

Determinación de uso del suelo mediante imágenes LandSat Campaña 2008/9 – Departamentos Marcos Juárez y Unión

Ing. Agr. Florencio Moore
Área de Economía, Estadísticas e Informática. EEA INTA Marcos Juárez.
fmoore@mjuarez.inta.gov.ar

Abstract. National statistics of a country's agricultural sector can be provided by an adequate, periodic survey based on probability sampling methods. Satellite images could give a significant source of information to achieve this goal. The methodology of Area Sampling to generate annual estimates of agricultural commodities produces a quantity of field data that can be used for image classification.

This survey confirms that the classification could be highly accurate, because of the quantity of field information available, the uniformity of a restricted area and the knowing of it.

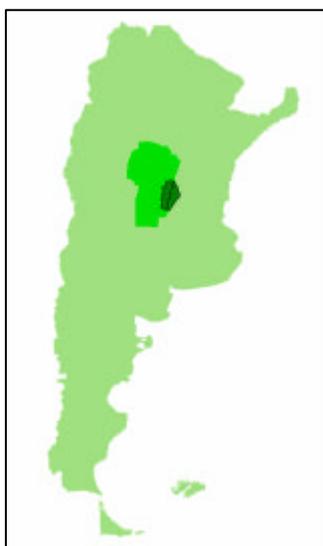
The image classification produces a map of soil coverage, that improve the knowing of crops distribution in each region, and how that could change from year to year, besides the validation of data obtained from the sampling method.

Keywords : remote sensing, image processing, agricultural statistics, crop estimates.

1. Introducción

El presente trabajo utiliza imágenes LandSat para identificar el uso del suelo para la campaña 2008/9 en los departamentos Marcos Juárez y Unión. El objetivo principal fue obtener el área de los cultivos más importantes en la región (trigo, soja y maíz) mediante una clasificación supervisada de las imágenes utilizando la información de campo obtenida en las recorridas realizadas por una muestra de segmentos distribuidas en el área de dichos departamentos. La determinación del uso del suelo de estos segmentos forma parte de un trabajo que procura estimar el área total destinada a los cultivos mencionados mediante la observación de una muestra representativa, trabajo que fue iniciado en la campaña 2007/08 y que se continúa en la presente 2008/09.

2. Materiales y métodos



El área de estudio fue la correspondiente a los departamentos Marcos Juárez y Unión de la provincia de Córdoba, como se observa en la imagen (aproximadamente dos millones de hectáreas).

En el trabajo se utilizaron cinco imágenes LandSat (path/row 228/82 y 228/83) de cinco fechas diferentes (4/10/08, 27/11/08, 24/01/09, 09/02/09 y 25/02/09). Las imágenes fueron georeferenciadas y se utilizaron seis bandas de cada fecha descartándose el infrarrojo termal, al superponerlas se generó una sola imagen con 30 bandas.

Utilizando el software ENVI (2002) se realizó una clasificación supervisada por el método de máxima probabilidad que calcula la probabilidad de que cada píxel con un valor radiométrico determinado pertenezca a las distintas categorías, asignando el mismo a aquella en la cual su probabilidad sea mayor (Chuvieco, 1996).

En la Figura 1 se puede observar una secuencia de dos imágenes, de las fechas 04/10/08 y 24/01/09, en las cuales el color rojo representa al verde de un cultivo en crecimiento, de manera que se percibe a simple vista la dominancia de cultivos de verano (fecha 24/01), fundamentalmente soja de primera.

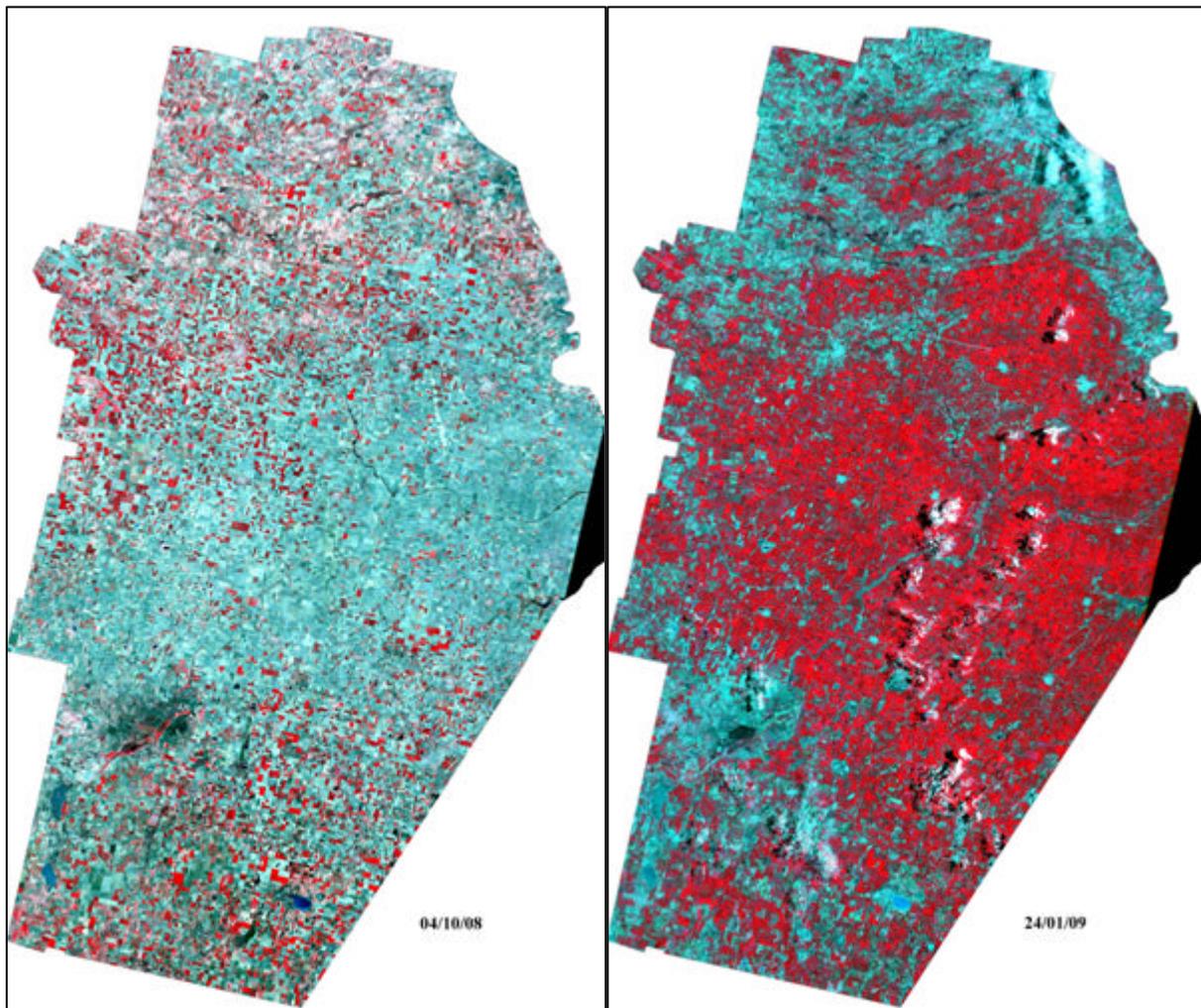


Figura 1. Secuencia imágenes LandSat (RGB 4,3,2 Falso Color Compuesto)

De la información de campo disponible se utilizó sólo una parte como área de entrenamiento, reservándose el resto para una posterior comprobación de la clasificación.

El cuadro 1 muestra la superficie de las áreas de entrenamiento seleccionadas y la superficie total de información disponible. Se realizaron dos comprobaciones de la imagen clasificada, la primera con los sitios de entrenamiento y la segunda con el total de la información de campo con que se contaba. Por ser relevantes para el análisis se incluyen las clases alfalfa y pastura natural (PN).

Cuadro 1. Superficie con información de campo

Uso suelo	Superficie (ha)	
	Sitios entren.	Total
Trigo	1009	4924
Maíz	1718	6094
Soja	2419	17771
Alfalfa	351	1933
P.Natural	356	1417
Total	5853	32139

3. Resultados y discusión

En el cuadro 2 se observa la Matriz de Confusión generada al contrastar la imagen resultante de la clasificación con la información de campo utilizada (sitios de entrenamiento, cuadro 1). Como se puede notar, el valor de precisión general es muy bueno (96,6%), y los valores de precisión individual para cada clase también, siendo el de la clase soja el menor con 95,5%. Debe notarse que se definieron dos clases para el maíz (maíz temprano sembrado en setiembre, maíz tardío sembrado en noviembre o diciembre). El valor del Índice Kappa es “muy bueno” (0,95). Este índice mide el grado de concordancia entre el resultado de la clasificación y los sitios utilizados como entrenamiento, cuando el acuerdo es mayor el índice se aproxima a uno y cuando es menor toma valores cercanos a cero (Cohen, 1960).

Los valores de separabilidad entre los sitios de entrenamiento seleccionados fueron adecuados ($>1,9$) en todos los casos (test Jeffries-Matusita y Divergencia Transformada). Este test prueba la separabilidad espectral de las áreas de entrenamiento utilizadas para cada clase.

Los registros obtenidos señalan una muy buena precisión en la clasificación cuando se contrastan con las zonas utilizadas como sitios de entrenamiento. Una mejor verificación se obtiene cuando se evalúan áreas que no fueron incluidas como sitios de entrenamiento, esto es posible cuando se cuenta con la información de campo suficiente que fue reservada para este propósito (columna 2 del cuadro 1). El resultado de la evaluación fue una buena medida de la precisión de la clasificación ya que la superficie utilizada para realizar esta comprobación fue amplia y además bien distribuída en toda el área de los departamentos.

Cuadro 2. Matriz de Confusión con sitios de entrenamiento.

Matriz de Confusión										
Precisión general = (81131/83982) 96.6052%										
Coeficiente kappa = 0.9589										
Clase	Verdad de campo (píxeles)									
	Trigo	Maiz	Soja	Alfa	PN	Monte	Agua	Urbano	Maiz-2	Total
No clasif.	0	39	38	0	0	0	14	0	0	91
Trigo	10887	0	168	1	2	3	0	0	23	11084
Maiz	0	14149	95	2	15	0	0	0	3	14264
Soja	2	73	25688	5	11	2	0	0	41	25822
Alfa	43	29	41	3781	49	1	0	0	7	3951
PN	17	155	333	95	3839	4	0	0	7	4450
Monte	0	0	0	0	0	2787	0	0	0	2787
Agua	0	0	0	0	0	0	7251	0	0	7251
Urbano	105	6	274	17	11	9	0	8870	5	9297
Maiz-2	160	678	239	0	24	5	0	0	3879	4985
Total	11214	15129	26876	3901	3951	2811	7265	8870	3965	83982
Clase	Verdad de campo (porcentajes)									
	Trigo	Maiz	Soja	Alfa	PN	Monte	Agua	Urbano	Maiz-2	Total
No clasif.	0.00	0.26	0.14	0.00	0.00	0.00	0.19	0.00	0.00	0.11
Trigo	97.08	0.00	0.63	0.03	0.05	0.11	0.00	0.00	0.58	13.20
Maiz	0.00	93.52	0.35	0.05	0.38	0.00	0.00	0.00	0.08	16.98
Soja	0.02	0.48	95.58	0.13	0.28	0.07	0.00	0.00	1.03	30.75
Alfa	0.38	0.19	0.15	96.92	1.24	0.04	0.00	0.00	0.18	4.70
PN	0.15	1.02	1.24	2.44	97.17	0.14	0.00	0.00	0.18	5.30
Monte	0.00	0.00	0.00	0.00	0.00	99.15	0.00	0.00	0.00	3.32
Agua	0.00	0.00	0.00	0.00	0.00	0.00	99.81	0.00	0.00	8.63
Urbano	0.94	0.04	1.02	0.44	0.28	0.32	0.00	100.00	0.13	11.07
Maiz-2	1.43	4.48	0.89	0.00	0.61	0.18	0.00	0.00	97.83	5.94
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Clase	Comision		Omission		Comision		Omission			
	(%)	(%)	(%)	(%)	(Píxeles)	(Píxeles)	(Píxeles)	(Píxeles)		
Trigo	1.78	2.92	197/11084	327/11214						
Maiz	0.81	6.48	115/14264	980/15129						
Soja	0.52	4.42	134/25822	1188/26876						
Alfa	4.30	3.08	170/3951	120/3901						
PN	13.73	2.83	611/4450	112/3951						
Monte	0.00	0.85	0/2787	24/2811						
Agua	0.00	0.19	0/7251	14/7265						
Urbano	4.59	0.00	427/9297	0/8870						
Maiz-2	22.19	2.17	1106/4985	86/3965						
Clase	Prod. Ac.		User Ac.		Prod. Ac.		User Ac.			
	(%)	(%)	(%)	(%)	(Píxeles)	(Píxeles)	(Píxeles)	(Píxeles)		
Trigo	97.08	98.22	10887/11214	10887/11084						
Maiz	93.52	99.19	14149/15129	14149/14264						
Soja	95.58	99.48	25688/26876	25688/25822						
Alfa	96.92	95.70	3781/3901	3781/3951						
PN	97.17	86.27	3839/3951	3839/4450						
Monte	99.15	100.00	2787/2811	2787/2787						
Agua	99.81	100.00	7251/7265	7251/7251						
Urbano	100.00	95.41	8870/8870	8870/9297						
Maiz-2	97.83	77.81	3879/3965	3879/4985						

El cuadro 3 muestra la Matriz de Confusión generada a partir de la comparación del total de la información de campo disponible y la imagen clasificada. Los resultados de la precisión general (83%) y el Índice Kappa (0,74) son satisfactorios considerando la importante cantidad de píxeles con información de campo incluidos en el análisis (342.961).

El análisis del cuadro nos permite detectar que la alfalfa es la clase con menor precisión en la clasificación, siendo ésta el 68,4%, por lo tanto la confianza en los resultados de la clasificación para dicho cultivo es menor. Esto se debe a los numerosos estados y condiciones que muestra el mismo a lo largo de su período de ocupación del lote, que de por sí es muy variable, y también lo es desde el punto de vista de su aprovechamiento, ya que puede ser consumida directamente por el ganado o mediante corte y enfardado, además si se encuentra como único cultivo o consociada.

Esta cantidad de estados que puede presentar este cultivo determinan una variación en el terreno muy difícil de captar mediante los sitios de entrenamiento.

Cuadro 3. Matriz de Confusión con la información de campo completa

Matriz de confusión						
Precisión general = (284845/342961) 83.0546%						
Coeficiente Kappa = 0.7406						
Verdad de campo (píxeles)						
Clase	maiz	soja	trigo	alfalfa	PNatural	Total
No clasif.	2551	2500	1529	1292	84	7956
Maiz	52854	6099	1044	343	313	60653
Soja	4857	163929	3671	645	676	173778
Trigo	909	6372	43779	371	518	51949
Alfa	1597	4373	1367	13638	905	21880
PN	2802	8150	1510	3638	10645	26745
Total	65570	191423	52900	19927	13141	342961
Verdad de campo (porcentajes)						
Clase	maiz	soja	trigo	alfalfa	PNatural	Total
No clasif.	3.89	1.31	2.89	6.48	0.64	2.32
Maiz	80.61	3.19	1.97	1.72	2.38	17.69
Soja	7.41	85.64	6.94	3.24	5.14	50.67
Trigo	1.39	3.33	82.76	1.86	3.94	15.15
Alfa	2.44	2.28	2.58	68.44	6.89	6.38
PN	4.27	4.26	2.85	18.26	81.01	7.80
Total	100.00	100.00	100.00	100.00	100.00	100.00
Clase	Comision (%)	Omision (%)	Comision (Píxeles)	Omision (Píxeles)		
Maiz	12.86	19.39	7799/60653	12716/65570		
Soja	5.67	14.36	9849/173778	27494/191423		
Trigo	15.73	17.24	8170/51949	9121/52900		
Alfa	37.67	31.56	8242/21880	6289/19927		
PN	60.20	18.99	16100/26745	2496/13141		
Clase	Prod. Acc. (%)	User Acc. (%)	Prod. Acc. (Píxeles)	User Acc. (Píxeles)		
Maiz	80.61	87.14	52854/65570	52854/60653		
Soja	85.64	94.33	163929/191423	163929/173778		
Trigo	82.76	84.27	43779/52900	43779/51949		
Alfa	68.44	62.33	13638/19927	13638/21880		
PN	81.01	39.80	10645/13141	10645/26745		

Las otras clases muestran valores mas adecuados (soja 85,6% y trigo 82,7%), aunque el maíz tiene un valor más bajo (80,6%) y una relativamente importante cantidad de píxeles clasificados como soja (7,4%), lo que indicaría una cantidad de este cultivo incluido erróneamente en la clase soja. Todas las clases presentan píxeles sin clasificar, siendo la alfalfa la de mayor valor con 6,4%. También todas tienen algún grado de confusión con la clase pastura natural (PN), pero es la alfalfa la que tiene el mayor porcentaje de píxeles (18,2%) que fueron incluidos en ésta clase erróneamente, mientras que un 6,8% de píxeles que eran PN fueron clasificados como alfalfa.

Al analizarse los errores de omisión (píxeles que no fueron clasificados en la clase a la que pertenecen) y de comisión (píxeles que se incluyen en la clase sin pertenecer a ella), se observan los porcentajes más altos en las clases alfalfa y PN. En maíz y soja se produjeron un mayor porcentaje de errores de omisión que de comisión, lo que indicaría que los valores de superficies obtenidos para dichos cultivos están por debajo del valor real. En trigo en cambio éstos porcentajes son similares (omisión 17,5 y comisión 15,7), de manera que los errores de exclusión e inclusión se equilibran en el resultado final. En la clase alfalfa existe una diferencia de 6% a favor de superficie incluida erróneamente, pero es en la PN en donde se registran valores muy altos, un 41% de diferencia a favor de los errores de comisión, es decir que fue incluida en la clasificación una importante superficie que no correspondía a dicha clase.

Debe advertirse que la presente campaña fue muy particular debido a la extensa sequía en los meses previos a la siembra de los cultivos de verano, lo que determinó que ésta se hiciera en forma escalonada a medida que la humedad lo permitía en cada región. Este hecho dificultó la identificación del maíz, que normalmente tiene una ventana de siembra muy acotada durante el fin del invierno e inicio de la primavera. También el trigo presentó problemas debido a que numerosos lotes estuvieron en muy malas condiciones, e incluso algunos fueron pastoreados o desecados al tener una perspectiva de rendimientos que no justificaba su cosecha.

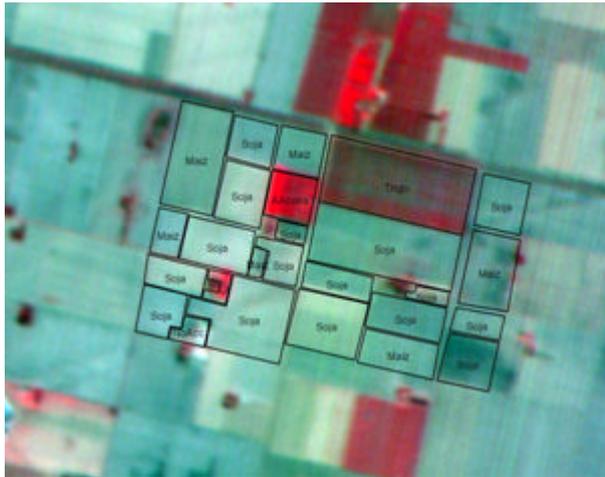


Figura 2.

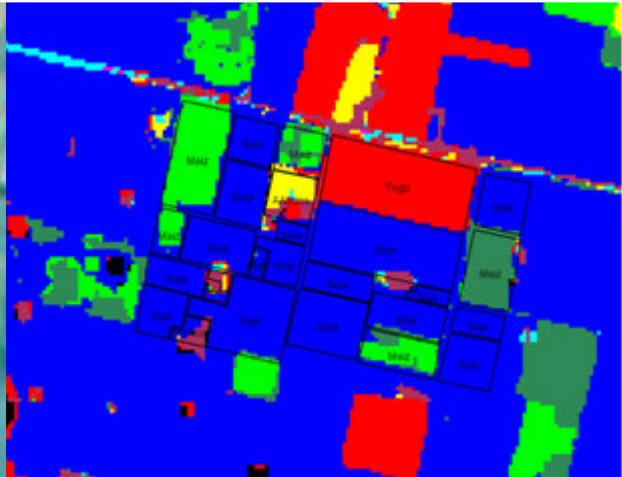


Figura 3.

La Figura 2 presenta un segmento superpuesto a una imagen Landsat (fecha 04/10/08), falso color compuesta (RGB 4,3,2), donde se señalan los cultivos que ocupan cada lote, información obtenida por medio de una recorrida realizada por los segmentos durante los meses del ciclo de los cultivos. Nótese que los lotes con cultivos de color verde se ven en la imagen de color rojo, para la fecha considerada pueden ser trigo o alfalfa, mientras que el resto, de color azulado-verde o celeste, son lotes en barbecho o recién sembrados.

La Figura 3 presenta el mismo segmento pero de la imagen clasificada. En ésta el color azul representa soja de primera, el rojo al trigo y soja de segunda, verde claro al maíz sembrado en fecha temprana, el verde oscuro es maíz tardío, el amarillo es alfalfa y el cian pastura natural.

El Cuadro 4 presenta los valores de superficie de cada cultivo obtenidos mediante la clasificación. La sumatoria del área total clasificada no se corresponde con la superficie de ambos departamentos ya que no se presentan en estos resultados las clases no relevantes como: monte, áreas urbanas, zonas bajas con agua y también una importante cantidad de píxeles que no lograron ser incluidos en ninguna de las clases propuestas.

En este cuadro los totales para las clases alfalfa y PN están sobrestimados. De acuerdo al análisis realizado anteriormente una parte importante de esta superficie sería principalmente soja de primera.

Cuadro 4. Resultados de la clasificación.

	Superficie (has)		
	Unión	M. Juárez	Total
Trigo	163,100	118,100	281,200
Soja	371,700	508,850	880,550
Maíz	87,500	172,600	260,100
Alfalfa	88,200	52,000	140,200
P.Natural	166,800	81,900	248,700
Total	877,300	933,450	1,810,750

La Figura 4 muestra la imagen obtenida mediante el proceso de clasificación supervisada, en ésta los colores representan las clases descritas en la Figura 3.

Un sector del departamento Marcos Juárez que en la imagen se observa de color negro, no fue incluido en el procesamiento por pertenecer a otra escena, el área tiene aproximadamente 40.000 hectáreas, y fue tenida en cuenta en el cálculo de las superficies de cada clase (Cuadro 4).

El análisis de la precisión de la clasificación muestra valores de los indicadores correctos para las clases trigo, maíz y soja, no así en las clases alfalfa y pastura natural. Para evitar este problema se hace necesario la identificación de las zonas con mayor concentración de dichas clases, principalmente áreas de menor aptitud agrícola, y enmascararlas con el objetivo de evitar su inclusión en la clasificación general.

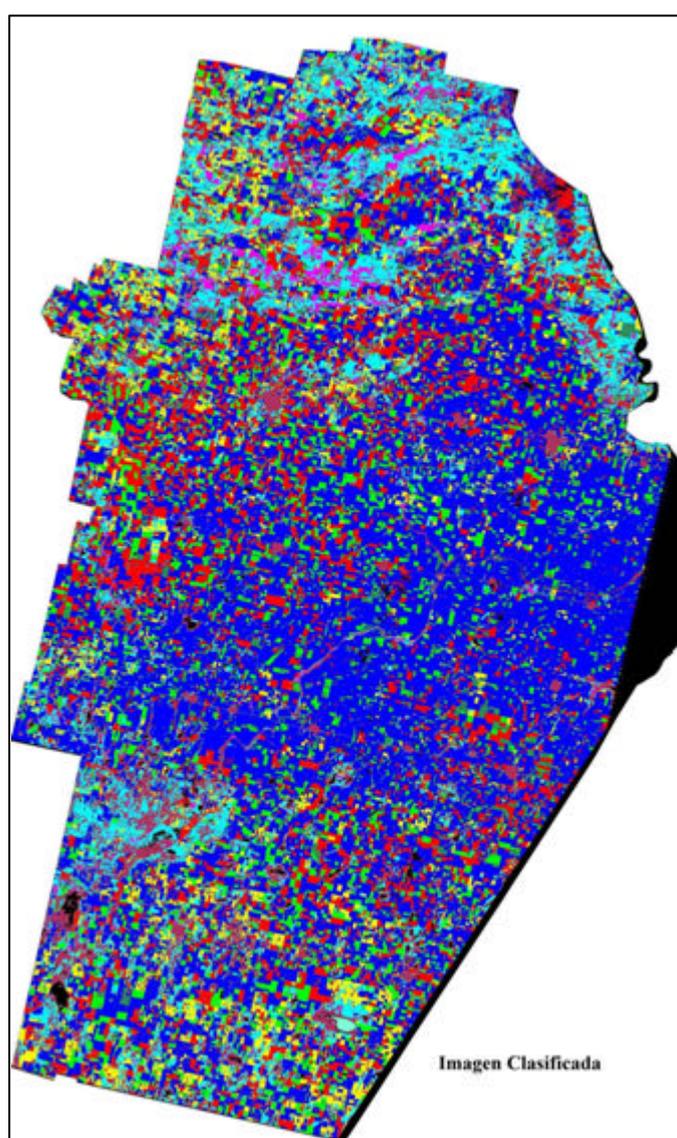


Figura 4 – Imagen Clasificada.

4. Conclusión

La metodología de clasificación de imágenes es un instrumento útil para la determinación del uso del suelo, especialmente cuando se cuenta con suficiente información de campo. En el presente trabajo la validación de la clasificación realizada mediante las matrices de confusión muestran una excelente precisión en ambos casos, cuando se comprueba mediante los sitios de entrenamiento y también cuando se usa el resto de la información de campo.

La estimación probabilística mediante muestreo de áreas genera una cantidad de información de campo que posibilita una clasificación supervisada con un buen grado de precisión cuando el área a clasificar es acotada, uniforme y conocida, y permite además la generación de un mapa del uso del suelo, que es un producto muy útil ya que identifica la distribución de los cultivos en las distintas zonas.

5. Bibliografía:

ENVI, v.3.6 (2002). Software de procesamiento y análisis de imágenes geoespaciales. (<http://www.itvis.com/envi>).

Chuvieco, E. (1996): Fundamentos de Teledetección Espacial - Madrid, Rialp, 568 Págs. 1996.

Cohen J. (1960). A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20:37-46