

Análisis del desempeño de técnicas de aprendizaje automático para identificar vegetación acuática con bandas de Sentinel-2

VIRGINIA VENTURINI^{1,2}✉; ZULEICA Y. MARCHETTI^{1,2}; ELISABET WALKER^{1,2} & GIANFRANCO FAGIOLI¹

¹Facultad de Ingeniería y Ciencias Hídricas, Universidad Nacional del Litoral. ²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

RESUMEN. En la Argentina se observa una frecuencia creciente de desastres naturales como desbordes de ríos, sequías extremas e incendios forestales naturales. Ante estas catástrofes, es fundamental gestionar de manera eficiente la toma de decisiones rápidas para minimizar los daños; esta es una preocupación latente tanto de los gobiernos locales y regionales como de la comunidad científica. La cuenca del río Paraná, en la Argentina, representa un recurso estratégico, ya que engloba la mayor riqueza fluvial y ecológica, además de grandes centros urbanos. Sin embargo, los eventos extremos que caracterizan la dinámica de los humedales afectan a los centros urbanos ubicados en sus cercanías. La presencia de vegetación acuática (libre o arraigada) enmascara las áreas anegadas y ocultan los primeros indicios de anegamiento, lo cual dificulta el monitoreo y la rápida detección de esas áreas. En este trabajo se utilizaron imágenes satelitales ópticas y modelos de aprendizaje automático para clasificar las diferentes coberturas del suelo en humedales del sistema fluvial del Paraná. Se hizo foco en ambientes donde coexisten el agua libre y la vegetación acuático-palustre, típicos de la región metropolitana de la ciudad de Santa Fe, considerando las limitaciones técnicas de los organismos tomadores de decisiones. Se usaron imágenes de la misión Sentinel-2 (S2) para entrenar y evaluar el rendimiento de distintos algoritmos de aprendizaje automático. Se utilizaron todas las bandas de las imágenes S2, unificando la resolución espacial en 10 m. Los resultados indicaron que las bandas de aerosoles costeros (B1) y dos bandas de infrarrojo medio (B11 y B12) son las que aportan más información para identificar las clases muestreadas. Por otra parte, el método random forest fue el de mejor desempeño para la clase vegetación acuática, de interés especial en este trabajo.

[Palabras clave: segmentación, Sentinel-2, aprendizaje automático, humedales fluviales]

ABSTRACT. Performance analysis of machine learning techniques to identify aquatic vegetation with Sentinel-2 bands. Natural disasters, such as river overflows, extreme droughts, natural forest fires are more frequently observed in Argentina. Faced with these catastrophes, efficient management is essential to make quick decisions to minimize damage, which is a latent concern in local and regional governments and in the scientific community. In Argentina, the Paraná River basin represents a strategic resource in itself, as it encompasses the greatest fluvial, ecological wealth and large urban centers. However, the extreme events that characterize the dynamics of the wetlands affect the urban centers located near them. The presence of aquatic vegetation (free or rooted) masks the flooded areas, hiding the first signs of flooding, making the monitoring and rapid detection of these areas difficult. In this work, optical satellite images and machine learning models were used to classify the different land covers in wetlands of the Paraná river system. The focus was on environments where free water and aquatic marsh vegetation coexist, characteristic of the metropolitan region of the city of Santa Fe, and considering the technical limitations of decision-making agencies. Therefore, the Sentinel-2 (S2) mission images were used to train and evaluate different machine learning algorithms. All bands of S2 images were used, unifying the spatial resolution to 10 m. The results indicated that the coastal aerosol bands (B1) and two mid-infrared bands (B11 and B12) provide the most information for the identification of the samples. Moreover, the random forest method showed the best performance for the aquatic vegetation class, which was of primary interest for this work.

[Keywords: segmentation, Sentinel-2, machine learning, river wetlands]

INTRODUCCIÓN

Los humedales en general, y las planicies de inundación de los ríos en particular, presentan características que hacen que el monitoreo de su anegamiento sea difícil de realizar por métodos tradicionales (registros *in situ*). Esto es aún más evidente durante un evento hídrico extremo dado que la vegetación, adaptada a la dinámica de los humedales, enmascara la presencia de agua (Klemas 2013; Sánchez-Espinosa and Schröder 2019). La información a tiempo cuasi-real relacionada con estos ambientes adquiere una importancia especial en la gestión de una emergencia de origen hídrico, en la que la comunidad y sus servicios sanitarios básicos (e.g., agua potable, cloaca, etc.) son amenazados (Li et al. 2020).

La definición de los humedales varía según las jurisdicciones, las organizaciones o las disciplinas, pero, en esencia, son áreas sujetas a inundaciones permanentes, periódicas o a saturaciones prolongadas del suelo (Tiner et al. 2015). Los humedales se pueden pensar como el nexo entre las tierras altas no saturadas y las aguas profundas, y así integran el mosaico del paisaje de llanura (National Wetlands Working Group 1997).

Los humedales fluviales, como los que componen las planicies de inundación de los grandes ríos americanos, son difíciles de individualizar y de representar en cartografía. En particular, presentan diferentes grados de conectividad hídrica, y de acuerdo con las fases hidrológicas, pueden estar completamente inundados sin diferenciación de límites o, estar completamente secos; la presencia de agua en superficie se puede limitar solo al curso de agua principal y sus secundarios (Kandus et al. 2018). En estos humedales, la conectividad hídrica se inicia cuando los ríos comienzan a incrementar su caudal, y cuando este excede el nivel de desborde, el área cubierta por agua aumenta hasta su máxima expresión (Marchetti et al. 2013; Marchetti et al. 2016). Así, el monitoreo del avance de la inundación se torna indispensable para su gestión. El monitoreo de los humedales sobre la base de datos recopilados en el terreno es muy informativo, pero resulta limitado en tiempo y espacio, y es dificultoso y costoso en términos logísticos, sobre todo en áreas extensas o de difícil acceso (Li et al. 2020). Por lo tanto, el inventario y el monitoreo de los humedales se suelen realizar con datos y técnicas de teledetección, a costos relativamente bajos y

con una cobertura espacial amplia (Tiner et al. 2015).

Durante las últimas cuatro décadas, se produjeron grandes cantidades de datos a partir de satélites de observación con diferentes sensores a bordo, los cuales se utilizaron para actualizar la cartografía y para seguir distintos fenómenos naturales a nivel mundial (Lang and McCarty 2008). Así, los enfoques de monitoreo y detección remota abordaron de manera efectiva gran parte de las dificultades propias de los enfoques tradicionales, como lo son la accesibilidad y repetitividad (Guo et al. 2017; Mahdavi et al. 2018). Un ejemplo de esto es la Misión Sentinel de la Agencia Espacial Europea (ESA, por sus siglas en inglés), cuyo objetivo es apoyar la generación de políticas para el monitoreo global del ambiente. La ESA ofrece imágenes de alta resolución espacial, tiempo de revisita y resolución espectral, provenientes de sensores ópticos, térmicos y de radar obtenidos de sus distintos satélites (Sánchez-Espinosa and Schröder 2019; Slagter et al. 2020).

Los algoritmos de aprendizaje automático se convirtieron en una parte integral en los estudios que usan información obtenida a partir de herramientas de teledetección, debido a su capacidad para realizar clasificaciones de coberturas de la superficie con diferentes niveles de complejidad (Mahdianpari et al. 2017). Los algoritmos de aprendizaje automático permiten extraer e interpretar datos satelitales de forma automática y robusta, con el fin de obtener información sobre la condición de los humedales. Las técnicas de aprendizaje automático superan a la clasificación basada en uno o varios índices porque permiten incorporar diferentes propiedades de la imagen, además de incorporar diferentes bandas e índices (Mirzazloumi et al. 2021). Los algoritmos más populares que se usan para este fin son los bosques aleatorios (RF, por sus siglas en inglés) (Breiman 2001) y las máquinas de soporte vectorial (SVM, por sus siglas en inglés) (Cortes and Vapnik 1995). Estos algoritmos de aprendizaje automático son técnicas flexibles que permiten ser entrenados y probados con combinaciones de datos obtenidos a partir de diversos tipos de sensores. Mahdianpari et al. (2017), Whyte et al. (2018), Berhane et al. (2018) y Akbari et al. (2021), entre otros, presentan ejemplos de aplicaciones exitosas de ambos métodos.

A pesar de la gran aplicabilidad de estas herramientas, el mapeo temático de paisajes

complejos como los de humedales del río Paraná es un desafío debido a la alta similitud de la respuesta espectral de algunas clases (Ozesmi and Bauer 2002). Por ejemplo, la vegetación sobre suelo seco y la vegetación sobre agua generan fácilmente confusión en el esquema de clasificación y dificultan el desempeño correcto de estos algoritmos, sobre todo para los usuarios que deben tomar decisiones sensibles en términos sociales a partir de los mapas generados. En tal sentido, la mayoría de los trabajos mencionados utilizan datos de imágenes de radar y ópticas, que requieren conocimientos avanzados sobre el procesamiento de imágenes que los usuarios finales podrían no disponer. Las imágenes Sentinel-2 (S2) son muy utilizadas con este fin porque disponen de una buena resolución espectral, espacial y tiempo de revisita. Sin embargo, en general solo se usan las bandas de resolución espacial 10 m y 20 m (Akbari et al. 2021; Li et al. 2021; Vanderhoof et al. 2021).

En este contexto, el presente trabajo tiene como objetivo estudiar distintos algoritmos de aprendizaje automático para segmentar áreas con vegetación acuática, utilizando todas las bandas de las imágenes de S2, y analizar la importancia de cada banda en la clasificación de estos ambientes complejos.

MATERIALES Y MÉTODOS

Área de estudio

Este trabajo se llevó a cabo en los humedales de la zona de influencia de la ciudad de Santa Fe, representativos del tramo medio de sistema fluvial del río Paraná, centrados en 31°37' S - 60°33' O (Figura 1). El río Paraná es considerado uno de los 10 ríos más importantes del mundo (Ashworth and Lewin 2012) y es el segundo en Sudamérica después del Amazonas (Latrubesse 2008). Su caudal medio histórico es 17000 m³/s (registrados en el puerto de la ciudad de Paraná), con extremos de 8000 y 60000 m³/s (Giacosa et al. 2000). Su régimen hidrológico se caracteriza por una sucesión de aguas altas y bajas, con inundaciones regulares (cada 1 a 3 años) y extremas (frecuencias variables) que modelan la dinámica de la vegetación característica (Carignan and Neiff 1992; Ceirano et al. 2000). La variabilidad espacio-temporal en la que este gran sistema fluvial se fue construyendo y modelando (manifiesta en sus geoformas, tipos de vegetación, redes de drenaje, etc.)

determina la presencia de diferentes tipos de humedales. Estos van desde los netamente lóticos (con agua en circulación en forma constante) a los leníticos (aguas quietas), e incluyen entre ambos extremos un amplio gradiente en términos de conectividad hidro-sedimentológica y características bióticas asociadas.

En cualquiera de sus expresiones, la vegetación es un elemento común que define la presencia y muchas veces la dinámica misma de los humedales. Entre algunos de los humedales frecuentes del sistema fluvial del Paraná aparecen bosques inundables (dominados por *Salix humboldtiana* y *Tessaria integrifolia*), bosques en galería con menor recurrencia de inundaciones, pero igualmente adaptados a inundaciones extraordinarias (dominados por *Albizia inundata*, *Nectandra angustifolia*, *Inga uraguensis*, *Croton urucurana* y *Erythrina cristagalli*, entre otros), pajonales (representados por *Coleataenia prionitis*), bajos, pantanos, bordes de laguna, cauces abandonados y otros cuerpos de agua colonizados por diferentes tipos de comunidades acuático-palustres que colonizan estos humedales en forma parcial o total dependiendo del nivel y la permanencia del agua (*Polygonum* spp., *Ludwigia peploides*, *Echinochloa* spp., *Paspalum* spp., *Panicum* spp., *Azolla filiculoides*, *Salvinia biloba* y *Eichhornia crassipes*) (Marchetti et al. 2013).

Datos de campo y de la misión Sentinel-2

La información de campo se recopiló en sucesivas campañas en las que se navegaron y accedieron a diferentes cursos de agua y lagunas del área de estudio. En cada campaña se georreferenció la vegetación acuática arraigada a las márgenes de los cursos de agua, y también la vegetación flotante en el interior de las lagunas. En cada caso, se estimaron de forma visual las dimensiones del área cubierta por la vegetación, se listaron las especies presentes y se asignó un valor de abundancia-cobertura a las especies presentes en 3 réplicas de superficie conocida. La metodología de trabajo con la vegetación acuática está descrita en Marchetti y Ramonell (2014). Estas campañas permitieron delinear polígonos sobre las imágenes satelitales correspondientes a fechas cercanas (~7 días), que en conjunto representan ~30% del total de polígonos (muestras) utilizados; el 70% restante se definió sobre la base de imágenes satelitales, buscando parámetros similares de textura, color y tonalidad.

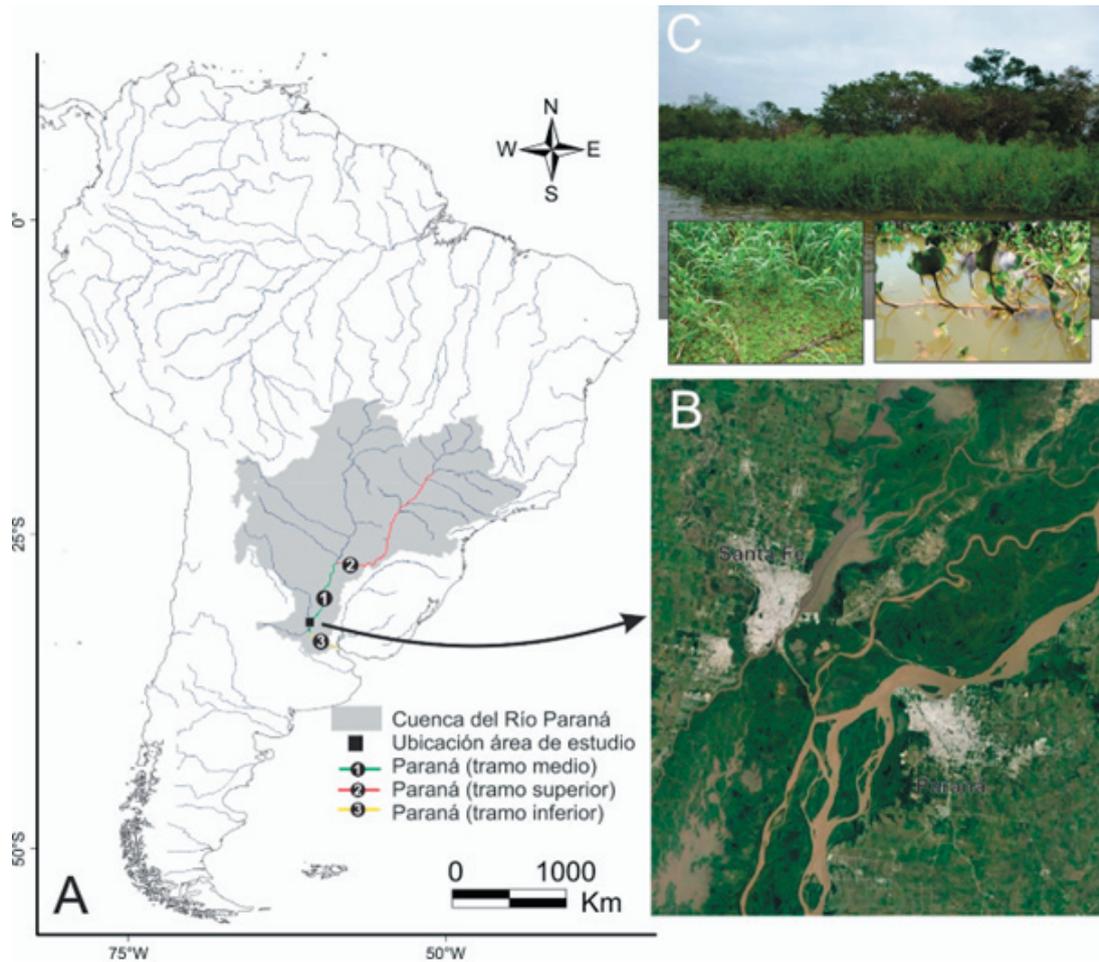


Figura 1. A). Ubicación del área de estudio dentro de la Argentina y de la Cuenca del Plata. B) Imagen color verdadero de los ambientes acuáticos muestreados. C) Vegetación acuática típica de la zona.

Figure 1. A) Location of the study area within Argentina and La Plata River Basin. B) True color image of the aquatic environments sampled. C) The typical aquatic vegetation of the study area.

Se utilizaron las imágenes Landsat 8 disponibles en el entorno Google Earth Pro, con las fechas más cercanas a las imágenes S2 seleccionadas; es decir, con una diferencia no mayor a 7 días. En cada imagen se digitalizaron de forma manual ~250 polígonos para cada clase (i.e., agua, vegetación acuática [vegetación flotante o arraigada a las márgenes o el lecho], ciudades y vegetación sobre tierra firme), lo cual resultó en casi 1000 polígonos. Si bien el objetivo de este trabajo es discriminar la clase vegetación acuática de las otras clases presentes en un ambiente de humedales, se agregó la clase ciudad por tener una respuesta espectral diferencial, que contraste y pueda usarse de control (Ball et al. 2017; Ozesmi and Bauer 2002).

Se extrajeron las coordenadas de los píxeles dentro de los polígonos definidos sobre

Landsat 8 y se identificaron en las imágenes S2 adquiridas en ~7 días de la fecha de visita de Landsat 8. Se unificó la resolución espacial de todas las bandas de S2 a fin de generar matrices apilables. Para esto, se cambió el tamaño de los píxeles de las bandas con el método de interpolación lineal (Hwang and Lee 2004) de forma tal que todas las bandas tuvieran píxeles de 10 m. Cabe aclarar que de esta forma no se modifica el valor de la reflectancia de los píxeles, es decir un píxel de 60 m (3600 m²) se dividió en 36 píxeles de 100 m² cuya reflectancia es igual al valor original (Chuvieco 1991; García et al. 2019).

En cada imagen se generaron ~250 polígonos por clase, de los que se extrajo el conjunto de datos que resultó en ~200000 píxeles-muestras para las clases agua, vegetación sobre tierra firme y ciudad; para la clase vegetación

acuática se contaba con solo 7000 muestras. Este desbalance entre las clases se resolvió descartando muestras al azar de las clases agua, ciudad y vegetación sobre tierra firme, de forma tal que las clases tuvieran un total de muestras geográficamente balanceadas y de ~7000 píxeles por clase. De esta forma se buscó que la evaluación de los modelos no esté afectada por un problema de representatividad o desbalance de clases. En resumen, la base de datos se conformó con cerca de 7000 muestras por clase, por imagen. Cabe mencionar que en este trabajo se hace énfasis en la clase vegetación acuática, ya que es la que más dificultades presenta para su muestreo en campo (etiquetado de los píxeles) y tiene una respuesta espectral resultado de la superposición de las respuestas espectrales de las clases agua y vegetación (Ozesmi and Bauer 2002).

Se reemplazaron las imágenes de Landsat 8 originalmente utilizadas para volcar los datos de campo con imágenes de mejor resolución espacial, como S2. Esta es una misión del programa Copernicus, que forma parte de la ESA. Esta misión está compuesta por 2 satélites que se encuentran en la misma órbita, desfasados 180° entre sí. El primero de los satélites fue puesto en órbita el 26 de junio del 2015 y el segundo se lanzó al espacio el 7 de marzo del 2017. El tiempo de revisita es de 10 días por satélite, lo que significa que se obtienen imágenes de una misma zona cada 5 días. Estos satélites llevan a bordo el instrumento multiespectral cuyo barrido es de 290 km (Drusch et al. 2012; Walker et al. 2021).

Las imágenes S2 están formadas por 13 bandas que cubren desde el espectro visible hasta el infrarrojo de onda corta (SWIR, por sus siglas en inglés). Cada una de las bandas de S2 tiene una resolución que puede ser de 10, 20 y 60 m (Tabla 1). Se seleccionaron imágenes S2 con un nivel de procesamiento 1C (LC1). Estas imágenes contienen correcciones geométricas que incluyen el registro multitemporal subpixel entre imágenes y la ortorectificación usando un DEM de 90 m (Planet DEM 90). También incluyen el procesamiento radiométrico utilizando los coeficientes para recuperar las radiancias por inversión de la transformación y, además, contienen máscaras de píxeles defectuosos (nubes y cirrus). Las imágenes seleccionadas para este trabajo cumplen con las siguientes condiciones: 1) se consideraron imágenes durante una condición hidrológica del río de aguas bajas-medias (i.e., que el nivel en el hidrómetro del puerto de Santa Fe-Paraná se mantenga en valores de 2 a 3 m), en la misma estación climática, ya que el foco de interés del análisis está en zonas donde coexisten el agua y la vegetación, previas a los eventos extremos; 2) la presencia de nubes, que contaminan la escena captada por el satélite, no debía superar el 10% de la imagen, y 3) la fecha de adquisición de la imagen debía estar próxima a alguna de las fechas en la que se hicieron campañas de muestreo en campo (i.e., siete días como máximo). Las imágenes seleccionadas que cumplían las condiciones explicadas arriba corresponden a las fechas 15-08-2017, 19-09-2017 y 10-08-2018. Las escenas seleccionadas estaban libres de nubes y el río Paraná en el hidrómetro del puerto de Santa Fe varió de 2.22 m a 2.72 m durante el período analizado.

Tabla 1. Descripción de las bandas de Sentinel-2.

Table 1. Description of Sentinel-2 bands.

Banda	Resolución espacial (m)	Longitud de onda central (nm)	Ancho de banda (nm)	Descripción
Banda 1	60	443	430-450	Azul-Violeta
Banda 2	10	490	450-520	Azul
Banda 3	10	560	540-570	Verde
Banda 4	10	665	650-680	Rojo
Banda 5	20	705	690-710	Borde Rojo 1
Banda 6	20	740	730-740	Borde Rojo 2
Banda 7	20	783	770-790	Borde Rojo 3
Banda 8	10	842	780-900	Infrarrojo cercano
Banda 8a	20	865	850-870	Infrarrojo cercano 2
Banda 9	60	940	930-950	Infrarrojo de onda corta (vapor de agua)
Banda 10	60	1375	1360-1390	Infrarrojo de onda corta (cirrus)
Banda 11	20	1610	1560-1650	Infrarrojo de onda corta 1
Banda 12	20	2190	2100-2280	Infrarrojo de onda corta 2

Modelos de inteligencia artificial

Actualmente, el aprendizaje automático es uno de los subcampos de la inteligencia artificial (IA) de más rápido crecimiento; su foco está puesto en el aprendizaje a partir de datos. Este progreso reciente en el aprendizaje automático fue impulsado tanto por el desarrollo de nuevos algoritmos y técnicas de aprendizaje estadístico como por el aumento de la disponibilidad de datos en línea y la computación de alto rendimiento a bajo costo. En este trabajo se analizaron tres de los métodos más populares con el fin de seleccionar el que tenga mejor desempeño para separar diferentes coberturas terrestres.

Máquina de soporte vectorial. SVM fue desarrollado por Vapnik et al. (1995). En esencia, una máquina de soporte vectorial es un algoritmo que intenta colocar una frontera de decisión de forma tal que se maximice la distancia de la misma a las clases presentes en la colección de datos (verdad de campo) que se dispone. Las bondades de esta metodología radican en las sólidas bases matemáticas de la teoría estadística detrás de este método. Las SVM demostraron resultados satisfactorios en una amplia gama de problemas del mundo real. Diferentes autores usaron este método para diferenciar vegetación acuática (Hunter et al. 2010; Singh et al. 2014; Tang et al. 2018; Piaser and Villa 2022).

Bosques aleatorios. Los RF son una combinación de árboles de decisión formado por nodos, en los que cada árbol depende de valores seleccionados aleatoriamente del conjunto de datos (Breiman 2001). El proceso de entrenamiento de RF se basa en generar un nuevo conjunto de datos, en el que sus valores se seleccionan aleatoriamente del conjunto de datos de entrenamiento. El clasificador usado para este estudio consiste en utilizar características seleccionadas al azar o una combinación de características en cada nodo para generar un árbol. Este nuevo grupo de datos puede tener un número de muestras menor al de entrenamiento e incluso tener valores repetidos. Luego se confecciona un árbol de decisión para este subconjunto. El procedimiento antes detallado se repite *N* veces generando *N* conjuntos con sus correspondientes árboles de decisión asociados a cada uno, quedando así conformado un bosque aleatorio con *N* árboles. Al momento de utilizar RF para predecir un valor, todos los árboles del bosque predicen una salida y el valor más popular es el que finalmente se

elige como salida del método. Husson et al. (2016), Piaser y Villa (2022) y de Grandpré et al. (2022) publicaron resultados prometedores usando este método.

Perceptrón multicapa. El MLP es un modelo computacional inspirado en el funcionamiento de las neuronas biológicas. Está compuesto por capas sucesivas de neuronas artificiales, las cuales reciben las señales de entrada y las ponderan mediante los pesos sinápticos aprendidos. Esta ponderación permite asignar una importancia a cada una de estas entradas. Los pesos sinápticos se van modificando en la etapa de aprendizaje y son los parámetros entrenables del modelo. El modelo más simple de MLP cuenta con 3 capas: una de entrada, una capa oculta y otra de salida. Todas las neuronas de la capa de entrada reciben los mismos datos de entrada, pero las neuronas de las capas ocultas y de salida reciben como entrada la salida de todas las neuronas de la capa anterior. Por este motivo, a estos modelos se los denomina redes totalmente conectadas. Este modelo es entrenado por medio del algoritmo de gradiente descendente. Diferentes autores aplicaron MLP para clasificar distintos ambientes (Dronova et al. 2012; Krishnaraj and Honnasiddaiah 2022; Pande-Chhetri et al. 2014).

Detalles de implementación

Este trabajo se realizó en el entorno de Google Colab utilizando Python, junto a las librerías Numpy para el manejo de matrices de datos, Sklearn para los modelos propuestos, Imageio para las imágenes satelitales y Seaborn para los gráficos. Para visualizar los resultados de estas clasificaciones de forma sencilla, se diseñó una interfaz gráfica que permite sectorizar la imagen para clasificar solo las regiones de interés del usuario y evitar así el consumo de tiempo computacional procesando áreas que no tienen interés. Esta interfaz, codificada en Python utilizando la librería PyQt5, tiene un menú muy simple que permite abrir una imagen S2, visualizarla en forma completa o focalizar en un sector de interés, seleccionar los sectores que se desean segmentar, clasificar y visualizar la imagen ya clasificada. Al presente se dispone de un prototipo básico del sistema —realizado para este trabajo— que demostró ser eficiente en términos computacionales. Por otro lado, cabe aclarar que la imagen se puede clasificar en forma completa o sectorizada según la capacidad computacional del usuario y el área de interés.

El clasificador RF tiene asociado un método que permite conocer qué bandas de S₂ tuvieron más importancia en la clasificación. La cuantificación del aporte de cada banda en RF se califica en un rango de 0 a 1.0; un valor de 0 indica que la banda no aporta a la clasificación, y uno de 1.0, el mayor peso posible.

También se utilizó el método de eliminación recursiva de características (RFE, por sus siglas en inglés) para chequear los resultados de RF. Este método también asigna pesos a las bandas, utilizando un clasificador de base, que en este caso fue SVM. Primero se entrena el estimador con todas las bandas, se eliminan las que menos importancia tienen, y así sucesivamente hasta que solo queden la cantidad de bandas deseadas.

Entrenamiento y prueba de los modelos

El proceso de entrenamiento de cada modelo implica la selección de los hiperparámetros que configuran la red para adaptarlo al problema y lograr el mejor desempeño utilizando solo los datos de entrenamiento y validación. Esta tarea se implementó en una búsqueda de grilla que combina una amplia variedad de valores de los parámetros y mediante sucesivas pruebas se seleccionan los más eficaces. La selección de la configuración más exitosa se realiza mediante el cálculo de la tasa de acierto (Ac, ver Ecuación 2), eligiendo la configuración que arroje el mayor Ac. Esto se repitió para los tres métodos propuestos en este trabajo. Específicamente, se optó por calibrar SVM con una función kernel lineal y la función de base radial, ambos con valores de C variando de 100 a 1000 y gamma variando de 0.3 a 30. En el caso de RF, se calibró la cantidad de árboles (variando entre 50 a 5000) con distintas cantidades de características en cada uno (Pal 2005). La topología de MLP quedó definida por una capa de entrada de 13 nodos, m capas ocultas con m neuronas en cada una y la capa de salida con las cuatro clases. Con este conjunto de hiperparámetros, se conformaron las grillas el entrenamiento de los métodos, en el caso de SVM la grilla involucraba 24 iteraciones, 12 iteraciones para RF y 15 para MLP, (Pal 2005; Millar and Richardson 2015; Berhane et al. 2018; García et al. 2019).

Una vez escogidos los hiperparámetros de los modelos, se prosiguió a entrenar y comparar los modelos. Para los experimentos se consideraron variaciones tanto en la dimensión temporal como espacial. La

dimensión temporal hace referencia a diferentes condiciones hidro-ambientales de las escenas de las tres imágenes S₂, ya que cada una fue obtenida en distintas fechas. Mientras que la dimensión espacial hace referencia a la heterogeneidad propia de los distintos sectores de una misma imagen, que son propias de estos paisajes de humedales. Se realizaron las siguientes pruebas para evaluar el desempeño de los métodos: Prueba 1) entrenamiento y test en dominio espacial y temporal fijo (hace referencia a tomar un sector de una imagen para entrenar y luego clasificar sobre otro sector de la misma imagen; este proceso se realiza en cada imagen por separado); Prueba 2) entrenamiento y test en dominio espacial fijo y temporal variante (hace referencia a entrenar con datos de un sector fijo tomando valores dentro del mismo en distintas fechas, y luego clasificar en las imágenes correspondientes a esas mismas fechas), y Prueba 3) entrenamiento y test en dominio espacial y temporal variante (hace referencia a entrenar con datos de diversos sectores tomando valores dentro de los mismos en distintas fechas, y luego clasificar en las imágenes capturadas en una fecha diferente).

Para evaluar el desempeño de los clasificadores en cada prueba propuesta, se utilizaron la exhaustividad y la tasa de acierto como métricas del error. Además, se calculó la matriz de confusión, que también evalúa el desempeño de los clasificadores (López et al. 2018).

La exhaustividad (Re, por sus siglas en inglés) por clase es la efectividad de un clasificador para identificar clases.

$$Re = VP / (VP + FN) \quad \text{Ecuación 1}$$

donde VP= verdadero positivo y FN= falso negativo.

La tasa de acierto (Ac, por sus siglas en inglés) expresa la tasa de acierto promedio del clasificador.

$$Ac = (VP + VN) / (VP + FP + VN + FN)$$

$$\text{Ecuación 2}$$

donde, verdadero positivo (VP) expresa que un elemento de clase relevante es correctamente clasificado, mientras que el verdadero negativo (VN) se refiere a los elementos de la clase irrelevante correctamente identificados.

Falso positivo (FP) hace referencia a que un elemento de la clase relevante es asignado a la clase irrelevante. Por último, Falso negativo (FN) significa que los elementos de la clase relevante no fueron identificados. En el caso de clasificación multiclase, la exactitud se computa directamente como la razón entre la cantidad de instancias bien clasificadas y el total de instancias analizadas.

RESULTADOS

SVM se aplicó en todos los casos con una función Kernel de base radial, con C variando entre 50-100 y γ 0.3 en todos los casos. En el caso de RF, los mejores resultados se encontraron con 100 a 150 árboles y con cuatro a seis características, dependiendo de la prueba. La topología de MLP quedó definida por una capa de entrada de 13 nodos, las capas ocultas variaron entre cuatro y cinco, con seis u ocho neuronas en cada una y la capa de salida con las cuatro clases.

En la Figura 2 se exponen los resultados obtenidos en las sucesivas pruebas realizadas para todas las clases muestreadas, y en la Figura 3 se muestran los resultados logrados para la clase vegetación acuática. En la Figura 2 se muestra el desempeño obtenido por los clasificadores, teniendo en cuenta el Ac para todas las clases juntas. En términos generales, se puede ver que MLP y SVM responden bien cuando se hace el entrenamiento y test en dominio espacial y temporal fijo (Prueba 1), pero todos los métodos pierden capacidad de generalización cuando se aumenta la variabilidad temporal de los datos en la Prueba 2. En efecto, durante la

Prueba 1, todos los clasificadores arrojaron un Ac superior a 0.95. Para la Prueba 2, solo RF mostró un valor medio de Ac superior a 0.95, mientras que los valores medios de los otros clasificadores cayeron por debajo de 0.9, aunque con una gran variabilidad en las diferentes realizaciones. En la Prueba 3, RF y SVM conservan una buen Ac (superior a 0.95), pero MLP cayó por debajo de 0.9.

La clase vegetación acuática arrojó valores de Re muy cercanos a 1 para el clasificador RF en todas las pruebas realizadas. Como se explicó antes, para la Prueba 2 se seleccionan áreas prefijadas e invariantes en la imagen S2 y se buscaron los valores de las bandas de todas las imágenes seleccionadas (para todas las fechas). Este proceso genera mayor dispersión en las características de cada clase, ya que contempla un problema de variabilidad del entorno. Sin embargo, podemos decir que RF identifica las características espectrales de cada clase en diferentes condiciones. En la Figura 3 se evidencia que SVM baja el rendimiento en la Prueba 2 para la clase vegetación acuática, la que presenta un Ac medio de 0.7, inferior a MLP y RF, aunque no sería un valor despreciable. También se observa que SVM tiene una dispersión de Ac superior a la observada para MLP y RF. La Prueba 3 dio buenos resultados para todos los clasificadores ($Ac > 0.9$), aunque RF sobresale por tener un Ac próximo a 1. A modo informativo, los valores de Re obtenidos para las otras clases fueron superiores a 0.96 para todas las pruebas diseñadas en este trabajo. Claramente, las clases con firmas espectrales bien definidas son más fáciles de identificar por los algoritmos utilizados y no presentan

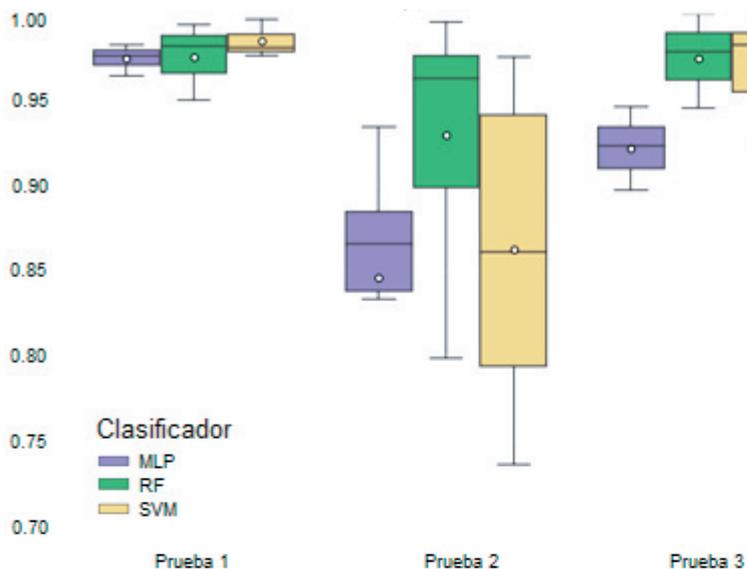


Figura 2. La tasa de aciertos de los diferentes modelos (valor medio y mediana en la caja, cuartiles en los límites de esta, y mínimo y máximo como extremos) teniendo en cuenta todas las clases. MLP hace referencia al método perceptrón multicapa, RF refiere a bosques aleatorios y SVM refiere a máquina de soporte vectorial.

Figure 2. Accuracy of the different models (mean value and median in the box, quartiles at the limits of the box, and minimum and maximum as extremes) considering all classes. MLP refers to the multilayer perceptron method, RF refers to random forest and SVM refers to support vector machine.

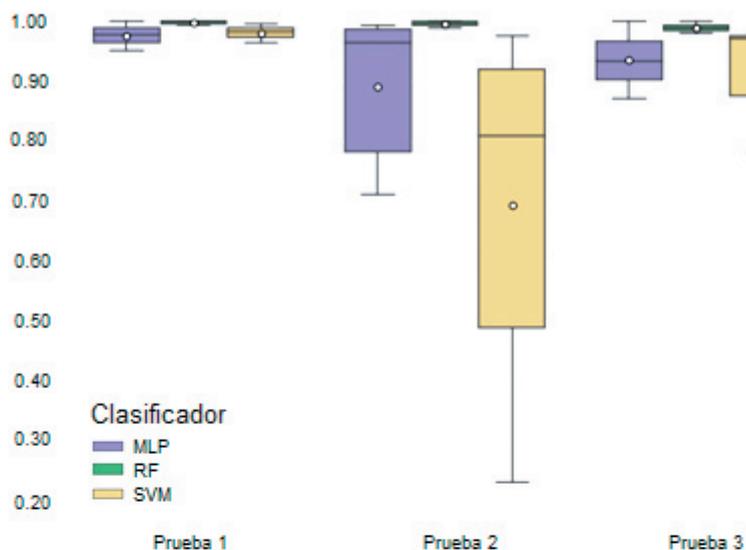


Figura 3. La exhaustividad de los diferentes modelos (valor medio y mediana en la caja, cuartiles en los límites de esta, y mínimo y máximo como extremos) de la clase vegetación acuática. MLP hace referencia al método perceptrón multicapa, RF refiere a bosques aleatorios y SVM refiere a máquina de soporte vectorial.

Figure 3. Recall of the different models (mean value and median in the box, quartiles at the limits of the box, and minimum and maximum as extremes) of the aquatic vegetation class. MLP refers to the multilayer perceptron method, RF refers to random forest and SVM refers to support vector machine.

problemas con los cambios naturales de estos sistemas complejos. En la Figura 4 se puede observar la matriz de confusión de la Prueba 3 para RF de 100 árboles con 5 características; se observa que las clases con firmas espectrales simples — como las clases agua y ciudad — no presentan confusión, mientras que las otras dos clases de vegetación presentan cierto nivel de confusión. La precisión global de esta Prueba fue de 0.9775, mientras que el valor del coeficiente K resultó en 0.9709; ambos valores demuestran que la clasificación resulto aceptable (Akbari et al. 2021).

Se seleccionaron tres áreas de interés dentro de la imagen S2 del 10-08-2018, cuando se

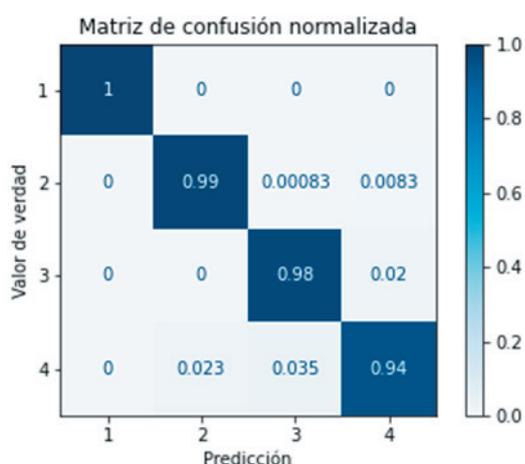


Figura 4. Matriz de confusión de los resultados de la Prueba 3. 1: Clase agua. 2: Clase ciudad. 3: Clase vegetación acuática. 4: Clase vegetación sobre tierra firme.

Figure 4. Confusion matrix of Test 3 results. 1: Water class. 2: City class. 3: Aquatic vegetation class. 4: Land vegetation class.

registraban 2.72 m de altura en el río Paraná a la altura de Santa Fe. En la Figura 5 se muestran los detalles de los tres sectores. En la Figura 5-A se perfila un sector de la ciudad de Santa Fe en el que se identificaron los parques más importantes de la ciudad en color verde y la laguna Setúbal en color azul. Esta imagen muestra cómo los clasificadores identifican sin problemas ambas clases. En la Figura 5-B se observa cierta confusión entre las clases ciudad y vegetación. La ruta provincial 2, a la altura de San Justo (Santa Fe), se perfila como vegetación y solo las estructuras más notorias se identifican como tales. Además, se observan zonas de vegetación acuática sobredimensionadas en algunos sectores. En contraste, la Figura 5-C muestra la zona propia del humedal cercano a la ciudad de Santa Fe, que fue correctamente clasificada en todas las clases. Si bien estos resultados son preliminares, se observa que RF puede diferenciar la vegetación acuática de las otras clases con una exactitud razonable.

La cuantificación del aporte de cada banda en RF dio como resultado que la B11 tiene un peso igual a 1.0, mientras que B1 y B12 tienen pesos de 0.9 y 0.8 respectivamente. Luego, se destaca la B2, con un peso de 0.6, y las restantes bandas tienen pesos inferiores o iguales a 0.4.

La verificación de los resultados anteriores usando el método RFE arrojó una selección de bandas similar a la de RF, escogiendo las bandas B1, B2, B9, B11 y B12 (todas con un peso igual a 1.0). La diferencia más notoria con RF está en B9, que tuvo un peso importante según RFE. Por lo demás, los resultados corroboran los encontrados con RF.

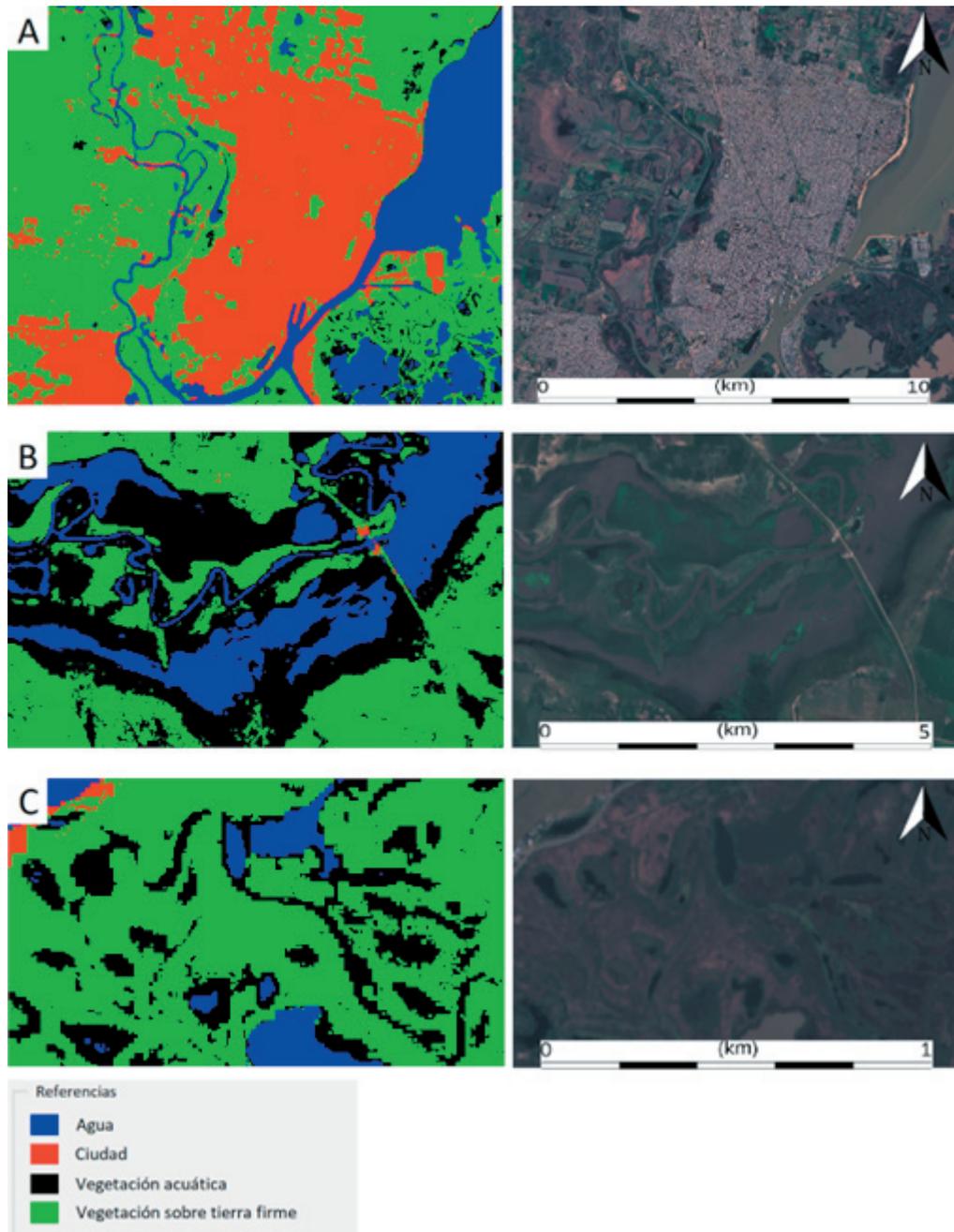


Figura 5. Clasificación resultante de aplicar random forest a la imagen de fecha 10-08-2018, en la que en rojo se observan ciudades; en verde, vegetación en tierra firme; en azul, agua, y en negro, vegetación acuática, para tres sectores: A) ciudad de Santa Fe y Laguna Setubal. B) Puente sobre el río Salado, en la Ruta Provincial 2. C) Zona frente a la ciudad universitaria de la Universidad Nacional del Litoral, conocido como Vuelta del Paraguayo.

Figure 5. Classification resulting from applying random forest to the image dated 10-08-2018, in which, in red, cities are observed; in green, land vegetation; in blue, water, and in black, aquatic vegetation, for three sectors: A) city of Santa Fe and Laguna Setubal. B) Bridge over the Salado River on Provincial Route 2. C) Area in front of the Universidad Nacional del Litoral campus, known as Vuelta del Paraguayo.

DISCUSIÓN

El monitoreo *in situ* y con métodos remotos de la vegetación acuática y no acuática que permanece inundada durante los eventos

de anegamiento representa un desafío aun con técnicas de aprendizaje automático (Munizaga et al. 2022). La vegetación y el agua son elementos claves en los humedales y, de acuerdo con los niveles inundación,

el agua puede resultar enmascarada por la vegetación hasta el punto de que su presencia solo comienza a ser perceptible por los satélites cuando la magnitud de la inundación alcanzó valores críticos.

Muchos autores trabajaron sobre el mapeo de los humedales utilizando información de diferentes misiones satelitales. Sin embargo, hasta el momento no se encuentra literatura sobre el mapeo con información de S₂ de la vegetación acuático-palustre del río Paraná, a la que en este trabajo denominamos clase 'vegetación acuática'. En este caso se utilizaron todas las bandas de S₂, sin discriminar por su resolución espacial, y se compatibilizaron las matrices mediante una interpolación lineal usando la mejor resolución. De este modo se logró incorporar la reflectancia de todas las bandas, sin descartar información a priori. Se destaca la simpleza de la segmentación realizada, ya que los métodos propuestos en este trabajo están disponibles en distintos softwares de uso libre, al igual que las imágenes S₂, que se pueden descargar de forma gratuita en diferentes plataformas.

Con Landsat Thematic Mapper y un método no supervisado, Kandus et al. (1999) encontraron una exhaustividad del 83% en el delta del río Paraná. Las autoras etiquetaron las clases encontradas con datos de campo y fotografía aérea, por lo que la exhaustividad encontrada en un sistema tan complejo como el área de estudio es muy buena. Pereira et al. (2017) encontraron una exactitud global de ~98.5% en el Pantanal, y resaltaron el detalle espectral y espacial logrado con la fusión de imágenes de diferentes misiones satelitales. En dicho trabajo, los autores utilizan una metodología mucho más compleja que la aquí propuesta. En la contribución de Pereira et al. (2017) solo se usaron las bandas de S₂ que tienen alta resolución espacial (Tabla 1), mientras que en este trabajo se logró una precisión similar sin fusionar imágenes de diferentes sensores y con todas las bandas de S₂. Ferral et al. (2019) realizaron una clasificación de los humedales asociados a la laguna Mar Chiquita utilizando la técnica de umbralización y fusión de diferentes misiones satelitales. Entre sus resultados informaron una exactitud global superior al 99% y reconstruyeron la variabilidad de los niveles de la laguna utilizando índices de humedad ópticos.

La interpolación realizada para llevar todas las bandas a 10 m permitió explorar el rol de

la B1 en la clasificación de áreas complejas con vegetación acuática. Cabe destacar que, en general, la literatura relacionada con la clasificación de vegetación acuática en particular, y de vegetación en general, descarta la B1 por tener una resolución espacial de 60 m (Pereira et al. 2017; Sánchez-Espinosa and Schröder 2019). La banda de aerosoles costeros, cuya longitud de onda es de 443 nm, tiene mucha mayor penetración en el agua que otras bandas ópticas (Poursanidis et al. 2019). El análisis de los histogramas de la B1 de las imágenes utilizadas muestra que su reflectancia en presencia de vegetación-agua es menor que la reflectancia en agua. Si bien las bandas de longitud de onda 1600-2200 nm (SWIR) son totalmente absorbidas por el agua, la banda visible que mejor penetra el agua es la de 440 nm, aproximadamente. Berhane et al. (2018) utilizó las bandas de S₂ junto a distintos índices de verdor para entrenar los métodos árboles de decisión y RF, y clasificar diferentes ambientes. Estos autores encontraron que la banda 1 (costal band) del satélite WorldView-2 estaba entre las que mayor poder de discriminación tenían. Traganos y Reinartz (2018), Traganos et al. (2018) y Poursanidis et al. (2019) usaron las bandas de S₂ para explorar la contribución (B1 en la banda de aerosoles costeros (B1 en los 443 nm) en la cartografía de los hábitats costeros marinos e incluso en batimetrías de tres lugares del Mediterráneo oriental. Estos autores encontraron una exactitud de ~87% en la mayoría de los casos. Estos trabajos sustentan nuestros resultados, en el sentido de que B1 también aportaría información valiosa en la clasificación de vegetación sumergida y sobre el agua cuando se la usa combinada con índices de verdor o con bandas sensibles a la clorofila como el infrarrojo cercano. Las bandas SWIR son muy absorbidas por zonas vegetadas donde hay agua, por lo que aporta información para diferenciar los estados húmedos o secos de diferentes clases (Ferral et al. 2019).

Es importante destacar que los problemas de variación del entorno espacial pueden afectar los resultados de una clasificación si cambian los niveles de agua y conectividad entre los humedales. En este trabajo se observó que la Prueba 2, en la que se fijaba el espacio y se incluían los datos de diferentes fechas en el entrenamiento de los métodos, fue la de mayor exigencia durante la clasificación, con resultados muy variables según los métodos seleccionados. Sobre la base de los resultados

de la Prueba 2, RF mostró ser el método más robusto para capturar la variabilidad temporal intra clase. Este resultado adquiere importancia debido a que la vegetación acuática cambia con los niveles hidrométricos del sistema fluvial y con la fenología de las plantas. La Prueba 3 mostró que se puede lograr un clasificador que sea aplicable a las imágenes de primavera durante una situación hidráulica de agua medias.

Si bien todos los métodos utilizados mostraron ser eficientes para detectar las clases bien definidas espectralmente (e.g., agua), no todos mostraron buenos resultados con la clase mixta vegetación acuática. En este sentido, cabe aclarar que incluso RF parecería confundir el agua cargada de clorofila por algas o con cierto tipo de vegetación acuática.

La resolución espacial de las imágenes utilizadas también son fuente de errores, ya que en más de un caso, los píxeles podrían no ser pura vegetación acuática. Si bien el cambio de Landsat 8 a S2 mejora notablemente la

definición de los ensamblajes monitoreados, se debe tener en cuenta que la pureza del píxel es importante para obtener buenos resultados con cualquier clasificador.

En la actualidad se está trabajando en la obtención de firmas espectrales de la vegetación acuática con un radiómetro y en una localización mejor de las muestras. También se observó que es necesario discriminar diferentes tipos de vegetación y así desglosar la clase vegetación acuática y en tierra firme en varias clases. Se espera que al incluir mayor cantidad de clases se logre una mayor diferenciación de coberturas.

AGRADECIMIENTOS. Queremos agradecer al Dr. Enzo Ferrante, por su colaboración y enriquecedora discusión de los resultados de este trabajo. Agradecemos a CONICET por el apoyo a la Dr. Marchetti y a la Dra. Venturini. Este estudio está enmarcado dentro del proyecto de investigación PICT 2017 Nro. 0935 - Agencia Nacional de Promoción Científica y Tecnológica.

REFERENCIAS

- Akbari, V., M. Simpson, S. Maharaj, A. Marino, D. Bhowmik, G. Prabhu, et al. 2021. Monitoring aquatic weeds in Indian wetlands using multitemporal remote sensing data with machine learning techniques. Pp. 6847-6850 en 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. <https://doi.org/10.1109/IGARSS47720.2021.9553207>.
- Ashworth, P. J., and J. Lewin. 2012. How do big rivers come to be different? *Earth-Science Reviews* 114(1-2):84-107. <https://doi.org/10.1016/j.earscirev.2012.05.003>.
- Ball, J. E., D. T. Anderson, and C. S. Chan. 2017. Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community. *Journal of Applied Remote Sensing* 11(4):042609-042609. <https://doi.org/10.1117/1.JRS.11.042609>.
- Berhane, T. M., C. R. Lane, Q. Wu, B. C. Autrey, O. A. Anenkhonov, V. V. Chepinoga, and H. Liu. 2018. Decision-tree, rule-based, and random forest classification of high-resolution multispectral imagery for wetland mapping and inventory. *Remote Sensing* 10(4):580. <https://doi.org/10.3390/rs10040580>.
- Breiman, L. 2001. Random forests. *Machine Learning* 45(1):5-32. <https://doi.org/10.1023/A:1010933404324>.
- Carignan, R., and J. J. Neiff. 1992. Nutrient dynamics in the floodplain ponds of the Paraná River (Argentina) dominated by the water hyacinth *Eichhornia crassipes*. *Biogeochemistry* 17(2):85-121. <https://doi.org/10.1007/BF00002642>.
- Ceirano, E., C. Paoli, and M. Schreider. 2000. Las inundaciones en el área de Santa Fe. Interpretación de sus efectos y simulación de los subsistemas Leyes-Setúbal y Ruta Nacional 168-Alto Verde. Pp. 239-297 en C. Paoli C. and M. Schreider (eds.). *El Río Paraná en su Tramo Medio. Una contribución al conocimiento y prácticas ingenieriles en un gran río de llanura*. Vol. 2. Universidad Nacional del Litoral: Santa Fe.
- Chuvieco, E. 1991. Fundamentos de teledetección espacial. *Estudios Geográficos* 52(203):371.
- Cortes, C., and V. Vapnik. 1995. Support-vector networks. *Machine Learning* 20(3):273-297. <https://doi.org/10.1007/BF00994018>.
- de Grandpré, A., C. Kinnard, and A. Bertolo. 2022. Open-Source Analysis of Submerged Aquatic Vegetation Cover in Complex Waters Using High-Resolution Satellite Remote Sensing: An Adaptable Framework. *Remote Sensing* 14(2):267. <https://doi.org/10.3390/rs14020267>.
- Dronova, I., P. Gong, N. E. Clinton, L. Wang, W. Fu, S. Qi, and Y. Liu. 2012. Landscape analysis of wetland plant functional types: The effects of image segmentation scale, vegetation classes and classification methods. *Remote Sensing of Environment* 127:357-369. <https://doi.org/10.1007/s11356-022-20386-9>.
- García, G., V. Venturini, M. Brogioni, F. Walker, and L. Rodríguez. 2019. Soil moisture estimation over flat lands in Argentinean Pampas region using Sentinel-1A data and non-parametric methods. *International Journal of remote sensing* 40(10):3689-3720. <https://doi.org/10.1080/01431161.2018.1552813>.
- Giacosa, R., C. Paoli, and P. Cacik. 2000. Conocimiento del régimen hidrológico. Pp. 71-103 en C. Paoli and M. Schreider (eds.). *El Río Paraná en su Tramo Medio. Una contribución al conocimiento y prácticas ingenieriles en un gran río de llanura*. Vol. 2. Universidad Nacional del Litoral: Santa Fe.

- Drusch, M., U. Del Bello, S. Carlier, O. Colin, V. Fernández, F. Gascón, B. Hoersch, C. Isola, et al. 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment* 120:25-36. <https://doi.org/10.1016/j.rse.2011.11.026>.
- Ferral, A., E. Luccini, A. Aleksinkó, and C. M. Scavuzzo. 2019. Flooded-area satellite monitoring within a Ramsar wetland Nature Reserve in Argentina. *Remote Sensing Applications: Society and Environment* 15:100230. <https://doi.org/10.1016/j.rsase.2019.04.003>.
- Guo, M., J. Li, C. Sheng, J. Xu, and L. Wu. 2017. A review of wetland remote sensing. *Sensors* 17(4):777. <https://doi.org/10.3390/s17040777>.
- Hunter, P., D. Gilvear, A. Tyler, N. Willby, A. and Kelly. 2010. Mapping macrophytic vegetation in shallow lakes using the Compact Airborne Spectrographic Imager (CASI). *Aquatic Conservation: Marine and Freshwater Ecosystems* 20(7):717-727. <https://doi.org/10.1002/aqc.1144>.
- Husson, E., F. Ecke, H. and Reese. 2016. Comparison of manual mapping and automated object-based image analysis of non-submerged aquatic vegetation from very-high-resolution UAS images. *Remote Sensing* 8(9):724. <https://doi.org/10.1007/s12665-013-2628-0>.
- Hwang, J. W., and H. S. Lee. 2004. Adaptive image interpolation based on local gradient features. *IEEE Signal Processing Letters* 11(3):359-362. <https://doi.org/10.1109/LSP.2003.821718>.
- Kandus, P., H. Karszenbaum, and L. Frulla. 1999. Land cover classification system for the Lower Delta of the Paraná River (Argentina): Its relationship with Landsat Thematic Mapper spectral classes. *Journal of Coastal Research* 909-926. URL: jstor.org/stable/4299011.
- Kandus, P., P. G. Minotti, N. S. Morandeira, R. Grimson, G. González Trilla, E. B. González, et al. 2018. Remote sensing of wetlands in South America: status and challenges. *International Journal of Remote Sensing* 39(4):993-1016. <https://doi.org/10.1080/01431161.2017.1395971>.
- Klemas, V. 2013. Remote sensing of emergent and submerged wetlands: an overview. *Int J Remote Sens* 34:6286-6320. <https://doi.org/10.1080/01431161.2013.800656>.
- Krishnaraj, A., and R. Honnasiddaiah. 2022. Remote sensing and machine learning based framework for the assessment of spatio-temporal water quality in the Middle Ganga Basin. *Environmental Science and Pollution Research* 1-20. <https://doi.org/10.1007/s11356-022-20386-9>.
- Lang, M. W., and G. W. McCarty. 2008. Remote sensing data for regional wetland mapping in the United States: trends and future prospects. Pp. 73-112 *en* R. E. Russo (ed.). *Wetlands: Ecology, Conservation and Restoration*. Nova Science Publishers, Inc., Hauppauge, New York.
- Latrubesse, E. M. 2008. Patterns of anabranching channels: the ultimate end-member adjustment of mega rivers. *Geomorphology* 101(1-2):130-145. <https://doi.org/10.1016/j.geomorph.2008.05.035>.
- Li, C., L. Zhou, and W. Xu. 2021. Estimating aboveground biomass using Sentinel-2 MSI data and ensemble algorithms for grassland in the Shengjin Lake Wetland, China. *Remote Sensing* 13(8):1595. <https://doi.org/10.3390/rs13081595>.
- Li, Z., H. Chen, J. C. White, M. A. Wulder, and T. Hermosilla. 2020. Discriminating treed and non-treed wetlands in boreal ecosystems using time series Sentinel-1 data. *International Journal of Applied Earth Observation and Geoinformation* 85:102007. <https://doi.org/10.1016/j.jag.2019.102007>.
- López, F. J. A., J. R. Avi, and M. V. A. Fernández. 2018. Control estricto de matrices de confusión por medio de distribuciones multinomiales. *Geofocus: Revista Internacional de Ciencia y Tecnología de la Información Geográfica* (21):6. <https://doi.org/10.21138/GF.591>.
- Mahdavi, S., B. Salehi, J. Granger, M. Amani, B. Brisco, and W. Huang. 2018. Remote sensing for wetland classification: A comprehensive review. *GIScience and Remote Sensing* 55(5):623-658. <https://doi.org/10.1080/15481603.2017.1419602>.
- Mahdianpari, M., B. Salehi, F. Mohammadimanesh, and B. Brisco. 2017. An assessment of simulated compact polarimetric SAR data for wetland classification using random forest algorithm. *Canadian Journal of Remote Sensing* 43(5):468-484. <https://doi.org/10.1080/07038992.2017.1381550>.
- Marchetti, Z. Y., M. S. Pereira, C. G. Ramonell, and E. M. Latrubesse. 2013. Vegetation and its relationship with geomorphologic units in the Paraná River floodplain, Argentina. *Journal of South American Earth Sciences* 46:122-136. <https://doi.org/10.1016/j.jsames.2013.03.010>.
- Marchetti, Z. Y., and C. G. Ramonell. 2014. Valoración preliminar de la retención de sedimentos por Hidrófitas en Cauces secundarios del Río Paraná Medio. *Aqua-LAC* 6(1):8-16. <https://doi.org/10.29104/phi-aqualac/2014-v6-1-02>.
- Marchetti, Z. Y., P. G. Minotti, C. G. Ramonell, F. Schivo, and P. Kandus. 2016. NDVI patterns as indicator of morphodynamic activity in the middle Paraná River floodplain. *Geomorphology* 253:146-158. <https://doi.org/10.1016/j.geomorph.2015.10.003>.
- Millard, K., and M. Richardson. 2015. On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping. *Remote Sensing* 7(7):8489-8515. <https://doi.org/10.3390/rs70708489>.
- Mirmazloumi, S. M., A. Moghimi, B. Ranjgar, F. Mohseni, A. Ghorbanian, S. A. Ahmadi, et al. 2021. Status and trends of wetland studies in Canada using remote sensing technology with a focus on wetland classification: a bibliographic analysis. *Remote Sensing* 13(20):4025. <https://doi.org/10.3390/rs13204025>.
- Munizaga, J., M. García, F. Ureta, V. Novoa, O. Rojas, and C. Rojas. 2022. Mapping Coastal Wetlands Using Satellite Imagery and Machine Learning in a Highly Urbanized Landscape. *Sustainability* 14(9):5700. <https://doi.org/10.3390/su14095700>.
- National Wetlands Working Group. 1997. *The Canadian Wetland Classification System*. National Wetlands Working

- Group, Waterloo, ON, Canada. URL: library.wur.nl/WebQuery/clc/1783071.
- Ozesmi, S. L., and M. E. Bauer. 2002. Satellite remote sensing of wetlands. *Wetlands Ecology and Management* 10: 381-402. <https://doi.org/10.1023/A:1020908432489>.
- Pal, M. 2005. Random forest classifier for remote sensing classification. *International journal of remote Sensing* 26(1): 217-222. <https://doi.org/10.1080/01431160412331269698>.
- Pande-Chhetri, R., A. Abd-Elrahman, and C. Jacoby. 2014. Classification of submerged aquatic vegetation in black river using hyperspectral image analysis. *Geomatica* 68(3):169-182. <https://doi.org/10.5623/cig2014-302>.
- Pereira, O. J. R., A. J. Melfi, and C. R. Montes. 2017. Image fusion of Sentinel-2 and CBERS-4 satellites for mapping soil cover in the Wetlands of Pantanal. *International Journal of Image and Data Fusion* 8(2):148-172. <https://doi.org/10.1080/19479832.2016.1261946>.
- Piaser, E., P. and Villa. 2022. Comparing machine learning techniques for aquatic vegetation classification using Sentinel-2 data. pp. 465-470 *en* 2022 IEEE 21st Mediterranean Electrotechnical Conference (MELECON). IEEE. <https://doi.org/10.1109/MELECON53508.2022.9843103>.
- Poursanidis, D., D. Traganos, P. Reinartz, and N. Chrysoulakis. 2019. On the use of Sentinel-2 for coastal habitat mapping and satellite-derived bathymetry estimation using downscaled coastal aerosol band. *International Journal of Applied Earth Observation and Geoinformation* 80:58-70. <https://doi.org/10.1016/j.jag.2019.03.012>.
- Sánchez-Espinosa, A., and C. Schröder. 2019. Land use and land cover mapping in wetlands one step closer to the ground: Sentinel-2 versus Landsat 8. *Journal of Environmental Management* 247:484-498. <https://doi.org/10.1016/j.jenvman.2019.06.084>.
- Singh, S. K., P. K. Srivastava, M. Gupta, J. K. Thakur, and S. Mukherjee. 2014. Appraisal of land use/land cover of mangrove forest ecosystem using support vector machine. *Environmental Earth Sciences* 71(5):2245-2255. <https://doi.org/10.1007/s12665-013-2628-0>.
- Slagter, B., N. E. Tsendbazar, A. Vollrath, and J. Reiche. 2020. Mapping wetland characteristics using temporally dense Sentinel-1 and Sentinel-2 data: A case study in the St. Lucia wetlands, South Africa. *International Journal of Applied Earth Observation and Geoinformation* 86:102009. <https://doi.org/10.1016/j.jag.2019.102009>.
- Tang, H., Z. Dai, Y. Jiang, T. Li, and C. Liu. 2018. PCG classification using multidomain features and SVM classifier. *BioMed Research International* 2018. <https://doi.org/10.1155/2018/4205027>.
- Tiner, R. W., M. W. Lang, and V. V. Klemas. 2015. *Remote sensing of wetlands: applications and advances*. CRC press, Taylor and Francis Group. <https://doi.org/10.1201/b18210>.
- Traganos, D., and P. Reinartz. 2018. Mapping Mediterranean seagrasses with Sentinel-2 imagery. *Marine Pollution Bulletin* 134:197-209. <https://doi.org/10.1016/j.marpolbul.2017.06.075>.
- Traganos, D., B. Aggarwal, D. Poursanidis, K. Topouzelis, N. Chrysoulakis, and P. Reinartz. 2018. Towards global-scale seagrass mapping and monitoring using Sentinel-2 on Google Earth Engine: The case study of the aegean and ionian seas. *Remote Sensing* 10(8):1227. <https://doi.org/10.3390/rs10081227>.
- Vanderhoof, M. K., T. J. Hawbaker, C. Teske, A. Ku, J. Noble, and J. Picotte. 2021. Mapping wetland burned area from Sentinel-2 across the Southeastern United States and its contributions relative to Landsat-8 (2016–2019). *Fire* 4(3):52. <https://doi.org/10.3390/fire4030052>.
- Vapnik, V., I. Guyon, and T. Hastie. 1995. Support vector machines. *Machine Learning* 20(3):273-297. <https://doi.org/10.1007/BF00994018>.
- Walker, E., D. C. Fonnegra Mora, and V. Venturini. 2021. Monitoreo de excesos hídricos en Santa Fe mediante la plataforma Google Earth Engine. RPIC 2021. XIX Edición de la XIX Reunión de Trabajo en Procesamiento de la Información y Control. San Juan- Argentina. 3-5 noviembre 2021.
- Whyte, A., K. P. Ferentinos, and G. P. Petropoulos. 2018. A new synergistic approach for monitoring wetlands using Sentinels-1 and 2 data with object-based machine learning algorithms. *Environmental Modelling and Software* 104: 40-54. <https://doi.org/10.1016/j.envsoft.2018.01.023>.