no tienen población en diseminado. Algunas otras cifras respecto a la estructura de núcleos y diseminados son de interés.

Un 84,3% de entidades singulares posee 1 solo núcleo (el 39,9% de los casos) o 1 solo diseminado (el 44,4% de los casos), mientras que un 13,3% de entidades singulares poseen 1 solo núcleo y 1 solo diseminado. Tan solo el 2,5% de entidades singulares posee más de un núcleo, pero representan un 17,4% de la población.

La mayor parte de la población se asienta en entidades singulares con 1 solo núcleo, el 81,1%; ya sea sin diseminado (el 38,9%) o con él (el 42,1%).

Si nuestro interés se centra en los municipios, el 53,7% de los mismos tiene 1 sola entidad singular; representando el 33,1% de la población. En lo referente a núcleos

y diseminados, el 24,2% de los municipios posee 1 solo núcleo (el 24% de los casos) o 1 solo diseminado (el 0,2% de los casos), mientras que un 27,3% de municipios poseen 1 solo núcleo y 1 solo diseminado. Sin embargo, un 44% de los municipios poseen más de un núcleo, y un 27,6% más de un diseminado. La intersección, es decir, aquellos municipios que tienen simultáneamente más de un núcleo y más de un diseminado asciende a un 23,1%, pero representan más de la mitad de la población (el 52,7%).

²² Estas cifras hacen referencia a unidades con población. El nomenclátor contiene un número de registros nulos no desdeñable, ya que incluye áreas excepcionalmente deshabitadas. En concreto en el 2010, aparecen 2.989 entidades singulares, 593 núcleos y 5.550 diseminados sin población. Todas las cifras del texto excluyen estos registros.

CUADRO 3.4: Entidades singulares, núcleos y diseminados por provincias. Nomenclátor del 2010							
Provincias	Entidades singulares		Núcleos		Diseminados		
	N.º	Porcentaje	N.º	Porcentaje	N.º	Porcentaje	
01	Álava	424	0,7	370	1,0	122	0,4
02	Albacete	307	0,5	387	1,1	185	0,5
03	Alacant /Alicante	288	0,5	673	1,8	250	0,7
04	Almería	619	1,1	551	1,5	486	1,4
05	Ávila	449	0,8	445	1,2	97	0,3
06	Badajoz	282	0,5	263	0,7	199	0,6
07	Balears, Illes	313	0,5	384	1,1	181	0,5
08	Barcelona	1.282	2,2	1.229	3,4	822	2,4
09	Burgos	1.160	2,0	1.239	3,4	131	0,4
10	Cáceres	326	0,6	318	0,9	219	0,6
11	Cádiz	210	0,4	200	0,5	179	0,5
12	Castelló/Castellón	356	0,6	272	0,7	257	0,7
13	Ciudad Real	166	0,3	167	0,5	129	0,4
14	Córdoba	340	0,6	269	0,7	269	0,8
15	Coruña, A	9.834*	16,8	4.013*	11,0	5.863	16,8
16	Cuenca	332	0,6	339	0,9	106	0,3
17	Girona	1.063	1,8	708	1,9	821	2,4
18	Granada	483	0,8	443	1,2	384	1,1
19	Guadalajara	509	0,9	495	1,4	155	0,4
20	Guipúzcoa	297	0,5	173	0,5	267	0,8
21	Huelva	215	0,4	197	0,5	134	0,4
22	Huesca	751	1,3	643	1,8	331	1,0
23	Jaén	372	0,6	289	0,8	281	0,8
24	León	1.395	2,4	1.405	3,9	230	0,7
25	Lleida	976	1,7	701	1,9	685	2,0
26	Rioja, La	249	0,4	247	0,7	73	0,2
27	Lugo	9.183	15,7	1.334	3,7	7.964*	22,9
28	Madrid	626	1,1	485	1,3	389	1,1
29	Málaga	251	0,4	362	1,0	232	0,7
30	Murcia	908	1,6	688	1,9	767	2,2
31	Navarra	878	1,5	662	1,8	432	1,2
32	Ourense	3.567	6,1	2.587	7,1	1.055	3,0
33	Asturias	6.271	10,7	3.120	8,6	3.504	10,1
34	Palencia	463	0,8	464	1,3	117	0,3

CUADRO 3.4 (cont.): Entidades singulares, núcleos y diseminados por provincias. Nomenclátor del 2010

Provincias	Entidades singulares		Núcleos		Diseminados	
	N.º	Porcentaje	N.º	Porcentaje	N.º	Porcentaje
35 Palmas, Las	515	0,9	851	2,3	303	0,9
36 Pontevedra	6.115	10,4	2.295	6,3	4.097	11,8
37 Salamanca	804	1,4	554	1,5	444	1,3
38 Santa Cruz de Tenerife	571	1,0	774	2,1	329	0,9
39 Cantabria	924	1,6	983	2,7	195	0,6
40 Segovia	385	0,7	386	1,1	88	0,3
41 Sevilla	237	0,4	340	0,9	188	0,5
42 Soria	493	0,8	496	1,4	4	0,0
43 Tarragona	468	0,8	555	1,5	273	0,8
44 Teruel	347	0,6	322	0,9	154	0,4
45 Toledo	350	0,6	432	1,2	197	0,6
46 València/Valencia	471	0,8	761	2,1	317	0,9
47 Valladolid	268	0,5	321	0,9	111	0,3
48 Vizcaya	568	1,0	345	0,9	429	1,2
49 Zamora	513	0,9	517	1,4	188	0,5
50 Zaragoza	393	0,7	421	1,2	182	0,5
51 Ceuta	3	0,0	4	0,0	3	0,0
52 Melilla	1	0,0	1	0,0		
España	58.571	100,0	36.480	100,0	34.818	100,0

* Es el valor máximo.

Nota: Las casillas vacías tienen valor 0.

Fuente: Elaboración propia a partir del nomenclátor (INE 2010).

CUADRO 3.5: Población en núcleos y diseminados por provincias. Nomenclátor del 2010

Provincias	Núcleo		Diseminado		Distribución porcentual	
	Habitantes	Porcentaje	Habitantes	Porcentaje	Núcleo	Diseminado
01 Álava	314.359	0,7	2.993	0,2	99,1	0,9
02 Albacete	395.965	0,9	5.717	0,4	98,6	1,4
03 Alacant /Alicante	1.823.837	4,0	102.448	6,4	94,7	5,3
04 Almería	641.617	1,4	53.943	3,4	92,2	7,8
05 Ávila	170.621	0,4	1.275	0,1	99,3	0,7
06 Badajoz	685.454	1,5	6.683	0,4	99,0	1,0
07 Balears, Illes	1.014.507	2,2	91.542	5,7	91,7	8,3
08 Barcelona	5.441.356	12,0	69.791	4,4	98,7	1,3
09 Burgos	373.078	0,8	1.748	0,1	99,5	0,5
10 Cáceres	408.073	0,9	7.010	0,4	98,3	1,7
11 Cádiz	1.180.433	2,6	56.306	3,5	95,4	4,6

CUADRO 3.5 (cont.): Población en núcleos y diseminados por provincias. Nomenclátor del 2010							
Provincias	Núcleo		Diseminado		Distribución porcentual		
	Habitantes	Porcentaje	Habitantes	Porcentaje	Núcleo	Diseminado	
12	Castelló/Castellón	569.517	1,3	34.757	2,2	94,2	5,8
13	Ciudad Real	525.585	1,2	3.868	0,2	99,3	0,7
14	Córdoba	779.057	1,7	26.051	1,6	96,8	3,2
15	Coruña, A	1.047.631	2,3	98.827	6,2	91,4	8,6
16	Cuenca	215.735	0,5	1.981	0,1	99,1	0,9
17	Girona	714.599	1,6	38.447	2,4	94,9	5,1
18	Granada	891.856	2,0	26.216	1,6	97,1	2,9
19	Guadalajara	246.858	0,5	4.705	0,3	98,1	1,9
20	Guipúzcoa	664.294	1,5	42.969	2,7	93,9	6,1
21	Huelva	505.713	1,1	12.368	0,8	97,6	2,4
22	Huesca	221.650	0,5	6.916	0,4	97,0	3,0
23	Jaén	656.044	1,4	14.717	0,9	97,8	2,2
24	León	496.334	1,1	2.950	0,2	99,4	0,6
25	Lleida	418.298	0,9	21.470	1,3	95,1	4,9
26	Rioja, La	320.283	0,7	2.132	0,1	99,3	0,7
27	Lugo	253.387	0,6	100.117	6,3	71,7	28,3
28	Madrid	6.446.967	14,2	11.717	0,7	99,8	0,2
29	Málaga	1.542.663	3,4	66.894	4,2	95,8	4,2
30	Murcia	1.342.460	3,0	119.519	7,5	91,8	8,2
31	Navarra	626.511	1,4	10.413	0,7	98,4	1,6
32	Ourense	324.972	0,7	10.247	0,6	96,9	3,1
33	Asturias	1.032.532	2,3	51.809	3,2	95,2	4,8
34	Palencia	168.293	0,4	4.217	0,3	97,6	2,4
35	Palmas, Las	1.061.343	2,3	29.262	1,8	97,3	2,7
36	Pontevedra	748.162	1,6	214.310	13,4	77,7	22,3
37	Salamanca	348.988	0,8	4.631	0,3	98,7	1,3
38	Santa Cruz de Tenerife	975.629	2,1	52.285	3,3	94,9	5,1
39	Cantabria	583.135	1,3	9.115	0,6	98,5	1,5
40	Segovia	163.206	0,4	1.062	0,1	99,4	0,6
41	Sevilla	1.895.757	4,2	21.340	1,3	98,9	1,1
42	Soria	95.249	0,2	9	0,0	100,0	0,0
43	Tarragona	782.806	1,7	25.614	1,6	96,8	3,2
44	Teruel	143.215	0,3	2.062	0,1	98,6	1,4
45	Toledo	690.845	1,5	7.114	0,4	99,0	1,0
46	València/Valencia	2.532.969	5,6	48.178	3,0	98,1	1,9
47	Valladolid	529.905	1,2	3.735	0,2	99,3	0,7

CUADRO 3.5 (cont.): Población en núcleos y diseminados por provincias. Nomenclátor del 2010

Provincias	Núcleo		Diseminado		Distribución porcentual	
	Habitantes	Porcentaje	Habitantes	Porcentaje	Núcleo	Diseminado
48 Vizcaya	1.103.801	2,4	49.923	3,1	95,7	4,3
49 Zamora	191.108	0,4	3.106	0,2	98,4	1,6
50 Zaragoza	963.406	2,1	9.846	0,6	99,0	1,0
51 Ceuta	78.003	0,2	2.576	0,2	96,8	3,2
52 Melilla	76.034	0,2			100,0	0,0
España	45.424.100	100,0	1.596.931	100,0	96,6	3,4

Nota: Las casillas vacías tienen valor 0.

Fuente: Elaboración propia a partir del nomenclátor (INE 2010).

de Pontevedra, con un total de 291 núcleos y 30 diseminados adscritos a otras tantas entidades singulares. Pero el municipio con más diseminados es Vilalba, en la provincia de Lugo, que con una población de 15.327 habitantes acumula 415 diseminados y 68 núcleos de población. En el otro extremo, los municipios más poblados de España, Madrid y Barcelona, tienen solamente un núcleo poblacional y ningún diseminado. Claramente esta dispersión no solo tiene que ver con la forma en la que se asienta la población sobre el territorio. Aunque el Reglamento de Población y Demarcación territorial (RD 2612/1996, de 2 de abril) establece que los Ayuntamientos revisarán, al menos una vez al año, la relación de entidades y núcleos de población, no hace explícitas las reglas geográficas que deben gobernar dicha revisión, lo que permite criterios heterogéneos y poco relacionados con la distribución de la población.

Pero el principal problema para la utilización del nomenclátor del 2010 en la elaboración de una *grid* de población procede de la ausencia de una cartografía directamente asociada al mismo. Y ello en una doble vertiente. Por una parte, el nomenclátor no está, en principio, georreferenciado, por lo que para un municipio dado nos indica cuántas unidades poblacionales tiene y cuál es la población de cada unidad, incluyendo el diseminado, pero no dónde están situadas

dentro del término municipal. Por otra parte, las entidades del nomenclátor no tienen superficie asociada a la que asignar la población y la información sobre los contornos urbanos para todos los núcleos existentes no es una información que esté directamente disponible en las estadísticas oficiales. Ambos aspectos son esenciales para la elaboración de una *grid* de población.

Estas afirmaciones deben ser matizadas, puesto que cierta información al respecto sí existe, pero es difícil relacionarla de forma inequívoca con el nomenclátor del 2010 a una fecha de referencia fija.

Buscar información para georreferenciar el nomenclátor implica escudriñar la cartografía. Un problema inicial importante con la producción cartográfica es que carece de fecha de referencia clara, típicamente su producción tiene un determinado ciclo, pero no está fechada en el sentido en el que lo está la población²³. En la práctica esto implica importantes problemas de ajuste con las fuentes demográficas cuando tratamos con grandes áreas.

El CNIG ha puesto recientemente (junio del 2011) a disposición del público en su Centro de Descargas el nomenclátor geográfico de municipios

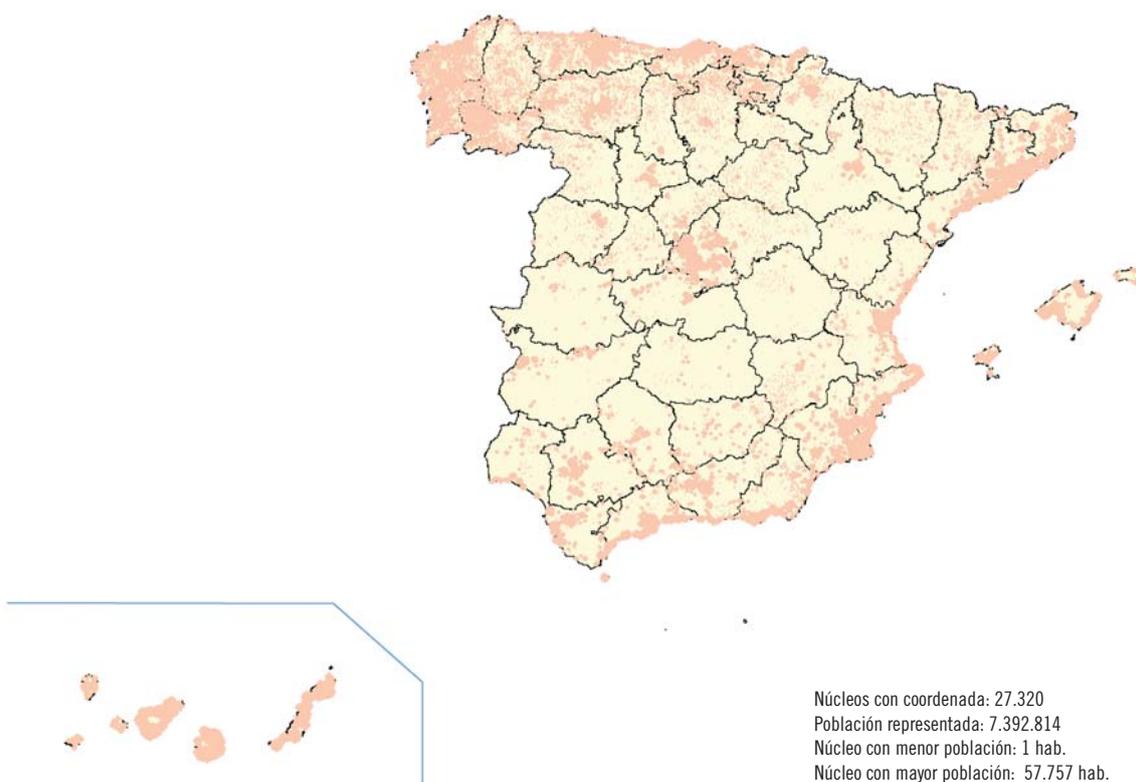
²³ Por ejemplo, la Base Cartográfica Numérica escala 1:25.000 (BCN25) nace en 1995 y tuvo un ciclo aproximado de unos diez años. Desde 2006 está siendo sustituida por la Base Topográfica Nacional escala

1:25.000 (BTN25) cuyo ciclo aún no ha acabado, por lo que actualmente todavía no está disponible para toda España.

y entidades de población, que contiene coordenadas geográficas de las entidades de población. Sin duda alguna esta información, aunque incompleta y confusa en ocasiones, es tremendamente útil²⁴. El mapa 3.10 muestra un mapa de círculos proporcionales a la población para los núcleos del nomenclátor del 2010 de los que se dispone de coordenadas en dicha base de datos. Este simple ejercicio es de escasa calidad por varias razones, de los 36.480 núcleos del nomenclátor del 2010 solo disponemos de coordenadas en la base de datos del IGN para 27.320; pero lo más sorprendente es que la población representada en el mapa 3.10 es de

solo 7.392.814 residentes. La razón es que solo los núcleos más pequeños tienen coordenada asignada; el núcleo con mayor población en el mapa 3.10 tiene 57.757 habitantes. Es cierto que muchas de estas coordenadas pueden recuperarse, porque en realidad han sido asignadas a la correspondiente entidad singular, y el código para el núcleo eliminado de la base de datos, pero esto obliga a costosas tareas de edición por parte del usuario²⁵.

Una coordenada supone disponer de un punto en el que concentrar la población, pero no es suficiente para la construcción de la *grid*.



Mapa 3.10 Núcleos de población para los que se dispone de coordenada. Nomenclátor del 2010

Fuente: INE (2010) y elaboración propia.

²⁴ La base de datos de junio del 2011 contiene 66.386 coordenadas diferentes. Se incluyen coordenadas para los diseminados, aunque la mayoría de estas coinciden con las de un núcleo o las de una entidad singular. De acuerdo con el resumen descriptivo de la base de datos, «las coordenadas asignadas a todas las unidades poblacionales [...] han sido dadas a centroides [...]» (página 10 del documento descriptivo de la base de datos), lo que implica calcular polígonos. Nada se dice de la

disponibilidad de dichos polígonos o sobre cómo han sido construidos, lo que no es obvio, ni en el caso de los diseminados, ni en el caso de entidades singulares con más de un núcleo de población.

²⁵ Existen además algunos errores de localización evidentes, como la localización de un núcleo del municipio de Sueca, en la provincia de Valencia, unos 20 kilómetros mar adentro.

Necesitamos además contornos urbanos, polígonos, sobre los que asignar la población.

En principio estos contornos urbanos podrían obtenerse de la cartografía. La Base Topográfica Nacional de España escala 1:25.000 (BTN25) (IGN 2010b) incluye una capa de *Entidades de población* que hace referencia a estos contornos, definidos como «[...] elementos superficiales que representan los límites imaginarios de la población [...]», o «[...] en el caso de edificaciones aisladas... un perímetro que englobe el conjunto de edificaciones [...], siempre y cuando haya dos o más edificaciones» (IGN 2010b, 65). El mapa 3.11 muestra un ejemplo de la captura de estas entidades.

Sin embargo, el manejo directo de dicha información para grandes áreas es prácticamente inviable, y ello por varias razones. En primer lugar, la unidad de distribución de dicha base de

datos es la hoja del Mapa Topográfico Nacional 1:25.000 (MTN25) lo que implica manejarse con más de 4.000 hojas para toda España, que sería necesario empalmar para obtener una cartografía digital continua. En segundo lugar, el campo correspondiente al código INE de la entidad en dicha base de datos está vacío, con lo que una asignación directa entre el nomenclátor del 2010 y los polígonos de dicha capa es inviable. En tercer lugar, de forma similar a lo que ocurre con el nomenclátor geográfico de municipios y entidades de población, no todas las entidades tienen un contorno definido, con una clara preferencia sobre las entidades de pequeñas dimensiones, y sin que parezca existir un criterio definido.

El mapa 3.12 muestra algunos municipios del área metropolitana de Valencia en los que se sobrepone la capa de *Entidades de población de la base cartográfica numérica* escala 1:25.000 (BCN25)²⁶

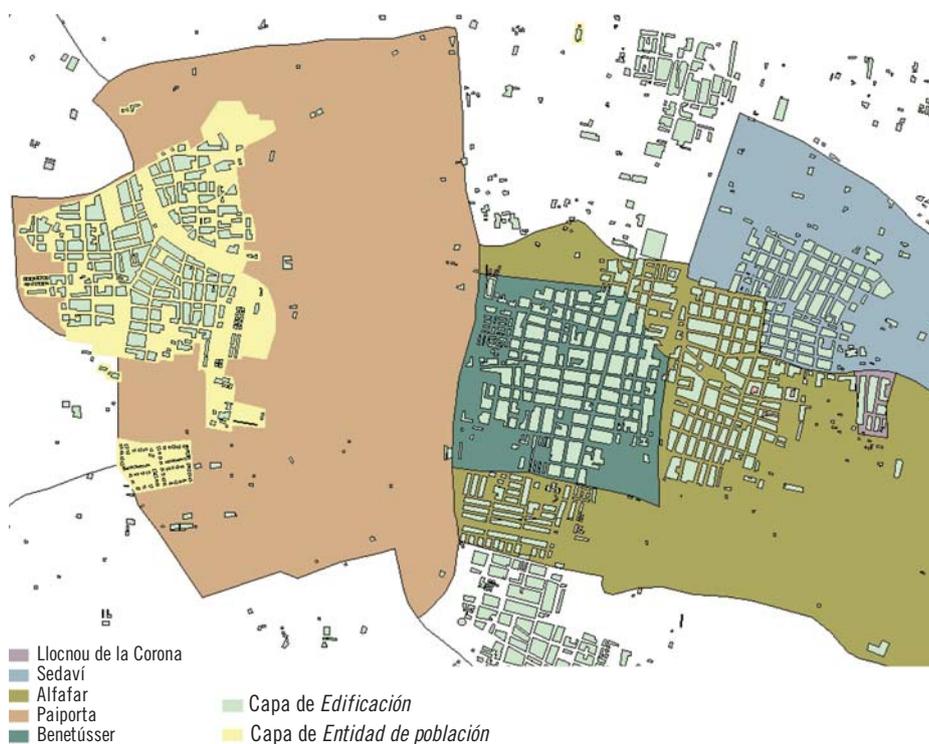


Mapa 3.11 Capa de *Entidades de población* de la BTN25. Método de captura

Fuente: IGN (2010b).

²⁶ La BCN25 es la versión antigua de la BTN25 y actualmente en proceso de sustitución. Sin embargo, no toda

España está disponible en la BTN25.



Mapa 3.12 Capa de *Entidades de población* de la BTN25. Método de captura

Fuente: Elaboración propia.

junto con la capa de *Edificación* de la misma base de datos²⁷. Solo para el municipio de Paiporta se recoge adecuadamente el contorno de su casco urbano, mientras que para el resto de municipios únicamente se observan las edificaciones. A partir de ellas podrían generarse contornos urbanos mediante criterios de densidad de edificios (Goerlich y Cantarino 2011), aunque esta sería una tarea ardua y tediosa para el conjunto nacional²⁸.

Alternativamente, es posible encontrar contornos urbanos en la capa de *Núcleos de población* de

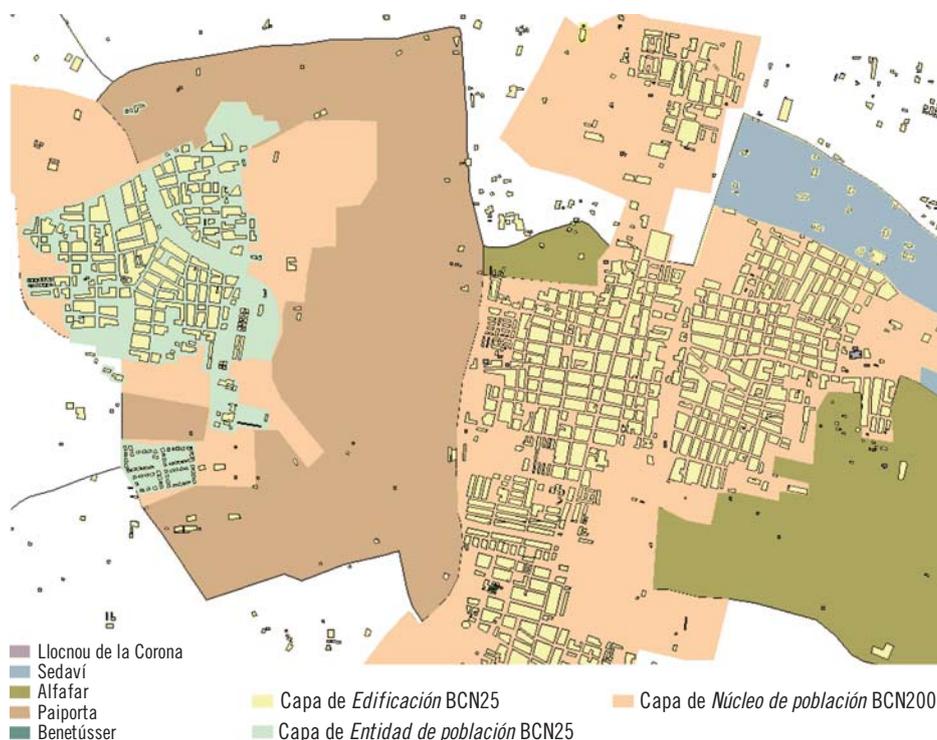
la *base cartográfica nacional* a escala 1:200.000 (BCN200) (IGN 2011a). Dicha base de datos se ofrece de forma continua para toda España, e incluye el código INE correspondiente. El mapa 3.13 añade, a la anterior, la capa de *Núcleos de población* de la BCN200, observamos que esta sí engloba los núcleos de todos los municipios, aunque la escala es algo reducida para la elaboración de una *grid* de 1 km² de resolución que es nuestro objetivo, de hecho los contornos de la capa de núcleos de la BCN200 tienden a mostrar unos límites demasiado holgados respecto a lo

²⁷ Definida como: «Construcción fija, hecha con materiales resistentes, para habitación humana o para otros usos». (IGN 2010a, 69).

²⁸ Obsérvese que los criterios mediante los cuales el IGN genera contornos urbanos no son explícitos. Sin duda alguna la idea de densidad de las edificaciones juega un papel importante, aunque una inspección visual de la cartografía no permite establecer patrones claros. Ciertamente, la densidad de edificios de una determinada área no es un indicador totalmente exacto de la densidad de población de dicha área, no todos los edificios son de uso residencial, ni todos aquellos que lo son albergan población residente. Sin embargo, es razonable

considerar que la correlación entre ambas variables será elevada y que aquellas zonas sin edificios, o con muy baja densidad de estos, pueden ser consideradas directamente como áreas no habitadas o con una probabilidad muy reducida de serlo, sin embargo, la BTN25 recoge en ocasiones polígonos para entidades de población con muy poca densidad de edificios, lo que por otra parte está en consonancia con las entidades del nomenclátor del 2010.

Desde un punto de vista más general sería conveniente establecer criterios claros y específicos sobre los que construir contornos urbanos en un contexto GIS de forma automática (Chaudhry y Mackaness 2008).



Mapa 3.13 Capa de *Entidades de población* de la BCN25 y *Núcleos de población* de la BCN200

Fuente: Elaboración propia.

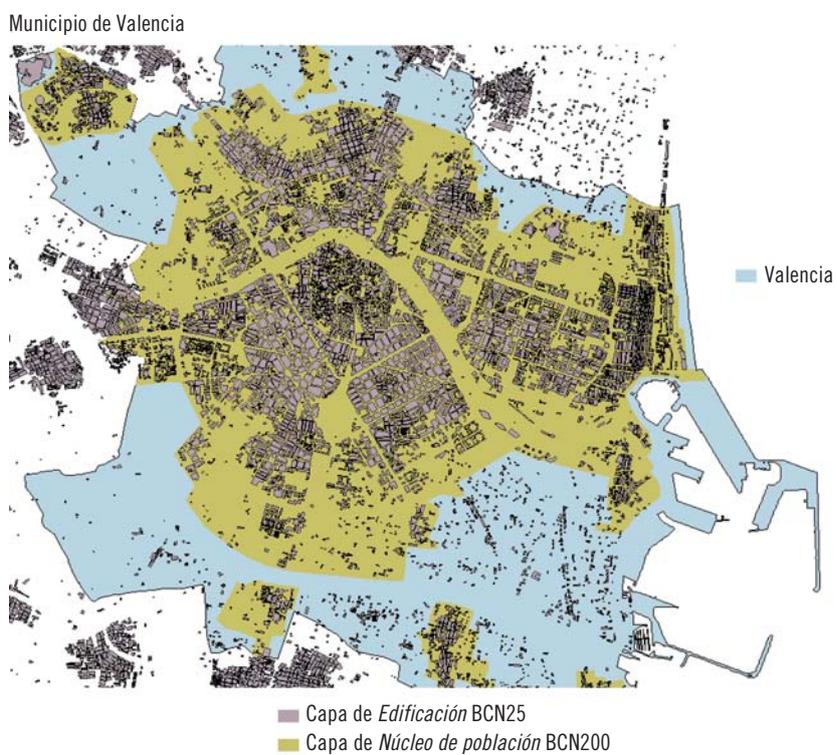
que parecen ser realmente los contornos urbanos. El mapa 3.14 confirma este hecho para el núcleo principal del municipio de Valencia.

El mapa 3.15 muestra la capa de *Núcleos de población* de la BCN200 para toda España; esta sería la geografía de los asentamientos en un método dasimétrico binario (Langford y Unwin 1994; Langford 2007) que localizara la población sobre estos polígonos, como el utilizado por Goerlich y Cantarino (2011) para la Comunitat Valenciana. El capítulo 6 ofrece un mapa similar de elaboración propia, que se ajusta mejor a nuestras necesidades.

Sin embargo, esta base de datos está sujeta a problemas similares a los mencionados para el nomenclátor geográfico de municipios y entidades de población. La capa de *Núcleos de población* de la BCN200 incluye 29.704 códigos INE diferentes, bastantes menos que los núcleos del nomenclátor del 2010. Muchos de estos registros corresponden a núcleos, pero otros a entidades singulares o colectivas, e incluso en 12 casos a

diseminados. La delimitación de contornos urbanos es problemática en estas situaciones, salvo en condiciones particulares como por ejemplo entidades con un solo núcleo de población.

Puede resultar de interés que la BCN200 ofrezca una capa de puntos de diseminados. Aunque solo contiene 19.610 códigos INE diferentes, dispone de más de 30.000 registros, porque en muchos casos para un mismo diseminado se ofrecen varios puntos. Esta capa representa una aproximación cartográfica a la distribución espacial del diseminado que, sin embargo, es de poca utilidad a efectos de elaboración de una *grid* de población, ya que con la información del nomenclátor del 2010 no es posible asignar la población de un diseminado a varios puntos. La representación visual de dicha capa se ofrece en el mapa 3.16; téngase en cuenta que los puntos no indican la importancia demográfica y que se han representado todos en el mismo tamaño. El mapa de asentamientos urbanos potenciales que ofreceremos en el capítulo 6 incluye, por construcción, la población en diseminado.



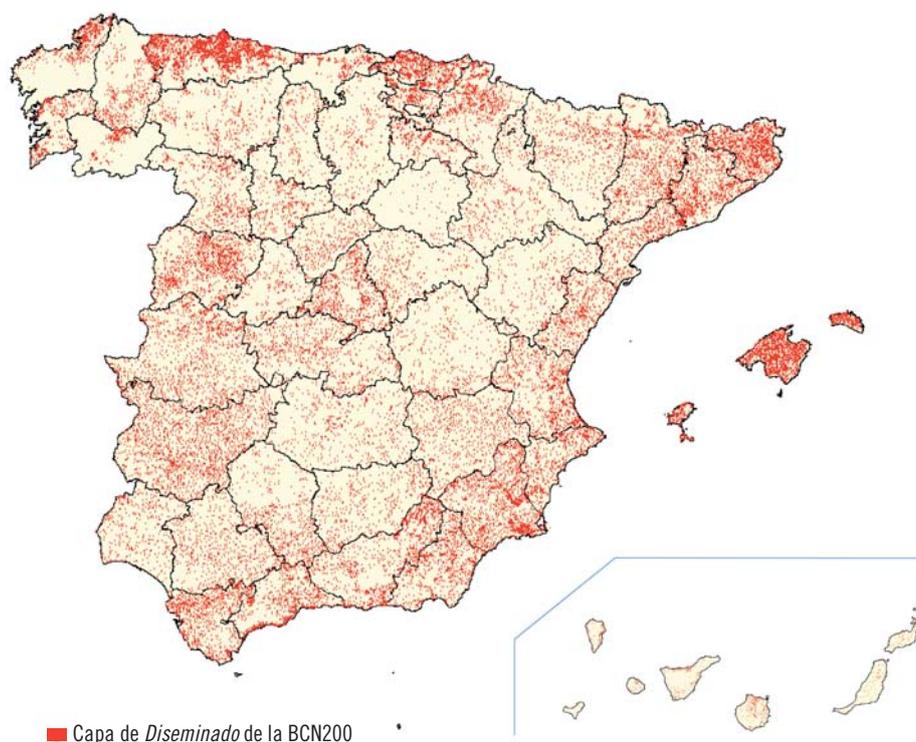
Mapa 3.14 Capa de Edificación de la BTN25 y de Núcleos de población de la BCN200

Fuente: Elaboración propia.



Mapa 3.15 Capa de Núcleos de población de la BCN200

Fuente: Elaboración propia.



Mapa 3.16 Capa de *Diseminados* de la BCN200

Fuente: Elaboración propia.

Las reflexiones de este apartado en lo referente a las posibilidades de asociar una cartografía al nomenclátor, con el fin último de elaborar una *grid*, tienen todas ellas una clara conclusión final. Si bien es cierto que el nomenclátor constituye una fuente de información muy detallada para este objetivo, las posibilidades de georreferenciación del mismo, con la información pública actualmente disponible, son prácticamente nulas. Aunque se han hecho enormes progresos en los últimos años, la geoestadística oficial tiene aquí una importante tarea pendiente, y de la que el INE ya ha tomado el testigo (Maldonado 2011).

3.3. ESTADÍSTICAS INFRAMUNICIPALES: SECCIONES CENSALES

La otra fuente de información de población inframunicipal es la constituida por las seccio-

nes censales, unas 35.000 en todo el territorio nacional²⁹. Al contrario que las unidades poblacionales del nomenclátor del 2010, las secciones censales son una subdivisión del término municipal, y en consecuencia tienen territorio asignado, lo que facilita su manipulación mediante GIS a efectos del estudio de la distribución espacial de la población. Sin embargo, las secciones censales no han sido diseñadas para este fin, sino como un instrumento auxiliar en la gestión del padrón, en los trabajos censales y en el diseño del muestreo de las encuestas del INE.

La Orden de 11 de julio de 1997 sobre comunicaciones electrónicas entre las Administraciones Públicas referentes a la información de los padrones municipales establece que el número de electores por sección debe estar comprendido entre 500 y 2.000, por lo que el tamaño poblacional de una sección censal

²⁹ Como unidad intermedia entre el municipio y las seccio-

nes censales están los distritos, que no consideraremos.

no debería superar los 2.500 habitantes de derecho. En la práctica hay numerosos casos que exceden estos límites. El cuadro 3.6 ofrece información sobre el número de secciones por provincia y estadísticos descriptivos sobre los tamaños poblacionales por sección. A este nivel de agregación la relación entre número de secciones y población es evidente, con un coeficiente de correlación de 0,993, mientras que el coeficiente de correlación entre el número de municipios y la población provincial es solamente de 0,079.

Así pues, aunque en términos de población las secciones censales son relativamente homogéneas (por construcción), se trata de una subdivisión territorial de los términos municipales; por lo que los municipios más pequeños están constituidos por una sola sección censal. A nivel nacional el 71,6% de los municipios solo tienen una sección censal, aunque en términos de población representan solo el 6,4%. Por tanto, desde el punto de vista de las áreas rurales no representan una alternativa clara frente a los datos municipales para la construcción de una *grid*

CUADRO 3.6: Secciones censales (SSCC) y estadísticos de tamaño poblacional por provincias (2010)

Provincias	Secciones censales		Promedio de SSCC por municipio	Estadísticos descriptivos de tamaño poblacional			
	N.º	Porcentaje		Mínimo	Mediana	Media	Máximo
01 Álava	257	0,7	5,0	175	1.199	1.235	3.746
02 Albacete	301	0,8	3,5	58	1.316	1.334	2.566
03 Alacant /Alicante	1.225	3,4	8,7	38	1.353	1.572	15.941
04 Almería	416	1,2	4,1	84	1.622	1.672	5.762
05 Ávila	308	0,9	1,2	18	233	558	2.734
06 Badajoz	551	1,5	3,4	93	1.187	1.256	2.697
07 Balears, Illes	620	1,7	9,3	276	1.710	1.784	4.180
08 Barcelona	3.563	10,0	11,5	24	1.472	1.547	8.890
09 Burgos	577	1,6	1,6	6	373	650	2.685
10 Cáceres	417	1,2	1,9	12	962	995	2.805
11 Cádiz	887	2,5	20,2	479	1.317	1.394	3.717
12 Castelló/Castellón	437	1,2	3,2	16	1.373	1.383	3.964
13 Ciudad Real	425	1,2	4,2	103	1.201	1.246	2.945
14 Córdoba	618	1,7	8,2	374	1.252	1.303	2.703
15 Coruña, A	917	2,6	9,8	534	1.203	1.250	2.812
16 Cuenca	308	0,9	1,3	16	379	707	2.847
17 Girona	525	1,5	2,4	89	1.409	1.434	5.049
18 Granada	649	1,8	3,9	146	1.344	1.415	2.959
19 Guadalajara	387	1,1	1,3	12	137	650	3.103
20 Guipúzcoa	545	1,5	6,2	120	1.242	1.298	2.820
21 Huelva	359	1,0	4,5	50	1.370	1.443	4.189
22 Huesca	289	0,8	1,4	24	540	791	2.847
23 Jaén	515	1,4	5,3	351	1.248	1.302	2.960
24 León	459	1,3	2,2	56	1.042	1.088	2.609
25 Lleida	390	1,1	1,7	57	1.085	1.128	3.128
26 Rioja, La	336	0,9	1,9	9	1.026	960	2.975
27 Lugo	302	0,8	4,5	207	1.076	1.171	2.605
28 Madrid	4.224	11,9	23,6	49	1.431	1.529	5.539

CUADRO 3.6 (cont.): Secciones censales (SSCC) y estadísticos de tamaño poblacional por provincias (2010)								
Provincias	Secciones censales		Promedio de SSCC por municipio	Estadísticos descriptivos de tamaño poblacional				
	N.º	Porcentaje		Mínimo	Mediana	Media	Máximo	
29	Málaga	1.006	2,8	10,1	127	1.421	1.600	8.058
30	Murcia	1.081	3,0	24,0	575	1.246	1.352	8.680
31	Navarra	552	1,5	2,0	17	1.171	1.154	5.407
32	Ourense	283	0,8	3,1	447	1.090	1.185	2.611
33	Asturias	875	2,5	11,2	170	1.190	1.239	2.669
34	Palencia	283	0,8	1,5	11	332	610	2.620
35	Palmas, Las	674	1,9	19,8	620	1.505	1.618	6.916
36	Pontevedra	743	2,1	12,0	556	1.238	1.295	2.648
37	Salamanca	527	1,5	1,5	18	370	671	2.413
38	Santa Cruz de Tenerife	641	1,8	11,9	578	1.452	1.604	6.484
39	Cantabria	460	1,3	4,5	72	1.248	1.288	2.700
40	Segovia	271	0,8	1,3	24	297	606	2.752
41	Sevilla	1.292	3,6	12,3	350	1.421	1.484	3.344
42	Soria	215	0,6	1,2	9	103	443	2.506
43	Tarragona	530	1,5	2,9	42	1.458	1.525	5.193
44	Teruel	284	0,8	1,2	8	211	512	2.452
45	Toledo	519	1,5	2,5	5	1.359	1.345	2.842
46	València/Valencia	1.799	5,0	6,8	36	1.385	1.435	3.401
47	Valladolid	571	1,6	2,5	28	941	935	2.453
48	Vizcaya	929	2,6	8,3	112	1.197	1.242	2.715
49	Zamora	320	0,9	1,3	40	400	607	2.646
50	Zaragoza	876	2,5	3,0	12	1.149	1.111	3.133
51	Ceuta	53	0,1	53,0	708	1.430	1.520	3.462
52	Melilla	38	0,1	38,0	644	1.869	2.001	4.182
	España	35.629	100,0	4,4	5	1.282	1.320	15.941

Fuente: Elaboración propia a partir del INE (2010).

de población. El cuadro 3.7 muestra la distribución provincial de municipios según el número de secciones censales.

Las secciones censales pueden ser de mucha utilidad, sin embargo, para ajustar la precisión de la *grid* en las grandes ciudades y núcleos urbanos de tamaño medio, ya que en estos casos ofrecen información sobre la localización de la población en áreas muy pequeñas. Un 60% de secciones tienen una superficie inferior al km², todas ellas en grandes ciudades o áreas metropolitanas, donde es posible encontrar secciones con superficie inferior a media hectárea. El mapa 3.17 muestra

la ganancia en resolución al pasar del municipio al nivel de sección censal para la ciudad de Valencia.

El principal inconveniente es que, dado que deben cubrir todo el territorio municipal, su forma tiene poco que ver con la distribución de la población en muchos casos. De hecho, cuando hay más de una sección censal en un municipio siempre hay al menos una sección cuya densidad no es comparable a la del resto, ya que actúa de forma residual. Un ejemplo sencillo con un municipio con dos secciones censales servirá para ilustrar lo que queremos decir.

CUADRO 3.7: Municipios según número de secciones censales por provincias y población que albergan a nivel nacional (2010)

		Municipios según número de secciones censales						
Provincias		1	2 a 5	6 a 20	21 a 100	101 a 500	> 500	España
01	Álava	44	4	2		1		51
02	Albacete	58	23	4	1	1		87
03	Alacant/Alicante	76	22	30	11	2		141
04	Almería	61	30	8	2	1		102
05	Ávila	233	14		1			248
06	Badajoz	86	64	11	2	1		164
07	Illes Balears	14	33	16	3	1		67
08	Barcelona	147	86	50	23	4	1	311
09	Burgos	357	11	1	1	1		371
10	Cáceres	178	34	7	2			221
11	Cádiz	7	12	15	8	2		44
12	Castellón/Castelló	102	21	8	3	1		135
13	Ciudad Real	54	29	15	4			102
14	Córdoba	18	33	21	2	1		75
15	Coruña, A	6	47	34	6	1		94
16	Cuenca	217	19	1	1			238
17	Girona	167	39	11	4			221
18	Granada	96	51	19	1	1		168
19	Guadalajara	272	14	1	1			288
20	Guipuzkoa	46	18	20	3	1		88
21	Huelva	36	29	13		1		79
22	Huesca	188	8	5	1			202
23	Jaén	23	54	16	4			97
24	León	178	25	6	1	1		211
25	Lleida	199	27	4	1			231
26	Rioja, La	154	14	5		1		174
27	Lugo	13	44	9	1			67
28	Madrid	88	48	18	18	6	1	179
29	Málaga	47	33	9	10	1		100
30	Murcia	6	6	23	8	2		45
31	Navarra	218	45	7	1	1		272
32	Ourense	40	46	5	1			92
33	Asturias	27	26	19	4	2		78
34	Palencia	176	12	2	1			191
35	Palmas, Las	2	7	19	5	1		34
36	Pontevedra	1	29	28	3	1		62
37	Salamanca	348	10	3		1		362
38	Santa Cruz de Tenerife	4	29	16	3	2		54
39	Cantabria	55	34	9	3	1		102
40	Segovia	198	9	1	1			209
41	Sevilla	10	54	33	7		1	105

CUADRO 3.7 (cont.): Municipios según número de secciones censales por provincias y población que albergan a nivel nacional (2010)

Provincias	Municipios según número de secciones censales						España
	1	2 a 5	6 a 20	21 a 100	101 a 500	> 500	
42 Soria	176	6		1			183
43 Tarragona	137	32	11	3			183
44 Teruel	225	8	2	1			236
45 Toledo	115	75	12	2			204
46 València/Valencia	149	61	42	13		1	266
47 Valladolid	197	22	5		1		225
48 Vizcaya	61	25	17	8	1		112
49 Zamora	241	4	2	1			248
50 Zaragoza	256	31	5		1		293
51 Ceuta				1			1
52 Melilla				1			1
España	5.807	1.457	620	183	43	4	8.114
Porcentaje	71,57	17,96	7,64	2,26	0,53	0,05	100,00
Población	2.997.822	6.258.003	9.983.328	10.837.393	10.538.634	6.405.851	47.021.031
Porcentaje	6,38	13,31	21,23	23,05	22,41	13,62	100,00

Nota: Las casillas vacías tienen valor 0.

Fuente: Elaboración propia a partir del INE (2010).



Mapa 3.17 Secciones censales frente a municipios en las grandes ciudades. Municipio de Valencia (46250). 596 secciones censales

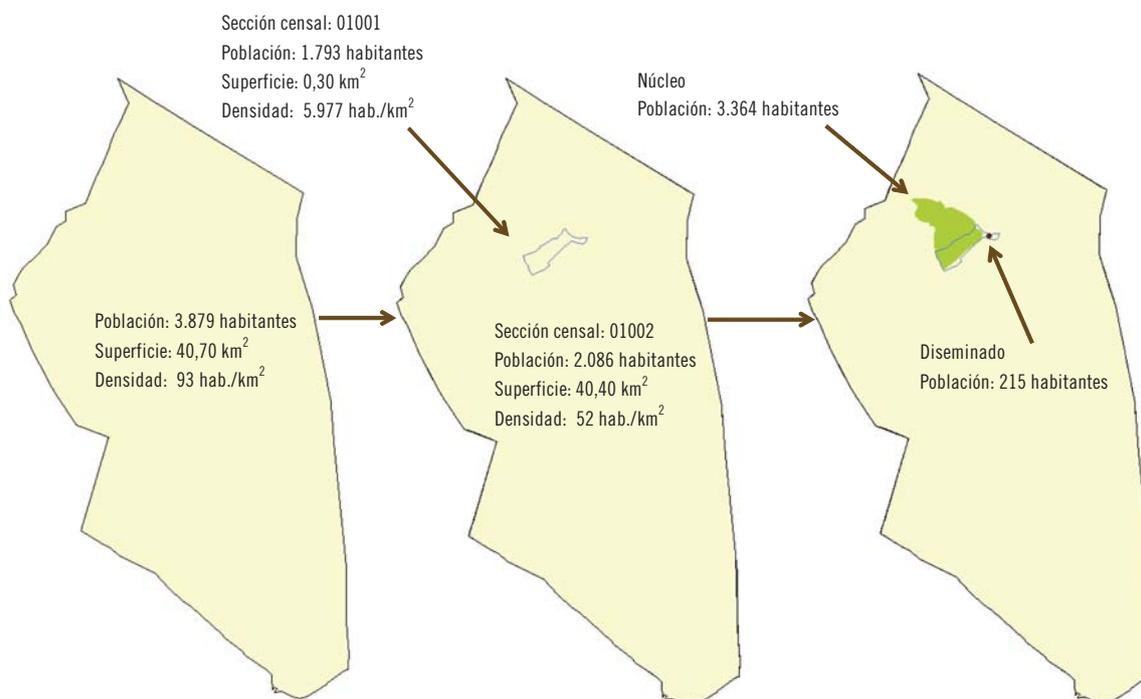
Fuente: INE (2010) y elaboración propia.

El municipio de Villar del Arzobispo, en la provincia de Valencia, tiene 3.879 habitantes y una superficie de 40,7 km², lo que arroja una densidad de 93 habitantes por km², justo en el promedio nacional. Su término municipal se divide en dos secciones censales con poblaciones similares, la primera con 1.793 habitantes y la segunda con 2.086. Sin embargo, este homogéneo reparto de la población no va acompañado de una división similar en cuanto a la superficie. La primera sección, que engloba parte del núcleo de población, tiene una superficie de 30 ha, mientras que la segunda aglutina la otra parte del núcleo y el resto del término municipal, con lo que su superficie es la práctica totalidad de dicho término, 40,40 km². Como consecuencia, el cálculo de densidades arroja valores no comparables y carentes de sentido, 5.977 habitantes por km² para la sección que engloba parte del núcleo, muy superior a la propia población del municipio a pesar de tener una superficie superior a los 40 km², y 52 habitantes por km² para la segunda sección censal. Sin información auxiliar sobre la ubicación de la población de esta sección censal no es posible

mejorar en estos casos. El mapa 3.18 muestra gráficamente este ejemplo. En el último mapa imponemos el contorno del núcleo urbano de la capa de *Núcleos de población* de la BCN200, así como la coordenada a la que dicha base cartográfica asigna el diseminado, que en este caso resulta estar dentro de la sección censal más pequeña.

El mensaje de este apartado es claro. Las secciones censales no han sido diseñadas para facilitar el estudio de los asentamientos humanos, y no representan una mejora sustancial respecto a la información proporcionada por los datos municipales en áreas rurales, sin embargo representan una mejora potencial muy importante en áreas urbanas. El 93,6% de la población reside en municipios con más de una sección censal, y el 80,3% lo hace en municipios con más de cinco.

A pesar de ello tanto los datos de población por secciones censales, como los de los municipios, requieren de información auxiliar sobre contornos urbanos que sirva de base en la



Mapa 3.18 Densidades de población a partir de secciones censales. Villar del Arzobispo (46258)

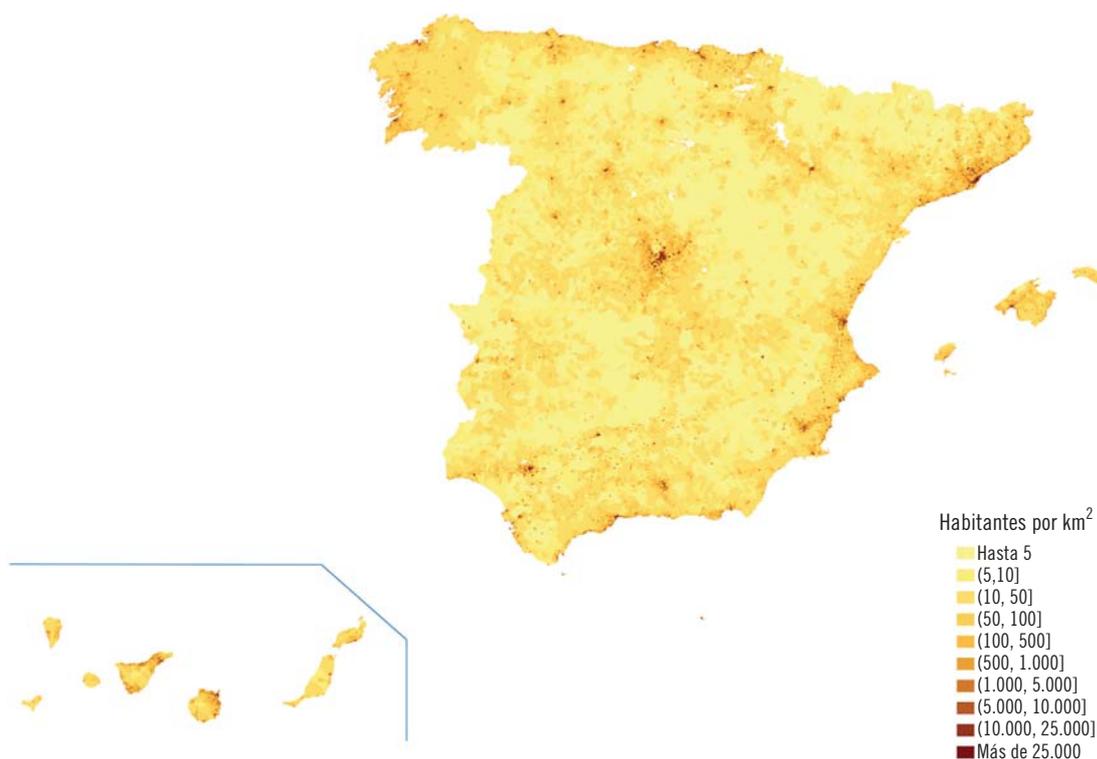
Fuente: INE (2010) y elaboración propia.

elaboración de una *grid* de población. Hemos expuesto con detalle en el apartado anterior las razones por las cuales estos no pueden provenir de las bases cartográficas actualmente disponibles, por lo que deberemos dirigir la atención a la información sobre coberturas del suelo. Los dos capítulos que siguen examinan este tipo de información en relación con el problema que nos ocupa.

Al igual que hicimos con las densidades municipales, el mapa 3.19 muestra la densidad de población calculada a partir de los datos de secciones censales, con los mismos cortes que el mapa 3.4. Hemos añadido, sin embargo, un intervalo más al final del histograma, ya que si la mayor densidad de población a nivel municipal supera ligeramente los 20.000 habitantes, a

nivel de sección censal se superan los 300.000 habitantes por km², y en algo más de 100 casos los 100.000 habitantes³⁰ por km².

La comparación de ambos mapas, 3.4 y 3.19, muestra un hecho conocido, y que ya se apreciaba en el mapa 3. Al reducir la dimensión de las unidades de análisis aumenta la concentración espacial de la población. El mapa 3.19 es notablemente más claro que el mapa 3.4, sobre todo en áreas centrales de la Península, al tiempo que las áreas oscuras se reducen en superficie y se intensifican en color. Aunque no se aprecia en el tamaño de la imagen, un aumento de la escala del mapa permite distinguir multitud de pequeños puntos oscuros rodeados de amplias zonas mucho más claras, es el efecto que muestra el mapa 3.18, se trata de secciones censales de núcleos



Mapa 3.19 Densidad de población por sección censal. Padrón del 2010

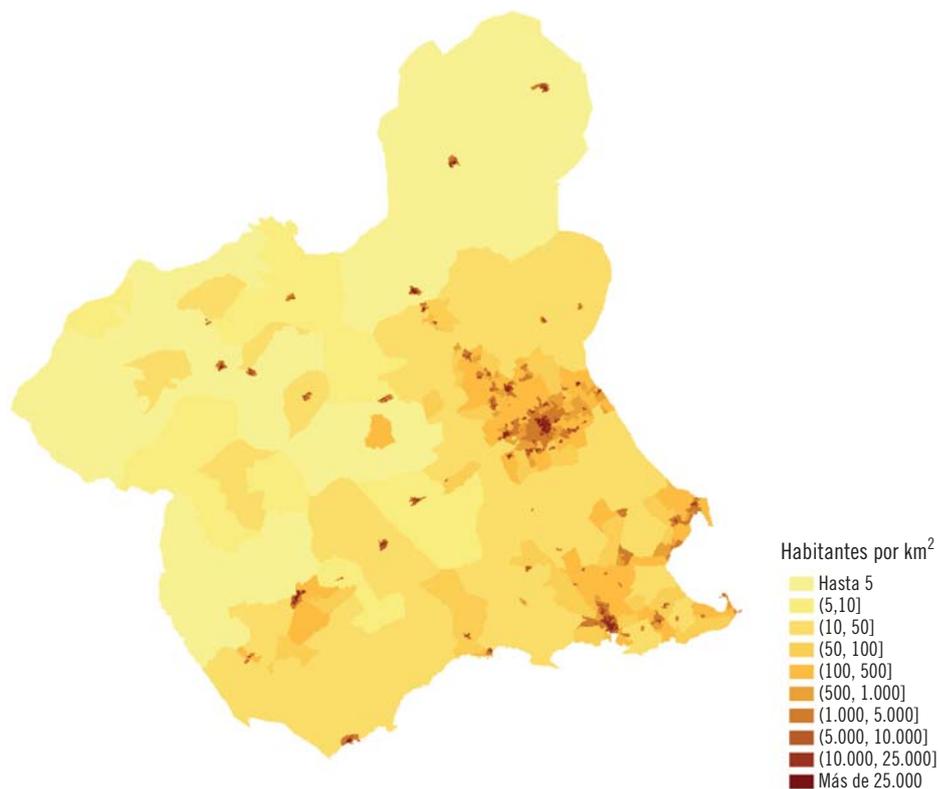
Fuente: INE (2010) y elaboración propia.

³⁰ Estas densidades se obtienen en secciones de superficie mínima, en muchos casos inferior a la hectárea. Ya hemos llamado la atención sobre lo absurdo de ciertos cálculos de densidades cuando la unidad de referencia (km²) es

mayor que la superficie sobre la que se calcula la densidad, puesto que en este caso no estamos distribuyendo población entre superficie, sino aumentando la escala hasta la unidad de referencia.

urbanos puros y reducido tamaño, rodeadas de otras mucho más grandes que juegan el papel de *completar* el término municipal. El mapa 3.20 muestra este efecto para la provincia de Murcia.

Como tendremos ocasión de comprobar en el capítulo 6 el mapa 3.19 todavía está lejos de ser una aproximación realista a la distribución espacial de la población en España.



Mapa 3.20 Densidad de población por sección censal. Padrón del 2010. Región de Murcia

Fuente: INE (2010) y elaboración propia.

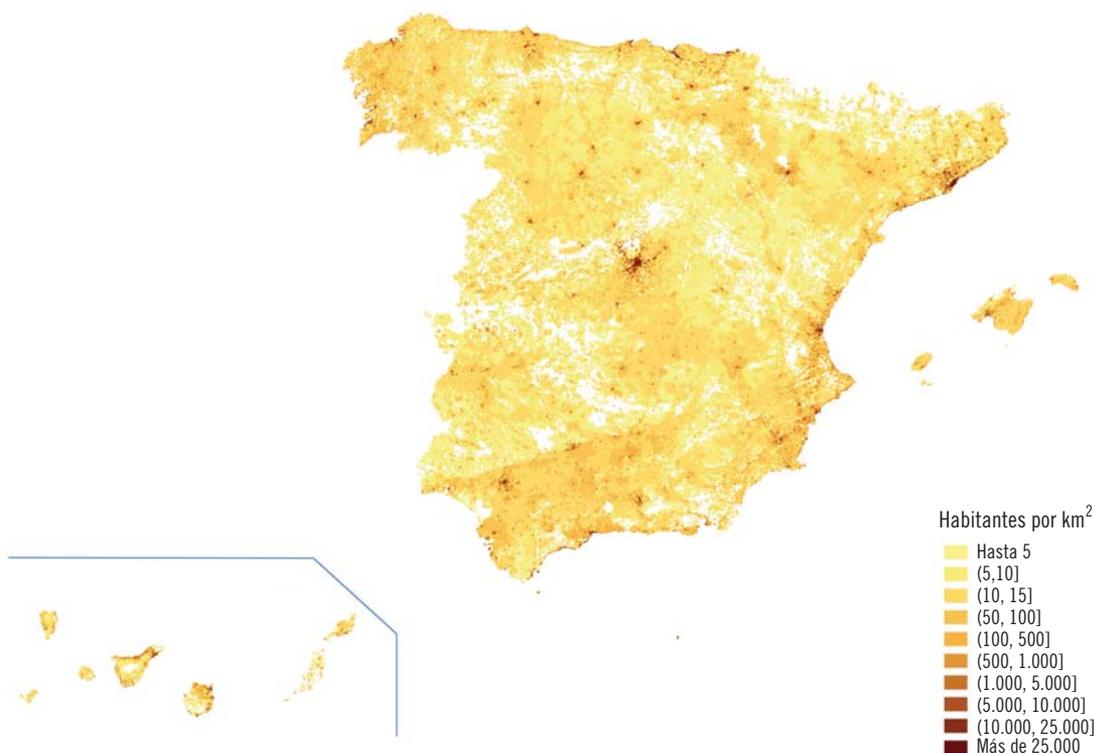
4

CORINE Land Cover y la distribución de la población

Como hemos mencionado en la Introducción, la única *grid* de población residente a escala nacional disponible actualmente para España ha sido realizada por métodos de desagregación espacial (*downscaling*), y procede de un ejercicio europeo realizado por el JRC e impulsado por Eurostat. La fecha de referencia de la población es el censo del 2001, y la resolución espacial es de 1 ha. La población de partida son los datos a nivel de municipio, y como información auxiliar se utilizan los datos de coberturas del suelo de

CORINE Land Cover (CLC). El mapa 4.1 muestra la representación de esta *grid* de densidad de población agregada a una resolución de 1 km².

Una inspección visual de este mapa muestra algunas zonas de España despobladas, algo que no es posible con los mapas de coropletas de los mapas 3.4 y 3.19 que acaban distribuyendo la población de forma uniforme sobre el polígono al que está asignada. Al mismo tiempo se observa una mayor concentración de la población



Mapa 4.1 Grid de población del JRC-Eurostat-AEMA. Censo del 2001

Fuente: EFGS (2012) y elaboración propia.

que en los mapas de densidades municipales o por secciones censales. Esta impresión visual descansa fundamentalmente en la aparición de áreas no habitadas, más que en un aumento de la propia concentración en las áreas habitadas. En estos lugares las diferencias no son excesivas desde un punto de vista visual (Goerlich y Cantarino 2010b), aunque aparecen al manejar los datos concretos e integrarlos con otras estadísticas. Un hecho fundamental del mapa 4.1 es que no hay límites administrativos que condicionen, al menos en principio, con la representación cartográfica.

Dada la resolución de la *grid*, celdas de 1 × 1 kilómetros, el mapa 4.1 muestra amplias zonas habitadas: el 85% de las celdas representadas contienen población. En gran parte de España esto contrasta con la observación directa del paisaje. En muchos lugares, la población se encuentra altamente concentrada en unas pocas grandes ciudades³¹ o en pequeños núcleos claramente definidos, y donde grandes áreas del interior peninsular están prácticamente vacías en términos demográficos. Por esta razón sería de esperar, a priori, que una *grid* de población para España mostrara amplias zonas deshabitadas, si el criterio de localización de la población es el de residencia. La razón de que el mapa 4.1 muestre una población más dispersa de lo que un observador de la geografía esperaría a priori hay que buscarla fundamentalmente en la resolución de CLC, así como en los métodos de desagregación espacial utilizados (Gallego 2010). A ello dedicamos este capítulo.

4.1. EL PROYECTO CORINE LAND COVER

El proyecto CLC tiene como objetivo fundamental la creación de una base de datos europea a escala 1:100.000 sobre coberturas del territorio mediante la interpretación de imágenes recogidas por satélite. El apéndice 1 ofrece información detallada sobre aspectos técnicos y coberturas de CLC.

La CLC es una base de datos de coberturas del suelo de tipo jerárquico, que divide el territorio en polígonos y les asigna una cobertura única de un total de 44 clases al máximo nivel de desagregación, considerándose hasta tres niveles. El mapa 1.1 ofrece los mapas de CLC2006 para Valencia y Burgos al nivel 1, y donde las coberturas artificiales aparecen en morado. En principio es de esperar que sobre ellas se asiente la mayor parte de la población, pero una comparación simple con el mapa 4.1 indica claramente que la *grid* del JRC dispersa la población mucho más allá de las zonas artificiales.

Esta sobredispersión es mucho más evidente si tenemos en cuenta que solo una parte de estas coberturas artificiales se clasifica como zona urbana (clase 1.1 al nivel 2). A su vez las zonas urbanas se desagregan a nivel 3 en dos clases: tejido urbano continuo y tejido urbano discontinuo (clases 1.1.1 y 1.1.2, respectivamente). Sería de esperar que la práctica totalidad de la población fuera asignada a las zonas urbanas, si estas englobaran también los núcleos de población de reducidas dimensiones.

Sin embargo, un elemento crítico de CLC para los ejercicios de desagregación espacial y transferencia de datos entre sistemas zonales que hemos presentado en el capítulo 1 es que la unidad mínima cartografiable (*minimum mapping unit*, MMU) es de 25 ha. Parcelas más pequeñas son incluidas en polígonos etiquetados con la cobertura dominante, y si no existe una clase dominante el polígono se etiqueta como *heterogéneo*. Esto ocurre en un 17,8% de la superficie nacional (clase 2.4 al nivel 2). Por este motivo, la superficie recogida por CLC2006 como artificial es muy reducida, un 2%; estimando la superficie clasificada como zona urbana en tan solo el 1,25% del total nacional (clase 1.1 al nivel 2).

Como expusimos en el capítulo 1 la resolución de la información auxiliar es vital para obtener buenos resultados en la elaboración de mapas dasimétricos, por lo que la razón de la sobredispersión que parece observarse en el mapa 4.1 hay que

³¹ Según el padrón del 2010 los 62 municipios existentes de más de 100.000 habitantes concentran el 39,8% de la po-

blación española; y los 145 municipios de más de 50.000 habitantes, más de la mitad de la población (el 52,4%).

buscarla fundamentalmente en la baja resolución del CLC para un ejercicio de esta naturaleza, y en menor medida en los métodos estadísticos utilizados.

4.2. DESAGREGACIÓN ESPACIAL (DOWNSCALING) CON CLC

El método de desagregación espacial utilizado para la elaboración de la *grid* de población elaborada por el JRC se basa, a grandes rasgos, en agregar CLC a nueve clases diferenciadas según densidades de población. A partir de aquí, se modela estadísticamente la densidad de población a nivel municipal para cada tipo considerado de cobertura del suelo, y se procede a aplicar un algoritmo iterativo de reparto. Finalmente, se impone la restricción de que las celdas que constituyen el municipio deben sumar la población municipal original (restricción picnofiláctica, Tobler 1979).

Diferentes modelos estadísticos para estas densidades, valores iniciales de partida en el algoritmo y la disponibilidad de información adicional, producen resultados ligeramente diferentes. El método que menor error parece producir, validado frente a *bottom-up grids* para los países en los que se dispone de ellas, es una variante del propuesto por Eicher y Brewer (2001), que partiendo de una densidad uniforme a nivel municipal dentro de las coberturas susceptibles de soportar población va modificando iterativamente estas densidades a partir de umbrales determinados exógenamente, y ajustado el exceso o el defecto de población entre clases hasta que se alcanza la convergencia (Gallego 2009). La *grid* finalmente distribuida por la AEMA se ajusta a este método.

Sin embargo, la ausencia de zonas urbanas en un gran número de municipios, que son las que soportan inicialmente la mayor densidad de población, junto con la restricción de volumen a nivel municipal, hace que, independientemente del método utilizado, se acabe dispersando la población sobre coberturas agrícolas y forestales, a las que por otra parte no se les supone densidad nula. Gallego (2010, 463) observa

que CLC2000 no reporta áreas urbanas en el 29% de los municipios europeos, y como veremos a continuación, la situación es mucho más extrema para España. Por esta razón, e independientemente de las diferencias en la modelización, todos los métodos producen resultados similares, por lo que finalmente el mapa derivado no se aleja sustancialmente del de coropletas a partir de información municipal, aunque las diferencias aparecen cuando se relacionan espacialmente los datos de población con otras variables.

España representa un caso extremo de esta situación en el contexto europeo. Las reducidas dimensiones de muchos municipios españoles van acompañadas de núcleos urbanos pequeños, algunos de ellos minúsculos. Debido a esta situación, CLC2006 no reporta zona urbana (clases 1.1.1 o 1.1.2) en más de la mitad de ellos (56,7%), debido a que no contienen ninguna parcela urbana superior a las 25 ha. Esta situación es, además, tremendamente heterogénea por provincias. El cuadro 4.1 ofrece esta información, junto con la población y superficie municipal afectadas. En diez provincias CLC2006 no recoge zona urbana en más de un 75% de sus municipios, y en dos de ellas, Soria y Burgos, el porcentaje supera el 90%. Claramente en estos casos la población se dispersará de forma similar a un mapa de coropletas tradicional, con muy ligeras variaciones.

Las diferencias son todavía más llamativas si nos restringimos a las superficies artificiales (clase 1). En la mitad de los municipios españoles (50,3%) CLC2006 no recoge ningún tipo de cobertura artificial. En seis provincias CLC2006 no informa de cobertura artificial en más de un 75% de sus municipios, y en Soria el porcentaje alcanza a prácticamente el 90%. Tomadas en un sentido literal, estas cifras indicarían que la mayor parte de los municipios de estas provincias están deshabitados, y la población se concentra en muy pocos lugares. Lo que sucede en la práctica es que los núcleos de población son de dimensiones mínimas y no son recogidos por CLC dada su resolución (MMU de 25 ha). Estas simples cifras muestran lo poco fiable que puede ser CLC para la determinación de superficies según determinadas

CUADRO 4.1: Municipios en los que CLC2006 no informa de clase urbana (clases 1.1.1 o 1.1.2). Población según el padrón del 2010 y superficie asociada

Provincias	Municipios		Población		Superficie	
	N.º	Porcentaje	Habitantes	Porcentaje	Km²	Porcentaje
01 Álava	38	74,5	28.002	8,8	2.034,16	68,7
02 Albacete	40	46,0	26.500	6,6	4.683,68	31,4
03 Alacant /Alicante	49	34,8	26.097	1,4	887,90	15,3
04 Almería	64	62,7	56.002	8,1	3.496,45	39,8
05 Ávila	210	84,7*	47.116	27,4	5.569,31	69,2
06 Badajoz	27	16,5	12.775	1,8	1.470,13	6,8
07 Balears, Illes	7	10,4	5.898	0,5	107,30	2,1
08 Barcelona	77	24,8	35.188	0,6	2.387,93	30,9
09 Burgos	338	91,1*	70.946	18,9	11.398,48	81,3
10 Cáceres	106	48,0	45.866	11,0	4.760,97	24,0
11 Cádiz	6	13,6	8.423	0,7	325,87	4,4
12 Castelló/Castellón	94	69,6	44.506	7,4	3.823,93	57,7
13 Ciudad Real	20	19,6	8.244	1,6	2.237,68	11,3
14 Córdoba	14	18,7	17.350	2,2	898,48	6,5
15 Coruña, A	29	30,9	77.823	6,8	2.572,98	32,4
16 Cuenca	176	73,9	37.640	17,3	9.955,85	58,1
17 Girona	86	38,9	31.346	4,2	2.132,82	36,1
18 Granada	91	54,2	101.408	11,0	5.298,89	41,9
19 Guadalajara	235	81,6*	26.228	10,4	8.632,36	70,9
20 Guipúzcoa	42	47,7	42.342	6,0	659,38	34,5
21 Huelva	24	30,4	12.995	2,5	1.502,10	14,8
22 Huesca	146	72,3	37.455	16,4	8.321,34	53,3
23 Jaén	27	27,8	32.142	4,8	2.494,41	18,5
24 León	124	58,8	64.881	13,0	9.615,76	61,8
25 Lleida	139	60,2	51.093	11,6	7.210,27	59,3
26 Rioja, La	138	79,3*	32.970	10,2	3.375,51	67,1
27 Lugo	32	47,8	64.817	18,3	4.226,15	42,9
28 Madrid	25	14,0	4.607	0,1	604,22	7,5
29 Málaga	53	53,0	71.415	4,4	2.052,85	28,4
30 Murcia	2	4,4	1.503	0,1	85,32	0,8
31 Navarra	174	64,0	55.911	8,8	5.270,00	53,8
32 Ourense	60	65,2	88.946	26,5	4.806,57	66,1
33 Asturias	30	38,5	34.159	3,2	3.007,33	28,4
34 Palencia	152	79,6*	27.929	16,2	5.195,19	64,5
35 Palmas, Las	4	11,8	14.862	1,4	207,88	5,1

CUADRO 4.1 (cont.): Municipios en los que CLC2006 no informa de clase urbana (clases 1.1.1 o 1.1.2). Población según el padrón del 2010 y superficie asociada

Provincias	Municipios		Población		Superficie	
	N.º	Porcentaje	Habitantes	Porcentaje	Km ²	Porcentaje
36 Pontevedra	22	35,5	89.954	9,3	1.634,68	36,4
37 Salamanca	303	83,7*	75.297	21,3	9.516,76	77,1
38 Santa Cruz de Tenerife	8	14,8	17.867	1,7	440,89	13,0
39 Cantabria	32	31,4	21.417	3,6	2.291,17	43,6
40 Segovia	166	79,4*	39.830	24,2	4.275,07	62,4
41 Sevilla	7	6,7	7.461	0,4	349,04	2,5
42 Soria	174	95,1*	28.542	30,0	8.566,23	83,1
43 Tarragona	90	49,2	46.137	5,7	2.399,03	38,1
44 Teruel	184	78,0*	31.022	21,4	9.386,31	63,4
45 Toledo	43	21,1	16.096	2,3	1.706,87	11,1
46 València/Valencia	76	28,6	50.356	2,0	2.564,07	23,7
47 Valladolid	166	73,8	41.124	7,7	4.625,37	57,0
48 Vizcaya	48	42,9	47.881	4,2	1.065,62	48,1
49 Zamora	200	80,6*	62.576	32,2	7.902,35	74,8
50 Zaragoza	200	68,3	41.671	4,3	6.865,52	39,7
51 Ceuta						
52 Melilla						
España	4.598	56,7	1.962.616	4,2	194.898,43	38,6

* Indica porcentaje de municipios superior al 75.

Nota: Las casillas vacías tienen valor 0.

Fuente: Elaboración propia a partir del INE (2010) y CLC2006 (IGN 2010c).

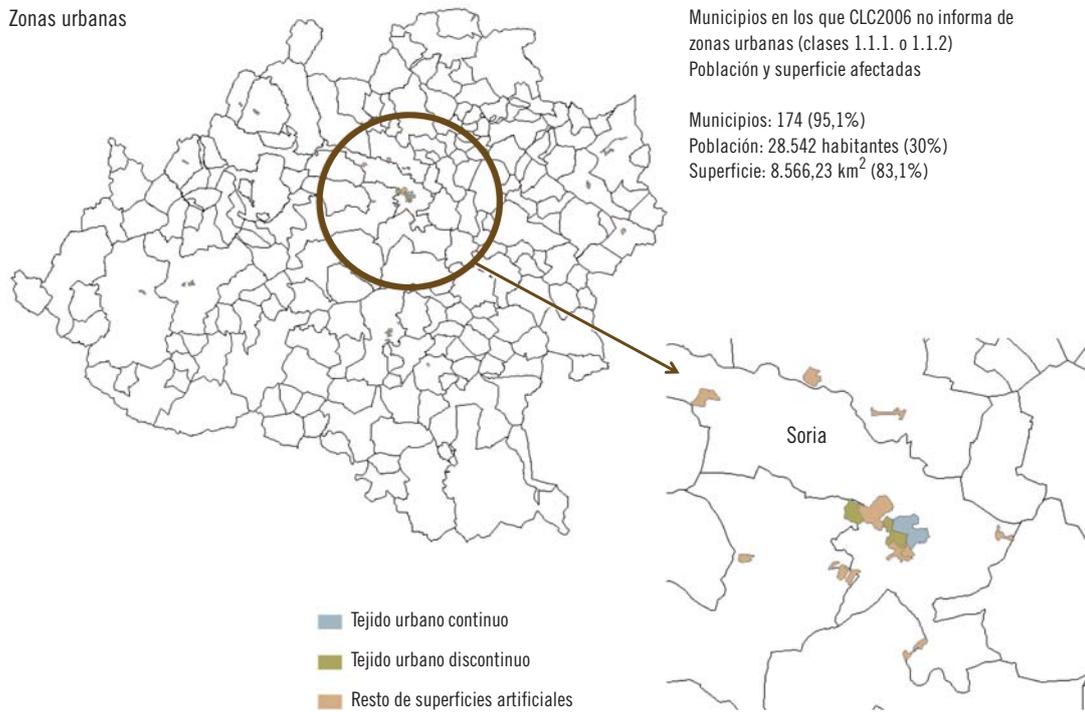
coberturas del suelo (Gallego y Bamps 2008). Los mapas 4.2 y 4.3 ilustran visualmente las coberturas artificiales de CLC2006 sobre los contornos municipales en los casos extremos de Soria y Burgos.

A nivel nacional los municipios para los que CLC2006 no muestra zona urbana (56,7%) representan el 4,2% del total de población y el 38,6% de superficie. Claramente son las zonas rurales, con baja densidad de población y núcleos pequeños, las más afectadas por los métodos de desagregación espacial utilizados. La distribución de la población utilizando CLC como información auxiliar acabará dispersando la población sobre amplias zonas del territorio que no están realmente habitadas, ante la imposibilidad de encontrar, dentro del municipio,

una zona urbana a la que asignar la mayor parte de los residentes.

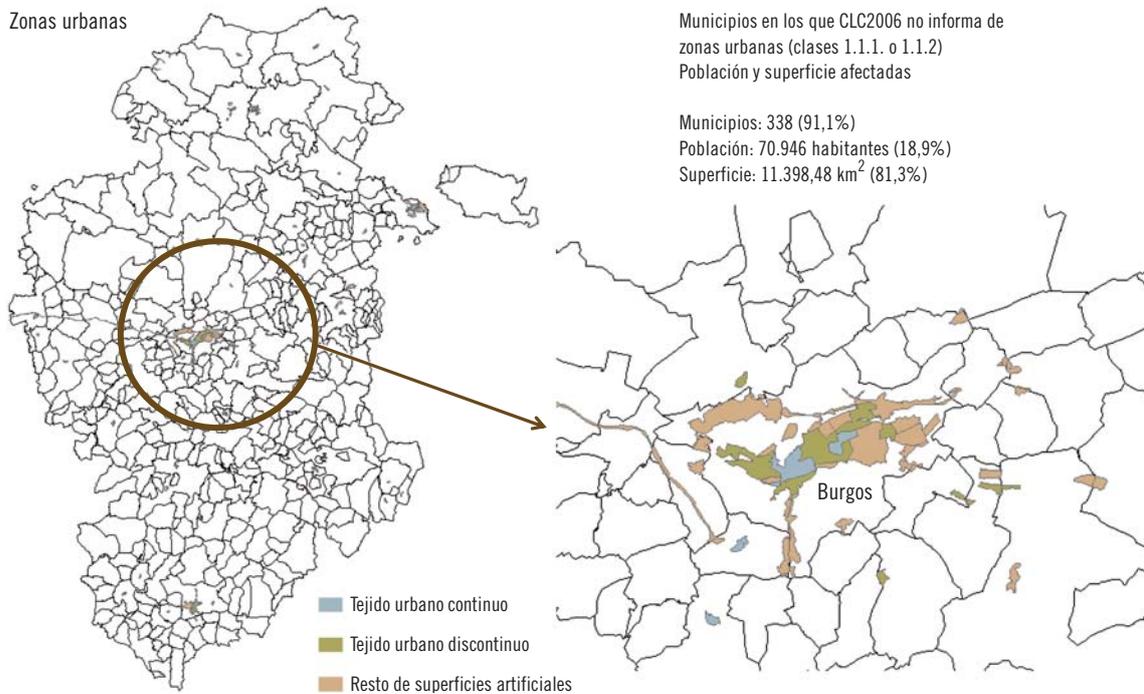
El ejercicio de la utilización de CLC para la desagregación de la población ha supuesto un paso importante, no solo en mejorar la localización de la misma sobre el territorio, sino sobre todo, en liberarla de límites administrativos predefinidos y poder combinarla de forma natural con información medioambiental. Al mismo tiempo, estas aplicaciones han puesto de manifiesto algunos hechos que son, actualmente, bien conocidos:

1. Los métodos de desagregación espacial (*downscaling* o *top-down*) presentan una tendencia a infraestimar la población urbana, mientras que, por el contrario, tienden



Mapa 4.2 Coberturas urbanas y artificiales de CLC2006. Soria

Fuente: Elaboración propia a partir del INE (2010) y CLC2006 (IGN 2010c).



Mapa 4.3 Coberturas urbanas y artificiales de CLC2006. Burgos

Fuente: Elaboración propia a partir del INE (2010) y CLC2006 (IGN 2010c).

a sobrestimar la población rural (Gallego 2010).

2. La calidad y resolución del mapa de coberturas utilizado como información auxiliar para la desagregación es más importante que la elección concreta del algoritmo de reparto (Martin, Tate y Langford 2000).
3. Los métodos dasimétricos producen resultados de baja calidad en situaciones donde las unidades administrativas de partida son altamente heterogéneas (Gallego 2010).

Por tanto, la resolución de CLC, la tendencia a la concentración de la población residente en núcleos pequeños y bien definidos, la elevada heterogeneidad de los municipios españoles, y su escasa dimensión, acaban provocando que la *grid* de población a partir de esta información sobre coberturas del suelo tienda a generar una población mucho más dispersa de lo que se observa en la realidad. Esto es cierto para Europa en

general, pero el problema es mucho más acusado en España.

El ejercicio de desagregación espacial de la población con CLC también ha puesto de manifiesto la necesidad de aumentar su resolución para determinados ejercicios. Aunque la resolución cartográfica de la *grid* del JRC es homogénea (1 ha), su calidad en términos de distribución de la población no lo es, siendo más pobre en áreas rurales con baja densidad de población que en áreas urbanas. Por esta razón existe actualmente un proyecto para mejorar CLC2006 a partir de la información de la capa de sellado del suelo (*soil sealing*) proporcionada por AEMA (2012b), con el objetivo último de generar una *grid* de población a nivel europeo, con fecha de referencia 2006, que se ajuste mejor a la realidad (Batista e Silva 2011). Como veremos en los dos capítulos siguientes, en España disponemos de información con suficiente resolución como para generar una *grid* de población a nivel nacional bastante ajustada a la realidad y con un enorme potencial de aplicaciones.

5

El Sistema de Información de Ocupación del Suelo en España: SIOSE

El proyecto sobre el Sistema de Información de Ocupación del Suelo en España (SIOSE) nace del esfuerzo y coordinación realizados por la Administración General del Estado (AGE) y las comunidades autónomas en la producción de CLC. Su objetivo es coincidente con CLC, la creación de una base de datos a nivel nacional y escala 1:25.000 sobre coberturas del territorio mediante la interpretación de imágenes recogidas por satélite. El apéndice 2 ofrece información detallada sobre aspectos técnicos y coberturas del SIOSE.

Podríamos decir que es la falta de resolución de CLC en muchos aspectos, lo que hace plantearse la posibilidad de mejorar el mismo tipo de producto a partir de fuentes de información primaria similares, y en muchos casos idénticas. De hecho, el SIOSE, dentro del Plan Nacional de Observación del Territorio en España (PNOT), dirigido y coordinado por el IGN-CNIG, integra a escala 1:25.000 la información de bases de datos de ocupación del suelo existentes en las comunidades autónomas y la AGE, con los objetivos de: 1) producir información homogénea para todo el territorio nacional; y 2) evitar duplicidades, reduciendo costes, en los procesos de producción de información sobre ocupación del suelo.

Las características fundamentales del SIOSE que nos afectan en lo que hace referencia a este trabajo son básicamente dos. En primer lugar, la unidad mínima cartografiable (MMU), que aunque variable en función de la cobertura, es de solo 1 ha para las superficies artificiales. En segundo lugar, la técnica de asignación de información temática a cada polígono.

Reducir la MMU de 25 ha en CLC a 1 ha en el SIOSE para las coberturas artificiales representa una gran mejora en términos de resolución para cualquier ejercicio de desagregación espacial de la población. A este nivel capturamos prácticamente todos los conjuntos urbanos de cualquier municipio, por pequeño que sea. Pero la característica más innovadora del SIOSE es la estructura de la información sobre ocupación del suelo.

La mayoría de los sistemas de información sobre ocupación del suelo utilizados en el análisis territorial se basan en clasificaciones fijas, y asignación de cada polígono en que se divide el territorio, según unos determinados criterios de homogeneidad, a una sola clase a partir de una serie de reglas. Esta es la filosofía seguida por CLC: dadas 44 clases, al máximo nivel de desagregación, cada polígono pertenece a una clase, y solo a una. Se trata de esta forma de un modelo jerárquico de asignación unívoca. Por el contrario en el SIOSE subyace un modelo de datos con filosofía de orientación a objetos. Ello permite, a partir de dos entidades fundamentales: el polígono y un listado inicial de coberturas simples, describir las coberturas presentes dentro de cada polígono. El objetivo final de la base de datos no es clasificar los polígonos, sino describirlos de acuerdo con una serie de relaciones, reglas de consistencia y atributos. El SIOSE, por tanto, no clasifica los polígonos en función de una nomenclatura, sino que posibilita la asignación de una o varias coberturas del suelo a un único polígono, mediante porcentajes de ocupación y atributos. Se trata de una información mucho más compleja que las clasificaciones temáticas tradicionales,

pero al mismo tiempo mucho más versátil y abierta, ya que es el investigador, en función de sus propias necesidades, el que debe diseñar la nomenclatura que necesita.

Mientras que en los modelos de tipo jerárquico, como CLC, existe un número determinado de clases distintas, siempre fijo, y el análisis de resultados suele hacerse en función de la superficie que ocupa cada clase, en los modelos orientados a objetos cada polígono almacena información sobre todas las coberturas presentes en el mismo, que se almacenan según determinadas reglas, con información sobre el porcentaje de ocupación de cada cobertura dentro del polígono. Por ejemplo, en el SIOSE se representan aquellas coberturas con una superficie de al menos el 5% de la superficie del polígono. Esto significa que, si la MMU para las superficies artificiales es de 1 ha, se representarán todos aquellos conjuntos de edificaciones, que es una de las coberturas simples consideradas, con una superficie superior a los 500 m². Por tanto, no se trata solo de un aumento notable de la resolución respecto a CLC, sino de una mejora sustancial en la descripción del territorio representado.

El número de coberturas simples del SIOSE es relativamente reducido, 40, de las cuales 7 corresponden a coberturas artificiales (apéndice 2). En lo que a nosotros respecta, la cobertura simple de edificación (EDF, 101) es la base de nuestro ejercicio de desagregación, tal y como expondremos en el capítulo siguiente. Estas coberturas simples se combinan de diferentes formas, constituyendo lo que el SIOSE denomina coberturas compuestas. Estas pueden ser combinaciones arbitrarias de coberturas simples, o formar parte de una clasificación de coberturas compuestas predefinidas incluidas en el SIOSE. Se trata de un conjunto de coberturas compuestas a las que se les ha dado una identificación por su representatividad dentro del territorio. Por ejemplo, dentro de las coberturas compuestas predefinidas de artificial compuesto de tipo urbano mixto encontramos las de casco (UCS, 811), ensanche (UEN, 812) y discontinuo (UDS, 813), informativas del tipo de estructura urbana de la que estamos hablando, y que jugarán un papel fundamental en la desagregación

espacial de la población, o los asentamientos agrícolas residenciales (AAA, 703), en un contexto más rural. Todas estas coberturas compuestas predefinidas están formadas por sus correspondientes coberturas simples, con sus porcentajes de ocupación y atributos. Adicionalmente, las coberturas compuestas pueden anidarse, con lo que la estructura de la información puede ser realmente compleja.

El conjunto de coberturas compuestas predefinidas no es un listado cerrado, pero un acuerdo sobre las mismas otorga homogeneidad al tratamiento de la información. Destaca, asimismo, el enorme esfuerzo del SIOSE en el tratamiento de las coberturas artificiales, ya que del total de 45 coberturas compuestas predefinidas consideradas, 41 forman parte del artificial compuesto, con un gran detalle no solo en lo que a estructura urbana se refiere, sino también en lo referente a coberturas relacionadas con la actividad económica, el equipamiento dotacional o las infraestructuras (apéndice 2).

Un ejemplo visual ayudará a entender la estructura de la información en el SIOSE y la utilización que le hemos dado (Villa 2009). El mapa 5.1 muestra un polígono urbano. En términos de la clasificación jerárquica de CLC este polígono sería etiquetado como *tejido urbano discontinuo*, clase 1.1.2 al nivel 3 de desagregación, dentro de las *zonas urbanas*, clase 1.1 al nivel 2 de desagregación. Esta sería toda la información que obtendríamos de CLC.

El SIOSE clasificaría el polígono como de *artificial compuesto: urbano mixto: discontinuo* (UDS, 813), que es una de las coberturas compuestas predefinidas. El descriptor del polígono aportaría, sin embargo, mucha más información, y nos diría el porcentaje de ocupación dentro del polígono de cada una de las coberturas simples presentes en él. Así, por ejemplo, nos informaría de que solo un 50% de la superficie está realmente ocupada por edificaciones (EDF, 101), a las que asignaría un atributo, vivienda unifamiliar aislada (va, 23); el resto pertenece a otras coberturas simples: un 15% de viales (VAP, 104), 30% de cubierta vegetal (FDC, 312, y MTR, 320) y un 5% de lámina de agua artificial (LAA, 103). De esta forma, a partir de un



1 polígono homogéneo:

CLC: Clase 1.1.2: Tejido urbano discontinuo.

SIOSE: Artificial compuesto: Urbano mixto: Discontinuo.

(100% de la superficie del polígono)

Coberturas dentro del polígono:

- Edificios (50%)
- Viales (15%)
- Arbolado caducifolio (20%)
- Matorral (10%)
- Lámina de agua artificial (5%)

Mapa 5.1 CLC versus SIOSE: zona urbana

Fuente: Villa (2009).

criterio genérico de homogeneidad que define el polígono obtenemos mucha más información de las coberturas que realmente contiene.

Resulta de interés destacar que, dada la estructura de la información en el SIOSE, disponemos de dos tipos de superficie que aportan información complementaria. Por una parte, tenemos la superficie del polígono, proporcionada directamente por el SIOSE en la base de datos, y por otra, la superficie de la cobertura simple dentro del polígono, obtenida como el producto de su porcentaje de ocupación por la superficie del polígono. Así, por ejemplo, si el polígono del mapa 5.1 ocupara una superficie de 10 km², la superficie realmente edificada sería solo de la mitad, 5 km², ya que la superficie de las edificaciones se estima en la mitad de la del polígono.

Como hemos visto reiteradamente en casos anteriores, a la hora de modelar la densidad de población las superficies sobre las que computar la misma juegan un papel decisivo en

los resultados que obtenemos. En el ejercicio que se describe con detalle en el capítulo siguiente llamaremos *superficie neta edificada* a la realmente ocupada por la edificación dentro del polígono, los 5 km² del ejemplo anterior; y esta será la que entrará en el cómputo de las densidades de población. De esta forma se aísla al máximo el efecto tamaño asociado a los polígonos donde se ubica la población residente. Dado que no es posible situar con exactitud las coberturas simples dentro de los polígonos SIOSE, una vez estimada la población, a partir de la modelización con la superficie neta edificada, esta se asignará al polígono, y será la superficie de este la que entre en el cómputo final de la *grid*.

Los mapas 5.2 y 5.3 muestran dos ejemplos adicionales de descriptores de polígonos SIOSE en el contexto de zonas agrícolas y forestales, que aunque no nos son de utilidad directa en nuestra aplicación, muestran la complejidad de la información de este modelo de coberturas del suelo.



Mapa 5.2 CLC versus SIOSE: zonas agrícolas

Fuente: Villa (2009).

1 polígono no homogéneo:

Cultivos herbáceos:

(72% de la superficie del polígono)

Coberturas dentro del polígono:

- Suelo desnudo (60%)
- Cultivos herbáceos distintos del arroz (40%)

Cultivos leñosos: Frutales:

(28% de la superficie del polígono)

Coberturas dentro del polígono:

- Suelo desnudo (12%)
- Frutales cítricos (80%)
- Matorral (2%)
- Embalses (6%)



Mapa 5.3 CLC versus SIOSE: zonas forestales

Fuente: Villa (2009).

1 polígono no homogéneo:

Arbolado forestal: Frondosas:

(55% de la superficie del polígono)

Coberturas dentro del polígono:

- Suelo desnudo (10%)
- Frondosas caducifolias (90%)

Arbolado forestal: Coníferas:

(45% de la superficie del polígono)

Coberturas dentro del polígono:

- Suelo desnudo (5%)
- Coníferas (95%)

6

La distribución de la población a partir del SIOSE: una *grid* de densidad de población para España

Presentamos ahora un ejercicio similar al expuesto en el capítulo 4, que utiliza como base la población municipal y CLC, pero a partir de una mayor resolución en las bases de datos. Como origen de la población utilizamos las secciones censales, ya hemos observado que estas no suponen una clara mejora sobre los datos municipales en contextos de baja densidad de población, pero son de gran utilidad en las grandes ciudades. Como información auxiliar sobre coberturas del suelo utilizamos el SIOSE, descrito en el capítulo anterior, y claramente superior a CLC en cuanto a cantidad de información y resolución.

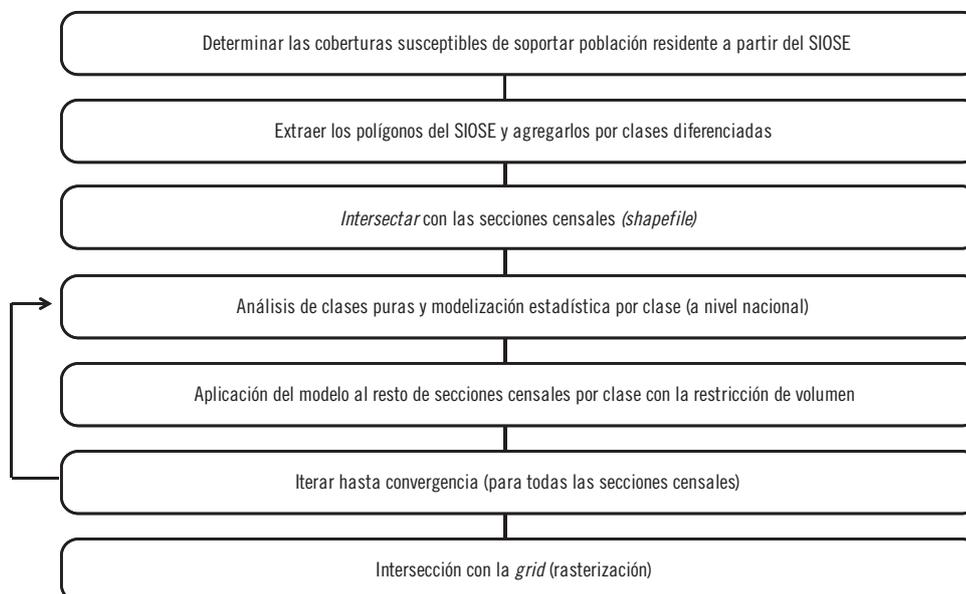
La distribución de la población se basa en una idea intermedia entre atribuir densidades diferenciales a coberturas diferentes y el método de la variable límite (*limiting variable method*) de Eicher y Brewer (2001), ya refinado por Gallego (2009) (*limiting variable method* «tuned»). En cualquier caso nosotros adaptamos la metodología a las especificidades de nuestra información. El diagrama de flujo de trabajo simplificado se ofrece en el esquema 6.1 y consta de los siguientes pasos:

1. Determinar las coberturas del SIOSE2005 susceptibles de soportar población residente.
2. Extraer dicha información y agregarla de acuerdo a las diferentes clases consideradas.
3. *Intersectar* el fichero resultante con el fichero vectorial de secciones censales del INE.
4. Examinar las secciones censales que constituyen clases puras y modelar su densidad.

5. Aplicar el modelo del paso anterior al resto de secciones censales e introducir la restricción de volumen (Tobler 1979).
6. Una vez que disponemos de poblaciones para todas las clases dentro de cada una de las secciones censales podemos, opcionalmente, iterar el modelo del paso 4 al conjunto completo de clases, hasta alcanzar un criterio de convergencia.
7. Imponer la *grid* de resolución de 1 km² sobre los resultados del paso 6.

Finalmente, podemos validar nuestros métodos gracias a la disposición de un fichero con el 98,8% de población georreferenciada para la Comunidad de Madrid en nuestro año de referencia, 2010. A partir de este fichero construimos una *bottom-up grid* que comparamos, mediante un estadístico simple de discrepancias, con la *grid* que elaboramos nosotros por los métodos de desagregación espacial descritos en este capítulo. Ensayamos diversas variedades en la modelización, para finalmente decantarnos por la *grid* que nos parece más razonable.

Gallego (2009, 2010) valida sus resultados para Europa frente a los países para los que dispone de una *bottom-up grid* (Austria, Dinamarca, Finlandia, Países Bajos, Irlanda del Norte y Suecia), y elige para su distribución al público aquel método que presenta menores discrepancias de acuerdo con su indicador. La información de la *grid* de estos países también entra en la modelización del método finalmente adoptado (Gallego 2009). Nuestros procedimientos son similares y, como veremos, podemos mejorar sustancialmente los resultados para España a partir de la información pública disponible.



Esquema 6.1 Diagrama simplificado de flujo de trabajo

Fuente: Elaboración propia.

6.1. DETERMINACIÓN DE LAS COBERTURAS SUSCEPTIBLES DE SOPORTAR POBLACIÓN RESIDENTE

Dada la estructura de datos del SIOSE2005, los polígonos susceptibles de soportar población residente deben contener edificaciones. Puesto que la unidad mínima cartografiada es de 1 ha, y el modelo de datos representa coberturas con superficie de al menos un 5% de la superficie del polígono (IGN 2011a, 10), esto indica que en el SIOSE debemos tener presente conjuntos de edificaciones de al menos 500 m², toda superficie edificada de menor tamaño no es representada en el SIOSE y la excluimos a priori como soporte de población.

A partir de esta idea restringimos los *polígonos susceptibles de soportar población residente* a

aquellos que tienen la cobertura artificial simple *edificación* (EDF, 101), con *atributos de tipo edificación* (20)³²:

- Edificio aislado (ea, 21)
- Edificio entre medianeras (em, 22)
- Vivienda unifamiliar. Aislada (va, 23)
- Vivienda unifamiliar. Adosada (vd, 24)

dentro de las coberturas compuestas predefinidas de *asentamiento agrícola residencial* (AAR, 703) y *artificial compuesto* (800): *urbano mixto* (810) de tipo: *casco* (UCS, 811), *ensanche* (UEN, 812) y *discontinuo* (UDS, 813).

En particular se excluye población en la cobertura compuesta predefinida de *huerta familiar* (UER, 704), que obligatoriamente debe incluir *edificación* (Equipo Técnico Nacional del SIOSE [ETN SIOSE] 2011d, 22). Consideramos que

³² Por razones obvias se excluyen los edificios *nave* (nv, 25) y los de *en construcción* (ec, 28), aunque parte de estos sean residenciales.

Obsérvese que parte de estos edificios estarán destinados a usos comerciales y no residenciales, pero no hay forma de distinguirlos con la información actualmente disponible.

la edificación en este tipo de cobertura compuesta no es, en su mayor parte, de carácter residencial.

Se excluye, asimismo, la posibilidad de población residente en coberturas compuestas predefinidas de tipo industrial (820), primario (830), terciario (840), o equipamiento/dotacional (850), como por ejemplo *comercial y oficinas* (TCO, 841), *camping* (TCG, 844), *educativo* (EDU, 854), *penitenciario* (EPN, 855) o *deportivo* (EDP, 858), donde sí puede haber población empadronada. Sin embargo, esta población debe ser escasa y no tenemos información a priori sobre ella, de forma que probablemente cometamos un menor error excluyendo población en estas coberturas que incluyéndola de forma generalizada. Un tratamiento genérico y uniforme de este tipo de coberturas en relación con la población residente es difícil.

Nuestro supuesto es ciertamente discutible, pero la elección refleja el necesario *trade-off* entre asegurarnos de que situamos la mayor parte de la población en su residencia, y una excesiva dispersión de la misma como consecuencia de considerar que muchas otras coberturas soportan población residente de forma generalizada. Esto provocaría un efecto parecido, aunque no tan acusado, al descrito en el capítulo 4 en relación con la CLC.

Obtenemos cierta justificación a nuestro punto de partida si cruzamos la información del padrón georreferenciado de Madrid para el 2010 con las coberturas del SIOSE2005. En este caso el 97,2% de la población se encuentra dentro de los polígonos que hemos supuesto soportan población residente³³. El resto, 2,8%, se encuentra en coberturas de carácter muy

diverso. Predominan las de artificial compuesto de tipo industrial, terciario o de equipamiento, como complejos sanitarios o educativos, aunque también coberturas de tipo vegetal a las que es difícil atribuir población con métodos generales de desagregación espacial. Entendemos que nuestro supuesto es razonable y que extenderlo a otras coberturas acabaría generando peores resultados. Ya comentamos en el capítulo 1 que no es posible competir con una *bottom-up grid*, si bien un análisis detallado de coberturas constituye una vía potencial de mejora de nuestros resultados.

Los polígonos que potencialmente soportan población residente se extraen de la base de datos del SIOSE2005. El mapa de estos polígonos se muestra en el mapa 6.1, y constituye nuestra alternativa a los contornos urbanos del mapa 3.15 construido a partir de la BCN200. Los polígonos habitables del SIOSE2005 están mucho mejor adaptados a nuestras necesidades, han sido obtenidos con una metodología uniforme y poseen una fecha de referencia clara. En estos polígonos hay que distribuir la población residente, de esta forma el mapa 6.1 nos da una primera vista de la localización de la población sobre el territorio en el sentido de que fuera de dichos polígonos no vamos a asignar población.

Estos polígonos son agregados en clases a partir de un criterio de densidad. Dada la estructura de la información resulta claro que el elemento que más influye en la densidad de población es el tipo de edificación. En consecuencia consideramos cuatro clases diferentes: edificio aislado o entre medianeras, y vivienda unifamiliar, aislada o adosada. Para cada una de estas clases consideraremos un modelo diferente que determine umbrales iniciales distintos en el algoritmo de reparto.

³³ Dicho porcentaje asciende hasta el 98% con los datos georreferenciados del padrón del 2006; más cercano

a la fecha de referencia del SIOSE2005.



Mapa 6.1 Polígonos susceptibles de soportar población residente: SIOSE2005

Fuente: SIOSE2005 (ETN SIOSE 2012f) y elaboración propia.

6.2. INTERSECCIÓN CON EL FICHERO VECTORIAL DE SECCIONES CENSALES

Los polígonos representados en el mapa 6.1, agregados por clase, se *intersectan* geoméricamente con el fichero en formato vectorial (*shapefile*) de secciones censales del INE.

Esta intersección reveló algunas deficiencias en la geometría de la información geográfica que utilizamos y que implicará que acabemos perdiendo 233.847 personas en el proceso de construcción de la *grid* respecto a la población del padrón del 2010, 47.021.031 habitantes, lo que representa un 0,5%.

Esta pérdida de población se debe a tres causas principales³⁴:

1. No existe una correspondencia exacta entre las secciones censales de la explotación del padrón municipal del INE a fecha de referencia 1 de enero, y el fichero de secciones censales proporcionado por el mismo organismo. En concreto, existen 7 secciones censales, 2 en Cáceres y 5 en Murcia, que no tienen correspondencia en el fichero vectorial, y cuya población se pierde al no disponer de territorio al que asignarla. Ello implica una pérdida de 6.197 habitantes.
2. La intersección geométrica entre las secciones censales y el SIOSE2005 es nula en términos de las coberturas que se supone albergan población residente. En principio esto podría deberse a un supuesto de partida erróneo respecto a las coberturas que

³⁴ Un detalle minucioso de estos errores geoméricos y las consecuencias que tienen en el proceso de construcción de la *grid* está disponible si se solicita a los auto-

res. A pesar de que algunos errores eran subsanables se prefirió no manipular la información geográfica de base.

soportan población³⁵, pero se comprobó que en la práctica totalidad de los casos el problema estaba en: 1) malas asignaciones del SIOSE en lo referente, normalmente, al tipo de edificación, por ejemplo, etiquetar como nave edificios residenciales; o 2) problemas con los lindes de las secciones censales, que en algunos casos son claramente erróneos³⁶.

3. El SIOSE2005 no asigna atributo a la cobertura simple *edificación* (EDF, 101) en unos pocos casos que afectan a 15 secciones censales. Puesto que en todos los casos la sección censal correspondiente tiene otros polígonos con edificación y atri-

buto de carácter residencial, este error del SIOSE2005 no supone pérdida de población, aunque sí un desplazamiento de la misma a otras coberturas dentro de la misma sección censal.

La práctica totalidad de la población perdida por estos motivos podría solucionarse con una mejora de la información geográfica de base. En total, la población que perdemos se concentra en 183 secciones censales, de las cuales 23 casos corresponden a un municipio completo.

El cuadro 6.1 ofrece el detalle de esta población distinguiendo por provincias, así como el

CUADRO 6.1: Secciones censales cuya población se pierde en el proceso de elaboración de la *grid* y población afectada (2010)

Provincias	Secciones censales			Población perdida		
	Número	Porcentaje sobre el total		Habitantes	Porcentaje sobre el total	
		Provincial	Nacional		Provincial	Nacional
01 Álava	3	1,2	1,6	6.890	2,2	2,9
02 Albacete	2	0,7	1,1	3.694	0,9	1,6
03 Alacant/Alicante	1	0,1	0,5	1.788	0,1	0,8
04 Almería	8	1,9	4,4	15.222	2,2	6,5
05 Ávila	3	1,0	1,6	2.576	1,5	1,1
06 Badajoz	2	0,4	1,1	2.154	0,3	0,9
07 Balears, Illes						
08 Barcelona	3	0,1	1,6	2.284	0,0	1,0
09 Burgos	1	0,2	0,5	1.663	0,4	0,7
10 Cáceres	6	1,4	3,3	3.542	0,9	1,5
11 Cádiz	5	0,6	2,7	7.188	0,6	3,1
12 Castelló/Castellón						
13 Ciudad Real						
14 Córdoba	7	1,1	3,8	10.550	1,3	4,5
15 Coruña, A	1	0,1	0,5	898	0,1	0,4
16 Cuenca						
17 Girona						

³⁵ Existe algún caso atribuible a este problema, por ejemplo, el del municipio de Orreaga/Roncesvalles, en Navarra, cuyos 24 habitantes empadronados en el 2010 residen en edificaciones dentro de una cobertura compuesta de tipo religioso, y que nosotros excluimos a priori. En consecuencia, esta población la perdemos en el proceso de construcción de la *grid*. Por las razones expuestas en el texto es difícil dar una respuesta general a estos casos específicos.

³⁶ Al margen de Vegaviana, en Cáceres, cuyo municipio no aparece en el fichero vectorial de secciones censales, los casos más llamativos son Albarracín (Teruel), cuyo término municipal es atribuido en su práctica totalidad a Frías de Albarracín, Mohedas de Granadilla (Cáceres) y Mombeltrán (Ávila), por idénticos motivos.

Provincias	Secciones censales			Población perdida		
	Número	Porcentaje sobre el total		Habitantes	Porcentaje sobre el total	
		Provincial	Nacional		Provincial	Nacional
18 Granada						
19 Guadalajara						
20 Guipúzcoa						
21 Huelva	45	12,5	24,6	63.250	12,2	27,0
22 Huesca						
23 Jaén	3	0,6	1,6	5.692	0,8	2,4
24 León	1	0,2	0,5	154	0,0	0,1
25 Lleida						
26 Rioja, La	1	0,3	0,5	2.179	0,7	0,9
27 Lugo						
28 Madrid	14	0,3	7,7	20.748	0,3	8,9
29 Málaga	11	1,1	6,0	17.123	1,1	7,3
30 Murcia	19	1,8	10,4	22.481	1,5	9,6
31 Navarra	1	0,2	0,5	24	0,0	0,0
32 Ourense						
33 Asturias						
34 Palencia	2	0,7	1,1	54	0,0	0,0
35 Palmas, Las	3	0,4	1,6	3.885	0,4	1,7
36 Pontevedra	4	0,5	2,2	3.297	0,3	1,4
37 Salamanca	2	0,4	1,1	241	0,1	0,1
38 Santa Cruz de Tenerife	9	1,4	4,9	11.227	1,1	4,8
39 Cantabria	4	0,9	2,2	3.645	0,6	1,6
40 Segovia	1	0,4	0,5	889	0,5	0,4
41 Sevilla	3	0,2	1,6	5.565	0,3	2,4
42 Soria	1	0,5	0,5	191	0,2	0,1
43 Tarragona						
44 Teruel	2	0,7	1,1	1.192	0,8	0,5
45 Toledo	3	0,6	1,6	4.062	0,6	1,7
46 València /Valencia	1	0,1	0,5	1.171	0,0	0,5
47 Valladolid	3	0,5	1,6	1.439	0,3	0,6
48 Vizcaya						
49 Zamora	1	0,3	0,5	513	0,3	0,2
50 Zaragoza	7	0,8	3,8	6.376	0,7	2,7
51 Ceuta						
52 Melilla						
España	183	0,5	100,0	233.847	0,5	100,0

Nota: Las casillas vacías tienen valor 0.

Fuente: Elaboración propia a partir del INE (2010) y SIOSE2005 (ETN SIOSE 2012f).

número de secciones censales afectadas. En 17 provincias no perdemos población y con excepción de Huelva, con una pérdida de población del 12,2%, que representa el 27% de la población que perdemos a nivel nacional, el resto de casos son de escasa magnitud. Pérdidas que superan el 2% de la población solo se observan en Álava (2,2%) y Almería (2,2%), y el 1% se sobrepasa en cinco provincias: Ávila (1,5%), Córdoba (1,3%), Málaga (1,1%), Murcia (1,5%) y Santa Cruz de Tenerife (1,1%).

El caso de la provincia de Huelva se debe a un problema de asignación del SIOSE2005 en la capital de provincia, donde se encuentran las 45 secciones censales afectadas, al etiquetar como nave (nv, 25) varias coberturas de edificación en lugar de asignarles un atributo de edificación no industrial.

6.3. ANÁLISIS DE SECCIONES CENSALES PURAS

Una vez efectuada la intersección geométrica, en algunas secciones censales solo encontramos una clase de cobertura, esto es, un tipo particular de edificación, independientemente de la cobertura compuesta donde esté situada, ya que lo definitorio de la clase es el atributo de la cobertura simple *edificación* (EDF, 101).

Llamamos a estos casos *secciones censales puras*, ya que en esta situación no hay problema de reparto entre clases. El SIOSE nos sirve para localizar mejor la población dentro de la sección censal, de forma que dados los polígonos SIOSE toda la población de la sección censal es asignada a esa superficie. Debido a la elevada resolución de nuestras bases de datos las secciones censales puras son más de la mitad del total (55,5%), englobando al 50,7% de la población.

El cuadro 6.2 ofrece el detalle de las secciones censales puras por provincias y clases. Solo para

la clase *edificios* entre medianeras (em, 22), que representan los 2/3 del total, existen secciones censales puras en todas las provincias. Para el conjunto de edificios se observa un mayor volumen de secciones censales puras en las provincias con mayor población, tendencia que no es evidente en el caso de las viviendas unifamiliares. A nivel del conjunto de clases, las secciones censales puras oscilan entre un máximo del 73,8% en Madrid, y un mínimo del 14,6% en Lleida.

Por su parte el cuadro 6.3 recoge la población situada en las secciones censales del cuadro anterior. A nivel nacional, la población perfectamente localizada, al nivel de precisión de nuestra información, es ligeramente superior a la mitad (50,7%), y oscila entre un máximo del 67% en Madrid y un mínimo del 17,2% en Lleida, acorde con los extremos en el número de secciones censales puras por provincias. En cuatro provincias la población residente en secciones censales puras supera el 60%: Pontevedra (66,2%), Zaragoza (65,5%), A Coruña (64,4%) y Valencia (64,2%). En otras provincias el problema del reparto es mucho más acusado, aunque además de Lleida solo Girona (18,3%) presenta porcentajes de población en secciones censales puras inferiores al 20%.

El mensaje fundamental de estos cuadros es simple. El SIOSE ayuda en la localización y limita el problema del reparto a la mitad de la población, aunque existe una gran heterogeneidad por provincias.

Como hemos señalado anteriormente el criterio básico de agregación en clases parte de la idea de densidad diferenciada, por lo que el tipo de edificación parecía el rasgo más relevante. Un análisis de varianza con inferencia robusta (White 1980) mostró que las diferencias en densidades por tipo de edificación son significativamente diferentes, incluso cuando introducimos efectos fijos provinciales y variables condicionantes a nivel de sección censal³⁷.

³⁷ Como se expone en el epígrafe siguiente, en algunos casos se consideró variabilidad a nivel de sección censal mediante la incorporación al análisis de tres variables: el

porcentaje de población de 65 años o más, el porcentaje de población extranjera y el porcentaje de residentes de la sección censal nacidos en el mismo municipio.

CUADRO 6.2: Secciones censales puras por provincia y clase (2010)

Provincia	Edificio		Vivienda unifamiliar		Total SSCC puras	
	Aislado ea, 21	Medianeras em, 22	Aislada va, 23	Adosada vd, 24	Número	Porcentaje sobre el total
01 Álava	45	86	10		141	54,9
02 Albacete	2	137		1	140	46,5
03 Alacant/Alicante	34	531	141	17	723	59,0
04 Almería		105	5	48	158	38,0
05 Ávila	22	21	12	121	176	57,1
06 Badajoz	1	137	24	107	269	48,8
07 Balears, Illes		192	24	121	337	54,4
08 Barcelona	389	1.501	90	16	1.996	56,0
09 Burgos	19	132	7	254	412	71,4
10 Cáceres		72	37	118	227	54,4
11 Cádiz	7	318	58	35	418	47,1
12 Castelló/Castellón		225	19	1	245	56,1
13 Ciudad Real	2	137	1		140	32,9
14 Córdoba		154	38	156	348	56,3
15 Coruña, A	14	248	367	3	632	68,9
16 Cuenca	4	139			143	46,4
17 Girona	1	56	20	10	87	16,6
18 Granada	3	126	30	54	213	32,8
19 Guadalajara		178	12	4	194	50,1
20 Guipúzcoa	71	79	31	1	182	33,4
21 Huelva	1	52	8	65	126	35,1
22 Huesca	5	87	31	13	136	47,1
23 Jaén	38	37	16	191	282	54,8
24 León	11	97	5	50	163	35,5
25 Lleida	3	42	10	2	57	14,6**
26 Rioja, La	3	140	9	2	154	45,8
27 Lugo		49	112		161	53,3
28 Madrid	215	2.717	88	99	3.119	73,8*
29 Málaga		367	49	130	546	54,3
30 Murcia	6	496	31	129	662	61,2
31 Navarra		168	24	2	194	35,1
32 Ourense		4	106	5	115	40,6
33 Asturias		390	145		535	61,1
34 Palencia	16	28	6	129	179	63,3
35 Palmas, Las	115	184	27	24	350	51,9
36 Pontevedra	7	145	366	4	522	70,3

CUADRO 6.2 (cont): Secciones censales puras por provincia y clase (2010)

Provincia	Edificio		Vivienda unifamiliar		Total SSCC puras	
	Aislado ea, 21	Medianeras em, 22	Aislada va, 23	Adosada vd, 24	Número	Porcentaje sobre el total
37 Salamanca	3	115	2	194	314	59,6
38 Santa Cruz de Tenerife	47	179	66	7	299	46,6
39 Cantabria	79	84	76		239	52,0
40 Segovia	9	50	6	99	164	60,5
41 Sevilla		417	136	187	740	57,3
42 Soria	1	38	3	113	155	72,1
43 Tarragona	25	113	2	4	144	27,2
44 Teruel	8	115	26	46	195	68,7
45 Toledo	1	137	5	4	147	28,3
46 València/Valencia	44	1.061	67	20	1.192	66,3
47 Valladolid	44	205	8	101	358	62,7
48 Vizcaya	177	272	38		487	52,4
49 Zamora	1	38	1	156	196	61,3
50 Zaragoza	163	348	73	49	633	72,3
51 Ceuta	8	11	1	1	21	39,6
52 Melilla	1	9	1	1	12	31,6
España	1.645	12.769	2.470	2.894	19.778	55,5

*73,8 es el valor máximo.

**14,6 es el valor mínimo.

Nota: Las casillas vacías tienen valor 0.

Fuente: Elaboración propia a partir del INE (2010) y SIOSE2005 (ETN SIOSE 2012f).

CUADRO 6.3: Población en secciones censales puras por provincia y clase (2010)

Provincia	Edificio		Vivienda unifamiliar		Total SSCC puras	
	Aislado ea, 21	Medianeras em, 22	Aislada va, 23	Adosada vd, 24	Población	Porcentaje sobre el total
01 Álava	59.520	97.884	11.382		168.786	53,2
02 Albacete	3.576	176.117		1.448	181.141	45,1
03 Alacant/Alicante	51.294	685.926	184.620	27.370	949.210	49,3
04 Almería		164.162	3.200	61.136	228.498	32,9
05 Ávila	8.348	22.282	1.727	24.686	57.043	33,2
06 Badajoz	1.348	187.882	23.482	122.360	335.072	48,4
07 Balears, Illes		310.335	39.532	209.367	559.234	50,6
08 Barcelona	530.844	2.192.374	108.368	21.185	2.852.771	51,8
09 Burgos	25.321	143.092	1.957	50.546	220.916	58,9
10 Cáceres		86.984	28.949	89.835	205.768	49,6
11 Cádiz	8.816	392.533	73.067	43.699	518.115	41,9
12 Castelló/ Castellón		300.792	26.796	1.212	328.800	54,4

CUADRO 6.3 (cont): Población en secciones censales puras por provincia y clase (2010)							
Provincia	Edificio		Vivienda unifamiliar		Total SSCC puras		
	Aislado ea, 21	Medianeras em, 22	Aislada va, 23	Adosada vd, 24	Población	Porcentaje sobre el total	
13	Ciudad Real	2.483	153.642	648	156.773	29,6	
14	Córdoba		201.532	39.870	201.878	443.280	55,1
15	Coruña, A	23.035	312.849	398.469	3.787	738.140	64,4
16	Cuenca	6.730	54.811			61.541	28,3
17	Girona	1.064	101.523	18.388	16.535	137.510	18,3
18	Granada	2.351	156.660	36.220	66.976	262.207	28,6
19	Guadalajara		60.849	14.367	7.789	83.005	33,0
20	Guipúzcoa	90.695	91.294	21.227	1.285	204.501	28,9
21	Huelva	1.448	71.529	11.551	97.420	181.948	35,1
22	Huesca	3.081	81.478	21.266	3.204	109.029	47,7
23	Jaén	51.736	51.704	16.750	238.420	358.610	53,5
24	León	16.514	109.963	4.807	28.480	159.764	32,0
25	Lleida	3.466	69.456	1.529	1.322	75.773	17,2**
26	Rioja, La	3.298	142.174	894	2.577	148.943	46,2
27	Lugo		67.365	115.103		182.468	51,6
28	Madrid	278.395	3.790.031	114.605	146.433	4.329.464	67,0*
29	Málaga		498.338	84.287	184.097	766.722	47,6
30	Murcia	7.352	625.442	38.071	163.666	834.531	57,1
31	Navarra		233.319	7.466	1.032	241.817	38,0
32	Ourense		6.105	133.508	4.482	144.095	43,0
33	Asturias		504.995	129.989		634.984	58,6
34	Palencia	20.767	27.923	487	43.212	92.389	53,6
35	Palmas, Las	162.742	282.062	36.814	29.170	510.788	46,8
36	Pontevedra	9.544	185.689	437.151	4.580	636.964	66,2
37	Salamanca	3.817	121.032	1.231	47.252	173.332	49,0
38	Santa Cruz de Tenerife	59.296	259.555	94.754	11.527	425.132	41,4
39	Cantabria	103.730	100.945	75.016		279.691	47,2
40	Segovia	8.396	27.610	3.126	28.109	67.241	40,9
41	Sevilla		541.624	192.183	256.760	990.567	51,7
42	Soria	740	19.362	1.438	15.044	36.584	38,4
43	Tarragona	41.320	157.424	4.416	1.249	204.409	25,3
44	Teruel	10.796	43.807	5.849	10.458	70.910	48,8
45	Toledo	2.842	158.585	7.620	7.261	176.308	25,3
46	València/Valencia	72.865	1.466.058	94.526	24.251	1.657.700	64,2
47	Valladolid	52.608	202.481	4.006	51.456	310.551	58,2
48	Vizcaya	217.355	308.628	27.899		553.882	48,0

CUADRO 6.3 (cont): Población en secciones censales puras por provincia y clase (2010)						
Provincia	Edificio		Vivienda unifamiliar		Total SSCC puras	
	Aislado ea, 21	Medianeras em, 22	Aislada va, 23	Adosada vd, 24	Población	Porcentaje sobre el total
49 Zamora	1.877	43.293	222	52.823	98.215	50,6
50 Zaragoza	218.597	367.273	38.203	13.674	637.747	65,5
51 Ceuta	10.322	16.327	780	2.916	30.345	37,7
52 Melilla	1.350	19.306	1.294	644	22.594	29,7
España	2.179.679	16.494.406	2.739.110	2.422.613	23.835.808	50,7

* 67,0 es el valor máximo.

** 17,2 es el valor mínimo.

Nota: Las casillas vacías tienen valor 0.

Fuente: Elaboración propia a partir del INE (2010) y SIOSE2005 (ETN SIOSE 2012f).

Se experimentó con la consideración de clases puras a partir de los tipos de edificación por cobertura compuesta, pero el número de clases puras no fue suficiente para una modelización por clase adecuada, por lo que finalmente mantuvimos las mencionadas cuatro clases.

Sin duda alguna el tamaño municipal también afecta a la densidad por tipo de edificación; dado que no disponemos de información sobre la altura de los edificios deberíamos esperar que estas clases mostraran mayores densidades en las grandes ciudades que en los municipios de pequeño tamaño. Por esta razón, podría pensarse en estratificar por tamaño de municipio y definir clases por tipo de edificación dentro de cada estrato. Los cuadros siguientes exploran esta idea.

El cuadro 6.4 muestra el número de secciones censales puras por tamaño del municipio y tipo de edificación, y ofrece un patrón claro. El grueso de secciones censales puras para la clase de edificios se encuentra en municipios de tamaño grande, por encima de los 50.000 habitantes; mientras que se observa la tendencia contraria en las viviendas unifamiliares, sobre todo en las adosadas, que se encuentran concentradas en los municipios de tamaño muy reducido.

Por su parte el cuadro 6.5 recoge la población en las secciones censales del cuadro anterior, y el cuadro 6.6 la densidad en habitantes por hectárea de superficie neta edificada, distinguiendo por tamaño de municipio y clase de edificación. Este cuadro confirma la intuición que hemos mencionado anteriormente, la densidad

CUADRO 6.4: Número de secciones censales puras por tamaño municipal y clase (2010)						
Tamaño municipal	Edificio		Vivienda unifamiliar		Total SSCC puras	
	Aislado ea, 21	Medianeras em, 22	Aislada va, 23	Adosada vd, 24	Número	Porcentaje sobre el total
Hasta 2.000	32	791	485	1.427	2.735	45,9
(2.000; 10.000]	18	272	650	410	1.350	28,0**
(10.000; 50.000]	180	1.777	729	550	3.236	41,7
(50.000; 500.000]	925	5.856	510	441	7.732	66,7
Más de 500.000	490	4.073	96	66	4.725	85,8*
España	1.645	12.769	2.470	2.894	19.778	55,5

* 85,8% es el valor máximo.

** 28% es el valor mínimo.

Fuente: Elaboración propia a partir del INE (2010) y SIOSE2005 (ETN SIOSE 2012f).

CUADRO 6.5: Población en secciones censales puras por tamaño de municipio y clase (2010)
 (habitantes)

Tamaño municipal	Edificio		Vivienda unifamiliar		Total SSCC puras	
	Aislado ea, 21	Medianeras em, 22	Aislada va, 23	Adosada vd, 24	Población	Porcentaje sobre el total
Hasta 2.000	5.843	172.395	214.820	371.995	765.053	27,2
(2.000; 10.000]	26.761	372.683	729.492	518.674	1.647.610	23,2**
(10.000; 50.000]	250.664	2.539.832	975.639	798.007	4.564.142	36,6
(50.000; 500.000]	1.226.425	7.897.345	679.417	658.980	10.462.167	61,6
Más de 500.000	669.986	5.512.151	139.742	74.957	6.396.836	83,6*
España	2.179.679	16.494.406	2.739.110	2.422.613	23.835.808	50,7

* 83,6 es el valor máximo.

** 23,2 es el valor mínimo.

Fuente: Elaboración propia a partir del INE (2010) y SIOSE2005 (ETN SIOSE 2012f).

CUADRO 6.6: Densidad en secciones censales puras por tamaño de municipio y clase (2010)
 (hab./ha)

Tamaño municipal	Edificio		Vivienda unifamiliar		Total SSCC puras	
	Aislado ea, 21	Medianeras em, 22	Aislada va, 23	Adosada vd, 24	Densidad	España = 100
Hasta 2.000	31,74	40,09	33,82	38,27	37,22	17,70**
(2.000; 10.000]	389,74	145,58	43,76	110,09	68,62	32,63
(10.000; 50.000]	523,78	295,38	66,54	170,54	160,61	76,37
(50.000; 500.000]	567,25	463,33	88,69	226,62	351,37	167,07
Más de 500.000	633,09	620,25	300,28	456,30	604,90	287,63*
España	551,59	398,51	59,79	109,21	210,31	
España = 100	262,28*	189,49	28,43**	51,93		100,00

* 287,63 / 262,28 son los valores máximos.

** 17,70 / 28,43 son los valores mínimos.

Fuente: Elaboración propia a partir del INE (2010) y SIOSE2005 (ETN SIOSE 2012f).

crece con el tamaño municipal: partiendo de densidades relativamente similares en los municipios de hasta 2.000 habitantes el crecimiento se produce en todas las clases, aunque lo hace de forma muy diferente. En los edificios el crecimiento es brusco al principio, sobre todo en los de tipo aislado, mientras que en las viviendas unifamiliares aisladas el aumento en la densidad es mucho más gradual hasta llegar a los municipios de más de 500.000 habitantes. Las viviendas adosadas representan un caso intermedio.

De nuevo un análisis de varianza con inferencia robusta (White 1980) mostró que las diferencias

en densidades por tipo de edificación y estratos de tamaño municipal son significativamente diferentes. Al igual que en el caso anterior esta conclusión resiste la introducción de efectos fijos provinciales y variables condicionantes a nivel de sección censal. Sin embargo, el número de observaciones en las que se basa es muy reducido para algunas celdas (cuadro 6.4), por lo que finalmente decidimos mantener las cuatro clases asociadas al tipo de edificación e introducir el tamaño municipal en la modelización del algoritmo de reparto. A ello dedicamos el epígrafe siguiente.

6.4. MODELIZACIÓN ESTADÍSTICA Y GENERALIZACIÓN AL RESTO DE SECCIONES CENSALES

El algoritmo de reparto es muy sencillo y se basa en manipular la propia definición de densidad de población, $Densidad = Población / Superficie$, reescrita en términos de la población, que es nuestra variable de interés, $Población = Densidad \times Superficie$. Dada la estructura de la información en el SIOSE, la superficie manejada debe entenderse como superficie edificada neta, en el sentido definido en el capítulo anterior, es decir, la superficie edificada dentro de la superficie del polígono donde se encuentra la edificación.

Así pues, llamando P a la población, S a la superficie y d a la densidad tenemos que $d = P / S$ implica $P = d \times S$. Esta identidad se verifica para cada clase, c , y cada unidad administrativa, secciones censales en nuestro caso, sc . Por tanto, $P_c = d_c \times S_c$ y $P^{sc} = d^{sc} \times S^{sc}$. Dada una partición exhaustiva de la superficie habitable de las secciones censales en diferentes clases, la relación anterior se verifica para cada sección censal y clase, $P_c^{sc} = d_c^{sc} \times S_c^{sc}$, de forma que

$$P^{sc} = \sum_c P_c^{sc} = \sum_c d_c^{sc} \times S_c^{sc} \quad (6.1)$$

La ecuación (6.1) es una identidad. Para las secciones censales puras, el número de elementos en el sumatorio es igual a la unidad, y no hay problema de reparto, observamos S_c^{sc} y $P_c^{sc} = P^{sc}$. Para el resto de secciones censales, puesto que disponemos de S_c^{sc} a partir de la intersección geométrica del SIOSE2005 y el fichero vectorial de secciones censales, diferentes supuestos sobre d_c^{sc} implican diferentes algoritmos de reparto. Además, el algoritmo debe incorporar la restricción de volumen o picnofiláctica (Tobler 1979), de forma que $P^{sc} = \sum_c P_c^{sc}$. Esta población es conocida al proceder de las estadísticas disponibles para las unidades administrativas. La idea es modelar d_c^{sc} a partir de las secciones censales puras y aplicar el modelo resultante al resto de secciones censales. Existen múltiples posibilidades.

El caso más sencillo es suponer que d_c^{sc} es constante para cada c , hasta un factor de proporcionalidad,

$$d_c^{sc} = \theta_c \times \lambda^{sc} \quad (6.2)$$

El coeficiente θ_c depende solamente de la clase, el tipo de edificación por ejemplo, y puede ser estimado a partir de las secciones censales puras de forma separada para cada c . Llamaremos a estos coeficientes *umbrales de reparto*. Por ejemplo, el cuadro 6.6 ofrece las densidades por tipo de edificación en la última línea y constituyen un candidato natural para θ_c . Estas densidades se aplican a las clases del resto de secciones censales. Puesto que la población resultante a nivel de sección censal no coincidirá con la cifra proporcionada por el padrón, el factor λ^{sc} asegura que al final del proceso se cumple esta restricción de volumen, y $P^{sc} = \sum_c P_c^{sc}$ después del reparto. Si los coeficientes θ_c son conocidos (o estimados), sustituyendo (6.2) en (6.1) obtenemos

$$\begin{aligned} P^{sc} &= \sum_c \theta_c \times \lambda^{sc} \times S_c^{sc} = \lambda^{sc} \times \sum_c \theta_c \times S_c^{sc} \Rightarrow \\ &\Rightarrow \lambda^{sc} = \frac{P^{sc}}{\sum_c \theta_c \times S_c^{sc}} \end{aligned} \quad (6.3)$$

lo que equivale a un ajuste proporcional de las discrepancias iniciales para una estimación de θ_c dada.

De esta forma, partiendo de unos umbrales de reparto, que coinciden con las densidades por clase de las secciones censales puras, acabamos con unas densidades por clase finales que vienen dadas por $d_c^{sc} = \theta_c \times P^{sc} / \sum_c \theta_c \times S_c^{sc}$.

Una vez disponemos de poblaciones para todas las secciones censales y clases, P_c^{sc} , es posible reestimar de nuevo los coeficientes θ_c e iterar hasta convergencia.

El proceso es idéntico si estratificamos las clases por tamaños municipales. En este caso podemos tomar como valor de los coeficientes θ iniciales las densidades del cuadro 6.6. Si bien en este

caso disponemos de clases y estratos, digamos θ_{ch} , aunque conceptualmente podemos considerar que disponemos simplemente de $c \times h$ clases. Este es básicamente el método CLC-iterativo inicialmente ensayado por Gallego y Peedell (2001) y descrito en Gallego (2010); convenientemente adaptado a la estructura de nuestra información, y que denominaremos *SIOSE-simple*, en una etapa o con iteración, y sin estratificación o con ella por tamaños municipales.

Aunque la densidad de población se supone constante por sección censal y clase, esto solo es cierto en términos de la superficie neta edificada, y no en términos de la superficie del polígono SIOSE donde se sitúa la edificación. Esta es una característica interesante que deriva de la estructura de información en el SIOSE.

Un supuesto simplificador del método que acabamos de describir es que la ratio entre densidades de población de diferentes clases dentro de una sección censal es constante, e igual para todas las secciones censales:

$$\frac{d_c^{sc}}{d_{c'}^{sc}} = \frac{\theta_c}{\theta_{c'}}, \quad \forall sc \quad (6.4)$$

Si estratificamos por tamaños municipales, la afirmación (6.4) es cierta dentro de cada estrato. Esta es una limitación importante que deriva de la separabilidad introducida en (6.2). Para escapar a esta restricción podemos simplemente suponer que

$$d_c^{sc} = \theta_c^{sc} \times \lambda^{sc} \quad (6.5)$$

donde el factor λ^{sc} juega el mismo papel que antes de asegurar la restricción de volumen al final del proceso de reparto y θ_c^{sc} es un umbral inicial de reparto para cada sección censal y clase obtenida mediante modelización.

De la misma forma que las densidades por clase dependen en gran medida de los tamaños municipales, debe existir una elevada correlación entre densidad por clase a nivel de sección censal y densidad de población a nivel municipal. Recuérdese que la densidad se define aquí en términos de la superficie neta edificada del SIOSE, y no se trata, por tanto, de la densidad de población municipal referida a su término municipal. En términos de esta superficie, municipios con elevadas densidades deben estar formados por secciones censales con elevadas densidades. Transformando dichas densidades a logaritmos, observamos que para las secciones censales puras, dicha correlación es: 0,68, para los edificios aislados; 0,80, para los edificios entre medianeras; 0,77, para las viviendas unifamiliares aisladas; y 0,91, para las adosadas.

Por tanto, podemos estimar el siguiente modelo log-log para las secciones censales puras,

$$\log d_c^{sc} = \alpha + \beta \log d^m + u \quad (6.6)$$

donde d^m es la densidad municipal en términos de la superficie neta edificada³⁸. El modelo (6.6) puede ser estimado de forma separada para cada clase, y su proyección para la densidad puede ser utilizada como umbral de reparto por sección censal y clase para el resto de secciones censales³⁹. Estos umbrales varían a nivel de término municipal, θ_c^m en (6.5), que actúa como variable de estratificación continua. Al igual que antes, una vez disponemos de población para todas las secciones censales y clases, el procedimiento puede ser iterado hasta que los coeficientes de las ecuaciones de regresión converjan.

Para hacer que los umbrales iniciales de reparto varíen a nivel de sección censal, θ_c^{sc} , podemos introducir variabilidad a este nivel en (6.6):

$$\log d_c^{sc} = \alpha + \beta \log d^m + \gamma'x^{sc} + u \quad (6.7)$$

³⁸ También se ensayó la modelización con la variable población a nivel municipal en lugar de con la densidad, d^m . Los resultados fueron muy similares en todos los casos y no se ofrecen, ya que finalmente se prefirió mantener la modelización en términos de densidades.

³⁹ La normalidad fue rechazada en todos los modelos estimados según el estadístico de Jarque y Bera (1980), por ello la proyección para recuperar los niveles a partir de los logaritmos incorpora la corrección no paramétrica indicada en Wooldridge (2003, cap. 6).

La información disponible a nivel de sección censal es escasa y puramente demográfica, pero el porcentaje de población de 65 años o más, el porcentaje de población extranjera y el porcentaje de residentes de la sección censal nacidos en el mismo municipio resultaron significativos en (6.7)⁴⁰.

En resumen, el algoritmo de reparto es el siguiente:

1. Estimamos (6.7) para las secciones censales puras de forma separada para cada tipo de edificación, es decir, para cada clase;
2. Proyectamos el modelo estimado para obtener umbrales de reparto por sección censal y clase, θ_c^{sc} , que aplicamos al resto de secciones censales. De esta forma obtenemos una cifra de población para cada clase dentro de cada sección censal que no es clase pura;
3. Ajustamos las discrepancias respecto al valor conocido de la población de la sección censal de forma proporcional, λ^{sc} ; lo que permite satisfacer la restricción de volumen; y finalmente,
4. Una vez disponemos de poblaciones para todas las clases de todas las secciones censales reestimamos (6.7) con el total de observaciones por clase hasta la convergencia.

Los métodos basados en la estimación de (6.6) son similares a los utilizados por Gallego (2009); si bien nuestra aplicación del umbral a las diferentes clases no procede de forma secuencial, sino de forma simultánea para todas las clases consideradas⁴¹. Adicionalmente su modelización se basa en *bottom-up grids* de referencia para otros países europeos, ya que la resolución de CLC no genera suficientes unidades administrativas con coberturas puras. Nosotros no tenemos esta limitación.

Es conveniente señalar dos aspectos antes de proseguir. Por una parte, el algoritmo de reparto opera

a nivel nacional, los modelos se estiman para todo el país, e ignoran la heterogeneidad provincial que hemos mostrado en el epígrafe 6.3. Dado que no disponemos de secciones censales puras en todas las provincias y clases, una forma de introducir la heterogeneidad a nivel provincial sería considerar un modelo de datos de panel para las secciones censales puras con efectos fijos provinciales y variables ficticias aditivas y multiplicativas para las diferentes clases. Experimentamos con este tipo de modelos, pero no proporcionaron mejores resultados que los ya descritos⁴². Por otra parte, la elección final del algoritmo de reparto no descansa en una medida de bondad del ajuste, sino que obedece al intento de superar las limitaciones conceptuales asociadas a los algoritmos más sencillos, SIOSE-simple. Básicamente al hecho de que extrapolar de forma simple las densidades de las secciones censales puras provoca que la ratio entre densidades de población de diferentes clases dentro de una sección censal sea constante, e igual para todas las secciones censales consideradas (6.4). No obstante volveremos sobre esta cuestión en el epígrafe 6.6.

6.5. GENERACIÓN DE LA GRID

El proceso descrito en el epígrafe anterior desagrega la población de cada sección censal en diferentes clases y la localiza de forma exacta dentro de los límites de la sección con la información auxiliar del SIOSE2005. Disponemos de esta forma de una capa vectorial de población procedente de la intersección del SIOSE con los contornos de las secciones censales. Básicamente la capa mostrada en el mapa 6.1 a la que hemos asignado población a los diferentes polígonos por los métodos expuestos en el epígrafe anterior. Esta es una información extremadamente útil que, sin embargo, no constituye el objetivo final del estudio, y que es la representación de la población en una *grid* de 1 km² de resolución ajustada a los estándares europeos (INSPIRE 2010b).

⁴⁰ Salvo las dos últimas variables para la clase de edificios aislados.

⁴¹ Como hemos señalado en el capítulo 4, estos métodos constituyen la base de la versión 5 de la *grid* europea del JRC-EEA, y que actualmente distribuye la EEA en su sitio web (Gallego 2009).

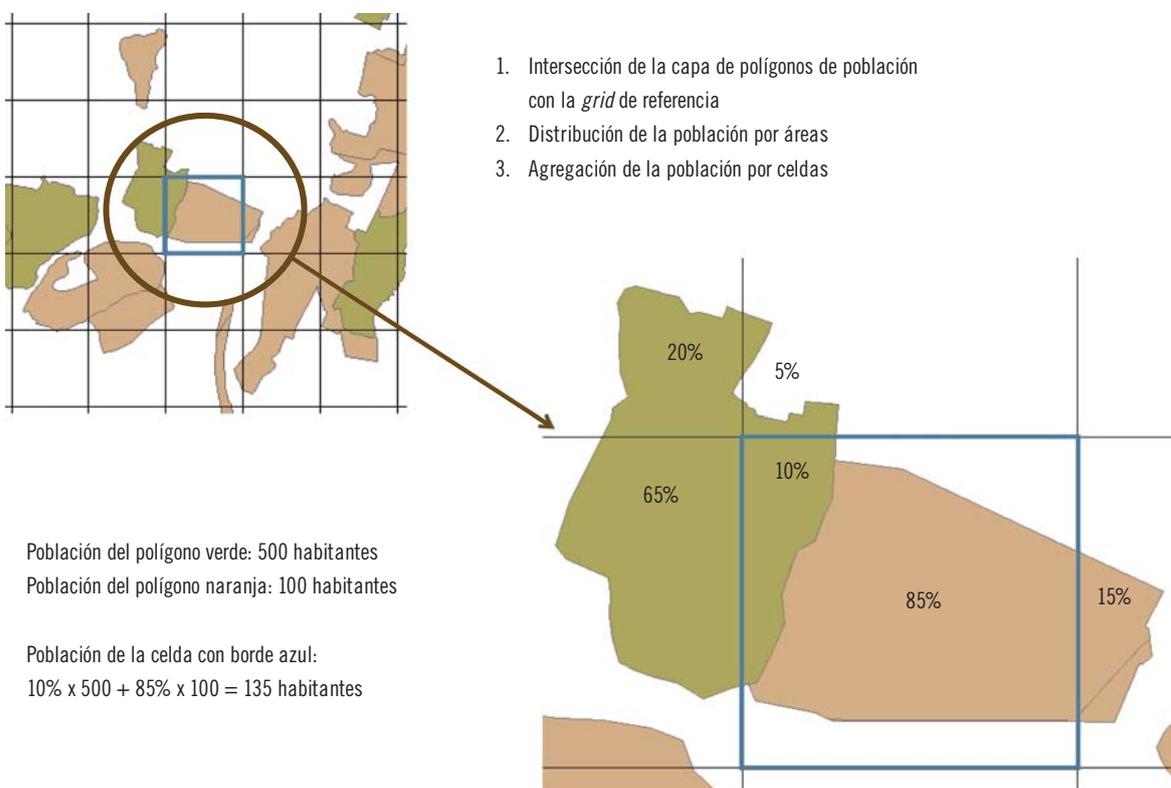
⁴² Además la iteración en estos modelos no constituye una extensión natural, puesto que al incorporar las secciones censales que tienen más de una clase, la estructura de los datos deja de ser de panel.

La generación de la *grid* a partir de esta capa es, sin embargo, inmediata. En primer lugar, debemos *interseccionar* geométricamente la *grid* de referencia con el fichero vectorial de población construido a partir del SIOSE2005 y las secciones censales. Esta capa contiene cifras de población por sección censal y tipo de edificación (clase). En segundo lugar, efectuamos una simple distribución de la población en las celdas por área tras la intersección con la *grid* (*areal weighting*). Obsérvese que estas áreas no son las superficies netas edificadas que entraban en el cómputo de las densidades por clase, sino las superficies de los polígonos SIOSE donde están ubicadas las edificaciones. Finalmente, agregamos la población por celdas de la *grid*. El mapa 6.2 muestra gráficamente el proceso.

De esta forma todo el proceso de construcción de la *grid* se realiza en vectorial y respeta escrupulosamente la transferencia de datos entre sistemas zonales, es decir, a partir de la *grid*,

y la información sobre áreas de la intersección, podemos recuperar la población por límites administrativos respetando la restricción de volumen de forma exacta.

Este cálculo final, mediante ponderación por áreas, es mucho más fino que los algoritmos habituales de rasterización de información vectorial normalmente implementados en el *software* de GIS, y que se basan en determinar el polígono que ocupa el área máxima dentro de la celda o donde se ubica el centro de la misma. Es además la única forma de no perder población en el proceso de rasterización y está en línea con las propuestas para la desagregación de datos socioeconómicos a formato de *grid* (Milego y Ramos 2010). Desde el punto de vista de la representación cartográfica es indiferente representar el mapa resultante en vectorial o en *raster*, si bien para el almacenamiento de la información el formato vectorial presenta claras ventajas.



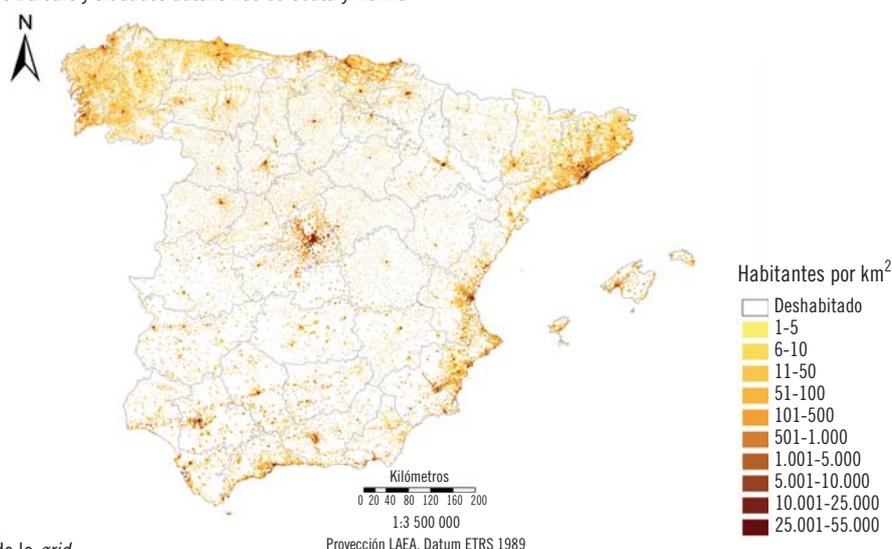
Mapa 6.2 Generación de la *grid*: distribución proporcional por área

Fuente: SIOSE2005 (ETN SIOSE 2012f) y elaboración propia.

El mapa 6.3 representa la *grid* de densidad de población para España que se deriva del proceso de cálculo que acabamos de describir, así como algunos estadísticos asociados a la *grid*⁴³. Según nuestros cálculos el porcentaje de superficie habitada no llega al 20% del área terrestre, y mientras que la densidad de población por km² se sitúa en torno a los 93 habitantes, la den-

sidad por km² habitado es considerablemente mayor, rozando los 500 habitantes. La celda con mayor densidad de población contiene 54.228 habitantes⁴⁴. El mapa 6.3 constituye una representación más realista de la distribución de la población sobre el territorio que la que hemos presentado en mapas anteriores en este informe.

a) Península, Illes Balears y ciudades autónomas de Ceuta y Melilla¹



Estadísticos de la *grid*

Fecha de referencia de la población	01/01/2010
Área total - km ²	506.011
Área terrestre - km ²	502.541
Población total - Padrón del 2010	47.021.031
Población en la <i>grid</i>	46.787.184
Población en la <i>grid</i> - %	99,5%
Celdas en la <i>grid</i> habitadas	94.440
Celdas en la <i>grid</i> habitadas - % sobre el área terrestre	18,8%
Habitantes por km ² de área terrestre	93
Habitantes por km ² habitado	495
Máxima población en una celda habitada	54.228

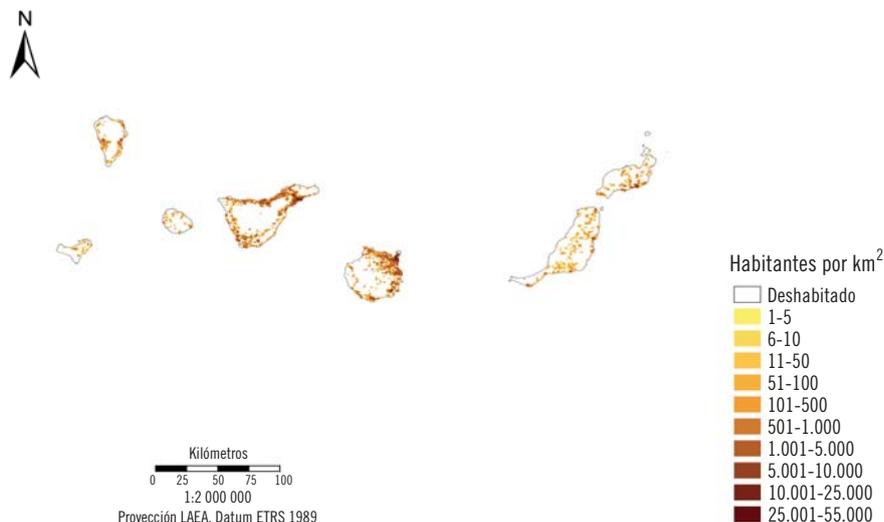
Mapa 6.3 *Grid* de población para España (2010)

¹ El área total se obtiene de ArcGIS 9.3 a partir del fichero de líneas de límite municipales del IGN (mayo 2011), originariamente en coordenadas geográficas, WGS84 (World Geodetic System 1984) para Canarias y ETRS89 (European Terrestrial Reference System 1989) para el resto, y proyección LAEA (sistema de coordenadas Lambert Azimutal Equal Area). El área terrestre deduce del área total la superficie de *cobertura de agua* (500) del SIOSE2005.

⁴³ El área terrestre, utilizada en los cálculos de densidad, sigue las recomendaciones de Eurostat (1999b) y excluye de la superficie total las coberturas correspondientes a láminas de agua. Para ello se dedujo el porcentaje de la *cobertura de agua* (500) de SIOSE2005, que asciende a un 0,69% del total (apéndice 2).

⁴⁴ Aun así, esta densidad es menor de la que se obtiene mediante el cálculo mecánico de distribuir población entre superficie a partir de las secciones censales, debido a la escasa superficie de algunas de ellas, tal y como ya mencionamos al final del capítulo 3.

b) Islas Canarias²



Estadísticos de la *grid*

Fecha de referencia de la población	01/01/2010
Área total - km ²	506.011
Área terrestre - km ²	502.541
Población total - Padrón del 2010	47.021.031
Población en la <i>grid</i>	46.787.184
Población en la <i>grid</i> - %	99,5%
Celdas en la <i>grid</i> habitadas	94.440
Celdas en la <i>grid</i> habitadas - % sobre el área terrestre	18,8%
Habitantes por km ² de área terrestre	93
Habitantes por km ² habitado	495
Máxima población en una celda habitada	54.228

Mapa 6.3 (cont.) Grid de población para España (2010)

² El área total se obtiene de ArcGIS 9.3 a partir del fichero de líneas de límite municipales del IGN (mayo 2011), originariamente en coordenadas geográficas, WGS84 para Canarias y ETRS89 para el resto, y proyección LAEA. El área terrestre deduce del área total la superficie de *cobertura de agua* (500) del SIOSE2005.

Fuente: SIOSE2005 (ETN SIOSE 2012f) y elaboración propia.

6.6. VALIDACIÓN

La elección final del algoritmo de distribución de la población se realizó a partir de las propiedades del método, y tras comprobar que los diferentes algoritmos generaban resultados similares. Es necesario, sin embargo, evaluar la bondad del ajuste de los resultados por algún procedimiento. Ello solo es posible frente a un estándar que aceptemos como verdadera distribución de la población, y que en nuestro caso viene dado necesariamente por una *bottom-up grid*.

Gallego (2009, 2010) evalúa sus resultados frente a los datos de referencia proporcionados por institutos oficiales de estadística de diferentes países que disponen de *bottom-up grids* construidas a partir de ficheros de población georreferenciados. Para ello parte de un indicador absoluto de discrepancias definido como:

$$\Delta = \sum_j |P_j - P_j^{ref}| \quad (6.8)$$

donde j indexa las celdas de la *grid*, y el superíndice *ref* se refiere a la población utilizada como referencia.

El valor de Δ oscila entre 0, cuando la distribución por celdas de la población es idéntica a la de la población de referencia, y 2 veces la población en la *grid*, cuando toda la población está en celdas diferentes. Obsérvese que se trata de un indicador de desajuste espacial un tanto ingenuo, ya que la contribución a la discrepancia es independiente del error en la distancia que cometemos al localizar incorrectamente la población, es decir, la contribución es la misma si a una persona la situamos en la celda contigua de donde debe estar o en la otra parte del país. Aunque la distancia es esencial en un indicador de discrepancias con contenido geográfico (O'Sullivan y Wong 2007) mantendremos (6.8) como indicador de referencia por motivos de comparabilidad.

Dada la dependencia de (6.8) respecto al tamaño de las poblaciones bajo comparación resulta útil reescalarlo al intervalo [0, 1], y definir el indicador de discrepancias relativo como:

$$\delta = \frac{\Delta}{2 \cdot \sum_j P_j} = \frac{\sum_j |P_j - P_j^{ref}|}{2 \cdot \sum_j P_j} \quad (6.9)$$

De esta forma $100 \times \delta$ puede interpretarse como el porcentaje de población que situamos de forma incorrecta sobre el territorio, donde la precisión de la incorrección en la localización viene determinada por la resolución de la *grid*. El cuadro 6.7 ofrece los resultados para este indicador mostrados por Gallego (2009, 2010) para los diversos métodos ensayados a partir de datos municipales y CLC, y cuya *grid* para España se ofrece en el mapa 4.1 para el último método mostrado en el cuadro 6.7⁴⁵.

De los diferentes métodos ensayados por Gallego (2009, 2010) hay una cierta preferencia por los métodos de variable límite propuestos por Eicher y Brewer (2001), si bien estas mejoras no son uniformes. La versión *refinada* de Gallego (2009) es similar a los métodos de regresión que hemos expuesto en el epígrafe anterior. En cualquier caso, el cuadro 6.7 indica claramente que las mejoras de los diferentes métodos son escasas en algunos países, aunque apreciables en otros. La heterogeneidad en la distribución de la población en los diferentes países se traslada a una disparidad en las mejoras de los diferentes algoritmos de reparto cuando estos se aplican de forma uniforme. La conclusión lógica es que, probablemente, no exista un algoritmo uniformemente mejor en todas las situaciones.

CUADRO 6.7: Índice normalizado de discrepancia, $100 \times \delta$, frente a *bottom-up grids* en diversos países europeos. *Grid* JRC-Eurostat-AEMA

Métodos de desagregación espacial	Austria	Dinamarca	Finlandia	Suecia	Países Bajos	Irlanda del Norte
Población del país	8,03	5,35	5,18	8,88	15,99	1,69
Municipios (coropletas)	55,79	56,82	65,54	70,27	58,54	33,14
CLC - iterativo	28,33	38,04	52,51	45,33	22,30	21,30
CLC - LUCAS simple	27,33	37,10	48,84	45,55	28,24	21,01
CLC - LUCAS logit	27,09*	36,92	48,55	45,44	22,14	18,93
CLC - EM	28,02	37,20	49,42	45,50	29,05	20,12
CLC variable límite simple	27,15	22,06	32,63	37,78	19,20*	19,23
CLC variable límite «refinado»	27,21	20,47*	30,89*	35,70*	19,32	17,75*

* Valor mínimo, siendo 17,75 el menor.

Fuente: Gallego (2009, 2010).

⁴⁵ Gallego (2009, 2010) nunca normaliza su indicador, lo

que dificulta su comparabilidad, y su interpretabilidad.

Por otra parte, los errores de localización cometidos se sitúan en un límite inferior en torno al 20% en el mejor de los casos.

Para validar nuestros métodos disponemos de un fichero de población georreferenciado para la Comunidad de Madrid y el año 2010. Dicho fichero cubre el 98,8% de la población del padrón, y nos permite construir una *bottom-up grid* que usamos como base de comparación. A partir de ella calculamos el estadístico δ , aunque dado que las poblaciones georreferenciadas y en la *grid* no son exactamente las mismas, la normalización adoptada es la siguiente:

$$\delta' = \frac{\Delta - \Delta_{\min}}{\Delta_{\max} - \Delta_{\min}} \quad (6.10)$$

La razón es sencilla. El valor máximo de (6.8) es 2 veces la población cuando la población en la *grid* y la de referencia son exactamente las mismas. En caso contrario, el valor máximo resulta ser la suma de las poblaciones, $\Delta_{\max} = \sum_j P_j + \sum_j P_j^{ref}$. El

valor mínimo ahora no es 0, porque si situamos toda la población correctamente en sus celdas, no todos los sumandos en (6.8) se anulan. El valor mínimo resulta ser la diferencia (en valor absoluto) de las poblaciones, $\Delta_{\min} = |\sum_j P_j - \sum_j P_j^{ref}|$. En estas circunstancias la normalización que acota el índice Δ en el intervalo [0, 1] es precisamente (6.10), lo que representa tanto un cambio de origen como de escala.

El cuadro 6.8 ofrece el índice normalizado de discrepancias, $100 \times \delta'$, para los algoritmos de distribución expuestos en el epígrafe anterior aplicados a la Comunidad de Madrid. Incluye además el índice resultante de construir la *grid* a partir de la densidad municipal, mapa 3.4, y por secciones censales, mapa 3.19. El resultado a nivel municipal se corresponde con la primera línea del cuadro 6.7 para los ejercicios de Gallego (2009, 2010). Estos dos métodos de distribución no incorporan información auxiliar sobre usos del suelo. Adicionalmente, mostramos los resultados que obtendríamos si, una vez seleccionados del SIOSE2005 los polígonos susceptibles de soportar población residente, aplicáramos un método dasimétrico binario (Langford y Unwin 1994;

CUADRO 6.8: Índice normalizado de discrepancia, $100 \times \delta'$, frente a una *bottom-up grid* en la Comunidad de Madrid

a) Población de la Comunidad de Madrid					
Padrón del 2010	Población en la <i>grid</i>		Población georreferenciada		
	Habitantes	Porcentaje	Habitantes	Porcentaje	
6.458.684	6.437.936	99,7	6.378.775	98,8	
b) Índice					
				Índice	
Municipios (coropletas)				69,02	
Secciones censales (coropletas)				12,03	
Dasimétrico binario: Superficie polígono				5,35	
Dasimétrico binario: Superficie neta edificada				4,61	
				Índice	
				Sin iterar	
				Iterativo	
Densidad simple por clases				4,04	3,99
Densidad simple por clases con estratificación por tamaños municipales				3,97	4,40
Regresión <i>log-log</i> densidades: ecuación (6.6)				4,67	4,48
Regresión <i>log-log</i> densidades con variabilidad a nivel de sección censal: ecuación (6.7)				4,53	4,41*

* Método elegido.

Fuente: Elaboración propia a partir del INE (2010), SIOSE2005 (ETN SIOSE 2012f) y datos facilitados por el Instituto de Estadística de la Comunidad de Madrid (IEM).

Langford 2007) en el que la población solo se asigna a estos polígonos, proporcionalmente a su superficie (*areal weighting*). Dada la estructura de la información en el SIOSE, la distribución proporcional por áreas puede hacerse atendiendo a dos criterios, en función de la superficie del polígono donde se encuentran las edificaciones residenciales, o en función de la superficie neta edificada.

La inspección del cuadro 6.8 revela varias conclusiones de interés. Los resultados del índice de discrepancia a nivel municipal son comparables con los mostrados por Gallego (2009, 2010), y muestran una incorrecta localización de la población en el entorno del 70%. Sin embargo, el descenso en la escala del análisis a nivel de sección censal mejora el ajuste de forma notable, reduciendo el índice a un 12%. Por tanto, solo con una desagregación suficiente en los datos demográficos, sin información auxiliar sobre coberturas del suelo, podemos obtener discrepancias notablemente inferiores a las obtenidas para la *grid* del JRC-Eurostat-AEMA, considerando los mejores métodos posibles. Este es un resultado importante, porque muestra que una difusión de datos demográficos a nivel de sección censal mejoraría de forma notable los resultados que se podrían obtener en este tipo de ejercicios.

Si además incorporamos la información del SIOSE, entonces el ajuste mejora de forma importante. Un método dasimétrico binario, asignando la población a los polígonos que incluyen las coberturas susceptibles de soportar población residente, es suficiente para reducir a la cuarta parte los errores de asignación que se obtienen con datos municipales y CLC. Si además esta asignación proporcional se realiza en función de la superficie neta edificada, en lugar de la

superficie del polígono, entonces los errores de localización se sitúan por debajo del 5%⁴⁶. A partir de aquí, todas las mejoras que se obtienen por distinción de tipos de edificación, estratificación por tamaños municipales, o modelización estadística de las densidades son puramente marginales. La iteración tiende a mejorar ligeramente la bondad del ajuste, aunque esto tampoco es uniforme para los diferentes métodos considerados. Otros métodos ensayados, o la consideración del tamaño municipal en lugar de la densidad produjeron resultados totalmente comparables.

La conclusión es sencilla, y ha sido señalada reiteradamente por varios autores: la calidad y resolución de la información de partida es mucho más importante para obtener buenos resultados que la elección concreta de los algoritmos de reparto (Martin, Tate y Langford 2000). Este argumento es válido no solo para la información auxiliar sobre coberturas del suelo, SIOSE frente a CLC, sino también para la fineza de la información demográfica utilizada, secciones censales frente a municipios.

Por esta razón, el método finalmente elegido para la elaboración de la *grid* representada en el mapa 6.3 no descansó solo en el valor del estadístico de discrepancia, sino en las propiedades del mismo, tal y como se indica en el epígrafe 6.4. Obsérvese, además, que la comunidad utilizada para validar los resultados, Madrid, es en cierta forma una comunidad atípica: con el 67% de la población en secciones censales puras, el problema de reparto se reduce a un 33% de la población. Esto explica el buen comportamiento de los métodos dasimétricos binarios en este caso, al mismo tiempo que merma las potencialidades de la modelización estadística⁴⁷.

⁴⁶ La mejora en la bondad del ajuste que va desde el mapa de coropletas de secciones censales al método dasimétrico binario según la superficie del polígono, puede identificarse como la ganancia asociada a la información auxiliar de coberturas del suelo en un modelo de tipo jerárquico como CLC. Por su parte, la que va desde este último método, a una asignación binaria en función de la superficie neta edificada, puede identificarse como la que se obtiene del modelo de datos orientado a objetos del SIOSE, para una resolución dada. Bajo

esta interpretación, resulta obvio que lo determinante es aumentar la resolución de la información de partida.

⁴⁷ Recuérdese que, como mencionamos en el epígrafe 6.1, un 2,8% de la población georreferenciada está fuera de las coberturas que hemos supuesto potencialmente habitables, con lo que los márgenes de error del cuadro 6.8 son realmente reducidos. Un examen detallado de estas coberturas podría mejorar el reparto en el caso de Madrid, pero es difícil generalizar a otros territorios y saber en qué grado se ven afectados los resultados.

Resumen y conclusiones: usos potenciales

Una determinación precisa de la localización de la población sobre el territorio, no directamente relacionada con los límites administrativos en los que se divide un país, es absolutamente esencial para numerosas cuestiones prácticas de organización social, y va más lejos de la simple representación cartográfica o del mero estudio descriptivo de la geografía humana.

Este informe ha presentado un ejercicio metodológico de elaboración de una *grid* de población residente, para la totalidad del Estado español, a partir de las secciones censales y del SIOSE como información auxiliar para la desagregación espacial. Al mismo tiempo hemos inspeccionado más generalmente el estado de las bases de datos geográficas para la integración de este tipo de información con estadísticas de carácter socioeconómico. Aunque se observa un enorme progreso en esta dirección por parte de las instituciones responsables de la información geográfica en los últimos años, la conclusión general es que hay bastantes lagunas de homogeneización, y que se trata de una asignatura pendiente por parte de la estadística oficial que deberá ser abordada en un futuro próximo.

En ausencia de estadísticas demográficas en formato de *grid*, y dado el calendario previsto para la actualización del SIOSE, los métodos aplicados en este estudio pueden ser rutinariamente utilizados para actualizar la *grid* de población. Un usuario puede razonablemente preguntarse por la utilidad de esta información más allá del mapa ofrecido en el capítulo anterior (mapa 6.3), lo que nos lleva a examinar los usos potenciales de la información elaborada.

Las líneas de investigación presentes y futuras en el campo de la adopción y utilización de las rejillas cartográficas como sistema de referencia zonal para el almacenamiento de estadísticas demográficas y socioeconómicas pueden agruparse en tres grandes bloques.

1. *Mejoras en los métodos de desagregación espacial para la elaboración de mapas dasimétricos.* Los métodos estadísticos utilizados no son especialmente novedosos, descansan sobre técnicas aplicadas anteriormente al mismo tipo de problema (Gallego 2010) y nuestro mérito descansa básicamente en la utilización de bases de datos con mucha mayor resolución, al mismo tiempo que se cubren amplias zonas del territorio.

Los métodos de desagregación espacial están, sin embargo, en la infancia y existe aquí un amplio margen para la experimentación. Tanto en lo que se refiere a incorporación de información auxiliar, por ejemplo, la ortofotografía aérea o las capas de sellado del suelo de alta resolución, como en lo referente a nuevas técnicas estadísticas. En particular, los modelos estadísticos utilizados son relativamente simples y, curiosamente, no incorporan de forma directa la geografía. Por esta razón, los métodos de econometría espacial (Anselin 1988), los modelos geográficamente no estacionarios, como la regresión geográficamente ponderada (*geographically weighted regression* [GWR], Fotheringham, Brunson y Charlton 2002) y la geoestadística más generalmente, merecen cierta atención en este campo.

Por otra parte, un análisis minucioso de las coberturas que soportan población residente puede generar algunas mejoras en los ejercicios de desagregación espacial, habida cuenta de que estas coberturas constituyen el soporte sobre el que se distribuye la población.

2. *Incorporación de otro tipo de información al mismo sistema zonal.* La población debe ser solo la primera variable objeto de transformación al formato de *grid*, pero no la única si queremos lograr una verdadera interoperabilidad en la utilización de estadísticas de origen diverso. Una extensión directa es la incorporación a la *grid* de otras características de la población: sexo, edad, nacionalidad, etcétera. Esto es directo en las *bottom-up grids*, pero los métodos de *downscaling* requerirán nuevas técnicas de desagregación, en las que restricciones adicionales de volumen deberán ser incorporadas al proceso.

Otra extensión de interés es cambiar el criterio de localización de la población, y en lugar del de residencia, utilizar, por ejemplo, el del lugar de trabajo (*workplace*). Ello lleva directamente a las denominadas *grids* de día (Hofstetter 2011), en lugar de las *grids* de noche, asociadas al lugar de residencia de la población. En esta misma dirección, en nuestro país sería conveniente el desarrollo de modelos estacionales de *grids*, ya que algunos lugares ven crecer considerablemente su población en períodos puntuales y porciones del territorio muy concretas, como es el caso del litoral mediterráneo o las estaciones de esquí.

El nexo de conexión entre estas *grids* es, sin duda, la movilidad de la población. Sería deseable modelarla, en la misma forma en que se modela su movilidad entre unidades administrativas, municipios o regiones,

y a partir de aquí desarrollar modelos de *grid* espacio-temporales en tiempo continuo (Bhaduri *et al.* 2007; Martin 2010; Bright y Bhaduri 2010).

Pero sobre todo, es necesario ir más allá de la demografía y elaborar estadísticas económicas sobre este formato. Los modelos de localización industrial y de nueva geografía económica serían mucho más fáciles de contrastar sobre esta base común de recopilación estadística (Duranton y Overman 2005). Si bien es cierto que este tipo de formato para las estadísticas económicas está todavía en sus inicios y quedan muchas cuestiones prácticas por resolver, ya existe un prototipo de base de datos mundial centrada en el *gross cell product* (GCP) en lugar de en el *gross domestic product* (GDP) desarrollada en la Universidad de Yale por el profesor William D. Nordhaus (2006, 2008), la llamada G-Econ, y la Comisión Europea está experimentando con las *grids* en el desarrollo de su nueva base de datos regional ESPON Database 2013 (Milego y Ramos 2010, 2011)⁴⁸.

3. *El desarrollo de aplicaciones potenciales.* La disponibilidad de información estadística en formato de *grid* presenta numerosas aplicaciones prácticas de interés. Quizá el ejemplo más directo sería el de la determinación de la población en riesgo ante determinadas catástrofes naturales (Tralli *et al.* 2005). En estos casos una localización exacta de población es absolutamente esencial. En consecuencia la evaluación de daños, el diseño y efectividad de políticas de intervención depende en gran medida de la localización de la población y otros activos físicos y económicos (Chen *et al.* 2004).

Las cuestiones relacionadas con el análisis de la accesibilidad a determinadas infraes-

⁴⁸ La resolución espacial de muchas variables de naturaleza económica no es lo suficientemente elevada como para su desagregación directa a una *grid* de 1 km², ya que solo están disponibles a nivel regional. En estos casos es posible aplicar una estrategia en dos etapas: en la primera etapa se efectúa una desagregación por

métodos estadísticos tradicionales a unidades administrativas más pequeñas, por ejemplo, los municipios (Chasco 1997; Elbers, Lanjouw y Lanjouw 2003; Elbers *et al.* 2003), y en la segunda etapa se desagrega a un formato de *grid*.

estructuras y servicios no pueden basarse en *situar* la población de un municipio o región en un solo punto, ya sea este el núcleo principal o el centroide del polígono (Gutiérrez, Gómez Cerdá y García Palomares 2009). Esta relación entre accesibilidad y localización es fundamental en el desarrollo de los modelos dinámicos de movilidad *intradía* a los que hemos hecho referencia en el punto anterior, lo que a su vez tiene interés en la gestión y planificación de infraestructuras urbanas y en el análisis de la dinámica de las grandes urbes (Smith 2011).

En un mundo en el que la agricultura pierde cada vez más contenido informativo sobre el carácter rural de los asentamientos humanos (Collantes 2007) una *grid* de población puede ser la base sobre la que construir nuevos indicadores de ruralidad a pequeña escala, así como delinear zonas rurales o urbanas directamente relacionadas con la geografía y no con los lindes administrativos (Librecht *et al.* 2004). Al mismo tiempo es posible construir índices de heterogeneidad rural/urbana para regiones (NUTS 2 o 3) mediante la agregación del carácter atribuido a cada celda particular de la *grid*, y estudiar las características que unas zonas tienen sobre otras (Dijkstra y Poelman 2008). Todo ello en la línea de superar la excesiva dependencia que la densidad de población regional tiene en la dicotomía rural/urbano (OCDE 1994).

La información sobre coberturas del suelo del SIOSE puede ser utilizada, de forma mucho más precisa que la de CLC, para redefinir las zonas urbanas morfológicas (*urban morphological zones*, UMZ) delimitadas por la AEMA (Simon, Fons y Milego 2010), y a partir de aquí una *grid* de población puede proporcionar una aproximación GIS a la delimitación de aglomeraciones urbanas. Algo sobre lo que ya trabaja actualmente la Comisión Europea dentro de los proyectos *Urban Audit* y *Urban Atlas* (Eurostat 2010; Poelman 2011).

Los estudios de impacto medioambiental requieren con frecuencia fijar la población afectada por una determinada acción, y la zona afectada puede tener poco que ver con los lindes administrativos en los que las estadísticas tradicionales están disponibles. Por ejemplo, la generación de mapas de ruidos (Vinkx y Visée 2008), o el impacto del riesgo ante incendios forestales (Barbosa *et al.* 2008). Como hemos mencionado repetidamente a lo largo del informe, el estudio de la presión de la actividad humana sobre el medio ambiente requiere con frecuencia combinar estadísticas procedentes de diferentes sistemas zonales, y la disponibilidad de una *grid* de población facilita enormemente el trabajo, al tiempo que incrementa la precisión de los efectos estimados (Weber y Christophersen 2002).

En definitiva, la relación entre el hábitat humano y su entorno requiere de sistemas de información flexibles en los que la combinación y el análisis de estadísticas se realicen con facilidad y eficiencia. En este contexto los GIS constituyen un entorno natural, y las *grids* resultan ser un sistema de referencia zonal tremendamente versátil, ya sea como sistema final de almacenamiento de la información o como paso intermedio entre la transferencia de datos entre diferentes sistemas zonales.

Como conclusión final, podemos decir que la *grid* de población que hemos presentado en este estudio tiene un enorme potencial para la investigación de las relaciones entre la localización de los asentamientos urbanos y sus relaciones con el medio ambiente. Este potencial se verá reforzado cuando dispongamos de más períodos temporales, si bien algunas de las aplicaciones mencionadas son actualmente objeto de investigación.

La información generada a lo largo de este informe, en particular la *grid* de población para el 2010 en formato GIS, se encuentra disponible a los investigadores que lo soliciten a través de los autores.

Apéndices

A.1. COBERTURAS DE CORINE LAND COVER

El programa *CORINE (Coordination of Information on the Environment)* fue establecido por la Comisión Europea en 1985 (CE/338/85) con el objetivo de recopilar, coordinar y homogeneizar la información sobre el estado del medio ambiente y los recursos naturales a nivel europeo, y todo ello en el contexto de los Sistemas de Información Geográfica (GIS)⁴⁹.

El proyecto sobre coberturas del suelo, *CORINE Land Cover* (CLC), tiene como objetivo fundamental la creación de una base de datos europea a escala 1:100.000 sobre coberturas del territorio mediante la interpretación de imágenes recogidas por los satélites Landsat y Spot⁵⁰. Desde 1995 CLC es responsabilidad de la Agencia Europea del Medio Ambiente (AEMA; European Environment Agency, EEA), quien distribuye los datos en su sitio web. Actualmente existen tres ediciones de CLC con fechas de referencia: 1990, CLC90; 2000, CLC2000; y 2006, CLC2006. Asimismo, está prevista una nueva actualización para 2012 dentro del programa GMES (*Global Monitoring for Environment and Security*).

⁴⁹ GIS: *Geographical Information System*. Los acrónimos utilizados en el texto se corresponden a su original en inglés, por la internacionalización de su uso y para evitar equívocos.

⁵⁰ Otros proyectos dentro del programa *CORINE* son *CORINE Air*, *CORINE Coastal Erosion* o *CORINE Biotopes*.

⁵¹ Este es el fichero inicial de distribución, descargado el 13/09/2010, posteriormente el CNIG ha sustituido el fichero único por dos ficheros, uno para la Península, Illes Balears, y las ciudades autónomas de Ceuta y Melilla, con las características técnicas señaladas, y otro para Canarias, en sistema de referencia geodésico REGCAN95 (compatible con WGS84) y proyección UTM en el huso 28N.

Los datos de CLC utilizados en este informe corresponden a la edición de 2006 (CLC2006), y proceden del Centro de Descargas del Centro Nacional de Información Geográfica (CNIG). El fichero de trabajo es único para toda España, incluida Canarias, y se ofrecen en formato vectorial *shapefile* de ESRI (Environmental Systems Research Institute), sistema de referencia geodésico ETRS89 (*European Terrestrial Reference System* 1989) y proyección UTM (*Universal Transverse Mercator*) en el huso 30N⁵¹.

Las principales características técnicas de CLC2006 son las siguientes:

- *Escala cartográfica*: 1:100.000.
- *Unidad mínima cartografiable* (MMU): 25 ha.
- *Anchura mínima de elementos lineales*: 100 m.
- *Modelo de datos*: jerárquico: 3 niveles y 44 clases al nivel 3⁵².
- *Sistema geodésico de referencia*: ETRS89.
- *Sistema cartográfico de proyección*: UTM (*universal transverse mercator*), huso 30 extendido.

El cuadro A.1.1 ofrece la descripción de contenidos de los 3 niveles y las 44 coberturas de CLC.

El sistema de referencia geodésico oficial para Canarias es REGCAN95, completamente compatible con los sistemas ETRS89 y WGS84. De hecho REGCAN95 no está implementado en ArcGIS 9.3, por lo que el IGN, cuando ofrece datos georreferenciados para Canarias incluye un archivo de georreferenciación en WGS84, cuya transformación a ETRS89 es inmediata para la precisión con la que trabajamos.

⁵² Existe una versión de CLC con cinco niveles y mayor resolución, no disponible con generalidad para todo el territorio (IGN 2002), pero que ha sido utilizada en el desarrollo del Atlas Urbano Europeo (*Urban Atlas*).

CUADRO A.1.1: Coberturas de CORINE Land Cover al nivel 3

1. Superficies artificiales	
1.1.	Zonas urbanas
1.1.1.	Tejido urbano continuo
1.1.2.	Tejido urbano discontinuo
	Se incluyen: Estructura urbana abierta
	Urbanizaciones exentas y/o ajardinadas
1.2.	Zonas industriales, comerciales y de transportes
1.2.1.	Zonas industriales o comerciales
	Se incluyen: Zonas industriales
	Grandes superficies de equipamientos y servicios
1.2.2.	Redes viarias, ferroviarias y terrenos asociados
	Se incluyen: Autopistas, autovías y terrenos asociados
	Complejos ferroviarios
1.2.3.	Zonas portuarias
1.2.4.	Aeropuertos
1.3.	Zonas de extracción minera, vertederos y de construcción
1.3.1.	Zonas de extracción minera
1.3.2.	Escombreras y vertederos
1.3.3.	Zonas en construcción
1.4.	Zonas verdes artificiales, no agrícolas
1.4.1.	Zonas verdes urbanas
1.4.2.	Instalaciones deportivas y recreativas
	Se incluyen: Campos de golf
	Resto de instalaciones deportivas y recreativas
2. Zonas agrícolas	
2.1.	Tierras de labor
2.1.1.	Tierras de labor en secano
2.2.2.	Terrenos regados permanentemente
	Se incluyen: Cultivos herbáceos en regadío
	Otras zonas de irrigación
2.2.3.	Arrozales
2.2.	Cultivos permanentes
2.2.1.	Viñedos
	Se incluyen: Viñedos en secano
	Viñedos en regadío
2.2.2.	Frutales
	Se incluyen: Frutales en secano
	Frutales en regadío, y dentro de ellos:
	Cítricos
	Frutales tropicales
	Otros frutales en regadío
2.2.3.	Olivares
	Se incluyen: Olivares en secano
	Olivares en regadío
2.3.	Prados y praderas
2.3.1.	Prados y praderas

CUADRO A.1.1 (cont.): Coberturas de CORINE Land Cover al nivel 3

2.4.	Zonas agrícolas heterogéneas
2.4.1.	Cultivos anuales asociados con cultivos permanentes
Se incluyen:	Cultivos anuales asociados con cultivos permanentes en secano
	Cultivos anuales asociados con cultivos permanentes en regadío
2.4.2.	Mosaico de cultivos
Se incluyen:	Mosaico de cultivos en secano, y dentro de ellos:
	Mosaico de cultivos anuales con prados o praderas en secano
	Mosaico de cultivos permanentes en secano
	Mosaico de cultivos anuales con cultivos permanentes en secano
	Mosaico de cultivos en regadío, y dentro de ellos:
	Mosaico de cultivos anuales con prados o praderas en regadío
	Mosaico de cultivos permanentes en regadío
	Mosaico de cultivos anuales con cultivos permanentes en regadío
	Mosaico de cultivos mixtos en secano y regadío
2.4.3.	Terrenos principalmente agrícolas, pero con importantes espacios de vegetación natural y seminatural
Se incluyen:	Mosaico de cultivos agrícolas en secano con espacios significativos de vegetación natural y seminatural
	Mosaico de cultivos agrícolas en regadío con espacios significativos de vegetación natural y seminatural
	Mosaico de prados o praderas con espacios significativos de vegetación natural y seminatural
2.4.4.	Sistemas agroforestales
Se incluyen:	Pastizales, prados o praderas con arbolado adhesionado
	Cultivos agrícolas con arbolado adhesionado
3.	Zonas forestales con vegetación natural y espacios abiertos
3.1.	Bosques
3.1.1.	Bosques de frondosas
Se incluyen:	Perennifolias
	Caducifolias y marcescentes
	Otras frondosas de plantación
	Mezcla de frondosas
	Bosques de ribera
	Laurisilva macaronésica
3.1.2.	Bosques de coníferas
Se incluyen:	Bosques de coníferas con hojas aciculares
	Bosques de coníferas con hojas de tipo cupresáceo
3.1.3.	Bosque mixto
3.2.	Espacios de vegetación arbustiva y/o herbácea
3.2.1.	Pastizales naturales
Se incluyen:	Pastizales supraforestales, y dentro de ellos:
	Pastizales supraforestales templado-oceánicos, pirenaicos y orocantábricos
	Pastizales supraforestales mediterráneos
	Otros pastizales, y dentro de ellos:
	Otros pastizales templado oceánicos
	Otros pastizales mediterráneos

CUADRO A.1.1 (cont.): Coberturas de CORINE Land Cover al nivel 3

3.2.2.	Landas y matorrales mesófilos
Se incluyen:	Landas y matorrales en climas húmedos
	Vegetación mesófila
	Fayal-breza macaronésico
3.2.3.	Matorrales esclerófilos
Se incluyen:	Matorrales esclerófilos mediterráneos, y dentro de ellos:
	Grandes formaciones de matorral denso o medianamente denso
	Matorrales subarbustivos o arbustivos muy poco densos
	Matorrales xerófilos macaronésicos
3.2.4.	Matorral boscoso de transición
Se incluyen:	Matorral boscoso de frondosas
	Matorral boscoso de coníferas
	Matorral boscoso de bosque mixto
3.3.	Espacios abiertos con poca o sin vegetación
3.3.1.	Playas, dunas y arenales
Se incluyen:	Playas y dunas
	Ramblas con poca o sin vegetación
3.3.2.	Roquedo
Se incluyen:	Rocas desnudas con fuerte pendiente (acantilados, etc.)
	Afloramientos rocosos y canchales
	Coladas lávicas cuaternarias
3.3.3.	Espacios con vegetación escasa
Se incluyen:	Xeroestepa subdesértica
	Cárcavas y/o zonas en proceso de erosión
	Espacios orófilos altitudinales con vegetación escasa
3.3.4.	Zonas quemadas
3.3.5.	Glaciares y nieves permanentes
4.	Zonas húmedas
4.1.	Zonas húmedas continentales
4.1.1.	Humedales y zonas pantanosas
4.1.2.	Turberas y prados turbosos
4.2.	Zonas húmedas litorales
4.2.1.	Marismas
4.2.2.	Salinas
4.2.3.	Zonas llanas intermareales
5.	Superficies de agua
5.1.	Aguas continentales
5.1.1.	Cursos de agua
Se incluyen:	Ríos y cauces naturales
	Canales artificiales
5.1.2.	Láminas de agua
Se incluyen:	Lagos y lagunas
	Embalses
5.2.	Aguas marinas
5.2.1.	Lagunas costeras
5.2.2.	Estuarios
5.2.3.	Mares y océanos

Fuente: IGN (2010a).

El fichero de trabajo consta de 155.770 polígonos (155.732 descontando los 38 de la clase 5.2.3. Mares y océanos), y ofrece la superficie por polígono (en ha). El tamaño medio por polígono resulta ser de 3,3 km², pero la dispersión en tamaños es extremadamente elevada, con una desviación típi-

ca de 42 km², un tamaño mínimo de tan solo 0,53 áreas y un tamaño máximo de 7.695 km².

La superficie por clase y su estructura porcentual para toda España se muestran en el cuadro A.1.2 en km². Dada la proyección y huso

CUADRO A.1.2: Superficies de las coberturas de CORINE Land Cover 2006 al nivel 3			
Coberturas CORINE Land Cover - Nivel 3		Km²	Porcentaje
1. Superficies artificiales		10.195	2,00
1.1. Zonas urbanas		6.340	1,25
1.1.1. Tejido urbano continuo		2.937	0,58
1.1.2. Tejido urbano discontinuo		3.403	0,67
1.2. Zonas industriales, comerciales y de transportes		1.979	0,39
1.2.1. Zonas industriales o comerciales		1.468	0,29
1.2.2. Redes viarias, ferroviarias y terrenos asociados		212	0,04
1.2.3. Zonas portuarias		108	0,02
1.2.4. Aeropuertos		191	0,04
1.3. Zonas de extracción minera, vertederos y de construcción		1.528	0,30
1.3.1. Zonas de extracción minera		756	0,15
1.3.2. Escombreras y vertederos		59	0,01
1.3.3. Zonas en construcción		713	0,14
1.4. Zonas verdes artificiales, no agrícolas		348	0,07
1.4.1. Zonas verdes urbanas		67	0,01
1.4.2. Instalaciones deportivas y recreativas		282	0,06
2. Zonas agrícolas		254.042	49,94
2.1. Tierras de labor		121.124	23,81
2.1.1. Tierras de labor en seco		97.637	19,19
2.1.2. Terrenos regados permanentemente		22.036	4,33
2.1.3. Arrozales		1.451	0,29
2.2. Cultivos permanentes		36.001	7,08
2.2.1. Viñedos		8.394	1,65
2.2.2. Frutales		8.930	1,76
2.2.3. Olivares		18.677	3,67
2.3. Prados y praderas		6.515	1,28
2.3.1. Prados y praderas		6.515	1,28
2.4. Zonas agrícolas heterogéneas		90.403	17,77
2.4.1. Cultivos anuales asociados con cultivos permanentes		1.427	0,28
2.4.2. Mosaico de cultivos		38.863	7,64
2.4.3. Terrenos principalmente agrícolas, pero con importantes espacios de vegetación natural y seminatural		25.078	4,93
2.4.4. Sistemas agroforestales		25.036	4,92
3. Zonas forestales con vegetación natural y espacios abiertos		240.017	47,18
3.1. Bosques		91.796	18,04
3.1.1. Bosques de frondosas		37.779	7,43
3.1.2. Bosques de coníferas		38.913	7,65
3.1.3. Bosque mixto		15.103	2,97

CUADRO A.1.2 (cont.): Superficies de las coberturas de CORINE Land Cover 2006 al nivel 3			
Coberturas CORINE Land Cover - Nivel 3		Km²	Porcentaje
3.2.	Espacios de vegetación arbustiva y/o herbácea	135.389	26,61
3.2.1.	Pastizales naturales	26.716	5,25
3.2.2.	Landas y matorrales mesófilos	9.452	1,86
3.2.3.	Matorrales esclerófilos	52.387	10,30
3.2.4.	Matorral boscoso de transición	46.834	9,21
3.3.	Espacios abiertos con poca vegetación o sin ella	12.832	2,52
3.3.1.	Playas, dunas y arenales	481	0,09
3.3.2.	Roquedo	2.418	0,48
3.3.3.	Espacios con vegetación escasa	9.368	1,84
3.3.4.	Zonas quemadas	554	0,11
3.3.5.	Glaciares y nieves permanentes	11	0,00
4.	Zonas húmedas	1.114	0,22
4.1.	Zonas húmedas continentales	550	0,11
4.1.1.	Humedales y zonas pantanosas	544	0,11
4.1.2.	Turberas y prados turbosos	6	0,00
4.2.	Zonas húmedas litorales	565	0,11
4.2.1.	Marismas	298	0,06
4.2.2.	Salinas	198	0,04
4.2.3.	Zonas llanas intermareales	69	0,01
5.	Superficies de agua	5.973	0,66
5.1.	Aguas continentales	3.000	0,59
5.1.1.	Cursos de agua	492	0,10
5.1.2.	Láminas de agua	2.508	0,49
5.2.	Aguas marinas	2.972	0,07
5.2.1.	Lagunas costeras	210	0,04
5.2.2.	Estuarios	137	0,03
5.2.3.	Mares y océanos	2.625	-
Total		511.340	-
Superficie total de España		508.715	100,00

Nota: La superficie constaba como campo (en ha), para cada polígono, en el fichero descargado. La superficie total de España no considera la cobertura 5.2.3. Mares y océanos. La estructura porcentual no tiene en cuenta dicha cobertura y se calcula sobre la superficie total de España.

Fuente: IGN (2010c).

geográfico del fichero de trabajo es necesario tener en cuenta que la proyección UTM no conserva superficies, y que no todas las superficies de este trabajo utilizan el mismo sistema de proyección y/o huso geográfico. El apéndice 3 comenta brevemente estas cuestiones que están frecuentemente ausentes, tanto en trabajos que incorporan información geográfica, como incluso en documentos de carácter más técnico (IGN 2007a).

A.2. EL MODELO DE DATOS DEL SIOSE: COBERTURAS SIMPLES, COMPUESTAS Y ATRIBUTOS

A pesar del gran avance que ha constituido CLC como primera base de datos de ocupación del suelo de ámbito europeo y duración prolongada en el tiempo, permitiendo estudiar la evolución de la cubierta terrestre europea en los últimos

20 años, su falta de resolución en algunos aspectos clave hizo pronto patente la necesidad de una mayor información que la proporcionada por CLC, al menos a escala nacional. Por esta razón en 2005 la Dirección General del Instituto Geográfico Nacional (IGN) puso en marcha el denominado Sistema de Información de Ocupación del Suelo de España (SIOSE), con el objeto de generar una base de datos de ocupación y uso del suelo para toda España a escala 1:25.000 con imágenes de referencia correspondientes a 2005 (SIOSE2005). Aunque los datos de ocupación han sido ampliamente desarrollados en dicha base de datos, la parte de usos está actualmente todavía en elaboración.

El SIOSE se enmarca dentro del Plan Nacional de Observación del Territorio en España (PNOT), que dirige y coordina el IGN/CNIG. Para la producción del SIOSE se dividió el territorio nacional en áreas con una ocupación del suelo homogénea en cuanto a cobertura del suelo se refiere, siguiendo un complejo modelo de datos previamente establecido, y que ha sido descrito de forma somera en el capítulo 5 de este informe. Estas áreas se representaron gráficamente mediante polígonos, a partir de la fotointerpretación de las imágenes del satélite Spot5 para el año 2005, que constituye la referencia geométrica y temporal de la primera versión de la base de datos (SIOSE2005). No obstante, se incorporaron al proyecto multitud de bases de datos, incluyendo las ortofotografías del Plan Nacional de Ortofotografía Aérea (PNOA), imágenes del satélite Landsat y otras bases de datos de ocupación del suelo de mayor escala aportadas por las comunidades autónomas en un sistema de integración multiescala. Adicionalmente se realizó un amplio trabajo de campo con el objetivo de comprobar la realidad existente en el terreno, y que ha dado lugar a una base de datos de fotografías georreferenciadas y uniformemente distribuidas por todo el

territorio nacional, que pueden visualizarse en el visor de imágenes IBERPIX del IGN.

En la actualidad solo está disponible el SIOSE2005, que es la base de datos utilizada en este estudio, estando ya en producción el SIOSE2009 y previendo una nueva actualización a 2012 coincidente con la producción de CLC.

Los datos del SIOSE2005 utilizados en este informe fueron suministrados por el CNIG, mediante descarga *ftp* tras su solicitud⁵³. La distribución es por comunidades autónomas, excepto Andalucía que se distribuye en cuatro fragmentos, y el formato es GeoDataBase (GDB) personal compatible con ArcGIS de ESRI, es decir, una base de datos Access con la geometría de polígonos integrada en formato vectorial⁵⁴. Cada comunidad autónoma incorpora un fichero adicional de metadatos en formato *xml*. Adicionalmente se suministra una base de datos Access que contiene, a nivel nacional, diferentes tablas planas con información de las diferentes coberturas, y que no incorpora la geometría. Su finalidad es la realización de consultas sencillas sobre coberturas del suelo que no requieran salida gráfica. El sistema de referencia geodésico es ETRS89, excepto Canarias donde se utiliza el WGS84 (*World Geodetic System 1984*), y la proyección es UTM en el huso correspondiente, es decir, 28N para Canarias; 29N, para Galicia; 31N, para Cataluña e Illes Balears; y 30N, para el resto de España.

Las principales características técnicas del SIOSE2005 son las siguientes:

- *Escala cartográfica*: 1:25.000.
- *Unidad mínima cartografiable* (MMU):
 - 0,5 ha: Cultivos forzados, coberturas húmedas, playas, vegetación de ribera y acantilados marinos.
 - 1 ha: Zonas urbanas y láminas de agua.

⁵³ Desde aquí queremos agradecer al equipo del SIOSE, en particular a M.^a Elena Caballero y Julián Delgado, su disponibilidad a suministrarnos la información provisional de la Comunitat Valenciana, antes de que la base de datos estuviera finalizada, lo que nos permitió familiarizarnos con esta compleja estructura de organizar la

información sobre coberturas del suelo. La información del SIOSE2005 puede ser descargada gratuitamente desde el Centro de Descargas del CNIG.

⁵⁴ El *software* GIS compatible con el SIOSE se describe en SIOSE (2011b).

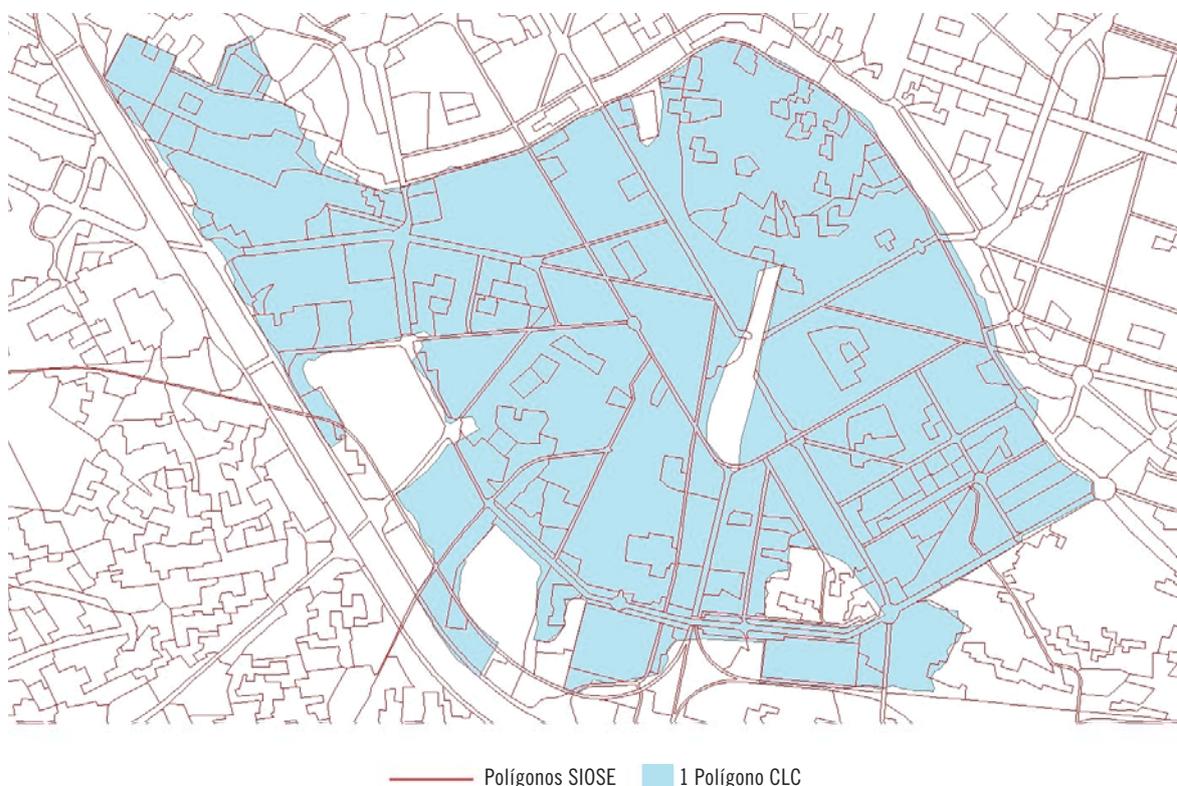
- 2 ha: Zonas agrícolas, forestales y naturales.
- *Anchura mínima de elementos lineales*: 15 m.
- *Modelo de datos*: Orientado a objetos: 40 clases simples y 46 clases compuestas predefinidas.
- *Sistema geodésico de referencia*: ETRS89, excepto Canarias que utiliza WGS84.
- *Sistema cartográfico de proyección*: UTM, huso correspondiente a cada comunidad autónoma (Canarias H28, Galicia H29, Cataluña e Illes Balears H31 y el resto de la Península, junto con Ceuta y Melilla, H30).

El mapa A.2.1 ofrece una imagen visual de la diferente resolución geométrica del SIOSE frente a CLC.

La característica más innovadora del SIOSE es su modelo de datos orientado a objetos, es decir, frente al modelo jerárquico tradicional seguido por CLC de asignación de una única cobertura a

todo un polígono en función de la cobertura dominante, el SIOSE se basa en un modelo de datos con filosofía de orientación a objetos. Partiendo de dos entidades fundamentales: el polígono y la cobertura asociada al mismo, el SIOSE describe las coberturas de cada polígono a partir de una clasificación inicial de coberturas simples, y mediante una serie de parámetros, atributos, relaciones y reglas de consistencia (ETN SIOSE 2011a, 2011c, 2011d, 2011e).

De esta forma, el SIOSE no se limita a una mera clasificación de polígonos en función de una única nomenclatura, por ejemplo, la que aparece en el cuadro A.1.1 relativa a CLC, sino que posibilita la asignación de una o múltiples coberturas, incluyendo coberturas compuestas, susceptibles de anidación, para un único polígono, mediante porcentajes de ocupación y atributos, y todo ello sujeto a unas reglas precisas. Por tanto, el SIOSE ofrece una información mucho más compleja que las clasificaciones temá-



Mapa A.2.1 Comparación de la resolución geométrica entre SIOSE2005 y CLC2006

Fuente: Elaboración propia a partir de CLC2006 (ING 2010c) y SIOSE2005 (ETN SIOSE 2012f).

ticas tradicionales a las que estamos acostumbrados, pero a su vez es mucho más versátil, ya que puede ser adaptada a las necesidades del propio investigador⁵⁵.

El cuadro A.2.1 ofrece el listado de coberturas simples y sus agrupaciones. Solo las coberturas que tienen etiqueta definida son coberturas simples SIOSE2005, 40 en total. Se ofrece

CUADRO A.2.1: Coberturas simples del SIOSE2005					
Código	Nombre	Etiqueta	Km²	Porcentaje	
100	Cobertura artificial		18.702	3,70	
101	Edificación	EDF	5.651	1,12	
102	Zona verde artificial y arbolado urbano	ZAU	1.633	0,32	
103	Lámina de agua artificial	LAA	731	0,14	
104	Vial, aparcamiento o zona peatonal sin vegetación	VAP	4.626	0,91	
111	Otras construcciones	OCT	817	0,16	
121	Suelo no edificado	SNE	3.921	0,77	
131	Zonas de extracción o vertido	ZEV	1.323	0,26	
200	Cultivos		190.622	37,67	
210	Cultivos herbáceos		133.998	26,48	
211	Arroz	CHA	1.338	0,26	
212	Cultivos herbáceos distintos de arroz	CHL	132.660	26,22	
220	Cultivos leñosos		49.720	9,83	
221	Frutales		13.458	2,66	
222	Frutales cítricos	LFC	3.404	0,67	
223	Frutales no cítricos	LFN	10.054	1,99	
231	Viñedo	LVI	10.771	2,13	
232	Olivar	LOL	24.330	4,81	
241	Otros cultivos leñosos	LOC	1.161	0,23	
290	Prados	PRD	6.904	1,36	
300	Pastizal	PST	89.716	17,73	
310	Arbolado forestal		91.779	18,14	
311	Fronosas		50.872	10,05	
312	Fronosas caducifolias	FDC	23.012	4,55	
313	Fronosas perennifolias	FDP	27.860	5,51	
316	Coníferas	CNF	40.907	8,08	
320	Matorral	MTR	85.455	16,89	
330	Terrenos sin vegetación		25.302	5,00	
331	Playas, dunas y arenales	PDA	354	0,07	
333	Suelo desnudo	SDN	15.764	3,12	

⁵⁵ El documento resumen sobre el SIOSE2005 del IGN (2011a) ofrece información sencilla sobre explotaciones alternativas de la información contenida en el SIOSE,

incluyendo la nomenclatura diseñada para la visualización de los datos del SIOSE en el visor de cartografía IBERPIX del IGN.

CUADRO A.2.1 (cont.): Coberturas simples del SIOSE2005				
Código	Nombre	Etiqueta	Km²	Porcentaje
334	Zonas quemadas	ZQM	873	0,17
335	Glaciares y nieves permanentes	GNP	5	0,00
336	Ramblas	RMB	378	0,07
350	Roquedo		7.928	1,57
351	Acantilados marinos	ACM	162	0,03
352	Afloramientos rocosos y roquedos	ARR	6.706	1,33
353	Canchales	CCH	698	0,14
354	Coladas lávicas cuaternarias	CLC	362	0,07
400	Coberturas húmedas		994	0,20
410	Humedales continentales		110	0,02
411	Zonas pantanosas	HPA	74	0,01
412	Turberas	HTU	25	0,00
413	Salinas continentales	HSA	11	0,00
420	Humedales marinos		884	0,17
421	Marismas	HMA	740	0,15
422	Salinas marinas	HSM	144	0,03
500	Cobertura de agua		3.470	0,69
510	Aguas continentales		3.312	0,65
511	Cursos de agua	ACU	874	0,17
512	Láminas de agua		2.438	0,48
513	Lagos y lagunas	ALG	217	0,04
514	Embalses	AEM	2.221	0,44
520	Aguas marinas		158	0,03
521	Lagunas costeras	ALC	85	0,02
522	Estuarios	AES	73	0,01
523	Mares y océanos	AMO	0	0,00
Superficie total de España			506.040	100,00

Fuente: ETN SIOSE (2011e) e IGN (2011a).

también la superficie de cada cobertura simple y su distribución porcentual; por definición la suma de estas superficies proporciona la superficie total de España.

El cuadro A.2.2 ofrece las coberturas compuestas predefinidas dentro del modelo de datos SIOSE2005. Se trata de un tipo especial de coberturas a las que se les ha asignado un nombre identificativo por su representatividad en el territorio, no es pues una clasificación cerrada, y otros países podrían identificar otras cobertu-

ras compuestas predefinidas en función de sus peculiares características. Al igual que en el cuadro A.2.1 solo las coberturas con etiqueta definida son coberturas compuestas predefinidas SIOSE2005. Destaca la gran desagregación prestada al artificial compuesto. Se ofrece igualmente la superficie de cada cobertura compuesta predefinida y su distribución porcentual. Por definición estas superficies no tienen por qué sumar la superficie de España, al contrario de lo que sucede con las coberturas simples.

CUADRO A.2.2: Coberturas compuestas predefinidas del SIOSE2005				
Código	Nombre	Etiqueta	Km ²	Porcentaje
701	Dehesa	DHS	24.373	58,97
702	Olivar viñedo	OVD	247	0,60
703	Asentamiento agrícola residencial	AAR	1.765	4,27
704	Huerta familiar	UER	393	0,95
800	Artificial compuesto		14.556	35,22
810	Urbano mixto		6.323	15,30
811	Casco	UCS	1.629	3,94
812	Ensanche	UEN	2.643	6,39
813	Discontinuo	UDS	2.051	4,96
820	Industrial		1.799	4,35
821	Polígono industrial ordenado	IPO	949	2,30
822	Polígono industrial sin ordenar	IPS	330	0,80
823	Industria aislada	IAS	520	1,26
830	Primario		1.617	3,91
831	Agrícola/Ganadero	PAG	727	1,76
832	Forestal	PFT	22	0,05
833	Minero extractivo	PMX	859	2,08
834	Piscifactoría	PPS	9	0,02
840	Terciario		266	0,64
841	Comercial y oficinas	TCO	144	0,35
842	Complejo hotelero	TCH	51	0,12
843	Parque recreativo	TPR	27	0,07
844	Camping	TCG	44	0,11
850	Equipamiento/Dotacional		1.238	3,00
851	Administrativo institucional	EAI	128	0,31
852	Sanitario	ESN	43	0,10
853	Cementerio	ECM	54	0,13
854	Educación	EDU	174	0,42
855	Penitenciario	EPN	11	0,03
856	Religioso	ERG	37	0,09
857	Cultural	ECL	47	0,11
858	Deportivo	EDP	280	0,68
859	Campo de golf	ECG	230	0,56
860	Parque urbano	EPU	234	0,57

CUADRO A.2.2 (cont.): Coberturas compuestas predefinidas del SIOSE2005				
Código	Nombre	Etiqueta	Km ²	Porcentaje
870	Infraestructuras		3.313	8,02
880	Transporte		2.815	6,81
881	Red viaria	NRV	2.256	5,46
882	Red ferroviaria	NRF	337	0,82
883	Portuario	NPO	68	0,16
884	Aeroportuario	NAP	154	0,37
890	Energía		137	0,33
891	Eólica	NEO	69	0,17
892	Solar	NSL	7	0,02
893	Nuclear	NCL	8	0,02
894	Eléctrica	NEL	21	0,05
895	Térmica	NTM	18	0,04
896	Hidroeléctrica	NHD	7	0,02
897	Gaseoducto/Oleoducto	NGO	7	0,02
900	Telecomunicaciones	NTC	5	0,01
910	Suministro de agua		280	0,68
911	Depuradoras y potabilizadoras	NDP	35	0,08
913	Desalinizadoras	NDS	242	0,59
912	Conducciones y canales	NCC	3	0,01
920	Residuos		76	0,18
921	Vertederos y escombreras	NVE	63	0,15
922	Plantas de tratamiento	NPT	13	0,03
Superficie total			41.334	100,00

Fuente: ETN SIOSE (2011e) e IGN (2011a).

A pesar de su similar nombre, la superficie de *artificial compuesto* (código 800) no tiene por qué corresponderse con la superficie de *coberturas artificiales* (código 100), de hecho esta última muestra una superficie notablemente superior a la primera. La razón es sencilla una vez familiarizados con el modelo de datos SIOSE2005, las coberturas compuestas de *asentamiento agrícola residencial* (código 703) y *huerta familiar* (código 704) deben incluir obligatoriamente *edificaciones* (código 101), que es una cobertura simple

del grupo de *coberturas artificiales* (código 100) (ETN SIOSE 2011c, 2011d).

Finalmente, el cuadro A.2.3 ofrece los atributos existentes dentro del SIOSE2005, y el tipo de los mismos. Estos atributos pueden ir asociados a coberturas simples, y es posible encontrar en ellas más de uno. Resulta de interés observar como los atributos aportan información adicional sobre las coberturas, pero no sobre los usos; por ejemplo, en el atributo de *tipo de edificación*,

CUADRO A.2.3: Atributos de las coberturas simples del SIOSE2005			
Código	Tipo	Nombre	Etiqueta
10	Distribución espacial		
11		Asociación	A
12		Mosaico regular	M
13		Mosaico regular	I
20	Tipo edificación		
21		Edificio aislado	ea
22		Edificio entre medianeras	em
23		Vivienda unifamiliar aislada	va
24		Vivienda unifamiliar adosada	vd
25		Nave	nv
28	En construcción	En construcción	ec
30	Irrigación		
31		Secano	sc
32		Regadío regado	rr
33		Regadío no regado	rn
35	Abanclado	Abanclado	ab
36	Es forzado	Forzado	fz
40	Plantación	Plantación	pl
41	Formación de ribera	Formación de ribera	fr
44	Función de cortafuegos	Función de cortafuegos	fc
45	Cortas	Cortas	ct
46	Procedencia de cultivo	Procedencia de cultivos	pc
47	Alta montaña	Alta montaña	am
48	Es roturado no agrícola	Roturado no agrícola	ra
49	Es zona erosionada	Zonas erosionadas	ze
50	Es cuaternaria	Cuaternarias	cu

Fuente: ETN SIOSE (2011e).

se indica si se trata de un *edificio aislado* o una *vivienda unifamiliar adosada*, pero no informa de si el uso dado a la edificación es residencial o comercial, lo que podría ser de importancia en muchos ámbitos.

El concepto de *uso* hace referencia al tipo de actividad socioeconómica y a las particularidades legales que se dan en el polígono de referencia. El modelo de datos sobre usos está actualmente en desarrollo en ETN SIOSE (2011e, 12). En muchos casos el uso no es directamente deducible mediante fotointerpretación, requiriendo

información adicional, así como convenciones y acuerdos en aras de la comparabilidad en el tiempo y en el espacio. Una base de datos sobre usos del suelo, más allá de las coberturas físicas, es esencial para pasar de una foto fija, como la que hemos dibujado en este estudio, a modelos dinámicos de movilidad de la población (Martin 2010; Bright y Bhaduri 2010; Hofstetter 2011).

Las superficies de los cuadros A.2.1 y A.2.2 se han tomado del IGN (2011a) y han sido obtenidas en el sistema geodésico, proyección y huso

geográfico en el que las bases de datos fueron generadas. Al igual que en el caso de CLC debe hacerse notar que la proyección UTM no conserva superficies.

La base de datos completa ocupa 13,7 GB y consta de 2.477.593 polígonos, por lo que el tamaño medio por polígono es de 0,2 km². Frente a la resolución temática de las 44 clases de CLC, el número de coberturas diferentes que aparecen en el SIOSE2005 es de 820.632; entendiendo por coberturas diferentes combinaciones distintas de clases, es decir, existen en el SIOSE2005, 820.632 polígonos diferentes.

A.3. SISTEMAS DE REFERENCIA GEOGRÁFICA EN LA UE. INSPIRE

El proceso de armonización y convergencia con Europa abarca multitud de aspectos, muchos de los cuales no son perceptibles directamente por el ciudadano. Al contrario de lo que sucede con la utilización de una moneda común, el euro, cuyos efectos son claramente palpables en la vida cotidiana, la armonización estadística llevada a cabo bajo las directrices de Eurostat, no es algo que tenga una incidencia directa sobre los ciudadanos, pero esta homogeneización en el tratamiento de la información es absolutamente esencial para progresar de forma coordinada. La percepción de la realidad depende de lo que medimos y cómo lo medimos, y para compararnos y poner en marcha políticas comunes es absolutamente esencial que esta medición estadística se realice de forma homogénea.

Con la información geográfica sucede lo mismo que con la estadística, debemos alcanzar un acuerdo sobre el sistema geográfico de referencia que debemos utilizar, si queremos que nuestra medida de la geografía tenga una base común sobre la que compartir la información debemos y, en consecuencia, sea comparable en todo el territorio europeo. Históricamente los diversos países han desarrollado sistemas de referencia geodésicos locales a partir de los cuales construían su cartografía. Estos sistemas representaban bien el territorio de una determinada área, pero no eran comparables con otras zonas.

En España, hasta el RD 1071/2007 de 27 de julio, estuvo vigente el sistema geodésico de referencia regional ED50 (*European Datum 1950*), basado en el elipsoide internacional de Hayford 1924, pero a partir de ese momento se adoptó el sistema de referencia geodésico global europeo, el ETRS89, basado en el elipsoide GRS80 (*Geodetic Reference System 1980*).

De forma similar a lo que ocurre cuando cambiamos la base de la Contabilidad Nacional (CN) o del índice de precios al consumo (IPC), una vez adoptado el nuevo sistema de referencia geodésico es necesario resolver el problema de la migración de toda la cartografía al nuevo sistema, un proceso que no es sencillo y que todavía no ha finalizado en nuestro país (IGN 2007b, 2007c, 2007d, 2007e). Precisamente porque el cambio de un sistema de referencia geográfico a otro es muy costoso en la práctica, los países han sido reacios a dicha migración lo que ha jugado en contra de la estandarización a nivel internacional.

En la actualidad, todavía conviven en nuestro sistema cartográfico los dos sistemas mencionados, el ED50, en proceso de extinción, y el ETRS89. Aunque la cartografía sobre coberturas del suelo, tanto CLC2006 como SIOSE2005, se encontraban ya en ETRS89, el fichero vectorial de secciones censales suministrado por el INE estaba en ED50 y proyección UTM en el huso 30N (WGS84 y proyección UTM en el huso 28N, para Canarias), por lo que hubo que proceder a su transformación al sistema ETRS89 como paso preliminar a cualquier análisis. Siguiendo las recomendaciones del IGN (2007b, 2007c, 2007d) el proceso de transformación de ED50 a ETRS89 se efectuó mediante la aplicación de la rejilla NTV2 (*National Transformation Version 2*), implementada en ArcGIS 9.3. Por su parte, el fichero de líneas de límite municipales del CNIG, a partir de cual se han dibujado los mapas del informe con excepción de la *grid*, tenía Canarias en fichero aparte en WGS84 (compatible con REGCAN95 y ETRS89), por lo que hubo que convertir todas las líneas de límite a un fichero único en ETRS89. En ambos casos, la información original procedente del CNIG estaba en coordenadas geográficas (sin proyección).

Las bases para la homogenización del sistema geográfico de referencia en la Unión Europea (UE) se establecen en el *workshop* sobre Sistemas de Referencia Espacial para Europa que, bajo los auspicios de la Comisión Europea y el Centro Común de Investigación (JRC, *Joint Research Center*), tuvo lugar en Marne-La Vallée (Francia) el 29-30 de noviembre de 1999 (Annoni y Luzet 2000). A este *workshop* inicial seguirían otros dos que han sido fundamentales en la armonización posterior. El primero sobre *Proyecciones Cartográficas para Europa* fue celebrado en Marne-La Vallée el 14-15 de diciembre de 2000 (Annoni *et al.* 2003); y el segundo sobre *Grids de Referencia para Europa*, en Ispra (Italia) el 27-29 de octubre de 2003 (Annoni 2005). Fueron estos encuentros los que sentaron los pilares del sistema geográfico de referencia europeo que todavía hoy está en construcción.

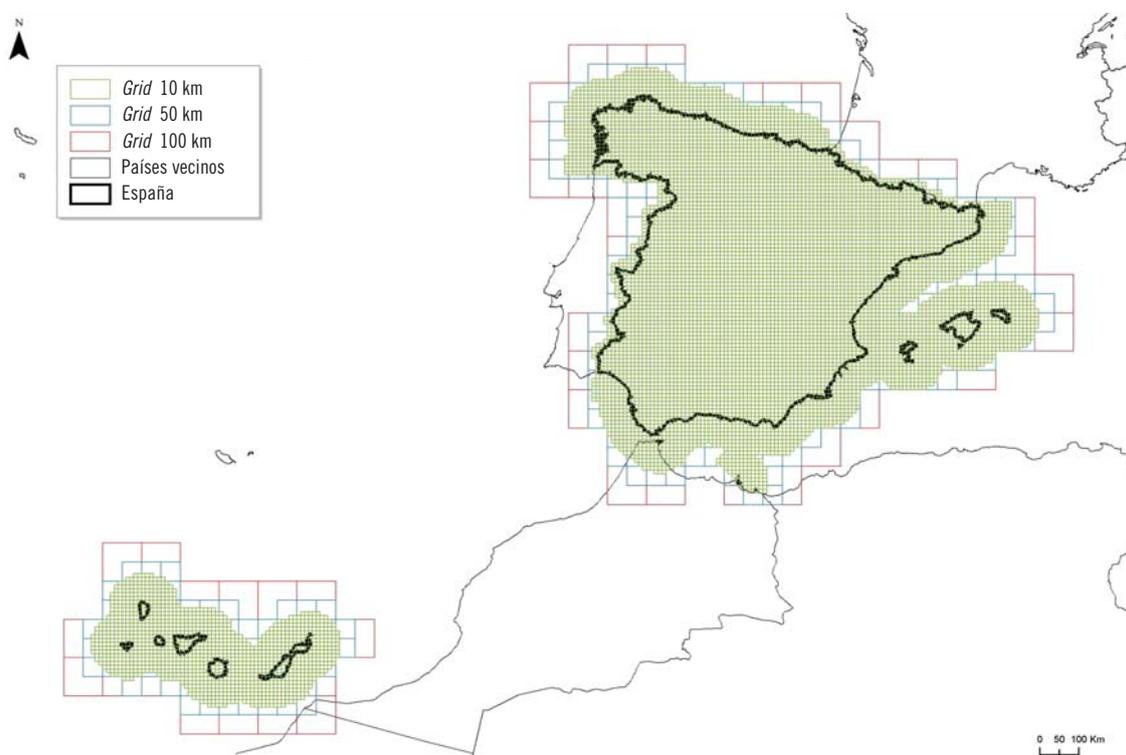
Las recomendaciones de estos *workshops* a la Comisión Europea que nos son de aplicación en este informe fueron las siguientes:

- Adopción del ETRS89 como *datum* geodésico. Expresar y almacenar información geográfica, en tanto en cuanto sea posible, en coordenadas elipsoidales con referencia al elipsoide GRS80.
- Dado que el almacenamiento de información geográfica en coordenadas elipsoidales no es suficiente para multitud de aplicaciones prácticas, para las que se requieren proyecciones cartográficas y sistemas de coordenadas planas, se establecieron tres sistemas de coordenadas cartesianas en función de la aplicación concreta:
 - El sistema de coordenadas Lambert Azimutal Equal Area, ETRS89-LAEA, para análisis estadístico espacial y presentación de resultados cuando se requiera la verdadera área de los polígonos en la representación.
 - El sistema de coordenadas Lambert Conformal Conic, ETRS89-LCC, para cartografía a escala igual o menor a 1:500.000.
 - El sistema de coordenadas Transverse Mercator, ETRS89-TMzn, para cartografía a escala superior a 1:500.000, donde *zn* se refiere al huso o zona en cuestión.

La proyección ETRS89-TMzn es idéntica a la UTM, en realidad esta es una especialización de la primera, por lo que en la práctica utilizamos esta en el huso correspondiente, que abarca 6° de longitud.

Estos trabajos acabaron por traducirse en una directiva del Parlamento Europeo y del Consejo (2007/2/CE) por la que se establece una Infra-estructura de Información Espacial en la Comunidad Europea (INSPIRE, *Infrastructure for Spatial Information in Europe*), que entró en vigor el 15 de mayo de 2007 y está siendo implementada en varias etapas, hasta su completa implementación en 2019. La directiva INSPIRE aspira a crear una *infraestructura de datos espaciales* (IDE) en el ámbito de la UE, con el fin último de intercambiar información geográfica, fundamentalmente de carácter medioambiental, y facilitar el acceso público a la información espacial en Europa. INSPIRE recoge en sus directrices las recomendaciones anteriores sobre el sistema geográfico de referencia, las proyecciones y los sistemas de coordenadas planas que acabamos de mencionar (INSPIRE 2010a, 2010b). Asimismo, ha sido transpuesta al ordenamiento jurídico español mediante la Ley 14/2010 o Ley de las Infraestructuras y Servicios de Información Geográfica en España (LISIGE).

Puesto que para la elaboración de una *grid* de densidad poblacional la representación exacta de la verdadera área es fundamental, la *grid* de este estudio utiliza el sistema de coordenadas cartesianas ETRS89-LAEA, con una resolución de 1 × 1 km. De otra forma la población estimada por celda no representaría una densidad constante por unidad de superficie en el producto final. Esta *grid*, centrada en 10° de longitud este (meridiano central) y 52° de latitud norte por motivos de homogeneidad con Europa, está disponible en formato vectorial *shapefile* en la AEMA (*reference grids*) o en el sitio web del European Forum for Geostatistics (EFGS). Ambas *grids* son idénticas, excepto por la codificación de las celdas, aunque por motivos de comparabilidad con otros trabajos hemos utilizado la del EFGS. Su representación para diversas resoluciones puede verse en el mapa A.3.1.



Mapa A.3.1 Grid de referencia para España en diversas resoluciones

Fuente: AEMA (2012a).

Para entender la importancia de esta proyección en el ejercicio que hemos realizado en este informe debemos tener presente que un sistema de coordenadas cartesiano representa sobre una superficie plana (dos dimensiones) una realidad aproximada por el elipsoide de referencia (tres dimensiones). Esta representación de la superficie terrestre en dos dimensiones causa necesariamente cierta distorsión en la forma, el área, la escala (distancias) y/o los ángulos (direcciones) de la realidad representada. Diferentes proyecciones generan diversos tipos de distorsiones, por lo que unas son más apropiadas que otras en función del uso que se les vaya a dar a los mapas (Kennedy y Kopp 2000).

La proyección LAEA mantiene las áreas y en consecuencia es la apropiada para un ejercicio de densidad de población, o incluso para representar límites político-administrativos si el mantenimiento de la verdadera área es una característica importante que se debe preservar. Este sistema de proyección no es recogido por el RD 1071/2007,

de 27 de julio, que regula el sistema geodésico de referencia oficial en España, que sí contempla, sin embargo, los otros dos sistemas mencionados, ETRS89-LCC y ETRS89-TM. La razón es una preferencia por una mayor precisión en la forma representada que en el área, ya que en ambos casos se hace énfasis en la representación cartográfica, más que en el análisis espacial. La proyección UTM es muy precisa en la representación de pequeñas superficies y, dentro del huso correspondiente, la distorsión es mínima para grandes superficies. También son mínimas, sin embargo, las distorsiones en el área dentro de cada huso. Por estas razones, la proyección UTM ha ganado una tremenda popularidad en la representación cartográfica, siendo probablemente la más utilizada en grandes escalas. Sin embargo, esta proyección no mantiene, por construcción, el área constante en todo el mapa, por lo que una *grid* de población en UTM no sería consistente en la representación de los habitantes por unidad de superficie, ya que el área de las celdas variaría a lo largo de la representación cartográfica.

La razón para proponer dos proyecciones conformes⁵⁶ alternativas para la representación cartográfica radica en la diferente escala de representación. En UTM la distorsión, incluso la de la forma, crece cuando nos alejamos del meridiano central del huso correspondiente. Para representar grandes áreas necesitamos una escala pequeña, y en ocasiones abarcamos varios husos de la proyección UTM. En estas condiciones este tipo de proyección puede presentar distorsiones importantes, incluso en la forma, por lo que para una escala igual o menor a 1:500.000 se propone una proyección alternativa, LCC.

Las superficies que aparecen en la cartografía de la *grid* del capítulo 6 se han obtenido de la cartografía manejada una vez proyectada a ETRS89-LAEA.

Ya hemos comentado al mencionar las superficies de CLC2006 y SIOSE2005 en los apéndices anteriores que estas se habían obtenido a partir de la proyección UTM y que esta no conserva superficies. En general, se considera que las discrepancias que pudieran encontrarse respecto a una proyección que mantenga las áreas son despreciables al nivel de precisión con el que trabajamos (IGN 2011a, 12); sin embargo, resulta ilustrativo examinar brevemente esta cuestión, ya que en el caso de CLC2006, la superficie se obtuvo a partir de la proyección UTM en el huso 30N extendido, es decir, extendiendo la proyección de dicho huso a zonas que caen fuera de él (Canarias H28, Galicia H29 y Cataluña e Illes Balears H31). El cuadro A.3.1 ofrece la superficie de España y las provincias de Canarias sumi-

nistrada por el INE en su sección de Territorio⁵⁷, y las obtenidas a partir de diversas proyecciones del fichero de líneas de límite municipales del CNIG. Tomando como base la cifra del INE, la proyección UTM en el huso 30N sobrestima la superficie en unos 500 km², lo que representa alrededor de un 1%; sin embargo, más de la mitad de este *error* se debe a las provincias de Canarias, que al estar en el huso 28N se encuentran relativamente alejadas del meridiano central del huso 30N. Si manteniendo la proyección UTM re proyectamos el fichero en el huso 28N, entonces los *errores* a nivel nacional se traducen en una sobrestimación considerable de la superficie nacional, pero las provincias de Canarias prácticamente no sufren distorsión respecto a las cifras del INE. Al igual que antes, lo que sucede es que la mayor parte de la Península está muy alejada del meridiano central del huso 28N, lo que provoca una mala estimación de la superficie. Finalmente, si proyectamos el fichero en LAEA entonces obtenemos unas cifras de superficie prácticamente iguales a las ofrecidas por el INE, tanto a nivel nacional como para las provincias de Canarias, ya que esta proyección no deforma las áreas, con lo que evitamos distorsiones entre las diferentes partes del territorio que se van a representar⁵⁸.

Por esta razón, las áreas de la *grid* son de elaboración propia después de proyectar la cartografía a ETRS89-LAEA; las de coberturas del suelo proceden de las propias bases de datos y las cifras de superficie y densidades de población que aparecen en el capítulo 2 proceden del INE, sección de Territorio, para el año 2010.

⁵⁶ Una proyección se dice que es conforme cuando mantiene las formas en una cierta área local.

⁵⁷ Dicha superficie es la ofrecida para el 2010, ya que el INE no es consistente a lo largo del tiempo en las cifras que ofrece como superficie total del territorio nacional. Así, para 2009 indica una superficie de 505.938,13 km² y para 2008, de 505.991,06 km². Son diferencias pequeñas, pero llamativas para una característica fija del territorio. La menor superficie en 2009 puede rastrearse hasta la provincia de Badajoz, por lo que no puede deberse a un problema de linderos marinos.

⁵⁸ Las pequeñas diferencias pueden deberse a multitud de causas. No hemos encontrado información sobre el sistema de cálculo de las superficies ofrecidas por el INE, y tampoco hemos encontrado datos de superficie en el sitio web del IGN (excepto para las islas en las Tablas de Datos Geográficos Generales del Atlas Nacional de España), pero la precisión asociada a la escala puede ser un factor relevante (1:25.000 en el caso de las líneas de límite municipales), además del problema de ausencia de un criterio unificado para la delimitación de las líneas de costa (Prendes y Lorenzo 2011).

CUADRO A.3.1: Medidas alternativas de superficie según diversas proyecciones					
(km²)					
Código	Área geográfica	INE	UTM huso 30N	UTM huso 28N	LAEA
	España	505.985,90	506.512,96	518.050,95	506.010,51
35	Palmas, Las	4.065,78	4.194,46	4.063,58	4.065,95
38	Santa Cruz de Tenerife	3.381,17	3.536,72	3.378,32	3.377,45
5	Canarias	7.446,95	7.731,18	7.441,89	7.443,40

Fuente: INE: Sección Territorio 2010. Para el resto: elaboración propia a partir del fichero de líneas de límite municipales del CNIG (mayo 2011), originariamente en coordenadas geográficas, WGS84 para Canarias y ETRS89 para el resto, y proyecciones y cálculos de superficie realizados en ArcGIS 9.3.

Bibliografía

AEMA (Agencia Europea del Medio Ambiente). Portal de AEMA. Disponible en internet: <http://www.eea.europa.eu/es> [consulta: marzo del 2012a].

—. «EEA Fast Track Service Precursor on Land Monitoring - Degree of soil sealing 100m». Disponible en internet: <http://www.eea.europa.eu/data-and-maps/data/eea-fast-track-service-precursor-on-land-monitoring-degree-of-soil-sealing-100m-1> [consulta: marzo del 2012b].

—. «EEA reference grid». Disponible en internet: <http://www.eea.europa.eu/data-and-maps/data/eea-reference-grids-1> [consulta: marzo del 2012c].

ANNONI, Alessandro (ed.). *European Reference Grids*. EUR Report n.º 21494 (2005). [También en *Proceedings* del European Commission-Joint Research Centre Workshop *European Reference Grids*, Ispra (Italia), 27-29 de octubre de 2003]. [Disponible en internet: <http://www.ec-gis.org/sdi/publist/pdfs/annoni2005eurgrids.pdf>].

ANNONI, Alessandro y Claude LUZET (eds.). *Spatial Reference Systems for Europe*. EUR Report n.º 19575 (2000). [También en *Proceedings* del European Commission-Joint Research Centre Workshop *Spatial Reference Systems for Europe*, Marne-La Vallée (Francia), 29-30 de noviembre de 1999]. [Disponible en internet: <http://www.crs-geo.eu/>].

ANNONI, Alessandro, Claude LUZET, Erich GUBLER y Johannes IHDE (eds.). *Maps Projections for Europe*. EUR Report n.º 20120 (2003). [También en *Proceedings* del European Commission-Joint Research Centre Workshop *Maps Projections for Europe*, Marne-La Vallée (Francia), 14-15 de diciembre de 2000]. [Disponible en internet: <http://www.ec-gis.org/sdi/publist/pdfs/annoni-et-al2003eur.pdf>].

ANSELIN, Luc. *Spatial Econometrics: Methods and Models*. *Studies in Operational Regional Science*. Dordrecht/Boston/London: Kluwer Academic Publishers, 1988.

BARBOSA, Paulo, Andrea CAMIA, Jan KUCERA, Giorgio LIBERTA, Ilaria PALUMBO, Jesús SAN-MIGUEL-AYANZ y Guido SCHMUCK. «Assessment of forest fire impacts and emissions in the European Union based on the European Forest Fire Information System». *Developments in Environmental Sciences* 8 (2008): 197-208.

BATISTA E SILVA, Filipe. «The effect of ancillary data in population dasymetric mapping: A test case using the original and a modified version of *CORINE Land Cover*». Ponencia presentada en el *European Forum for Geostatistics* 2011, Lisboa (Portugal), 12-14 de octubre del 2011. Disponible en internet: <http://www.efgs.info/workshops/efgs-2011-lisboa-portugal>.

BHADURI, Budhendra, Edward BRIGHT, Philip COLEMAN y Jerome DOBSON. «LandScan: Locating people is what matters». *Geoinformatics* 5, n.º 2 (2002): 34-37. Disponible en internet: <http://www.ornl.gov/sci/landscan/>.

BHADURI, Budhendra, Edward BRIGHT, Philip COLEMAN y Marie L. URBAN. «LandScan USA: A high-resolution geospatial and temporal modeling approach for population distribution and dynamics». *GeoJournal* 69 (2007): 103-117. Disponible en internet: <http://www.ornl.gov/sci/landscan/>.

BRACKEN, Ian y David MARTIN. «The generation of spatial population distributions from census centroid data». *Environment and Planning A* 21, n.º 4 (1989): 537-543.

—. «Linkage of the 1981 and 1991 UK censuses using surface modelling concepts». *Environment and Planning A* 27, n.º 3 (1995): 379-390.

BRIGHT, Eddie y Budhendra BHADURI. «Characterizing High-Resolution Population Distributions: a LandScan Experience». Ponencia presentada en el *European Forum for Geostatistics* 2010, Tallin (Estonia), 5-7 de octubre del 2010. Disponible en internet: <http://www.efgs.info/workshops/efgs-2010-tallinn-estonia/>.

- BURCHFIELD, Marcy, Henry G. OVERMAN, Diego PUGA y Matthew A. TURNER. «Causes of sprawl: A portrait from space». *Quarterly Journal of Economics* 121, n.º 2 (mayo 2006): 587-633.
- CAPDEVILA I SUBIRANA, Joan. *Historia del deslinde de la frontera Hispano-Francesa. Del tratado de los Pirineos (1659) a los tratados de Bayona (1856-1868)*. Madrid: Centro Nacional de Información Geográfica, Instituto Geográfico Nacional, Ministerio de Fomento, 2009. Disponible en internet: <http://www.ign.es/ign/layout/In/libDigitalesPublicaciones.do>.
- CASADO ALONSO, Hilario. «La economía en las Españas medievales (c. 1000-c. 1450)». En F. Comín, M. Hernández y E. Llopis (eds.). *Historia Económica de España. Siglos X-XX*. Barcelona: Editorial Crítica/Historia del Mundo Moderno (2002):13-50.
- CHASCO YRIGOREN, Coro. «Estimación de la renta familiar disponible municipal y regional de 1997». Documento Metodológico. Madrid: Instituto Lawrence R. Klein, Universidad Autónoma de Madrid, 1997.
- CHAUDHRY, Omair y William A. MACKANESS. «Automatic identification of urban settlement boundaries for multiple representation databases». *Computers, Environment and Urban Systems* 32 (2008): 95-108.
- CHEN, Keping, John McANENEY, Russell BLONG, Roy LEIGH, Laraine HUNTER y Christina MAGILL. «Defining area at risk and its effect in catastrophe loss estimation: A dasymetric mapping approach». *Applied Geography* 24, n.º 2 (abril 2004): 97-117.
- CIESIN (Center for International Earth Science Information Network). *Gridded Population of the World (GPW). Versión 3*. Palisades, NY: CIESIN y Columbia University, 2005. Disponible en internet: <http://sedac.ciesin.org/plue/gpw/>.
- COLLANTES GUTIÉRREZ, Fernando. «La desagrarización de la sociedad rural española, 1950-1991». *Historia Agraria: Revista de Agricultura e Historia Rural* 42 (agosto 2007): 251-276.
- DE COS GUERRA, Olga. «Valoración del método de densidades focales (kernel) para la identificación de los patrones espaciales de crecimiento de la población de España». *GeoFocus* 4 (2004): 136-165.
- DE COS GUERRA Olga y Pedro REQUES VELASCO. «Los cambios en los patrones territoriales de la población española (1900-2001)». *Papeles de Economía Española* 104 (2005): 167-192.
- DENTON, Frank T. «Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization». *Journal of American Statistical Association* 66, n.º 333 (marzo 1971): 99-102.
- DIJKSTRA, Lewis y Hugo POELMAN. «Remote rural regions: How proximity to a city influences the performance of rural regions». *Regional Focus* 1/2008, Bruselas: EC-DG REGIO, 2008.
- DOBSON, Jerome E., Edward A. BRIGHT, Phillip R. COLEMAN, Richard C. DURFEE y Brian A. WORLEY. «LandScan: A global population database for estimating populations at risk». *Photogrammetric Engineering and Remote Sensing* 66, n.º 7 (2000): 849-857.
- DURANTON, Gilles y Henry G. OVERMAN. «Testing for localization using micro-geographic data». *Review of Economic Studies* 72, n.º 4 (octubre 2005): 1077-1106.
- EFGS (European Forum for Geostatistics). Portal de EFGS. Disponible en internet: <http://www.efgs.info/> [consulta: marzo del 2012].
- EICHER, Cory L. y Cynthia A. BREWER. «Dasymetric mapping and areal interpolation: Implementation and evaluation». *Cartography and Geographic Information Science* 28 (2001): 125-138.
- ELBERS, Chris, Jean O. LANJOUW y Peter LANJOUW. «Micro-level estimation of poverty and inequality». *Econometrica* 71, n.º1 (enero 2003): 355-364.
- ELBERS, Chris, Peter LANJOUW, Johan MISTIAEN, Berk ÖZLER y Kenneth SIMLER. «Are neighbours equal? Estimating local inequality in three developing countries». WIDER Discussion Paper n.º 2003/52, Helsinki: World Institute for Development Economics Research, julio 2003.
- ESPON (European Observation Network for Territorial Development and Cohesion). *The Modifiable Areas Unit Problem. Final Report*. ESPON Scientific Support Project n.º 3.4.3. Luxemburgo: European Observation Network for Territorial Development and Cohesion Programme, Comisión Europea, noviembre 2006. Disponible en internet: http://www.espon.eu/main/Menu_Projects/Menu_ESPON2006Projects/Menu_StudiesScientificSupportProjects/maup.html.
- . *Mapping Guide. Cartography in ESPON 2013. Technical Report, ESPON 2013 Database*. Luxemburgo: European Observation Network for Territorial Development and Cohesion Programme, Comisión Europea, 2010. Disponible en internet: http://www.espon.eu/main/Menu_ScientificTools/MappingGuide/.

- Esri. MapaBase 2.0. Esri España Geosistemas, S. A. Disponible en internet <http://www.mapabase.es/mapabase2/> [consulta: marzo del 2012].
- ETN SIOSE (Equipo Técnico Nacional del Sistema de Información de Ocupación del Suelo en España). *Estructura de la Base de Datos SIOSE-MF2*. Versión 2.1, Instituto Geográfico Nacional. Disponible en internet: <http://www.siose.es/siose/> [consulta: 29 de marzo del 2012a].
- . *Ejemplos de salidas gráficas de consultas sobre T_VALORES*. Versión 1.1, Instituto Geográfico Nacional. Disponible en internet: <http://www.siose.es/siose/> [consulta: 29 de marzo del 2012b].
- . *Documento Técnico SIOSE 2005*. Versión 2, Instituto Geográfico Nacional. Disponible en internet: <http://www.siose.es/siose/> [consulta: 1 de febrero del 2012c].
- . *Manual de Fotointerpretación SIOSE*. Versión 2, Instituto Geográfico Nacional. Disponible en internet: <http://www.siose.es/siose/> [consulta: 1 de febrero del 2012d].
- . *Anexo I. Descripción del Modelo de Datos y Rótulo SIOSE2005*. Versión 2, Instituto Geográfico Nacional. Disponible en internet: <http://www.siose.es/siose/> [consulta: 1 de febrero del 2012e].
- . Sistema de Información de Ocupación del Suelo de España (SIOSE). Base de datos disponible en internet: <http://www.siose.es/siose/> [consulta: 1 de febrero del 2012f].
- EUROSTAT. *GISCO Desktop Mapping Guide*. Versión 2. Geographical Information System for the Commission Directorate D-Unit D2 Regional Statistics and Geographical Information, febrero 1999a. Disponible en internet: http://epp.eurostat.ec.europa.eu/portal/page/portal/gisco_Geographical_information_maps/publications/manuals.
- . *Recommendations for a harmonized definition of calculation of surface area of territorial units. Methods and Nomenclature. Theme 1*. Luxemburgo: Comisión Europea, Eurostat, 1999b.
- . *Guidelines for Geographic Data Intended for the GISCO Reference Database*. Witney, Reino Unido: Lov-ell Johns Ltd, november 2005. Disponible en internet: http://epp.eurostat.ec.europa.eu/portal/page/portal/gisco_Geographical_information_maps/publications/manuals.
- . *Regions in the European Union. Nomenclature of Territorial Units for Statistics. NUTS 2006/EU-27*. N.º KS-RA-07-020-EN, Luxemburgo: Comisión Europea, Eurostat, 2007. Disponible en internet: http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction.
- . *Eurostat regional yearbook 2010. Eurostat Statistical Books*. N.º KS-HA-10-001-EN, Luxemburgo: Comisión Europea, Eurostat, 2010. Disponible en internet: http://epp.eurostat.ec.europa.eu/portal/page/portal/product_details/publication?p_product_code=KS-HA-10-001.
- EUROSTAT-Geographical Information System at the Commission (GISCO). *Disaggregated Population by 1 km² grid*. EuroGeographics Association for the Administrative Boundaries Cartography. Luxemburgo: Eurostat-GISCO, 2008. Disponible en internet: <http://epp.eurostat.ec.europa.eu/cache/GISCO/posters/classification%20ESTAT.pdf>.
- FISHER, Peter F. y Mitchel LANGFORD. «Modelling the errors in areal interpolation between zonal systems by Monte Carlo simulation». *Environment and Planning A* 27, n.º 2 (1995): 211-224.
- FLOWERDEW, Robin, Mick GREEN y Evangelos KEHRIS. «Using areal interpolation methods in GIS». *Papers in Regional Science* 70, n.º 3 (1991): 303-315.
- FOTHERINGHAM, A STEWART, Chris BRUNSDON y Martin CHARLTON. *Geographically Weighted Regression. The Analysis of Spatially Varying Relationships*. Chichester (Reino Unido): John Wiley & Sons, Ltd., 2002.
- GALLEGO, F. Javier. «Population density grid of EU-27+, version 5: Summary of the downscaling method». Documento metodológico. Copenhagen: European Environment Agency, 2009. Disponible en internet: <http://www.eea.europa.eu/data-and-maps/data/population-density-disaggregated-with-corine-land-cover-2000-2>.
- . «A population density grid of the European Union». *Population and Environment* 31, n.º 6 (julio 2010): 460-473.
- GALLEGO, F. Javier y Catharina BAMPS. «Using CORINE land cover and the point survey LUCAS for area estimation». *International Journal of Applied Earth Observation and Geoinformation* 10 (2008): 467-475.
- GALLEGO, F. Javier, y Steve PEEDELL. «Using land cover to map population density». En *Towards agri-environmental indicators. Integrating statistical and administrative data with land cover information*. Topic report

n.º 6, Luxemburgo: DG Agriculture, DG Environment, Joint Research Center, European Environment Agency, Eurostat (2001): 94-105.

GARCÍA GONZÁLEZ, Juan Antonio y Francisco CEBRIÁN ABELLÁN. «La interpolación como método de representación cartográfica para la distribución de la población: Aplicación a la provincia de Albacete». Ponencia presentada en el *XII Congreso Nacional de Tecnologías de la Información Geográfica*, Granada, 19-23 de septiembre de 2006.

GOERLICH, Francisco J. e Isidro CANTARINO. «Un índice de rugosidad del terreno a escala municipal a partir de modelos de elevación digital de acceso público». Documento de Trabajo n.º 7/2010. Bilbao: Fundación BBVA, julio 2010a. Disponible en internet: http://www.fbbva.es/TLFU/dat/DT_7_2010.pdf.

—. «A population grid for Spain: Experiences in assembling population and cartographic data from publicly available sources». Ponencia presentada en el *European Forum for Geostatistics 2010*, Tallin (Estonia), 5-7 de octubre del 2010b. Disponible en internet: <http://www.efgs.info/workshops/efgs-2010-tallinn-estonia/>.

—. «Cartografía y Demografía: Una grid de población para la Comunidad Valenciana». *Revista de Estadística Española* 53, n.º 176 (primer cuatrimestre 2011): 127-161. Disponible en internet: <http://www.ine.es/>.

GOERLICH, Francisco J. y Matilde MAS. «Algunas pautas de localización de la población española a lo largo del siglo XX». *Investigaciones Regionales* 12 (primavera 2008a): 5-33.

—. «Empirical evidence of population concentration in Spain». *Population-E* 63, n.º 4 (2008b): 635-650.

—. «Drivers of agglomeration: Geography versus History». *The Open Urban Studies Journal* 2 (2009): 2842. Disponible en internet: <http://www.bentham.org/open/tousj/openaccess2.htm>.

GOERLICH, Francisco J., Matilde MAS, Joaquín AZAGRA y Pilar CHORÉN. *La Localización de la Población Española sobre el Territorio. Un Siglo de Cambios. Un estudio basado en series homogéneas (1900-2001)*. Bilbao: Fundación BBVA, 2006.

GÓMEZ ANDRÉS, Marisol y Francisco Javier SÁNCHEZ MARTÍNEZ. «Sistema Nacional de Cartografía de Zonas Inundables (SNCZI)». Ponencia presentada en la *Conferencia Europea Esri 2011*, MARM/Subdirección General de Gestión Integrada del Dominio Público Hidráulico, Ma-

drid, 26-28 de octubre del 2011. Disponible en internet: <http://ficheros.esri.es/conferencia2011/conferencia/medioambiente/marm-national-mapping-system.pdf>.

GRASLAND, Claude, Hélène MATHIAN y Jean-Marc VINCENT. «Multiple analysis and maps generalization of discrete social phenomena: Statistical problems and political consequences». *Statistical Journal of the United Nations Economic Commission for Europe* 17, n.º 2 (enero 2000): 157-188.

GUTIÉRREZ PUEBLA, Javier, Gabriel GÓMEZ CERDA y Juan Carlos GARCÍA PALOMARES. «Estudio de accesibilidad del PEIT (Plan Estratégico de Infraestructuras y Transportes del Ministerio de Fomento)». Madrid: Departamento del Geografía Humana. Universidad Complutense de Madrid, 2009.

HARVEY, Andrew C. *Forecasting, structural time series models and the Kalman filter*. Cambridge: Cambridge University Press, 1989.

HOFSTETTER, Claudia. «Day Time Population of the Canton of Zurich». Ponencia presentada en el *European Forum for Geostatistics 2011*, Lisboa, 12-14 de octubre del 2011. Disponible en internet: <http://www.efgs.info/workshops/efgs-2011-lisboa-portugal>.

INE (Instituto Nacional de Estadística). *Densidad de población de los municipios españoles. Mapas provinciales. Censos de población y viviendas 1991*. Madrid, 1994.

—. *Proyecto de los censos demográficos 2011*. Madrid: Subdirección General de Estadísticas de la Población, febrero 2011. Disponible en internet: http://www.ine.es/censos2011/censos2011_proyecto.pdf.

—. «Estadísticas demográficas. Padrón municipal del 2010». Madrid, 2010.

—. «Estadísticas demográficas. Sección Territorio». Disponible en internet: <http://www.ine.es/jaxi/menu.do?type=pcaxis&path=%2Ft43%2Fa011&file=inebase&L=0> [consulta: marzo del 2012].

INSPIRE (Infrastructure for Spatial Information in Europe). «D2.8.1.1 INSPIRE Specification on Coordinate Reference Systems: Guidelines». INSPIRE Thematic Working Group Coordinate Reference Systems and Geographical Grid Systems. Versión 3.1 (2010-04-26), Infrastructure for Spatial Information in the European Community. Luxemburgo: Comisión Europea, 2010a. Disponible en internet: <http://inspire.jrc.ec.europa.eu/index.cfm/pageid/2>.

- . «D2.8.1.2 INSPIRE Specification on Geographical Grid Systems: Guidelines». INSPIRE Thematic Working Group Coordinate Reference Systems and Geographical Grid Systems, Version 3.0.1 (2010-04-26), Infrastructure for Spatial Information in the European Community. Luxemburgo: Comisión Europea, 2010b. Disponible en internet: <http://inspire.jrc.ec.europa.eu/index.cfm/pageid/2>.
- IGN (Instituto Geográfico Nacional) *CORINE 2000. Descripción de la Nomenclatura del CORINE Land Cover al nivel 5.º*. Madrid, diciembre 2002.
- . *Necesidad de un nuevo «Datum»*. Versión 1.0. Madrid: Grupo de trabajo para la transición a ETRS89, Consejo Superior Geográfico, octubre 2007a.
- . *Análisis de la problemática del cambio de datum geodésico a ETRS89*. Versión 1.0. Madrid: Grupo de trabajo para la transición a ETRS89, Consejo Superior Geográfico, octubre 2007b.
- . *Análisis de los sistemas de cambio*. Versión 1.0. Madrid: Grupo de trabajo para la transición a ETRS89, Consejo Superior Geográfico, octubre 2007c.
- . *Herramientas para facilitar el cambio*. Versión 1.0. Madrid: Grupo de trabajo para la transición a ETRS89, Consejo Superior Geográfico, octubre 2007d.
- . *El sistema ETRS89 y la nueva cartografía (Relación entre las series cartográficas y «mapas continuos»)*. Versión 1.0. Madrid: Grupo de trabajo para la transición a ETRS89, Consejo Superior Geográfico, octubre 2007e.
- . *Metodología de Producción de la Base de Datos CLC-Change 2000-2006*. Madrid, enero 2010a. Disponible en internet: <http://centrodedescargas.cnig.es/CentroDescargas/>.
- . *BTN25 Base Topográfica Nacional de España 1:25000. Anexo C Diccionario de Datos y Normas de Captura BTN25*. Versión 2.1, Madrid: Subdirección General de Producción Cartográfica, enero 2010b. <http://centrodedescargas.cnig.es/CentroDescargas/>.
- . Fichero vectorial de CORINE Land Cover 2006 del Centro de Descargas del Centro Nacional de Información Geográfica (CNIG) [consulta: 13 de septiembre del 2010c].
- . *Sistema de Información de Ocupación del Suelo en España —SIOSE2005—*. Documento Resumen. Madrid, 10 de mayo del 2011a.
- . *Modelo de Datos de la Base Cartográfica Nacional 1:200.000 BCN200. Versión 4.0*, Madrid, enero 2011b. Disponible en internet: <http://centrodedescargas.cnig.es/CentroDescargas/>.
- . «Fondos Cartográficos del Instituto Geográfico Nacional. España Siglos XVI y XIX». Madrid. Disponible en internet: <http://www2.ign.es/fondoscartograficos/> [consulta: marzo del 2012].
- JABLONSKI, Radoslaw. «Lessons learned from disaggregating population data by using different data sources». Ponencia presentada en el *European Forum for Geostatistics 2011*, Lisboa, 12-14 de octubre del 2011. Disponible en internet: <http://www.efgs.info/workshops/efgs-2011-lisboa-portugal>.
- JARQUE, Carlos M. y Anil K. BERA. «Efficient tests for normality, homocedasticity and serial independence of regression residuals». *Economics Letters* 6, n.º 3 (1980): 255-259.
- KAYADJANIAN, Maxime y Claude VIDAL. «Reassignment of the farm structure survey's data». En AEMA, ed. *Towards agri-environmental indicators. Integrating statistical and administrative data with land cover information*. Topic report n.º 6. Luxemburgo: Oficina de Publicaciones de la Unión Europea, capítulo 5 (2001): 75-95.
- KENNEDY, Melita y Steve KOPP. *Understanding Map Projections. GIS by ESRI™. ESRI ArcGIS™*. Redlands, CA: Environmental Systems Research Institute, Inc. (ESRI), 2000.
- LANGFORD, Mitchel. «Rapid facilitation of dasymetric-based population interpolation by means of raster pixel maps». *Computers, Environment and Urban Systems* 31, n.º 1 (enero 2007): 19-32.
- LANGFORD, Mitchel y David J. UNWIN. «Generating and mapping population density surfaces within a geographical information system». *Cartographic Journal* 31, n.º 1 (1994): 21-26.
- LIBRECHT, Ireen, Diedrik TIRRY, Francisco J. GALLEGO, Jos Van ORSHOVEN, Stefaan Vande WALLE y Jean Michel TERRES. *A GIS approach to measure the rural character*. n.º EUR 21182 EN. Bruselas: Joint Research Centre, Comisión Europea, 2004. Disponible en internet: <http://publications.jrc.ec.europa.eu/repository/handle/11111/1111/8004?mode=full>.
- MALDONADO CECILIA, José Luis. «Sistema Integrado de Población y Territorio: SIG_PT». Ponencia presentada en el curso *El padrón y el censo electoral: procedimien-*

tos de gestión, diferencias e interrelaciones organizado por el Instituto Nacional de Estadística, Madrid, 11 a 13 de abril del 2011.

MARTIN, David. «An assessment of surface and zonal models of population». *International Journal of Geographical Information Systems* 10, n.º 8 (1996): 973-989.

—. «Progress report: 24-hour gridded population models». Ponencia presentada en el *European Forum for Geostatistics 2010*, Tallin (Estonia), 5-7 de octubre del 2010. Disponible en internet: <http://www.efgs.info/workshops/efgs-2010-tallinn-estonia>.

MARTIN, David, Nicholas J. TATE y Mitchel LANGFORD. «Refining population surface models: Experiments with Northern Ireland Census data». *Transactions in GIS* 3 (2000): 285-301.

MILEGO, Roger y María José RAMOS. *Disaggregation of socioeconomic data into a regular grid and combination with other types of data*. Technical Report, ESPON Database 2013, European Observation Network for Territorial Development and Cohesion Programme, Bruselas: Comisión Europea. Disponible en internet: http://www.espon.eu/main/Menu_Projects/Menu_ScientificPlatform/espon-database2013.html [consulta: diciembre del 2010].

—. «Disaggregation of socioeconomic data and combination with other spatial data types by means of the OLAP technology». Ponencia presentada en el *European Forum for Geostatistics 2011*, Lisboa, 12-14 de octubre de 2011. Disponible en internet: <http://www.efgs.info/workshops/efgs-2011-lisboa-portugal>.

MINISTERIO DE FOMENTO. *Anuario Estadístico 2006*. Madrid: Dirección General de Programación Económica, Subdirección General de Estadísticas y Estudios, 2007. Disponible en internet: http://www.fomento.gob.es/MFOM/LANG_CASTELLANO/ESTADISTICAS_Y_PUBLICACIONES/INFORMACION_ESTADISTICA/EstadisticaSintesis/Anuario/2006/default.htm.

—. Construcción de edificios 2003-2007. Licencias municipales de obra. Serie Estadísticas. Madrid: Dirección General de Programación Económica. Subdirección General de Estadísticas y Estudios. 2008. Disponible en internet: <http://www.fomento.gob.es/NR/rdonlyres/8DA003C3-1-E26-4630-B2A8-B427ADB8AAD6/70811/CONSTRUCCIONDEEDIFICIOS2007.pdf>.

—. Construcción de edificios 2004-2008. Licencias municipales de obra. Serie Estadísticas. Madrid: Dirección General de Programación Económica. Subdirección General de Estadísticas y Estudios. 2009. Disponible

en internet: <http://www.fomento.gob.es/NR/rdonlyres/A15F150B-283D-47C3-A7C5-815E8052F4D4/70813/CONSTRUCCIONDEEDIFICIOS2008.pdf>.

—. Construcción de edificios 2005-2009. Datos recogidos de las licencias de obra concedidas por los ayuntamientos. Serie Estadísticas. Madrid: Inspección General de Fomento. Subdirección General de Tecnologías de la Información. Administración Electrónica de Estadísticas. 2010. Disponible en internet: <http://www.fomento.gob.es/NR/rdonlyres/F3A5D4AE-A798-4399-B4F2-5-D0DE2DF7878/75803/ConstruccionEdificios2009.pdf>.

MUGURUZA CAÑAS, Carmen y José Miguel SANTOS PRECIADO. «La importancia de las unidades de análisis en el modelo de ecología factorial». *Espacio, Tiempo y Forma* 2, Revista de la Facultad de Geografía e Historia de la UNED (1989): 87-101.

NAREDO, José Manuel, Óscar CARPINTERO y Carmen MARCOS. *Patrimonio inmobiliario y balance nacional de la economía española, 1991-2004*. Ensayos, Madrid: Fundación de las Cajas de Ahorros (FUNCAS), 2005.

NORDHAUS, William D. «Alternative approaches to spatial rescaling». Versión 2.2.2. Mimeo. New Haven, CT: Universidad de Yale, febrero 2002. Disponible en internet: <http://gecon.yale.edu/research-papers>.

—. «Geography and macroeconomics: New data and new findings». En *Proceedings of the National Academy of Sciences of the United States of America* 103, n.º 10 (marzo 2006): 3510-3517. Disponible en internet: <http://gecon.yale.edu/research-papers>.

—. «New metrics for environmental economics: Gridded economic data». *The Integrated Assessment Journal, Bridging Sciences and Policy* 8, n.º 1 (2008): 73-84.

OCDE (Organización para la Cooperación y Desarrollo Económicos). *Creating rural indicators for shaping territorial policy*. París, 1994.

OPENSHAW, Stan. *The Modifiable Areal Unit Problem*. CATMOG (Concepts and Techniques in Modern Geography) n.º 38, Norwich, Reino Unido: Geo Books, 1984.

O'SULLIVAN, David y David W. S. WONG. «A surface-based approach to measuring spatial segregation». *Geographical Analysis* 39, n.º 2 (abril 2007): 147-168.

PÉREZ MOREDA, Vicente. «El legado demográfico del antiguo régimen». En E. Llopis, coord. *El Legado Económico del Antiguo Régimen en España*, Barcelona: Editorial Crítica/Historia del Mundo Moderno (2004): 121-146.

- POELMAN, Hugo. «Population distribution grid uses in the context of regional and urban analysis at European scale». Ponencia presentada en el *European Forum for Geostatistics 2011*, Lisboa, 12-14 de octubre del 2011. Disponible en internet: <http://www.efgs.info/workshops/efgs-2011-lisboa-portugal>.
- PRENDES RUBIERA, Nicanor y Albarito LORENZO ALONSO. «Línea de costa. Criterios de generación, elaboración y definición». Ponencia presentada en la *Conferencia Europea Esri 2011*, CEDEX – Ministerio de Fomento, Madrid, 26-28 octubre 2011. Disponible en internet: <http://ficheros.esri.es/conferencia2011/conferencia/produccioncartografica/cedex-linea-costas.pdf>.
- RASE, Daniel. «Dealing with the modifiable areal unit problem. Spatial transformation methods for the analysis of geographic data». En *Towards agri-environmental indicators. Integrating statistical and administrative data with land cover information*. Topic report n.º 6. Luxemburgo: Oficina de Publicaciones de la Unión Europea, capítulo 2 (2001): 25-38.
- REHER, David S. «Ciudades, procesos de urbanización y sistemas urbanos en la Península Ibérica, 1550-1991». En M. Guardia, F. J. Monclús y J. Oyón (dirs.). *Atlas histórico de ciudades europeas*. Barcelona: Centre de Cultura Contemporània de Barcelona y Salvat (1994): 1-29.
- . «Perfiles regionales de la población española en tiempos históricos». En J. A. Fernández Cordón y J. Leal Maldonado (eds.). *Análisis Territorial de la Demografía Española 2006*. Madrid: Fundación Fernando Abril Martorell (julio de 2006): 53-83.
- REQUES VELASCO, Pedro y Vicente RODRÍGUEZ RODRÍGUEZ. *Atlas de la Población Española. Análisis de base municipal*. Santander: Servicio de Publicaciones de la Universidad de Cantabria, CSIC y ESRI-España, 1998.
- RODRÍGUEZ PASCUAL, Antonio F., Gema MARTÍN-ASÍN LÓPEZ y Beatriz ASTUDILLO MUÑOZ. «La Base Cartográfico Numérica 1:25.000 (BCN25)». *Mapping*, Revista Internacional de Ciencias de la Tierra, abril-mayo de 1997. Disponible en internet: http://www.mappinginteractivo.com/plantilla-ante.asp?id_articulo=782.
- SANTOS PRECIADO, José y Francisco J. GARCÍA LÁZARO. «El método dasimétrico-picnofiláctico: Un procedimiento para la desagregación de datos censales». Ponencia presentada en *IX Conferencia Iberoamericana de SIG*, Cáceres, 24-26 de septiembre de 2003.
- SCHMID, Calvin F. y Earle H. MACCANNELL «Basic problems, techniques and theory of isopleth mapping». *Journal of the American Statistical Association* 50 (1955): 220-239.
- SIMON, Alejandro, Jaume FONS y Roger MILEGO. *Urban Morphological Zones. Version F2v0. Definition and procedural steps. Final Report*. Barcelona y Copenhague: European Topic Centre, Land Use and Spatial Information (ETC LUSI), European Environment Agency (EEA), 2010. Disponible en internet: <http://www.eea.europa.eu/data-and-maps/data/urban-morphological-zones-2006-umz2006-f3v0/>.
- SMITH, Duncan Alexander. «Polycentricity and sustainable urban form. An intra/urban study of accessibility, employment and travel sustainability for the strategic planning of the London Region». Tesis doctoral. Londres: Center for Advanced Spatial Analysis & Department of Geography, University College London, agosto 2011. Disponible en internet: http://casa.ucl.ac.uk/duncan/DASmith_PhDthesis_2011.pdf.
- SUÁREZ VEGA, Rafael, Dolores R. SANTOS PEÑATE y Pablo DORTA GONZÁLEZ. «Generación de un modelo superficial de la población de Gran Canaria». En L. Hernández y J. M. Parreño (eds.). *Tecnologías de la Información Geográfica para el Desarrollo Territorial*. Las Palmas de Gran Canaria: Servicio de Publicaciones y Difusión Científica de la ULPGC (2008): 183-193.
- TOBLER, Waldo R. «Smooth pycnophylactic interpolation for geographical regions». *Journal of the American Statistical Association* 74, n.º 367 (septiembre 1979): 519-530.
- TRALLI, David M., Ronald G. BLOM, Victor ZLOTNIKI, Andrea DONNELLAN y Diane L. EVANS. «Satellite remote sensing of earthquake, volcano, flood, landslide and coastal inundation hazards». *ISPRS Journal of Photogrammetry and Remote Sensing* 59, n.º 4 (2005): 185-198.
- VIDAL, Claude, Javier GALLEGO y Maxime KAYADJANIAN. «Geographical use of statistical data. Methodological overview». En AEMA, ed. *Towards agri-environmental indicators. Integrating statistical and administrative data with land cover information*. Topic report n.º 6, Luxemburgo: Oficina de Publicaciones de la Unión Europea, capítulo 1 (2001): 11-24.
- VILLA, Guillermo. «CLC-00. Data model. Application and examples». Ponencia presentada en 2nd EIONET OODM Working Group Meeting. Madrid: Instituto Geográfico Nacional, 10-11 de diciembre de 2009. Disponible en internet: <http://inspire.jrc.ec.europa.eu/index.cfm/pageid/241/referenceid/32916>.

- VINKX, K. y T. VISÉE. «Usefulness of population files for estimation of noise hindrance effects». ICAO Committee on Aviation Environmental Protection. CAEP/8 Modelling and Database Task Force (MODTF) 4th Meeting. Sunnyvale, CA, 20-22 de febrero de 2008. Disponible en internet: <http://noise.eionet.europa.eu>.
- VINUESA ANGULO, Julio. *El Desarrollo Metropolitano de Madrid: Sus Repercusiones Geodemográficas*. Madrid: Instituto de Estudios Madrileños, 1976.
- . «La vivienda vacía en España: Un despilfarro social y territorial insostenible». Ponencia presentada en el *X Coloquio Internacional de Geocrítica. Diez años de cambios en el mundo, en la geografía y en las ciencias sociales. 1999-2008*. Barcelona: Universidad de Barcelona, 26-30 de mayo de 2008. Disponible en internet: <http://www.ub.es/geocrit/-xcol/74.htm>.
- WEBER, Norbert y Tim CHRISTOPHERSEN. «The influence of non-governmental organizations on the creation of Natura 2000 during the European Policy process». *Forest Policy and Economics* 4, n.º 1 (2002): 1-12.
- WHITE, Halbert A. «A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity». *Econometrica* 48, n.º 4 (mayo 1980): 721-746.
- WONG, David W. «The Modifiable Areal Unit Problem (MAUP)». En J. Janelle, B. Warf y K. Hansen (eds.). *WorldMinds: Geographical Perspectives on 100 Problems*. Cambridge, MA: Kluwer Academic Publishers, n.º 93 (2004): 571-575.
- WOOLDRIDGE, Jeffrey M. *Introductory Econometrics. A Modern Approach*. Segunda edición. Mason, OH: Thomson-South-Western, 2003.
- YUN CASALILLA, Bartolomé. «El siglo de la hegemonía castellana (1450-1590)». En F. Comín, M. Hernández y E. Llopis (eds.). *Historia Económica de España. Siglos X-XX*. Barcelona: Editorial Crítica/Historia del Mundo Moderno (2002): 51-84.
- ZOIDO NARANJO, Florencio y Andrés ARROYO PÉREZ. «La población de España». En A. Arroyo Pérez (coord.). *Tendencias demográficas durante el siglo XX en España*. Madrid: Universidad de Sevilla e Instituto Nacional de Estadística (1 de febrero 2004): 17-75. Disponible en internet: http://www.ine.es/prodyser/pubweb/tend_demo_s20/tend_demo_s20.htm.

Índice de cuadros, esquemas y gráficos

Cuadro 3.1:	Distribución de municipios por tamaño. Superficie ocupada y población en el 2010 ...	44
Cuadro 3.2:	Distribución provincial de municipios por tamaño	45
Cuadro 3.3:	Municipios, población y estadísticos de tamaño poblacional por provincias	46
Cuadro 3.4:	Entidades singulares, núcleos y diseminados por provincias. Nomenclátor del 2010	54
Cuadro 3.5:	Población en núcleos y diseminados por provincias. Nomenclátor del 2010	55
Cuadro 3.6:	Secciones censales (SSCC) y estadísticos de tamaño poblacional por provincias	64
Cuadro 3.7:	Municipios según número de secciones censales por provincias y población que albergan a nivel nacional.....	66
Cuadro 4.1:	Municipios en los que CLC2006 no informa de clase urbana (clases 1.1.1 o 1.1.2). Población según el padrón del 2010 y superficie asociada	74
Cuadro 6.1:	Secciones censales cuya población se pierde en el proceso de elaboración de la <i>grid</i> y población afectada	87
Cuadro 6.2:	Secciones censales puras por provincia y clase.....	90
Cuadro 6.3:	Población en secciones censales puras por provincia y clase.....	91
Cuadro 6.4:	Número de secciones censales puras por tamaño municipal y clase.....	93
Cuadro 6.5:	Población en secciones censales puras por tamaño de municipio y clase.....	94
Cuadro 6.6:	Densidad en secciones censales puras por tamaño de municipio y clase	94
Cuadro 6.7:	Índice normalizado de discrepancia, $100 \times \delta$, frente a <i>bottom-up grids</i> en diversos países europeos. <i>Grid</i> JRC-Eurostat-AEMA.....	101
Cuadro 6.8:	Índice normalizado de discrepancia, $100 \times \delta'$, frente a una <i>bottom-up grid</i> en la Comunidad de Madrid	102
Cuadro A.1.1:	Coberturas de <i>CORINE Land Cover</i> al nivel 3.....	110
Cuadro A.1.2:	Superficies de las coberturas de <i>CORINE Land Cover</i> 2006 al nivel 3	113
Cuadro A.2.1:	Coberturas simples del SIOSE2005.....	117

Cuadro A.2.2:	Coberturas compuestas predefinidas del SIOSE2005	119
Cuadro A.2.3:	Atributos de las coberturas simples del SIOSE2005.....	121
Cuadro A.3.1:	Medidas alternativas de superficie según diversas proyecciones.....	126
Esquema 1.1:	Aumentando la precisión de la localización espacial con información auxiliar	31
Esquema 1.2:	Transferencia de datos entre sistemas zonales: <i>areal weighting</i>	31
Esquema 1.3:	Dificultades con la aplicación simplista de <i>areal weighting</i>	32
Esquema 1.4:	Transferencia de datos entre sistemas zonales: <i>areal weighting</i> con información auxiliar.....	33
Esquema 6.1:	Diagrama simplificado de flujo de trabajo	84
Gráfico 2.1:	Superficie sobre el terreno que ocupará la edificación residencial. 2006/2007 y 2008/2009	39
Gráfico 3.1:	Distribución de municipios por tamaño	41

Índice de mapas

Mapa 1:	Densidad de la población de España por partidos judiciales. Censo de 1860.....	16
Mapa 2:	Contigüidad de los núcleos de población al margen de los lindes administrativos.....	17
Mapa 3:	Densidad de población. Padrón del 2010. Efecto de la escala de representación	18
Mapa 4:	Población georreferenciada: aproximación <i>bottom-up</i>	21
Mapa 5:	Mapa dasimétrico: aproximación <i>top-down</i> o <i>downscaling</i>	21
Mapa 1.1:	Coberturas de suelo de <i>CORINE Land Cover</i> 2006	26
Mapa 1.2:	Densidad de población municipal 2010.....	27
Mapa 1.3:	Imposición de una <i>grid</i> sobre otros sistemas zonales	28
Mapa 1.4:	<i>Grid</i> de población elemental de baja resolución	29
Mapa 3.1:	Localización de Emperador (46117, Valencia).....	42
Mapa 3.2:	Localización de Llocnou de la Corona (46152, Valencia)	42
Mapa 3.3:	Localización de Cáceres (10037) y Lorca (30024, Murcia).....	43
Mapa 3.4:	Densidad de población municipal. Padrón del 2010	49
Mapa 3.5:	Condominios	50
Mapa 3.6:	Fragmentación de Xàtiva (46145, Valencia)	50
Mapa 3.7:	Enclaves tangentes al contorno principal: Mohedas de Granadilla (10124, Cáceres)..	51
Mapa 3.8:	Polígonos <i>donut</i> : Albarracín (44009, Teruel)	51
Mapa 3.9:	Municipios sin contigüidad con otros municipios españoles	52
Mapa 3.10:	Núcleos de población para los que se dispone de coordenada. Nomenclátor del 2010 ..	58
Mapa 3.11:	Capa de <i>Entidades de población</i> de la BTN25. Método de captura	59
Mapa 3.12:	Capa de <i>Entidades de población</i> de la BTN25. Método de captura	60
Mapa 3.13:	Capa de <i>Entidades de población</i> de la BTN25 y <i>Núcleos de población</i> de la BCN200	61

Mapa 3.14:	Capa de <i>Edificación</i> de la BTN25 y de <i>Núcleos de población</i> de la BCN200	62
Mapa 3.15:	Capa de <i>Núcleos de población</i> de la BCN200.....	62
Mapa 3.16:	Capa de <i>Diseminados</i> de la BCN200	63
Mapa 3.17:	Secciones censales frente a municipios en las grandes ciudades. Municipio de Valencia (46250). 596 secciones censales	67
Mapa 3.18:	Densidades de población a partir de secciones censales. Villar del Arzobispo (46258)	68
Mapa 3.19:	Densidad de población por sección censal. Padrón del 2010	69
Mapa 3.20:	Densidad de población por sección censal. Padrón del 2010. Región de Murcia	70
Mapa 4.1:	<i>Grid</i> de población del JRC-Eurostat-AEMA. Censo del 2001.....	71
Mapa 4.2:	Coberturas urbanas y artificiales de CLC2006. Soria	76
Mapa 4.3:	Coberturas urbanas y artificiales de CLC2006. Burgos	76
Mapa 5.1:	CLC <i>versus</i> SIOSE: zona urbana	81
Mapa 5.2:	CLC <i>versus</i> SIOSE: zonas agrícolas	82
Mapa 5.3:	CLC <i>versus</i> SIOSE: zonas forestales	82
Mapa 6.1:	Polígonos susceptibles de soportar población residente: SIOSE2005	86
Mapa 6.2:	Generación de la <i>grid</i> : distribución proporcional por área.....	98
Mapa 6.3:	<i>Grid</i> de población para España.....	99
Mapa A.2.1:	Comparación de la resolución geométrica entre SIOSE2005 y CLC2006	116
Mapa A.3.1:	<i>Grid</i> de referencia para España en diversas resoluciones.....	124

Fundación **BBVA**

Plaza de San Nicolás, 4
48005 Bilbao
España
Tel.: +34 94 487 52 52
Fax: +34 94 424 46 21

Paseo de Recoletos, 10
28001 Madrid
España
Tel.: +34 91 374 54 00
Fax: +34 91 374 85 22

publicaciones@fbbva.es
www.fbbva.es



ISBN 978-84-92937-39-4



9 788492 937394