

# A methodology for the characterization of land use using medium-resolution spatial images

## Una metodología para la caracterización del uso del suelo mediante imágenes de media resolución espacial

Cristian Guevara-Ochoa<sup>1,4</sup>; Bruno Lara<sup>2,4\*</sup>; Luis Vives<sup>1</sup>; Erik Zimmermann<sup>3,4</sup>; Marcelo Gandini<sup>2</sup>

<sup>1</sup>Instituto de Hidrología de Llanuras “Dr. Eduardo Jorge Usunoff” (IHLLA). República de Italia 780, CC 47 (B7300). Azul, Buenos Aires, Argentina.

<sup>2</sup>Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA), Facultad de Agronomía, Laboratorio de Investigación y Servicios en Teledetección de Azul (LISTA). República de Italia 780, CC 47 (B7300). Azul, Buenos Aires, Argentina.

<sup>3</sup>Centro Universitario Rosario de Investigaciones Hidroambientales (CURIHAM). Riobamba 245 bis (2000). Rosario, Santa Fe, Argentina.

<sup>4</sup>Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Av. Rivadavia 1917, (C1033AAJ). Ciudad Autónoma de Buenos Aires, Argentina.

\*Corresponding author: brunolara73@gmail.com, tel.: +54 (02281) 433291.

### Abstract

**Introduction:** The characterization of land uses represents one of the essential inputs for the management of natural resources at different scales.

**Objective:** To develop a methodology to characterize land use in the upper creek basin from the Azul stream (Buenos Aires, Argentina), through the fusion of satellite images with a medium spatial resolution.

**Materials and methods:** A time-series of 23 images was used from the Normalized Difference Vegetation Index (NDVI) of the MODIS-Terra satellite (product MOD13Q1) for the period May 2015 - May 2016. Landsat 8 images were used to discriminate some categories difficult to classify with NDVI-MODIS. The final cover map was validated regarding verification points independent to the classification process; its accuracy was evaluated by means of the Kappa statistic.

**Results and discussion:** The NDVI time series allowed to recognize the phenological patterns of the covers and land use of greater representativeness in the region. Seven land cover were discriminated; the agricultural uses represented 81.5 % of the surface, double-crop wheat-soya (soybean in Argentina) system predominated (39.4 %). The overall accuracy of the final map was high (88.9 %, Kappa coefficient = 0.86).

**Conclusion:** The methodology used has the advantage of being quick and replicable, to characterize the land uses of a given region and to evaluate its potential changes over time.

### Resumen

**Introducción:** La caracterización de los usos del suelo representa uno de los insumos indispensables para el manejo de los recursos naturales a diferentes escalas.

**Objetivo:** Desarrollar una metodología para caracterizar el uso del suelo en la cuenca superior del arroyo del Azul (Buenos Aires, Argentina), a través de la fusión de imágenes satelitales de media resolución espacial.

**Materiales y métodos:** Se utilizó una serie temporal de 23 imágenes del índice de vegetación de diferencia normalizada (NDVI, por sus siglas en inglés) del satélite MODIS-Terra (producto MOD13Q1) para el periodo mayo 2015 - mayo 2016. Además, se emplearon imágenes Landsat 8 para discriminar algunas categorías difíciles de clasificar con NDVI-MODIS. El mapa final de coberturas se validó considerando puntos de verificación independientes al proceso de clasificación; su precisión se evaluó a través del estadístico Kappa.

**Resultados y discusión:** La serie temporal de NDVI permitió reconocer los patrones fenológicos de las coberturas y usos del suelo de mayor representatividad en la región. Se discriminaron siete coberturas; los usos agrícolas representaron 81.5 % de la superficie, siendo el sistema de doble cultivo trigo-soya (soja en Argentina) el predominante (39.4 %). La precisión global del mapa final fue alta (88.9 %, coeficiente Kappa = 0.86).

**Conclusión:** La metodología empleada tiene la ventaja de ser rápida y replicable, para caracterizar los usos del suelo de una región determinada y evaluar sus cambios potenciales a lo largo del tiempo.

**Keywords:** Remote sensing; NDVI series; MODIS satellite; Landsat satellite; phenological signature.

**Palabras clave:** Sensores remotos; serie NDVI; satélite MODIS; satélite Landsat; firma fenológica.

## Introduction

The alteration of the earth's surface, due to human actions, has generated changes in land uses. For this reason, the characterization of land cover and land use is key to understanding anthropic-environmental interactions at different spatio-temporal scales (Müller & Munroe, 2014). Spatial patterns and their changes have significant implications when analyzing, for example, water balance, climate change, deforestation, sedimentation, loss of biodiversity, and biogeochemical processes of the Earth (DeFries & Eshleman, 2004).

In recent decades, accelerated world population increase has been observed, accompanied by the intensification of agricultural activities and expansion of the agricultural frontier (Eitzinger, Orlandini, Stefanski, & Naylor, 2010). These structural transformations have caused the alteration of hydrological processes (Breuer et al., 2009; Fohrer, Haverkamp, Eckhardt, & Frede, 2001; Li, Liu, Zhang, & Zheng, 2009), imbalances in carbon and nitrogen cycles (Murty, Kirschbaum, Mcmurtrie, & Mcgilvray, 2002; Post & Kwon, 2000) and losses of species and ecosystems at high rates, despite efforts to conserve biological diversity (Pettorelli et al., 2016).

In Argentina, particularly in the Pampean Region, there has been an agriculturalization process characterized by the increase of agricultural areas in traditional livestock use areas. This process began in the 1960s and has accelerated dramatically during the last 20 to 25 years, sustained mainly by structural economic changes and the incorporation of the technological package associated with the cultivation of genetically modified soya (soybean in Argentina) (Manuel-Navarrete et al., 2009). These changes strongly impact the structure and functioning of ecosystems, are perceived at different scales (Baldi, Guerschman, & Paruelo, 2006; Guerschman, Paruelo, & Burke, 2003; Lara & Gandini, 2014) and generate alterations in productivity patterns, stability and sustainability (Viglizzo, Roberto, Lértora, Gay, & Bernardos, 1997).

The availability and use of satellite images have allowed the analysis of large areas, acquiring information at very different temporal and spatial scales (Horning, Robinson, Sterling, Turner, & Spector, 2010). The use of these images is associated with the description of land cover /land uses, by characterizing the heterogeneity of the landscape and its spatial pattern (Cabello & Paruelo, 2008; Paruelo, 2008). Such application is based on the response forms of the Earth's surface elements at different wavelengths of the electromagnetic spectrum (Chuvieco, 1990), although often one of the main limitations is the scarce field information or its inadequate capture. In this context, several spectral indices derived from

## Introducción

La alteración de la superficie terrestre, a causa de las acciones humanas, ha generado que cambien los usos del suelo. Por esta razón, la caracterización de las coberturas y usos del suelo es clave para comprender las interacciones antrópico-ambientales a diferentes escalas espacio-temporales (Müller & Munroe, 2014). Los patrones espaciales y sus cambios tienen implicaciones significativas cuando se analizan, por ejemplo, el balance hídrico, el cambio climático, la deforestación, la sedimentación, la pérdida de biodiversidad, y los procesos biogeoquímicos de la Tierra (DeFries & Eshleman, 2004).

En las últimas décadas se ha observado el aumento acelerado de la población mundial, acompañado de la intensificación de actividades agrícolas y de expansión de la frontera agropecuaria (Eitzinger, Orlandini, Stefanski, & Naylor, 2010). Estas transformaciones estructurales han provocado la alteración de procesos hidrológicos (Breuer et al., 2009; Fohrer, Haverkamp, Eckhardt, & Frede, 2001; Li, Liu, Zhang, & Zheng, 2009), desbalances en los ciclos del carbono y nitrógeno (Murty, Kirschbaum, Mcmurtrie, & Mcgilvray, 2002; Post & Kwon, 2000) y pérdidas de especies y ecosistemas a tasas altas, a pesar de los esfuerzos por conservar la diversidad biológica (Pettorelli et al., 2016).

En Argentina, particularmente en la región pampeana, se ha experimentado un proceso de agriculturización caracterizado por el incremento de las áreas agrícolas en zonas de uso tradicionalmente ganadero. Dicho proceso comenzó en la década de 1960 y se ha acelerado de manera dramática durante los últimos 20 a 25 años, sostenido principalmente por cambios económicos estructurales y la incorporación del paquete tecnológico asociado al cultivo de soya (soja en Argentina) genéticamente modificada (Manuel-Navarrete et al., 2009). Estos cambios impactan fuertemente la estructura y funcionamiento de los ecosistemas, se perciben a diferentes escalas (Baldi, Guerschman, & Paruelo, 2006; Guerschman, Paruelo, & Burke, 2003; Lara & Gandini, 2014) y generan alteraciones en los patrones de productividad, estabilidad y sostenibilidad (Viglizzo, Roberto, Lértora, Gay, & Bernardos, 1997).

La disponibilidad y uso de imágenes satelitales han permitido el análisis de grandes extensiones de superficie, adquiriendo información a escalas temporal y espacial muy diferentes (Horning, Robinson, Sterling, Turner, & Spector, 2010). El uso de dichas imágenes está asociado a la descripción de tipos de coberturas/usos del suelo, mediante la caracterización de la heterogeneidad del paisaje y su patrón espacial (Cabello & Paruelo, 2008; Paruelo, 2008). Tal aplicación está sustentada en las formas de respuesta de los elementos de la

remote sensors have been proposed as estimators of vegetation presence and condition, based on the observation of spectral signatures of typical leaf reflectance. In these, the energy reflected in the visible range of the spectrum is very low, due to the high absorption of photosynthetically active pigments, with maximum absorption at blue (470 nm) and red (670 nm) wavelengths; almost all of the near infrared radiation (NIR) is scattered (reflected and transmitted) with very little absorption, due to the mesophyll structure of leaves (Solano, Didan, Jacobson, & Huete, 2010). As a result, the contrast between red and NIR responses is a sensitive measure of the vegetation amount, with maximum differences over dense and minimal vegetation where there is little or no cover (Solano et al., 2010). One of the most widely used spectral indices is the Normalized Difference Vegetation Index (NDVI), which is a linear estimator of the fraction of absorbed photosynthetically active radiation intercepted by vegetation (fAPAR) (Wang et al., 2004). From the data provided by satellite sensors of high temporal resolution (short revisit time, i.e. MODIS), hypertemporal series of NDVI can be constructed, from which it is possible to extract valuable information about the vegetation phenology (Reed, Schwartz, & Xiangming, 2009). In this way, it is possible to use differences in seasonal patterns to characterize various types of vegetation (summer crops, winter crops, perennial vegetation, etc.). On the other hand, the available cover mapping and land use on a large scale (Blanco et al., 2013; Eva et al., 2004) show inconsistencies due to their continental approximation and are insufficient to be applied on a more detailed scale. In this sense, the objective of this study was to develop a methodology to characterize land use in the upper creek basin of Del Azul in Buenos Aires, Argentina, by means of medium spatial resolution images from Landsat and MODIS satellites.

## Materials and methods

### Study area

The upper creek basin of Del Azul (Figure 1) is located in the center of the province of Buenos Aires, Argentina, between 59° 48' - 60° 0' W and 37° 01' 48'' - 37° 09' 36'' S, draining an area of 1 024 km<sup>2</sup>. This basin originates towards the South in the town of Chillar (60 km south of the city of Azul); its main tributaries are Videla stream with an area of 120 km<sup>2</sup> and Santa Catalina stream with 140 km<sup>2</sup>.

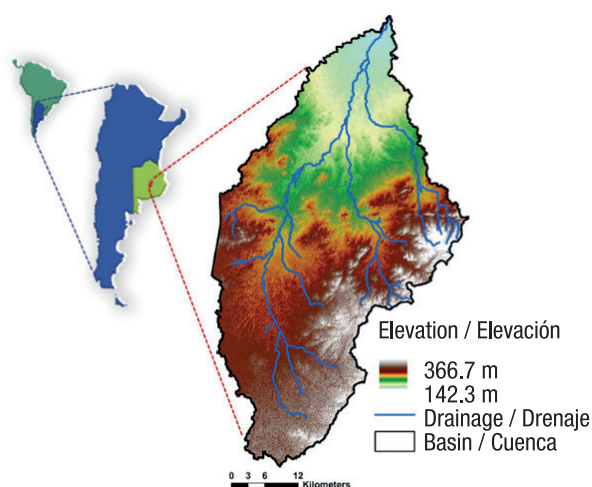
The altitude in the basin varies from 142 to 367 m. The highest part, located towards the southeast, is on the Tandilia's System, where slopes are greater than 6 %. Towards the northwest, the basin is transformed into a transitional plain, in the medium-low sector, with slopes of less than 1 %. The basin limits to the North with the city of Azul, to the West with the city of

superficie terrestre a distintas longitudes de onda del espectro electromagnético (Chuvieco, 1990), aunque a menudo una de las limitantes principales es la escasa información de campo o su captura inadecuada. En este contexto, varios índices espectrales derivados de sensores remotos han sido propuestos como estimadores de la presencia y condición de la vegetación, basados en la observación de las firmas espectrales de reflectancia típica de las hojas. En éstas, la energía reflejada en el rango visible del espectro es muy baja, debido a la alta absorción de los pigmentos fotosintéticamente activos, con absorciones máximas en las longitudes de onda del azul (470 nm) y del rojo (670 nm); casi la totalidad de la radiación del infrarrojo cercano (IRC) es dispersada (reflejada y transmitida) con muy poca absorción, debido a la estructura del mesófilo de las hojas (Solano, Didan, Jacobson, & Huete, 2010). Como resultado, el contraste entre las respuestas del rojo y del IRC es una medida sensible de la cantidad de vegetación, con máximas diferencias sobre la vegetación densa y mínimas donde existe poca o nula cobertura (Solano et al., 2010). Uno de los índices espectrales más utilizados es el índice de vegetación de diferencia normalizada (NDVI, por sus siglas en inglés), el cual es un estimador lineal de la fracción de radiación fotosintéticamente activa interceptada por la vegetación (fAPAR) (Wang et al., 2004). A partir de los datos brindados por los sensores satelitales de alta resolución temporal (tiempo corto de revisita, i. e. MODIS) pueden construirse series hipertemporales de NDVI, de las cuales es posible extraer información valiosa acerca de la fenología de la vegetación (Reed, Schwartz, & Xiangming, 2009). De esta manera, es posible utilizar las diferencias en los patrones estacionales para caracterizar diversos tipos de vegetación (cultivos de verano, invierno, vegetación perenne, etc.). Por otro lado, los mapeos disponibles de coberturas y usos del suelo a gran escala (Blanco et al., 2013; Eva et al., 2004) presentan inconsistencias debido a su aproximación continental y resultan insuficientes para poder ser aplicados a una escala más detallada. En este sentido, el objetivo del trabajo fue elaborar una metodología para la caracterización del uso del suelo en la cuenca superior del arroyo del Azul en Buenos Aires, Argentina, mediante el uso de imágenes de media resolución espacial de los satélites Landsat y MODIS.

## Materiales y métodos

### Área de estudio

La cuenca superior del arroyo del Azul (Figura 1) se ubica en el centro de la provincia de Buenos Aires, Argentina, entre 59° 48' - 60° 0' LO y 37° 01' 48'' - 37° 09' 36'' LS, drenando una superficie de 1 024 km<sup>2</sup>. Dicha cuenca nace hacia el sur en la localidad de Chillar (60 km al sur de la ciudad de Azul); sus principales afluentes son el arroyo Videla con una superficie de 120 km<sup>2</sup> y el de Santa Catalina con 140 km<sup>2</sup>.



**Figure 1. Upper creek basin of Del Azul located in the center of the province of Buenos Aires, Argentina.**

**Figura 1. Cuenca superior del arroyo del Azul ubicada en el centro de la provincia de Buenos Aires, Argentina.**

Olavarría, to the East with the city of Tandil and to the South with the town of Chillar. Mean annual rainfall is 913 mm. The highest rainfall occurs in the spring-summer period with a monthly average of around 100 mm, and the lowest occurs in autumn-winter with a monthly average of 50 mm. December to March are the months with the highest temperatures with an average of 20 °C, while from June to August are the months with the coldest temperatures with an average of 8 °C.

#### Satellite data

Two sources of medium spatial resolution satellite images were used, contributed by the National Aeronautics and Space Administration (NASA). First, we used a time-series of 23 NDVI-MODIS images from the product MOD13Q1, scene h13v12 ([https://lpdaac.usgs.gov/dataset\\_discovery/modis/modis\\_products\\_table/mod13q1](https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13q1)) that included one year of data, from May 2015 to May 2016. Although it does not adjust to the calendar year, this time window allows obtaining information on the growing seasons of land cover in the southern hemisphere; for example, the beginning of the growth of winter season crops such as wheat and the end of the harvest of summer crops such as soybeans and corn can be appreciated. This data set consists of 16-day compositions with the maximum NDVI values estimated at a spatial resolution of 250 m. Additionally, two Landsat 8 images (one for winter and one for summer) were used from Path/Row 225/086 (<https://landsat.usgs.gov/landsat-8>). Only the reflective bands (blue, green, red, near infrared and two medium infrared bands) having a spatial resolution of 30 m were used, while the thermal bands were rejected. This selection of spectral bands is due to the fact that covers

La altitud en la cuenca varía de 142 a 367 m. La parte más alta, ubicada hacia el sureste, se encuentra sobre el Sistema de Tandilia, donde las pendientes son mayores de 6 %. Hacia el noroeste, la cuenca se transforma en una llanura de transición, en el sector medio-bajo, con pendientes menores de 1 %. La cuenca limita al norte con la ciudad de Azul, al oeste con la ciudad de Olavarría, al este con la ciudad de Tandil y al sur con la ciudad de Chillar. La precipitación media anual es de 913 mm. Las mayores precipitaciones se presentan en el periodo primavera-verano con un promedio mensual alrededor de 100 mm, y las más bajas ocurren en otoño-invierno con una media mensual de 50 mm. Los meses de diciembre a marzo son los de mayor temperatura con un promedio de 20 °C, mientras que los meses de junio a agosto son los más fríos con un promedio de 8 °C.

#### Datos satelitales

Se utilizaron dos fuentes de imágenes satelitales de media resolución espacial, aportadas por la agencia espacial norteamericana (NASA). En primer lugar, se utilizó una serie temporal de 23 imágenes NDVI-MODIS provenientes del producto MOD13Q1, escena h13v12 ([https://lpdaac.usgs.gov/dataset\\_discovery/modis/modis\\_products\\_table/mod13q1](https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13q1)) que comprendió un año de datos, desde mayo de 2015 a mayo de 2016. Aunque no se ajusta al año calendario, esta ventana temporal permite obtener información de las estaciones de crecimiento de las coberturas en el hemisferio sur; por ejemplo, se puede apreciar el inicio del crecimiento de cultivos de estación invernal como el trigo y el fin de la cosecha de cultivos de verano como soja y maíz. Este conjunto de datos consiste en composiciones de 16 días con los valores máximos de NDVI calculados a una resolución espacial de 250 m. Adicionalmente, se



to be identified show a greater dynamic range in the reflective bands and not in the thermal bands (Lara & Gandini, 2014).

### Classifying and obtaining land cover

From the time series of NDVI-MODIS images, an unsupervised classification was made using the ISODATA (Iterative Self-Organizing Data Analysis) algorithm, which makes the partition of an n-dimensional image into a number of conglomerates according to the established parameters. In general, the process begins with a number of random centroids and, subsequently, each of the pixels is assigned to that conglomerate whose centroid is closest, in terms of the euclidean distances. This is an iterative process in which the means (centroids) are recalculated and the pixels are reclassified with respect to the new means. The process continues until the number of pixels in each class changes less than the proposed change threshold or until the configured maximum number of iterations is reached (Tou & Gonzalez, 1974).

In this procedure we considered 50 iterations, a tolerance threshold of 5 % and maximum standard deviation of 1. In this way, 15 classes were generated, from which the mean value of NDVI was obtained for each of the 23 dates contemplated in the temporary series. The average values correspond to the phenological signatures, which were associated, based on the annual behavior, to each one of defined land cover and land uses: soybean, corn, double wheat-soybean cultivation systems, grassland (natural and cultivated) and mixed areas (Figure 2). For example, crops were associated with classes that had one or two periods of strong growth and a period with low NDVI values. To discriminate between crops, the unimodal curves with maximum values of NDVI in February were classified as soybeans, and those that showed a similar pattern, but with the maximum values of NDVI in January and February were classified as corn. Those classes that showed a maximum during September-October (characteristic of winter crops) and, in addition, a maximum in February, were categorized as wheat-soybeans. On the other hand, classes that did not have a unimodal or bimodal average pattern were classified as grassland and mixed areas. Grassland, which include cultivated perennial grassland and natural grassland, were characterized by maintaining an average value of NDVI greater than 0.4 throughout growing season; while mixed areas correspond to an average seasonal value of less than 0.4. The latter correspond to urban areas, bodies of water and rocky outcrop areas with active vegetation.

Probably the mixed areas that did not differentiate from the MODIS images are associated with mixed pixels at that scale, which would make it difficult to

emplearon dos imágenes Landsat 8 (una de invierno y otra de verano) del Path/Row 225/086 (<https://landsat.usgs.gov/landsat-8>). Solo se usaron las bandas reflectivas (azul, verde, rojo, infrarrojo cercano y dos bandas de infrarrojo medio) que poseen una resolución espacial de 30 m, mientras que las bandas térmicas fueron descartadas. Esta selección de bandas espectrales se debe a que las coberturas a identificar presentan un rango dinámico mayor en las bandas reflectivas y no en las bandas térmicas (Lara & Gandini, 2014).

### Clasificación y obtención de coberturas del suelo

A partir de la serie temporal de imágenes NDVI-MODIS se hizo una clasificación no supervisada con el algoritmo ISODATA (*Iterative Self-Organizing Data Analysis*), el cual hace la partición de una imagen de n-dimensiones en un número de conglomerados de acuerdo con los parámetros establecidos. Por lo general, el proceso comienza dando un número de centroides aleatorios y, posteriormente, cada uno de los píxeles es asignado a aquel conglomerado cuyo centroide se encuentra más cercano, en términos de las distancias euclidianas. Este es un proceso iterativo en el cual las medias (centroides) se recalculan y los píxeles se reclasifican con respecto a las nuevas medias. El proceso continúa hasta que el número de píxeles en cada clase cambia menos que el umbral de cambio propuesto o hasta que se alcanza el número máximo de iteraciones configurado (Tou & Gonzalez, 1974).

En este procedimiento se consideraron 50 iteraciones, un umbral de tolerancia del 5 % y desviación estándar máxima de 1. De esta forma, se generaron 15 clases de las cuales se obtuvo el valor medio de NDVI para cada una de las 23 fechas contempladas en la serie temporal. Los valores medios corresponden a las firmas fenológicas, las cuales fueron asociadas, con base en el comportamiento anual, a cada una de las coberturas y usos del suelo definidas: soya, maíz, sistemas de doble cultivo trigo-soya, pasturas (naturales y cultivadas) y áreas mezcladas (Figura 2). Por ejemplo, los cultivos se asociaron a las clases que presentaron uno o dos periodos de marcado crecimiento y un periodo con bajos valores de NDVI. Para discriminar entre cultivos, las curvas unimodales con valores máximos de NDVI en febrero se clasificaron como soya, y las que mostraban un patrón similar, pero con los valores máximos de NDVI en enero y febrero se clasificaron como maíz. Aquellas clases que presentaron un máximo durante septiembre-octubre (característico de cultivos invernales) y, además, un máximo en febrero, se categorizaron como trigo-soya.

Por otro lado, las clases que no presentaron un patrón promedio unimodal o bimodal se clasificaron como pasturas y áreas mezcladas. Las pasturas, que incluyen

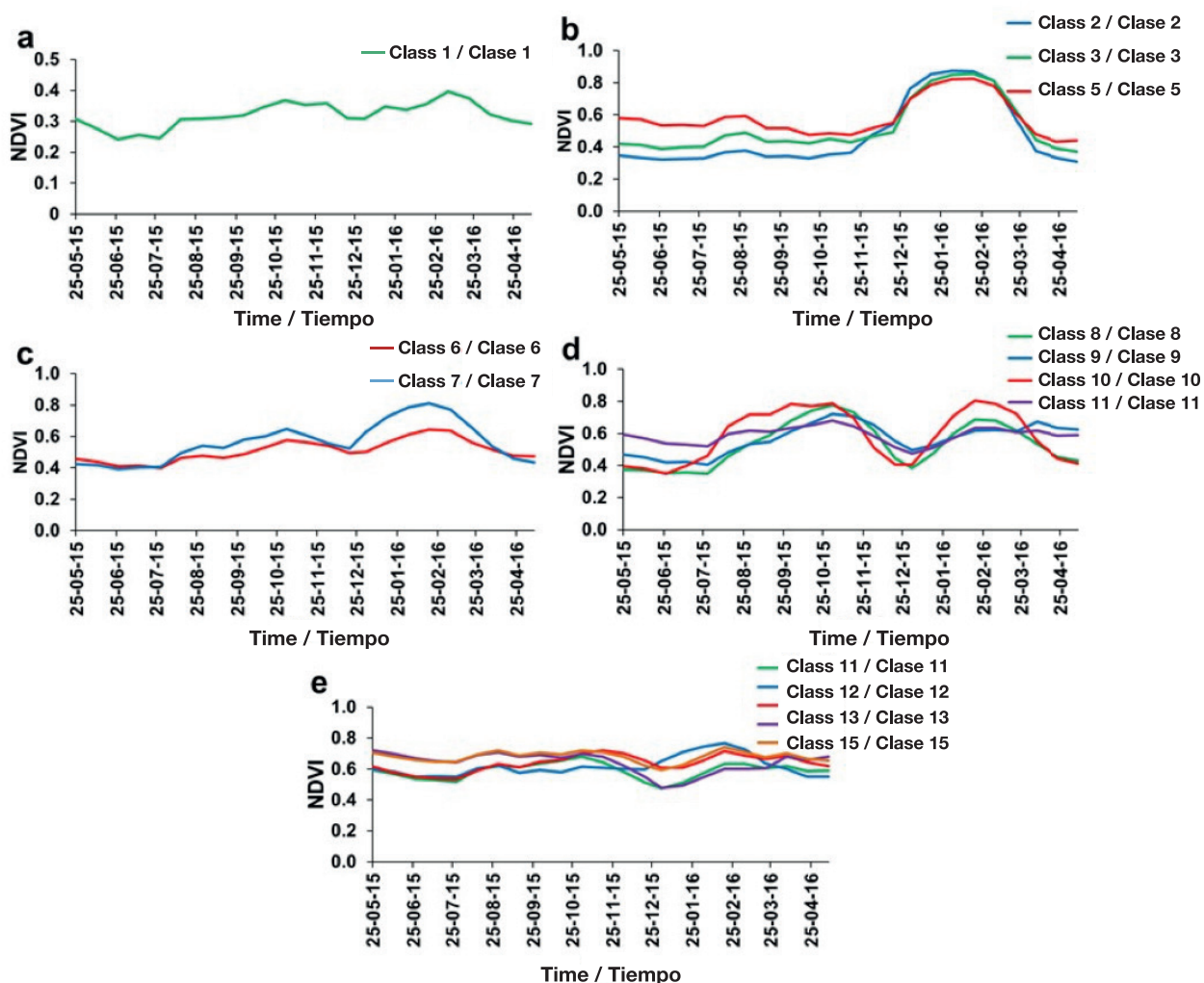


Figure 2. Seasonal curves of the Normalized Difference Vegetation Index (NDVI) from May 2015 to May 2016 for each of the 15 classes resulting from the unsupervised classification, scene h13v12 from the MODIS satellite. a) mixed area, b) soybean, c) corn, d) wheat-soybean and e) grassland.

Figura 2. Curvas estacionales del índice de vegetación de diferencia normalizada (NDVI) en el periodo mayo del 2015 a mayo del 2016 para cada una de las 15 clases resultantes de la clasificación no supervisada, escena h13v12 del satélite MODIS. a) área mezclada, b) soja, c) maíz, d) trigo-soya y e) pastura.

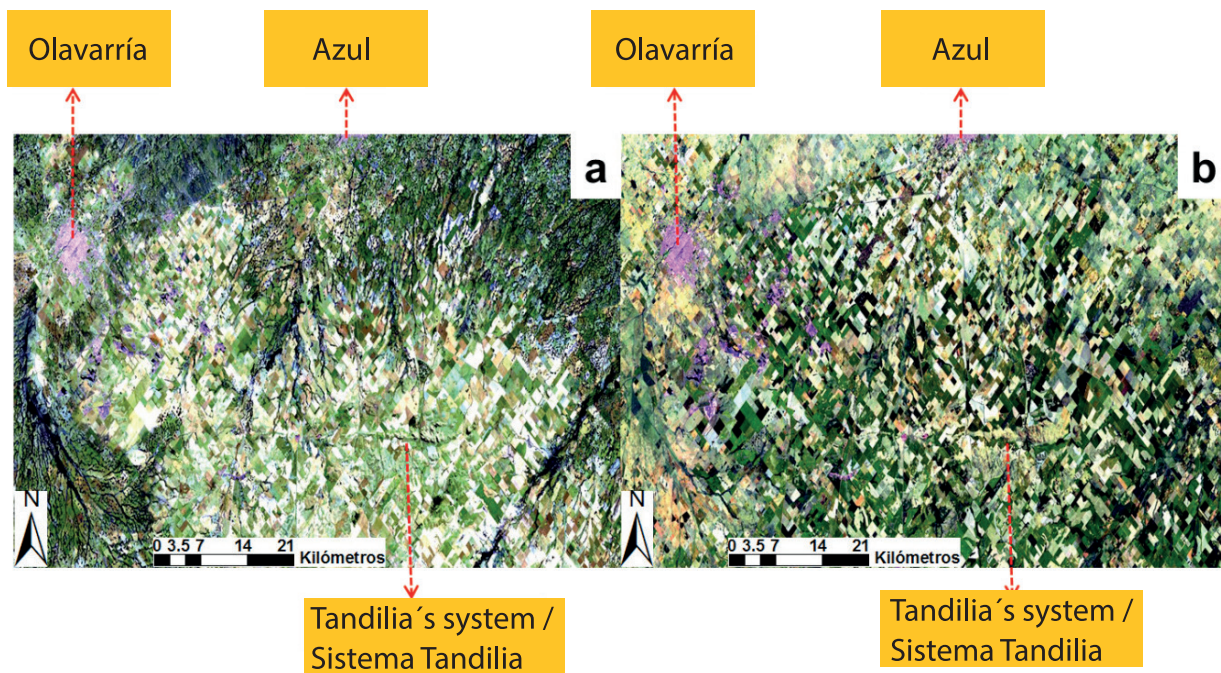
find pure pixels to differentiate between covers (Lara & Gandini, 2016). For this reason, in order to discriminate the areas, two Landsat 8 images (winter and summer season) were used in a combination 7, 6, 4, which allows to penetrate atmospheric particles, smoke and haze (Figure 3). Subsequently, the two data sources were merged, resampling the classification made with MODIS images previously, at a spatial resolution of 30 m to obtain the final map of land cover and land uses of the upper creek basin of Del Azul.

#### Validation of the cover map

The final cover map was validated regarding independent verification points to the classification process, taken during the 2015-2016 agricultural

pastizales perennes cultivados y pastizales naturales, se caracterizaron por mantener un valor promedio de NDVI mayor de 0.4 a lo largo de la estación de crecimiento; mientras que las áreas mezcladas corresponden a un valor estacional promedio menor de 0.4. Estas últimas corresponden a zonas urbanas, cuerpos de agua y áreas de afloramiento rocoso con vegetación activa.

Es probable que las áreas mezcladas que no lograron diferenciarse a partir de las imágenes MODIS estén asociadas a píxeles mixtos a esa escala, lo cual dificultaría hallar píxeles puros para diferenciar entre coberturas (Lara & Gandini, 2016). Por esta razón, para poder discriminar las áreas, se utilizaron dos imágenes Landsat 8 (época invernal y estival) en una combinación 7, 6, 4, que permite penetrar



**Figure 3.** Landsat 8 images used. (a) Winter season, August 15, 2015. (b) Summer season, January 22, 2016. Path/ Row 225/086. Composition of RGB bands: 7, 6, 4. The vegetation appears in dark green; urban areas are gray, greenish blue or purple; water is black or dark blue; and rocky outcroppings appear in red or yellow.

**Figura 3.** Imágenes Landsat 8 utilizadas. (a) Época invernal, 15 de agosto de 2015. (b) Época estival, 22 de enero de 2016. Path/Row 225/086. Composición de bandas RGB: 7, 6, 4. La vegetación aparece en tonos de verde oscuro; las áreas urbanas son de color gris, azul verdoso o púrpura; el agua es de color negro o azul oscuro; y los afloramientos rocosos aparecen en tonos de rojo o amarillo.

campaign. We used the Kappa statistic (Congalton, 1991) which, based on a confusion matrix, takes into account commission errors (or inclusion) and errors of omission (or exclusion); the classification is more accurate the closer it is to a value of 1. The equation for the Kappa statistic is as follows:

$$k = \frac{N \sum_{i=1}^r (x_{ii}) - \sum_{i=1}^r (x_{i+} * x_{+i})}{N^2 \sum_{i=1}^r (x_{i+} * x_{+i})}$$

where,

N = total number of reference pixels

$x_{ii}$  = number of observations in row  $i$  and column  $i$

$x_{i+}$  = marginal total of row  $i$

$x_{+i}$  = marginal total of column  $i$

$r$  = number of rows in the matrix

Nominally, the value of the index varies between -1 and 1, but the following scale can be considered to assess a classification: 0.05 to 0.20 = very low accuracy; 0.20 to 0.40 = low accuracy; 0.40 to 0.55 = reasonable accuracy; 0.55 to 0.70 = good accuracy; 0.70 to 0.85 = very good accuracy; 0.85 to 0.99 = excellent accuracy; and 0.99 to 1 = perfect accuracy.

partículas atmosféricas, humo y neblina (Figura 3). Posteriormente, las dos fuentes de datos se fusionaron remuestreando la clasificación realizada con imágenes MODIS previamente, a una resolución espacial de 30 m, para obtener el mapa final de coberturas y usos del suelo de la cuenca superior del arroyo del Azul.

#### Validación del mapa de coberturas

El mapa final de coberturas se validó considerando puntos de verificación independientes al proceso de clasificación, tomados durante la campaña agrícola 2015-2016. Se utilizó el estadístico Kappa (Congalton, 1991) que, basado en una matriz de confusión, toma en cuenta los errores de comisión (o inclusión) y los errores de omisión (o exclusión); la clasificación es más precisa cuanto más cercana se encuentre a un valor de 1. La ecuación para el estadístico Kappa es la siguiente:

$$k = \frac{N \sum_{i=1}^r (x_{ii}) - \sum_{i=1}^r (x_{i+} * x_{+i})}{N^2 \sum_{i=1}^r (x_{i+} * x_{+i})}$$

donde,

N = número total de píxeles de referencia



## Results and discussion

Figure 4 shows the final map of land cover and land uses of the upper creek basin of Del Azul, Argentina. The map was obtained from the combination of medium spatial resolution images, to discriminate crop types and grassland areas based on their characteristic phenological patterns. Table 1 indicates the area covered by each of the land uses; the double wheat-soybean crop system is the most widespread, representing 39.4 % (40 470.4 ha) of the study area. The sum of the areas covered by all agricultural uses reaches 81.5 %, which reflects the magnitude of the agriculturalization process in the region, taking into account that the natural vegetation was natural grassland (Matteucci, 2012). Although there are no official data, agricultural statistics of the Ministerio de Agroindustria de Argentina (Sistema Integrado de Información Agropecuaria [SIIA], 2016) reflect to a certain extent, the predominance of the same crops in the 2015-2016 campaign for the department of Azul, such as soybean (124 700 ha), wheat (132 110 ha) and corn (43 300 ha); it must be taken into account that in these data there are areas of agricultural use that are external to the study area.

$x_{ii}$  = número de observaciones en la fila  $i$  y en la columna  $i$   
 $x_{i+}$  = total marginal de la fila  $i$   
 $x_{+i}$  = total marginal de la columna  $i$   
 $r$  = número de filas en la matriz

Nominalmente, el valor del índice varía entre -1 y 1, pero puede considerarse la siguiente escala para valorar una clasificación: 0.05 a 0.20 = muy baja precisión; 0.20 a 0.40 = baja precisión; 0.40 a 0.55 = precisión razonable; 0.55 a 0.70 = buena precisión; 0.70 a 0.85 = muy buena precisión; 0.85 a 0.99 = excelente precisión; y 0.99 a 1 = precisión perfecta.

## Resultados y discusión

La Figura 4 muestra el mapa final de coberturas y usos del suelo de la cuenca superior del arroyo del Azul, Argentina. El mapa se obtuvo a partir de la combinación de imágenes de media resolución espacial, para discriminar tipos de cultivos y áreas de pastizales con base en sus patrones fenológicos característicos. El Cuadro 1 indica la superficie cubierta por cada uno de los usos del suelo; el sistema de doble cultivo trigo-soya es el más extendido representando 39.4 % (40 470.4 ha) del área de estudio. La suma de las superficies cubiertas

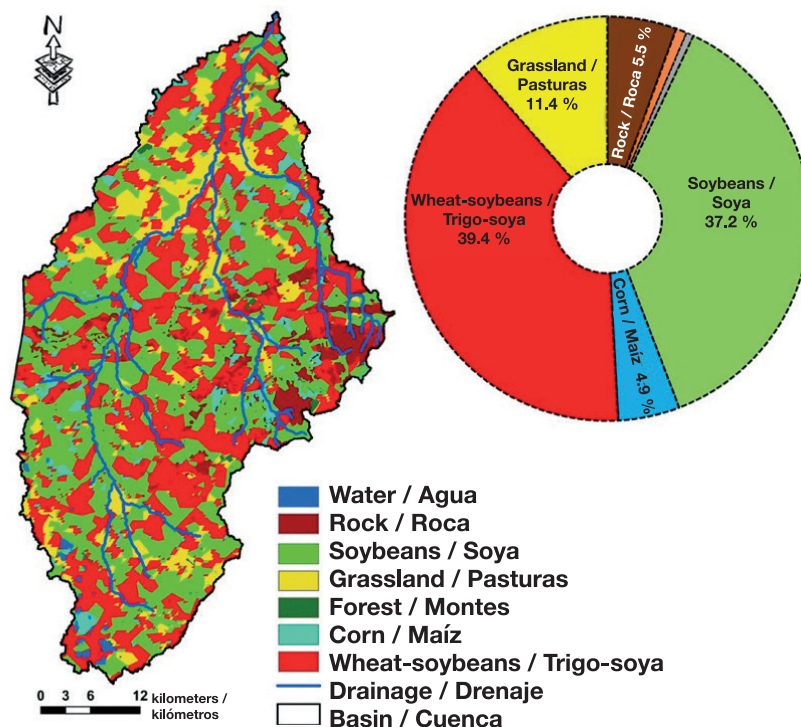


Figure 4. Final map of land uses in the upper creek basin of Del Azul (Buenos Aires, Argentina) for the period 2015-2016, using MODIS and Landsat 8 images.

Figura 4. Mapa final de usos del suelo en la cuenca superior del arroyo del Azul (Buenos Aires, Argentina) para el periodo 2015-2016, mediante el uso de imágenes MODIS y Landsat 8.



**Table 1. Area covered by land uses in the upper creek basin of Del Azul in Argentina.****Cuadro 1. Superficie cubierta por los usos del suelo en la cuenca superior del arroyo del Azul en Argentina.**

Land use/Usos del suelo	Study area/Área de estudio	
	(ha)	(%)
Rock/Roca	5 694.32	5.55
Forest/Montes	903.15	0.88
Water/Agua	593.19	0.58
Soybeans/Soya	38 171.65	37.20
Corn/Maíz	5 043.80	4.92
Wheat-soybeans/Trigo-soya	40 470.43	39.44
Grassland/Pasturas	11 725.06	11.43
Total	102 601.6	100

In general, the cultivated areas are in the highest part of the basin, since there predominate soils of greater suitability and with few restrictions for agricultural activity, although in recent years, the advance of crops has also occurred on areas of traditional livestock use (Lara & Gandini, 2014). Grassland (which include natural and cultivated grassland) are associated, mainly, with the rocky outcroppings of the Tandilia's system and in low areas of the basin. Probably, the areas close to the rocky outcrops represent areas of natural grassland where it is not possible to till the soil, while those located in low areas of the basin represent, mostly, cultivated grassland that are usually used in areas with drainage and salinization limitations. The current methodology did not allow to obtain a clear separation in the phenological signatures of these grasslands, due to the overlap of the start and end dates of growing seasons was frequent.

The map of land uses showed good success levels, with a total accuracy of 88.9 % and a Kappa coefficient of 0.86. According to Table 2, the lowest precision of the producer corresponded to the uses of corn and soybeans with values of 69.9 and 76.0 %, respectively. Precisely, the greatest confusion in the classification occurs between these two uses, which is to be expected since their phenological signatures are similar (Figure 2). In spite of this, the accuracy of the final map, obtained from a methodology based on phenological signatures derived from NDVI and without field information, satisfactorily represented the spatio-temporal patterns of the main uses of the basin in a given period. Undoubtedly, the combination of data obtained at a higher spatial resolution contributed to achieving greater precision.

por todos los usos agrícolas alcanza 81.5 %, lo cual refleja la magnitud del proceso de agriculturización en la región, teniendo en cuenta que la vegetación natural era el pastizal natural (Matteucci, 2012). A pesar de que no existen datos oficiales, las estadísticas agropecuarias del Ministerio de Agroindustria de Argentina (Sistema Integrado de Información Agropecuaria [SIIA], 2016) reflejan en cierta medida, el predominio de los mismos cultivos en la campaña 2015-2016 para el departamento de Azul, tales como soya (124 700 ha), trigo (132 110 ha) y maíz (43 300 ha); se debe tener en cuenta que en estos datos existen áreas de uso agrícola que son externas al área de estudio.

En general, las áreas cultivadas se encuentran en la porción más alta de la cuenca, ya que ahí predominan los suelos de mayor aptitud y con pocas restricciones para la actividad agrícola, aunque en los últimos años, el avance de los cultivos también se ha dado sobre áreas de uso tradicionalmente ganadero (Lara & Gandini, 2014). Las áreas de pasturas (que incluyen a los pastizales naturales y las pasturas cultivadas) se encuentran asociadas, principalmente, a los afloramientos rocosos del Sistema de Tandilia y en zonas bajas de la cuenca. Probablemente, las zonas cercanas a los afloramientos rocosos representan zonas de pastizal natural donde no es posible labrar el suelo, mientras que las ubicadas en zonas bajas de la cuenca representan, en su mayoría, pasturas cultivadas que suelen utilizarse en zonas con limitaciones de drenaje y salinización. La metodología actual no permitió obtener una separación clara en las firmas fenológicas de estos pastizales, debido a que el solapamiento de las fechas de inicio y final de las estaciones de crecimiento fue frecuente.

**Table 2. Confusion matrix for the map of land uses of the upper creek basin of Del Azul, Argentina, in the period 2015-2016. The diagonal contains the correctly classified pixels (5 240) and the errors are located outside the diagonal; along the horizontal are commission (inclusion) errors and along the vertical are errors of omission (exclusion).**

**Cuadro 2. Matriz de confusión para el mapa de usos del suelo de la cuenca superior del arroyo del Azul, Argentina, en el periodo 2015-2016. La diagonal contiene los pixeles clasificados correctamente (5 240) y por fuera de la diagonal se encuentran los errores; a lo largo de la horizontal son errores de comisión (inclusión) y a lo largo de la vertical son errores de omisión (exclusión).**

Classified data/ Datos Clasificados	Field data/Datos de campo							Total	User accuracy (%) / Precisión del usuario (%)
	Water/Agua	Rock/Roca	Soybeans/Soya	Grassland/ Pasturas	Wheat-soybeans/ Trigo-soya	Corn/Maíz	Forest/Montes		
Water/Agua	532	0	0	0	0	0	0	532	100
Rock/Roca	0	486	0	0	0	0	0	486	100
Soybeans/Soya	0	0	1 258	0	48	177	0	1 483	84.83
Grassland/Pasturas	0	0	0	735	0	0	14	749	98.13
Wheat-soybeans/ Trigo-soya	0	0	65	0	1 689	0	0	1 754	96.29
Corn/Maíz	0	0	331	0	0	412	0	743	55.45
Forest/Montes	2	0	0	14	0	0	128	144	88.89
Total	534	486	1 654	749	1 737	589	142	5 891	
Producer accuracy (%) / Precisión del usuario (%)	99.62	100	76.05	98.13	97.23	69.95	90.14		

Total accuracy:  $(5\ 240/5\ 891) * 100 = 88.95\ %$  / Precisión total:  $(5\ 240/5\ 891) * 100 = 88.95\ %$   
 Kappa statistic = 0.86 / Estadístico Kappa = 0.86

The methodology developed has the advantage of being relatively fast and replicable with few inputs, to be used in different regions in order to map the most representative uses; however, it has the disadvantage that the annual behavior of the uses to be discriminated must be known *a priori*, which could be avoided by taking field data in pure pixels to obtain the phenological signature.

The characterization of land cover and land uses represents one of the essential inputs for the monitoring and management of natural resources at different scales, such as semi-distributed hydrological modeling (Arnold et al., 2012); however, in Argentina it is usually one of the missing information in departmental or regional management bodies. Therefore, when strategies of sustainable management of natural resources are required at a local scale, the existing large-scale procedures (Blanco et al., 2013; Eva et al., 2004) are deficient, and that is when the methodology

El mapa de usos del suelo mostró buenos niveles de aciertos, con una precisión total de 88.9 % y un coeficiente Kappa de 0.86. De acuerdo con el Cuadro 2, las precisiones más bajas del productor correspondieron a los usos maíz y soya con valores de 69.9 y 76.0 %, respectivamente. Justamente, la mayor confusión en la clasificación se da entre estos dos usos, lo cual es esperable ya que sus firmas fenológicas son similares (Figura 2). A pesar de esto, la precisión del mapa final, obtenido a partir de una metodología basada en firmas fenológicas derivadas de NDVI y sin información de campo, representó de manera satisfactoria los patrones espacio-temporales de los principales usos de la cuenca en un periodo determinado. Sin duda, la combinación de datos obtenidos a una mayor resolución espacial contribuyó a lograr una precisión mayor.

La metodología desarrollada posee la ventaja de ser relativamente rápida y replicable con pocos insumos, para ser utilizada en diferentes regiones con el fin

developed in this study can provide useful tools. On the other hand, in regions of great dynamism of land uses, such as the one analyzed here, the methodology would allow to evaluate and monitor cover changes and, in particular, of the agricultural frontier, practically in real time.

## Conclusions

The methodology developed in this study allowed the characterization of the land cover and land uses of the upper creek basin of Del Azul, Argentina, using phenological signatures derived from NDVI-MODIS time series. The mixed areas that could not be differentiated were categorized from Landsat 8 images, to improve the accuracy of the final map. The classification method used is a valid alternative when precise field data are not available, but requires knowledge about the seasonal behavior of the covers to be mapped. In this way, the proposal presented can be applicable and reproducible in different regions and scales, becoming a valuable tool for monitoring changes in land use at the local and regional scales.

## Acknowledgments

The authors thank the CONICET (for providing doctoral scholarships for Guevara-Ochoa and Lara) and to the Secretaría de Ciencia y Técnica of UNCPBA for the funding provided to carry out this study.

*End of English version*

## References / Referencias

- Arnold, J. G., Kiniry, J. R., Srinivasan, R., Williams, J. R., Haney, E. B., & Neitsch, S. L. (2012). Soil & water assessment tool. Input/Output documentation version 2012. Retrieved from <http://swat.tamu.edu/media/69296/SWAT-IO-Documentation-2012.pdf>
- Baldi, G., Guerschman, J. P., & Paruelo, J. M. (2006). Characterizing fragmentation in temperate South America grasslands. *Agriculture, Ecosystems and Environment*, 116(3-4), 197-208. doi: 10.1016/j.agee.2006.02.009
- Blanco, P. D., Colditz, R. R., López, S. G., Hardtke, L. A., Llamas, R. M., Mari N. A., ...Arroyo, V. B. (2013). A land cover map of Latin America and the Caribbean in the framework of the SERENA project. *Remote Sensing of Environment*, 132, 13-31. doi: 10.1016/j.rse.2012.12.025
- Breuer, L., Huisman, J. A., Willems, P., Bormann, H., Bronstert, A., Croke, B. F. W., & Kite, G. (2009). Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: Model intercomparison with current land use. *Advances in Water Resources*, 32(2), 129-146. doi: 10.1016/j.advwatres.2008.10.003

de mapear los usos de mayor representatividad; sin embargo, tiene la desventaja de que el comportamiento anual de los usos a discriminar se debe conocer *a priori*, lo que podría evitarse tomando datos de campo en píxeles puros para obtener la firma fenológica.

La caracterización de las coberturas y usos del suelo representa uno de los insumos imprescindibles para el monitoreo y manejo de los recursos naturales a diferentes escalas, como por ejemplo en la modelación hidrológica semidistribuida (Arnold et al., 2012); sin embargo, en Argentina suele ser una de las informaciones ausentes en los organismos de gestión departamental o regional. Por lo tanto, cuando se requieren estrategias de manejo sustentable de los recursos naturales a escala local, los procedimientos existentes a gran escala (Blanco et al., 2013; Eva et al., 2004) resultan deficientes, y es entonces cuando la metodología desarrollada en este trabajo puede aportar herramientas útiles. Por otro lado, en regiones de gran dinamismo de usos del suelo, como la aquí analizada, la metodología permitiría evaluar y monitorear los cambios de coberturas y, en particular, de la frontera agropecuaria, prácticamente en tiempo real.

## Conclusiones

La metodología desarrollada en el trabajo permitió la caracterización de las coberturas y usos del suelo de la cuenca superior del arroyo del Azul, Argentina, mediante el uso de firmas fenológicas derivadas de series temporales de NDVI-MODIS. Las áreas mezcladas que no pudieron diferenciarse fueron categorizadas a partir de imágenes Landsat 8, para mejorar la precisión del mapa final. El método de clasificación empleado es una alternativa válida cuando no se dispone de datos precisos de campo, pero requiere conocimiento acerca del comportamiento estacional de las coberturas a ser mapeadas. De esta manera, la propuesta presentada puede ser aplicable y reproducible en diferentes regiones y escalas, convirtiéndose en una herramienta valiosa para el monitoreo de los cambios de usos del suelo a escala local y regional.

## Agradecimientos

Al CONICET (becas doctorales de Guevara-Ochoa y Lara) y a la Secretaría de Ciencia y Técnica de la UNCPBA por la financiación brindada para la realización de este trabajo.

*Fin de la versión en español*



- Cabello, J., & Paruelo, J. M. (2008). La teledetección en estudios ecológicos. *Ecosistemas*, 17(3), 1–3. Retrieved from <https://www.revistaecosistemas.net/index.php/ecosistemas/article/viewFile/77/74>
- Chuvieco, E. (1990). *Fundamentos de teledetección espacial*. Madrid, España: Ed. Rialp, S. A.
- Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1), 35–46. doi: 10.1016/0034-4257(91)90048-B
- DeFries, R., & Eshleman, K. N. (2004). Land-use change and hydrologic processes: a major focus for the future. *Hydrological Processes*, 18(11), 2183–2186. doi: 10.1002/hyp.5584
- Eitzinger, J., Orlandini, S., Stefanski, R., & Naylor, R. E. L. (2010). Climate change and agriculture: introductory editorial. *Journal of Agricultural Science*, 148(5), 499–500. doi: 10.1017/S0021859610000481
- Eva, H., Belward A., De Miranda E., Di Bella C., Gond V., Huber O., ...Fritz, S. (2004). A land cover map of South America. *Global Change Biology*, 10(5), 731–744. doi: 10.1111/j.1529-8817.2003.00774.x
- Fohrer, N., Haverkamp, S., Eckhardt, K., & Frede, H. G. (2001). Hydrologic response to land use changes on the catchment scale. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 26(7-8), 577–582. doi: 10.1016/S1464-1909(01)00052-1
- Guerschman, J. P., Paruelo, J. M., & Burke, I. (2003). Land use impacts on the normalized difference vegetation index in temperate Argentina. *Ecological Applications*, 13(3), 616–628. doi: 10.1890/1051-0761(2003)013[0616:liiotn]2.0.co;2
- Horning, N., Robinson, J., Sterling, E., Turner, W., & Spector, S. (2010). *Remote sensing for ecology and conservation: A handbook of techniques*. New York, USA: Oxford University.
- Lara, B., & Gandini, M. (2014). Quantifying the land cover changes and fragmentation patterns in the Argentina Pampas, in the last 37 years (1974-2011). *GeoFocus*, 14, 163–180. Retrieved from [http://geofocus.rediris.es/2014/Articulo9\\_2014.pdf](http://geofocus.rediris.es/2014/Articulo9_2014.pdf)
- Lara, B., & Gandini, M. (2016). Caracterización del comportamiento fenológico de las coberturas vegetales en un sector de la Pampa Deprimida (Argentina): una aproximación utilizando series temporales de NDVI. *Ambiência*, 12(4), 765–783. doi: 10.5935/ambiencia.2016.04.01
- Li, Z., Liu, W. Z., Zhang, X. C., & Zheng, F. L. (2009). Impacts of land use change and climate variability on hydrology in an agricultural catchment on the Loess Plateau of China. *Journal of Hydrology*, 377(1), 35–42. doi: 10.1016/j.jhydrol.2009.08.007
- Manuel-Navarrete, D., Gallopin, G. C., Blanco, M., Díaz-Zorita, M., Ferraro, D. O., Herzer, H., ...Viglizzo, E. F. (2009). Multicausal and integrated assessment of sustainability: The case of agriculturization in the Argentine Pampas. *Environment, Development and Sustainability*, 11(3), 621–638. doi: 10.1007/s10668-007-9133-0
- Matteucci, S. (2012). Ecorregión Pampa. In J. Morello, S. Matteucci, A. Rodríguez, & M. Silva (Eds.), *Ecorregiones y complejos ecosistémicos argentinos* (pp. 391–446). Buenos Aires, Argentina: Orientación Gráfica Editora. Retrieved from [https://www.researchgate.net/profile/Silvia\\_Matteucci/publication/268447997\\_Ecorregion\\_Pampa/links/54f608650cf2ca5efefdc9.pdf](https://www.researchgate.net/profile/Silvia_Matteucci/publication/268447997_Ecorregion_Pampa/links/54f608650cf2ca5efefdc9.pdf)
- Müller, D., & Munroe, D. K. (2014). Current and future challenges in land-use science. *Journal of Land Use Science*, 9(2), 133–142. doi: 10.1080/1747423X.2014.883731
- Murty, D., Kirschbaum, M. U., Mcmurtrie, R. E., & Mcgilvray, H. (2002). Does conversion of forest to agricultural land change soil carbon and nitrogen? A review of the literature. *Global Change Biology*, 8(2), 105–123. doi: 10.1046/j.1354-1013.2001.00459.x
- Paruelo, J. M. (2008). La caracterización funcional de ecosistemas mediante sensores remotos. *Ecosistemas*, 17(3), 4–22. Retrieved from <http://revistaecosistemas.net/index.php/ecosistemas/article/view/83>
- Pettorelli, N., Wegmann, M., Skidmore, A., Sander, M., Dawson, T. P., Fernandez, M., ...Penner, J. (2016). Framing the concept of satellite remote sensing essential biodiversity variables: challenges and future directions. *Remote Sensing in Ecology and Conservation*, 2(3), 122–131. doi: 10.1002/rse2.15
- Post, W. M., & Kwon, K. C. (2000). Soil carbon sequestration and land-use change: processes and potential. *Global Change Biology*, 6(3), 317–327. doi: 10.1046/j.1365-2486.2000.00308.x
- Reed, B. C., Schwartz, M. D., & Xiangming X. (2009). Remote sensing phenology: Status and the way forward. In A. Noormets (Ed.), *Phenology of ecosystem processes* (pp. 231–246). New York, USA: Springer. doi: 10.1007/978-1-4419-0026-5\_10
- Sistema Integrado de Información Agropecuaria (SIIA). (2016). Retrieved June 19, 2017, from <http://www.agroindustria.gob.ar/datosagroindustriales/>
- Solano, R., Didan, K., Jacobson A., & Huete, A. (2010). MODIS Vegetation Index User's Guide (MOD13 Series) C5. Retrieved from [https://vip.arizona.edu/documents/MODIS/MODIS\\_VI\\_UsersGuide\\_01\\_2012.pdf](https://vip.arizona.edu/documents/MODIS/MODIS_VI_UsersGuide_01_2012.pdf)
- Tou, J. T., & Gonzalez, R. C. (1974). *Pattern recognition principles*. Massachusetts, USA: Addison-Wesley Publishing Company.
- Viglizzo, E. F., Roberto, Z. E., Lértora, F., Gay, E. L., & Bernardos, J. (1997). Climate and land-use change in field-crop ecosystems of Argentina. *Agriculture, Ecosystems & Environment*, 66(1), 61–70. doi: 10.1016/S0167-8809(97)00079-0
- Wang, Q., Tenhunen, J., Dinh, N. Q., Reichstein, M., Vesala, T., & Keronen, P. (2004). Similarities in ground- and satellite-based NDVI time series and their relationship to physiological activity of a Scots pine forest in Finland. *Remote Sensing of Environment*, 93(1-2), 225–237. doi: 10.1016/j.rse.2004.07.006