

Limitaciones de los métodos convencionales de estimación de la fiabilidad de los mapas temáticos derivados de teledetección

Guillermo Castilla⁽¹⁾, Salomón Montesinos⁽²⁾

⁽¹⁾Sección de Teledetección y SIG, IDR, Universidad de Castilla La Mancha. Campus universitario s/n, 02071 Albacete. Guillermo.Castilla@uclm.es.

⁽²⁾Geosys SL. Sector Foresta, 23 Locales 7 y 8, 28760 Tres Cantos, Madrid. smontesinos@geosys.es.

Resumen

La matriz de confusión y el error medio cuadrático (RMSE) son lugares comunes en la estimación de respectivamente la precisión temática y posicional de los mapas derivados de imágenes de teledetección. Sin embargo, ambos presentan algunas carencias que son poco conocidas y que aquí se tratan de poner de relieve, en particular las referentes a la distribución espacial de los errores. El principal argumento de este trabajo es que los métodos convencionales asumen que los errores están distribuidos aleatoriamente sobre el territorio, lo que en la mayoría de los casos está lejos de ser cierto. Se proponen algunas alternativas metodológicas y finalmente se insiste en la necesidad de metadatos para poder valorar la calidad de estos mapas.

1. Introducción

Los mapas temáticos son modelos geográficos que representan la distribución sobre un territorio de un determinado fenómeno, como la vegetación, el uso del suelo o la geología. Fijado un nivel de detalle tanto cartográfico como temático, serán errores aquellas desviaciones entre la representación y lo representado que no se puedan explicar por los mecanismos de generalización espacial y categórica aplicados. Por tanto habrá dos tipos de errores que, si bien interrelacionados, conviene distinguir: el error posicional y el temático, cuya abundancia y magnitud determinarán la mayor o menor fiabilidad del mapa. En este artículo se cuestiona la idoneidad de los métodos convencionales de estimación de la fiabilidad de los mapas temáticos derivados de *ortoimágenes* (tanto de satélite como aéreas), enfatizándose sus carencias a la hora de describir la distribución espacial de los errores.

1.1. Campos geográficos y objetos geográficos

Antes de continuar, es necesario hacer notar que la definición de error es dependiente del modelo conceptual en que se apoye la representación del

fenómeno estudiado. Existen dos modelos principales de mapas temáticos, que se pueden asociar respectivamente a dos disciplinas de diferente origen y antigüedad pero con trayectorias convergentes: la teledetección digital por un lado y la cartografía y fotogrametría por otro.

En el primer caso el territorio se divide según una malla regular cuyas cuadrículas se asignan individualmente, normalmente por un método semi-automático, a una determinada clase temática. Bajo esta perspectiva, el fenómeno cartografiado constituye un *campo geográfico*, es decir una variable regionalizada que puede tomar tantos valores como clases hay en la leyenda del mapa, siendo ese valor función de la posición geográfica. La representación digital más usada para este modelo es el *ráster*, en que cada cuadrícula viene representada por un píxel cuyo color es el asignado a la clase temática a la que se supone pertenece.

En el segundo caso, el territorio se halla ocupado por un mosaico de recintos de forma y tamaño diversos que, a diferencia de las cuadrículas, pertenecen necesariamente a una clase diferente de la de sus vecinos colindantes. Estos recintos, al constituir un todo espacialmente unitario, semánticamente homogéneo, y diferente de lo que le rodea, pueden verse como objetos geográficos, de forma que cada clase temática estaría constituida por el conjunto de objetos que poseen las características distintivas de esa clase. A diferencia del modelo anterior, cuya generación está bastante automatizada, la creación de este tipo de modelos a partir de *ortoimágenes* es tan compleja que de momento sólo puede hacerse mediante *fotointerpretación*. La representación digital habitual para el modelo de objetos es la *capa vectorial de polígonos*.

Aparentemente no hay contradicción entre ambos modelos, pues sus respectivas representaciones, ráster y vectorial, son intercambiables. Sin embargo, desde el punto de vista conceptual, el modelo de objetos se sitúa en un

nivel de abstracción superior que el de campos, y por tanto el tipo de información que transmite y su inherente error son diferentes. La distinción principal entre ambos modelos es su concepto de clase [1]. Para el primero las clases se refieren a tipos de materiales (p.ej. *agua*), y por tanto tienen la consideración de *nombres incontables*. Las clases así concebidas tienen la particularidad de que una porción arbitrariamente delimitada (pongamos una cuadrícula) dentro de una zona con un determinado material sigue siendo un referente del nombre de ese material. En contraposición, las clases en el segundo modelo se refieren a tipos de objetos, y por consiguiente se puede hablar de ellas como *nombres contables* (p. ej. *lagos*, en que una parte arbitrariamente delimitada dentro de un lago ya no es referente del nombre *lago*). En adelante, pese a no existir como se ha explicado una relación biunívoca entre el modelo conceptual y su representación digital, se denominará al modelo de campos *mapa ráster* y al de objetos *mapa de polígonos*, por ser una terminología más familiar.

2. Fiabilidad posicional

En los mapas ráster, la fiabilidad posicional está relacionada con el grado de confianza que se tiene en que la posición real en el terreno de la cuadrícula se ajuste razonablemente a las coordenadas geográficas del píxel correspondiente, por tanto depende sobre todo de la calidad de la corrección geométrica aplicada a la imagen de la cual se ha derivado el mapa. La precisión posicional se estima a través del error medio cuadrático (RMSE) de un conjunto de puntos claramente identificables en la imagen y cuyas coordenadas se conocen por una fuente de mayor precisión. Esta estimación asume que los errores posicionales están distribuidos al azar por la imagen, lo que por desgracia está lejos de ser cierto en el caso de terrenos accidentados, donde los errores suelen ser más abundantes en las zonas de mayor pendiente. En cualquier caso la incertidumbre sobre la posición de una cuadrícula suele ser bastante más benigna que la referente a su clase, por lo que aquí se se trata de pasada.

En los mapas de polígonos, la fiabilidad posicional tiene que más que ver, asumiendo una correcta corrección geométrica de la imagen, con el grado de confianza que se tiene en que el límite entre dos polígonos ha sido bien delineado. Esto se estima en la práctica a través del RMSE de una muestra de puntos situados sobre arcos que correspondan a bordes bien definidos en la imagen, de forma que se pueda tomar una referencia fiable.

Sin embargo, esta estimación está sesgada hacia bordes de origen antrópico, ya que los bordes naturales suelen ser difusos. Por tanto el método convencional, más que estimar lo bien que los bordes de los polígonos esquematizan la distribución espacial del fenómeno estudiado, lo que en realidad estima es estimar la bondad del pulso del intérprete a la escala a la que visualizó la imagen. Además, a diferencia del caso anterior, la incertidumbre aquí es inseparable de la temática, pues de todos los bordes existentes en la imagen, únicamente se delinean aquellos que separan clases que figuran en la leyenda. Por tanto la fiabilidad de los bordes es proporcional, por un lado, a lo diferentes que se vean en esa parte de la imagen las clases separadas por ese borde y a cuán rápida sea la transición entre ellas, y por otro, a la dificultad que se tenga para discernir entre esas clases. Como no todos los bordes son igual de nítidos, su fiabilidad debe ser estimada independientemente para cada arco de la capa vectorial.

Esto podría hacerse mediante la definición de una *banda ypsilon* [2], dentro de la cual hay una probabilidad del 99% de que se sitúe el borde 'verdadero'. Sin embargo este método apenas se usa en la práctica. Como ejemplo, Green y Hartley [3] calcularon la anchura de las bandas ypsilon intersectando capas vectoriales delineadas sobre la misma escena por diferentes intérpretes y midiendo la anchura media de los polígonos de manga (i.e. espurios) creados por la intersección. Por desgracia, este método no sólo no es operativo para trabajos de producción cartográfica, sino que además se basa en una hipótesis poco realista: que dado un territorio y un nivel de detalle espacial y categórico, se puede crear una representación *egg-yolk* [4] de éste, consistente en un conjunto de polígonos solapados semejante a una bandeja llena de huevos fritos, en la que la yema de cada huevo es el núcleo del polígono por el que ningún intérprete ha trazado un borde, y las claras son las diversas bandas ypsilon que rodean a cada núcleo. Desafortunadamente, en la práctica no será difícil encontrar algunos polígonos delineados por el intérprete A que son cortados por la mitad por un borde delineado por el intérprete B y viceversa.

Una alternativa más viable que estamos investigando para estimar la fiabilidad posicional de los arcos de estos mapas, es expresar ésta como una medida combinada de la notoriedad tanto visual como semántica. La notoriedad visual de un arco dado se podría estimar como la magnitud media del gradiente de los píxeles que atraviesa ese arco; y la semántica podría ser el valor de algún índice de

similitud temática entre las clases separadas por el arco. Adicionalmente, la precisión posicional se podría estimar para una muestra de polígonos (la misma que se use para estimar la precisión temática). La pregunta que se haría en cada polígono es, ¿están los bordes de este polígono delineados de tal forma que se puede concebir como una unidad coherente con el esquema clasificatorio del mapa? La respuesta podría darse cualitativamente mediante una escala nominal o incluso cuantitativamente calculando las bandas ϵ derivadas del trabajo de varios intérpretes a los que se les diera independientemente la tarea de mejorar la delineación de los polígonos de la muestra, que en virtud de lo expuesto anteriormente no debieran ser adyacentes.

3. Fiabilidad temática

La fiabilidad temática se asocia en el caso de los mapas ráster al grado de confianza que se tiene en encontrar efectivamente la clase temática asignada a un píxel dado si se visita la cuadrícula de terreno correspondiente a ese píxel. Esta incertidumbre se evalúa usando una tabla de contingencia o acuerdo entre valores observados en el campo y pronosticados en el mapa, la famosa *matriz de confusión*. Esta matriz se obtiene de una muestra de píxeles cuya clase real sobre el terreno es conocida. La principal carencia de esta matriz es que solo refleja el error medio sobre el territorio cartografiado, mientras que la probabilidad de clasificación errónea varía de forma acusada de un lugar a otro [5]. Resulta sorprendente que tratándose de una variable regionalizada, la norma establecida de estimación de errores deje completamente de lado el aspecto geoestadístico.

Además de considerar que los errores están distribuidos aleatoriamente por el territorio, los estimadores de fiabilidad derivados de la matriz de confusión dan igual peso a todos los desacuerdos entre clases, mientras que está claro que unos errores son más graves que otros dependiendo de la importancia y significación de las clases. Una solución a esto sería dar a cada error un peso según una escala nominal que iría de 'absolutamente erróneo' a 'error comprensible y no importante' [6].

Otra debilidad es el modo de seleccionar los píxeles de la muestra. Aun en el caso de que se aplique algún tipo de diseño estadístico, las zonas de verificación se suelen tomar por bloques en lugar de por píxeles individuales. Esta práctica puede producir un sesgo optimista en los estimadores de la precisión, ya que estos bloques se suelen escoger

bien alejados de los bordes, de forma que representen a un única clase, con lo que se dejan sistemáticamente de lado en la muestra los píxeles mezclados, que son más proclives a estar mal clasificados [7]. Además, por efecto de la autocorrelación espacial, los valores de los píxeles dentro del bloque son interdependientes más que independientes, por lo que al tomar bloques se viola la hipótesis de independencia del muestreo estadístico.

Otro asunto en el que no se suele ahondar es la fuente de la mal llamada 'verdad terreno', es decir, los datos que se usan como referencia para verificar la clasificación. Muchas veces a causa de limitaciones logísticas y financieras se usan imágenes de resolución superior o mapas supuestamente más precisos en lugar de recurrir a las siempre costosas inspecciones sobre el terreno, lo que aumenta la incertidumbre sobre la validez de esa estimación de la fiabilidad. Por ejemplo, si se usa un mapa de polígonos, más detallado pero disponible sólo en algunas zonas de la imagen, como 'verdad terreno' para estimar la fiabilidad de un mapa ráster obtenido por clasificación digital, serán fuente de incertidumbre, además de otros aspectos, la compatibilidad entre las leyendas y la contingencia de inclusiones menores que la Unidad Mínima Cartografiable (UMC) en los polígonos, sobre todo si ésta es bastante mayor que el tamaño del píxel, como suele ser común.

Una alternativa más apropiada para la estimación de la fiabilidad es el modelo de error propuesto por Goodchild *et al* [8], en el que cada píxel tiene asociado un vector de probabilidades m -dimensional que da la probabilidad de que el píxel pertenezca a cada clase de 1 a m . Entonces se considera que el mapa es una realización de un proceso estocástico definido por esos vectores, es decir, que sobre un gran número de realizaciones, la proporción de veces que un píxel es asignado a cada una de las clases convergerá hacia la probabilidad de esa clase en ese píxel. El modelo además incluye una serie de parámetros que describen la dependencia espacial de los resultados locales de la realización, de forma que los píxeles vecinos se hallarán en general correlacionados. A modo de ejemplo, una forma de obtener una realización sería efectuar la clasificación sobre una submuestra aleatoria de píxeles de entrenamiento, de forma que se puedan ajustar los parámetros geoestadísticos y obtener realizaciones ya directamente con el modelo que se asemejen a las obtenidas mediante este procedimiento.

En el caso de los mapas de polígonos, la fiabilidad temática está relacionada con el grado de confianza que se tiene en que al inspeccionar en campo un determinado polígono se llegaría a la conclusión de que la clase que mejor lo representa es la que ha sido asignada en el mapa. De igual forma que en el anterior caso, esta incertidumbre se puede evaluar a través de la matriz de confusión, sólo que aquí se usan como unidades de muestreo polígonos completos en lugar de parcelas, ya que la asignación del polígono a una clase se ha hecho considerando al polígono como un todo, obviando las inclusiones menores que la UMC que pueda tener, y teniendo en cuenta incluso lo que le rodea. La verificación por tanto debería hacerse por un medio (p.ej. vuelos de baja altura) que permita una evaluación global de los polígonos de la muestra. Sin embargo, esto puede acarrear problemas tanto operativos como de diseño de muestreo, sobre todo si hay gran disparidad de tamaño de polígonos, pues la estimación de fiabilidad se debe referir al área total más que al número de polígonos [9].

4. Conclusiones

Los métodos convencionales de estimación de la fiabilidad de los mapas temáticos derivados de teledetección presentan ciertas limitaciones, algunas de ellas muy relevantes, como su incapacidad para dar información sobre la distribución espacial de los errores. Existen métodos alternativos que pueden dar respuesta a estas carencias, pero apenas son conocidos o no pasan de ser propuestas. En cualquier caso, a pesar de la aparente objetividad de las estimaciones cuantitativas sean del tipo que sean, es importante que se interpreten con cuidado porque no es difícil llegar a conclusiones erróneas apoyándose únicamente en los números [10]. La conclusión es que la mejor forma de valorar la fiabilidad de un mapa es a través de sus metadatos, que hoy por hoy siguen sin usarse y exigirse todo lo que se debiera. De hecho es común encontrar, si acaso hay metadatos, que el bloque de *información sobre calidad de los datos* del archivo XML de ISO 19115 esté vacío. Sería deseable corregir esta situación y que este último no sólo describa profusamente el material y métodos empleados en la producción del mapa, sino también la localización de las muestras empleadas en la estimación de la fiabilidad.

5. Agradecimientos

Este trabajo se ha llevado a cabo dentro del proyecto 'Harmonised Techniques and Representative River Basin Data for Assessment

and Use of Uncertainty Information in Integrated Water Management (*HarmoniRiB*)', financiado por el programa de Energía, Medioambiente y Desarrollo Sostenible de la Comisión Europea (contrato EVK1-CT-2002-00109).

6. Referencias

- [1] Castilla, G., *Object-oriented analysis of Remote Sensing images for land cover mapping: conceptual foundations and a segmentation method to derive a baseline partition for classification*, Tesis doctoral, Univ. Pol. Madrid, 2003.
- [2] Blakemore, M., "Generalization and error in spatial databases", *Cartographica*, 21, 131-139, 1984.
- [3] Green, D.R., and Hartley, S., "Integrating photointerpretation and GIS for vegetation mapping: some issues of error", *Vegetation mapping: from patch to planet*, Wiley & sons, 103-134, 2000.
- [4] Cohn, A.G., and Gotts, N.M., "The 'egg-yolk' representation of regions with indeterminate boundaries", *Geographic objects with indeterminate boundaries*, Taylor & Francis, 171-188, 1996.
- [5] Goodchild, M.F. (1994) "Integrating GIS and remote sensing for vegetation analysis and modelling: methodological issues", *J. Veg. Sci.*, 5, 615-626, 1994.
- [6] Gopal, S., and Woodcock, C.E., "Theory and methods for accuracy assessment of thematic maps using fuzzy sets", *Photogrammetric Eng. and Remote S.*, 60(2), 181-188, 1994.
- [7] Congalton, R.G., and Plourde, L.C., "Sampling methodology, sample placement, and other important factors in assessing the accuracy of remotely sensed forest maps", *Proc. 4th Int. Symp. on spatial accuracy assessment in natural resources and environmental sciences*, 117-124, 2000.
- [8] Goodchild, M.F., Sun, G., and Yang, S., "Development of an error model for categorical data", *Int. J. GIS*, 6 (2), 87-104, 1992.
- [9] Stheman, S.V., and Czaplewski, R.L., "Design and analysis for thematic map accuracy assessment: fundamental principles", *Remote S. Env.*, 64 (3), 331-344, 1998.
- [10] Foody, G.M., "Status of land cover classification accuracy assessment", *Remote S. Env.*, 80, 185-201, 2002.