

METODOLOGÍAS PARA LA ACTUALIZACIÓN AUTOMÁTICA DE LAS BASES DE DATOS DE OCUPACIÓN DEL SUELO: EL CASO DEL PROYECTO CORINE LAND-COVER

A. AROZARENA*, E. BAYARRI** & A. SEBASTIÁN*

* ÁREA DE TELEDETECCIÓN DEL INSTITUTO GEOGRÁFICO NACIONAL. C/ IBÁÑEZ DE ÍBERO 3, 28003-MADRID. E-MAIL: TELEDETE_IGN.ESPANA*MAD.SERVICOM.ES

** SERVICIO REGIONAL DEL INSTITUTO GEOGRÁFICO NACIONAL EN ARAGÓN. C/ COSO 55, 5º Y 9º. 50001-ZARAGOZA. E-MAIL: EBAYARRI*DATABASEDM.ES

RESUMEN

Esta comunicación plantea uno de los principales problemas relacionados con la cartografía de ocupación del suelo: el de su actualización. Tras analizarse los distintos enfoques existentes en la Unión Europea, se presentan algunos de los resultados y conclusiones obtenidos de los proyectos Europeos en ejecución bajo la tutela del European Topic Centre on Land Cover.

P.C.: Cartografía de Ocupación del Suelo. Actualización. Fotointerpretación. Clasificación. Generalización raster. Agencia Europea del Medio Ambiente.

SUMMARY

In this paper are presented some of the problems that the updating of land cover/use maps implies. The main approaches adopted are explained. Following, some findings from an European project that aimed at locating and assessing the land cover/use map updating experiences are given. Finally there is a list of conclusions

K.W.: Land cover/use cartography. Updating. Photo-interpretation. Classification. Raster generalization. Environmental European Agency.

INTRODUCCIÓN Y OBJETIVOS

El desarrollo de los sistemas informáticos, el avance de la tecnología espacial y las continuas caídas de los precios de equipos, programas y datos han conducido a que la técnica cartográfica haya sufrido tal revolución que podamos considerar como obsoletos los procesos clásicos y a la mayoría de los productos generados en ellos. El mapa actual es mucho más que sólo un dibujo sobre papel con el que representar la realidad ya que, al entenderse como base de datos geográficos, es, ante todo, información. Entre las características concretas de las bases cartográficas numéricas hay que citar su complejidad (información geométrica, topológica y temática), su alto grado de organización (datos codificados y estructurados en niveles o capas) y la facilidad de acceso (fiabilidad en el intercambio entre formatos; gran variabilidad de soportes). La cartografía temática en general, y la de ocupación del suelo en particular, no han sido ajenas a esta revolución tecnológica.

Esta comunicación emplea el término ocupación de suelo al objeto de englobar dos conceptos, *cobertura del suelo* (Land Cover) y *usos del suelo* (Land Use), frecuentemente

diferenciados en la bibliografía especializada pero de muy difícil separación desde un punto de vista práctico. Por cobertura entendemos a los materiales que cubren la superficie de la tierra en un área determinada en tanto que el término uso hace referencia a la utilización que el hombre hace de dichos materiales.

Disponer de información bien inventariada (fiabilidad y actualidad) sobre la ocupación del suelo permite analizar los datos de partida, verificar las hipótesis de base y generalizar los resultados finales (AGUILÓ et al., 1984). Pese a esto, y a pesar de que las exigencias de calidad son cada vez mayores, la cartografía temática sigue presentando deficiencias ya que se piensa que su precisión geométrica puede ser menor que la de los mapas topográficos.

En el caso concreto de la cartografía de ocupación, como consecuencia de la importancia que tiene la fecha de captura de la información temática, se presenta el problema añadido de la necesidad de actualización. Las bases de datos sobre ocupación serán válidas sólo en la medida de que sean actuales los datos que contienen. Dado que las alteraciones de la cubierta vegetal implican variaciones importantes de la ocupación, es necesario diseñar mecanismos para la puesta al día de los datos de ocupación así como definir la frecuencia de actualización.

El objeto de esta comunicación es el de plantear los problemas asociados a la actualización de las bases cartográficas sobre ocupación del suelo. Ya que el proceso de producción de cartografía está totalmente automatizado, cabe preguntarse si pueden diseñarse mecanismos automáticos de actualización de este tipo de bases, cual es la precisión geométrica y la calidad temática de los productos derivados, cual es el coste económico de la producción de datos fiables, cual es la dependencia tecnológica que conllevan tales procedimientos, etc.

METODOLOGÍAS O ENFOQUES DE ACTUALIZACIÓN

La necesidad de actualización de la cartografía de ocupación del suelo es una preocupación del Instituto Geográfico Nacional (IGN) que comparten todos los países de la Unión Europea. Es por ello que la Comisión, a través de su organismo investigador (JRC) y de la Agencia Europea de Medio Ambiente, ha buscado los apoyos técnicos, económicos y administrativos necesarios para abordar la situación. El interés despertado es tal que se están investigando dos planteamientos conceptuales diferentes: por un lado se analizan las posibilidades que ofrece la foteointerpretación (métodos vectoriales) y, por otro, se investiga la viabilidad de los métodos automáticos basados en las técnicas de clasificación y generalización (métodos raster).

La anterior división en técnicas vectoriales y raster es, como todo intento de clasificación, bastante teórica ya que las experiencias que se están realizando en este campo podrían definirse como métodos mixtos, con los que se pretende aprovechar las ventajas de cada grupo.

Conviene matizar que las denominadas técnicas automáticas no lo son en su totalidad pues dependen, en mayor o menor grado, de la intervención humana. Sin embargo, la aparición de los Sistemas de Información Geográfica (SIG) orientados a objeto y el desarrollo de algoritmos basados en el conocimiento (OOSTEROM, 1996), han abierto nuevas líneas de investigación que apuntan hacia lo que será la automatización total del proceso de actualización.

Por metodología vectorial se entiende al conjunto de procedimientos que, basándose en los criterios de creación de la base de datos CORINE Land-Cover, pueden ser adoptados para la actualización de cualquier base vectorial. En esencia consiste en la interpretación (con ayuda de cartografía digital y de imágenes de satélite) de los cambios producidos desde que se creó la base de datos original (JRC 1992 y 1993). El proceso se estructura como sigue:

- *Compilación de información.* En especial la base de datos original, las imágenes que se emplearon para su creación, una imagen reciente y los posibles metadatos existentes.

- *Detección de errores y corrección de la base de datos original.* Detección y corrección de errores geométricos (cambio de huso, case de hojas, cases de polígonos, etc.) y errores semánticos (polígonos contiguos con el mismo código, etiquetas duplicadas, etc.).

- *Tratamiento digital de la nueva imagen.* Corrección geométrica y radiométrica.

- *Integración en el sistema de la información disponible.* Antes de la actualización, el sistema debe estar provisto de la información de base necesaria (base de datos original depurada de errores, imagen vieja e imagen nueva corregida geométrica y radiométricamente) así como de la información exógena de interés (topografía digital, bases de datos temáticos ...).

- *Actualización.* Tras la detección de cambios (comparación en pantalla de la base de datos existente y su imagen correspondiente con la imagen reciente) se procede a actualizar en el sentido estricto del término. A la vista de los cambios puede ser necesario modificar líneas o borrar o crear nuevos polígonos de manera que, tras cada modificación, será preciso reconstruir la topología de la base de datos siempre y cuando el sistema no sea capaz de hacerlo de forma automática o no se trabaje en modo espagueti (reconstrucción de toda la topología al final de la actualización).

- *Validación de la base de datos actualizada.*

El método permite el empleo, a modo de información exógena, de los resultados de clasificaciones automáticas de imágenes ya que, si se logra encontrar una buena correspondencia entre clases espectrales y los distintos tipos de cubierta, se contará con una información de gran valor para interpretar los cambios. Pese a esto, JRC recomienda no dedicar mucho tiempo a las clasificaciones, no sólo por su incapacidad para discriminar usos del suelo sino, sobre todo, porque CLC tiene una orientación claramente vectorial.

- *Metodología raster.* Existen experiencias de producción del CLC a partir de clasificación supervisada y posterior generalización automática. Los fundamentos teóricos en que se apoyan pueden encontrarse en LI & OPENSHAW (1992) que diseñan algoritmos para la simplificación de fronteras entre polígonos, en GOFFREDO et al. (1993) donde se exponen los principios de los procesos de generalización de datos vectoriales para generar bases CORINE, en SCHYLBERG (1992) que describe el proceso básico de generalización de objetos cartográficos en ambientes raster, en MULLER y ZESHEN (1992) donde se describen procedimientos para generalizar a partir de múltiples polígonos de tamaño muy pequeño o en MONMONIER (1983) donde se justifica la ventaja de emplear estrategias raster para representación a pequeñas escalas.

En Holanda (LINDEN et. al., 1990) se ha llevado a cabo un experimento para generar CORINE Land-Cover a partir de clasificación supervisada de una imagen Landsat, reclasificación y, mediante métodos vector, eliminación de polígonos menores de 1 ha. y posterior fusión de teselas hasta generar recintos de tamaño mínimo de 25 ha.

En el Reino Unido hay un gran interés en la obtención del CLC mediante técnicas semiautomáticas ya que existe cartografía de la totalidad del país a partir de una clasificación automática de datos Landsat TM. Este mapa (LCMGB), obtenido mediante clasificación supervisada de imágenes de verano e invierno y posterior empleo de máscaras para mejorar los resultados, es una base de datos raster que identifica 25 tipos de ocupación con unidad cartografiable mínima de 0,125 ha. (FULLER & GROOM, 1993;). Su existencia condujo a plantearse la viabilidad de emplear procedimientos de generalización espacial (complementado con el apoyo de información exógena) al objeto de alcanzar las especificaciones CLC. En este sentido, se han desarrollado técnicas semiautomáticas para generar los datos CORINE a partir de las clases LCMGB sobre píxeles de 25 m., remuestreo a tamaño de pixel de 100 m. y filtrado mediante métodos raster a polígonos de no menos de 25 ha. (FULLER, GROOM & JONES, 1994).

En Suecia se han ejecutado experiencias raster similares, proponiéndose la producción del CLC mediante un proceso escalonado de generalización desde píxeles de 25 m. (procedentes de clasificación automática de datos TM) a parcelas de 5 ha y, desde éstas, a recintos de 25 ha. (PAULSSON, 1994).

En Finlandia (JAAKKOLA, 1995) se han realizado experiencias relacionadas con la conversión de datos existentes sobre uso del suelo. La propuesta implica el empleo de una gran variedad de datos de entrada en formatos variados, combinación de todos ellos en una cobertura única y generalización automática de ésta hasta alcanzar las especificaciones propias del CORINE Land-Cover. En este momento el proceso es totalmente operativo, implicando un tiempo de proceso de apenas una hora para la generalización de los datos del área de estudio (1200 Km²). Una ventaja añadida al procedimiento es su viabilidad para generar distintas bases con resoluciones diferentes obteniéndose muy buenos parámetros en control de calidad.

- *Ventajas e Inconvenientes.* Dado el estado actual de automatización de cada método, no es posible inclinarse por uno u otro. En realidad, cada una de las soluciones presenta ventajas e inconvenientes que serán quienes inclinen la selección hacia uno u otro procedimiento.

Entre las ventajas de los métodos vectoriales cabe resaltar que se trata de aproximaciones intuitivas, muy ligadas al proceso humano, para las que existen buenos profesionales; la facilidad en la formación de técnicos; el que, por basarse en metodologías existentes, suponen menores tiempos de aprendizaje o que impliquen menores inversiones. Entre los inconvenientes hay que citar la mayor subjetividad de los resultados; los problemas que, debidos a dicha subjetividad, se presentan entre las zonas de contacto; la dificultad de una buena discriminación de las clases mixtas; la acumulación de errores a lo largo de las distintas fases o el implicar un mayor cansancio (menor concentración) para los operadores.

Los procesos raster presentan, por contra, una serie de ventajas entre las que cabe destacar su menor subjetividad; mayor rapidez; facilidad para la automatización; menor fatiga, mayor homogeneidad y coherencia de los resultados; mayor precisión geométrica; facilidad de controlar la calidad temática; viabilidad para generar productos múltiples, capacidad para efectuar actualizaciones periódicas; menores costes (a largo plazo) de los procesos de actualización así como el favorecer la comparación de resultados entre países o regiones.

EXPERIENCIAS EUROPEAS DE ACTUALIZACIÓN DE CARTOGRAFÍA

El IGN y el Centro Nacional de Información Geográfica de Portugal firmaron en 1996 un contrato para la ejecución de una serie de proyectos en el marco del Centro Temático Europeo Land-Cover. Dado que uno de dichos proyectos pretendía localizar y valorar las experiencias europeas en actualización de cartografía de ocupación del suelo (ETC/LC IGN, 1996) se envió un cuestionario a todos aquellos organismos relacionados con la cartografía temática. Analizados los resultados, a quienes afirmaron haber actualizado cartografía de ocupación de suelo se les envió un nuevo cuestionario: por un lado para obtener una descripción general de la cartografía actualizada (extensión superficial, escala, unidad mínima cartografiada, número y tipo de clases en la leyenda, frecuencia de actualización, número de versiones y formas de almacenarlas, etc.) y, por otro, para conocer las herramientas empleadas, la metodología seguida y los resultados obtenidos. Si bien el 68 % de los organismos respondieron a la petición, se constató que sólo el 39% de las respuestas se referían a actualizaciones de cartografía de ocupación del suelo propiamente dicha.

Se han realizado experiencias en países tan diferentes como Finlandia o España y, aunque pocas, son suficientes para demostrar que, incluso bajo un mismo enfoque (vectorial o raster), la actualización puede abordarse desde planteamientos muy distintos. El estudio de la información recibida permite hacer el siguiente análisis sinóptico de la situación:

- las experiencias de actualización, en el estricto sentido de la palabra, son realmente escasas.
- Las experiencias existentes, aunque pocas, son muy variadas por lo que es necesario un gran esfuerzo de simplificación para clasificarlas en una u otra de las dos metodologías ya descritas.
- En general, los procesos se basan en fotointerpretación si la escala es mayor de 1:50.000 o en técnicas semiautomáticas de clasificación para escalas menores de 1:100.000.
- Los países que tienen que enfrentarse a actualizaciones de gran extensión aplican técnicas automáticas o semiautomáticas por razones económicas y de productividad.
- La frecuencia de actualización depende de la disponibilidad de datos y de factores de tipo económico. Rara vez se considera el dinamismo de las categorías de la leyenda (frecuencia de cambios). En todos los casos el periodo entre actualizaciones fue mayor de 4 años.
- El tipo de datos a utilizar depende de la escala de la base de datos. Existe una convivencia absoluta entre fotografía aérea e imágenes de satélite si bien la fotografía tiene mayor importancia en actualizaciones a escalas mayores que 1:25.000. Por otra parte, se tiende a la integración de la información digital ya disponible.
- El valor a partir del cual se considera necesario modificar la base de datos original varía entre 1/5 y 2/5 de la superficie de la unidad mínima cartografiada.
- Las clases mixtas son las que presentan mayores problemas a la hora de la actualización.
- La tendencia general es la de archivar cada nueva versión de base de datos.
- Es difícil comparar las experiencias por cuanto están influenciadas por el organismo que las ha efectuado y por el tipo y características de la ocupación del suelo en cada país o región.
- Es difícil comparar las metodologías empleadas ya que no es usual validar los resultados.

CONCLUSIONES

- 1) La producción de datos sobre cubierta y usos está experimentando un crecimiento importante debido a la facilidad de almacenamiento y manejo de información geográfica en ambientes GIS.
- 2) La enorme producción de datos digitales y la necesidad de integrar información preexistente y variable en cuanto a fuente, contenido y escala (resolución) son causas que han hecho necesario el concurso de los métodos de generalización automática (GURNELL et al. 1996).
- 3) Las tendencias futuras en el desarrollo de las capacidades GIS y la plena operatividad de los nuevos sistemas orientados al objeto (USERY, 1996) son hitos que favorecerán la automatización de los procesos. Sin embargo, autores como SPIESS (1995) sostienen que nunca serán operativas las soluciones batch al 100% y los sistemas expertos multi-propósito por lo que, aceptadas las necesidades de automatización, debería profundizarse en el desarrollo de métodos interactivos basados en el conocimiento.
- 4) La presión de los campos científico y comercial favorece el que en este momento exista un amplio cuerpo de investigación trabajando en la automatización de los procesos. Las principales ventajas de alcanzar tal objetivo serían la de la rapidez, la posibilidad de generar múltiples productos desde una única base, la certeza de trabajar con datos consistentes y homogéneos, la fiabilidad en la geometría de los elementos y la facilidad para contrastar la precisión temática.

BIBLIOGRAFÍA

- AGUILÓ, M. et al. (1984). *Guía para la elaboración de estudios del medio físico: contenido y metodología*. Centro de Estudios de Ordenación del Territorio y Medio Ambiente. Manuales 3 (2ª edición).
- ETC/LC IGN (1995). *Assessment on Updating experiences on updating process*. ETC/LC IGN Final draft Report 1995.
- FULLER, R. & GROOM, G. (1993) *The land cover map of Great Britain*. Mapping Awareness, 7(9): 18-20
- FULLER, R.; GROOM, G. & JONES, A.R. (1994). *The land cover map of Great Britain: An automated classification of Landsat Thematic Mapper data*. Photogrammetric Engineering and Remote sensing, 60(5): 553-562
- GOFFREDO, S.; WILKINSON, G.G. & FISHER, P.F (1993). Spatial generalisation of thematic maps derived from satellite imagery for operational land cover mapping, towards operational applications. Annual Conference of the Remote Sensing Society. RSS, Reading, UK.
- GURNELL, A.M.; ANGOLD, P. & GREGORY, K.J. (1994) *Clasificación de rivers corridors: Issues to be addressed in developing and operational methodology*. Aquatic Conservation/Marine and Freshwater Ecosystems, 4(3):219-231
- JRC, 1992. *Updating of the CORINE land cover database*. Annual Report'92. ECSC-EC-EAEC, Brussels, Luxembourg.
- JRC, 1993. *Monitoring the condition and development of vegetation at a continental scale*. Annual Report'93. ECSC-EC-EAEC, Brussels, Luxembourg.
- JAAKKOLA, O. (1995) *Automatic Generalization of land cover data*. Proceedings of the 17th International Cartographic Conference, vol 2. Barcelona, Spain. (pp. 1984-1989).
- LI, Z. & OPENSHAW, S. (1992). *Algorithms for automated line generalisation based on natural principle of objective generalisation*. International Journal of Geographical Information Systems 6(5): 373-389.
- LINDEN, V.H.; JANSSEN, L. & ZEEUW, C. (1990). *Integrated use of ERDAS and ARC/INFO for land cover mapping purposes*. CORINE Technical Papers.
- MONMONNIER, M.S. (1983). *Raster-mode area generalization for land use and land cover maps*. Cartographica 20(4): 65-91
- MULLER, J.C. & ZESHEN, W. (1992). *Area-patch generalisation: a competitive approach*. The Cartographic Journal 29:137-144
- OOSTERON, P.V. (1996). *Applying reactive data structures in a interactive multi scale GIS*. In M. Molenaar (Ed): *Methods for generalization of geo-databases*. Publications in Geodesy, 43. Nederlandse Commissie voor Geodesy.
- PAULSSON, B. (1994). *A computerized generalization approach for land cover maps*. XFF640. Swedish Space Corporation
- SCHILBERG, L. (1992). *Cartographic amalgamation of area objects*. Proceedings of ISPRS Commission IV 1992.
- SPIESS, E. (1995). *The need for generalization in a GIS environment*. In: J.C. Muller, J.P. Lagrange & R. Weibel (Eds): *GIS and Generalization*. Taylor & Francis. London.
- USERY, E.L. (1996). *A feature based geographic information system model*. Photogrammetric Engineering and Remote Sensing, 62(7): 833-838