

CONSECUENCIAS EN LA CARTOGRAFÍA DE CULTIVOS MEDITERRÁNEOS DE LA COMBINACIÓN DE DATOS RÁSTER-VECTOR: ENRIQUECIMIENTO DE UNA BASE SIGPAC A TRAVÉS DE TELEDETECCIÓN

Consequences on Mediterranean crop mapping by combining raster and vector data: enrichment of a SIGPAC database using remote sensing

Serra, P.¹, Moré, G.² and Pons, X.^{1,2}

Recibido 07/12/2006; aceptado el 15/03/2007

Resumen. *El objetivo de este artículo es el estudio de las implicaciones del enriquecimiento de cartografía vectorial a partir de cartografía ráster obtenida mediante teledetección. El banco de pruebas es una base vectorial SIGPAC de seis comarcas catalanas enriquecida con un mapa de cultivos mediterráneos ráster elaborado a partir de la clasificación mixta de siete imágenes Landsat TM del 2004. Se han ensayado dos niveles de fidelidad en el clasificador y dos niveles de pureza en la composición del polígono. Los resultados se analizan a nivel global y por cultivo, observándose, entre otras cuestiones, que el enriquecimiento vectorial suele ser positivo a ambos niveles. Asimismo, la exigencia de mayor fidelidad y una pureza con restricción conlleva mayor exactitud aunque un menor número de polígonos etiquetados, lo que sólo significa más error si estos se consideran errores.*

Palabras clave: *Cultivos mediterráneos, teledetección, enriquecimiento SIGPAC, matrices de confusión, combinación analítica de capas.*

Abstract. *The objective of this article is to analyze the consequences of enriching vector cartography from raster cartography obtained using remote sensing. The testing ground is a SIGPAC vector base of six Catalan regions enriched with a raster map of Mediterranean crops elaborated from the mixed classification of seven Landsat TM images of year 2004. Two levels of fidelity in the classifier and two levels of purity in the polygon composition have been tested. The outcomes are analyzed at global level and per crop, giving as a result, among other issues, that the vector enrichment is usually positive at both levels. Likewise, the exigency of greater fidelity and a purity with restriction yields greater exactitude despite of a smaller number of labeled polygons, which only means more error if they are considered errors.*

Key Words: *Mediterranean crops, remote sensing, SIGPAC enrichment, error matrices, analytical combination of layers.*

¹ Departamento de Geografía. Edificio B, Campus de la Universidad Autónoma de Barcelona, 08193-Cerdanyola del Vallès (Barcelona). Tel.: 00 34 93581 3273; Fax: 00 34 93 581 2001.
E-mail: pere.serra@uab.cat

² Centre de Recerca Ecològia i d'Aplicacions Forestals (CREAF). Edificio C, Universidad Autónoma de Barcelona, 08193-Cerdanyola del Vallès (Barcelona).

Serra, P., Moré, G and Pons, X

INTRODUCCIÓN

En los últimos años uno de los campos con más aplicaciones de la teledetección espacial con imágenes de satélite ha sido el agrícola, especialmente por la gran ventaja que presenta respecto a otras herramientas de observación terrestre. En efecto, por una parte la teledetección espacial proporciona la adquisición continuada de imágenes a lo largo del año, lo cual permite calcular el estado fenológico de las plantas a través de su respuesta espectral y radiométrica almacenada en las imágenes. Por otra parte, la riqueza multispectral es en sí misma riqueza de información multivariante que permite abordar la generación de cartografía temática, en este caso de cultivos, con un excelente material estadístico de partida. Así, pues, y a pesar de las dificultades técnicas que puede presentar, la cartografía actualizada de mapas de cultivos son de gran utilidad para diversas administraciones públicas, que ya la están usando, como son el Departamento de Agricultura, Ganadería y Pesca de la Generalitat de Catalunya o la Agencia Catalana del Agua.

Sin embargo, el producto final de los trabajos obtenidos con teledetección suele almacenarse siguiendo un modelo ráster ya que resulta el más inmediato a partir de la información de captación de las imágenes de satélite. Este hecho puede dificultar su interés o aplicabilidad ya que mucha información, sobretudo la gestionada por las administraciones públicas, se encuentra en modelos vectoriales. Las diferencias en dichos modelos pueden conllevar discrepancias más o menos significativas en, por ejemplo, la cuantificación de las superficies de polígonos o el cálculo de las longitudes de elementos lineales (Burrough & McDonell, 1998). Además, en la mayoría de entornos SIG no es posible efectuar de forma inmediata consultas combinadas sobre ambos modelos.

Una posible solución consiste en vectorizar el ráster final; existen dos opciones, siendo una la vectorización siguiendo la forma y tamaño de píxel del ráster original, con lo cual se obtiene el mismo producto pero en modelo vectorial (se observa la típica escalera en los bordes de los polígonos debido a los píxeles), y siendo la segunda el mismo proceso pero aplicando una generalización (suavizado) de sus límites con la consiguiente modificación de los datos originales. Además, este proceso de transformación tampoco garantiza una combinación de capas sin problemas ya que suele producir, al menos, la existencia de micropolígonos (*slivers*) que dificultan el análisis final.

Sin embargo, existe otra solución de mayor interés que consiste en «enriquecer» una capa vectorial preexistente, como por ejemplo el catastro, con la información obtenida en un ráster derivado de teledetección (Aplin *et al.*, 1999). Este método tiene la ventaja que los límites de la base vectorial no se modifican y, por tanto, no aparecen los problemas anteriormente mencionados. Para realizar dicha operación desde el programa MiraMon (Pons, 2006) se aplica el módulo

«combinación analítica de capas» (Pesquer *et al.*, 2000) y la opción «transferencia de campos estadísticos». Dicha opción calcula los estadísticos requeridos por el usuario (moda, % de pureza, etc.) a partir de los píxeles situados en el interior de cada polígono y coloca sus valores en sendos campos que enriquecerán la base de datos de la capa vectorial. De esta forma, estos nuevos datos podrán ser consultados en la capa vectorial enriquecida, tal como se detallará en el correspondiente apartado metodológico. En este trabajo se transferirán los campos estadísticos al Sistema de Información Geográfica de la Política Agraria Común (SIGPAC), capa en formato vectorial gestionada por el Departamento de Agricultura, Ganadería y Pesca de la Generalitat de Catalunya, a partir del mapa de cultivos obtenido con las imágenes de satélite.

Un aspecto primordial para la combinación ráster-vector objetivo de nuestro trabajo consiste en que los mapas de cultivos tengan elevada fiabilidad ya que la información final se combinará con la información vectorial anteriormente detallada. Para ello se especificará brevemente la metodología aplicada, en concreto el proceso de clasificación automática a través de un clasificador híbrido y, posteriormente, se detallará la exactitud temática del mapa obtenido. En teledetección la exactitud temática se suele conocer a partir de las matrices de confusión donde se comparan los resultados obtenidos con áreas conocidas (denominadas áreas de entrenamiento cuando se emplean para la clasificación automática y áreas de test cuando sirven para cuantificar los errores de la clasificación), normalmente derivadas de trabajo de campo o de otras fuentes cartográficas (Khorram, 1999; Foody, 2002). En nuestro caso las áreas de entrenamiento y de test corresponden a parcelas de cultivos conocidas. Las matrices de confusión se calculan a partir de las áreas de test comparando píxel a píxel (y, por tanto, siendo posible calcular el número de píxeles bien clasificados), que es la opción clásica (Barbosa *et al.*, 1996; De Wit *et al.*, 2004), o polígono a polígono (y, por tanto, obteniéndose el número de polígonos correctamente clasificados), que es una opción menos empleada (Aplin *et al.*, 1999; Martínez *et al.*, 2001) posiblemente debido a su mayor dificultad. En este estudio se compararán las dos opciones con el objetivo de analizar qué diferencias se producen, o sea, si el producto vectorial enriquecido es mejor o peor que el ráster original. La hipótesis de investigación parte de que la comparación polígono a polígono debería ser más exacta debido a la eliminación de píxeles aislados (*spurious pixels*), propios del efecto «sal y pimienta» típico de las clasificaciones píxel a píxel, es decir de la aparición de píxeles mixtos responsables de una mayor fragmentación de los resultados (Chuvieco, 2002) como se detallará en el apartado metodológico. Un aspecto importante es que se parte de la base que dentro de cada polígono debe existir solamente un tipo de cultivo, de modo que los posibles píxeles aislados que aparezcan serán fruto de errores en el proceso de clasificación, ya

sean debido a presencia de ruido en la señal radiométrica, a la contaminación por píxeles contiguos, a posibles imprecisiones cartográficas o incluso a una mala delimitación del polígono.

ÁREA DE ESTUDIO Y MATERIALES

El área de estudio corresponde a seis comarcas del centro-oeste de Cataluña (nordeste de España), en la provincia de Lleida, con una superficie total de 348.533 ha (figura 1). Se trata de unas comarcas con un elevado predominio de los usos agrícolas, tanto herbáceos como leñosos, cubriéndose todas ellas con una imagen Landsat *Thematic Mapper* (TM correspondiente a la pasada 198 y fila 31). Las imágenes Landsat TM poseen siete canales espectrales: tres correspondientes al visible (entre 0,45 y 0,69 μm) y tres infrarrojos (0,76 a 2,35 μm) con un píxel de 30m * 30m y un canal térmico (10,4 a 12,5 μm) con un píxel de 120m * 120m. Debido a la peor resolución espacial del canal térmico y a que los valores de temperatura corresponden sólo al momento de paso del sensor, es decir, son valores muy concretos en el tiempo, este canal no se usó en este trabajo. Para conseguir un mejor conocimiento del comportamiento fenológico intraanual de los distintos cultivos, se aplicó una aproximación multitemporal; en concreto se emplearon siete imágenes del 2004: 16 de mayo, 1 y 17 de junio, 19 de julio, 4 de agosto, 23 de octubre y 8 de noviembre.

Una vez que los metadatos de las imágenes originales fueron convenientemente importados, el siguiente paso fue la corrección geométrica de las imágenes con el objetivo de evitar las distorsiones provocadas por la captación del sensor y por el relieve. El método empleado corresponde al desarrollado por Palá y Pons (1995) en el cual se tiene en cuen-

ta el relieve a través de la inclusión de un modelo digital del terreno. Durante el proceso se usaron una media de 26 puntos de control y 12 de test y se obtuvo para cada imagen un *Root Mean Squared*¹ (RMS) por debajo de un píxel, unos 20 metros.

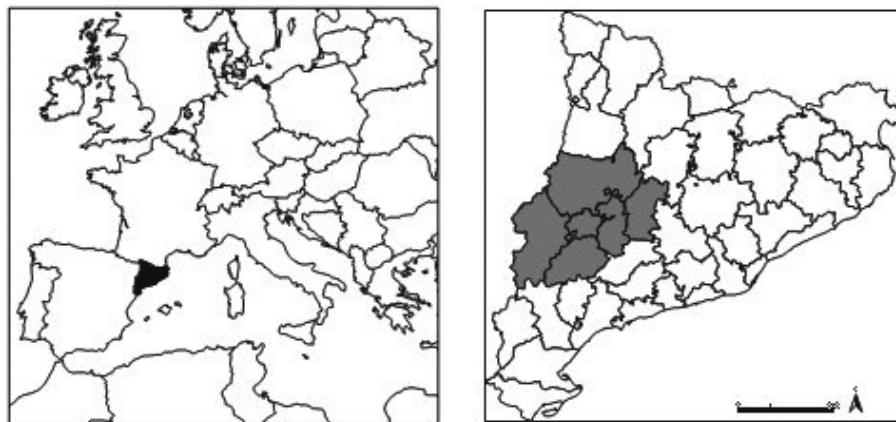
El segundo paso fue la corrección radiométrica, en la cual los DNs (*digital numbers*) fueron convertidos en reflectancias usando los parámetros de calibración del sensor y otros factores como los efectos atmosféricos, el ángulo de incidencia solar según el relieve, etc. (Pons y Solé-Sugrañes, 1994). El resultado de las imágenes corregidas presentaba un rango coherente de reflectancias.

Para evitar, durante el proceso de clasificación, confusiones radiométricas y espectrales provenientes de otras cubiertas como las forestales o las urbanas, se aplicó una máscara para seleccionar únicamente las cubiertas agrícolas. Esta discriminación se obtuvo a partir de la propia base SIG-PAC, donde se seleccionaron sólo los usos agrícolas, enmascarándose el resto en todas las imágenes como área no agrícola, y, por tanto, no incluyéndose en el análisis.

CLASIFICACIÓN

Una vez las imágenes estaban corregidas se inició la fase de clasificación automática; en teledetección este término, a diferencia de los métodos de fotointerpretación visual, hace referencia a la obtención automática de mapas categóricos gracias al tratamiento estadístico de los datos digitales de las imágenes de satélite (Chuvienco, 2002). Un requisito imprescindible en esta fase es la adopción de la leyenda, o sea, de los cultivos a discriminar. A través de los conocimientos adquiridos en el trabajo de campo y en trabajos previos (Serra *et al.*, 2003), el mapa de cultivos del 2004 constó de los

FIGURA 1: Área de estudio situada al nordeste de España (en negro) y centro-oeste de Cataluña (en gris).



1. Indicador de la calidad de la corrección geométrica a través de la raíz cuadrada de la media del cuadrado de las desviaciones entre los valores observados y los estimados (Chuvienco, 2002)

Serra, P., Moré, G and Pons, X

siguientes cultivos: cereales de invierno, arroz, maíz de regadío, alfalfa de regadío, otros cultivos herbáceos, frutales, olivos, viñedos, pastos y barbecho.

La aplicación del clasificador híbrido consiste en dos módulos del programa MM (Pons, 2006): IsoMM y ClsMix, que corresponden a una combinación entre la clasificación no supervisada (el primer módulo donde el programa clasifica automáticamente las clases espectrales creando *clusters* o grupos homogéneos de píxeles) y la clasificación supervisada (el segundo módulo a través de áreas de entrenamiento). En el módulo ISOMM los *clusters* se forman por asignaciones interactivas de píxeles n-dimensionales. Estas asignaciones están basadas en la distancia mínima de un píxel respecto a todos los centroides de los *clusters* iniciales. El grupo inicial de centroides, las semillas, se obtiene en la fase anterior a la obtención de los *clusters*. El módulo presenta tres opciones para obtener las semillas iniciales: i) a lo largo de la diagonal multivariante, calculada a partir de todas las variables originales, ii) una distribución aleatoria en todo el espacio multivariante, iii) una distribución basada en una muestra espacialmente equidistante sobre la imagen (por ejemplo una semilla cada 50 píxeles, etc.). Después de cada iteración los centros de *cluster* son actualizados al centroide de todos los píxeles asignados. El método empleado en este trabajo corresponde a la primera opción.

Una de las características más interesantes de ISOMM es que admite cientos de imágenes. La principal utilidad de esta propiedad es que permite el uso de series satelitarias de elevada resolución temporal y de variables climáticas y topográficas. Asimismo, el módulo permite obtener un elevado número de categorías estadísticas (hasta 32.767) que pueden ser eliminadas o modificadas por el algoritmo según diferentes parámetros configurados por el usuario: la distancia mínima entre dos *clusters* válidos y el mínimo número de píxeles por *cluster* para considerar un *cluster* válido. En el primer caso, los *clusters* se fusionan cuando la distancia es menor que un mínimo valor definido por el usuario, mientras que en el segundo caso un *cluster* es eliminado si su área total es menor a un umbral establecido por el usuario. Finalmente, el módulo requiere la introducción del número de *clusters* deseado, el máximo número de iteraciones antes de terminar y un valor umbral de estabilidad de los resultados para finalizar el algoritmo.

En este estudio todas las bandas mencionadas anteriormente fueron introducidas, a excepción de la térmica por los motivos anteriormente expuestos. Además, con el objetivo de discriminar la actividad clorofílica de los cultivos a lo largo del año, para cada fecha se incluyó el *Normalized Difference Vegetation Index* (NDVI). El NDVI es el índice más empleado por la información que proporciona sobre los cultivos (Lyon *et al.*, 2003), siendo considerado un índice de verdor porque es una estimación de la actividad clorofílica de las plantas.

En la segunda fase del proceso de clasificación, CLSMIX asigna cada clase espectral (obtenida con ISOMM) a una temática a través de dos parámetros: fidelidad y representatividad. Por una parte, la fidelidad es la proporción umbral a partir de la cual se acepta una clase espectral como parte de una clase temática, o sea, la proporción de la clase espectral que coincide con la clase temática. Por ejemplo, 0,8 equivaldrá a que el 80% o más de la clase espectral se encuentra dentro de la temática (según las áreas de entrenamiento) y, en consecuencia se asignará a ella. Por otra parte, se aplica simultáneamente la proporción umbral a partir de la cual se acepta una clase espectral como parte de una temática, o sea, la proporción de la categoría que está formada por una determinada clase espectral. Por ejemplo, 0,02 equivaldrá a que el 2% o más de una categoría está formada por una determinada clase espectral, asignándose ésta a la clase temática correspondiente.

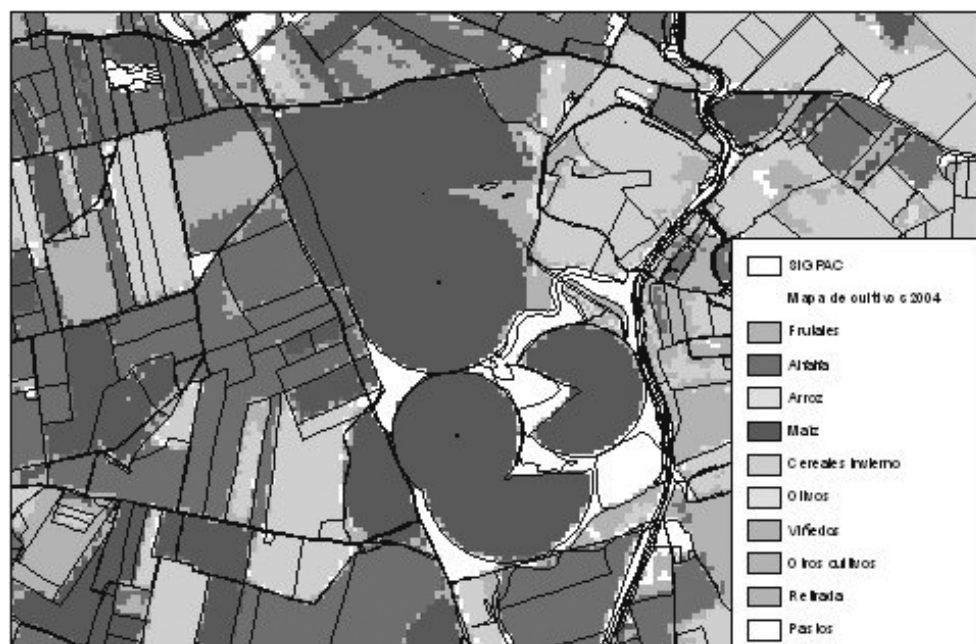
En consecuencia, cuando un píxel es clasificado el módulo elige la categoría que tiene la asignación más razonable: i) la correspondencia espacial entre la clase espectral y la área de entrenamiento de esa categoría (por tanto, la clase espectral está dentro del área de entrenamiento), ii) la clase espectral está principalmente incluida en la categoría (o sea, una importante proporción de la clase espectral corresponde a la categoría temática), y iii) la clase espectral no es una parte insignificante de la categoría. Contrariamente, un píxel no se clasificará si no existe una clase espectral clara o debido a que el píxel pertenece a una clase que está pobremente representada.

En este trabajo se han contrastado dos valores de fidelidad: 0,31 (o sea, un 31% o más de la clase espectral debe estar dentro de las áreas de entrenamiento de una determinada clase temática) y, otra mucho más restrictiva, 0,51 (más de la mitad de la clase espectral debe estar dentro de las áreas de entrenamiento de una clase temática) con una representatividad de 0,0001 (como mínimo la clase espectral tiene que representar un 0,01% de la categoría temática). Tal como se ha detallado anteriormente, una fidelidad más restrictiva comportará la existencia de más píxeles sin clasificar pero tal vez un menor porcentaje de error en los píxeles clasificados, aspecto que se contrastará en los resultados finales.

INTEGRACIÓN RÁSTER-VECTOR

Una vez que se obtuvo el mapa de cultivos de 2004 con dos fidelidades distintas, el siguiente paso fue la integración de la información ráster con el mapa vectorial digital correspondiente al SIGPAC, creado a partir del Reglamento (CE) n° 1593/2000, del Consejo de 17 de Julio de 2000, que establece la obligación de crear un Sistema Gráfico Digital de Identificación de Parcelas Agrícolas (MAPA 2006). El SIGPAC es un registro público de identificación de parcelas agrícolas y actualmente es una referencia obligatoria para identificar

FIGURA 2: Superposición del mapa de cultivos y de la capa SIGPAC.



una explotación agraria declarada en la Declaración Única Agraria (DUN) (DARP, 2006). La información disponible está elaborada a partir de ortofotos y la superposición de las parcelas catastrales que forman los recintos, definidos como la porción continua de terreno que, dentro de una parcela catastral, tiene un mismo uso o aprovechamiento del terreno. Los usos agrarios SIGPAC que se distinguen para toda Cataluña son: cítricos, entorno de olivos, frutales, invernaderos, asociación olivo-frutal, olivos, pastos arbolados, pastos arbustivos, tierra campá (que equivaldría a cultivos herbáceos), huerta, asociación viñedo-frutal, viñedos, asociación olivo-viñedo y zona de concentración parcelaria. Así, pues, el objetivo era enriquecer los mencionados usos con la leyenda obtenida a través de teledetección, especialmente en relación a los cultivos herbáceos, que como vemos incluye sólo un uso, tierra campá.

El uso del módulo CombiCapa de MiraMon permite la integración de dos capas ráster, una capa ráster y una vectorial o dos capas vectoriales. En nuestro caso, la opción fue la combinación entre una capa ráster (el mapa de cultivos del 2004) y otra vectorial (SIGPAC) que corresponde al enfoque mixto del mencionado módulo (figura 2). En este caso se calculan los valores estadísticos de los píxeles del ráster dentro de cada polígono y estos valores se transfieren a los registros correspondientes de la tabla principal del polígono, sin alterar la geometría de los objetos gráficos. En los límites del polígono, los valores del ráster se tendrán en cuenta según la combinación de los criterios del área y del punto central, o sea, si el punto central del ráster y la mayoría de su área se mantienen dentro del vector. Este aspecto es de una gran importancia ya que el efecto frontera puede alte-

rar los resultados. Los estadísticos implementados en un ráster categórico son: el número total de píxeles, la moda, el porcentaje de la moda sobre el total, el índice de Shannon (que calcula la diversidad de los valores), el valor mínimo, el valor máximo y el rango. En nuestro caso los campos retenidos fueron el número total de píxeles, la moda y el porcentaje total de la moda respecto al total de píxeles (excluyendo los NODATA). Los valores no agrícolas del ráster no se incluyeron en el análisis, tal como se ha detallado anteriormente.

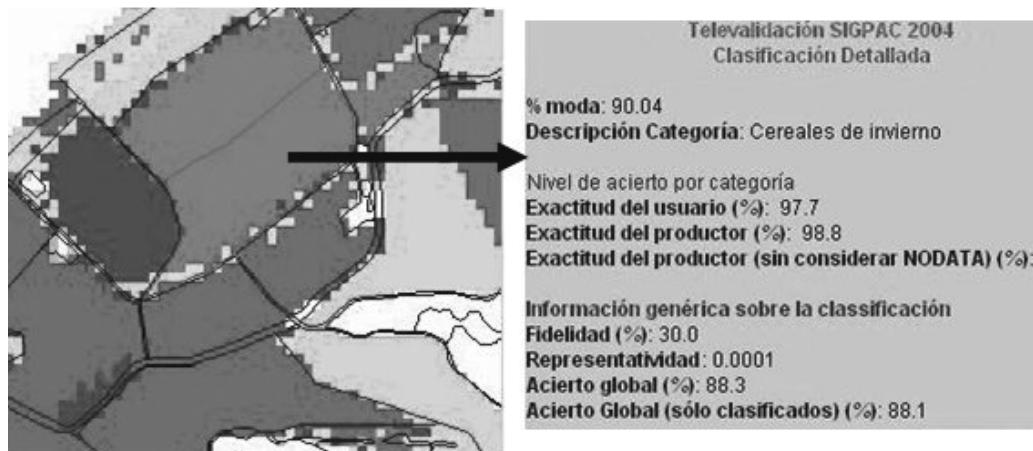
RESULTADOS

La figura 3 es un ejemplo de enriquecimiento de un recinto ocupado por cereales de invierno. Como se puede apreciar, el 90,04% de los píxeles estaban incluidos en ese cultivo (según el porcentaje de moda) dentro sus límites. Además se incluye el tipo de fidelidad y de representatividad, así como el acierto por categoría y el global. La figura 4 es la misma zona pero ya con el recinto enriquecido como producto final. Obsérvese la nitidez del resultado final comparada con la heterogeneidad del mapa de cultivos en formato ráster.

En el momento de aplicar el estadístico moda al conjunto del recinto SIGPAC, un aspecto fundamental es el umbral aplicado a partir del cual se enriquecerá el polígono. Porcentajes muy elevados (del 90, 80 o 70%) no comportan incertidumbres sobre su idoneidad, pero modas del 20, 30 o 40% posiblemente sí ya que valores bajos incrementan las dudas debido a la presencia de recintos más heterogéneos a causa del método de clasificación o al propio recinto (a ve-

Serra, P., Moré, G and Pons, X

FIGURA 3: Ejemplo de enriquecimiento de la capa SIGPAC, añadiéndose como nuevos campos la información generada. La flecha indica el recinto correspondiente a la información detallada.



ces por ejemplo puede darse combinaciones de cultivos como cereales de invierno y alfalfa, etc.). En algunos ejemplos de enriquecimiento de parcelas agrarias, por ejemplo en Martínez et al., 2001, al menos el 80% del polígono tenía que estar ocupado por un determinado cultivo. Por este motivo se detallarán los resultados en el caso de no aplicar ninguna restricción en el valor moda y de aplicar un valor umbral del 50%, o sea, que el valor moda del recinto fuese igual o superior al 50%, quedando en caso contrario el recinto sin clasificar (NODATA).

RESULTADOS GLOBALES

Una vez enriquecidos los recintos SIGPAC según una fidelidad de 0,31 y una de 0,51 y una representatividad de 0,01, se calculó la fiabilidad del mapa de cultivos (pixel a pixel) y

de los recintos enriquecidos del SIGPAC (polígono a polígono) a través de las matrices de confusión. En la tabla 1 se muestra un ejemplo de matriz; como se puede observar los resultados no ponderan la superficie de los distintos cultivos, opción que se considerará en futuros trabajos. Las matrices de confusión permiten obtener el porcentaje de exactitud temática global dividiendo el número total de píxeles o polígonos clasificados correctamente (la suma de la diagonal) entre el número total de píxeles o polígonos. La exactitud temática de los cultivos individuales incluye la exactitud del productor y del usuario. La exactitud del productor se obtiene de la división entre el número total de píxeles o polígonos correctamente clasificados en un determinado cultivo y el número total de píxeles o polígonos asignados a tal cultivo por las áreas de test (normalmente en las columnas). Esta exactitud indica los errores de omisión, es decir, los elementos que de-

FIGURA 4: Ejemplo de diversos recintos enriquecidos.

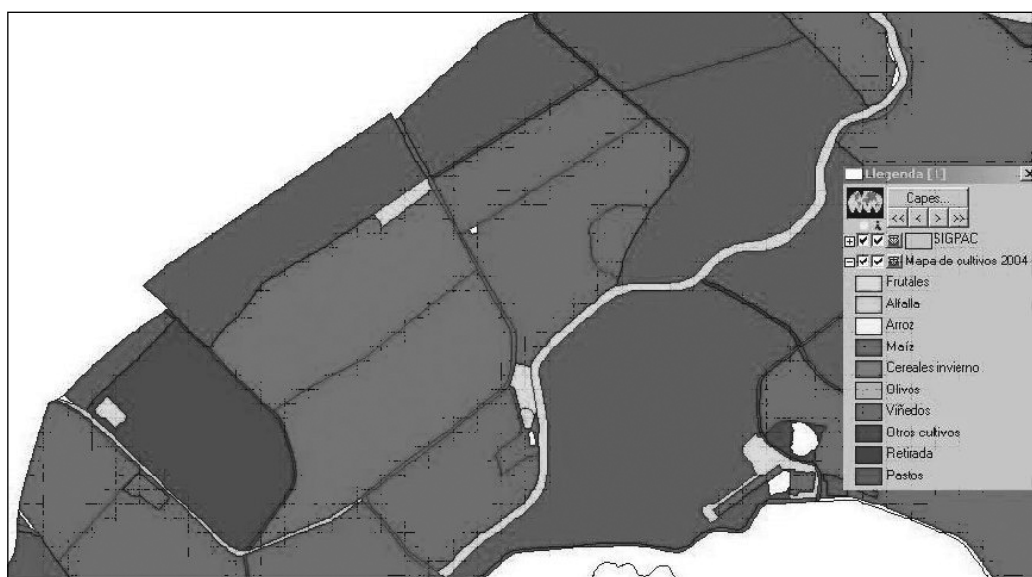


TABLA 1. Ejemplo de matriz de confusión, en píxeles, con una fidelidad de 0,51.

Verdad Terreno \ Clasificación	Cereales invierno	Arroz	Maíz	Alfalfa	Otros cultivos	Frutales	Olivos	Viñedos	Pastos	Barbecho	Total	EC (%)	EU (%)
Sin clasificar	141	2	67	366	292	1161	1143	6	292	1056	4526		
Cereales invierno	26823	0	2	8	41	0	0	0	84	493	27451	2.3	97.7
Arroz	0	84	0	0	0	0	0	0	0	0	84	0.0	100.0
Maíz	0	0	6129	0	15	1	0	9	1	17	6172	0.7	99.3
Alfalfa	9	0	12	7467	436	42	0	0	1	1	7968	6.3	93.7
Otros cultivos	0	0	0	157	655	135	0	15	17	15	994	34.1	65.9
Frutales	4	0	19	19	72	4513	199	48	31	367	5272	14.4	85.6
Olivos	0	0	5	0	0	95	166	0	0	105	371	55.3	44.7
Viñedos	0	0	1	0	0	71	5	615	0	30	722	14.8	85.2
Pastos	119	0	0	0	11	474	0	0	312	100	542	42.4	57.6
Barbecho	69	0	20	0	2	0	666	7	12	1908	3158	39.6	60.4
Total	27165	86	6255	8017	1524	6492	2179	700	750	4092	57260		
EO (%)	1.3	2.3	2.0	6.9	57.0	30.5	92.4	12.1	58.4	53.4			
EP (%)	98.7	97.7	98.0	93.1	43.0	69.5	7.6	87.9	41.6	46.6			
EP sin clasificar no como error (%)	99.3	100.0	99.0	97.6	53.2	84.7	19.2	88.6	68.1	62.8			

EO: error de omisión; EP: exactitud del productor; EC: error de comisión; EU: exactitud del usuario.

berían haberse clasificado como una categoría y se han clasificado como otra. Por otra parte, la exactitud del usuario se calcula de la misma forma pero considerando las filas, y obteniéndose los errores de comisión, es decir, los elementos que se han clasificado como una categoría que en realidad corresponden a otra. En la tabla 1, además, se calcula la exactitud del productor excluyendo los píxeles sin clasificar.

El umbral más ampliamente usado para aceptar una clasificación en teledetección corresponde a una exactitud temática global del 85% de los píxeles (Campbell, 2002). La tabla 2 sintetiza los resultados obtenidos; como se observa todos los porcentajes son superiores al 85%. Con la fidelidad 0,31 se clasifican más píxeles que con la fidelidad 0,51, concretamente un 15,2% más, mientras que, de los píxeles incluidos en el cálculo de la exactitud temática, un 7,6% se excluían en el caso de la fidelidad más restrictiva al permanecer como no clasificados. Aparentemente, pues, los resul-

tados eran peores con 0,51 ya que se clasificaba menos área y con menor exactitud temática global (del 88,2 al 85,1%). Pero, en realidad, si se excluyen como error los píxeles sin clasificar (generados por las dudas en CLSMIX), el mejor resultado se obtiene con la fidelidad 0,51, el 92,3% respecto al 88,5% con 0,31. En definitiva, clasificamos menos píxeles pero con mejor exactitud.

La tabla 3 muestra los resultados correspondientes al análisis de los polígonos enriquecidos. Como se observa, en la opción sin restricciones, aplicando una fidelidad de 0,31 el número total de polígonos clasificados es mayor (22), ya que la opción 0,51 genera más polígonos sin clasificar al ser más restrictivo, al igual que el número de polígonos bien clasificados (27 más). Los porcentajes de exactitud temática en un caso y otro son similares, 89,7 y 88,7 respectivamente, mientras que la exactitud sin considerar los polígonos no clasificados como error es prácticamente la misma en los dos ca-

TABLA 2. Exactitudes del mapa de cultivos en píxeles.

	Fidelidad 0,31 en píxeles	Fidelidad 0,51 en píxeles
Número total de píxeles clasificados	8 604 879	7 469 101
Número total de píxeles usados en el cálculo de la exactitud temática	57 260	57 260
Número total de píxeles usados en el cálculo de la exactitud temática excluyendo los sin clasificar	57 073	52 734
Número de píxeles bien clasificados	50 527	48 672
Exactitud temática total (%)	88,2	85,1
Exactitud temática total (%) excluyendo los píxeles sin clasificar	88,5	92,3

Serra, P., Moré, G and Pons, X

TABLA 3. Exactitudes del mapa de cultivos, en polígonos.

	Fidelidad 0,31 en polígonos	Fidelidad 0,51 en polígonos
Número total de polígonos clasificados	73 662	73 640
Número total de polígonos usados en el cálculo de la exactitud temática	2 819	2 819
Sin restricciones valor moda		
Número total de polígonos usados en el cálculo de la exactitud temática excluyendo los no clasificados	2 819	2 816
Número de polígonos bien clasificados	2 529	2 502
Exactitud temática total (%)	89,7	88,7
Exactitud temática total (%) excluyendo los polígonos no clasificados	89,7	88,8
Valor moda \geq 50%		
Número total de polígonos usados en el cálculo de la exactitud temática excluyendo los no clasificados	2 685	2 476
Polígonos no clasificados	134	343
Número de polígonos bien clasificados	2 483	2 366
Exactitud temática total (%)	88,1	83,9
Exactitud temática total (%) excluyendo los polígonos no clasificados	92,5	95,6

Los resultados de la tabla, además, resume los resultados en el caso de aplicar un valor de moda \geq 50%. Como se detalla, la exactitud temática global disminuye con una moda más restrictiva (ya que en los anteriores casos no se aplica un mínimo porcentaje de moda). En el caso de considerar una fidelidad de 0,31, el acierto disminuye hasta el 88,1%, mientras que en el caso de la fidelidad de 0,51 disminuye hasta el 83,9%. Contrariamente al caso sin restricciones, donde la exactitud temática sin considerar los polígonos no clasificados como error era prácticamente la misma, en el caso de excluir los polígonos no clasificados como error, los porcentajes incre-

mentan hasta el 92,5 y 95,6% respectivamente. Este hecho permite concluir que una fidelidad restrictiva y una moda restrictiva al 50%, excluyendo los polígonos sin clasificar como error, producen el acierto total significativamente más elevado.

RESULTADOS POR CULTIVOS

En el análisis final de los resultados por cultivos se presentarán sólo los que afectan a la exactitud del productor (los errores de omisión) ya que la exactitud del usuario (los

FIGURA 5. Resultados de la exactitud del productor con dos fidelidades (0,31 y 0,51) considerando y sin considerar los píxeles no clasificados como error. CI = cereales de invierno y OC = otros cultivos.

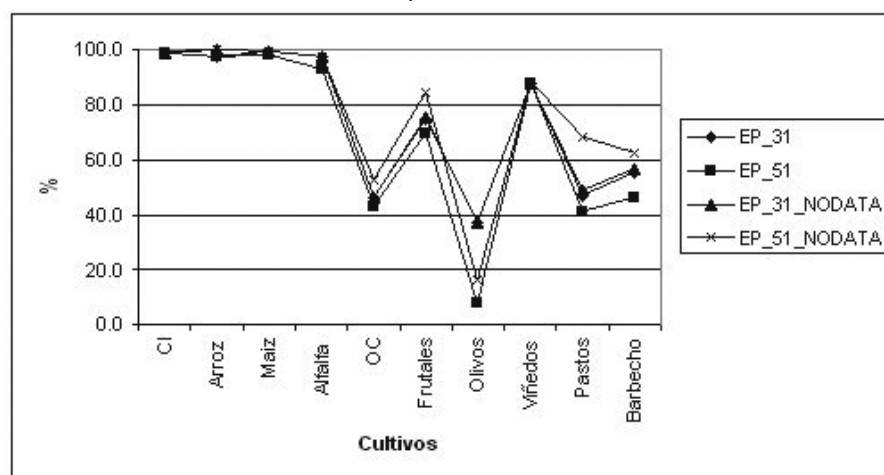
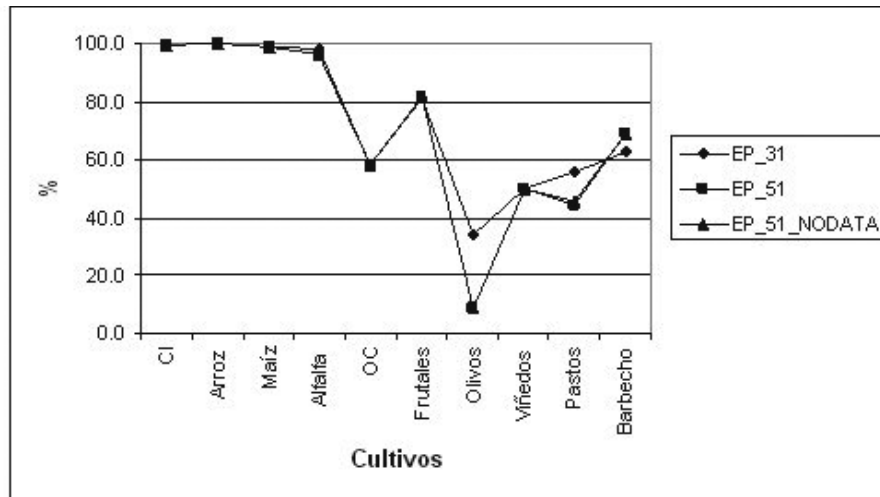


FIGURA 6. Resultados de la exactitud del productor con dos fidelidades (0,31 y 0,51) considerando y sin considerar los polígonos no clasificados como error (NODATA) en el caso de la fidelidad 0,51. CI = cereales de invierno y OC = otros cultivos.



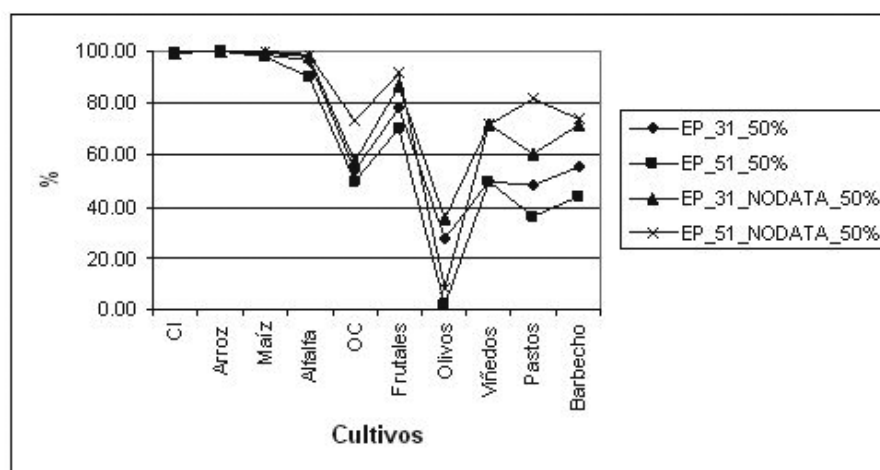
errores de comisión) no está afectada por los píxeles o polígonos sin clasificar y las diferencias entre las dos fidelidades son mínimas.

La figura 5 resume los resultados de la exactitud del productor con una fidelidad de 0,31 y 0,51 considerando los píxeles sin clasificar como error y sin error (opción NODATA). Como se aprecia, el porcentaje de acierto de los cultivos herbáceos es claramente superior al resto de cultivos, especialmente de los cereales de invierno, del arroz, del maíz y de la alfalfa donde los porcentajes son muy elevados y, por tanto, se ven débilmente afectados por la fidelidad y los píxeles sin clasificar. Sin embargo el grupo «otros cultivos» presenta unos valores manifiestamente inferiores en todos los casos, mientras que los frutales presentan los valores más elevados (> 80%) en el caso de una fidelidad de 0,51 sin considerar los

píxeles no clasificados como errores. Los valores de los viñedos también son elevados (> 80%) pero exactamente igual en todos los casos, mientras que en los olivos los mejores resultados son con una fidelidad de 0,31 tanto sin considerar los píxeles no clasificados como error o no. De todas formas, es el cultivo con los peores resultados. Finalmente, los resultados de pastos y barbecho mejoran con una fidelidad restrictiva y sin considerar los no clasificados como errores.

La figura 6 sintetiza los resultados de la exactitud del productor con una fidelidad de 0,31 y 0,51 considerando los polígonos sin clasificar como error y sin error (NODATA). En la figura sólo aparece representada la opción NODATA de la fidelidad 0,51, a pesar de presentar valores muy parecidos, ya que la fidelidad 0,31 presenta exactamente los mismos resultados. Como se observa, se mantiene el alto porcentaje de

FIGURA 7. Resultados de la exactitud del productor con dos fidelidades (0,31 y 0,51) considerando y sin considerar los polígonos no clasificados como error (NODATA) con una restricción de la moda igual o superior al 50%. CI = cereales de invierno y OC = otros cultivos.



Serra, P., Moré, G and Pons, X

acierto de los principales cultivos herbáceos respecto al resto de cultivos. Sin embargo, los resultados difieren en el caso de los olivos, pastos y barbecho, ya que en los dos primeros la mejor opción es la fidelidad de 0,31 mientras que en el tercero la mejor opción es la fidelidad de 0,51.

La figura 7 resume los resultados después de aplicar una restricción de la moda igual o superior al 50% y además considerando los píxeles no clasificados como error (NODATA) o no. En esta opción un polígono es clasificado si al menos el 50% de los píxeles clasificables tienen la misma moda. Como se observa, a parte de los cuatro cultivos que siempre se clasifican muy bien (cereales de invierno, arroz, maíz y alfalfa), los mejores resultados en el caso los otros cultivos, frutales y pastos se obtienen aplicando una moda restringida $\geq 50\%$ y no considerando los píxeles sin clasificar como error. En los viñedos y barbecho esta opción es muy parecida a la fidelidad 0,31, mientras que ésta última es claramente mejor en el caso de los olivos (a pesar de seguir siendo el cultivo peor clasificado).

CONCLUSIONES

La aplicación de la teledetección ha permitido la obtención de un mapa de cultivos actualizado y de una exactitud temática elevada. Como hemos visto, los valores de la fidelidad pueden ser más restrictivos o menos y, en consecuencia, los resultados pueden ser diversos. La conclusión más clara cuando se trabaja con píxeles es que una fidelidad menos restrictiva (0,31) produce inicialmente resultados más fiables pero un análisis más detallado nos indica que la opción más restrictiva (en nuestro caso la opción 0,51) es la que produce los resultados más fiables si los píxeles no clasificados no se consideran errores.

Otro aspecto importante es el porcentaje de moda aceptado en el momento de enriquecer un recinto SIGPAC. Como hemos visto, el porcentaje aplicado ($\geq 50\%$) afecta los resultados debido a que sin restricción la presencia de un solo píxel dentro de un polígono (cuya presencia puede deberse al efecto de frontera o al efecto «sal y pimienta») ya es suficiente para clasificarlo totalmente. Con ello evitamos tales efectos, filtrándose la contaminación de píxeles aislados, concordando con la hipótesis inicial del trabajo.

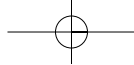
Otra conclusión es que el porcentaje global elevado de acierto se debe a la gran fiabilidad obtenida en la alfalfa, arroz, maíz y cereales de invierno. Con las imágenes empleadas estos cultivos son plenamente diferenciados. Al ser cultivos muy bien discriminados, la fidelidad y el porcentaje de la moda adoptado no les afecta tanto como al resto de cultivos. En el caso de los frutales, la opción enriquecida NODATA y una moda $\geq 50\%$ es la que presenta mejores resultados posiblemente debido a que son parcelas con una elevada presencia de píxeles frontera y de píxeles impuros o mezclados con otros cultivos. Esta misma opción produce

también los mejores resultados en el caso de los otros cultivos, pastos y barbecho (éste último en menor medida). De todos los cultivos clasificados los únicos cultivos que presentan un porcentaje de exactitud peor en el caso de un enriquecimiento del SIGPAC son los viñedos, claramente, y los olivos posiblemente debido a la atomización de las parcelas y a las dudas en el clasificador que provocan elevada confusión con otros cultivos.

Por todo ello, la opción de enriquecimiento de información vectorial con el módulo CombiCapa de MiraMon a través de una fidelidad restrictiva y con un porcentaje elevado de moda, sin tener en cuenta los píxeles no clasificados como error, parece una buena opción para obtener como producto final de mayor calidad.

BIBLIOGRAFÍA

- APLIN, P., ATKINSON, P.M. y CURRAN, P.J.: «Fine spatial resolution simulated satellite sensor imagery for land cover mapping in the United Kingdom». *Remote Sensing of Environment*, 68, (1999), pp. 206-216.
- BARBOSA, P.M., CASTERAD, M.A. y HERRERO, J.: «Performance of several Landsat 5 Thematic Mapper TM image classification methods for crop extent estimates in an irrigation district». *International Journal of Remote Sensing*, 17, (1996), pp. 3665-3674.
- BURROUGH, P. y McDONNELL, R.: *Spatial information systems and geostatistics*. Oxford, University Press, 1998.
- CAMPBELL, J.B.: *Introduction to remote sensing*. New York, The Guilford Press, 2002.
- CENTRO DE INVESTIGACIÓN ECOLÓGICA Y APLICACIONES FORESTALES (CREAF), Mapa de Cubiertas del Suelo de Catalunya. Disponible en: <http://www.creaf.uab.es/mcsc/>
- CONGALTON, R.G., BALOGH, M., BELL, C., GREEN, K., MILLIKEN, J.A. y OTTMAN, R.: «Mapping and monitoring agricultural crops and other land cover in the Lower Colorado River Basin». *Photogrammetric Engineering and Remote Sensing*, 64, (1998), pp. 1107-1113.
- CRIST, E. P., LAURIN, R. y CICONE, R.C.: «Vegetation and soils information contained in transformed Thematic Mapper data». En *Proceedings of IGARSS' 86 Symposium*, Paris, European Space Agency, (1986), pp. 1465-70.
- CHUVIECO, E. Teledetección ambiental. Barcelona, Ariel, 2002
- DEPARTAMENTO DE AGRICULTURA, GANADERÍA Y PESCA (DARP) (2006). SIGPAC: Sistema de Información Geográfica. <http://www.gencat.net/darp/c/serveis/sigpac/csigpac1.htm>
- DE WIT, A.J.W. y CLEVERS, G.P.W.: «Efficiency and accuracy of per-field classification for operational crop mapping». *International Journal of Remote Sensing*, 25, (2004), pp. 4091-4112.
- FOODY, G.M.: «Status of land cover classification accuracy assessment». *Remote Sensing of Environment*, 80, (2002), pp. 185-201.
- KAUTH, R.J. y THOMAS, G.S.: «The tasselled cap—a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat». En: *LARS: Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*. West Lafayette, Purdue University Press, (1976), 4B-41-4B-51.
- KHORRAM, S.: *Accuracy assessment of remote sensing-derived change detection*. Maryland, American Society for Photogrammetry and Remote Sensing, 2004.
- LYON, J.G.; WARD, A.; ATHERTON, B.C.; SENAY, G. y KRILL, T.: «Remote sensing and GIS for site-specific farming». En: Lyon, J.G. (Ed.): *GIS for water resources and watershed management*. Boca Raton, CRC Press, (2003), pp. 251-253.
- MINISTERIO DE AGRICULTURA, PESCA Y ALIMENTACIÓN (MAPA) (2006). Sistema de Información Geográfica de la Política Agraria Común (SIGPAC). <http://www.mapa.es/es/sig/pags/sigpac/intro.htm>



CONSECUENCIAS EN LA CARTOGRAFÍA DE CULTIVOS MEDITERRÁNEOS DE LA COMBINACIÓN DE DATOS RÁSTER-VECTOR: ENRIQUECIMIENTO DE UNA BASE ...

- MARTINEZ, C. y CALERA, A.: «Irrigated crop area estimation using Landsat TM imagery in La Mancha, Spain». *Photogrammetric Engineering and Remote Sensing*, 67, (2001), pp. 1177-1184.
- PALÀ, V., y PONS, X.: «Incorporation of relief in polynomial-based geometric corrections». *Photogrammetric Engineering and Remote Sensing*, 61, (1995), pp. 935-944.
- PESQUER, LL, MASÓ, J. y PONS, X.: «Herramientas de análisis combinado raster/vector en un entorno SIG». *IX Congreso del Grupo de Métodos Cuantitativos, Sistemas de Información Geográfica y Teledetección*, 20-22 Septiembre 2000. Alcalá de Henares, Madrid.
- PONS, X. (2006). MiraMon. Geographic Information System and Remote Sensing software. Centre de Recerca Ecològica i Aplicacions Forestals, CREAf. Bellaterra. <http://www.creaf.uab.es/miramon>
- PONS, X. y SOLÉ-SUGRAÑES, L.: «A simple radiometric correction model to improve automatic mapping of vegetation from multispectral satellite data». *Remote Sensing of Environment*, 48, (1994), pp. 191-204.
- SERRA, P., MORÉ, G. y PONS, X.: «Application of a hybrid classifier to discriminate Mediterranean crops and forests. Different problems and solutions». En: *Proceedings of the XII International Cartographic Conference*, 9-16 Julio 2005, A Coruña, CD-ROM.
- SERRA, P., PONS, X. y SAURÍ, D.: «Post-classification change detection with data from different sensors. Some accuracy considerations». *International Journal of Remote Sensing*, 24, (2003), pp. 3311-3340.

