



Teledetección y redes neuronales aplicadas al mapeo de coberturas del suelo de la cuenca del Matanza-Riachuelo, Buenos Aires, Argentina

Author(s): Rodrigo Martín Becerra Serial, Daniela Czibener and Paulina Esther Nabel

Source: *Revista Geográfica*, No. 146 (JULIO-DICIEMBRE 2009), pp. 125-152

Published by: [Pan American Institute of Geography and History](#)

Stable URL: <http://www.jstor.org/stable/40996845>

Accessed: 11-12-2015 19:29 UTC

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Pan American Institute of Geography and History is collaborating with JSTOR to digitize, preserve and extend access to *Revista Geográfica*.

<http://www.jstor.org>

Teledetección y redes neuronales aplicadas al mapeo de coberturas del suelo de la cuenca del Matanza-Riachuelo, Buenos Aires, Argentina

Rodrigo Martín Becerra Serial*
Daniela Czibener*
Paulina Esther Nabel*

Abstract

The Matanza-Riachuelo river basin is a particularly complex area in terms of both its current environmental situation and the diversity of land use and of materials covering its surface. By analyzing satellite images, it is possible to produce up-to-date base cartography, essential for the environmental management of the basin.

We used a multitemporal series of Landsat 5 (TM) images for this work, which correspond to an annual cycle, as well as a set of derived bands, including NDVI, brightness, greenness, and humidity (tasseled cap transformation), and bands of scatter measurements of data in the annual cycle (e.g. variance), amongst others. After the bands had been selected, two neural networks (multilayer perceptrons) were trained using the Backpropagation algorithm. The first of these was used to classify the image into three main groups of ground cover: water, waterproofed, and vegetation. The second network was used to classify the vegetation into subclasses: evergreen forest, deciduous forest, grass, winter plant association, and summer plant association.

The overall exactitude of the resulting map was 95.13%. The most common types of ground cover were grasses and waterproofed surfaces. The latter are mainly associated with urbanized zones and take up almost a quarter of the river basin area. Two types of vegetation related with agricultural activity are also found in large proportions. Along the urban-rural axis, the diversity of ground cover changes: there is minimal diversity (maximum uniformity) in urban areas, maxi-

* Grupo de Investigaciones Geoambientales (GIGA), Museo Argentino de Ciencias Naturales "Bernardino Rivadavia", Av. Ángel Gallardo 470, 4° piso, C1405DJR, Ciudad Autónoma de Buenos Aires, Argentina, correo electrónico: rodbece@gmail.com

imum heterogeneity in periurban areas, and low values in rural areas, without reaching the level of uniformity shown in the city.

The results of this work encourage the use of satellite images and non-parametric classifiers like neural networks to produce cartography in complex environments such as that of the river basin in question. This cartography could be useful as a tool for the environmental management of said basin.

Key words: *The Matanza-Riachuelo River Basin, Satellite Images, Neural Networks, Types of Vegetation.*

Resumen

La cuenca del Matanza-Riachuelo es un área particularmente compleja tanto por su situación ambiental actual como por la diversidad de usos del suelo y de materiales que cubren la superficie. Por medio de análisis de imágenes satelitales es posible producir cartografía de base actualizada, imprescindible para la gestión ambiental de la cuenca.

En este trabajo, usamos una serie multitemporal de imágenes Landsat 5 (TM) correspondientes a un ciclo anual, más un conjunto de bandas derivadas, entre ellas NDVI, brillo, verdor y humedad (transformación tasseled cap), y bandas de medidas de dispersión de los datos en el ciclo anual (ej.: varianza), entre otras. Luego de una selección de bandas, se entrenaron dos redes neuronales (perceptrones multicapa) mediante el algoritmo Backpropagation. La primera de ellas sirvió para clasificar la imagen en tres grandes grupos de cobertura del suelo: agua, impermeabilizado y vegetación. La segunda red se usó para clasificar a la vegetación en subclases: bosque de perennifolias, bosque de caducifolias, pasto, asociación vegetal invernal y asociación vegetal estival.

El mapa resultante mostró una exactitud general del 95.13%. Las clases de coberturas más comunes son los pastos y las superficies impermeabilizadas. Éstas últimas se asocian principalmente con zonas urbanizadas y ocupan casi un cuarto del área de la cuenca. Dos coberturas vegetales relacionadas con usos agropecuarios también se encuentran en importante proporción. En el eje urbano-rural, la diversidad de coberturas del suelo cambia: hay mínima diversidad (máxima uniformidad) en áreas urbanas, máxima heterogeneidad en el periurbano, y valores bajos en las zonas rurales sin llegar al nivel de uniformidad de la ciudad.

Los resultados de este trabajo alientan el uso de imágenes satelitales y de clasificadores no paramétricos como las redes neuronales para producir cartografía en ambientes complejos como el de la cuenca estudiada. Dicha cartografía podría ser útil como herramienta para la gestión ambiental de la misma.

Palabras clave: *Cuenca del Matanza-Riachuelo, imágenes satelitales, redes neuronales, coberturas vegetales.*

Introducción

La cuenca del río Matanza-Riachuelo, ubicada al NE de la provincia de Buenos Aires, abarca tanto zonas del Área Metropolitana de Buenos Aires como sectores rurales. Cobija alrededor de 4 millones de habitantes, sobre una superficie de más de 2,000km². La presión por contaminación en la cuenca ha generado un importante deterioro del río principalmente en su porción baja; lo cual a su vez ha determinado que su saneamiento y manejo sustentable sean objetivos prioritarios de la agenda ambiental nacional, provincial y de los municipios involucrados. En este contexto, la disponibilidad de cartografía actualizada es un aspecto fundamental para la evaluación de la situación ambiental de la cuenca y para su gestión ambiental.

La teledetección es una disciplina que incluye el procesamiento y análisis de imágenes digitales, captadas mayormente desde plataformas satelitales. Su finalidad es la obtención de información de la superficie terrestre, usando la radiación electromagnética que ésta refleja o emite (de Jong *et al.*, 2004).

En la aproximación cuantitativa al análisis de imágenes satelitales, la extracción de información se base en lo siguiente: los valores de brillo de los píxeles indican la intensidad de reflexión o emisión de radiación electromagnética, en el momento de toma de la imagen. Además, cada material de la superficie refleja o emite radiación de una manera particular y diferente del resto de las coberturas (Richards y Jia, 2006). El reconocimiento del patrón de respuesta espectral (o firma espectral) de algunos píxeles en la imagen y su asociación con algún tipo de cobertura (ej.: agua, vegetación, etc.), permite luego la generalización al resto de los píxeles de la imagen, lo que implica el agrupamiento e identificación de clases de píxeles que representan tipos de coberturas particulares. El proceso de discriminación de grupos de píxeles según su respuesta espectral y la identificación de cada uno como perteneciente a un tipo de cobertura se conoce como clasificación (Richards y Jia, 2006).

Uno de los aspectos clave para lograr exactitud en el mapeo de coberturas del suelo, es la correcta elección del tipo de datos satelitales (Lu y Weng, 2007). Cuando el objetivo del análisis es la elaboración de mapas de escala regional, lo más apropiado es el uso de imágenes de resolución espacial media, como las obtenidas por el satélite Landsat 5 (píxel=30m) (Lu y Weng, 2007). Actualmente una de las formas más efectivas de mejorar la exactitud de una clasificación es la incorporación de diferentes tipos de datos al análisis. Por ejemplo, los datos espectrales (intensidad de la radiación captada por el sensor en el satélite) pueden complementarse con datos de variabilidad temporal de la respuesta espectral. La incorporación de la variación temporal se realiza mediante el análisis simultáneo de imágenes de distintas fechas (ej.: estaciones del año). Esto es particularmente útil cuando las coberturas analizadas son vegetales, ya que las plantas modifican de manera relativamente predecible su respuesta espectral dependiendo del estado fenológico en el

que se encuentren (brote, floración, senescencia, etc.). Se ha demostrado que el uso de imágenes multitemporales en una clasificación genera una mayor exactitud que la obtenida con imágenes de una sola fecha (Lunetta y Balogh, 1999; Guerschman *et al.*, 2003).

Otro aspecto fundamental que afecta la exactitud de una clasificación es el método de clasificación usado. En regiones con gran heterogeneidad de coberturas (como la cuenca del Matanza-Riachuelo) el uso de clasificadores paramétricos (por ejemplo ISODATA o *maximun likelihood*) podría resultar inadecuado ya que comúnmente el supuesto de distribución normal de los valores espectrales no se cumple (Lu y Weng, 2007). Asimismo, otra desventaja de los clasificadores paramétricos es la dificultad que implica la integración de datos espectrales y datos auxiliares provenientes por ejemplo de mapas temáticos. Frente a estos, los clasificadores no paramétricos como las redes neuronales o los árboles de decisión son preferibles, ya que al no utilizar parámetros estadísticos para separar clases espectrales, no requieren que se cumpla el supuesto de normalidad de la muestra para entrenamiento del clasificador (Lu y Weng, 2007). Los clasificadores no paramétricos son especialmente útiles cuando se quiere incorporar datos no espectrales al proceso de clasificación y está demostrado que, en paisajes complejos, pueden generar mejores resultados (Paola y Schowengerdt, 1995; Foody, 2002). Específicamente, el uso de redes neuronales, además de su naturaleza no paramétrica, posibilita una fácil inclusión de diferentes tipos de datos y además, los valores resultantes son difusos, esto es, valores numéricos relacionados con la probabilidad de pertenencia a las distintas clases, lo cual permite establecer los límites entre las mismas según el ajuste logrado con distintos valores de corte (Paola y Schowengerdt, 1995).

Redes neuronales y teledetección

Las Redes Neuronales Artificiales (RNA) o simplemente Redes Neuronales (RN) son un paradigma computacional alternativo al modelo procedural. Está inspirado en la neurociencia, aunque no trata de ser biológicamente realista en detalle (Hertz, 1991).

Las RN son capaces de resolver problemas de clasificación y reconocimiento de patrones, por su capacidad de adquirir conocimiento a través de la experiencia en aplicaciones donde no es sencillo contar con un modelo que pueda ser programado. La información original consiste en una cantidad de muestras previamente seleccionadas y clasificadas por un experto. La flexibilidad inherente a las redes neuronales, posibilita descubrir similitudes entre patrones, que muchas veces escapan a la observación humana.

Las redes neuronales están compuestas por neuronas, que son las unidades básicas de procesamiento, y por conexiones entre las mismas. McCulloch y Pitts

(1943), propusieron un modelo simple de neurona, que consiste en calcular la suma pesada de las entradas de la unidad de procesamiento, dando como salida 0 ó 1, en función de que el resultado de la suma sea mayor o menor a un cierto umbral. En la Figura 1 se muestran las neuronas de entrada representadas por los valores X_1, X_2, \dots, X_n , los pesos de las conexiones de cada una de estas neuronas W_1, W_2, \dots, W_n que alimentan a la neurona Y , la cual luego de realizado el cálculo genera la salida A .

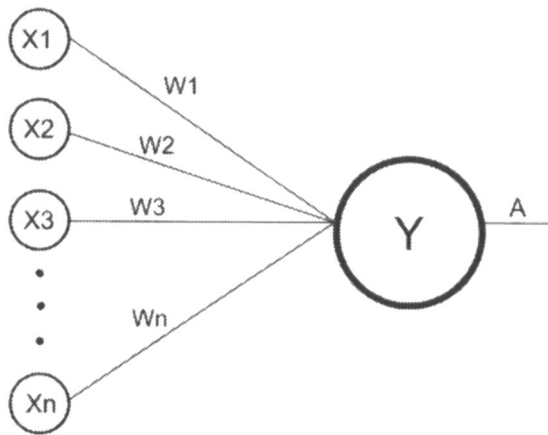


Figura 1. Esquema simple de neurona de McCulloch y Pitts (1943).

El cálculo que se realiza en Y es la suma pesada de las entradas:

$$F = \left(\sum_{i=1}^n (X_i * W_i) \right) = A$$

La función $F(x)$ para este caso es una función escalón, que determina si el valor que puede tomar A es 0 ó 1, dependiendo del resultado obtenido. Los pesos de las conexiones sinápticas constituyen la fuerza entre pares de neuronas. Estos pesos, de no ser nulos (iguales a 0), representan una sinapsis excitatoria o inhibitoria, de lo contrario no hay sinapsis. Una generalización del cálculo de McCulloch y Pitts, consiste en reemplazar la función escalón por una función de activación o transferencia (Russell y Norvig, 1996).

El éxito de las redes neuronales reside en la capacidad que poseen, luego de ser entrenadas con una cantidad de ejemplos, de reconocer situaciones nuevas en base a interpolar y extrapolar lo aprendido (Hertz, 1991). El aprendizaje está basado en el ajuste sucesivo de los pesos de las conexiones, que puede realizarse de dos modos:

- *Supervisado*: el aprendizaje esta dado por comparación directa entre la salida de la red y la respuesta correcta, y la posterior corrección de los pesos cuando la salida es incorrecta.
- *No supervisado*: la única información disponible está en la correlación de los datos o señales de entrada. Se espera entonces que la red “descubra” categorías de estas correlaciones, y que produzca señales de salida correspondientes con las categorías de las entradas.

Para entrenar una red, no hace falta conocer cada detalle del cálculo, simplemente se tiene que entrenar un conjunto representativo de los datos. Esto puede ayudar a resolver problemas cuyas reglas serían muy complicadas de conocer de antemano (Hertz, 1991).

Para este trabajo se usó el modelo de red de perceptrón, que se basa en el aprendizaje supervisado. Inicialmente, se debe contar con los datos de entrada y sus correspondientes salidas, conjunto denominado de entrenamiento. Para clasificar cada píxel los datos de entrada son los valores numéricos de las bandas incluidas en el análisis y la salida es el tipo de cobertura del suelo al que corresponde en la realidad. Definido dicho conjunto, la red es entrenada comparando la salida generada por el cálculo con la respuesta correcta. La diferencia surgida entre el valor esperado y el calculado, se propaga corrigiendo el peso de las conexiones. A esto se le llama propagación del error hacia atrás.

El entrenamiento comienza asignando pesos aleatorios a las conexiones. Luego se realizan varias etapas o ciclos presentando de a uno el patrón de entrada, calculando la salida y calculado el error entre la salida esperada y la calculada. Luego se actualizan los pesos de las conexiones, para minimizar la diferencia con la salida correcta. Si los pesos convergen, el entrenamiento es exitoso, con lo cual se puede decir que el conjunto de entrenamiento fue aprendido y la red puede responder a las entradas de estos patrones con un error menor al umbral establecido inicialmente. Para poder determinar la capacidad de generalización de la red en base a lo aprendido, se la somete a una prueba con datos que no estuvieron presentes en la etapa de entrenamiento.

En el modelo de red neuronal de perceptrón existe un conjunto de neuronas de entrada (capa de entrada) cuya única función es distribuir los valores del conjunto de entrenamiento en el resto de la red. En el caso de los perceptrones multicapa (Multilayer Perceptron) puede haber uno o más niveles de neuronas intermedias (capas ocultas), seguidos de la capa de salida, que contiene el resultado del cálculo (Figura 2). Las unidades de las capas intermedias son comúnmente llamadas unidades ocultas. Los perceptrones que no poseen unidades intermedias son llamados Perceptrones Simples.

Convencionalmente, las redes neuronales aplicadas en teledetección son definidas con una capa de nodos de entrada, los cuales distribuyen las entradas en el nivel siguiente para su procesamiento, y una capa de salida la cual provee la información del etiquetado de clases. Entre las capas de entrada y salida puede haber una o más capas ocultas (capas de nodos de procesamiento), aunque normalmente el uso de una sola capa oculta es suficiente (Richards y Jia, 2006).

Objetivo

Este trabajo tiene como objetivo la producción de un mapa de coberturas del suelo en la cuenca del río Matanza-Riachuelo (Argentina), mediante la clasificación de imágenes satelitales. Se espera que el producto pueda utilizarse en la evaluación de la situación ambiental del área y en la elaboración de planes de manejo y saneamiento. Además, este estudio propone evaluar la eficacia del uso de redes neuronales en la clasificación espectral de imágenes en áreas ambientalmente heterogéneas, incorporando información satelital multitemporal.

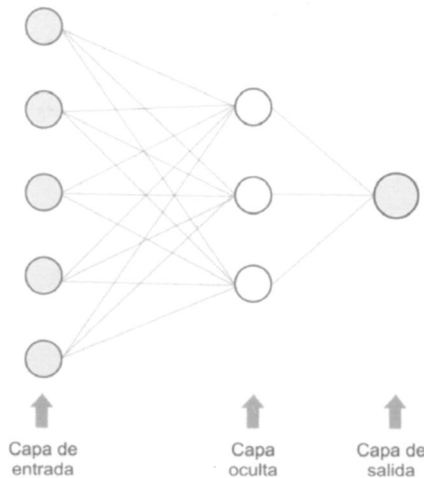


Figura 2. Modelo de red neuronal Perceptrón Multicapa.

Área de estudio

La cuenca del Matanza-Riachuelo se encuentra ubicada al NE de la provincia de Buenos Aires, Argentina (Figura 3). La porción baja de la misma abarca sectores urbanizados del Área Metropolitana de Buenos Aires (AMBA) y la cuenca alta

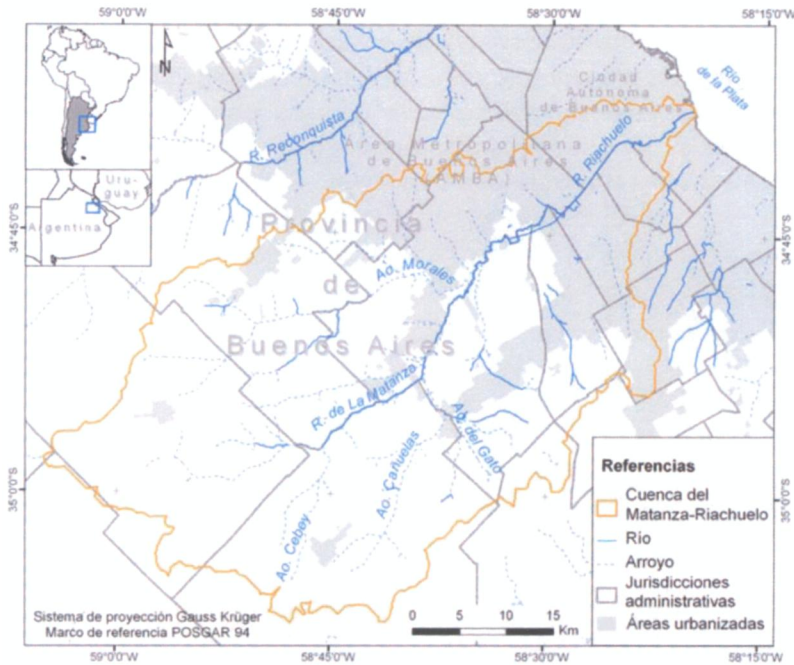


Figura 3. Ubicación del área de estudio.

comprende zonas rurales. La porción media incluye a la zona de interfase entre el paisaje urbano y el rural, denominada periurbano (Mörtberg *et al.*, 2007).

El Matanza-Riachuelo es un río de llanura caracterizado por su baja pendiente y su recorrido meandroso. Sus aguas escurren principalmente en dirección SO-NE. La topografía llana predominante determina que en caso de lluvias abundantes, sus aguas desborden ocasionando inundaciones en su llanura aluvial. Comúnmente se lo denomina Riachuelo en su porción más baja, cuando atraviesa áreas urbanas, y Matanza en su tramo medio y alto, desde el periurbano y hacia su cabecera. Entre sus tributarios más importantes se pueden mencionar a los arroyos Morales y Cañuelas.

El área de la cuenca incluye a una gran variedad de tipos de coberturas, lo cual genera respuestas espectrales variadas y complejas que aumentan el grado de dificultad en la producción de mapas a partir de imágenes satelitales. Esta variabilidad de tipos de coberturas es consecuencia de la existencia de múltiples usos del suelo en las porciones baja, media y alta. La parte inferior de la cuenca, con usos principalmente residenciales e industriales, presenta coberturas impermeables de materiales como cemento, asfalto, chapa y tejas; aunque no es despreciable la cobertura de vegetación del arbolado urbano, parques y jardines. En el otro extremo los sectores

altos de la cuenca, presentan un uso predominante agrícola y pecuario, y consecuentemente el dominio de coberturas vegetales. Sin embargo, dada la cercanía con la ciudad, éstas se ven frecuentemente interrumpidas por la presencia de materiales impermeables propios de las zonas urbanizadas. Por último, el periurbano, como franja de transición, se caracteriza por la gran heterogeneidad de usos de suelo, entre ellos residencial, industrial, áreas recreativas y terrenos baldíos. En estos últimos se desarrollan comunidades vegetales seminaturales, en muchos casos dominadas por plantas leñosas exóticas como la acacia negra (*Gleditsia triacanthos*) (Morello y Mateucci, 1999).

Métodos

Delimitación del área de estudio

El área de la cuenca del Río Matanza-Riachuelo fue delimitada usando información topográfica de una base vectorial de curvas de nivel de 1.25m de equidistancia, previamente digitalizada a partir de cartas topográficas 1:50,000 del IGM (Instituto Geográfico Militar) y publicada en el Atlas Ambiental de Buenos Aires <www.atlasdebuenosaires.gov.ar>.

Los valores de cota de las curvas se usaron para generar un Modelo Digital de Elevación (MDE) del tipo grilla regular con un tamaño del píxel de 30m, mediante un método de interpolación global (tendencia polinómica) para generar los valores altimétricos intermedios, recomendado como método de interpolación para análisis regionales (Borough, 1986). Para esta tarea se usó la herramienta “topo to raster” de ArcGIS (ESRI). Luego, se delimitó el área de la cuenca identificando las líneas divisorias de agua, utilizando la herramienta “basin” del mismo programa.

Mapeo de coberturas del suelo mediante clasificación de imágenes satelitales

Se generó un mapa de coberturas del suelo luego del procesamiento y análisis de información satelital multitemporal, contenida en 4 imágenes Landsat 5 (TM) (path: 225 y row:84). Las imágenes seleccionadas tuvieron una cobertura de nubes menor al 5% y corresponden a fechas representativas de las cuatro estaciones del año: 26 de octubre de 2006: primavera, 14 de enero de 2007: verano, 4 de abril de 2007: otoño y 23 de junio de 2007: invierno.

Los procedimientos que se llevaron a cabo sobre las imágenes para generar el mapa de coberturas de suelo, fueron los siguientes:

- a) Preprocesamiento: correcciones radiométricas y geométricas
- b) Definición de un modelo conceptual de coberturas a identificar
- c) Selección de muestras de entrenamiento, correspondientes a clases de coberturas definidas en el modelo conceptual

- d) Generación de nuevas variables (bandas) y selección del conjunto a utilizar en la clasificación
- e) Clasificación de las imágenes por medio de redes neuronales, usando las bandas seleccionadas
- f) Cuantificación del error de la clasificación.

Estos procedimientos se describen con mayor detalle en las secciones siguientes.

Preprocesamiento de las imágenes

A fin de reducir el tamaño de los archivos y acortar tiempos de procesamiento, la escena original se recortó a un área menor que abarcara la superficie de la cuenca. Además, las bandas térmicas de las cuatro imágenes fueron eliminadas ya que el tamaño del píxel de las mismas es demasiado grande (120m) como para aportar información valiosa para identificar unidades pequeñas.

Las 6 bandas multispectrales de las 4 imágenes recortadas fueron calibradas radiométricamente, llevando los valores de conteo (valores crudos denominados también *digital numbers*) a radiancia (Richards y Jia, 2006) y corregidas atmosféricamente convirtiendo los valores digitales a reflectancia en superficie utilizando los datos de ganancia y sesgo (provistos en el archivo *header* de cada imagen) de acuerdo al modelo propuesto por Stumpf (1992). De esta manera se calculó la reflectancia en superficie removiendo el efecto de dispersión molecular (Rayleigh), producto de la interacción de la radiación electromagnética con los componentes moleculares de la atmósfera.

La corrección geométrica se realizó georreferenciando mediante el corregistro imagen a imagen con base en otra imagen georreferenciada, de la base de datos geográfica del Atlas Ambiental de Buenos Aires. Para este procedimiento se usaron 25 puntos de control en cada imagen, obteniéndose un error cuadrático medio menor a un píxel en todos los casos. El método de interpolación usado fue polinomial de primer orden. El método de reasignación de valores numéricos a los píxeles (remuestreo) fue el de vecino más cercano, a fin de preservar la información original. El corregistro final píxel a píxel se realizó mediante la superposición de capas (*layer stacking*) de las 4 imágenes (6 bandas cada una). Este procedimiento generó una imagen multitemporal de 24 bandas (6 por estación).

Definición de un modelo conceptual de coberturas del suelo

Se estableció un modelo conceptual general con tres clases de coberturas: a) agua, b) impermeabilizado y c) vegetación. La clase agua estaba representada en el recorte de la escena mayormente en el Río de la Plata, pero también en lagos artificiales (embalse Roggero y lagos de Lugano en la ciudad de Buenos Aires) y en el

Riachuelo en su porción inferior cerca de la desembocadura. La clase impermeabilizado correspondía a píxeles heterogéneos pero mayormente ocupando zonas urbanizadas y conteniendo materiales impermeables como asfalto, cemento, tejas, chapa y otros. La clase vegetación estuvo representada por píxeles que contaran en algún momento del año con cobertura vegetal, pudiendo presentar suelo desnudo en alguna época. En la clase vegetación se incluyen varias subclases.

Las “subclases” conceptuales de coberturas vegetales, que representan subconjuntos más homogéneos, fueron: bosque de perennifolias, bosque de caducifolias, pastos, asociación vegetal invernal y asociación vegetal estival. En este modelo conceptual las coberturas a mapear se presentan entonces en dos niveles jerárquicos: uno superior, más general, de las clases agua, impermeabilizado y vegetación, y otro inferior, más detallado, con las subclases de vegetación.

Generación de las muestras de entrenamiento

Las muestras de píxeles usadas para el entrenamiento del clasificador (red neuronal) se generaron por distintas vías según el tipo de cobertura. Para las clases agua e impermeabilizado las muestras de entrenamiento fueron obtenidas directamente en la imagen, teniendo en cuenta la facilidad para identificar píxeles correspondientes a estas coberturas. Los píxeles se definieron usando 9 polígonos para la clase agua y 6 para la clase urbano, distribuidos en la imagen de manera homogénea. En total se contó con un conjunto de alrededor de 800 píxeles para cada una de estas dos clases, la mitad de los cuales se destinaron al entrenamiento de la red neuronal y la otra mitad a la evaluación de la exactitud de la clasificación (punto f).

Considerando que la clase vegetación está formada por varias subclases de difícil identificación por simple observación de la imagen, fue necesario contar con información de campo para establecer la distribución espacial de puntos correspondientes a estas coberturas. A fin de generar esta información, se realizaron relevamientos de campo en el área de estudio en los meses de junio y julio de 2008, en los que se localizaron puntos conteniendo estas coberturas vegetales que fueron georreferenciados con navegador satelital (GPS). En estas salidas se registraron tres tipos de coberturas: bosques de caducifolias, bosques de perennifolias y pastos, que fueron consideradas constantes en el periodo entre la fecha de toma de las imágenes y la de toma de las muestras a campo.

Los puntos de GPS fueron superpuestos sobre la imagen, generándose luego polígonos que abarcaran superficies de reflectancia homogénea alrededor de dichos puntos. Además de estos puntos tomados en el campo, se contó con un mapa de vegetación del predio de la Comisión Nacional de Energía Atómica en Ezeiza (Ana Faggi, datos no publicados), cuya información sirvió para ubicar píxeles correspondientes a los distintos subtipos de vegetación. Tanto los polígonos generados a partir de puntos de GPS como los surgidos de este mapa de vegetación fueron su-

perpuestos sobre la imagen multitemporal para seleccionar las muestras de píxeles para entrenamiento de la red neuronal, o bien para validación de la clasificación.

Las muestras de entrenamiento de las clases asociación vegetal invernal y asociación vegetal estival (definidas *a priori* en el modelo conceptual), fueron generadas analizando los patrones de variación temporal del índice verde normalizado (NDVI) (Tucker, 1979). El valor del NDVI es indicador de la presencia de tejidos verdes, actividad fotosintética y productividad vegetal (Guerschman *et al.*, 2003, Richards y Jia, 2006). Los colores resultantes desplegando la combinación de bandas NDVI de primavera, verano y otoño permitieron establecer visualmente la presencia de dos grandes grupos de vegetación ocupando principalmente las parcelas agrícolas. En uno de ellos el NDVI presentó un valor máximo en verano, mientras que en el otro hubo dos picos de NDVI, en primavera y otoño, respectivamente. De esta manera las muestras de entrenamiento para estas subclases se tomaron de las parcelas en la imagen pertenecientes a alguno de los dos grupos.

La Figura 4 muestra la diferencia en el comportamiento del NDVI de las subclases asociación vegetal estival y asociación vegetal invernal y el de la otra subclase de vegetación herbácea: los pastos. La marcha anual del NDVI en cada subclase es producto de los patrones de cobertura vegetal a lo largo del año y de las especies que dominan en cada subclase de cobertura.

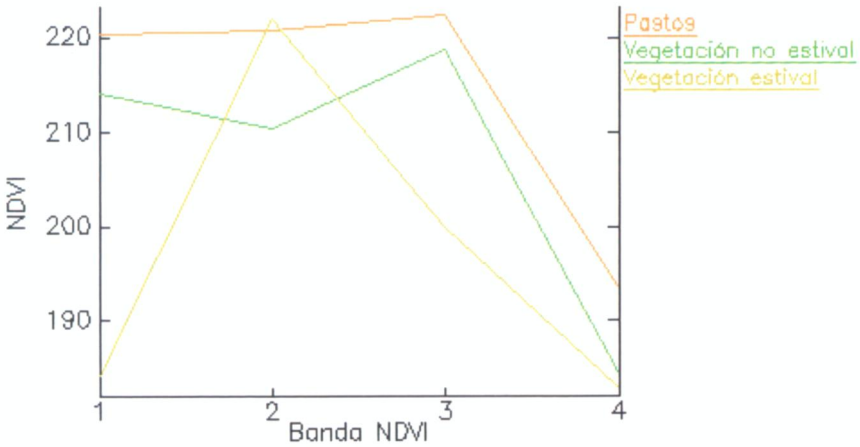


Figura 4. Variación estacional del promedio del índice verde normalizado (NDVI) para las muestras de las clases de vegetación herbácea: pastos, asociación vegetal estival y asociación vegetal invernal. Las bandas de NDVI 1, 2, 3 y 4 corresponden a primavera (octubre de 2006), verano (enero de 2007), otoño (abril de 2007) e invierno (junio de 2007), respectivamente. Los valores de NDVI se encuentran reescalados a 8 bits (0-255).

Generación y selección de bandas para la clasificación espectral

En primer lugar, a partir de las 24 bandas de la imagen multitemporal se generaron las siguientes bandas adicionales:

- NDVI para cada estación (4 nuevas bandas)
- Desviación estándar (en el eje temporal): para cada píxel se calculó el desvío estándar de las 6 bandas multispectrales en el eje temporal (a cada banda le corresponden 4 valores, uno para cada estación). Este procedimiento generó 6 nuevas bandas
- Promedio y varianza del NDVI, usando las cuatro bandas del periodo anual considerado (resultado: 2 nuevas bandas)
- Brillo, Verdor y Humedad de la transformación Tasseled Cap (Kauth y Thomas, 1976), 3 bandas para cada estación, en total 12 nuevas bandas
- Además se calcularon bandas de desvío estándar en el tiempo del Brillo, Verdor y Humedad (3 bandas).

Un resumen de las bandas consideradas en el análisis (en total 51) puede verse en la Tabla 1.

Tabla 1
Bandas totales consideradas para el análisis

	<i>Primavera</i>	<i>Verano</i>	<i>Otoño</i>	<i>Invierno</i>	
Bandas TM originales	B1-P	B1-V	B1-O	B1-I	DesvEst B1
	B2-P	B2-V	B2-O	B2-I	DesvEst B2
	B3-P	B3-V	B3-O	B3-I	DesvEst B3
	B4-P	B4-V	B4-O	B4-I	DesvEst B4
	B5-P	B5-V	B5-O	B5-I	DesvEst B5
	B7-P	B7-V	B7-O	B7-I	DesvEst B7
Índice verde	NDVI - P	NDVI - V	NDVI - O	NDVI - I	Prom NDVI y Var NDVI
Transformación tasseled cap	Brillo - P	Brillo - V	Brillo - O	Brillo - I	DesvEst Brillo
	Verdor - P	Verdor - V	Verdor - O	Verdor - I	DesvEst Verdor
	Humedad - P	Humedad - V	Humedad - O	Humedad - I	DesvEst Humedad

Abreviaturas: B: banda, P: primavera, V: verano, O: otoño, I: invierno, DesvEst: desviación estándar, Prom: promedio, Var: varianza.

De acuerdo con el modelo conceptual definido cuyas clases de coberturas se presentan en dos niveles jerárquicos diferentes, la clasificación se realizó asimismo en dos etapas. En cada una de ellas se usó un conjunto particular de bandas, eliminando las que menos contribuyeran a la separación de las clases de coberturas, representadas por las muestras de entrenamiento. El aporte de las bandas a la exactitud de la clasificación se midió a través del índice de Jeffries-Matussita (Richards y Jia, 2006). Este índice mide la separabilidad de dos muestras usando un determi-

nado conjunto de bandas. A medida que aumenta el número de bandas consideradas el valor del índice también aumenta indicando que las muestras son más separables. Sin embargo, a partir de cierto punto la adición de más bandas al conjunto no hará aumentar sustancialmente la separabilidad de las muestras, con lo que es preferible no incluirlas ya que la información que aportan es presumiblemente redundante y su único efecto es hacer más lento el proceso de clasificación.

Para definir el conjunto de bandas óptimo a utilizar en cada etapa de la clasificación, se estableció primero un tamaño óptimo del conjunto, siendo el tamaño el número de bandas consideradas. Luego se determinó, para ese tamaño, la combinación de bandas que proporcionara una separabilidad espectral máxima de las muestras. A esta combinación la denominamos combinación óptima. Así, el *conjunto óptimo* fue aquel en el que hubiera un tamaño óptimo y una combinación óptima.

A fin de determinar el *conjunto óptimo* para cada tamaño de conjunto y para cada combinación específica de un tamaño dado, se calculó el índice de Jeffries-Matusita entre todos los pares de muestras de entrenamiento. Luego, con los valores del índice se calculó el promedio y el mínimo. Ambos valores fueron tenidos en cuenta para evaluar la eficacia de un conjunto de bandas en separar las muestras; el promedio como medida de la efectividad general del conjunto y el mínimo como indicador de la ausencia de pares de muestras poco separables.

Tanto para el promedio como para el mínimo de índices un valor elevado indica mayor separabilidad de las muestras de entrenamiento. De acuerdo con esto, para cada conjunto de tamaño n la mejor combinación fue aquella para la que ambas medidas eran simultáneamente las más elevadas. Esto se midió calculando el promedio entre las dos, al cual denominamos promedio compuesto (PComp).

El valor n (= número de bandas) usado para la clasificación se determinó en cada caso evaluando si la ganancia en separabilidad al incluir una nueva banda justificaba el incremento en cantidad de datos a analizar y por ende el tiempo de procesamiento.

Antes de definir el conjunto de bandas que sería el de entrada en cada red se realizaron preselecciones entre conjuntos de bandas de información semejante (ej.: las 24 originales, las 9 de desviación estándar en el eje temporal, etc.). Las bandas de NDVI de cada estación y su promedio y varianza no fueron sometidas a preselección sino que se integraron directamente en el conjunto de entrada para entrenamiento de las redes, por considerar que podían aportar información importante para separar clases espectrales. En última instancia la selección de las bandas de entrada final se realizó entre 24 bandas preseleccionadas. El mismo procedimiento se realizó para ambas redes.

Para el mapa de coberturas generales de la cuenca (etapa 1), las bandas seleccionadas fueron 6 (tamaño óptimo=6): la banda 4 de octubre de 2006, la banda 7 de

enero de 2007, la banda promedio de NDVI, la banda varianza de NDVI, la banda brillo de la transformación tasseled cap de la imagen de enero de 2007, y la banda verdor de la transformación tasseled cap de la imagen de abril de 2007. Para la etapa 2 en la que se generó un mapa de coberturas vegetales, las bandas usadas fueron la banda 5 de enero de 2007, la banda 4 de abril de 2007, la banda 5 de junio de 2007, la banda NDVI de octubre de 2006, la banda Brillo de la transformación tasseled cap de octubre de 2006 y la banda Verdor de la transformación tasseled cap de junio de 2007.

Clasificación de las imágenes por medio de redes neuronales

Para cada una de las etapas de clasificación antes mencionadas se entrenaron dos perceptrones multicapa mediante el algoritmo de Backpropagation. En este tipo de redes neuronales de aprendizaje supervisado, cada patrón del conjunto de entrenamiento consta de su correspondiente salida esperada. El patrón de entrenamiento esta representado por los n valores de un píxel, correspondientes a las n bandas seleccionadas. La cantidad de neuronas de entrada se corresponde con el número de bandas seleccionadas.

Primera Red

Del conjunto de píxeles de las muestras de entrenamiento de las clases agua, impermeabilizado y vegetación, se tomaron en forma aleatoria 500 píxeles de cada una, para entrenar un perceptrón multicapa con backpropagation.

Se realizaron numerosas pruebas con distinto número de capas ocultas y distinto número de neuronas en dichas capas. Finalmente la arquitectura seleccionada, por haber tenido el mejor comportamiento, fue de 6 neuronas en la capa de entrada 30 neuronas en la capa oculta y 1 en la capa de salida.

Las 6 neuronas de la capa de entrada son la banda 4 de octubre de 2006, la banda 7 de enero de 2007, la banda promedio de NDVI, la banda varianza de NDVI, la banda brillo de la transformación tasseled cap de la imagen de enero de 2007, y la banda verdor de la transformación tasseled cap de la imagen de abril de 2007.

La función de transferencia para el primer nivel fue la tangente hiperbólica y la del nivel de salida la función sigmoidea. La función de entrenamiento fue el gradiente descendente con momento y factor de aprendizaje adaptativo. El factor de aprendizaje fue 0.8 y se entreno durante 20000 épocas consiguiendo un error menor a 10^{-3} (ECM=0.000248396).

Los píxeles rotulados que no entraron dentro de las muestras para entrenar la red se utilizaron como conjunto de prueba.

En la Figura 5a se presenta el escalón que marca los valores esperados para cada uno de los píxeles de entrenamiento. Los primeros 500 valores corresponden a la

clase agua, cuyo valor esperado fue 0; los segundos 500 corresponden a los píxeles de la clase urbano, con un valor esperado de 0.5 y los últimos 500 pertenecen a la clase vegetación, con un valor esperado de 0.99. Estos valores están superpuestos con la salida producida por la red para cada uno de los píxeles. Se observa una importante coincidencia entre los valores de salida de la red y los valores esperados, lo cual indica que la red aprendió correctamente a clasificar los píxeles del conjunto de entrenamiento.

En la Figura 5b se presenta la salida de la red para todos los píxeles rotulados, incluyendo además de los del conjunto de entrenamiento (5a), aquellos que no fueron vistos por la red previamente. De esta manera, vemos gráficamente que la capacidad de generalización de la red fue muy buena.

La arquitectura elegida por tener el mejor comportamiento constó de 6 neuronas en la capa de entrada, 35 neuronas en la capa oculta y 1 en la capa de salida. Las 6 bandas usadas fueron la banda 5, de enero de 2007; la banda 4, de abril de 2007; la banda 5, de junio de 2007; la banda NDVI, de octubre de 2006; la banda brillo, de la transformación *tasseled cap* de octubre de 2006 y la banda verdor de la transformación *tasseled cap* de junio de 2007.

En el entrenamiento de la segunda red, la función de transferencia para el primer nivel y el nivel de salida, así como la función de entrenamiento, el factor de aprendizaje y la cantidad de épocas para el entrenamiento, fueron los mismos que para la primera. En la segunda red también se consiguió un error menor a 10^{-3} (ECM= 0.000571059).

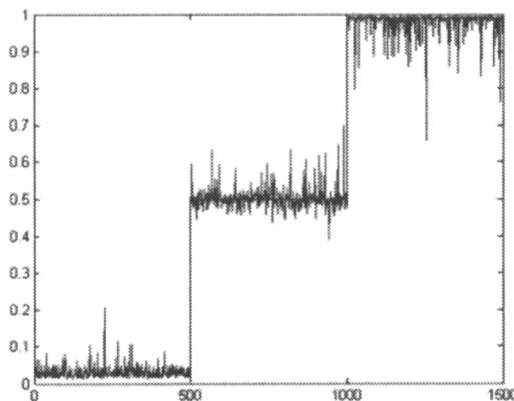


Figura 5a. Testeo. Valores de salida (de 0 a 1) de los píxeles usados para la etapa de entrenamiento de la primera red (1,500 píxeles) y su relación con los valores de salida esperados (0, 0.5 y 0.99).

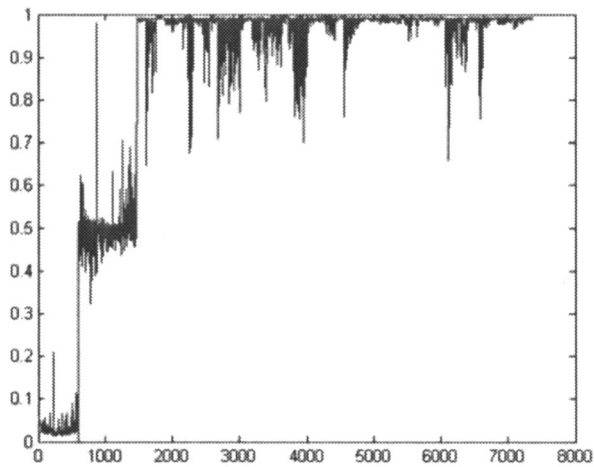


Figura 5b. Generalización. Valores de salida (de 0 a 1) producida por la red 1 entrenada incluyendo píxeles no utilizados en la etapa de entrenamiento y su relación con los valores de salida esperados (0, 0.5 y 0.99).

Los píxeles rotulados que no entraron dentro de las muestras para entrenar la red se utilizaron como conjunto de prueba.

En la Figura 6a se ven los valores esperados para cada una de las clases y los valores devueltos por la red entrenada, siendo los 350 primeros correspondientes a bosques de perennifolias (con valor esperado de salida 0) y los siguientes a bosques de caducifolias (valor esperado= 0.25), pastos (valor esperado= 0.50), asociación vegetal invernal (valor esperado= 0.75) y asociación vegetal estival (valor esperado = 0.99). En la Figura 6b se muestra la salida de la red para todos los píxeles rotulados, incluyendo además de los presentados en la Figura 6a, aquellos que no fueron utilizados por la red en la etapa de entrenamiento.

Generación del mapa de salida

Luego de haber clasificado los píxeles usando la red 1 entrenada, los valores de salida resultantes continuos entre 0 y 1 fueron reescalados a valores enteros entre 0 y 100. Para establecer las clases de coberturas se establecieron rangos de valores de acuerdo al ajuste observado en los límites geográficos entre clases en el mapa resultante. Los píxeles asignados a la clase vegetación fueron sometidos a la segunda red usando las bandas seleccionadas previamente. A los valores de salida de la segunda red se le sumó el valor fijo 100, con el objeto de poder mapear la salida de la primera y segunda red en forma conjunta.

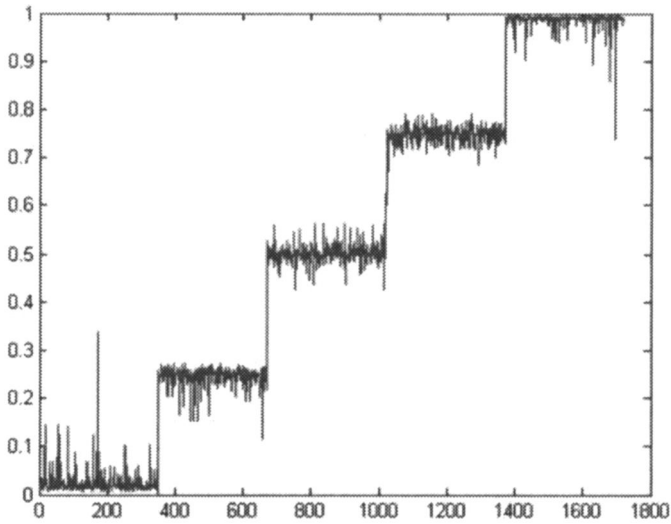


Figura 6a. Testeo. Valores de salida (de 0 a 1) de los píxeles usados para la etapa de entrenamiento de la segunda red (1750 píxeles) y su relación con los valores de salida esperados (0; 0.25; 0.5; 0.75 y 0.99).

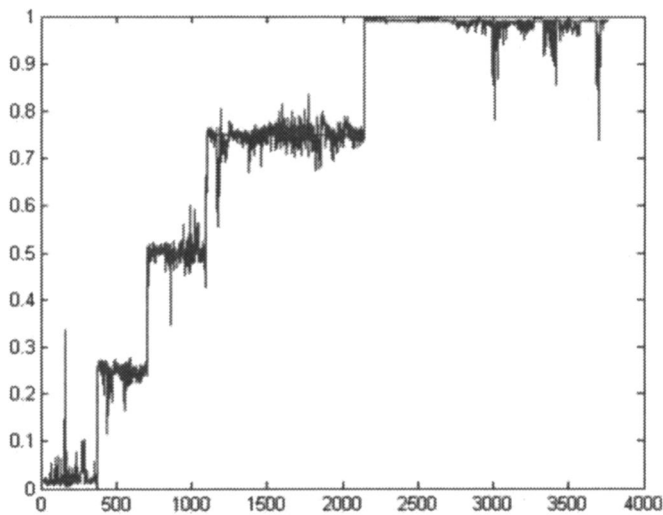


Figura 6b. Generalización. Valores de salida (de 0 a 1) producidos por la red 2 (entrenada) incluyendo píxeles no utilizados en la etapa de entrenamiento y su relación con los valores de salida esperados (0; 0.25; 0.5; 0.75 y 0.99).

Los valores esperados para las clases y subclases del esquema conceptual de clasificación fueron los siguientes: agua: 0, impermeabilizado: 50, bosques de perennifolias: 100, bosques de caducifolias: 125, pastos: 150, asociación vegetal invernal: 175 y asociación vegetal estival: 199.

En la Figura 7 se muestra el circuito del píxel desde la imagen original, hasta llegar a la clasificación generada por las redes ya entrenadas. A partir de las cuatro imágenes originales (Landsat 5 TM) más las bandas generadas posteriormente (Tabla 1) se define el *conjunto óptimo* de bandas a usar como entrada en cada red. Para cada píxel se toman las bandas del conjunto óptimo para alimentar a la red 1, o sea que a cada píxel le corresponde un conjunto de datos numéricos que son los correspondientes a las bandas usadas. Luego del procesamiento usando la red 1 entrenada se produce la salida "S1". Si la respuesta obtenida es "agua" o "impermeabilizado", se produce esa salida como clasificación resultante. En cambio si la respuesta obtenida por la red 1 es "vegetación", se toman las bandas del *conjunto óptimo* de la red 2 y la salida producida será entonces "S2", completando la clasificación. El resultado es un mapa de las coberturas de tipos de suelo con 7 clases: agua, impermeabilizado, bosque de perennifolias, bosque de caducifolias, pastos, asociación vegetal invernal y asociación vegetal estival.

Resultados

La superficie de la cuenca del río Matanza-Riachuelo se calculó en 206,502.21 ha. El mapa de coberturas resultante de la clasificación multispectral (Figura 8) muestra la distribución espacial de las coberturas del suelo en la cuenca. La matriz de error (o confusión) para la evaluación de la exactitud de la clasificación (Tabla 2), presenta en la diagonal la cantidad de píxeles correctamente clasificados por clase (concordancia entre verdad de campo y mapa). Fuera de la diagonal se muestra la cantidad de píxeles que fueron clasificados incorrectamente, con los porcentajes entre paréntesis. La exactitud general de la clasificación fue del 95.13% (2,971 píxeles correctamente clasificados sobre 3,123 evaluados). El coeficiente Kappa, que ajusta la proporción de píxeles correctamente clasificados a la probabilidad de acierto por azar (Congalton *et al.*, 1983; Congalton y Green, 1999), fue de 0.94.

La distribución porcentual de las distintas coberturas en el área de la cuenca se muestra en la Tabla 3. En la Tabla 4 se presenta una caracterización de las distintas clases de coberturas presentes en la Figura 8.

Tabla 2
Matriz de error de la clasificación

Clase	Verdad de campo							Total
	AGU	IMP	PER	CAD	PAS	AVI	AVE	
AGU	595 (99,83)	0	0	0	0	0	0	595
IMP	1 (0,17)	513 (100)	0	0	0	2 (0,44)	0	516
PER	0	0	369 (98,93)	0	0	0	0	369
CAD	0	0	3 (0,8)	332 (100)	51 (11,81)	0	0	386
PAS	0	0	1 (0,27)	0	313 (72,45)	1 (0,22)	0	315
AVI	0	0	0	0	66 (15,28)	436 (94,99)	5 (1,2)	507
AVE	0	0	0	0	2 (0,46)	20 (4,36)	413 (98,8)	435
Total	596	513	373	332	432	459	418	3123

Símbolos: AGU: agua, IMP: impermeabilizado, PER: bosque de perennifolias, CAD: bosque de caducifolias, PAS: pastos, AVI: asociación vegetal invernal, AVE: asociación vegetal estival.

Tabla 3
Superficie de las distintas coberturas identificadas en el área de la cuenca del Matanza-Riachuelo (año 2006-2007)

Cobertura	Píxeles	Hectáreas	Porcentaje
Agua	3,712	334.08	0.16
Impermeabilizado	528,859	47,597.31	23.05
Bosque de perennifolias	29,383	2,644.47	1.28
Bosque de caducifolias	71,068	6,396.12	3.1
Pastos	802,509	72,225.81	34.98
Asociación vegetal invernal	539,067	48,516.03	23.49
Asociación vegetal estival	319,871	28,788.39	13.94
Total	2,294,469	206,502.21	1.00

Análisis del mapa de coberturas

A lo largo del eje urbano-rural (sentido NE-SO), el mapa muestra en la cuenca baja un predominio de cobertura impermeabilizada, que disminuye hacia el periurbano principalmente por la aparición de una mayor proporción de coberturas vegetales. La mancha urbana (sustrato impermeabilizado) se extiende formando brazos a lo largo de rutas de acceso a la ciudad de Buenos Aires, que se ubican en las zonas más elevadas. Contrariamente, la mancha de vegetación ingresa en el sector urbano principalmente siguiendo las zonas bajas del valle de inundación del Matanza-Riachuelo, que en la cuenca media (periurbano) está ocupado principalmente por bosques de caducifolias y perennifolias. Hacia el sector rural esta zonificación en la distribución de coberturas relacionada con la posición topográfica, es mucho menos evidente.

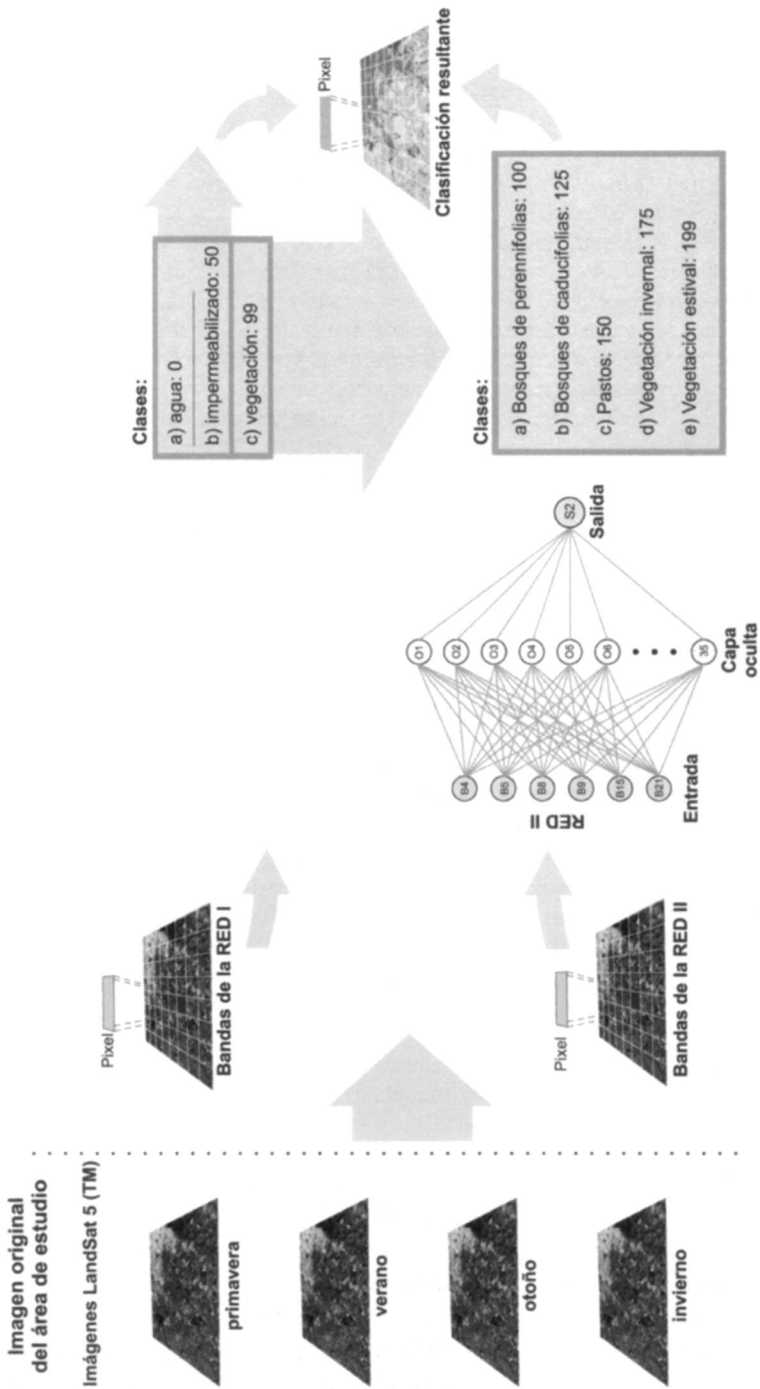


Figura 7. Circuito del pixel desde la imagen original hasta la clasificación resultante.

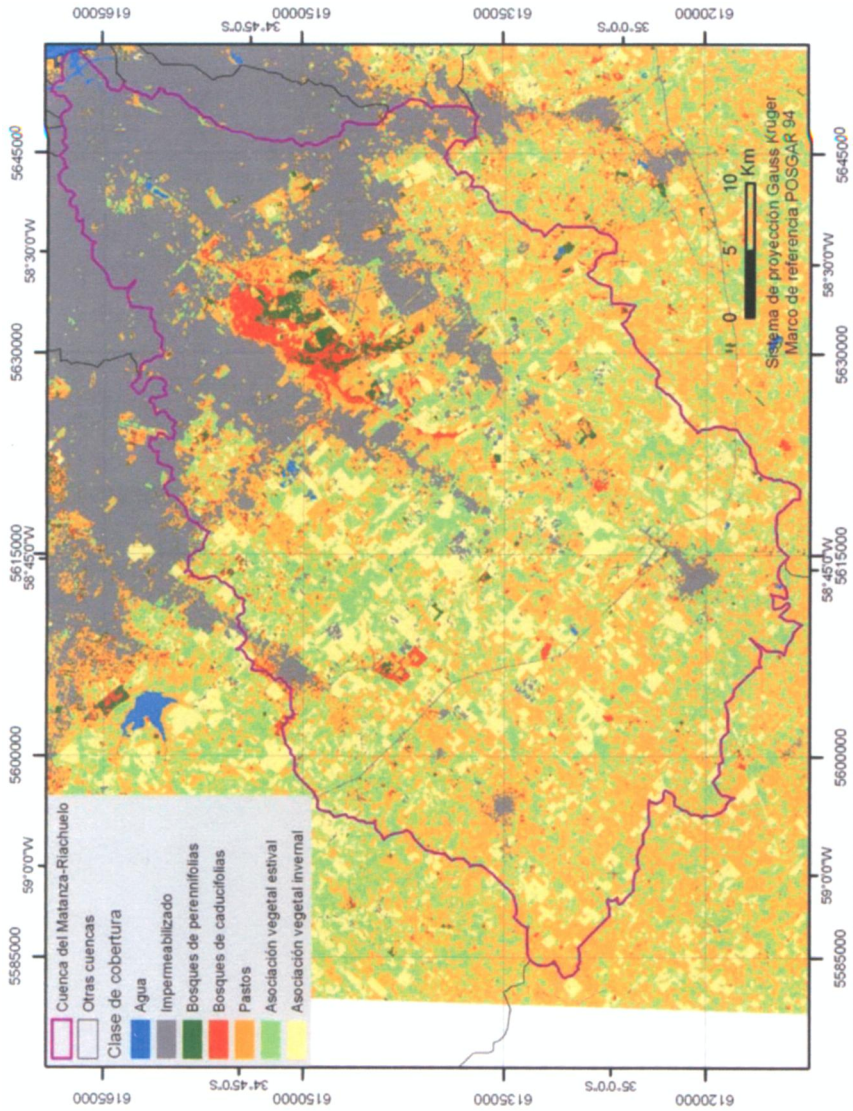


Figura 8. Mapa de coberturas del suelo de la cuena del Matanza-Riachuelo.

Tabla 4
Caracterización de las distintas coberturas en la cuenca del Matanza- Riachuelo

<i>Clase</i>	<i>Descripción</i>
Agua	Representada por cuerpos y cursos de agua permanente de la cuenca. En la cuenca baja se distinguen el Riachuelo en su desembocadura y los Lagos de Lugano. En el periurbano existen piletas y lagunas artificiales (partido de Ezeiza) y cavas inundadas remanentes de la extracción minera de tosca, especialmente en las cercanías de la localidad de Gonzalez Catán en La Matanza. Hacia la cuenca alta también aparecen cavas asociadas a las ciudades satélites y en menor medida lagunas naturales.
Impermeabilizado	Cobertura presente principalmente en zonas urbanizadas de la baja cuenca (AMBA) y de las ciudades satélites (Cañuelas, General Rodríguez). Desde el AMBA se extienden hacia el SO brazos de suelo impermeabilizado (asfaltado, construido o compactado) siguiendo las rutas de acceso a la ciudad de Buenos Aires. Hacia la cuenca alta se presentan píxeles progresivamente más dispersos, algunos sobre las rutas asfaltadas importantes como la ruta nacional 6 que se observa como una línea que atraviesa la cuenca en sentido NO-SE.
Bosques de perennifolias	Formados mayormente por especies de eucaliptos y pinos y por casuarinas, se encuentran presentes especialmente en los bosques de Ezeiza. Hacia la cuenca alta se presentan en parches más o menos pequeños en plantaciones casi siempre asociadas a construcciones.
Bosques de caducifolias	Cobertura formada principalmente por la invasora acacia negra (<i>Gleditsia triacanthos</i>) y en menor medida por paraíso (<i>Melia azedarach</i>), que también se presenta en asociaciones boscosas espontáneas o bien en plantaciones, igual que el álamo (<i>Populus spp</i>). Se encuentran grandes masas de estos bosques en la zona inundable baja de la cuenca intermedia.
Pastos	Cobertura vegetal heterogénea dominada por gramíneas. Se caracteriza por la presencia de tejidos verdes vivos en primavera, verano y otoño. Se encuentra distribuida principalmente hacia las zonas altas de la cuenca, aunque también está presente en buena medida en el periurbano.
Asociación vegetal invernial	Cobertura vegetal con presencia de tejidos verdes en primavera y otoño. Se encuentra distribuida de en parcelas rectangulares de las zonas rurales altas, por lo que se la puede asociar a lotes con doble cultivo: cultivo de invierno (trigo, avena u otros), seguido de soja (soja de segunda).
Asociación vegetal estival	Cobertura vegetal con tejidos verdes en verano, que se distribuye principalmente en parcelas rectangulares del sector rural de la cuenca, alto y medio. Presenta gran actividad fotosintética en verano lo que permite relacionarla con lotes con cultivos de verano en el ciclo anual estudiado (soja o maíz principalmente).

Siguiendo un eje desde las áreas urbanas hacia el sector periurbano se espera un aumento de la cantidad de coberturas por unidad de área. En términos de diversidad de coberturas se presume un valor mínimo en la ciudad, máximo hacia el periurbano (mayor heterogeneidad de coberturas), e intermedio en las áreas rurales.

A fin de medir la variación de la heterogeneidad ambiental (cantidad de coberturas por unidad de área) a lo largo del eje urbano-rural se calculó el índice de diversidad de Shannon (Zar, 1999) en cuadrados de 300 m de lado (9 ha) para toda el área de la cuenca. Para verificar la variación de la tendencia general a lo largo del eje urbano-rural de esta variable, se realizó un muestreo sistemático de valores de diversidad equidistantes cada 1,200 metros que luego se ajustaron a una superficie mediante la función spline (Franke, 1982).

Los resultados del análisis de la variación de la diversidad de coberturas por unidad de área a lo largo del eje urbano-rural se muestran en la Figura 9. Los valores bajos se presentan hacia el NE (áreas urbanas) y hacia el SO (zonas rurales), mientras que los valores más altos se presentan en el sector periurbano.

Discusión

El mapa generado mediante el sistema de clasificación utilizado muestra una exactitud relativamente elevada, por lo que se lo puede considerar confiable. Si se toman las dos etapas de la clasificación por separado, durante el entrenamiento de las redes el error cuadrático medio de la primera (clasificación de tres coberturas) es menor que el de la segunda. Esto es consecuencia de la mayor similitud de respuesta espectral entre las subclases de vegetación en comparación con la existente entre el agua, la cobertura impermeable y la vegetación.

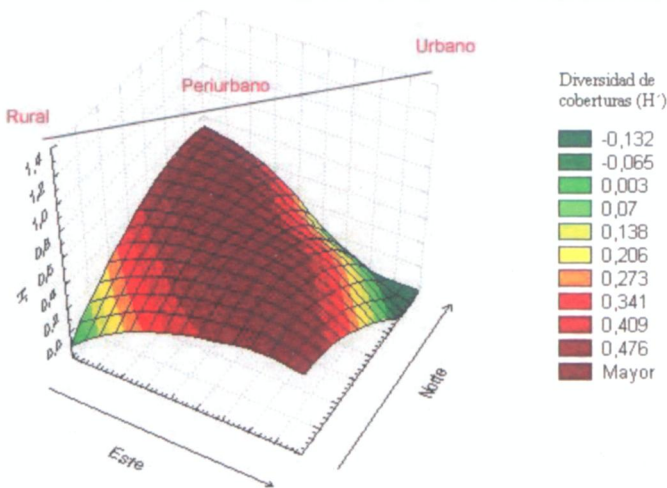


Figura 9. Variación geográfica del índice de diversidad de Shannon (H') ajustado mediante la función spline. Se observan valores bajos hacia el NE (áreas urbanas) y hacia el SO (zonas rurales), mientras que las áreas intermedias (periurbano) presentan los valores más altos.

Por otra parte, en el mapa resultante entre las subclases de vegetación se produjo mayor confusión relativa que entre estas y las coberturas no vegetales. Sin embargo, el porcentaje de píxeles bien clasificados fue alto para todas las subclases de vegetación, salvo para los pastos en los que casi un 30% de los píxeles usados en la verificación resultaron mal clasificados, y confundidos principalmente con asociaciones vegetales invernales y bosques de caducifolias. Este error se puede deber al efecto de la heterogeneidad real de la asociación vegetal denominada *a priori* como pasto. Probablemente esta subclase está formada por una mezcla de especies dispuestas en parches, en los que las dominantes resultan tener un comportamiento fenológico (ciclo de vida) tal que espectralmente se asemejan a la vegetación arbórea caducifolia o a la vegetación invernal.

La clase de cobertura menos representada en la cuenca fue el agua, a la que le siguen los bosques de perennifolias y los bosques de caducifolias que de manera conjunta representan menos del 5% de las coberturas de la cuenca. Sin embargo, es probable que debido a que los bosques de perennifolias están formados por especies arbóreas invasoras (ej.: acacia negra, *Gleditsia triacanthos*) su superficie cubierta se encuentre en aumento. Por el contrario, las especies que forman los bosques de perennifolias se caracterizan por su escasa capacidad de dispersión, por lo que posiblemente se trate de una clase de cobertura relativamente estable.

Cerca de un cuarto de la superficie de la cuenca se encuentra impermeabilizada y ese porcentaje se encuentra casi en su totalidad en la cuenca baja. Esto pone de manifiesto la gran presión antrópica a la que se encuentran expuestas las aguas del río. Dentro de la cuenca la clase pastos es la más abundante y las clases Asociación Vegetal Invernal (AVI) y Asociación Vegetal Estival (AVE) también se presentan en importante proporción.

La abundancia de pastos en grandes zonas de la cuenca alta y cerca de las divisorias de agua, podría estar relacionada con la escorrentía difusa de las zonas de nacientes, que por ende son propensas al anegamiento y poco aptas para la agricultura.

Las coberturas AVI y AVE se encuentran representadas mayormente en parcelas rectangulares de la cuenca alta y, por lo tanto, se las puede relacionar con el uso agrícola. De hecho, podrían tomarse como una misma clase si se hablara de “usos del suelo”, ya que por la alternancia de sistemas de rotación de cultivos, un lote, puede presentar en años consecutivos doble cultivo (comúnmente, cultivos de invierno, principalmente trigo, seguido de “soja de segunda”) en cuyo caso sería clasificado como AVI, o bien, un solo cultivo de verano, principalmente maíz o soja, clasificándose entonces como AVE.

Por otra parte, existen porciones de la clase pasto que corresponderían a parcelas ganaderas con pasturas implantadas, con lo que, sumando a éstas las proporciones

de coberturas relacionadas con cultivos (AVI y AVE) la cuenca podría estar utilizada en la actualidad con fines agropecuarios en valores cercanos al 70%.

A lo largo del gradiente urbano-rural, el patrón de variación de la diversidad de coberturas por unidad de área es un aspecto destacado que resulta del análisis del mapa. El método de cuantificación de la heterogeneidad ambiental (diversidad de coberturas), resultó apropiado para poner de manifiesto las diferencias estructurales existentes entre las distintas porciones de la cuenca y podría ser usado para caracterizar el paisaje, clasificar y delimitar las zonas de la cuenca, aunque es probable que sea necesario integrar otras variables para generar un mapa de zonas funcionales.

En este análisis se evidencia que la zona de mayor heterogeneidad del territorio corresponde al periurbano, que es precisamente la zona de mayor dinámica de transformación. El estudio del cambio en la heterogeneidad ambiental podría realizarse a partir del análisis de mapas multitemporales de coberturas de la cuenca. Dicho análisis serviría para establecer ritmos de crecimiento y entender los patrones de cambio en la distribución de coberturas, y serviría como herramienta en la toma de decisiones de manejo.

Conclusiones

El mapa de coberturas del suelo en la cuenca del río Matanza-Riachuelo obtenido mediante el análisis de una serie multitemporal de imágenes de resolución intermedia, clasificadas usando redes neuronales presento una exactitud del 95.13%, lo que significa que el resultado es altamente satisfactorio, sobre todo teniendo en cuenta la complejidad del ambiente estudiado.

El mapa resultante presenta en el área de la cuenca un 0.16% ocupado por agua, un 23.05% por suelos impermeabilizados, principalmente relacionados con la urbanización, mientras que la vegetación cubre el 76.79%. Además, entre las coberturas vegetales, aquellas que probablemente estén relacionadas con usos agropecuarios podrían llegar a cubrir el 70% de la cuenca.

A lo largo de un eje desde las áreas urbanas hacia el sector rural existen cambios en la diversidad de coberturas por unidad de área: el valor de diversidad es mínimo en la ciudad, crece hasta un máximo en el periurbano (máxima heterogeneidad de coberturas), y baja nuevamente en las áreas rurales, aunque alcanzando valores intermedios entre las dos zonas funcionales anteriores.

Si bien en este trabajo se verifica la utilidad de las redes neuronales en la rápida producción de insumos de base (mapas) para la gestión ambiental de la cuenca del Matanza-Riachuelo, o de otras áreas similares, hay que tener en cuenta que es necesario realizar monitoreos periódicos para verificar las tendencias en la disposición espacial de coberturas del suelo. En este sentido, queda pendiente la utilización de

imágenes satelitales de archivo (captadas durante los últimos 30 años), para reconstruir la historia de la ocupación de la cuenca.

Agradecimientos

Agradecemos a Ana Faggi y a Federico Denegri por su ayuda en los muestreos de campo, a Sandra Torrusio y su equipo por el asesoramiento en el trabajo de calibración de las imágenes, a Cristina Parpaglione por su apoyo en la utilización de redes neuronales y a Jorge Alba Posse por el diseño de figuras. Además agradecemos especialmente a la Comisión Nacional de Actividades Espaciales (CONAE) por la cesión de las imágenes Landsat 5 (TM).

Bibliografía

- Borrough, P.A., *Principles of Geographical Information Systems for Land Resources Assessment*, Oxford University Press, New York, 194 pp., 1986.
- Congalton, R.G., Oderwald, R.G. y Mead R.A., "Assessing Landsat classification accuracy using discrete multivariate statistical techniques", en *Photogram. Eng. Remote Sens.*, 49, pp. 1671-1678, 1983.
- Congalton, R.G. y Green K., *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, CRC Press, Boca Raton, 160 pp., 1999.
- Foody, G.M., "Hard and soft classifications by a neural network with a non-exhaustively defined set of classes", en *International Journal of Remote Sensing*, 23, pp. 3853-3864, 2002.
- Guerschman, J.P., Paruelo, J.M., Di Bella, C., Giallorenzi, M.C. y Pacin, F. "Land cover classification in the Argentine Pampas using multi-temporal Landsat TM data", en *International Journal of Remote Sensing*, 24, pp. 3381-3402, 2003.
- Kauth, R.J. y Thomas, G.S., "The tasseled Cap -a graphic description of the spectral-temporal development of agricultural crops as seen by LANDSAT", en: *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*, Purdue University of West Lafayette, Indiana, pp. 41-51, 1976.
- Jong, S.M. de, van der Meer, F.D. and Clevers, J.G.P.W., "Basics of remote sensing", en de Jong, S.M. y van der Meer, F.D. (Eds.) *Remote sensing image analysis: including the spatial domain*, pp. 1-15, Dordrecht, 2004.
- Lu, D. y Weng, Q., "A survey of image classification methods and techniques for improving classification performance", en *International Journal of Remote Sensing*, 28 (5), pp. 823-870, 2007.
- Lunetta, R. y Balogh, E., "Application of multi-temporal Landsat 5 TM imagery for wetland identification", en *Photogrammetric Engineering & Remote Sensing*, 65(11), pp. 1303-1310, 1999.

- Paola, J.D. y Schowengerdt, R.A., "A review and analysis of back propagation neural networks for classification of remotely sensed multispectral imagery", en *International Journal of Remote Sensing*, 16, pp. 3033-3058, 1995.
- Morello, J. y Matteucci, S., "Biodiversidad y fragmentación de los bosques en la Argentina", en Matteucci, S., Solbrig, O., Morello, J. y Halffer G. (Eds.) "Biodiversidad y uso de la tierra. Conceptos y ejemplos de Latinoamérica", Eudeba, pp. 463-498, Buenos Aires, 1999.
- Mörtberg, U.M., Balforsa, B. y Knolb W.C., "Landscape ecological assessment: a tool for integrating biodiversity issues in strategic environmental assessment and planning" en *Journal of Environmental Management*, 82, pp. 457-470, 2007.