

METODOLOGÍA PARA DEDUCIR RELACIONES DE LINAJE EN EL CATASTRO DE ESPAÑA

PILAR MORENO REGIDOR, TERESA ITURRIOZ AGUIRRE Y SANDRA MARTÍNEZ CUEVAS

Universidad Politécnica de Madrid | ETSI en Topografía, Geodesia y Cartografía
Autovía de Valencia km 7, E28031 Madrid (España)

mariapilar.moreno@upm.es ; teresa.iturrioz@upm.es ; sandra.mcuevas@ump.es

RESUMEN

En España, los datos catastrales de acceso público, obtenidos a través de la Sede Electrónica del Catastro (SEC), no incluyen información sobre la genealogía o las relaciones de linaje existente entre las parcelas, de forma que la gestión de la información histórica es muy limitada. Este artículo presenta un método para obtener las relaciones de linaje más frecuentes entre las parcelas (agregación y segregación) y propone un prototipo de estructura relacional para el almacenamiento y la gestión histórica de los datos catastrales de acceso público. El proceso de análisis para deducir el linaje se basa en superposiciones espacio-temporales junto con secuencias de sentencias SQL. El método proporciona un 70% de relaciones de agregación y segregación correctas; el resto presentan errores debidos, en general, a anomalías presentes en los propios datos catastrales.

Palabras clave: Catastro histórico, relaciones de linaje, superposiciones espacio-temporales

METHODOLOGY TO INFER LINEAGE RELATIONSHIPS IN THE SPANISH CADASTRE

ABSTRACT

Public cadastral data in Spain, available through the Electronic Office of Cadastre, do not include any information on the genealogy or lineage relationship among the parcels, so the management of historical information becomes very limited. This paper presents a method for obtaining the most common lineage relationships between plots (aggregation and segregation) and proposes a relational structure prototype for the storage and historic management of public cadastral data. The analysis process to deduce the lineage is based on spatio-temporal overlaps coupled with

SQL statements sequences. This method provides a 70% of correct aggregation and segregation relationships; the wrong cases are mainly caused by errors existing in the cadastral data themselves.

Keywords: historical cadastre, lineage relationships, spatio-temporal overlaps

1. Introducción

El Catastro español es un registro administrativo que contiene la descripción de los bienes inmuebles rústicos, urbanos y de características especiales, incluyendo sus características físicas, económicas y jurídicas. Estos datos lo convierten en un gran banco de datos del territorio disponible para todo usuario, del sector público o privado. Los datos catastrales son de naturaleza dinámica y los bienes inmuebles sufren continuos cambios, por lo que se precisa conocer tanto la información catastral actual como la histórica. Desde la Resolución de 23 de marzo de 2011, la Dirección General del Catastro (DGC) permite el acceso al servicio de descarga masiva de datos, espaciales y temáticos, a través de la Sede Electrónica del Catastro (SEC¹). Estos datos comprenden la versión actual y las versiones de las parcelas desde el año 2001. En la base de datos catastral accesible no se incluye información sobre la genealogía o las relaciones de linaje entre las parcelas, de forma que la gestión de la información histórica es muy limitada. Si un usuario necesitara los datos correspondientes a la evolución de un conjunto específico o bien de todas las parcelas de un municipio, tendría que diseñar un proceso de técnicas de análisis que, hasta la fecha, no ha sido descrito ni validado para el Catastro de España en ningún trabajo de investigación. Las cuestiones que se plantean son: ¿es posible, usando las funciones de un SIG o las de una base de datos espacial, obtener y almacenar la genealogía o linaje de las parcelas a partir de los datos catastrales de la SEC? Y, si es factible, ¿con qué grado de fiabilidad?

Este artículo presenta un método para obtener las relaciones de linaje más frecuentes (agregación y segregación) y propone un prototipo de estructura relacional para el almacenamiento y la gestión histórica de los datos catastrales de acceso público (SEC). El modelo relacional se plantea como una alternativa complementaria del actual modelo del Catastro de España, que permitiría realizar consultas sobre la genealogía y trazabilidad de las parcelas. El proceso de análisis para deducir el linaje se basa en superposiciones espacio-temporales entre las versiones de las parcelas junto con secuencias de sentencias SQL. El método ha sido validado empleando datos del Catastro de rústica y urbana de 19 municipios de las provincias de Madrid y Guadalajara. Tras revisar un 60% de las relaciones de agregación y segregación de la tabla de linaje de cada uno de los municipios, se ha comprobado que un 70% de ellas son correctas. El resto de las relaciones obtenidas presentan errores que, en su mayoría, tienen origen en las anomalías presentes en los datos catastrales. El método fue desarrollado inicialmente con los programas *GeoMedia Professional* v.6.1 y *ArcGIS* v.10, utilizando una base de datos relacional *Access* como soporte de almacenamiento. Posteriormente se ha implementado en *PostgreSQL*, con la extensión espacial *PostGIS*, mediante un *script* PL/PGSQL que permite generar rápidamente, en cuestión de algunos minutos, el linaje correspondiente al Catastro de un municipio.

El artículo está estructurado como sigue: el apartado 2 trata el estado de la cuestión; en el apartado 3 se analizan los cambios que generan relaciones de linaje entre las parcelas del Catastro de España; en el apartado 4 se presenta un prototipo de estructura relacional para almacenar los

datos catastrales y el linaje de las parcelas; en los apartados 5 y 6 se presenta el método para deducir las relaciones generadas por procesos de agregación y segregación entre parcelas y las limitaciones del algoritmo diseñado en su aplicación práctica al Catastro de España; y, por último, en el apartado 7 se presentan las conclusiones y líneas futuras de investigación.

2. Estado de la cuestión

Las entidades geográficas son entes dinámicos que están sometidos a diversos tipos de cambios que manifiestan su evolución a lo largo del tiempo. De ahí la necesidad de incorporar la dimensión temporal en los sistemas de información, más en concreto, en las bases de datos y en los SIG, lo que se ha puesto de manifiesto en los diversos métodos utilizados desde principios de los años 90 hasta hoy en día (Snodgrass, 1992; Langran, 1992; Worboys, 1992; Bohlen, 1995; Yuan, 1999; Rodríguez *et al.*, 2001; Choi *et al.*, 2008; etc.).

La dimensión temporal permite referenciar los cambios a que están sometidas las entidades geográficas. Generalmente se consideran una o dos de las siguientes dimensiones temporales: (a) tiempo válido, que registra el instante en que suceden los eventos del mundo real; y (b) tiempo de transacción, que registra el instante en que la información sobre el mundo real se introduce en el sistema. En función de las dimensiones temporales consideradas, los sistemas de información pueden clasificarse en cuatro tipos básicos: estáticos (no incluyen ninguna dimensión temporal), históricos (únicamente gestionan el tiempo válido), *rollback* (solo gestionan el tiempo de transacción) y bi-temporales (sistemas que permiten gestionar las dos dimensiones temporales).

El método más utilizado para representar la información temporal consiste en el uso de marcas de tiempo (Snodgrass, 1992), existiendo tres tipos básicos: (a) marcas de evento, que constan de un único identificador de la unidad de tiempo elegida (t_i); (b) marcas de intervalo, formadas por dos marcas de evento que indican los instantes de inicio y finalización del intervalo ($[t_{inicio}, t_{fin}]$) y (c) *spans* o períodos, que son marcas que indican duración temporal sin especificar los eventos de inicio y fin (un mes, una semana, etc.). Todas estas marcas se pueden asociar tanto a entidades temporales como a cualquiera de las propiedades (atributos) que caracterizan a dichas entidades (Stefanakis, 2003).

Las parcelas son un tipo de fenómeno geográfico espacio-temporal cuyas propiedades espaciales y temáticas evolucionan a lo largo del tiempo. Tradicionalmente, las bases de datos catastrales han sido sistemas estáticos que registraban el estado del parcelario en un determinado momento, de forma que en los procesos de actualización se perdían los datos correspondientes al estado anterior. Es obvio que esta información histórica es de gran interés para muchas organizaciones, tales como administraciones, autoridades legales, propietarios, etc. La necesidad de conservar la información histórica en los sistemas catastrales comienza a manifestarse en la década de los 90 (Vrana, 1989; Hunter y Williamson, 1990). A partir de entonces empieza un gran trabajo de investigación sobre diseño y gestión de bases de datos catastrales históricas (Arcieri *et al.*, 1999; Lemmen y Van Oosterom, 1995; Frank y Kemp, 2001; Peerbocus *et al.*, 2001; Arriaga, 2003; Nan *et al.*, 2006; Choi *et al.*, 2008; Song y Yang, 2013), lo que va acompañado del desarrollo de modelos de datos catastrales espacio-temporales. Hunter y Williamson (1990) proponen el uso de marcas temporales de intervalo, para indicar la fecha de creación/aparición y finalización/cese de

los objetos catastrales. En esta línea Peerbocus y Jomier (2004), Lemmen y Van Oosterom (1995), Van Oosterom (1997) y Van Oosterom y Lemmen (2001) hacen una propuesta para gestionar los datos históricos del Catastro de Holanda mediante la incorporación de dos nuevos atributos ($t_{\min}$ y $t_{\max}$) a su modelo de datos: cuando se crea un objeto, se asigna a $t_{\min}$ el tiempo correspondiente al instante del registro de datos o transacción y $t_{\max}$ se rellena con un valor especial; cuando un objeto se destruye, en $t_{\max}$ se registra el instante del registro/transacción y $t_{\min}$ conserva su valor. También hay que citar los trabajos de Peerbocus *et al.* (2001) y Peerbocus y Jomier (2004) sobre el almacenamiento y gestión de planes catastrales² en bases de datos de tipo multiversión.

Para almacenar correctamente la historia de las entidades, también es necesario registrar las relaciones existentes entre ellas a lo largo del tiempo. Las entidades geográficas, ya sean físicas (parcelas, carreteras) o abstractas (límites administrativos), tienen una identidad y un conjunto de propiedades, espaciales y temáticas, que pueden variar. Estos cambios pueden clasificarse en dos categorías básicas (Frank *et al.*, 2001): (a) cambios de vida y (b) cambios de movimiento. Los primeros se refieren al cambio en el estatus de existencia de las entidades: un ente puede aparecer y desaparecer (lagunas, zonas de especial protección), puede fusionarse con otros o dividirse en dos o más entes nuevos (unidades administrativas, objetos catastrales), etc. Los segundos son cambios en las propiedades temáticas y/o espaciales de las entidades, tales como: variaciones de geometría (rectificación de los límites de una parcela), de posición (movimiento de un vehículo), de forma y posición (propagación de un incendio); modificaciones en los atributos (cambio de titularidad de una vivienda, de uso de suelo de una parcela), etc. Si las entidades se analizan individualmente, solo se consideran los cambios de vida y movimiento. Sin embargo, cuando la información histórica resulta de interés (Stefanakis, 2003), también hay que considerar (c) los cambios de tipo "sucesión", que realmente son una secuencia de cambios de movimiento y/o de vida que afectan a un grupo de entidades. La sucesión permite conservar las relaciones entre las entidades a lo largo del tiempo, más en concreto, entre entidades antecesoras y sucesoras o descendientes. En Catastro, una parcela puede fusionarse a otras existentes o subdividirse y generar nuevas parcelas, creándose vínculos de parentesco del tipo padre-hijo o antecesor-sucesor. A este tipo de relaciones se les denomina relaciones de linaje y definen, a grandes rasgos, la genealogía de las parcelas, objetos que, por tanto, están sometidos a los tres tipos de cambio descritos.

Para el seguimiento de la historia o evolución de los objetos catastrales se usan, básicamente, dos métodos (Song y Yang, 2013). El primero, apenas desarrollado en la literatura, utiliza técnicas de análisis y superposición espacio-temporal entre parcelas para obtener, en un momento dado, información sobre el linaje de uno o varios objetos de interés. El segundo, tema de múltiples trabajos de investigación, consiste en usar vínculos y marcas temporales para almacenar explícitamente las relaciones padre-hijo según se van generando entre las parcelas. A continuación se citan algunos de los autores más importantes que han desarrollado sus trabajos aplicando este segundo método.

Spéry *et al.* (1999) proponen un modelo de metadatos de linaje, adaptado al Catastro de Francia, que, para representar la historia de las parcelas, asocia cada una de ellas a sus padres e hijos mediante un árbol de filiación. Los autores consideran que las parcelas están sometidas a dos categorías de cambios: cambios elementales (división, fusión, extracción, rectificación y eliminación de las parcelas) y cambios complejos (expropiaciones y reparcelación). Este modelo, ampliado en trabajos posteriores (Spéry *et al.*, 2001), debería permitir la creación del grafo

genealógico de una parcela. En esta misma línea, Peerbocus y Jomier (2004) indican que los cambios geométricos elementales, autorizados por la administración francesa (Ministerio de Economía, 1995), son cinco: (1) "extracción" o creación de una parcela a partir de la Propiedad Estatal; (2) integración de una parcela en la Propiedad Estatal; (3) rectificación de los límites entre dos parcelas; (4) división de una parcela en otras nuevas y (5) fusión de varias parcelas contiguas en una nueva. En su trabajo proponen la creación de un grafo, en una base de datos objeto-relacional, mediante un objeto genealógico donde se almacenan los identificadores internos (*id_obj_padre*, *id_obj_hijo*) de cada pareja de parcelas relacionadas. Un identificador (*id_obj_hijo*) corresponde al nuevo objeto catastral que se crea y el otro (*id_obj_padre*) al objeto del que procede. En el objeto genealógico pueden representarse las relaciones de linaje con cardinalidad 1:n, n:1 y n:m.

Para gestionar el Catastro urbano de Zuland (China), Nan *et al.* (2006) proponen una base de datos espacio-temporal organizada en tres partes o sub-bases interdependientes: dos para almacenar los datos espaciales y temáticos de los objetos catastrales (una para los objetos actuales y otra para los históricos), y la tercera para los datos de los procesos correspondientes a los trámites de gestión de la tierra. En esta propuesta se define una tabla para almacenar los cambios históricos (*History_change_table*). El contenido de esta tabla corresponde a las relaciones de linaje entre padres e hijos, por lo que permite reconstruir la historia de las parcelas. El modelo se ha implementado en un sistema gestor de base de datos relacional, utilizando diferentes técnicas (*Section Fast Indexing* (SFI), *Storage Factors of Variable Granularity* (SFVG)) para mejorar la eficiencia del sistema.

Con el fin de gestionar la historia de las entidades, Choi *et al.* (2008) proponen un modelo temporal basado en entidades (*Feature Based Temporal Model* (FBTM)). Junto con este modelo, los autores han diseñado un prototipo, *Feature History Management System* (FHMS), que puede utilizar como sistema de almacenamiento tanto una base de datos relacional (*Access*) como una objeto-relacional (*Oracle* con la extensión *Spatial Cartridge*). El prototipo emplea los datos catastrales, desde el año 2000 al 2005, de una zona del norte de Atenas. Los resultados obtenidos demuestran la validez del modelo para resolver consultas sobre la historia individual de las entidades (cambios en la geometría y atributos de una entidad -parcela- a lo largo de su período de vida) y sobre las relaciones entre las versiones de un grupo de entidades.

Por último, Song y Yang (2013) describen cuatro tipos de cambios espaciales que afectan a las parcelas del Catastro de China: (1) división de una parcela en dos o más partes; (2) unión o fusión de dos o más parcelas en una nueva; (3) reparcelación del terreno, de forma que se crean 'p' nuevas zonas catastrales a partir de las 'n' zonas existentes inicialmente y (4) modificación o rectificación del límite común entre dos zonas. Estos autores proponen un modelo catastral basado en el modelo de composición espacio-temporal. Al igual que en el trabajo de Nan *et al.* (2006), para gestionar los objetos catastrales se utiliza una base de datos histórica y otra actual, ambas definidas en un sistema objeto-relacional. Para implementar el modelo en un entorno SIG, estos autores han desarrollado un Sistema de Información Catastral Espacio-Temporal, donde los resultados obtenidos demuestran la validez del modelo en el almacenamiento y gestión de los objetos catastrales, así como en el seguimiento de sus cambios e historia.

Finalmente, indicar que existen otros trabajos sobre el linaje no catastral, realizados en el ámbito de las divisiones administrativas, tales como los de Kauppinen *et al.* (2008) en Finlandia,

Moreno Regidor, P., Iturrioz Aguirre, T. y Martínez Cuevas, S. (2014): "Metodología para deducir relaciones de linaje en el Catastro de España", *GeoFocus (Artículos)*, nº 14, p. 275-300. ISSN: 1578-5157

Gantner *et al.* (2013) en Suiza, y Lacasta *et al.* (2014) a nivel jurisdiccional europeo, en los que se aplican distintos enfoques semánticos para generar o explotar sus relaciones espacio-temporales.

En ninguno de los trabajos realizados en el ámbito del Catastro se describe un método de análisis espacio-temporal para la deducción de la genealogía de los objetos catastrales. Por esta razón, este artículo propone un método para obtener las relaciones de linaje más frecuentes en el Catastro de España (agregación y segregación) y un prototipo de estructura relacional que permita el almacenamiento y la gestión histórica de los datos accesibles a través de la Sede Electrónica del Catastro (SEC). El objetivo es facilitar, a los usuarios de estos datos catastrales, la obtención del linaje de los municipios de interés y su almacenamiento en una estructura alternativa a la actual del Catastro. De esta forma sería posible realizar todo tipo de consultas sobre la genealogía e historia de las parcelas, proceso actualmente imposible a partir de los datos de acceso público (SEC). El método propuesto ha sido validado mediante los datos catastrales, de rústica y urbana, de 19 municipios de las provincias de Madrid y Guadalajara.

3. Relaciones de linaje en el Catastro de España

Las parcelas catastrales son entidades que están sometidas a los tres tipos de cambio descritos en la sección anterior: (a) cambios de vida, (b) cambios de movimiento y (c) cambios de tipo sucesión. En España, la asignación de la referencia catastral (*National Cadastral Reference*, Directiva Inspire) a una parcela indica su registro o aparición como un nuevo objeto catastral (cambio de vida). A partir de ese momento puede experimentar variaciones en sus propiedades espaciales y/o temáticas (cambios de movimiento). Por último, la evolución de una parcela puede definirse como el conjunto de sucesivos estados desde su creación hasta que se elimina o desaparece de la base de datos catastral.

Los cambios espaciales y de vida de los objetos catastrales han sido estudiados en diferentes países y en varios trabajos de investigación (Spéry *et al.*, 1999; Peerbocus y Jomier, 2004; Song y Yang, 2013). Analizando los cambios desde el punto de vista de la sucesión o linaje, Stefanakis (2003) considera que, en una aplicación catastral, pueden aparecer los siguientes tipos: (a) agregación; (b) agrupación; (c) segregación y (d) división. Estos cambios, básicamente espaciales, son los que generan mayoritariamente las relaciones de linaje en el ámbito catastral.

En el Catastro de España, según el Art. 18 del Real Decreto 417/2006, las parcelas pueden verse afectadas por los siguientes procesos (cambios espaciales):

1. División: la geometría de una parcela se divide en dos o más geometrías nuevas. La referencia catastral de la finca matriz desaparece y se asigna una nueva referencia a cada una de las fincas resultantes.
2. Segregación: una parcela se divide en varias geometrías nuevas. Se mantiene la referencia de la finca sobre la que se practica la segregación y se asignan nuevas referencias al resto.
3. Agrupación: varias parcelas se fusionan en una geometría nueva. La referencia de las fincas agrupadas desaparece y se asigna una nueva a la finca resultante.
4. Agregación: varias parcelas se fusionan a otra, creando una nueva geometría. Se mantiene la referencia catastral de la finca a la que se unen.

5. **Expropiación:** una parcela existente "desaparece" pasando a formar parte del dominio público, diferenciándose dos casos:
 - Rectificación, si los límites comunes entre, al menos, dos geometrías cambian espacialmente. Equivale a aplicar una división a cada una de las geometrías de forma individual y eliminar una de las dos partes resultado de la división.
 - Eliminación, si la expropiación es total y la geometría desaparece.
6. **Aparición:** se genera una nueva parcela mediante un recinto procedente del dominio público.
7. **Concentración parcelaria** (reparcelación): es un cambio espacial compuesto de varios simples (aparición, desaparición, agregación, división, etc.). A partir de un conjunto de n parcelas se genera otro conjunto m, sin que exista, a nivel individual, una relación precisa de parentesco entre las geometrías nuevas y las previas. La superficie total de las nuevas geometrías ha de ser la misma que la de las anteriores.

Además de los cambios indicados, también pueden producirse modificaciones individuales en la geometría de las parcelas, es decir, una rectificación de sus límites que puede afectar o no a los límites de las parcelas adyacentes. En líneas generales, estas rectificaciones pueden tener su origen en procesos de revisión y actualización cartográfica y se caracterizan porque suelen producir, teóricamente, variaciones superficiales no superiores al 10% del área de las parcelas afectadas. Así, para estudiar la historia o evolución de las parcelas es necesario almacenar las relaciones de linaje que se generan entre ellas. Si se conserva un registro de cada una de estas relaciones se obtiene la genealogía de cada parcela. Dicha genealogía presenta una estructura muy sencilla basada en relaciones de parentesco del tipo padre-hijo (figura 1): una parcela padre, o antecesor, puede tener una o varias parcelas hijas (esta relación se produce en casos de división y segregación); y toda parcela hija, descendiente o sucesor, puede tener una o varias parcelas padre (casos de unión y agregación).

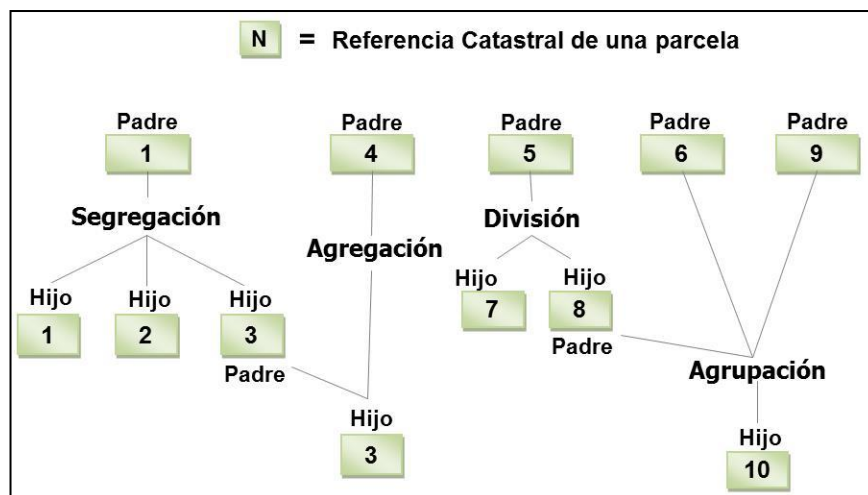


Figura 1. Ejemplo de relaciones padre-hijo entre parcelas

Las relaciones entre las parcelas padre e hijo (P:H), sin considerar los procesos donde interviene el dominio público, se pueden clasificar en tres tipos en función de su propiedad cardinal:

- (a) 1:n- Un padre tiene n hijos y cada hijo un solo padre (división, segregación)

- (b) n:1- Un hijo tiene n padres y cada padre un solo hijo (agrupación, agregación)
- (c) n:m- Un hijo tiene n padres y cada padre m hijos (reparcelación).

Los procesos de rectificación de límites entre dos objetos catastrales son un caso particular del tipo c (2:2), ya que cada parcela hijo tiene dos padres y cada padre tiene dos hijos que son los mismos.

Los procesos de expropiación y aparición quedan, de momento, fuera del ámbito de este trabajo. Esto es debido a que, en los datos de acceso público, no figura explícitamente la parcela asociada al dominio público –a excepción de los recintos de este tipo incluidos en el Catastro de rústica– lo que impide, en cierta medida, registrar dicha parcela en el linaje correspondiente a estos casos. A partir del análisis de las relaciones de linaje es posible diseñar una estructura de almacenamiento que permita analizar la evolución histórica de las parcelas.

4. Estructura de datos propuesta

Este trabajo propone un prototipo de estructura relacional para una base de datos catastral histórica. El prototipo se basa en el actual modelo de datos del Catastro de España, y se presenta como una alternativa que incluye la información del linaje para poder realizar un seguimiento o trazabilidad histórica de las parcelas. Esta estructura es similar a otros modelos, propuestos anteriormente (Oosterom y Lemmen, 2001; Stefanakis, 2003; Peerbocus y Jomier, 2004; Nan *et al.*, 2006; Choi *et al.*, 2008; Cömert y Alkan, 2004; Song y Yang, 2013), que han sido implementados en sistemas gestores de base de datos relacionales, objeto-relacionales u orientadas a objetos. En casi todos estos modelos, al igual que en esta propuesta, se define un objeto genealógico o una tabla para almacenar la información del linaje entre las parcelas, es decir, las relaciones del tipo padre-hijo comentadas anteriormente.

Con respecto a la incorporación del tiempo, dimensión básica en la gestión histórica, no se ha seguido la línea de los modelos que gestionan de forma independiente las dimensiones del espacio, tiempo y atributos –modelos de tres dominios (Claramunt y Theriault, 1996; Yuan, 1999) y modelo temporal basado en entidades (Choi *et al.*, 2008)–. En nuestro caso, la estructura que se plantea es semejante a la del modelo *Object-Oriented Spatiotemporal* (Worboys, 1992), ya que los objetos catastrales se representan como objetos espaciales bidimensionales con el tiempo como tercera dimensión. Por tanto, los datos temáticos son dependientes del espacio al almacenarse como atributos de la geometría de las parcelas. Además, para gestionar adecuadamente la evolución temporal de los objetos, se recurre al concepto de versión, una vez demostrada su eficacia como medio para almacenar los diferentes estados de las entidades (Rodríguez *et al.*, 1999 y 2001; Moro *et al.*, 2002).

En el modelo conceptual diseñado (figura 2), al igual que en el Catastro de España, cada vez que se registra un cambio en un objeto catastral se genera una versión del mismo. En la estructura relacional implementada, las versiones se almacenan como registros independientes en la tabla(s) de la base de datos. La referencia catastral es el identificador único de cada objeto (parcela) que se repite en todas sus versiones. A cada versión se le asocian dos marcas temporales (marcas de intervalo) que definen su período de validez: la fecha de creación y la de cese o desaparición. Estas

Moreno Regidor, P., Iturrioz Aguirre, T. y Martínez Cuevas, S. (2014): “Metodología para deducir relaciones de linaje en el Catastro de España”, *GeoFocus (Artículos)*, nº 14, p. 275-300. ISSN: 1578-5157

fechas son las que figuran en la base de datos utilizada como fuente (SEC), campos FechaAlta y FechaBaja respectivamente, y son tiempos de transacción, relativos a los registros de altas y bajas, con una granularidad temporal de un día. Estos campos, a su vez, se corresponden con los atributos temporales (*beginLifespanVersion*, *endLifespanVersion*) de la parcela catastral definidos en la Directiva Inspire *Data Specification on Cadastral Parcels* (2010).

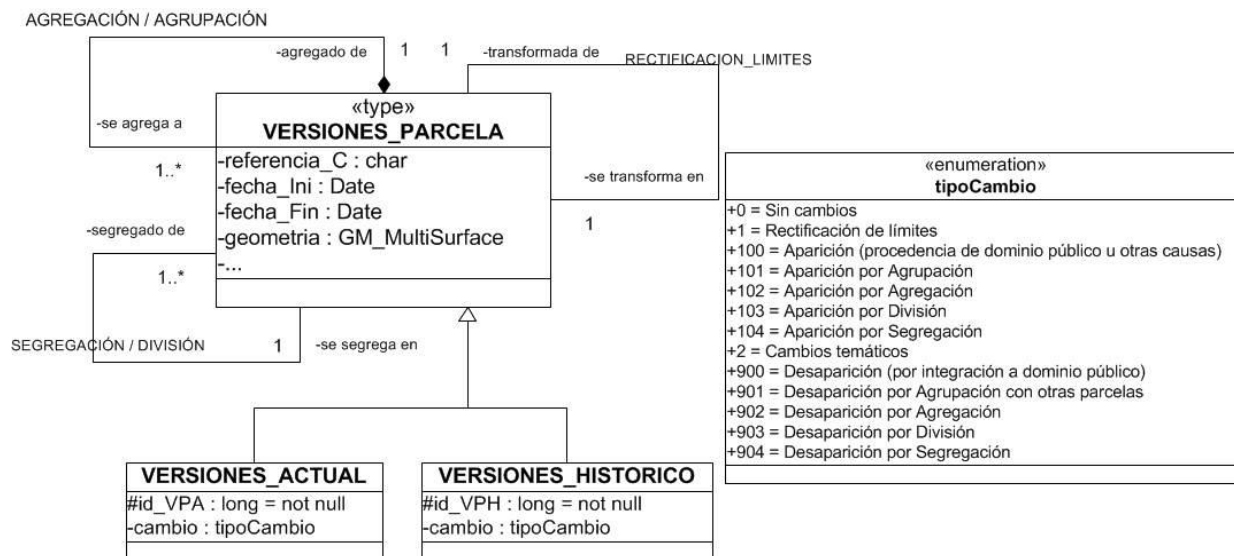


Figura 2. Diagrama UML del modelo conceptual de datos propuesto

El modelo relacional que se propone (figura 3) consta de cuatro tablas: Geometría, Parcelas_Actual, Parcelas_Histórico y Linaje.

En la tabla “Geometría”, cada geometría (recinto que define la forma y posición de la versión de una parcela) tiene un identificador único, *Id_Geom*, que permite su asociación con los registros de las tablas “Parcelas_Actual” y “Parcelas_Histórico”. Si hay varias versiones de una parcela con la misma geometría, estos datos se almacenarían una sola vez en la tabla Geometría.

En las tablas “Parcelas_Actual” y “Parcelas_Histórico” se almacenan los atributos de las versiones de las parcelas catastrales que existen o han existido anteriormente. Ambas tablas están compuestas por cinco campos básicos (*Id_VPA/Id_VPH*: clave primaria, *Referencia_C*: referencia catastral, *Fecha_Ini* y *Fecha_Fin*: fechas de inicio y fin del periodo de vigencia de la versión de una parcela; *Id_Geom*: clave externa con el identificador de la geometría correspondiente a la versión) y otros opcionales, que pueden añadirse según se necesiten, tales como la superficie, el uso u otros datos relacionados con la parcela.

Cuando la versión de una parcela sufre un cambio se elimina de la tabla “Parcelas_Actual” y se inserta en “Parcelas_Histórico”, donde se le asigna un nuevo identificador³ (campo *Id_VPH*) y el campo “Fecha_Fin” se actualiza con la fecha de baja (fecha del cambio). La nueva versión se inserta en “Parcelas_Actual” con la fecha de alta en el campo “Fecha_Ini” y un valor especial en el campo “Fecha_Fin”.

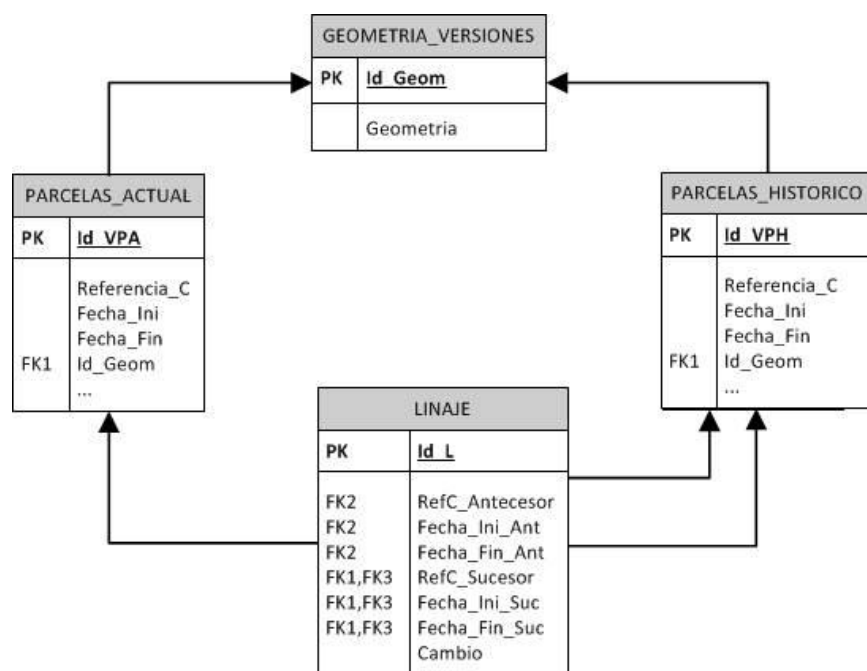


Figura 3. Modelo relacional propuesto

La tabla “Linaje” permite conocer la historia de las entidades del parcelario y se compone de campos relacionados con la identificación y los datos temporales asociados a las versiones antecesoras y sucesoras de las parcelas. Esta tabla podría simplificarse utilizando las referencias a los identificadores internos de las versiones de las parcelas padre e hijo (Id_VPA/Id_VPH), en vez de las referencias catastrales y las fechas de inicio y fin de las correspondientes versiones (RefC_Antecesor, F_Ini_Ant, F_Fin_Ant; RefC_Sucesor, F_Ini_Suc, F_Fin_Suc). No obstante, dado que los primeros son identificadores internos gestionados por el sistema y que sus valores pueden repetirse en las dos tablas, se descarta utilizar esta estructura simplificada.

El método que se propone en este trabajo para deducir las relaciones de linaje solo obtiene las correspondientes a los procesos de agregación y segregación. Las relaciones de agrupación y división se podrían generar fácilmente modificando el algoritmo propuesto. Los cambios correspondientes a aparición y desaparición se pueden obtener mediante secuencias de consultas SQL a las tablas “Parcelas_Actual” y “Parcelas_Histórico”.

5. Método para deducir relaciones de linaje (agregaciones y segregaciones)

Básicamente existen dos métodos (Song y Yang, 2013) para el seguimiento de la historia o evolución de los objetos catastrales: uno se basa en el uso de técnicas de análisis espacio-temporal para obtener información sobre el linaje de una o varias parcelas de interés (Arriaga, 2003; Zhang, 2006), y el otro consiste en usar vínculos temporales para registrar explícitamente las relaciones de

parentesco padre-hijo entre parcelas. En este trabajo se propone un método, basado en superposiciones espacio-temporales y sentencias SQL, para detectar las relaciones de linaje, de tipo agregación y segregación, en los datos de acceso público del Catastro de España (SEC). La información obtenida puede almacenarse en la tabla de linaje propuesta en el apartado anterior. El método podría extenderse fácilmente para detectar las relaciones del tipo agrupación y división de parcelas. No obstante, dado que estas relaciones son minoritarias en nuestro Catastro, se ha optado por centrar el estudio en los casos de agregación y segregación.

A continuación se describen los diez procesos básicos de los que consta el algoritmo. Cada uno de ellos implica el uso de una o varias funciones de análisis espacial y/o consultas SQL. La Figura 4 muestra un diagrama de flujo general del algoritmo.

(1) Superposición espacial del conjunto de versiones v_{ij} , $\forall 1 \leq j \leq n$, correspondientes a cada parcela p_i , siendo n el número de versiones existentes y v_{ij} la versión j de p_i . De las versiones superpuestas se puede calcular la superficie máxima y mínima, $\text{Area_máx}_{v_i} = \text{Max}(\text{Area}_{v_{ij}})$ y $\text{Area_mín}_{v_i} = \text{Min}(\text{Area}_{v_{ij}}) \mid \forall 1 \leq j \leq n$, siendo v_i el término que designa todas las versiones de una parcela p_i .

Cuando la diferencia entre los dos valores anteriores es superior a un cierto umbral se considera que la parcela puede haber sufrido un proceso de agregación o segregación. El umbral utilizado, siguiendo las indicaciones que figuran en el Real Decreto 417/2006, es el 20% del área mínima. No obstante, es conveniente revisar este valor para cada conjunto de datos catastrales.

$$(\text{Area_máx}_{v_i} - \text{Area_mín}_{v_i}) \geq (\text{Area_mín}_{v_i} * 0.20)$$

Si la variación entre la superficie máxima y mínima de las versiones de una parcela es superior a un 20% del área mínima, dichas versiones de la parcela se seleccionan como candidatas, obteniéndose el conjunto de datos A.

(2) Del conjunto A se seleccionan las versiones de mayor y menor superficie como elementos representativos del resto de versiones. Para ello, las versiones de cada parcela p_i se agrupan según su valor de superficie (k), y se calculan las fechas de alta y baja de cada grupo k , que sean las más recientes y las más antiguas, respectivamente:

$$\begin{aligned} (\text{F_Ini_Min}_{v_{ij}})_k &= \text{Min}(\text{F_Ini}_{v_{ij}})_{v_{ij} \in k}; (\text{F_Fin_Min}_{v_{ij}})_k = \text{Min}(\text{F_Fin}_{v_{ij}})_{v_{ij} \in k}; \\ (\text{F_Ini_Max}_{v_{ij}})_k &= \text{Max}(\text{F_Ini}_{v_{ij}})_{v_{ij} \in k}; (\text{F_Fin_Max}_{v_{ij}})_k = \text{Max}(\text{F_Fin}_{v_{ij}})_{v_{ij} \in k} \\ &\forall 1 \leq j \leq n, 1 \leq k \leq 3. \end{aligned}$$

De cada conjunto de versiones de una parcela se podrían obtener dos grupos. No obstante, dado que se ha constatado la existencia de más de un proceso de agregación o segregación en algunas parcelas del área de estudio –generalmente una segunda segregación–, en esta fase se han considerado tres grupos, los correspondientes a la superficie mínima, media y máxima. El número de grupos a obtener depende de la casuística del parcelario a tratar, por lo que se necesita un análisis previo de la zona de estudio. Por último, se obtiene el conjunto de datos B, que incluye un registro con los datos representativos de cada grupo k de versiones de una parcela.

(3) De los registros de cada parcela del conjunto B puede deducirse, comparando las superficies y las fechas de las versiones de sus diferentes grupos, si se trata de un caso de agregación o segregación. Se obtiene el conjunto C.

(4) Del conjunto B se seleccionan las versiones de superficie máxima y media de las parcelas que han sufrido un proceso de segregación. Se obtiene el conjunto B_SEG.

(5) Del conjunto B se seleccionan las versiones de superficie máxima y media de las parcelas que han sufrido un proceso de agregación. Se obtiene el conjunto B_AGR.

(6) Se realiza una intersección espacial de las versiones seleccionadas en B_SEG y B_AGR con el resto de versiones de las parcelas del conjunto de datos. Se obtiene el conjunto D, que incluye la geometría de intersección de cada versión seleccionada ($REFCAT_{pi}$) con cada una de las restantes ($REFCAT_{1_{pi}}$) junto con los atributos de ambas.

(7) Del conjunto D se seleccionan las geometrías cuya superficie sea superior a cierto umbral. En este caso, se ha fijado como umbral que la superficie de intersección (Sup_{Int}) sea superior al 20% de la superficie menor de las versiones de las parcelas intersecadas. Se obtiene el conjunto E.

(8) De cada parcela ($REFCAT_{pi}$) del conjunto E, candidata a sufrir un proceso de agregación (elementos del conjunto B_AGR), se eliminan las versiones de cada una de las parcelas con las que interseca, dejando únicamente la versión que verifica el criterio de ser consecutiva en el tiempo, en este caso anterior o precedente. La versión seleccionada será aquella cuyas fechas de alta y baja coincidan con las de la versión de la parcela candidata con la que interseca. En el resultado de la selección se ordenan secuencialmente en el tiempo las versiones intersecadas, de forma que la versión candidata, resultado de la agregación, aparece como parcela sucesora (RC_{SUC}). Se obtiene el conjunto F_AGR, al que también puede añadirse el tipo de cambio al que corresponden los datos.

(9) De cada parcela ($REFCAT_{pi}$) del conjunto E, candidata a sufrir un proceso de segregación (elementos del conjunto B_SEG), se eliminan las versiones de cada una de las parcelas con las que interseca, dejando únicamente la versión que verifica el criterio de ser secuencial en el tiempo, en este caso consecutiva o posterior. La versión seleccionada será aquella cuyas fechas de alta y baja coincidan con las de la versión de la parcela candidata con la que interseca. En el resultado de la selección se ordenan temporalmente las versiones intersecadas, de forma que la versión candidata, sujeto de una segregación, aparece como parcela antecesora (RC_{ANT}). Se obtiene el conjunto F_SEG, al que también puede añadirse el tipo de cambio al que corresponden los datos.

(10) A partir de los conjuntos F_SEG y F_AGR ya es posible generar la tabla de Linaje con la estructura propuesta (tabla 1).

Tabla 1. Ejemplo de la tabla linaje con datos de agregación/segregación para datos de Catastro de urbana

id_l	r_antecesor	f_ini_ant	f_fin_ant	r_sucesor	f_ini_suc	f_fin_suc	cambio
4	9227513VL5292N	2002-07-30	2003-05-14	9227513VL5292N	2003-05-14	2009-08-07	104
5	9227513VL5292N	2002-07-30	2003-05-14	9227529VL5292N	2003-05-14	2009-08-07	104
7	9027105VL5292N	2002-07-30	2005-06-03	9027105VL5292N	2005-06-03	2006-02-08	102
8	9027106VL5292N	2002-07-30	2005-06-03	9027105VL5292N	2005-06-03	2006-02-08	102

La primera versión del algoritmo descrito se desarrolló con los programas *GeoMedia Professional* v.6.1 y *ArcGIS* v.10, usando como soporte de almacenamiento una base de datos relacional *Access*. Para hacerlo extensible al mayor número de usuarios posible, se ha implementado posteriormente en *PostgreSQL* –con la extensión espacial *PostGIS*– mediante *scripts* PL/PGSQL. La figura 4 muestra un diagrama de flujo general del algoritmo.

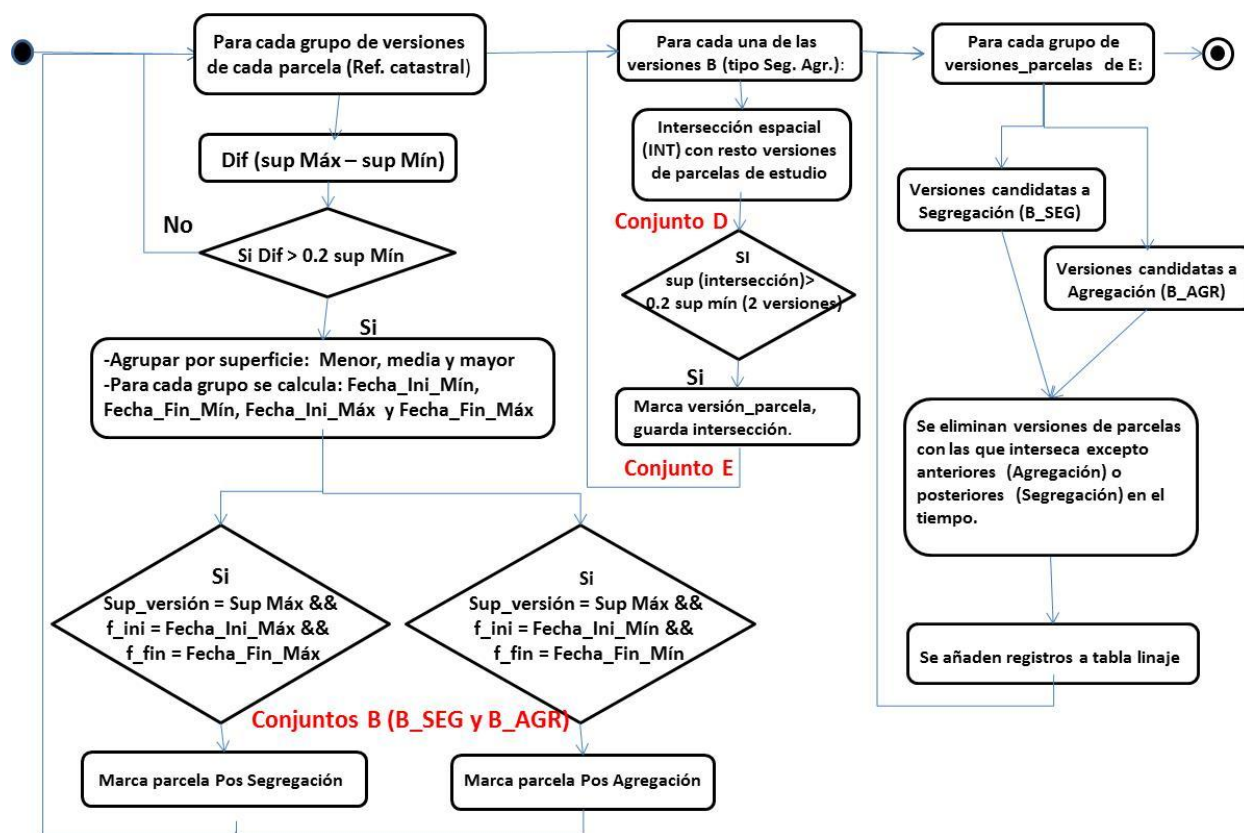


Figura 4. Diagrama de flujo del algoritmo

6. Análisis de resultados

El método se validó inicialmente con datos del Catastro de rústica y urbana de cinco municipios (Bustarviejo, Alovera, Azuqueca de Henares, La Acebeda, Buitrago). El tiempo de ejecución del *script* final es rápido y depende, básicamente, del número total de versiones de parcelas existentes en el Catastro analizado y de las posibles relaciones de parentesco entre ellas (figura 5). Por ejemplo, para un conjunto de 12.596 versiones, la tabla de linaje se generó en 254 segundos (poco más de cuatro minutos) y para otro conjunto de 5.256 versiones, el tiempo empleado fue de 139 segundos (dos minutos y medio).

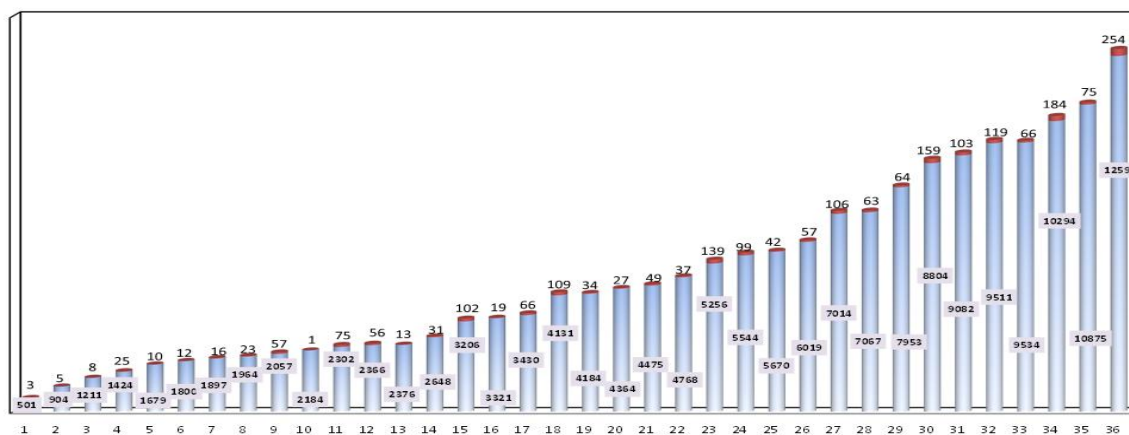


Figura 5. Tiempo de cálculo (en segundos, en color rojo) del linaje vs número total de versiones de parcelas del Catastro (de urbana o de rústica, en color azul) de un municipio

El análisis del linaje de las parcelas realizado en cada municipio (aproximadamente el 60% de los casos de agregación y segregación) reveló la existencia de dos fuentes de error que generaban relaciones incorrectas: (a) las anomalías presentes en los propios datos catastrales y (b) la definición en el algoritmo de criterios y tolerancias que no se ajustan a todas las situaciones particulares del Catastro.

(a) Las anomalías detectadas en los datos catastrales son la causa principal de los errores producidos en la deducción del linaje. Éstas son de diversa índole (se han podido clasificar siete categorías) y pueden estar presentes en las geometrías, en las fechas de registro, así como en la asignación de referencias catastrales de las versiones de parcelas. Por ejemplo, puede darse el caso de que en un período de tiempo existan diferentes versiones asociadas a una sola referencia catastral, o de que una misma parcela varíe de forma y posición, apareciendo en diferentes lugares y con diferentes geometrías, a lo largo del tiempo. Cuando la fuente de error es el propio Catastro, es difícil modificar el algoritmo para corregir los errores de linaje que se producirán.

(b) Para el segundo caso, ha sido posible introducir en el algoritmo un conjunto de sentencias SQL para detectar y corregir o eliminar la mayoría de las relaciones de linaje incorrectas.

El algoritmo, en su segunda versión, ha sido validado con el Catastro de los 14 municipios restantes (Patones, El Escorial, Cercedilla, San Agustín de Guadalix, El Vellón, Rascafría,

Torrelaguna, Miraflores, Soto del Real, Robregordo, Valdemanco, Ajalvir, Talamanca del Jarama, Valdeavero). En la figura 6 se muestra el porcentaje de los casos revisados que son correctos, tanto con la primera versión del algoritmo (color rojo) como con la segunda (color verde). Con esta última versión el número de relaciones correctas aumenta una media de casi el 20%, obteniéndose mejores resultados en el Catastro de urbana que en el de rústica. También se constata que las relaciones de linaje revisadas e incorrectas, aproximadamente un 28% de casos erróneos, se deben mayoritariamente (~80%) a las anomalías presentes en los datos catastrales, que no han podido eliminarse.

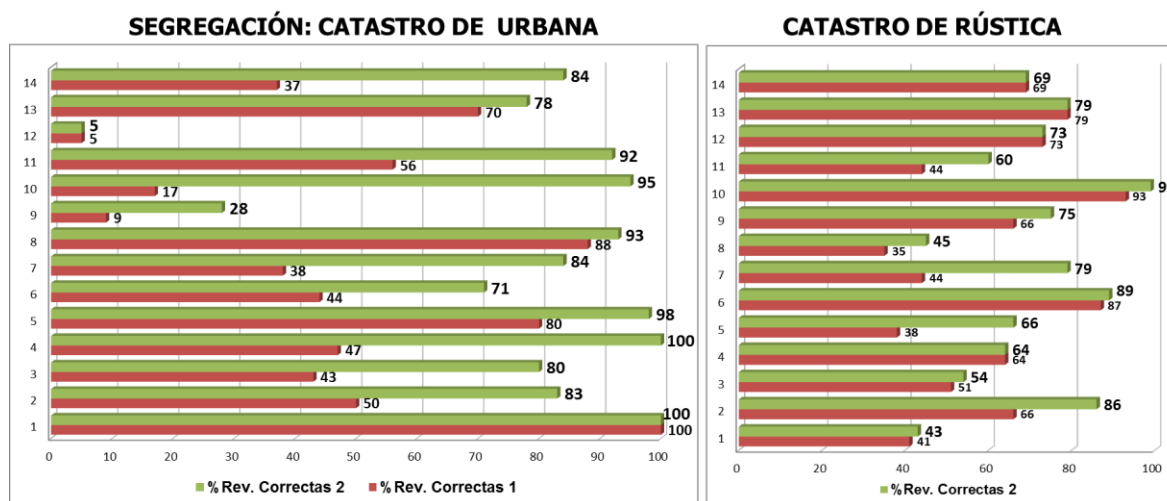


Figura 6. Casos revisados y correctos (en %) utilizando las versiones 1 (color rojo) y 2 (color verde) del algoritmo.

A continuación se describen los criterios empleados en el algoritmo que dan lugar a una serie de errores en la tabla de linaje, algunos de los cuales han podido ser detectados y corregidos. También se clasifican y se describen de forma breve las anomalías detectadas en los propios datos del Catastro.

6.1. Errores debidos al establecimiento de tolerancias

La fijación de tolerancias relativas a la proporción de superficie que debe variar en las versiones de una parcela, para considerar la existencia de un proceso de agregación o segregación, puede producir resultados incorrectos. En principio, según información aportada por la DGC, a efectos del linaje, toda variación superficial inferior al 10% puede ignorarse ya que puede deberse tanto a cambios reales como a procesos periódicos de actualización de la cartografía catastral. Por tanto, y teniendo en cuenta las indicaciones sobre agregaciones y segregaciones del Real Decreto 417/2006, se ha establecido como límite una variación en área superior al 20%, de esta forma, si la diferencia entre las superficies mínima y máxima de las versiones de una parcela es superior al 20% de la superficie mínima, la parcela pasa a ser analizada por el algoritmo.

El establecimiento de esta tolerancia, válida en la mayoría de los casos analizados, da lugar a dos tipos de errores: (a) la generación de relaciones de linaje inexistentes (el algoritmo detecta como

Moreno Regidor, P., Iturrioz Aguirre, T. y Martínez Cuevas, S. (2014): "Metodología para deducir relaciones de linaje en el Catastro de España", *GeoFocus (Artículos)*, nº 14, p. 275-300. ISSN: 1578-5157

agregaciones o segregaciones, cambios que realmente son rectificaciones de la geometría de las versiones), y (b) la no detección de relaciones de agregación/segregación existentes en el Catastro. Se ilustra con el siguiente Ejemplo 1.

Ejemplo 1: Segregación no registrada (Catastro de urbana). En este caso (Figura 7), la variación de superficie ($402-332=70 \text{ m}^2$) entre las versiones V1 y V2 de la parcela 6706627VL5160N es inferior al 20% de la mínima superficie correspondiente a V2 ($332 \cdot 0.2 = 66,4 \text{ m}^2$)

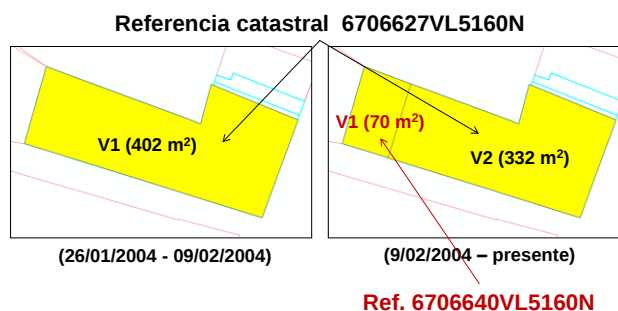


Figura 7. Segregación no registrada por el algoritmo de linaje

Además de lo anterior, el aumento/disminución de la superficie de una parcela por anexión/pérdida de suelo de dominio público –en el Catastro de urbana–, puede dar lugar a la detección de casos de agregación/segregación inexistentes. Este hecho vuelve a constatar la dificultad de establecer tolerancias de área que permitan seleccionar las parcelas candidatas a ser procesadas, o bien rechazar los cambios producidos por variaciones en la geometría. Se ilustra a continuación con el Ejemplo 2.

Ejemplo 2: Segregación inexistente registrada (Catastro de urbana) (figura 8). El aumento/disminución de la superficie de una parcela por anexión/pérdida de suelo de dominio público, puede dar lugar a la detección de casos de agregación/segregación inexistentes. La superficie de la versión V2 (3/4/2007 - 13/4/2007) de la parcela 6210701VL5161S ha disminuido considerablemente respecto de V1, terreno que pasa a dominio público. En el linaje, este cambio se ha registrado como una segregación de esta parcela en sí misma, sin que figure ningún otro descendiente, ya que el terreno que se incorpora al dominio público no se registra como una parcela.

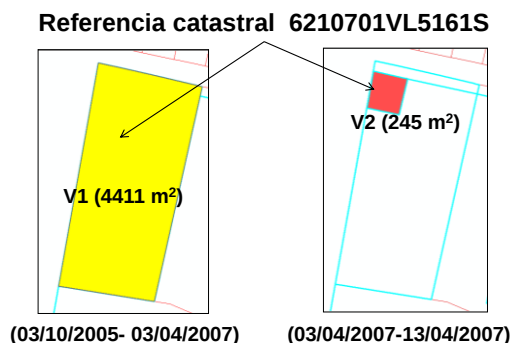


Figura 8. Pérdida de superficie en la versión de una parcela por integración de su terreno al dominio público

Los errores generados por aplicar los dos criterios anteriores pueden detectarse y corregirse fácilmente en la tabla de linaje, tal como se ha realizado en la segunda versión del algoritmo. Bastaría, por un lado, añadir la información del área de las versiones de las parcelas antecesora (padre) y sucesora (hija) y, por otro, diseñar y ejecutar una secuencia de sentencias SQL que eliminen las filas erróneas. El caso de las relaciones de linaje no detectadas, por excluir del proceso las parcelas cuya variación superficial sea inferior al umbral especificado, es el único sin una solución genérica que pueda incorporarse al algoritmo.

6.2. Errores debidos al denominado período de alegaciones (30 días)

Cuando se realiza una ponencia⁴ de valor en un municipio los datos catastrales se actualizan y los propietarios disponen de 30 días para interponer un recurso en caso de acuerdo. Durante este período, en el Catastro se pueden registrar diferentes versiones de una misma parcela según se presenten alegaciones, que no serán definitivas hasta que finalice dicho plazo. También se puede registrar la versión final de la parcela con una fecha de inicio no consecutiva con la anterior, pero dentro de este período.

Dada la casuística existente y empleando la información de la base datos catastral, resulta difícil saber cuál de las versiones generadas en dicho periodo es la definitiva. Por este motivo, el algoritmo diseñado inicialmente consideraba como versiones posteriores a una dada todas las que tuvieran su fecha de inicio en ese intervalo de tiempo. Esto dió lugar a múltiples errores en la tabla de linaje: por ejemplo, en el caso de agregaciones se registraban relaciones espurias entre una versión del hijo con diferentes versiones de cada uno de sus padres, cuando solo debía figurar una de ellas.

El criterio de considerar el período de alegaciones ha sido la causa del mayor porcentaje de los errores de linaje generados por el algoritmo diseñado. En principio, es de suponer que si existen múltiples versiones de una parcela en un intervalo temporal, la última es la definitiva. Si se aplicara este razonamiento –que no siempre es válido–, eliminando del proceso todas las versiones excepto la última registrada en este período, en muchos casos se generarían otros errores de inconsistencia. Tras analizar en detalle este tipo de situaciones, se ha constatado que la mejor solución es obviar este criterio en el algoritmo y considerar, como versión consecutiva a una dada, la versión inmediatamente posterior. Si existieran diferentes versiones posteriores, con idéntica fecha de inicio, se seleccionaría la de fecha de finalización posterior. En caso de existir varias versiones anteriores, con la misma fecha de finalización, se seleccionaría la de fecha de inicio posterior.

6.3. Anomalías detectadas en los datos catastrales

Las anomalías detectadas en los propios datos catastrales están presentes en las geometrías, en las fechas de registro y en la asignación de las referencias catastrales de las versiones de las parcelas. En este trabajo se han identificado siete tipos de anomalías.

Moreno Regidor, P., Iturrioz Aguirre, T. y Martínez Cuevas, S. (2014): "Metodología para deducir relaciones de linaje en el Catastro de España", *GeoFocus (Artículos)*, nº 14, p. 275-300. ISSN: 1578-5157

(a) Existencia, en el mismo período de tiempo, de diferentes versiones asociadas a la misma referencia catastral: cada versión tiene una geometría de forma y posición diferente al resto.

(b) Anomalías en la asignación de referencias catastrales, que consisten en:

(1) La asignación incorrecta de la referencia correspondiente a la parcela que delimita la zona urbana en el Catastro de rústica. Esta debería identificarse por sus últimos cuatro caracteres (xxxxxxxxx9000), pero, durante un período de tiempo, su geometría se asocia tanto a la referencia anterior como a la nueva, por lo que dicha zona urbana aparece duplicada como dos parcelas.

(2) Asignación incorrecta de las referencias (de xxxxxxxxxxx9001 a xxxxxxxxxxx9999) correspondientes a parcelas de tipo "recinto de dominio público o ajuste topográfico (descuentos)"⁵. Por ejemplo se han detectado casos en los que a un descuento de tipo camino, se le ha asociado la referencia de un descuento de hidrografía, situado a varios kilómetros de distancia.

(c) Existencia de procesos mixtos donde se mezclan cambios de difícil clasificación, que no pueden considerarse casos de agregación, segregación, agrupación o división. Este hecho dificulta la detección así como la catalogación del proceso de cambio.

(d) Existencia de versiones de una misma parcela cuyas geometrías varían de forma y posición a lo largo del tiempo. Se muestra un ejemplo en la figura 9.

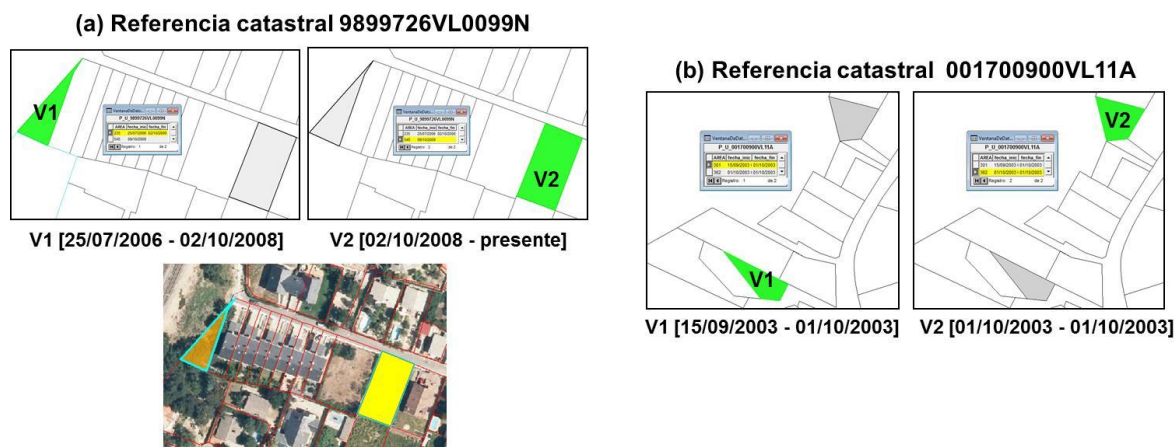


Figura 9. Cambio de ubicación de las versiones de una parcela en Catastro de urbana (izda.) y rústica (dcha.)

(e) Errores en el registro temporal de las versiones de las parcelas. Se deben a un registro incorrecto de las fechas de alta y baja de las versiones. Son de diversos tipos y pueden clasificarse según se describe en la tabla 2. La Figura 10 presenta a modo de ejemplo un caso en el que existen cinco versiones con la misma fecha de inicio (19/7/2010), finalizando además, tres de ellas (V2, V4 y V6), en ese mismo día.

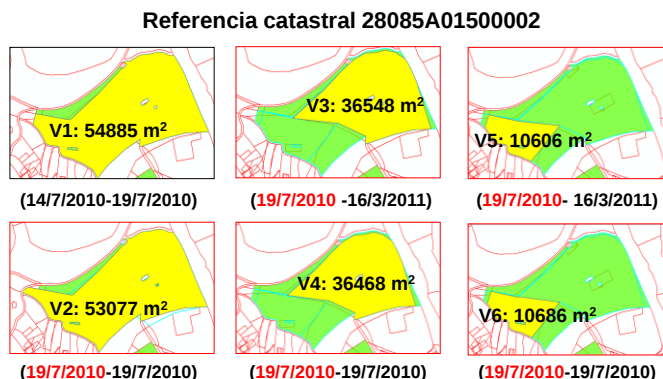


Figura 10. Múltiples versiones de una parcela con idéntica fecha de inicio.

(f) Posibles errores en los procesos de actualización y puesta al día de los datos catastrales. Muchos de los errores de actualización del Catastro suelen manifestarse como “versiones espurias” de las parcelas. Un ejemplo se muestra en la figura 11, donde es muy probable que la versión V3 (del 28/11/2005 al 28/11/2005) de la parcela 9195701VL0099N sea un error de actualización. Dadas las geometrías y fechas de registro existentes en la base de datos, el algoritmo detecta una falsa agregación en la versión V3 de esta parcela, resultado de la unión de su versión anterior (V2) con la V2 de la referencia 9195707VL0099N.

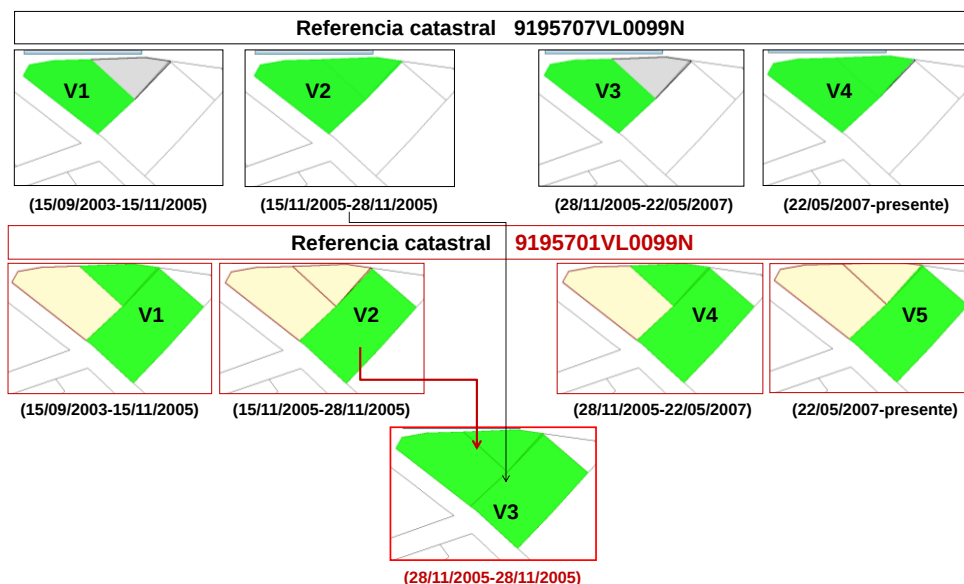
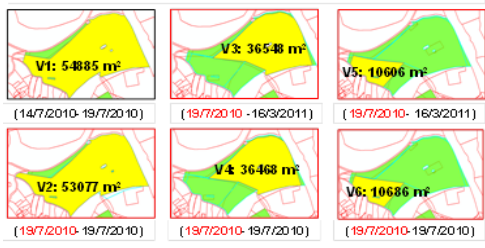


Figura 11. Ejemplo de posible agregación incorrecta

Moreno Regidor, P., Iturrioz Aguirre, T. y Martínez Cuevas, S. (2014): "Metodología para deducir relaciones de linaje en el Catastro de España", *GeoFocus (Artículos)*, nº 14, p. 275-300. ISSN: 1578-5157

Tabla 2. Errores en el registro temporal de las versiones de las parcelas

Errores en el registro temporal de las versiones de una parcela	
Registro de versiones consecutivas (V_i , V_{i+1}) anómalas	
Versiones discontinuas en el tiempo: la fecha de inicio de V_{i+1} no es inmediatamente posterior a la fecha de finalización de V_i	Períodos de V_1 y V_2 de ref: 000200300VK79B V_1 [2/11/2010 - 17/5/2011] V_2 [6/6/2011 - presente]
Versiones con solape temporal: la fecha de inicio de V_{i+1} es anterior a la fecha de finalización de V_i . Esta circunstancia suele darse cuando se produce una actualización masiva de los datos catastrales ¹	Actualización masiva de datos: 8/7/2005 Períodos de V_1 y V_2 de la parcela ref: 7925801VK7972N V_1 [22/11/2002 - 29/7/2005] V_2 [8/7/2008 - 3/7/2008]
Registro de versiones cuya fecha de finalización es anterior a su fecha de inicio	
Períodos para las versiones V_3 y V_6 de la parcela ref: 0205310VL1100N V_3 [20/3/2013 - 5/3/2013] V_6 [20/3/2013 - 6/3/2013]	
Existencia de varias versiones registradas el mismo día en momentos diferentes	
La fecha de inicio y fin del período de vigencia es la misma en todas las versiones: $(F_Ini_{V_i} = F_Ini_{V_{i+1}} = F_Ini_{V_{i+2}} = \dots = F_Fin_{V_i} = F_Fin_{V_{i+1}} \dots)$	Períodos para las versiones V_3 , V_4 y V_5 de la parcela ref: 0894421VL1009S V_3 [15/11/2004-15/11/2004] V_4 [15/11/2004-15/11/2004] V_5 [15/11/2004-15/11/2004]
La fecha de inicio es la misma en todas las versiones: $(F_Ini_{V_i} = F_Ini_{V_{i+1}} = \dots)$	Períodos para las versiones V_2 , V_3 , V_4 , V_5 y V_6 de la parcela ref: 28085A01500002 V_2 [19/7/2010-19/7/2010] V_3 [19/7/2010-16/3/2011] V_4 [19/7/2010-19/7/2010] V_5 [19/7/2010-16/3/2011] V_6 [19/7/2010-19/7/2010]
La existencia de múltiples versiones de una parcela, registradas con la misma fecha de inicio y/o fecha de cese, puede causar diversos errores en la tabla de linaje (véase Ejemplo 5, Anexo I). ¹ Ley 14/2010, de 5 de julio, sobre las infraestructuras y los servicios de información geográfica en España: "Cuando exista cartografía topográfica oficial inscrita en el Registro Central de Cartografía y debidamente actualizada, realizada por alguno de los agentes integrados en el Sistema Cartográfico Nacional, la cartografía catastral correspondiente se actualizará o realizará a partir de ella".	Referencia catastral 28085A01500002  V1: 54885 m² (14/7/2010-19/7/2010) V2: 53077 m² (19/7/2010-19/7/2010) V3: 36548 m² (19/7/2010-16/3/2011) V4: 36468 m² (19/7/2010-19/7/2010) V5: 10606 m² (19/7/2010-16/3/2011) V6: 10686 m² (19/7/2010-19/7/2010)

(g) Cambios en la referencia catastral de las parcelas. Se han detectado varios casos en los que la referencia catastral de una parcela ha sido sustituida por otra a partir de una fecha dada.

Como conclusión de esta sección 6, puede afirmarse que la mayor fuente de error en la generación de relaciones de linaje en el Catastro español es la presencia de anomalías e inconsistencias en los propios datos catastrales, lo que dificulta el diseño de un algoritmo genérico adaptable a todas las situaciones atípicas. En este trabajo, el proceso de validación ha dado una media de casos revisados erróneos inferior al 30%, por lo que el linaje generado puede considerarse fiable en un 70% de los casos. A modo de ejemplo se incluyen tres de las consultas realizadas sobre la tabla de linaje de los municipios de estudio, que permiten conocer la evolución de las parcelas y reconstruir su historia: (1) En un municipio dado, ¿qué versiones de parcelas han experimentado algún proceso de agregación/segregación en el período comprendido entre t_1 y t_2 ? (Figura 12). (2) ¿Cuál es el antecesor(es) de la versión n de la parcela X ? ¿Cuál su relación de parentesco? (Figura 13). (3) ¿Qué descendientes tiene la versión n de la parcela Y , y cuál es la relación de parentesco entre ellos?

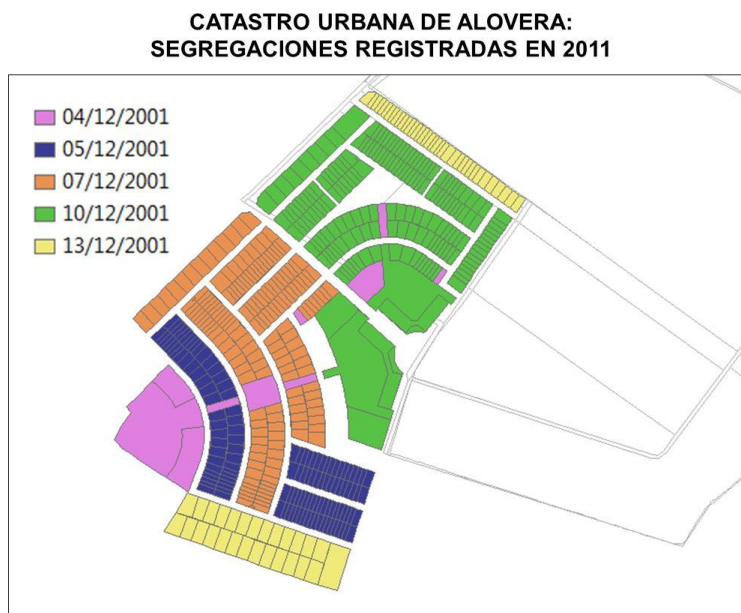
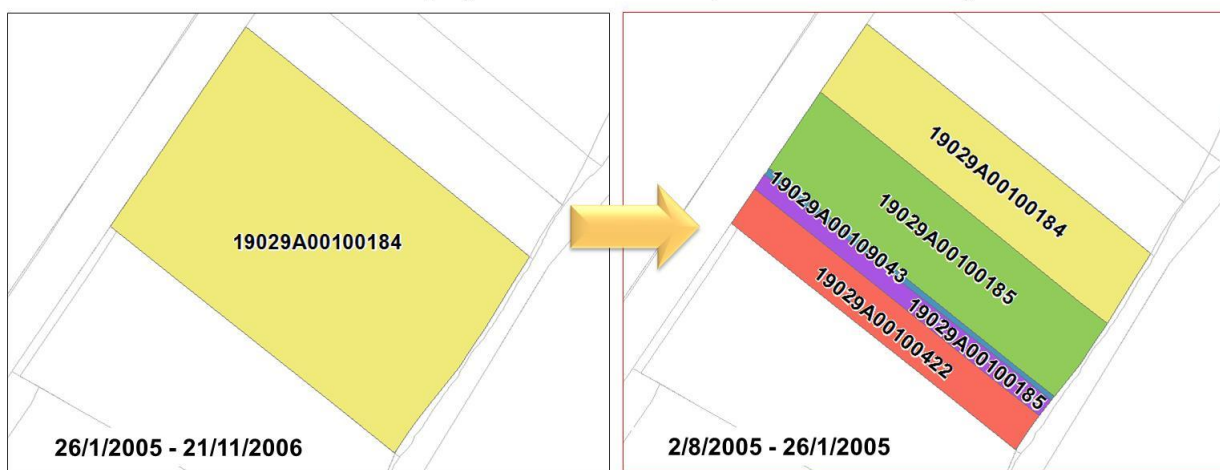


Figura 12. Parcelas del municipio de Alovera (Catastro de urbana) que han experimentado algún proceso de segregación entre enero y diciembre de 2011

CATASTRO RÚSTICA: ANTECESOR(ES) DE LA VERSIÓN V2 (26/1/2005 - 21/11/2006) DE 19029A00100184



id_l	r_antecesor	f_ini_ant	f_fin_ant	r_sucesor	f_ini_suc	f_fin_suc	cambio
177	19029A00100185	2001-08-02	2005-01-26	19029A00100184	2005-01-26	2006-11-21	102
178	19029A00100184	2001-08-02	2005-01-26	19029A00100184	2005-01-26	2006-11-21	102
179	19029A00100185	2001-08-02	2005-01-26	19029A00100184	2005-01-26	2006-11-21	102
180	19029A00109043	2001-08-02	2005-01-26	19029A00100184	2005-01-26	2006-11-21	102
207	19029A00100422	2004-09-09	2005-01-26	19029A00100184	2005-01-26	2006-11-21	102

Figura 13. Antecedentes de la versión V2 de 19029A00100184 (Catastro de rústica) y la relación de parentesco entre ellos (102 = agregación).

7. Conclusiones

Este artículo presenta un método para obtener las relaciones de linaje más frecuentes (agregación y segregación) en los datos catastrales de España de acceso público (SEC), y propone un prototipo de estructura relacional para el almacenamiento y la gestión histórica de dichos datos. El proceso para deducir las relaciones de linaje se basa en superposiciones espacio-temporales, entre las diferentes versiones de las parcelas, y secuencias de sentencias SQL. Con respecto al método propuesto: (1) ha sido validado en el Catastro de rústica y urbana de 19 municipios de las provincias de Madrid y Guadalajara; (2) la validación ha permitido detectar un conjunto de anomalías en los datos catastrales que son la causa de aproximadamente el 80% de los errores presentes en el linaje revisado. El restante 20% de los errores del linaje proceden de los criterios y tolerancias aplicados en el algoritmo y que pueden adaptarse, con mayor o menor éxito, a cada caso de estudio.

Para conocer la evolución histórica del Catastro, en términos de parentesco entre los objetos catastrales, es necesario utilizar información sobre el linaje. Cuando las relaciones de linaje no se almacenan explícitamente en el sistema de información, la alternativa consiste en recurrir, cada vez que sea necesario, a una secuencia de operaciones de superposición espacio-temporal. Hasta la fecha no se ha encontrado ninguna publicación que explique en detalle el método de análisis espacio-temporal necesario para deducir la información sobre el linaje.

En este trabajo se demuestra que dicho método ha de tener en cuenta las particularidades del Catastro objeto de estudio (procesos de actualización de datos, períodos de ponencias de valores...), y que la validez y precisión en sus resultados están condicionadas por el estado y fiabilidad de los datos catastrales originales. En el Catastro de España se han detectado siete tipos de anomalías en los datos, que representan irregularidades respecto de las normas y legislación catastral. Estas anomalías, que se han clasificado y tipificado en este trabajo, generan una serie de errores en el linaje, que no son fáciles de detectar ni corregir con procesos genéricos que puedan adaptarse a todas las situaciones particulares que se presentan.

Además, se precisa contrastar el método con datos reales para, por un lado, adecuar los criterios y tolerancias usados en el algoritmo a las características específicas del Catastro analizado y, por otro lado, diseñar un conjunto de secuencias SQL que permitan eliminar o reducir al máximo los errores generados. En nuestro caso, se diseñó una segunda versión del algoritmo que incorporaba sentencias SQL de corrección de un tipo de errores, los originados por la aplicación de una serie de criterios y tolerancias que no se ajustaban adecuadamente a los datos catastrales de los municipios analizados. De ahí la importancia de validar siempre el algoritmo diseñado con una muestra del conjunto de datos de estudio, ya que probablemente sea preciso modificar, en cada caso, algunos de esos criterios y tolerancias.

Una vez detectados y corregidos la mayoría de los errores debidos a los criterios usados en el algoritmo, se han revisado un 60% de las relaciones de agregación y segregación de la tabla de linaje de cada uno de los municipios de estudio y, se ha comprobado, que un 70% de ellas son válidas.

Como líneas futuras de investigación, se propone depurar y mejorar el algoritmo para que también pueda obtener las relaciones de linaje de agrupación y división, así como detectar automáticamente las relaciones atípicas por estar derivadas de datos catastrales probablemente anómalos. También se propone crear una interfaz de trabajo que permita al usuario especificar los valores de los criterios y tolerancias que son causa de error. La posibilidad de ejecutar el algoritmo con diferentes valores permitirá ajustar más rápidamente dicho algoritmo a las características específicas de los datos.

Por último, indicar que el método se desarrolló inicialmente con los programas GeoMedia Professional v.6.1 y ArcGIS v.10. Posteriormente, se ha implementado en PostgreSQL -con la extensión espacial PostGIS- mediante scripts PL/PGSQL, cuya ejecución es sumamente rápida. El tiempo invertido depende básicamente del número total de versiones existentes en el Catastro analizado y de las posibles relaciones de parentesco entre ellas.

Bibliografía

Arcieri, F.; Cammino, C.; Nardelli, E.; Talamo, M. y Venza, A. (1999): "The Italian cadastral information system: A real-life spatio-temporal DBMS" en M. H. Bohlen, C. S. Jensen & M. O. Scholl (Eds.): *Spatio-Temporal Database Management*. Berlin, Springer-Verlag, pp. 79-99.

Moreno Regidor, P., Iturrioz Aguirre, T. y Martínez Cuevas, S. (2014): "Metodología para deducir relaciones de linaje en el Catastro de España", *GeoFocus (Artículos)*, nº 14, p. 275-300. ISSN: 1578-5157

Arriaga Egüés, C. (2003): "SIG temporal: gestión histórica de un sistema de información catastral", *Mapping: Map and SIG Consulting*, 85, pp. 32-39.

Bohlen, M. H. (1995): "Temporal database system implementations", *Sigmod Record*, 24, pp. 16-30.

Boletín Oficial del Estado, núm. 97 (2006): *Real Decreto 417/2006 por el que se desarrolla el texto refundido de la Ley del Catastro Inmobiliario, aprobado por Real Decreto Legislativo 1/2004, de 5 de marzo*. [Consulta: 27-09-2014]. Disponible en:
<http://www.boe.es/boe/dias/2006/04/24/pdfs/A15650-15674.pdf>.

Choi, J. et al. (2008): "Innovations in individual feature history management. The significance of feature-based temporal model", *Geoinformatica*, 12, pp. 1-20 (doi: 10.1007/s10707-007-0019-y).

Claramunt, C. y Theriault, M. (1996): "Towards semantics for modelling spatio-temporal processes within GIS", *Advances in GIS Research I: Proceedings of the Seventh International Symposium on Spatial Data Handling (SDH'96)*, pp. 47-63.

Cömert, Ç. y Alkan, M. (2004): "The design and development of a temporal GIS for cadastral and land title data of Turkey", *Commission IV, WG IV/1* (Karadeniz Technical University, Geodesy and Photogrametry Department).

Frank, A.; Raper, J. y Cheylan, J. P. -eds.- (2001): *Life and Motion of Socio-Economic Units. GISDATA 8*. London, Taylor & Francis.

Frank, R. y Kemp, Z. (2001): "Integrated spatiotemporal analysis for environmental applications", en Peter J. Hall (Ed.), *Spatial information and the environment, Innovations in GIS 8*. London, Taylor & Francis, pp. 121-134.

Gantner, F.; Bettina Waldvogel, B.; Meile, R. y Patrick Laube, P. (2013): "The basic formal ontology as a reference framework for modeling the evolution of administrative units." *Transactions in GIS*, 17 (2), pp. 206-226

Hunter, G. J. y Williamson, I. P. (1990): "The development of a historical digital cadastral database", *International Journal Geographical Information Systems*, 4, 2, pp. 169-179 (doi: 10.1080/02693799008941538).

INSPIRE Thematic Working Group Cadastral Parcels (2010): "D2.8.I.6 INSPIRE Data Specification on Cadastral Parcels – Guidelines".

Kauppinen, T. et al. (2008): *Creating and using geospatial ontology time series in a semantic cultural heritage portal*. In: Proceedings of the ESWC 2008, Tenerife, Spain, Springer-Verlag

Lacasta, J.; Lopez-Pellicer, F.J.; Florczyk, A.; Zarazaga-Soria, F.J. y Nogueras-Iso, J. (2014): "Population of a spatio-temporal knowledge base for jurisdictional domains." *International Journal of Geographical Information Science*, 28, 9, pp. 1964-1987. DOI:10.1080/13658816.2014.911412

Langran, G. (1992): *Time in Geographic Information System*. London, UK: Taylor & Francis Ltd.

Lemmen, C. y Van Oosterom, P. (1995): "Efficient and automatic production of periodic updates of cadastral maps" en *JEC'95, Joint European Conference and Exhibition on Geographical Information*. The Netherlands: Hague, pp. 137-142.

Moreno Regidor, P., Iturrioz Aguirre, T. y Martínez Cuevas, S. (2014): "Metodología para deducir relaciones de linaje en el Catastro de España", *GeoFocus (Artículos)*, nº 14, p. 275-300. ISSN: 1578-5157

Moro, M. M.; Edelweiss, N. y dos Santos, C. S. (2002): "Temporal versions model" en *XV Workshop de Teses e Dissertações-CTD*, pp. 33-37.

Nan, L.; Renyi, L.; Guangliang, Z. y Jiong. X. (2006): "A spatial-temporal system for dynamic cadastral management", *Journal of Environmental Management*, 78, pp. 373-381.

Van Oosterom, P. (1997): "Maintaining consistent topology including historical data in a very large spatial database" en *Auto-Carto 13: Proceedings of the International Symposium on Computer-Assisted Cartography*. Seattle, WA, pp. 327-336.

Van Oosterom, P. y Lemmen, C. H. J. (2001): "Spatial data management on a very large cadastral database", *Computers Environment and Urban Systems*, 25 (4-5), pp. 509-528.

Peerbocus, M. A.; Bauzer-Medeiros, C.; Jomier, G. y Voisard, A. (2001): "Documenting changes in a spatiotemporal database", en *XVI Brazilian symposium on database*, Brazil, 10-24.

Peerbocus, A. y Jomier, G. (2004): "The management of the cadastral evolution using documented cadastral plans", *Computers, Environment and Urban Systems*, 28, pp. 487-509.

Rodriguez, L.; Ogata, H. y Yano, Y. (1999): "TVOO: A Temporal Versioned Object-Oriented data model", *Information Sciences*, 114 (1-4), pp. 281-300.

Rodriguez, L.; Ogata, H. y Yano, Y. (2001): "A temporal versioned object-oriented database schema", *Computers and Mathematics with Applications*, 41 (1-2), pp. 177-192.

Snodgrass, R. T. (1992): "Temporal databases" en Frank, A. U.; Campari, I. & Formentini, U. (Eds.): *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*. Berlin, Springer-Verlag, pp. 22-64.

Spéry, L.; Claramunt, C. y Libourel, T. (1999): "A lineage metadata model for the temporal management of a cadastre application" en *Proceedings of the International Workshop on Spatio-temporal Models and Language SDTDML'99*. Florence: The IEEE Computer Society, pp. 466-474.

Spéry, L.; Claramunt, C. y Libourel, T. (2001): "A Spatio-Temporal Model for the Manipulation of Lineage Metadata", *GeoInformatica*, vol. 5, nº 1, pp. 51-70.

Stefanakis, E. (2003): "Modelling the history of semi-structured geographical entities", *International Journal of Geographical Information Science*, 17, 6, pp. 517-546.

Song, W. y Yang, X. (2013): "A spatio-temporal cadastral data model based on space-time composite model" en *Geoinformatics, 21st International Conference on*. IEEE, pp. 1-4 (doi: 10.1109/Geoinformatics.2013.6626047).

Vrana, R. (1989): "Historical data as an explicit component of land information systems", *International Journal of Geographic Information Systems*, 3, pp. 33-49.

Worboys, M. F. (1992): "A model for spatio-temporal information" en *Proceedings of 5th International Symposium on Spatial Data Handling*. Charleston, South Carolina, USA, pp. 602-611.

Yuan, M. (1999): "Use of a three-domain representation to enhance GIS support for complex spatiotemporal queries", *Transactions in GIS*, 3, 2, pp. 137-159.

Zhang, N. (2006): *Spatio-temporal cadastral data model: geo-information management perspective in China*. M.Sc. Thesis, ITC, Enschede, the Netherlands.

Moreno Regidor, P., Iturrioz Aguirre, T. y Martínez Cuevas, S. (2014): “Metodología para deducir relaciones de linaje en el Catastro de España”, *GeoFocus (Artículos)*, nº 14, p. 275-300. ISSN: 1578-5157

¹ <http://www.sedecatastro.gob.es/>

² Cada versión de un plan catastral representa un estado concreto de la zona de estudio.

³ Cuando una versión de la tabla “Parcelas_Actual” pasa a formar parte del histórico, se inserta en “Parcelas_Histórico” con un nuevo valor en el campo “Id_VPH”, perdiendo el valor que tenía asignado anteriormente en “Id_VPA”.

⁴ Una ponencia de valor colectiva (general, parcial o simplificada) es el documento que recoge los criterios, módulos de valoración y demás elementos precisos para llevar a cabo la determinación del valor catastral.

⁵ (Circular 07.04/06, de 9 de junio, sobre criterios de asignación y modificación de la referencia catastral de los bienes inmuebles): “La referencia catastral de estos recintos de dominio público en suelo rústico, que dejarán de denominarse “descuentos”, tendrá la misma estructura que actualmente tienen los bienes inmuebles rústicos, y se les asignará un número de parcela comprendido entre 09001 y 09999”.

Es preciso diferenciar los recintos de dominio público de los corrientemente conocidos como “descuentos topográficos”, denominación que se evitará en adelante, sustituyéndola por la de ajustes topográficos. Como tales se entenderán aquellas parcelas ficticias que se añaden en la cartografía catastral para que el suelo rústico rellene todo el municipio, debido a limitaciones del sistema.