

Clasificación del uso de suelo en agricultura a partir de series temporales de imágenes LANDSAT

Miguel Nolasco¹, Enrique Willington¹ y Mónica Bocco¹

¹ Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba. Argentina.
mnolasco@agro.unc.edu.ar

Resumen. La utilización de los recursos naturales tiene como protagonista preponderante al sector agropecuario, en lo que respecta al uso del recurso suelo; siendo su conocimiento de gran importancia en la actualidad. La teledetección posibilita, a través de la clasificación de imágenes satelitales, conocer el uso de los suelos y monitorear sus cambios; en estos estudios las series multi-temporales representan una alternativa para mejorar las clasificaciones. El presente estudio fue realizado en la región central de la provincia de Córdoba, donde la soja y el maíz son los cultivos más importantes, por área sembrada y réditos económicos producidos. Los objetivos fue evaluar y comparar el desempeño de dos alternativas de clasificación, empleando el algoritmo de máxima verosimilitud, una que utiliza una serie temporal de imágenes satelitales y otra que clasifica imágenes individuales, para identificar la cobertura del suelo en el área de estudio. La clasificación a partir de datos multitemporales presenta valores superiores en los estadísticos de validación.

Palabras clave: clasificación; series multitemporales; cultivos agrícolas; LANDSAT TM.

1 Introducción

El uso de la tierra para la producción agropecuaria está relacionado con la utilización de los recursos naturales. En la República Argentina, el sector agropecuario está considerado uno de los principales recursos en su economía, tanto como generador de divisas por las exportaciones como abastecedor del mercado interno [1]. A nivel global, los datos del tipo de cobertura del suelo están entre los más importantes y utilizados para la gestión de los recursos naturales y en estudios sobre cambio climático, por esto es cada vez más necesario poseer datos exactos, detallados y periódicos [2,3]. Los mapas de cobertura no están fácilmente disponibles y además como ésta cambia muy rápidamente, la capacidad para monitorearla con precisión es importante [4]. A partir del estudio y conocimiento del uso de la tierra se pueden evaluar el cálculo de la producción agrícola, así como los requerimientos de insumos tales como fertilizantes, pesticidas, maquinarias y mano de obra [1].

El mapeo de la cobertura del suelo es una de las aplicaciones más importantes en teledetección, y se han realizado importantes investigaciones para mejorar los métodos de clasificación de las imágenes satelitales [5]. Entre las ventajas de la

teledetección respecto de otros métodos que obtienen datos a campo se encuentran un menor costo, la cobertura de grandes áreas y la posibilidad de repetir la obtención de los datos a lo largo del tiempo [6].

Para la clasificación de imágenes satelitales pueden utilizarse numerosos métodos, que se pueden dividir en paramétricos y no-paramétricos. Dentro de los primeros se encuentra el de Máxima verosimilitud (método basado en probabilidad) uno de los más utilizados para clasificar uso o cobertura del suelo, que en el área de estudio produce resultados con índices muy buenos de precisión [7].

Al clasificar el uso de suelo agrícola en zonas heterogéneas, debidas a los distintos cultivos, fechas de siembra, estado fenológico, tecnología utilizada, etc. se presentan complicaciones para definir las clases de uso. En este sentido, se ha encontrado que la máxima discriminación entre los cultivos ocurre en diferentes etapas del crecimiento y por lo tanto no es posible la captura de esas diferencias usando una sola fecha; en consecuencia, la aplicación de imágenes multitemporales, que proporcionan información adicional sobre la fenología de la vegetación, ha sido recomendada para mejorar la discriminación [1,8]. En Argentina, utilizando datos de Landsat TM multitemporales, de una misma estación de crecimiento, en [9] se muestra una clasificación de tipos de cobertura de suelo. Para el partido de Guaminí (Buenos Aires, Argentina) en [1] determinaron el uso de la tierra, realizando un estudio multitemporal, a partir de imágenes satelitales Landsat 5 TM y a partir de una imagen de satélite Landsat 7ETM+ en [10] clasificaron suelos de uso agrícola para determinar el tipo de labranza. Estos resultados se combinaron con los de imágenes Spot 4 de toma anterior, con el objetivo de mejorar la precisión de la clasificación mediante un análisis multi-temporal. Para clasificar cultivos agrícolas en las planicies centrales de EEUU investigaron la aplicabilidad de series temporales de datos de MODIS 250 m [3].

Los cultivos de soja (*Glycine max* (L.) Merrill) y maíz (*Zea mays* L.) son los más importantes en Argentina si se considera la superficie sembrada y los réditos económicos que producen; esta situación se repite en la provincia de Córdoba, segunda productora a nivel nacional en ambos cultivos [11].

Tabla 1. Superficies sembradas con Soja y Maíz (campaña 2011/2012)

Cultivo	Argentina	Córdoba
Soja	18.670.000 ha	5.014.000 ha
Maíz	5.000.000 ha	1.272.000 ha

Los objetivos de este trabajo fue evaluar y comparar el desempeño de dos alternativas de clasificación de cobertura del suelo de uso agrícola, una que clasifica las imágenes de fechas individuales y otra que utiliza una serie temporal que cubre la totalidad del ciclo de los cultivos, a partir de imágenes satelitales LANDSAT.

2 Materiales y métodos

2.1 Área de estudio

El estudio fue realizado en la región central de la provincia de Córdoba, Argentina (departamentos Río Segundo y Santa María, cuya ubicación aproximada es 32° S; 64° W). El área presenta un relieve ligeramente ondulado con una leve pendiente hacia el este, suelos de textura franco limosa clasificados como Haplustoles énticos y típicos.

El clima es clasificado como seco sub-húmedo, con precipitaciones anuales promedio de 800 mm, concentradas en el verano. El método de labranza predominante en la zona es el de siembra directa, siendo la principal actividad productiva la agricultura, predominando los cultivos de verano, soja y maíz. En menor medida se siembra trigo (*Triticum aestivum* L.) como cultivo de invierno. [12]

2.2 Datos a campo

El trabajo de campo previo consistió en un relevamiento de 27 lotes agrícolas del área, con utilización de GPS con el fin de obtener una base de datos correspondiente a los diferentes usos del suelo. Durante la campaña agrícola (noviembre 2010-marzo 2011) se relevaron lotes cultivados con soja y maíz, registrándose además época de siembra y el cultivo antecesor a fin de determinar las clases de cobertura para las clasificaciones.

2.3 Imágenes satelitales y clases de cobertura

Se utilizaron tres imágenes satelitales LANDSAT (provistas por CONAE), una tomada el 20/12/2010, correspondiente al inicio del ciclo de cultivo, otra con fecha del 21/01/2011 correspondiente a un estado intermedio de desarrollo y la tercera registrada en la fecha 02/03/2011, cuando los cultivos se encontraban en etapas avanzadas de su ciclo.

Las imágenes de diciembre y enero se referenciaron tomando como base la de marzo, que ya se encontraba referenciada en latitud y longitud. En forma simultánea se hizo un recorte del área de estudio en cada imagen y se creó una serie multitemporal a partir de las mismas. Sobre estos cuatro productos se realizaron las clasificaciones, utilizando las bandas 1, 2, 3, 4, 5 y 7 provistas por el satélite.

De acuerdo a los datos relevados durante el trabajo de campo se establecieron las siguientes clases de cobertura:

Cob 1: Maíz temprano (antecesor soja), emergido hasta 12/10/2010.

Cob 2: Maíz tardío (antecesor soja), emergido hasta el 12/01/2011.

Cob 3: Soja temprana (antecesor soja), emergida hasta 25/10/2010.

Cob 4: Soja intermedia (antecesor soja), emergida entre 26/10/2010 hasta 12/12/2010.

Cob 5: Soja tardía (antecesor trigo), emergida a partir del 13/12/2010.

Cob 6: Otras coberturas (camino, pasturas, ribera de río, etc).

En la imagen de diciembre la **Cob 2** (maíz tardío, aún no sembrado) poseía rastrojo de soja, con presencia de malezas en algunos casos y en los lotes correspondientes a **Cob 5** el cultivo de trigo se encontraba recién cosechado.

2.4 Clasificación supervisada

El algoritmo de clasificación supervisada utilizado fue el de máxima verosimilitud, y se realizaron cuatro clasificaciones: una correspondiente a cada una de las imágenes en forma individual y la cuarta que consideró la composición multitemporal de las tres imágenes anteriores.

Sobre los lote relevados se digitalizaron polígonos, con el fin de tomar muestras de su firma espectral y utilizarlos como regiones de interés. En las clasificaciones se utilizaron polígonos (como verdad de terreno) y otros distintos en la validación de la clasificación.

2.5 Análisis estadístico

La metodología estadística más utilizada para validar este tipo de modelos es el análisis de las matrices de confusión, con la que se evalúa la precisión de las clasificaciones. A partir de los valores de las matrices de confusión (datos observados y estimados para cada característica o propiedad que se clasifique) se calcularon distintos parámetros [7], cuyas fórmulas son:

Precisión global (Pg):

$$P = \frac{\sum_{i=1}^m x_{ii}}{N} \quad (1)$$

Precisión del productor (Pp):

$$P_p = \frac{x_{ii}}{x_{\Sigma i}} \quad (2)$$

Precisión del usuario (Pu):

$$P_u = \frac{x_{ii}}{x_{i\Sigma}} \quad (3)$$

Coefficiente Kappa (κ):

$$\kappa = \frac{N \sum_{i=1}^m x_{ii} - \sum_{i=1}^m x_{i\Sigma} x_{\Sigma i}}{N^2 - \sum_{i=1}^m x_{i\Sigma} x_{\Sigma i}} \quad (4)$$

Donde m = número total de clases, N = número total de píxeles en las m clases de referencia, x_{ii} = elementos de la diagonal principal de la matriz de confusión,

$x_{\Sigma i}$ = suma de los píxeles de la clase i de referencia y $x_{i\Sigma}$ = suma de los píxeles clasificados como la clase i .

El coeficiente kappa (κ) computa el acuerdo entre la imagen clasificada y la verdad de terreno, debido únicamente a la exactitud de la clasificación, suprimiendo el acuerdo que cabría esperar simplemente por azar. En [13] se propuso una escala para la interpretación del valor de dicho coeficiente donde valores menores a 0,40 se clasifican como pobres, 0,40-0,55 suficientes, 0,55-0,70 buenos, 0,70-0,85 muy buenos y mayores a 0,85 excelentes. Los demás parámetros tienen una escala de 0 a 100%.

3 Resultados y discusión

En la Tabla 2 se muestran los valores de los índices κ y las precisiones globales (en porcentaje) para cada imagen clasificada, al realizar la validación del algoritmo clasificador.

Tabla 2. Valores de los estadísticos para las clasificaciones realizadas.

Fecha Imagen	Índice Kappa (κ)	Precisión Global (%)
20/12/2010	0,89	91,30
21/01/2011	0,70	75,76
02/03/2011	0,81	85,13
Dic_Ene_Mar	0,99	98,92

En los resultados se observa que la imagen compuesta por las tres escenas (diciembre, enero y marzo) presenta los mayores valores de validación en los estadísticos κ y PG. En esta serie multitemporal la mayor cantidad de datos, que incluyen a los cultivos en sus distintas etapas de desarrollo (diferentes estados fenológicos, datos provistos por las imágenes de enero y marzo) y datos del rastrojo anterior (provistos por la imagen de diciembre), permitieron decididamente una mejor clasificación. En el trabajo de [14] se clasificaron uso de suelo en dos clases: cultivos de verano (soja y maíz) y no cultivos de verano, a partir de imágenes multitemporales de los satélites MODIS, para el Estado de Paraná (Brasil) en 4 campañas agrícolas y obtuvieron valores de κ similares a los obtenidos en el presente trabajo con los mismos cultivos, entre 0,86 y 0,90 y de PG entre 93 y 95%. Los resultados permiten afirmar que el análisis multitemporal posibilita mejorar la clasificación, logrando una excelente precisión espacial, eliminando el inconveniente de la variabilidad que introducen las fechas de siembra.

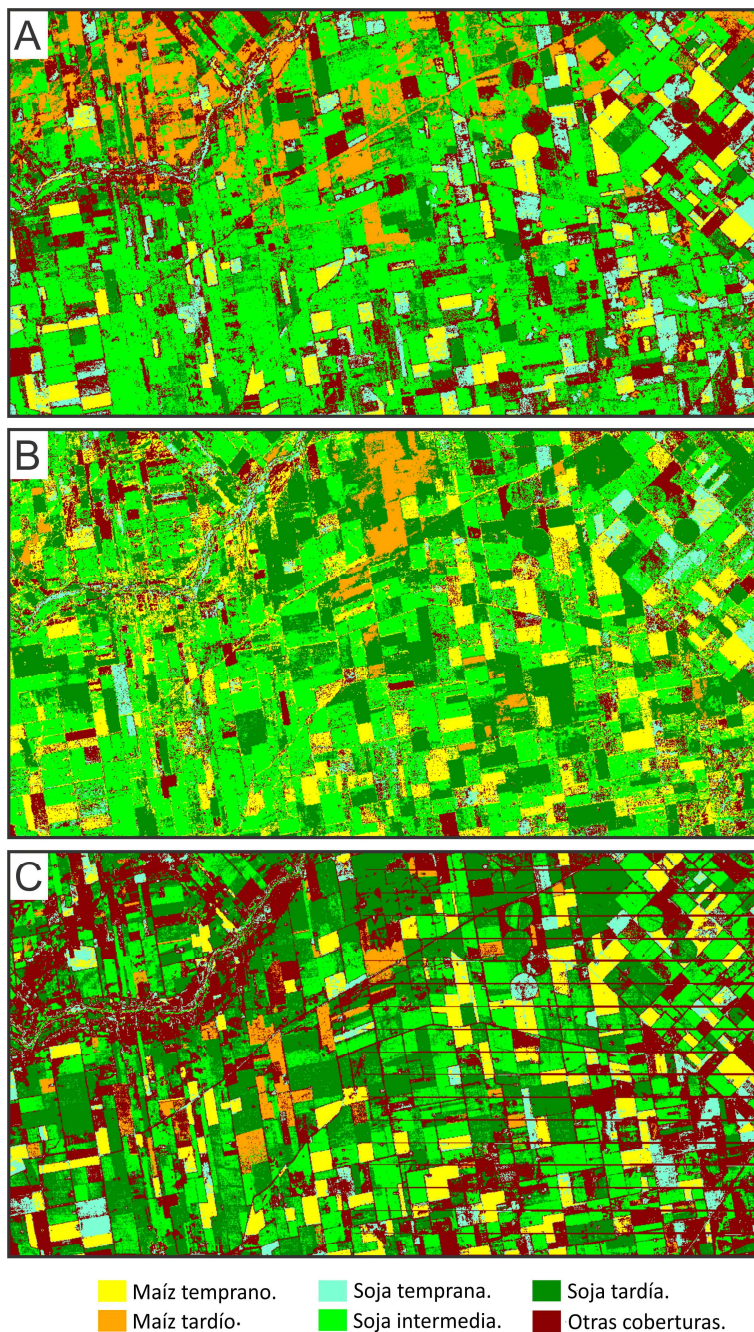


Figura 1. Imágenes clasificadas individualmente correspondientes a los meses de diciembre (A), enero (B) y marzo (C).

Comparando los estadísticos de las imágenes de fechas individuales (Figura 1) se puede afirmar que en los inicios de la campaña (diciembre) el algoritmo clasificador se desempeña mejor que en etapas más avanzadas (enero-marzo); esto podría deberse a que en diciembre los lotes correspondientes a soja tardía se encontraban con el cultivo anterior (trigo) recién cosechado, lo que permite una clasificación de la clase **Cob 5** muy precisa. Los mayores valores de κ y precisión global encontrados en la imagen que corresponde a diciembre, indicarían que la variabilidad entre clases que agrega la superficie aún ocupada por rastrojo para esta fecha, permite una mejor discriminación de los cultivos por fecha de siembra; sin embargo, esta separación que es fácil de determinar en etapas tempranas, ya no es posible en etapas avanzadas de los ciclos (marzo).

Los menores valores de los estadísticos se presentan para momentos en los que la superficie del suelo se encuentra ocupada por una interfaz rastrojo-cultivo (imagen de enero).

En la Figura 2 se muestra el resultado de la clasificación de la imagen multitemporal, para las clases de cobertura definidas. Esta clasificación fue la que mejores estadísticos presentó cuando se validó con los datos de campo del área de estudio.

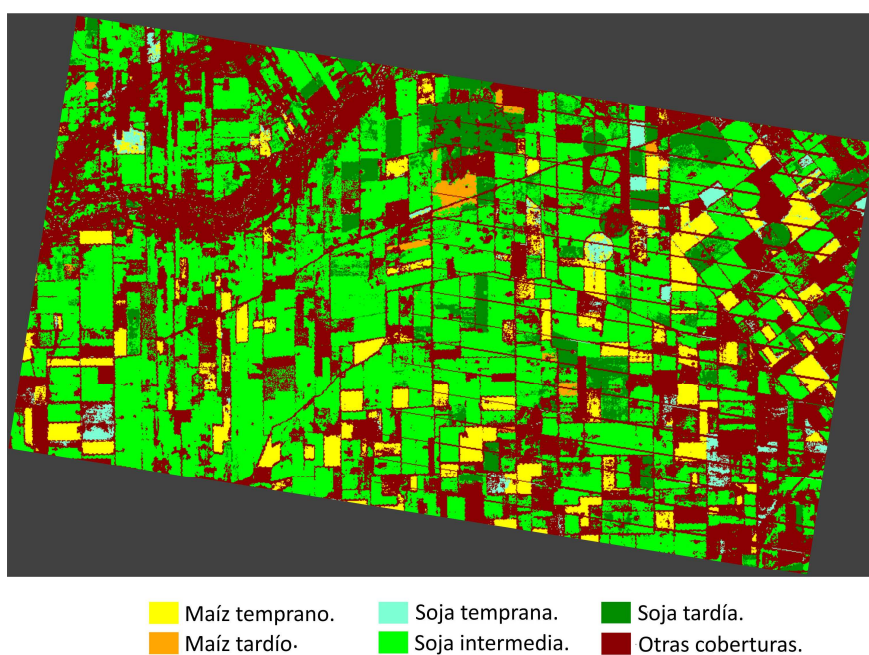


Figura 2. Imagen clasificada utilizando la serie multitemporal.

A partir de imágenes SPOT y LANDSAT TM, en [10] cuando clasificaron tipos de labranza a escala regional, también comprobaron que realizando un análisis multitemporal se obtenía mejor precisión en la clasificación, obteniendo valores

promedio de κ de 0,61 (SPOT) y 0,50 (LANDSAT) para imágenes individuales y 0,65 para datos multitemporales.

El valor de $\kappa = 0,98$ encontrado en el presente trabajo para la clasificación realizada a partir de una imagen multitemporal es categorizado como excelente según la escala propuesta por [13]. Realizando una clasificación no supervisada (método Isodata) sobre una serie multitemporal, de imágenes LANDSAT 5 TM del partido de Guaminí (Buenos Aires), en [1] obtuvieron un valor de $\kappa = 0,85$, separando en seis categorías (las cuales incluían soja, maíz y pasturas, entre otras).

La imagen multitemporal fue la única que aseguró valores mayores a 85% de precisión de clasificación, tanto productor como usuario en todas las clases de uso/cobertura. En la Tabla 3 se muestran los valores de P_p y P_u de las cuatro clasificaciones, para las clases de uso/cobertura del suelo.

Tabla 3. Valores (en %) de los estadísticos de P_p y P_u para distintas imágenes y clases de cobertura de suelo.

Tipo de Cobertura	Imagen 20/12/2010		Imagen 21/01/2011		Imagen 02/03/2011		Imagen Multitemporal Dic_Ene_Feb	
	P_p	P_u	P_p	P_u	P_p	P_u	P_p	P_u
	Cob 1	88,8	89,1	71,8	83,3	98,9	99,4	100,0
Cob 2	67,4	57,4	91,5	100,0	100,0	97,9	100,0	97,7
Cob 3	78,9	80,9	82,8	65,7	89,8	94,2	97,8	99,9
Cob 4	93,7	98,6	69,7	82,6	64,7	89,8	98,9	100,0
Cob 5	99,3	98,1	97,7	78,0	88,4	63,4	82,7	100,0
Cob 6	95,0	86,6	56,2	60,4	95,7	86,5	100,0	100,0

Utilizando datos multitemporales de LANDSAT TM y algoritmo clasificador de Máxima verosimilitud, en [9] al clasificar cobertura del suelo en 14 clases que incluían maíz y soja, obtuvieron valores de $\kappa = 0,55$ y $PG = 62,6\%$, con $P_p = 65\%$ y $P_u = 23\%$ para maíz y $P_p = 74\%$ y $P_u = 83\%$ para soja; y al disminuir la cantidad de clases de la clasificación, sus resultados mejoraron a valores de $\kappa = 0,84$ y $PG = 88,7\%$.

4 Conclusiones

El análisis multitemporal permite una precisión de la clasificación más eficiente de imágenes LANDSAT respecto de una clasificación realizada con datos una sola escena, logrando valores catalogados como excelentes en todas las clases de cobertura cuando se quiere identificar soja y maíz.

Trabajando con imágenes de fechas individuales, el algoritmo clasificador de máxima verosimilitud pierde capacidad de discriminación entre cultivos tempranos y tardíos, cuando estos últimos se encuentran desarrollados. Siendo la época del inicio de los ciclos el mejor momento para poder diferenciarlos.

Agradecimientos. Los autores expresan su gratitud a la Secretaría de Ciencia y Técnica – Universidad Nacional de Córdoba - Argentina (Secyt-UNC) por el apoyo financiero que permitió la realización de este trabajo.

Referencias

1. Marini, M. F., Vergara, M. F., Krüger, H.: Determinación del uso de la tierra en el partido de Guamini (Argentina) mediante un estudio multitemporal con imágenes Landsat. *Revista de Teledetección*. 27, 80--88 (2007)
2. Corts, R., Navarro, M., Pardo, J.A., Bisbal, M.: Clasificación Periódica de Coberturas Terrestres a Escala Regional con Imágenes Modis. *Revista Internacional de Ciencia y Tecnología de la Información Geográfica*. 10, 1--17 (2010)
3. Wardlow, B.D., Egbert, S.L., Kastens, J.H.: Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sensing of Environment*. 108, 290--310 (2007).
4. Foody, G.M.: Status of land cover classification accuracy assessment. *Remote Sensing of Environment*. 80, 185--201 (2002)
5. Lu, D., Batistella, M., Li, G., Moran, E., Hetrick, S., da Costa Freitas, C., Vieira Dutra, L., and Sidnei João Siqueira Sant'Anna.: Land use/cover classification in the Brazilian Amazon using satellite images. *Pesquisa Agropecuária Brasileira*. 47, 1185--1208 (2012)
6. Pal, M., Mather, P.M.: Assessment of the effectiveness of support vector machines for hyperspectral data. *Future Generation Computer Systems*. 20, 1215--1225 (2004).
7. Willington, E., Nolasco, M., Bocco, M.: Clasificación supervisada de suelos de uso agrícola en la zona central de Córdoba (Argentina): comparación de distintos algoritmos sobre imágenes Landsat. In: 42 JAIIO Congreso Argentino de AgroInformática, pp. 207--216. (2013)
8. Murthy, C. S., Raju, P.V., Badrinath, K.S.: Classification of wheat crop with multi-temporal images: performance of maximum likelihood and artificial neural networks. *International Journal of Remote Sensing*. 24, 4871--4890 (2003)
9. Guerschman, J.P., Paruelo, J.M., Di Bella, C., Giallorenzi, M.C., Pacin, F.: Land cover classification in the Argentine Pampas using multi-temporal Landsat TM data. *Int. J. Remote Sens.* 24, 3381--3402 (2003)
10. Saldaña Robles, N., Zavala Segoviano, A., Cabrera Sixto, J., M., Serwatowski, R., García Barrón, S., Gutiérrez Vaca, C., Flores Ortega, A., Flores García, A.: Discriminación de la superficie agrícola sembrada bajo labranza de conservación empleando análisis multitemporal. *Revista Mexicana de Ciencias Agrícolas Pub. Esp.* 4, 679--689 (2012)
11. Ministerio de Agricultura, Ganadería y Pesca de la Nación, <http://www.sii.gov.ar/index.php/series-por-tema/agricultura>. Consultado en abril 2014

12. Bocco, M., Ovando, G., Sayago, S., Willington, E.: Modelos simples para estimar cobertura de suelo agrícola a partir de Imágenes MODIS: aplicación a la zona central de Córdoba (Argentina). In: 42 JAIIO Congreso Argentino de AgroInformática, pp. 61--69. (2013)
13. Monserud, R.A., Leemans, R.: Comparing Global Vegetation Maps with the Kappa Statistic. *Ecological Modelling*. 62, 275--293 (1992)
14. Adriani Johann, J., Vieira Rocha, J., Garbellini Duft, D., Camargo Lamparelli, R. A.: Estimativa de áreas com culturas de verão no Paraná, por meio de imagens multitemporais EVI/Modis. *Pesquisa Agropecuária Brasileira*. 47, 1295--1306 (2012)