

Proyecto SIGECAH: diseño de un algoritmo de reconstrucción cartográfica asociado al Catastro de Ensenada

Laura García Juan^a, Ángel Álvarez de Miguel^b

^a*IULCE-UAM*

^b*Universidad de Alcalá de Henares*

E-mail: laura.garciaj@uam.es.

Resumen

En esta comunicación presentamos el primer prototipo de un módulo de simulación de escenarios cartográficos, concretamente desarrollaremos la estructura del motor basado en Autómatas Celulares. Tal y como se avanzó en comunicaciones anteriores, este módulo se encuentra integrado dentro de SIGECAH (Sistema de Gestión de Catastros Históricos), proyecto que tiene como objetivo el estudio y análisis de fuentes geohistóricas, y de manera más concreta, buscar y abrir nuevas vías de investigación que permitan avanzar en su estudio y en el manejo de la información en ellas contenida. En el desarrollo del prototipo se han integrado distintos métodos, tales como modelos probabilísticos y la evaluación multicriterio (sistemas de información geográfica). Debido a la complejidad del tema, y para facilitar la comprensión al lector mostraremos aquí el marco metodológico y la estructura del prototipo de un autómata celular, dejando para posteriores trabajos el desarrollo de aspectos más concretos.

Palabras clave: Fuentes Geo-Históricas, Catastro de Ensenada, Sistemas de Información Geográfica, Reconstrucción cartográfica, Autómatas Celulares.

1. Introducción

En los últimos años están surgiendo numerosas vías de investigación y trabajos que basan sus estudios en el empleo de Inteligencia Artificial con el objetivo de generar escenarios futuros. El crecimiento urbano es una de las áreas en las que más se ha empleado, sin olvidar otras como usos de suelo, vegetación... En las

mismas líneas existen estudios con carácter retrospectivo, donde destacaremos el de Camacho *et alii*. 2004, en el cual, a través de Sistemas de Información Geográfica y mediante la aplicación de evaluación multicriterio y multiobjetivo, se trata de obtener probabilidades de adecuación de un determinado uso de suelo en el pasado en un espacio concreto.

En nuestro caso buscamos soluciones globales para la carencia de cartografía asociada a fuentes catastrales históricas (también conocidas como fuentes geo-históricas de carácter textual) de singular relevancia, como puede ser el Catastro de Ensenada. Comenzamos nuestro trabajo estudiando en una línea retrospectiva, partiendo del presente, la evolución histórica de unos determinados espacios, analizando los procesos y variables implicados en el devenir histórico. Esta experiencia se tradujo en el establecimiento de una metodología que el lector puede consultar en García Juan *et alii*. 2012, y que es la base del prototipo que estamos presentando. La siguiente fase que estamos implementando y en la que se sitúa este prototipo da el salto desde un trabajo manual a un procedimiento más automático, basado en Inteligencia Artificial, concretamente empleando la estructura de un Autómata Celular.

El objetivo final del módulo no es obtener un software que nos ofrezca una solución única para una zona concreta, sino que tratamos de ofrecer una herramienta libre que posibilite el análisis de diferentes hipótesis obtenidas por los investigadores a partir del análisis de la información textual de diferentes fuentes geo-históricas. El procedimiento de trabajo comenzaría con el estudio de la información contenida en las fuentes, el establecimiento de las hipótesis sería la segunda etapa y, por último, nos encontraríamos con la traducción de esas hipótesis en reglas que serán las que rijan al autómata y que se codificarán en un algoritmo. En la fase de simulación, el investigador será un observador, en la que deberá ir valorando los resultados que se vayan obteniendo y analizando los cambios.

El prototipo está desarrollado para trabajar dos tipologías diferentes de datos: textuales y cartográficos, siendo la unión entre ambos la base del proyecto, aunque se presenten por separado para su mejor comprensión. Estas fuentes actuarán, bien como fuentes principales, o bien como fuentes auxiliares, empleadas de apoyo para la ejecución del módulo que nos ocupa.

En relación a las fuentes principales, actualmente estamos analizando dos:

- El *Catastro de Ensenada* (siglo XVIII). Esta fuente, que podemos clasificar como geo-histórica y textual, constituye un elemento singular, por la cantidad y calidad de la información que contiene, dentro del panorama europeo: abarca todo el territorio de la Corona de Castilla (unos 300.000 km² y casi 6,6 millones de habitantes). La fuente se ha empleado en distintas líneas de investigación y ha sido objeto de estudios en sí misma, resultando de gran valor para la investigación en múltiples áreas del conocimiento. Sin embargo, de cara a la investigación en Geografía presenta un problema importante: carece de cartografía técnica asociada a la información textual sobre el territorio. Ahora bien, a pesar de esa carencia, la fuente en sí misma puede considerarse un “mapa” que describe con gran exactitud las fincas rústicas y urbanas, su ubicación y los límites o lindes de los bienes inmuebles circundantes.
- Las *cédulas catastrales* asociadas a las *hojas kilométricas* del catastro topográfico-parcelario levantado para buena parte de las localidades madrileñas por los técnicos de la Junta General de Estadística en la segunda mitad del siglo XIX, según el proyecto técnico elaborado por Francisco Coello. La primera de las fuentes es de tipo textual; la segunda, textual y cartográfica.

Las fuentes auxiliares empleadas son fundamentalmente fuentes cartográficas de diferentes épocas, que se emplearán para ir ajustando el sistema. No podemos precisar un listado de fuentes concretas, pues no existen las mismas para todos los espacios, y éstas deben ser seleccionadas por cada investigador para cada caso concreto. Esa circunstancia se debe al hecho de que, en el panorama cartográfico español, no contamos con iniciativas que cubran la totalidad del territorio español, a excepción del catastro actual, y tampoco contamos con un registro fidedigno que nos informe de las fuentes existentes.

Concretamente, para el desarrollo del prototipo, se ha trabajado en cuatro zonas piloto, las mismas que han sido empleadas para el resto de los elementos de SIGECAH: Hoyo de Manzanares y Griñón (Madrid), Azuqueca de Henares (Guadalajara) y las ciudades de Granada y Soria. El Catastro de Ensenada ha constituido la fuente principal y, como fuentes auxiliares, hemos utilizado un amplio abanico: *hojas kilométricas, minutas y cédulas catastrales* del catastro topográfico-parcelario levantado por la Junta General de Estadística, *planimetrías* del Mapa Topográfico Nacional (la mayoría de la década de 1950 para nuestras zonas), fotografía aérea proveniente de varios vuelos y cartografía del catastro actual. Nos hemos centrado en analizar usos de suelo y simulación de escenarios cartográficos de tipo catastral.

2. Módulo de simulación SIGECAH: Diseño del autómata celular

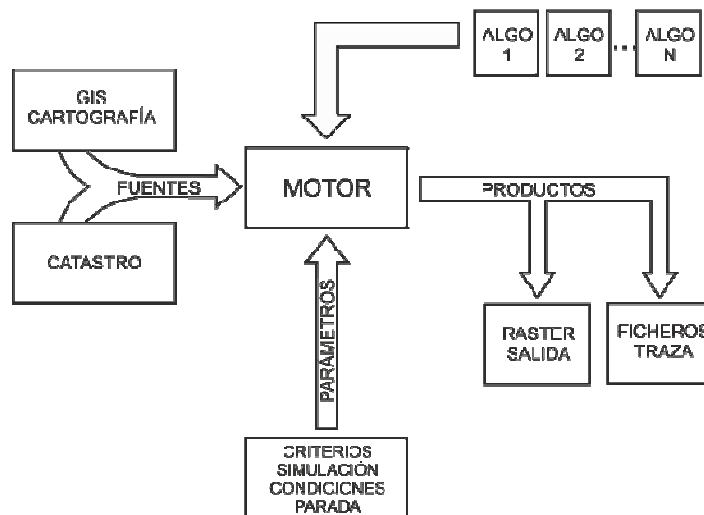


Fig. 1. Diagrama de módulos del prototipo

Como se ha indicado *ut supra*, para la ejecución del módulo de simulación/reconstrucción SIGECAH, usaremos algoritmos basados en el modelo de autómata celular. Para poder ejecutar el sistema, es necesario realizar una fase previa de análisis y de desarrollo de las hipótesis.

En esta fase pre-inicial el investigador debe seleccionar las fuentes, tanto principal como auxiliar, más adecuadas para la zona de estudio en la que quiere trabajar. En relación a las fuentes auxiliares debemos modelarlas para comenzar a definir los parámetros. Los datos a extraer se centran en dos grupos: información

textual y datos raster. Al final de proceso contaremos con una cartografía de partida y asociada a la misma distinguiremos dos bloques de elementos:

- *Elementos evaluables*: son aquellos factores que se han establecido como los más significativos y universales. Para cada una de las fuentes o cartografías auxiliares que empleemos en el análisis retrospectivo de nuestra línea cronológica debemos extraer tres capas: usos de suelo, toponimia y cambios/continuidades.

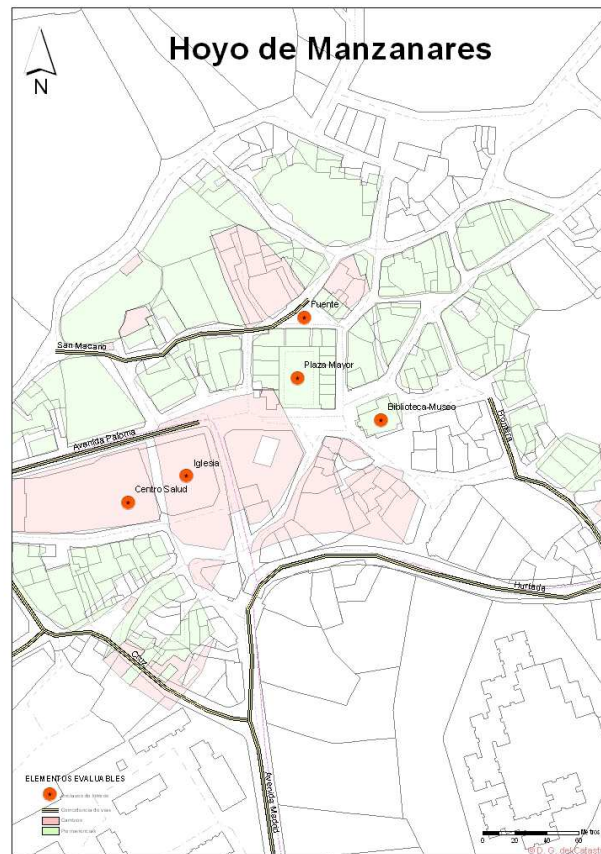


Fig. 2. Elementos evaluables correspondientes a una de las zonas piloto, Hoyo de Manzanares (Madrid).

- *Elementos de fricción*. hemos designado bajo este nombre a aquellas variables que utilizamos en conjunto con los elementos evaluables y que nos sirven para ajustar el resultado final:
 1. Enclaves de interés;
 2. Red viaria;
 3. Pendiente;
 4. Litología;
 5. Proximidad a la red hidrográfica;

6. Máscara de seguridad (axioma que indica que, cuanto más nos alejamos del núcleo urbano, la certidumbre se eleva).

Para facilitar diferentes tipos de uso, el motor de simulación tiene varios módulos posibles para el algoritmo del autómata celular de tal manera que cada algoritmo se puede ajustar para diferentes variantes del estudio (más preciso, más robusto etc.). El motor de la simulación opera sobre los datos de iniciales descritos y con las instrucciones facilitadas a través de los parámetros de configuración y la elección del algoritmo de los distintos disponibles. Los parámetros de configuración definen el comportamiento del autómata celular, como el tamaño de la vecindad y las condiciones de activación de las reglas para cada tipo de celda. Además, se definirán las condiciones de parada, elementos de fricción a considerar etcétera.

En relación a los algoritmos, inicialmente se trabajó sobre un único algoritmo, aunque finalmente se descartó esta vía debido fundamentalmente a la dificultad de cubrir todos los principios con los que se inició este módulo. Así pues, lo que se ha hecho es establecer una serie de diferentes módulos de funcionamientos similares, pero adaptados para diversos estudios. Contar con una librería de algoritmos posibilita el papel de investigación experimental del módulo. El usuario, a la hora de ejecutar el autómata, podrá optar por elegir su propio algoritmo o bien elegir uno entre los que se encuentren disponibles si alguno de éstos satisface sus hipótesis de trabajo.

La línea temporal gobernará de manera singular las transacciones de este autómata, que tratará de cubrir saltos cronológicos secuenciados por la introducción de las cartografías intermedias trabajadas en la fase anterior. Cada celda está sujeta a la evaluación de una serie de reglas de transición que asignan una probabilidad de cambio a un cierto estado, en función de los valores de su vecindad. Una vez evaluadas todas las reglas, el autómata generará un cambio al azar con una probabilidad dependiente del valor obtenido en cada regla, determinando así el siguiente estado que tendrá esta celda. Así mismo, cada regla tiene unos condicionantes en su activación (dependencias temporales y valores tabú) y sugerirá un valor que puede ser específico o bien una combinación lineal de diferentes valores obtenidos de la evaluación de su vecindad. Al introducirse una cartografía nueva, el autómata celular analizará patrones de cambios y continuidades con relación a la inicial. Tendrá en cuenta la superposición y la vecindad de los distintos elementos evaluables.

Así pues, como resultado, el sistema obtendrá un valor en el que se tendrá en cuenta la combinación de todos los elementos de acuerdo a unas reglas y una probabilidad de transición a este nuevo valor. Al no tener una componente determinista el algoritmo, los resultados obtenidos de su ejecución no serán siempre iguales, siendo necesario aplicar métodos estadísticos para obtener una significación en los análisis apropiada al tipo de estudio que se está realizando.

Para un conjunto de celdas dadas, la simulación consta de un número variable de transformaciones aplicando el algoritmo celular a cada una de las celdas representativas de la zona de estudio, el número de transformaciones dependerá, pues, de las reglas de activación o criterios de parada que se indiquen en los parámetros de configuración.

El último punto sería el relacionado con los productos de salida. A este respecto, debemos indicar que contamos con diferentes formatos que pueden ofrecer distintos tipos de datos:

- *Salida en raster* o en *vectorial*, que representa el estado final de la zona objeto de la simulación, en donde las distintas celdas representan el último estado al que han transitado.

- *Información textual*, compuesta por los ficheros de trazas (depuración), histogramas y referencias sobre valores como el RMS de ciertos elementos de fricción o topónimos, según especificó el usuario. Esta información ayuda a comprender cómo los criterios de parada han modelado la simulación de principio a fin.
- *Datos auxiliares*, como referencias necesarias para correr la herramienta integrada en el sistema SIGECAH.

Actualmente el prototipo tiene integrados un algoritmo en fase de pruebas, este trata de modelar los cambios geopolíticos que acontecen en el devenir de nuestra línea retrospectiva teniendo en cuenta las características de las distintas épocas históricas. Además cuenta con dos algoritmos en producción:

- *Algoritmo de uso de suelo*. Ofrece como resultado, el análisis del potencial de cambio de los usos de suelo a lo largo de la línea retrospectiva que haya establecido el investigador. Como regla principal, se ha propuesto que los espacios marcados con usos industrial y comercial irán desapareciendo para convertirse en espacios rurales o urbanos dependiendo de la hipótesis con la que se trabaje. En relación al crecimiento urbano, partimos de la premisa de una disminución progresiva del espacio urbano, aunque teniendo en cuenta que en raras ocasiones puede darse un proceso contrario, una mayor concentración urbana, cómo será el caso de los despoblados o núcleos de población venidos a menos. Otro valor presente es la proximidad a la red hidrográfica: cuanto más nos alejamos en nuestra línea temporal, más necesaria se hace la presencia de recursos hídricos para constatar un asentamiento urbano.
- *Algoritmo de simulación cartográfica catastral*. Muestra un escenario cartográfico con una probabilidad (nivel de certidumbre). Para su aplicación son importantes los ajustes de los parámetros, sobre todo prestando importancia al RMS obtenido de la georreferenciación de la cartografía histórica. La regla se centra en analizar, teniendo en cuenta este valor, la continuidad o no de una planta, asignándole mayor peso si en las celdas vecinas se encuentran elementos de fricción o elementos evaluables.

3. Integración del módulo de simulación con SIGECAH

SIGECAH constituye un marco común de investigación que presenta las interfaces de entrada, la visualización y un espacio de trabajo definido por un perfil de usuario. Todas estas funcionalidades se desarrollan a través de un cliente web, lo que implica que el usuario no tiene que contar con ningún tipo de software para la ejecución de los diferentes módulos.

En la imagen contigua el lector puede observar los componentes desarrollados en Joomla y cómo interactúan con el módulo de simulación a través de eventos. En este flujo de datos, la base de datos cumple un papel fundamental al ser el contenedor de todos los resultados obtenidos en el proceso.

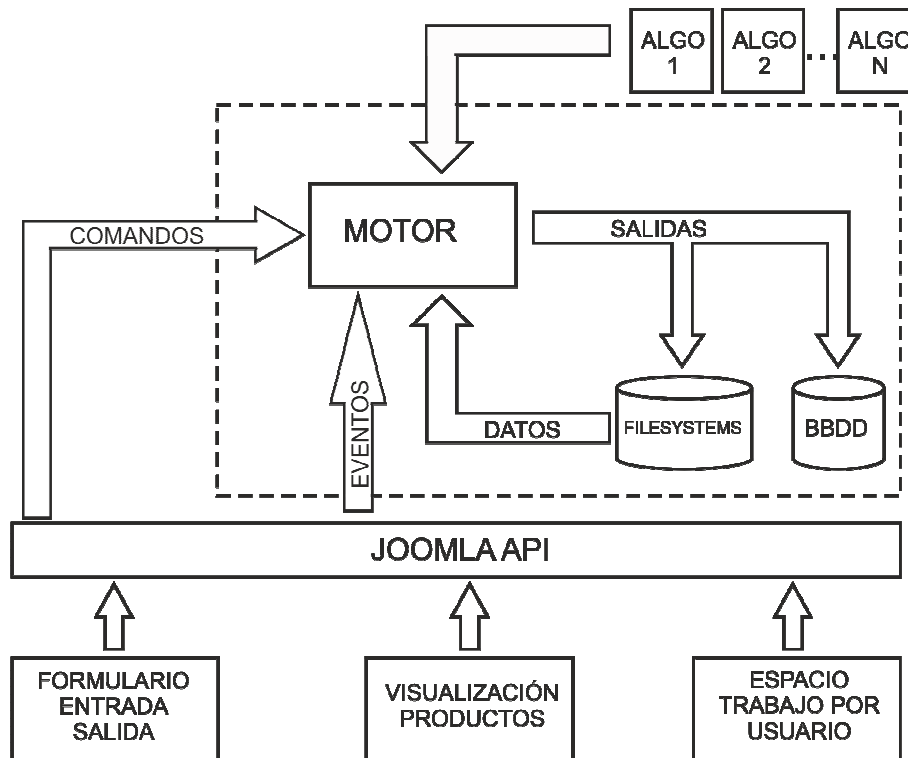


Fig. 3. Diagrama metodológico del módulo SIGECAH

Uno de los puntos que queremos destacar es la gestión de usuarios. Para acceder a los distintos módulos de trabajo, es necesario contar con un perfil individualizado para cada usuario. Este aspecto ofrece enormes ventajas no sólo en relación a la seguridad, ya que posibilita un control de usuarios, sino también en relación a la generación de directorios individualizados a través de los cuales los usuarios tienen acceso a los trabajos que realizan en el marco del proyecto contando con un directorio individualizado de salida.

4. Conclusiones

Las posibilidades de una herramienta como la que proponemos son múltiples y abarcan más allá del propio trabajo con fuentes geohistóricas ya que, aunque se empleen éstas para establecer los posibles escenarios, éstos pueden emplearse como fuente auxiliar para otras áreas de estudio.

Creemos en la necesidad de ejecutar herramientas propias que cubran las particularidades de algunas áreas de conocimiento como las planteadas en esta comunicación, empleando tecnologías y métodos ampliamente utilizados en otras, pero con menor incidencia en nuestro caso. Entre estas áreas de conocimiento, como hemos podido ver, las tecnologías de información geográfica ocupan un papel importante en unión con otras, como es el caso de la inteligencia artificial, bases de datos... El desarrollo de estas herramientas no se realiza

para un estudio concreto, sino que tienen como objetivo ofrecer un software que pueda ser empleado por toda la comunidad investigadora.

Agradecimientos

Este trabajo se realiza en el marco de los proyectos de investigación CEMU 2013, financiado por la Universidad Autónoma de Madrid y CSO2011-29027-C02-02/GEOG, financiado por la Dirección General de Investigación Científica y Técnica del Ministerio de Economía y Competitividad, dirigidos por los profesores Manuel Rivero y Concepción Camarero respectivamente.

Referencias

- Barreira González, P., Aguilera Benavente, F., Gómez Delgado, M. (2012): "Propuesta de un análisis de sensibilidad para validar modelos prospectivos de simulación de crecimiento urbano basados en autómatas celulares", en *GeoFocus*, 12, 303-328
- Bosque González, I. et alii. (2010): "Cartografía y demografía histórica en una IDE WMS del plano de Madrid de Facundo Cañada", *Revista Catalana de Geografia*, 40. [Disponible en red: <http://www.rcg.cat/articles.php?id=180>]
- Bosque Sendra, J., Gómez Delgado, M., Aguilera Benavente, F., Plata Rocha, W., Pascual Rosa, V. (2010): "Simulación de modelos de crecimiento urbano: métodos de comparación con los mapas reales", en *Tecnologías de la Información Geográfica: La Información Geográfica al servicio de los ciudadanos*. Sevilla, Secretariado de publicaciones Universidad de Sevilla, 1.000-1.013.
- Camacho Olmedo, M.T., Mulero Pérez, A., Naveros Santos, D., Salinas Sánchez, F. (2004): Ensayo de modelización retrospectiva del paisaje utilizando las técnicas de evaluación multicriterio y evaluación multiobjetivo. En "Territorio y medio ambiente: Métodos cuantitativos y Técnicas de Información Geográfica". Pp 303-316. GMCSIGT, AGE. Universidad de Murcia.
- Camacho Olmedo, M.T., Paegelow, M. y García Martínez, P. (2007): Modélisation géomatique rétrospective des paysages par évaluation multicritères et multiobjectifs. *Cybergéo (European Journal of Geography)*. 09-03-2007. Systèmes, modélisation & géostatistiques, N° 365, 24 pp.
- Camarero Bullón, C. (2002): "Averiguarlo todo de todos: el Catastro de Ensenada", en *Estudios Geográficos*, 248-239, 493-531.
- Camarero Bullón, C. (2002): El Catastro de Ensenada: magna averiguación fiscal para alivio de los vasallos y mejor conocimientos de los reinos. Madrid, Ministerio de Hacienda, 558 pp.
- Camarero Bullón, C. (2007): "La cartografía de los catastros españoles del siglo XVIII", en VV.AA: *La cartografía cadastral a Espanya (segles XVIII-XX)*. Barcelona, Institut Cartogàfic de Catalunya, 39-52.
- Camarero Bullón, C y Vidal Domínguez, M. J. (2013): "Los levantamientos topográfico parcelarios de la Junta General de Estadística en Soria (1867-1869). El plano de Soria de 1869", en *CT Catastro*, 76, 91-136.
- Capel Sáez, H. (2002): *La morfología de las ciudades. Vol.1: Sociedad, cultura y paisaje urbano*. Barcelona, Ediciones Serbal, 544 pp.
- Fernández-Wytenbach, A., Silbato, W., Moya, I., Dawood, Z., Galindo A. y Bernabé-Poveda, M.A. (2009): "Integración de Cartotecas Virtuales como herramienta de apoyo en la investigación histórica y social", en *Actas de las Jornadas de la Infraestructura de Datos Espaciales de España (JIDEE 2009)*. Murcia, 4-6 noviembre 2009.
- Ferrer Rodríguez, A. (2002): "La documentación del Catastro de Ensenada y su empleo en la reconstrucción cartográfica", en *CT Catastro*, 46, 99-110 (español) y pp. 161-166 (inglés).
- García Juan, L., Álvarez Miguel, A.J., Camarero Bullón, C. y Escalona, J. (2011): "Modelo de datos para la digitalización y gestión de fuentes catastrales geohistóricas textuales. Aplicación al Catastro de Ensenada", en *CT Catastro*, 72, 73-97.
- García Juan, L., Escalona, J. y Camarero Bullón, C. (2008): "Propuesta metodológica para la reconstrucción del parcelario antiguo mediante sistemas de información geográfica", en *CT Catastro*, 63, 203-214.
- García Juan, L., Álvarez Miguel, A.J., Camarero Bullón, C. y Escalona, J. (2012): "Generación de una metodología para la gestión y recreación cartográfica a partir de la información del Catastro de Ensenada", en *Geofocus*, 12, 268-302.
- García Juan, L., Álvarez Miguel, A.J. y Fernández Sánchez, N. (2012): "Pasos para la georreferenciación automática del Catastro de Ensenada", en *CT Catastro*, 75, pp. 55-71.
- Muro Morales, José Ignacio (2007): "Las realizaciones catastrales de la Junta General de Estadística", en VV.AA.: *150 aniversario de la creación de la Comisión de Estadística General del Reino*. Madrid, INE, 305-330.
- Muro Morales, José Ignacio, Nadal Piqué, Francesc y Urteaga, Luis (1992): "Los trabajos topográfico-catastrales de la Junta General de Estadística (1856-1870)", en *Ciudad y territorio: Estudios territoriales*, 94, 33-59.

- Muro Morales, J.I., Nadal Piqué, F. y Urteaga, L. (1996): *Geografía, estadística y catastro en España, 1856-1870*. Barcelona, Ediciones del Serbal.
- Nadal Piqué, Francesc (2007): "El proyecto catastral de Francisco Coello", en VV.AA.: *150 aniversario de la creación de la Comisión de Estadística General del Reino*. Madrid, INE, 287-304.
- Urteaga, Luis y Nadal Piqué, Francesc (2001): *Las series del mapa topográfico de España, escala 1:50.000*. Madrid, Centro Nacional de Información Geográfica, 368 pp.