

## Método de Clasificación «Por-parcela» para la Detección de Cultivos: Aplicación al Área Agrícola de Las Lajitas (Salta, Argentina), Campaña Agrícola de Invierno 2006.

Volante, José N.; Campos, Cristian J.; Noé, Yanina E. y Elena, Hernán J.

Instituto Nacional de Tecnología Agropecuaria (INTA), Estación Experimental Agropecuaria Salta, Laboratorio de SIG y Teledetección.

Ruta Nacional 68 Km 172 (4403) Cerrillos, Salta, Argentina.

Tel: +54 387 4902087. e-mail: jvolante@correo.inta.gov.ar

### RESUMEN

Se realizó el mapa de los cultivos de invierno de 2006, de un área agrícola de 720.000 ha, de Salta, (Argentina) mediante clasificación «por parcela». Este algoritmo permite integrar información de imágenes satelitales (multi-temporal y multi-sensor), con cartografía digital correspondientes a los contornos de las parcelas de cultivos. Se analizan las ventajas relativas del método y se propone una nueva forma de analizar las fuentes del *Error Global* de la cartografía, particionando éste en dos componentes *Error de Asignación Temática* (inherente a la clasificación) y el *Error de Digitalización de Parcela* (inherente a la cartografía vectorial utilizada).

Palabras claves: Clasificación por parcela, fiabilidad cartográfica, mapa de cultivos.

### ABSTRACT

The map of the 2006's winter cultures of an agricultural area of Salta (Argentina) was made by-parcel classification. This algorithm allows to integrate information of satellite images (multi-temporary and multi-sensor), with digital cartography corresponding to the parcels of cultures. The relative advantages of the method are analyzed. A new form to analyze the sources of the Global Error of the cartography, was made. There are two components: Thematic Allocation Error (inherent to the classification) and the Parcel Digitalization Error (inherent to the used cartography).

Key words: Classification by parcel, accuracy assessment, map of cultures.

### Introducción

El Laboratorio de Teledetección del INTA Salta en convenio con la CONAE (Comisión Nacional de Actividades Espaciales, Argentina) realiza desde el año 2000, cartografía de cultivos de las provincias del Noroeste Argentino (NOA: Jujuy, Salta, Tucumán, Santiago del Estero y Catamarca), (Volante et al., 2005). Con frecuencia semestral, se generan mapas de cultivos de secano de una extensa y heterogénea región del país, que comprende 5,3 millones de hectáreas agropecuarias, de las cuales 2,8 millones son destinadas a la agricultura, (<http://www.inta.gov.ar/prorenea>).

Los principales problemas presentados en la rutina operativa del mencionado trabajo, están estrechamente vinculados a los métodos con-

vencionales de clasificación (clasificación por-píxel), los que imponen un alto grado de interacción operador - datos. Este hecho trae aparejado que el sistema sea altamente dependiente de la capacidad del personal técnico y del tiempo operativo requerido para lograr productos cartográficos de calidad.

#### Clasificación orientada a objetos

En la década de los 70 ha surgido un nuevo concepto, el análisis de imágenes orientado a objetos, puesto en práctica con el desarrollo de software especializado diez años después, (Flanders, 2003). Este tipo de clasificación se basa en la identificación de objetos, definidos como, grupos de píxeles contiguos que poseen un comportamiento similar, de manera que, la variabilidad entre los píxeles que lo componen,

es inferior a la variabilidad entre el objeto y su entorno. Este concepto, impulsó el desarrollo de algoritmos y software con diferentes grados de complejidad (Flanders, 2003).

Más allá del método de clasificación orientado a objetos (como metodología analítica) el concepto elemental de «objeto», a sido utilizado de diversas formas (Dean, 2003; Flanders, 2003; Thunnissen, 2000; Turker, 2005).

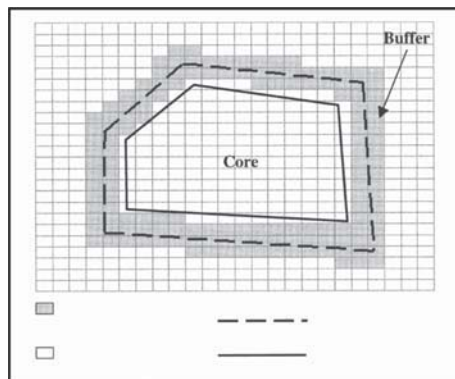
### Clasificación por-parcela

La clasificación convencional de imágenes de satélite se basa en la asignación de los píxeles individuales a determinadas clases temáticas. Este sistema además de su ineficiente capacidad de análisis del contexto, presenta un tipo de problema asociado a la resolución espacial de las imágenes. Los píxeles de borde o mixtos, representan mezclas de dos o más tipos de coberturas. Esto es motivo de las principales causas de error de asignación, y por consecuencia de la pérdida de exactitud en la cartografía temática generada. (Fisher 1997), (Fisher and Pathirana 1990). (Markham and Townsend 1981, Veregin and Sultana 1993).

Una solución práctica a este problema es aportada por el trabajo de Dean (2003), el que brinda una aproximación operativa utilizando el concepto de objeto, para la detección de cultivos dentro de las parcelas, aprovechando las características intrínsecas de los mismos: cubiertas vegetales, coetáneas y homogéneas que se desarrollan en espacios regulares de formas geométricas definidas (parcelas), las que son relativamente estables en el tiempo (años).

Esta metodología se basa en la clasificación de las imágenes utilizando información estadística descriptiva (media y varianza) de los píxeles de la zona núcleo de las parcelas, esto es, la zona central libre de píxeles de borde o mixtos. (Fig 1). Esta rutina de clasificación, utilizada de manera experimental en otras regiones del mundo, integra información de sensores remotos con cartografía digital (Dean, 2003).

Con la finalidad de encontrar una metodología operativa que permita realizar las rutinas de procesamiento de imágenes con la mayor eficiencia posible (máxima calidad cartográfica en menor tiempo operativo posible), se realizó un análisis comparativo de dos procedimientos: uno convencional de clasificación «por-píxel», y otro utilizando el concepto de clasificación «por-parcela». A continuación se presentan los resultados del análisis y se discuten las ventajas comparativas.



**Fig. 1.** Clasificación por-parcela. El tamaño del buffer es controlado por la distancia de contracción vinculada al tamaño del píxel. Sólo los píxeles alojados en el área núcleo, son utilizados en la caracterización estadística de la parcela. Tomado de Dean (2003).

### Materiales y método

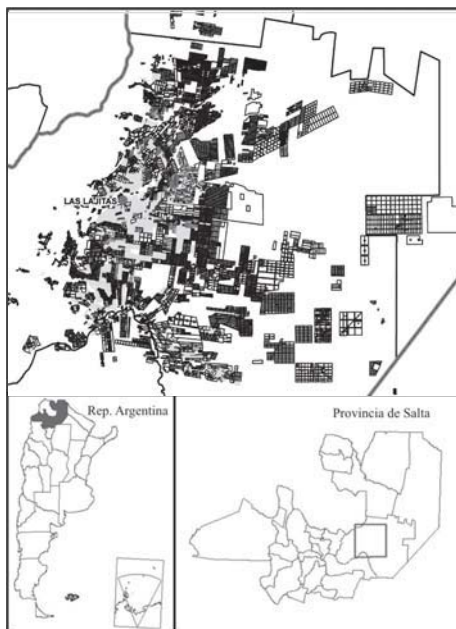
#### Área de estudio

El presente trabajo se realizó en la zona agrícola más importante de la Provincia de Salta (Argentina), ubicada entre los 24°00' y 25°30' de Latitud S y entre los 64°45' y 63°00' de Longitud O, en cuyo centro geográfico se encuentra la localidad de Las Lajitas. El área de estudio representa una superficie de 21.000 km<sup>2</sup>, en el que se desarrolla una importante actividad agropecuaria con una superficie de 720.000 ha.

Este parche se considera representativo de una extensa zona agro-ecológica denominada Umbral al Chaco la que ocupa la zona de transición entre las primeras estribaciones de las Sierras Subandinas al Oeste y la Llanura Chaqueña al Este, (Bravo, 1999).

El clima se caracteriza por poseer un régimen de lluvias estival muy concentrado, con un promedio anual que va de 600 a 800 mm. Las precipitaciones van en aumento de Este a Oeste por razones orográficas. Las elevadas temperaturas de verano provocan déficit hídricos que varían entre los 200 y 500 mm/año.

Los suelos tienen aptitud para agricultura de secano (73 %) con distintos grados de limitación: agrícola - ganadera (9 %), ganadero - agrícola (11 %) y ganadera (7 %).



**Fig. 2.** Las Lajitas, (Salta, Argentina). Área de estudio y mapa de cultivos de invierno, campaña agrícola 2006. En gris claro: trigo; gris oscuro: cártamo; negro: límites de parcelas no cultivadas.

La principal actividad de la zona de estudio, es la agricultura de secano. En verano se cultiva soja, maíz en sistemas de siembra directa, y poroto en menor proporción, mediante sistemas de labranza convencional. En invierno se produce trigo y cártamo, que por las características climáticas ya descritas, son altamente dependientes de la humedad almacenada en el suelo en el período estival. Esto hace que los rendimientos zonales promedio del cultivo de trigo, se encuentren en el orden de 800 kg/ha/año, (Volante, 2005).

La ganadería vacuna de cría y engorde completan el mosaico del paisaje agropecuario de esta zona, actividades que se encuentran en franco crecimiento. Durante los años 2005 y 2006 se han deforestado 127.000 ha, para emprendimientos agropecuarios.

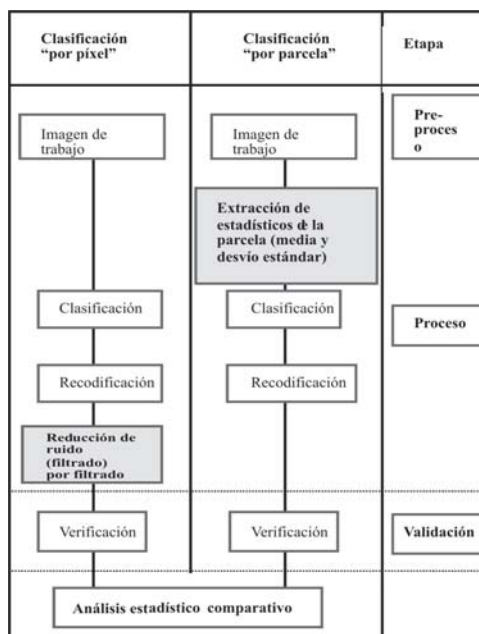
#### Esquema general del trabajo

A fin de poder analizar las ventajas que ofrece la rutina de clasificación «por parcela», se han realizado dos ensayos.

En primer lugar, con el objetivo de comparar sistemas de clasificación, se efectuaron dos rutinas (una por píxel y otra por parcela), sobre un mismo conjunto de datos multi-temporal. Para ello se construyó una «imagen de trabajo» formada por el apilamiento de las bandas 3, 4 y 5 del satélite Landsat 5 TM, de dos fechas distintas, [(26/08 y 11/09), (Tabla 1)]. En la figura 3 se presenta una sinopsis general en la que se muestran las principales etapas de ambos procesos.

**Tabla 1.** Imágenes utilizadas en el análisis.

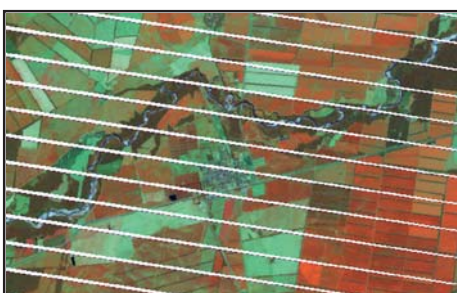
Imagen	Path/Row	Fecha
Landsat 7, SLC-off	230/077	17/07/2006
Landsat 5	230/077	26/08/2006
Landsat 5	230/077	11/09/2006



**Fig. 3.** Esquema general de los sistemas de clasificación «por píxel» y «por parcela»

En segunda instancia con el fin de analizar la potencialidad del método por parcela, se clasificó una imagen trabajo multi-temporal compuestas por 3 imágenes de dos sensores distintos. Para ello se agregó, a la imagen de trabajo utilizada en el primer ensayo, las bandas 3, 4 y 5, del satélite **Landsat 7 ETM+ (nivel 1G Standard,**

**SLC-off).** Es necesario mencionar que este producto contiene bandas de datos faltantes (valores cero), producidas por el mal funcionamiento de un instrumento denominado *Scan Line Corrector*, (Fig. 3). Desde el 11 de octubre de 2003, este instrumento está operando en modo «off», ([http://landsat7.usgs.gov/data\\_products/slc\\_off\\_data\\_products/index.php](http://landsat7.usgs.gov/data_products/slc_off_data_products/index.php))



**Fig. 4.** Imagen Lansat 7 ETM+ (nivel 1G Standard, SLC-off), del 17/07/2006. En el centro puede apreciarse la Localidad rodeada de parcelas agrícolas.

### Leyenda

La leyenda cartográfica fue definida a priori describiendo la totalidad de la diversidad de cultivos extensivos de secano que se realizan en la zona de estudio. La misma quedó definida en 4 categorías:

- \* Trigo (*Triticum aestiva*).
- \* Cártamo (*Cartamus tinctoria*).
- \* Barbecho (suelo no cultivado, en descanso) y pasturas para ganadería.
- \* Otros, no agropecuario (ambiente natural y modificado).

En la etapa de pre-procesamiento, las imágenes de satélite fueron georreferenciadas, corregidas radiométricamente y re-escaladas a ocho bit. Posteriormente se unieron las imágenes, en una sola imagen de trabajo mediante apilado de bandas.

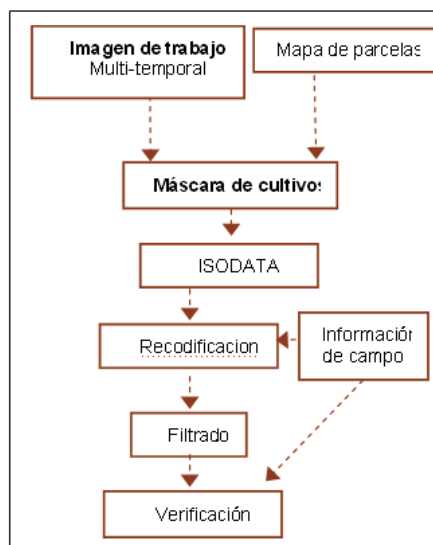
#### Mapa de parcelas

La digitalización del mapa de parcelas se realizó a partir de la utilización de imágenes pancromáticas Landsat 7 ETM+ del año 2002, (sin bandas faltantes, o sea con SLC-on). Se utilizaron estas imágenes para aprovechar la resolución espacial de las mismas (15 m). Se hizo una clasificación no supervisada y posterior re-

codificación de clases a fin de separar parcelas agropecuarias del resto del paisaje. Al resultado de ese proceso, se pasó a formato vectorial y mediante interpretación visual se digitalizaron modificaciones, actualizaciones y correcciones para lograr un mapa de polígonos de parcelas del área de estudio. La interpretación visual se realizó en pantalla a escala 1:75.000, sobre imágenes Landsat 5 de 2006, en composición RGB 453.

### Clasificación

En la figura 5 se describen los pasos realizados en la etapa de procesamiento de imágenes siguiendo la rutina de clasificación «por píxel». La clasificación no supervisada se realizó con 30 clases, y 99% de convergencia. Para recodificar las clases en las categorías establecidas como leyenda, se utilizó como referencia información de campo. Finalmente se aplicó un filtro de mayoría de ventana 3 x 3.



**Fig. 5.** Rutina de clasificación «por-píxel».

En la rutina de clasificación «por parcela» (Fig. 6), el mapa de los núcleos o zonas centrales de las parcelas se hizo a partir de realizar un buffer «hacia adentro» de 60 m, (2 píxeles), siguiendo el criterio utilizado por Dean (2003). Con este «mapa buffer» se extraen los datos estadísticos de los lotes (media y desviación estándar) con los que se realiza la clasificación. Se aplicó

así el algoritmo ISODATA con 30 clases y 99% de convergencia. Para recodificar las 30 categorías se utilizó información de campo transfiriendo la información del mapa buffer clasificado, al mapa de parcelas. Esta etapa se denomina reconstrucción de parcela.

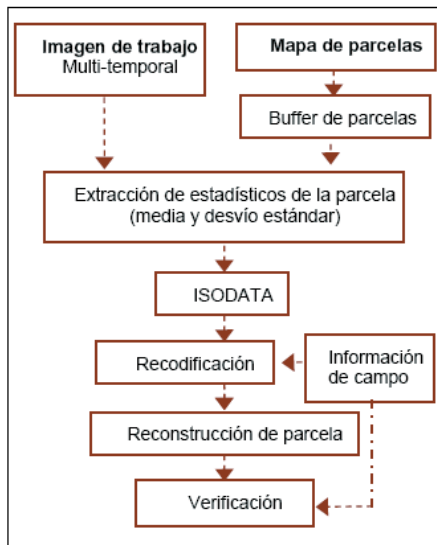


Fig. 6. Rutina de clasificación «por-parcela».

#### Verificación y análisis estadístico comparativo

Se recorrieron 173 km, por rutas y caminos vecinales del área de estudio (Fig. 6), en dónde se realizaron observaciones exhaustivas sobre la presencia/ausencia de cultivos (trigo y cártamo), barbechos o pasturas, y otros paisajes no agrícolas. En gabinete se digitalizaron los recorridos realizados sobre un área de influencia (buffer) de 800 m con lo que se generaron fajas de «verdad de campo» de 1600 m de ancho, que representan 27.740 ha, equivalentes a un 1,3% del área de estudio. Dentro de estas fajas se digitalizaron las parcelas o lotes agrícolas asignando a los polígonos la información observada en el campo. Como base para la digitalización de las parcelas se utilizaron imágenes georreferenciadas de alta resolución publicadas en Internet por Google Earth®.

Esta es una modificación al método descrito por Belart, (2001) en donde la información descrita por el autor es puntual. El mapa de cultivos digitalizado en las fajas de verdad de cam-

po, y se lo comparó mediante superposición cartográfica (overlay) con los mapas generados por las distintas rutinas de clasificación. Con los resultados se construyeron matrices de confusión, y se compararon los resultados a partir del estadístico Kappa, (Congalton, 2001).



Fig. 7. Itinerario realizado en el área de estudio y la faja de verificación. Rojo: trigo; verde: cártamo; amarillo: barbecho-pasturas; blanco: no agrícola

#### Resultados y discusión

La digitalización del mapa de parcelas se realizó en 80 horas de trabajo. Como resultado se obtuvo un mapa de 7.630 polígonos (parcelas) con una superficie promedio de 70 ha ( $\bar{A} \pm 54$  ha). (Figura 7). Para estimar la exactitud del mapa de lotes, se estimó el Error Cuadrático Medio (RMSE) cuantificado en 36 m.

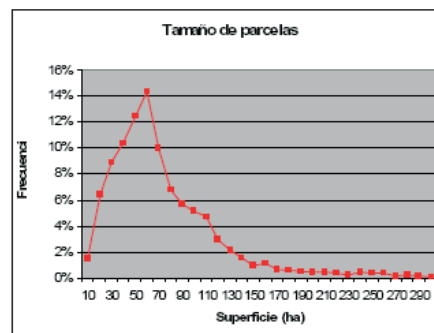


Fig. 8. Frecuencia relativa de tamaño de parcelas. Las Lajitas (Salta).

Con las rutinas descritas en el apartado «*metodología*», se obtuvieron 3 mapas de cultivos cuyas superficies, estimadas de cada clase presentan en la Tabla 2. El mapa obtenido mediante clasificación por píxel muestra lotes en los cuales se ve claramente el efecto «sal y pimienta» o «moteado» producido por la asignación errónea de píxeles o pequeños grupos de píxeles. Por este método la superficie asignada a cada clase es continua teniendo como unidad menor al píxel. En cambio, en los otros mapas, realizados mediante clasificación por parcela, la asignación de clases es discreta (polinomial), donde todo el lote es asignado a una clase solamente.

Los resultados de la verificación de los mapas se presentan en las Tablas 3, 4, 5 y 6. Al analizar las posibles fuentes de error del método por parcela, se observó que existe un tipo de error inherente al mapa de parcelas (RMSE = 36 m). Este error es cuantificable en términos de *Fiabilidad Global* (por medio de la matriz de confusión), por comparación del mapa digitalizado con el mapa del área de verificación (verdad de terreno). La *Fiabilidad Global* del mapa de lotes se estimó de esa forma en 90,6%. Este porcentaje representa el máximo valor que podría haberse obtenido mediante clasificación por parcela utilizando ese mapa de lotes. Expresado en términos del error (1- *Fiabilidad*), tenemos:

$$E_{Total} = E_{Mapa} + E_{Clas} [1]$$

En donde:

***E<sub>Total</sub>***: Error cartográfico total

***E<sub>Mapa</sub>***: Error de digitalización

***E<sub>Clas</sub>***: Error de asignación o clasificación

**Tabla 2.** Campaña agrícola de invierno 2006. Superficie en hectáreas.

Método	Trigo	Cártamo	Barbecho - Pasturas
Por-píxel	76.236	8.584	635.180
Por-parcela 2 fechas	70.780	11.935	637.285
Por-parcela 3 fechas	74.971	12.914	632.115

**Tabla 3.** Exactitud de los mapas.

Método de clasificación	Fiabilidad Global	I. C.
Por píxel 2 fechas	77,4%	±0,49%
Por parcela 2 fechas	80,8%	±0,46%
Por parcela 3 fechas	80,1%	±0,47%

**Tabla 4.** Matriz de confusión. Clasificación por píxel de imágenes Landsat 5 de 2 fechas.

Verdad Campo	Clasificación			Total	
	No agropecuario	Trigo	Cártamo		Barbecho-pasturas
No agropecuario	5620	909	5	1193	7727
Trigo	314	8135	118	1189	9756
Cártamo	33	740	145	34	953
Barbecho-pasturas	278	1427	22	7579	9305
Total	6245	11212	292	9997	27740

**Tabla 5.** Matriz de confusión. Clasificación por parcela de imágenes Landsat 5 de 2 fechas.

Verdad campo	Clasificación			Total	
	No agropecuario	Trigo	Cártamo		Barbecho-pasturas
No agropecuario	6066	781	81	796	7725
Trigo	404	8293	270	789	9756
Cártamo	50	266	603	34	953
Barbecho-pasturas	477	1300	71	7458	9307
Total	6998	10641	1026	9081	27740

**Tabla 6.** Matriz de confusión. Clasificación por parcela de imágenes Landsat 5 y 7 de 3 fechas.

Verdad campo	Clasificación			Total	
	No agropecuario	Trigo	Cártamo		Barbecho-pasturas
No agropecuario	6066	795	78	785	7725
Trigo	404	8430	208	713	9756
Cártamo	50	192	677	35	953
Barbecho-pasturas	477	1667	109	7054	9307
Total	6998	11084	1072	8586	27740

De acuerdo a la ecuación [2] los errores de clasificación de los métodos por parcela con imágenes de 2 y 3 fechas, serían 9,8% y 10,5% respectivamente. En el método por píxel solo se puede estimar el error total (22,6%).

En la figura 9 puede apreciarse claramente que en el método por píxel, el error de omisión del *cártamo* es una importante fuente de error total, en donde la principal confusión se produce con el trigo (Tabla 4). Los métodos por parcela mejoran significativamente la asignación de esa categoría. Esto podría deberse a que el método por parcela toma en cuenta la variabilidad de los niveles digitales de los píxeles en los lotes mediante el desvío estándar como variable de clasificación. El cultivo de cártamo, al tener una cobertura poco densa presenta una gran heterogeneidad de píxeles dentro de los lotes. El agregado de una tercera imagen reduce significativamente los errores de omisión y comisión del cártamo

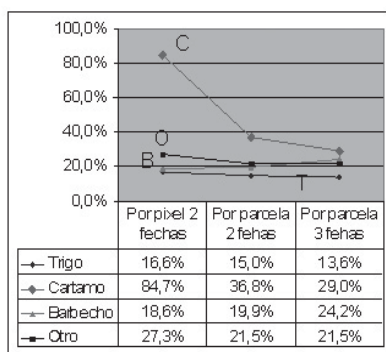


Fig. 9. Errores de Omisión.

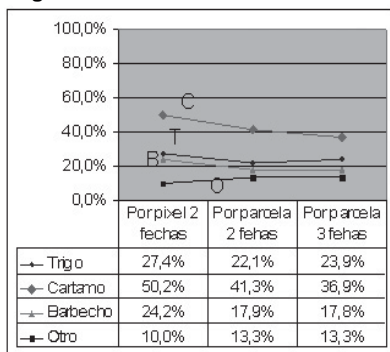


Fig. 10. Errores de Comisión.

Tabla 7. Análisis de calidad cartográfica mediante el estadístico Kappa.

Método de clasificación	Kappa	Varianza K	P
1) Por píxel 2 fechas	66,5%	1,210 E-05	< 0,001
2) Por parcela 2 fechas	71,9%	1,090 E-05	< 0,001
3) Por parcela 3 fechas	70,9%	1,114 E-05	< 0,001
Comparación		Z	P
1) Vs 2)		11,34	< 0,001
2) Vs 3)		2,13	0,01645

Los métodos por parcela han generado cartografía de cultivos con exactitud cartográfica significativamente superior al método por píxel. Las diferencias al incorporar una tercera imagen al análisis no son significativas para un nivel de confianza del 99%, (Tabla 7).

### Conclusiones

El resultado más importante de este trabajo es haber podido comprender las potencialidades del método de clasificación por parcela, para la realización de cartografía de cultivos, más allá de haber encontrado diferencias significativas en la exactitud cartográfica.

Una de las principales ventajas encontradas, en comparación con el método convencional de clasificación por píxel, es que el mismo permite la integración de información multi-temporal y multi-sensor. Esto es importante si se piensa que este método permite combinar imágenes de distintas resoluciones radiométricas y espaciales (tamaños de píxel). Esto es posible debido a que se trabaja con los estadísticos del lote. Este hecho permite utilizar la información de las imágenes Landsat 7 ETM+, de utilización restringida.

Permite por otra parte, conocer con mejor precisión las fuentes del error al poder medir la calidad de la cartografía de parcelas de forma independiente de su asignación temática.

Bajo el punto de vista operativo es necesario destacar dos aspectos. Un mismo mapa de parcelas (con pequeñas actualizaciones), puede ser utilizado durante largos períodos de tiempo, para confeccionar mapas de cultivos en forma rutinaria. Esto permitiría implementar Sistemas de Información Geográfica para evaluar la dinámica de cultivos. En los análisis de cambios multi-temporales, los errores de sus distintas capas, estarían vinculados solo a errores de asignación (errores de atributo). Los errores inherentes a los bordes (de las parcelas), sería igual a cero.

Un segundo aspecto es el relativo al tiempo. La confección de la cartografía de parcelas se recupera rápidamente, debido a que la rutina operativa es mucho más rápida que la realizada mediante clasificación por píxel.

### Agradecimientos

Agradecemos a la Comisión Nacional de Actividades Espaciales (CONAE) por la información satelital aportada para la realización del presente trabajo y en especial al Dr. Conrado Varotto, la Ing. Monica Rabolli y la Srta. Andrea Gullich, quienes siempre han brindado de manera desinteresada, su colaboración para la prosecución de actividades de investigación que desarrolla INTA Salta.

### Referencias

- Bravo, G.; Bianchi, A.R.; Volante, J.N.; Alderete, S.; Sempronii, G.; Vicini, L.; Fernández, M.; Lipshitz, H. y Piccolo, A. 1999. Regiones Agro-económicas del Noroeste Argentino. Primeras Jornadas de SIG. INTA-UNSE, Gobierno de Santiago del Estero. Disco Compacto.
- Belart, P.; Pinilla, C.; Ariza, F.J., 2001. Control de calidad de clasificaciones de imágenes por el método de los itinerarios. En: <http://www.mappinginteractivo.com>
- Congalton, R. and Green, K. 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Taylor and Francis Group. Boca Raton. 137 pp.
- Dean, A. M. 2003. An evaluation of per-parcel land cover mapping using maximum likelihood class probabilities. International Journal of Remote Sensing. 1-16, preview article.
- Flanders, D.; Hall-Beyer, M. and Pereverzoff, J. 2003. Preliminary evaluation of eCognition object-based software for cut block delineation and feature extraction. Can. J. Remote Sensing, Vol. 29, No. 4, pp. 441-452.
- Ming, D.; Luo, J.; Shen Z. 2005. Features Based Parcel Unit Extraction from High Resolution Image. IEEE, 1875-1878.
- Paoli, H.; Volante, J.; Noé, Y.; Campos, C. Vale, L.; Elena, H.; Morales, C. 2007. Monitoreo de cultivos extensivos del Noroeste Argentino a partir de sensores remotos. Campaña agrícola 2006. Cultivos de invierno. Instituto Nacional de Tecnología Agropecuaria (INTA), Estación Experimental Agropecuaria Salta; Argentina. <http://www.inta.gov.ar/prorenea>
- Thunnissen, H.; Wit, A. 2000. The National Land Cover Database of the Netherlands. ISPRS, Vol. XXXIII.
- Turker, M.; Arıkan, M. 2005. Sequential Masking Classification of Multi-temporal Landsat7 ETM+ images for field-based crop mapping in Karacabey, Turkey. International Journal of Remote Sensing, Vol. 26, Nº 17, 3813-3830.
- Volante, J.; Bianchi, A.; Paoli, H.; Noé, Y.; Elena, H. 2005. Análisis de la Dinámica del Uso del Suelo Agrícola del Noroeste Argentino Mediante Teledetección y Sistemas de Información Geográfica. Período 2000-2005. Instituto Nacional de Tecnología Agropecuaria (INTA), Estación Experimental Agropecuaria Salta; Argentina. <http://www.inta.gov.ar/prorenea>