

Clasificación de usos y coberturas del suelo del sudeste de la provincia de Buenos Aires utilizando Google Earth Engine

*Trabajo final presentado para optar al título de
Especialista en Teledetección y Sistemas de Información Geográfica aplicados al estudio de los
recursos naturales y la producción agropecuaria*

Alejandro Luis Salomón
Ingeniero Agrónomo - Universidad de Buenos Aires - 2018



Escuela para Graduados Ing. Agr. Alberto Soriano
Facultad de Agronomía - Universidad de Buenos Aires

TUTOR

Pablo Baldassini

Licenciado en Ciencias Ambientales (UBA)

Doctor en Ciencias Agropecuarias (UBA)

JURADO DE TRABAJO FINAL

Jurado

Carlos Marcelo Di Bella

Ingeniero Agrónomo (UBA)

Doctor en Ciencias Agropecuarias (INA-PG)

Jurado

Javier Houspanossian

Lic. Tecnología Ambiental (UNICEN)

Doctor en Ciencias Biológicas (UNC)

Fecha de defensa del Trabajo Final: 18 de marzo de 2022

Índice

Resumen	4
1. Introducción	5
2.1. Área de estudio	9
2.2 Clasificaciones de uso del suelo	10
2.2.1 Procesamiento de la información de campo	10
2.2.2 Procesamiento de imágenes satelitales y clasificación supervisada	12
2.2.3 Evaluación.....	16
2.2.4 Estimación de Área	17
3. Resultados y discusión	17
4. Conclusiones	29
5. Bibliografía.....	30
6. Anexo I: Matrices de confusión para las distintas clasificaciones	33

Resumen

La información sobre el uso y el relevamiento de los cambios de cobertura permite articular políticas para la explotación sustentable de los recursos. Los sensores remotos, montados en plataformas satelitales, proveen una cobertura completa de la superficie terrestre, y es posible modelar la información registrada para realizar predicciones. En este trabajo se realizó una clasificación supervisada para estimar el uso del suelo en los partidos de San Cayetano, Tres Arroyos y Necochea, ubicados en el sur de la provincia de Buenos Aires, para la campaña 2015-2016. El proceso de clasificación combinó la capacidad de procesamiento de la plataforma Google Earth Engine con información terrestre para cada subcampaña (invierno y verano) provista por el Método de relevamiento de Segmentos Aleatorios realizado por el Ministerio de Agricultura, Ganadería y Pesca, el cual consiste en el muestreo de áreas permanentes (segmentos) de 400 hectáreas. Estas áreas son visitadas periódicamente por técnicos que constatan los cultivos presentes y el porcentaje de ocupación de cada cultivo. Esta información se utilizó para entrenar el algoritmo de clasificación Random Forest, el cual combinó información de reflectancia superficial registrada en las bandas del espectro visible, infrarrojo cercano e infrarrojo medio del satélite Landsat 8. Se realizaron tres submuestreos (denominados total, balanceado y proporcional) con la intención de evaluar la respuesta del algoritmo frente a distintos volúmenes de datos. En este trabajo se evaluaron, además, distintas aproximaciones para la estimación de área, ya sea estimando el uso por subcampaña (invierno y verano), y a escala de campaña anual. Para los usos y coberturas obtenidas se estimaron los valores de superficie poblacional por cultivo, tanto de manera puntual como por intervalo de confianza del 95%, siguiendo la metodología propuesta por Olofsson et al. (2014). Los resultados sugieren que el muestreo balanceado permite obtener mejores precisiones que los otros dos tipos de muestreo, con una precisión general entre 77 y 92%, y precisiones del usuario y del productor superiores al 75% en la mayoría de los usos. No solo las precisiones difirieron entre clasificaciones (invierno, verano y anual), sino que también se observaron diferencias en cuanto a la superficie estimada por cada una de ellas. A su vez, las superficies estimadas presentaron importantes diferencias respecto a los datos reportados por el Ministerio de Agricultura, Ganadería y Pesca de la Nación en base a informantes calificados, ubicándose estos valores, en la gran mayoría de los casos, por fuera del

intervalo de confianza estimado por las clasificaciones. Estas diferencias observadas plantean la necesidad de incorporar nuevas metodologías de estimación que permitan un abordaje estadístico de manera de mejorar y complementar los métodos de estimación existentes.

Palabras claves: Clasificación Supervisada, Random Forest, Método de Segmentos Aleatorios, Pampa Austral

1. Introducción

Disponer de información relacionada con los usos y coberturas del suelo presentes a diferentes escalas de observación es fundamental para la planificación del territorio. El análisis a escala nacional, provincial y departamental, así como también la información sobre los cambios temporales en el uso del suelo, permite a los diferentes organismos públicos articular políticas a favor de la explotación racional y sustentable de la superficie agrícola (Hansen et al. 2000; IPCC 2000; McGuire et al. 2001), limitando así el impacto de las actividades humanas sobre el ecosistema (Cihlar 2000). En los últimos 40 años la explotación de nuevas tierras para la producción de alimentos y la intensificación del uso por unidad de superficie (mediante el uso de maquinarias, fertilizantes, y agroquímicos), significó la pérdida de áreas de vegetación nativa y, por consiguiente, la pérdida de biodiversidad, el cambio en los ciclos biogeoquímicos y la pérdida de servicios ecosistémicos (Daily 1997). Se estima que América Latina (en particular Argentina, Brasil e Uruguay) es la región con mayor cambio en el uso del suelo en ese sentido (FAOSTAT 2013; Baeza 2016).

En Argentina se usan métodos de estimación del uso agrícola basados en variables climáticas y político-cambiarías. Tal es el caso de la metodología de consulta a expertos utilizada por el Ministerio de Agricultura, Ganadería y Pesca de la Nación. Dicha metodología estima la superficie agrícola y el porcentaje ocupado por cada cultivo basándose en valoraciones subjetivas asociadas con las precipitaciones de la campaña y el precio de los commodities (MAGYP 2016). Sin embargo, la misma presenta una importante falencia asociada a la ausencia de un método estadístico que sustente o dé certezas de las valoraciones realizadas. De esta manera, las estimaciones obtenidas no permiten ser evaluadas respecto a la incertidumbre (intrínseca)

asociada a esa estimación. Es decir: se obtienen valores puntuales de superficie, en lugar de estimadores estadísticos (como la media de un conjunto de muestras, acompañado de un intervalo de confianza), los cuales permitirían atribuirle certidumbre a dicha estimación respecto al valor real (poblacional) de superficie para cada cultivo (Cap et al. 2010; Leavy et al. 2019).

El uso de sensores remotos, por su parte, permite obtener información que cumple con los objetivos estadísticos mencionados. Los sensores, montados en plataformas satelitales, permiten una cobertura completa de la superficie terrestre, y es posible modelar los datos recopilados por el sensor para realizar predicciones. Comparando cada pixel con sus vecinos, resulta posible agruparlos por su comportamiento espectral y/o fenológico en categorías o clases (tipos de cultivos o coberturas del suelo) definidas a priori y, de esta manera, clasificar los diferentes cultivos presentes en una región de interés. A esta técnica se la conoce con el nombre de “Clasificación supervisada”, en la cual se utilizan datos georreferenciados de usos y coberturas de suelo (denominados áreas de entrenamiento o de referencia) para entrenar algoritmos y generar modelos que permitan predecir la cobertura del resto de los píxeles del área de interés (Paruelo 2014). De esta manera es posible obtener mapas de cobertura con los cuales estimar la superficie agrícola.

Una limitante para la adopción de esta práctica a gran escala es la ausencia de datos de terreno (comúnmente llamada “verdad de campo”) que sean fiables respecto a las coberturas que declaran y que cumplan con algunos requisitos metodológicos recomendados al momento del muestreo. Olofsson et al. (2014) plantean que es necesario asegurar la aleatorización del proceso de selección de los puntos de muestreo, y recomienda por eso mismo, que los datos provengan de un muestreo probabilístico. Un muestreo probabilístico cumple dos requisitos: en primer lugar, debe conocerse la probabilidad de inclusión de cada unidad de muestreo y, en segundo lugar, la probabilidad de inclusión de cada unidad debe ser mayor que cero. Este prerrequisito permite realizar inferencias y generalizaciones utilizando los datos de muestreo para obtener así los parámetros poblacionales; además, permite conocer la incertidumbre de los estimadores de la muestra utilizada. Dichos autores recomiendan, por este motivo, utilizar estimadores muestrales que no sean sesgados. En Argentina desde el año 2013, se realiza bajo la órbita del Ministerio de Agricultura, Ganadería y Pesca de la Nación un muestreo de cultivos a nivel nacional llamado

“Método de Segmentos Aleatorios (MSA)”. Esta metodología consiste en un muestreo probabilístico estratificado, mediante el cual se muestrean áreas permanentes (segmentos) de 400 hectáreas. Estas áreas son visitadas periódicamente por técnicos que constatan los cultivos presentes en cada segmento y el porcentaje de ocupación de cada cultivo (Ministerio de Agricultura; Circa 2016). Estos datos, si bien relevan poca superficie porcentual del territorio agrícola, resultan de gran utilidad para el uso combinado con la información satelital que brinda cobertura completa del territorio con una muy buena resolución espacial y temporal, lo que hace posible obtener, de esta manera, una estimación total de la superficie de cada cultivo (Paruelo 2014; MAGYP 2016).

Históricamente, el costo de las imágenes, la limitación de los equipos informáticos para el pre-procesamiento (corrección geométrica, radiométrica y atmosférica (Jensen 1996; Toutin 2004; Lu & Weng 2007) y de los soportes de almacenamiento que se sumaban a la limitación de la ausencia de verdad terrestre mencionada anteriormente, hacían de las clasificaciones con sensores remotos una herramienta limitada (Hey et al. 2009): se realizaban para pocos años, con pocas imágenes o escenas, para una porción acotada del territorio y/o presentaban una baja resolución conceptual (cantidad de clases o categorías discriminadas). A modo de ejemplo, se puede mencionar el trabajo realizado por Guerschman et al. (2003), quienes llevaron a cabo una clasificación de tipos de cultivos para la campaña 1996/1997 del sur de Buenos Aires, usando las bandas espectrales de Landsat 5 de una escena en 4 fechas representativas a lo largo de la ventana temporal de la campaña agrícola. En esos mismos años, el Laboratorio de Análisis Regional y Teledetección (LART), junto con el Ministerio de Asuntos Agrarios, llevaron adelante una clasificación para la campaña 2003/2004 en la provincia de Buenos Aires utilizando el satélite SAC-C (disponible en <https://ced.agro.uba.ar/lart/areas.html>). Ambas clasificaciones, si bien difirieron en su extensión espacial, permitieron separar a grosso modo cultivos de verano, cultivos de invierno, doble cultivo y recursos forrajeros.

En los últimos años, la disponibilidad gratuita de imágenes de alta resolución espacial, el aumento en la capacidad de procesamiento de los equipos informáticos, y la disponibilidad de plataformas de procesamiento de datos en la nube, permitieron disponer de información georreferenciada sin que sea necesaria una gran infraestructura de procesamiento o de

almacenamiento de datos. Esto permitió, por ejemplo, llevar adelante la elaboración del Mapa Nacional de Cultivos por parte del INTA para las últimas tres campañas agrícolas (2018/2019, 2019/2020 y 2020/2021), a partir de imágenes Landsat 8 y herramientas de procesamiento en la nube (de Abelleira et al. 2019, 2020, 2021). Las clasificaciones se realizaron para cada una de las zonas PAS definidas por la Bolsa de cereales de Buenos Aires. El algoritmo de clasificación utilizado fue Random Forest, el cual se entrenó con información de índices de frecuencia y dinámica temporal del índice de vegetación. Este producto, si bien presenta una buena caracterización de los cultivos de verano, no es capaz de distinguir entre cultivos de invierno. En particular, en las primeras ediciones, para la porción sur de la provincia de Buenos Aires, no se indican los errores de omisión y de comisión para todas las clases, lo cual fue corregido en la edición 2020/2021 (de Abelleira et al., 2019, 2020, 2021).

A nivel mundial, existen diversos países que disponen de inventarios nacionales de cultivos basado en clasificaciones supervisadas como, por ejemplo, Canadá, y Estados Unidos. El mapa de Estados Unidos se realiza con cobertura nacional desde el 2008, con una resolución de 30 metros, utiliza imágenes Landsat 8, ISRO ResourceSat-2 LISS-3, y Sentinel 2, el algoritmo utilizado está basado en el uso de árboles de decisión y la verdad de campo corresponde a parcelas permanentes (CLU) que son visitadas por productores (https://www.nass.usda.gov/Research_and_Science/Cropland/sarsfaqs2.php#Section3_2.0). El mapa nacional de Canadá se realiza desde el 2009, también utiliza un árbol de decisión, y se basa en imágenes ópticas (Landsat-8, Sentinel2) y de radar (Radarsat-2), mientras que la verdad de campo es provista por las empresas aseguradoras agrícolas provinciales (<https://open.canada.ca/data/en/dataset/ba2645d5-4458-414d-b196-6303ac06c1c9>).

En este trabajo se buscó hacer un aporte a esta disciplina aprovechando las nuevas tecnologías disponibles (Big Data y procesamiento en la nube) (Gorelick et al. 2017), realizando un mapa de usos y coberturas del suelo a partir de combinar datos provenientes de un muestreo probabilístico con información provista por sensores remotos. A su vez, se buscó evaluar el algoritmo de clasificación utilizando distintos volúmenes de datos. El trabajo se enfocó en 3 departamentos ubicados en una región históricamente triguera del sudeste de la provincia de Buenos Aires en la campaña 2015/2016. La verdad de campo se muestreó de 3 formas distintas

para entrenar el algoritmo de clasificación. Se propuso, además, trabajar con dos aproximaciones: 1) entrenando un algoritmo para toda la campaña agrícola, generando un mapa que comprende el uso anual, y 2) entrenando el algoritmo para cada subcampaña, obteniendo así dos mapas parciales, uno para los cultivos de invierno y otro para los cultivos de verano, y posteriormente un mapa final anual resultante de la combinación de ambos. Para ambas aproximaciones se estimaron las superficies promedio de cada uso del suelo acompañado de un intervalo de confianza del 95%.

2. Materiales y métodos

2.1. Área de estudio

El área de estudio comprendió los departamentos bonaerenses de Tres Arroyos, San Cayetano y Necochea (Figura 1), los cuales se encuentran comprendidos en el complejo Pampa Interserrana de la Ecorregión Pampeana (Matteucci 2012), y dentro del Distrito Pampa Austral (Cabrera 1976). Los tres departamentos cubren un total de 13.657 Km². La región presenta un clima templado húmedo con una precipitación media anual de 700 mm bajo un régimen isohigro, y tiene una temperatura media anual de 14°C. Los suelos son más elevados en comparación con los del resto de la región pampeana, y pueden tener una capa de tosca a baja profundidad, o presentar rocas en superficie. La vegetación natural de la región conforma una pseudo estepa de mesófitas con predominancia del estrato arbustivo en los bordes serranos y cerros rocosos (Oyarzabal et al. 2018). Respecto al perfil productivo agrícola es una región de explotación mixta con aproximadamente el 60% de la superficie dedicada a la producción exclusivamente agrícola (Mosciaro 2009). Los cultivos principales son trigo, girasol, maíz y cebada cervecera en la parte central y norte de San Cayetano y Tres Arroyos (zona agrícola homogénea Tres Arroyos), y trigo y girasol en la parte costera de esas localidades (zona agrícola Orense). Respecto al departamento de Necochea, en la parte centro norte (zona agrícola homogénea Lobería) se siembra trigo, girasol, maíz y soja, compartiendo un uso similar a la parte sur la cual pertenece a la zona agrícola homogénea Necochea (Mosciaro 2009). En todo el distrito Austral se

observó un aumento en la superficie sembrada con cereales y oleaginosas de 37% en el período 1995–2008 (Mosciaro 2009). Los cultivos que más crecieron en superficie relativa fueron soja (de un 2% a un 37%) y cebada cervecera, la cual se incrementó un 180% (Mosciaro 2009).

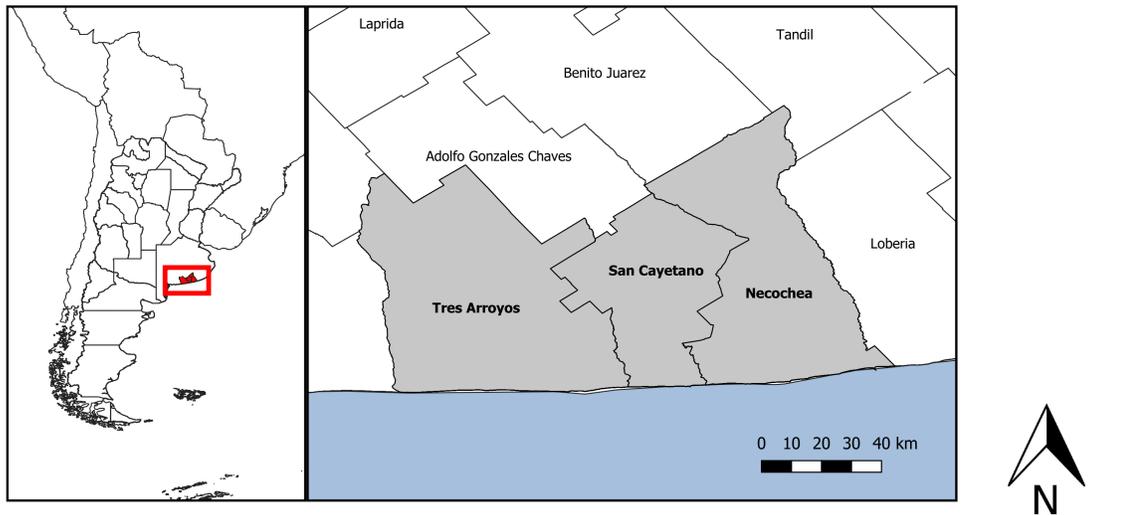


Figura 1: Mapa del área de estudio, la cual cubre los partidos de Tres Arroyos, San Cayetano y Necochea, ubicados en el sudeste de la Provincia de Buenos Aires.

2.2 Clasificaciones de uso del suelo

2.2.1 Procesamiento de la información de campo

Se contó con un total de 1437 muestras (lotes digitalizados) correspondientes al relevamiento a campo realizado por el MAGYP para la campaña 2015/2016. El área representada por los lotes del muestreo del MAGYP corresponde a un 2,36% de la superficie total del área de estudio (alrededor de 32 mil ha) (Figura 2). Del total, 417 muestras correspondieron a Necochea, 356 a San Cayetano y 664 a Tres arroyos. Estos polígonos brindan información de los usos y coberturas presentes en dos momentos del año: durante la campaña de invierno, y durante la campaña verano. De total, fueron descartadas 7 muestras clasificadas como NR (No Relevado), asociado a la falta de información, ya sea por imposibilidad de acceso o por omisión, dado que tampoco disponían de información complementaria que diera cuenta del uso del suelo en esa

campana (MAGYP 2016). También se eliminaron aquellas muestras que tuvieran información de cobertura temporalmente inconsistente como, por ejemplo, cultivos de invierno durante el relevamiento de verano, lotes con pastizal y campo natural luego de un cultivo anual sin haber pasado por una instancia de suelo desnudo o barbecho (sumando un total de 35). Posteriormente al filtrado se obtuvieron 1395 muestras totales: 399 para Necochea, 340 para San Cayetano, y 656 para Tres Arroyos.

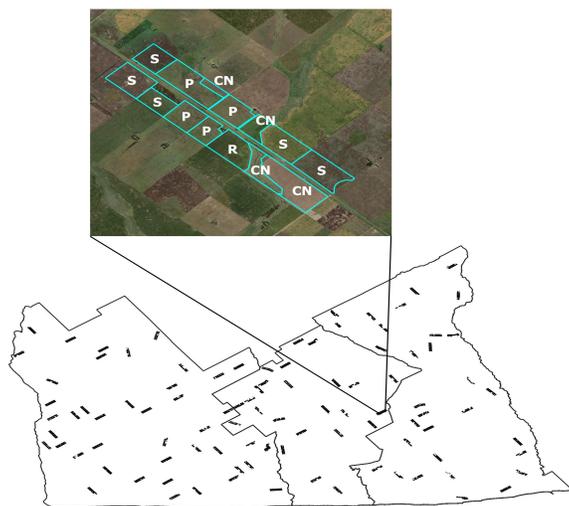


Figura 2: Distribución de los lotes muestrales del Método de Segmentos Aleatorios del MAGYP. Se muestra el detalle de un lote de 4x2 km con una imagen de muy alta resolución.

Los datos previamente filtrados permitieron discriminar tipos de uso (cultivos y coberturas) dentro de cada sub-campaña (invierno y verano) y, posteriormente como una combinación de los anteriores, se clasificaron los usos del suelo a escala anual. Así, se discriminaron 13 rotaciones anuales, 8 coberturas de invierno y 8 coberturas de verano (Tabla 1). Estas clases se eligieron por su representatividad, ya sea debido a la cantidad de muestras presentes en el área de estudio, como también por la validez conceptual de dicha clase, es decir: si la cobertura relevada coincidía con el uso y el aprovechamiento típico para la zona. En algunos casos se unificaron clases en una categoría con una menor resolución conceptual, como en el caso

de las clases Otro Cultivos de Invierno, la cual contenía los lotes cultivados con centeno, alpiste y lino; Otro Cultivo de Verano, donde se ubican los lotes con sorgo granífero; y la rotación Otro Doble Cultivo, que contenía diversas combinaciones de las dos anteriores. La clase Sin Cobertura engloba todos los lotes con barbecho, suelo desnudo y rastrojo.

Leyenda		
Campaña Anual	Campaña Invierno	Campaña Verano
Soja		Soja
Pastura	Pastura	Pastura
Cebada -Soja	Cebada	Soja
Maiz		Maíz
Trigo - Soja	Trigo	Soja
Sin Cobertura	Sin Cobertura	Sin Cobertura
Trigo	Trigo	
Campo Natural	Campo Natural	Campo Natural
Otro Doble Cultivo	Otro Cultivo Invierno	Otro Cultivo Verano
Girasol		Girasol
Verdeo	Verdeo	Verdeo
Cebada	Cebada	
Avena	Avena	

Tabla 1: Leyenda definida a partir de la información de los relevamientos a campo brindados por el Método de Segmentos Aleatorios (MAGYP 2016).

2.2.2 Procesamiento de imágenes satelitales y clasificación supervisada

Se utilizaron imágenes Landsat 8 de reflectancia de superficie (Colección LANDSAT/LC08/C02/T1_L2) correspondientes al periodo agosto 2015-mayo 2016, a las cuales se accedió a través de la plataforma Google Earth Engine (Gorelick et al. 2017). Se utilizó todo el espectro de reflectancia registrado por los sensores del satélite, tanto en el espectro visible, como en el infrarrojo cercano e infrarrojo medio, siendo las bandas utilizadas la B2 (blue), B3 (green),

B4 (red), B5 (Nir), B6 (SWIR 1) y B7 (SWIR 2). A su vez, se utilizó la banda *QA_PIXEL* para filtrar o descartar aquellos píxeles que presentaban mala calidad asociada a la presencia de nubes, sombras o aerosoles.. A partir de todas las imágenes filtradas resultantes se obtuvieron imágenes conformadas con el valor de la mediana de cada píxel para cada banda espectral para distintas ventanas temporales: Primavera (del 15/08/2015 al 30/11/2015), Verano (del 01/12/2015 al 20/02/2016) y Verano Tardío (21/02/2016 a 15/05/2016) de las cuales resultaba una colección de imágenes de mediana de todas las bandas anteriormente mencionadas. Las imágenes de mediana resultantes se eligieron o combinaron según la ventana temporal deseada para la clasificación, es así que para la Campaña anual se combinaron las imágenes de las 3 medianas (Invierno, Verano y Verano tardío). Las clasificaciones para la campaña de invierno y de verano utilizaron sus respectivas medianas de primavera, verano y verano tardío. Las fechas resultantes para cada clasificación fueron: Campaña Anual del 15/08/2015 al 15/05/2016, Campaña de invierno del 15/08/2015 al 30/11/2015, y Campaña de verano del 01/12/2015 al 15/05/2016.

Con la intención de poner a prueba el algoritmo de clasificación elegido, se construyeron 3 conjuntos de datos para cada ventana temporal (invierno, verano y anual), utilizando los siguientes criterios: el primero se construyó utilizando la información de todos los píxeles de la imagen de mediana (de dicha ventana temporal) que coincidieran con los lotes del método de segmentos aleatorios, es decir el total de los valores de reflectancia de la verdad de campo para cada clase (denominado muestreo *Total*); en el segundo se utilizó la misma cantidad de píxeles en todas las clases (muestreo *Balanceado*), donde el valor utilizado se estableció con el total de píxeles de la clase menos representada (i.e.: la clase con menor cantidad de datos); y en el tercero se realizó un muestreo que contemple la variabilidad entre clases, pero evitando valores extremos, el rango utilizado fue de 300 a 1050 píxeles por clase (muestreo *Proporcional*). La cantidad de píxeles resultantes para cada muestreo se muestra en la Tabla 2.

	Anual			Invierno			Verano		
	Total	Balanceado	Proporcional	Total	Balanceado	Proporcional	Total	Balanceado	Proporcional
Soja	118913	2220	1050				193969	2220	1050
Pastura	52694	2220	624	52694	2220	529	52694	2220	497
Cebada-Soja	57475	2220	655						
maíz	21744	2220	425				21744	2220	376
trigo soja	17580	2220	399						
Sin cobertura	4683	2220	316	167286	2220	1050	66487	2220	551
trigo	23940	2220	440	41521	2220	479			
campo natural	8078	2220	338	8078	2220	327	8078	2220	323
Otro cultivo/ DC	17612	2220	399	17612	2220	370	17612	2220	360
girasol	21945	2220	427				21945	2220	377
verdeo	2220	2220	300	2220	2220	300	2220	2220	300
cebada	26464	2220	456	83939	2220	671			
avena	11399	2220	359	11399	2220	342			

Tabla 2: Cantidad de Pixeles Landsat utilizados como verdad terrestre para cada clase. Valores definidos a partir de información provista por el Método de Segmentos Aleatorios discriminada para los tres tipos de muestreo de datos utilizados. (MAGYP 2016).

El algoritmo utilizado para el proceso de clasificación fue Random Forest. Este algoritmo consiste en un ensamble de árboles de decisión que entrega como resultado final la clase modal del conjunto total de resultados (Shelestov et al. 2017). Cada árbol evalúa el set de datos y entrega un valor de clase para cada pixel, del total de los resultados otorgados, se obtiene el valor del pixel clasificado final utilizando el valor de que tuvo la aparición más frecuente (i.e.: la moda). Una gran ventaja de esta forma de evaluación es la ausencia de sobreajuste, ya que los resultados convergen, pero, por otro lado, tiende a la generalización de los resultados (Breiman 2001). En este trabajo, con la intención de obtener para cada campaña (invierno, verano y anual) una clasificación que utilice cada uno de los 3 sub-sets de datos mencionados anteriormente (Balanceado, Proporcional, y Total), se entrenaron 9 algoritmos. En todos los casos, con el 70% de los puntos de muestreo del subset de datos, mientras que el conjunto de puntos restantes (30%) se reservó para la posterior evaluación de la clasificación obtenida. A todas las clasificaciones obtenidas se les aplicó un filtro de post-clasificación de mediana de 3x3 píxeles para eliminar el efecto sal y pimienta (Paruelo 2014). El flujo de trabajo de lo mencionado hasta aquí se muestra en la figura 3.

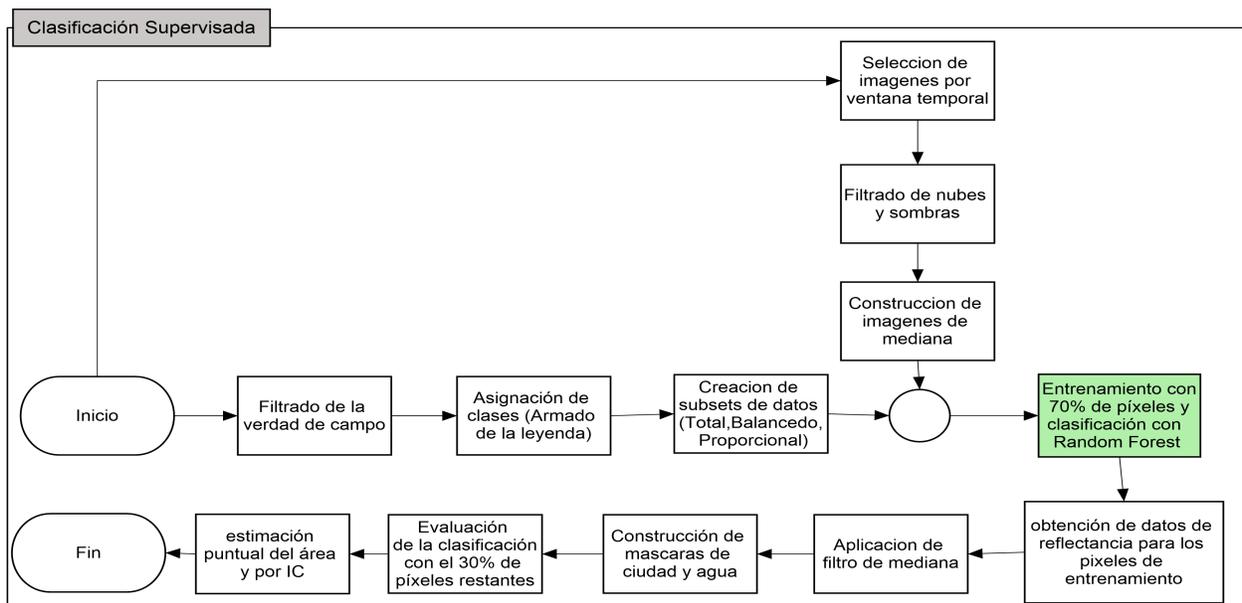


Figura 3: Flujo de trabajo para la obtención de las clasificaciones supervisadas para el Sudeste de la provincia de Buenos Aires en la campaña 2015/2016

Las categorías asociadas a áreas urbanas y cuerpos de agua no fueron incluidas dentro del proceso de clasificación supervisada, sino que fueron enmascaradas a partir de la utilización de diversos productos. Por un lado, se utilizó el producto MCD12Q1.006 MODIS Land Cover Type Yearly Global 500m para la identificación de zonas urbanas (correspondiente a la clase 13 de la banda LC_Type1). Por otro lado, se utilizó el producto JRC-Yearly History, el cual da cuenta de la presencia de agua a escala anual con una resolución espacial de 30 metros para la identificación de cursos y cuerpos de agua en el periodo clasificado.

Finalmente, con la intención de comparar dos metodologías de clasificación, se construyó una clasificación anual resultante de la sumatoria de las clasificaciones de la subcampana invierno y la subcampana verano. De esta clasificación se obtuvieron 64 clases que resultaron de combinar las 8 clases de uso del suelo de la campaña de invierno con las 8 clases de la campaña de verano. Estas clases se filtraron según su representatividad, validez conceptual y disponibilidad de verdad de campo para su validación. Luego del filtrado se obtuvieron 13 clases idénticas a las resultantes en la clasificación para la campaña anual (ver leyenda de Tabla 1) a las que se les agregó una clase llamada Otras Coberturas, donde se integraron las clases descartadas en el filtrado anterior (por ejemplo, combinaciones de: Cebada – Girasol, Avena – Soja, Trigo-Verdeo, etc).

2.2.3 Evaluación

Para evaluar la calidad de las clasificaciones realizadas, se realizó una matriz de confusión (Ver Anexo) para cada una de ellas, la cual permitió evaluar la precisión con la que fue identificada cada una de las clases discriminadas en los mapas resultantes. La matriz de confusión es una tabla cruzada de valores de campo con valores obtenidos en la clasificación, y que ubica en la diagonal de la tabla los aciertos del proceso, mientras que cualquiera de las otras ubicaciones dentro de la matriz se tratará de errores de asignación (Paruelo 2014). A partir de la misma se calculó la precisión general (obtenida como la sumatoria de la diagonal de la matriz – es decir, los aciertos de *todas* las clases- dividido el total de píxeles de validación), la precisión del usuario o error de comisión (calculada a partir de la sumatoria de los aciertos de cada clase

dividido por la cantidad total de píxeles de validación para dicha clase), y la precisión del productor o error de omisión (calculada como la sumatoria del total de aciertos de cada clase, dividido por la cantidad de píxeles de validación para dicha clase). La precisión del usuario o error de comisión indica la sobrestimación de los píxeles clasificados para una determinada clase, mientras que la precisión del productor o error de omisión indica la subestimación.

Para validar la clasificación obtenida por sumatoria de campaña invierno y campaña verano se unificaron los datos de validación de ambas clasificaciones. Esos datos fueron filtrados con la intención de dejar solo los valores únicos (sin repetición). Con este set de datos se construyó la matriz de confusión y se calcularon los indicadores mencionados en el párrafo anterior.

2.2.4 Estimación de Área

Se aplicó la metodología de Olofsson et al. (2014) para obtener una estimación insesgada del área o superficie de cada clase. Para ello, se combinó la información de las matrices de confusión anteriormente calculadas con la superficie relativa de cada una de las clases estimada de los mapas de clasificación resultantes por sumatoria de píxeles y la extrapolación a través del área individual de cada píxel Landsat (30 x 30 m). Estos valores de conteo de píxeles, se transformaron en valores de proporción de superficie muestreada respecto del área total de estudio. Contar con valores de proporción de área estimada por clase permite realizar una estimación puntual insesgada del área de cada clase y, a su vez, realizar una estimación por intervalo de confianza (utilizando, en este caso, un intervalo del 95%). Este procedimiento se realizó para la clasificación por campaña anual, y para la clasificación resultante de la suma de las clasificaciones para las campañas invierno y verano.

3. Resultados y discusión

Se observó, para todas las ventanas temporales, que el muestreo de puntos balanceado fue el que arrojó los mejores resultados en todas las clasificaciones; la clasificación con el total de puntos no arrojó resultados debido al error User Memory Exceeded de la plataforma Google

Earth Engine. Para la campaña anual, el muestreo balanceado obtuvo una precisión general (PG) de 0.89, frente a una precisión de 0.72 obtenida a través del muestreo proporcional (Tabla 3). Para la campaña de invierno, la precisión general fue de 0.77, frente a una precisión de 0.67 del muestreo proporcional (Tabla 4). Lo mismo sucedió con la campaña de verano, en donde se obtuvo un valor de 0.86 frente a 0.59 (Tabla 5). Respecto a las precisiones del productor y del usuario, comparando el promedio de todas las clases, se observó que para la clasificación anual el muestreo balanceado brindó valores promedios de 0.90 para precisiones del usuario (PU) y de 0.89 para precisiones del productor (PP), mientras que en el muestreo proporcional se observaron precisiones de 0.71 y 0.85, respectivamente (Tabla 3). Para la campaña de invierno los valores promedio de PP y PU para el muestreo balanceado fueron de 0,77 para ambas precisiones, frente a 0.61 y 0.68 del muestreo proporcional (Tabla 4). Para la campaña de verano, por su parte, las precisiones promedio fueron de 0.86 (tanto de PP y como de PU), superiores a los valores obtenidos de 0.68 y 0.75 en el muestreo proporcional (Tabla 5).

		Precisión general					
		Total	Balanceado	Proporcional			
			0,89	0,72			
		Precisión del productor		Precisión del usuario			
		Total	Balanceado	Proporcional	Total	Balanceado	Proporcional
Soja			0,86	0,93		0,86	0,79
Pastura			0,79	0,79		0,74	0,57
Cebada -Soja			0,89	0,84		0,83	0,65
Maiz			0,91	0,75		0,94	0,95
Trigo - Soja			0,91	0,80		0,91	0,80
Sin Cobertura			0,93	0,75		0,96	0,93
Trigo			0,89	0,73		0,94	0,84
Campo Natural			0,86	0,53		0,75	0,61
Otras Coberturas			0,83	0,50		0,95	0,93
Girasol			0,95	0,85		0,95	0,91
Verdeo			1,00	0,90		0,96	0,89
Cebada			0,90	0,79		0,94	0,79
Avena			0,87	0,58		0,93	0,86

Tabla 3. Valores de Precisión General, Precisión del Productor y Precisión del Usuario para la clasificación de la **campaña anual**, discriminado para los tres tipos de muestreo de los valores de entrenamiento.

Precisión general						
	Balanceado	Total	Proporcional			
	0,77		0,67			
Precisión del productor			Precisión del usuario			
	Balanceado	Proporcional	Total	Balanceado	Proporcional	Total
Trigo	0,77	0,63		0,71	0,63	
Cebada	0,68	0,75		0,64	0,60	
Avena	0,76	0,35		0,79	0,82	
Pastura	0,70	0,77		0,69	0,48	
Campo Natural	0,71	0,43		0,77	0,64	
Otro Doble Cultivo	0,66	0,28		0,85	0,70	
Verdeo	0,99	0,79		0,87	0,75	
Sin Cobertura	0,88	0,89		0,82	0,83	

Tabla 4: Valores de Precisión General, Precisión del Productor y Precisión del Usuario para la clasificación de la **campaña de invierno**, discriminado para los tres tipos de muestreo de los valores de entrenamiento.

Precisión general						
	Balanceado	Total	Proporcional			
	0,86		0,59			
Precisión del productor			Precisión del usuario			
	Balanceado	Total	Proporcional	Balanceado	Total	Proporcional
Soja	0,78		0,92	0,80		0,68
Maiz	0,93		0,64	0,87		0,76
Girasol	0,94		0,83	0,94		0,88
Pastura	0,75		0,67	0,75		0,56
Campo Natural	0,87		0,52	0,77		0,66
Otro	0,83		0,26	0,89		0,82
Verdeo	0,99		0,91	0,94		0,88
Sin Cobertura	0,76		0,65	0,90		0,76

Tabla 5: Valores de Precisión General, Precisión del Productor y Precisión del Usuario para la clasificación de la **campaña de verano**, discriminado para los tres tipos de muestreo de los valores de entrenamiento.

Estos resultados resultan promisorios si se los compara con los obtenidos por otros trabajos que abordan esta problemática para la misma región. Puede mencionarse el caso del Mapa Nacional de Cultivos realizado por el INTA, el cual realiza una clasificación para la misma zona de estudio (catalogada como PAS XII, siguiendo la metodología de la Bolsa de Cereales de Buenos Aires), y en la cual los resultados ofrecen mucha menor resolución conceptual en comparación con la que aborda este trabajo. El Mapa Nacional de Cultivos, en sus tres ediciones

(2018, 2019 y 2020), no discrimina entre cultivos de invierno, y otorgan resultados para una clase contenedora de dicha variabilidad llamada CI (Cultivo de invierno). Algo similar sucede con los cultivos de verano, para los cuales se construyeron clases que no contemplan si son de primera o segunda fecha de siembra y, por lo tanto, no contemplan la ocurrencia o no de algún cultivo antecesor dentro de la rotación. Con respecto a los indicadores obtenidos, el Mapa Nacional de Cultivos de la campaña 2018-2019 (de Abelleira et al. 2019) obtuvo una menor precisión general para la región que la que se obtuvo con la clasificación anual y el muestreo balanceado en este trabajo (0.85 vs 0.89). Respecto al Mapa Nacional de Cultivos 2019-2020 (de Abelleira et al. 2020), éste discrimina en cultivos de invierno y cultivos de verano, y si bien la exactitud general para los cultivos de invierno supera a la obtenida en este trabajo, no se indican valores de precisión del productor y del usuario discriminados por clase de cultivo. Respecto a la última versión realizada del MNC, se discriminan en este último los cultivos de verano, y se observa en general que las precisiones del productor y usuario son muy similares y en algunos casos levemente menores respecto a las obtenidas en este trabajo, sin embargo, se observó nuevamente que la precisión general del mapa nacional de cultivos fue superior a la aquí obtenida. Por otro lado, la Bolsa de Cereales de Buenos Aires, en conjunto con la Universidad de Maryland (Copati et al. 2021) también realizaron una clasificación para la zona PAS XII con el objetivo de discriminar áreas sembradas con Trigo y Cebada (agrupadas en una categoría llamada Winter Crops). La precisión general fue inferior a la obtenida en este trabajo (60% vs 77%) y no presentan resultados de precisión del productor ni del usuario.

Respecto a los patrones espaciales observados en el mapa para la clasificación anual, se observó una coincidencia con los representados en el mapa nacional de cultivos de la campaña 2018/2019, que fue la única edición que presentó un mapa para la campaña completa (representando los usos de invierno-verano en conjunto). Se observó que las áreas de ganadería del mapa nacional de cultivos coinciden con las clases verdeo, pastura y pastizal de la clasificación realizada en este trabajo. Además, los lotes agrícolas coinciden en su delimitación, aunque las coberturas pueden diferir entre ambas clasificaciones. Sin embargo, en términos generales se puede mencionar que el MNC es más homogéneo en lo que respecta a diversidad de cultivos, lo cual se debe a la no discriminación de los cultivos de invierno. También se observan

ciertos artefactos resultantes de la clasificación en dicho producto. Sumando las diferencias entre las precisiones obtenidas se considera que los resultados aquí obtenidos superan en algunos aspectos al MNC en su edición 2018-2019, como se mencionó anteriormente la PG del MNC fue de 0.85 vs 0.89 obtenido en este trabajo, la PU del MNC fue de 0.85 en promedio con muchos faltantes en algunas clases, en este trabajo la PU fue en promedio de 0.9. La PP presentó coincidencias respecto al valor, sin embargo, hay muchas clases no cuantificadas en el MNC, en ambos casos la PP fue cercana a 0.9.

Observando los mapas resultantes para cada subcampana, se observó en la campaña de invierno un predominio de la clase Sin Cobertura, la cual resultó superior a las demás con un área de 464 mil hectáreas. Le siguen en superficie las clases Pastura, Cebada y Trigo, con 197 mil hectáreas, 190 mil hectáreas, y 160 mil hectáreas, respectivamente. Esto se muestra en la Figura 4 y la Tabla 6, en esta última no están contempladas las áreas pertenecientes a las ciudades y los cuerpos de agua. Los valores de superficie mencionados son los obtenidos por el conteo de píxeles y la extrapolación a través del área individual de cada píxel. Respecto a la campaña de verano, se observó como clase mayoritaria al cultivo de soja con 446 mil hectáreas, el cual fue seguido por las clases Pastura y Sin Cobertura, con 229 mil hectáreas y 197 mil hectáreas, respectivamente (Figura 5 y Tabla 6). Del análisis de la matriz de confusión obtenida para la clasificación de la Campaña de invierno se observó que las clases Trigo y Cebada presentaron cierta confusión entre sí, al igual que las clases Otro Doble Cultivo y Cebada. Esto permitiría explicar los bajos valores de Precisión del Productor obtenidos en esas clases (Tabla 4).

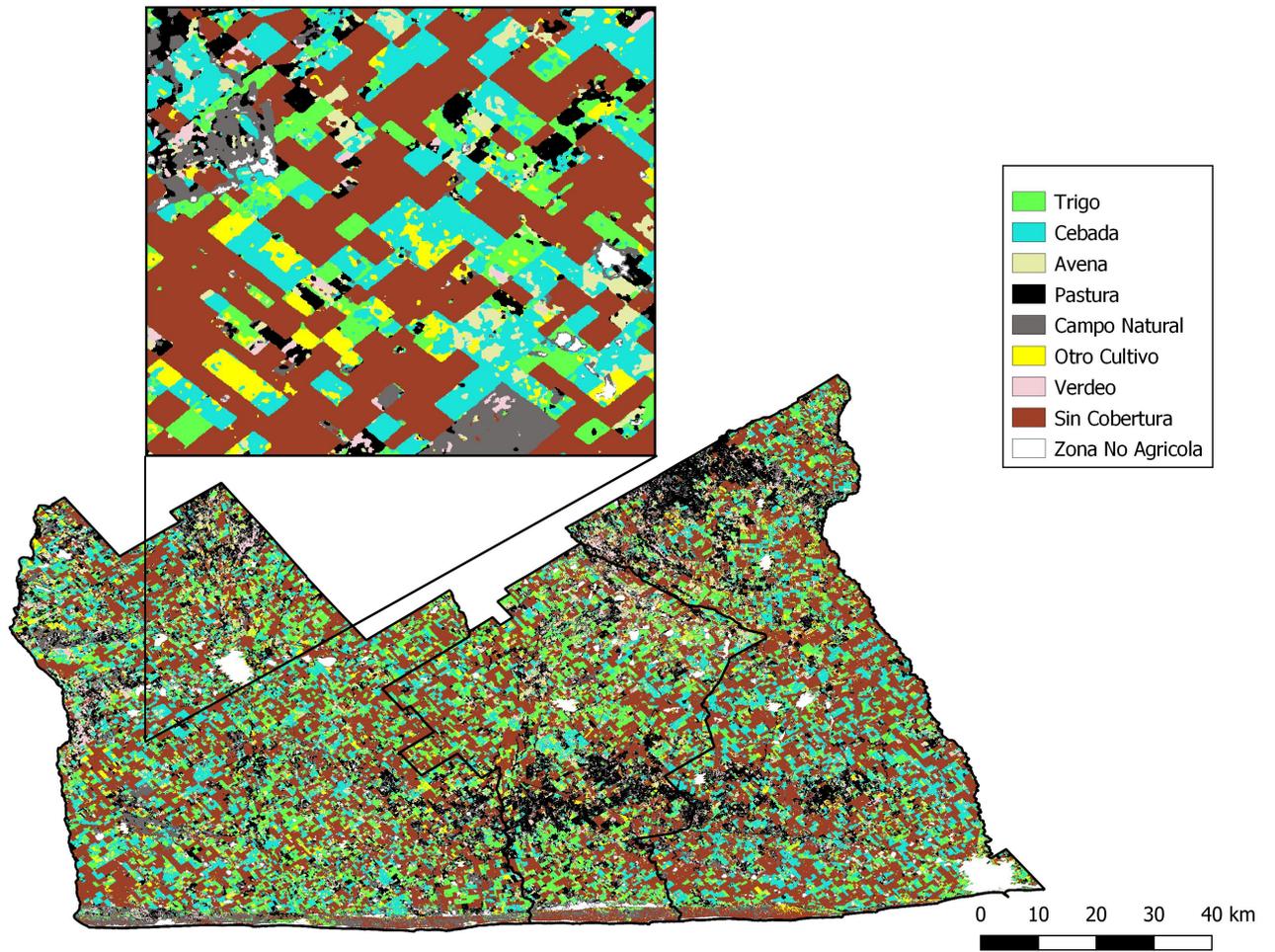


Figura 4: Clasificación Supervisada para la campaña de invierno resultante del muestreo balanceado para los departamentos de Tres arroyos, San Cayetano y Necochea (en ese orden). Se muestra un detalle (zoom) de 14 km (Oeste- Este) x 12 km (Norte-Sur).

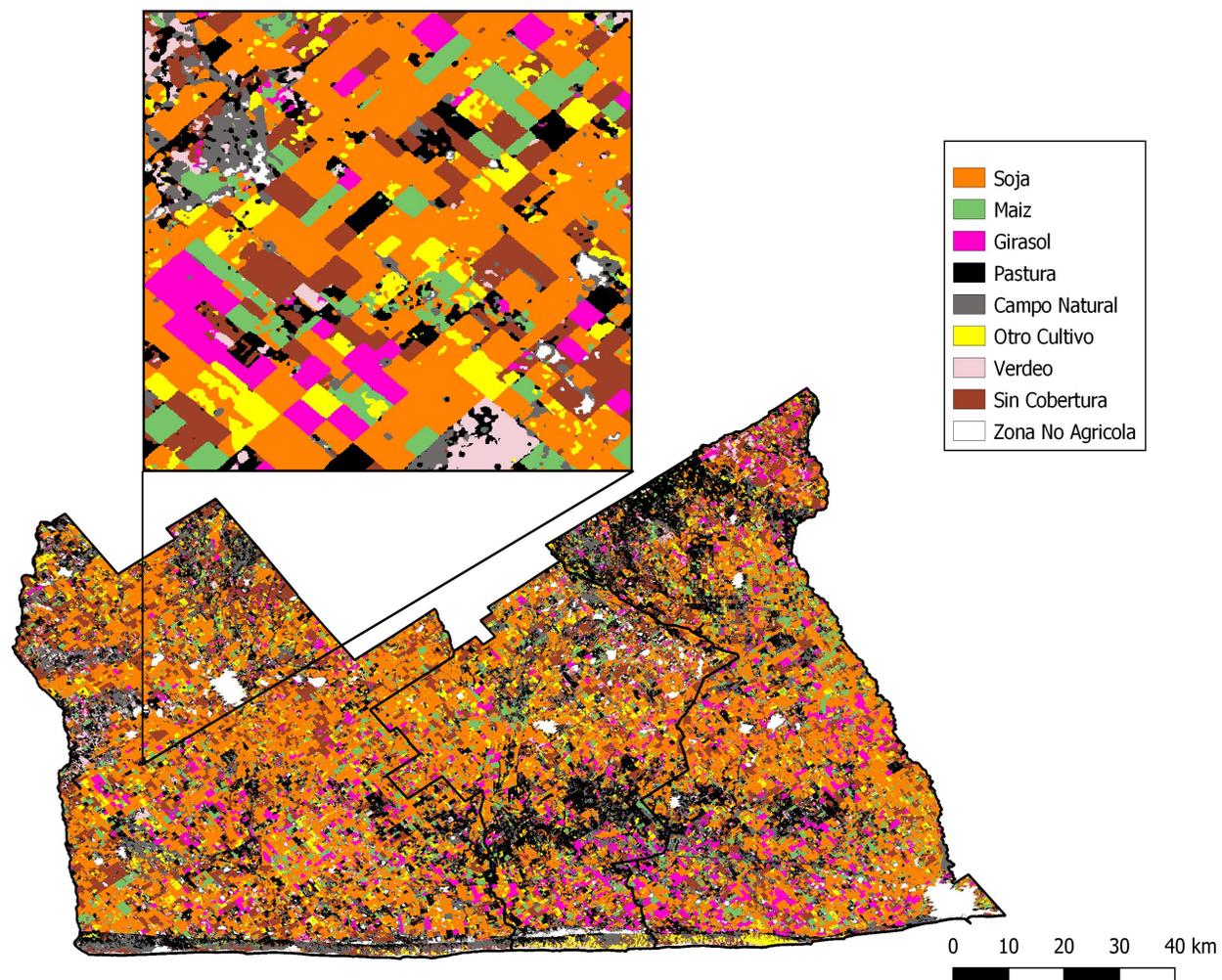


Figura 5: Clasificación Supervisada para la campaña de verano resultante del muestreo balanceado para los departamentos de Tres Arroyos, San Cayetano y Necochea (en ese orden). Se muestra un detalle (zoom) de 14 km (Oeste- Este) x 12 km (Norte-Sur).

Verano		Invierno	
Soja	445.923,00	Sin Cobertura	464.329
Pastura	229.404,00	Pastura	197.219
Sin Cobertura	197.101,00	Cebada	189.730
Campo Natural	138.410,00	Trigo	159.653
Maiz	95.281,00	Campo Natural	108.316
Girasol	86.947,00	Avena	73.547
Otro Cultivo	79.923,00	Otro Cultivo	65.300
Verdeo	28.265,00	Verdeo	43.312

Tabla 6: Superficie (en ha) representada en el mapa para cada clase. Datos obtenidos del conteo de píxeles LANDSAT en las clasificaciones de campaña de invierno y de verano. El área total clasificada asciende a 1.365.700 ha, siendo la superficie restante del área de estudio (64.446 ha) la correspondiente a las clases Agua y Urbano, enmascarados a partir del producto JRC y MODIS, respectivamente.

Comparando los mapas obtenidos para las rotaciones anuales con las dos aproximaciones planteadas por este trabajo, puede verse que la distribución de cultivos en el área de estudio fue muy similar en ambos casos (Figuras 6 y 7). Sin embargo, se observó el reemplazo de algunas coberturas por otras, como fue el caso de las clases Campo Natural y Pastura, que se vieron reemplazadas por la clase Verdeo en la clasificación resultante de la suma. Esto se ve reflejado en la Tabla 7, donde se observa un aumento del área ocupada por dicha clase. De la misma manera, se observó un reemplazo de varias clases nativas de la Clasificación Anual, por la clase Otras Coberturas, –la cual fue construida para eliminar combinaciones de cultivos poco probables o conceptualmente erróneas– mayormente las clases Cebada-Soja, Cebada, y Otro Doble Cultivo. En definitiva, el mapa resultante de la suma de las campañas de invierno y de verano resultó más homogéneo, debido a los pocos cambios en cobertura sufridos por las clases Soja y Girasol, y a la dilución de las otras clases mencionadas anteriormente. Comparando las superficies obtenidas por la metodología de Oloffson et al. (2014) en ambas clasificaciones (Tabla 7), se observó que las clases mayoritarias fueron las mismas en términos generales (Soja, Pastura, Campo Natural y

Cebada-Soja). Sin embargo, algunas clases sufrieron aumentos significativos (Otro Doble Cultivo, Verdeo y Avena), y otras disminuyen significativamente (Sin Cobertura y Cebada). Respecto a estos cambios es importante mencionar que las precisiones fueron ampliamente superiores en la clasificación Anual respecto a la clasificación resultante de la suma para todas las clases (Tabla 8).

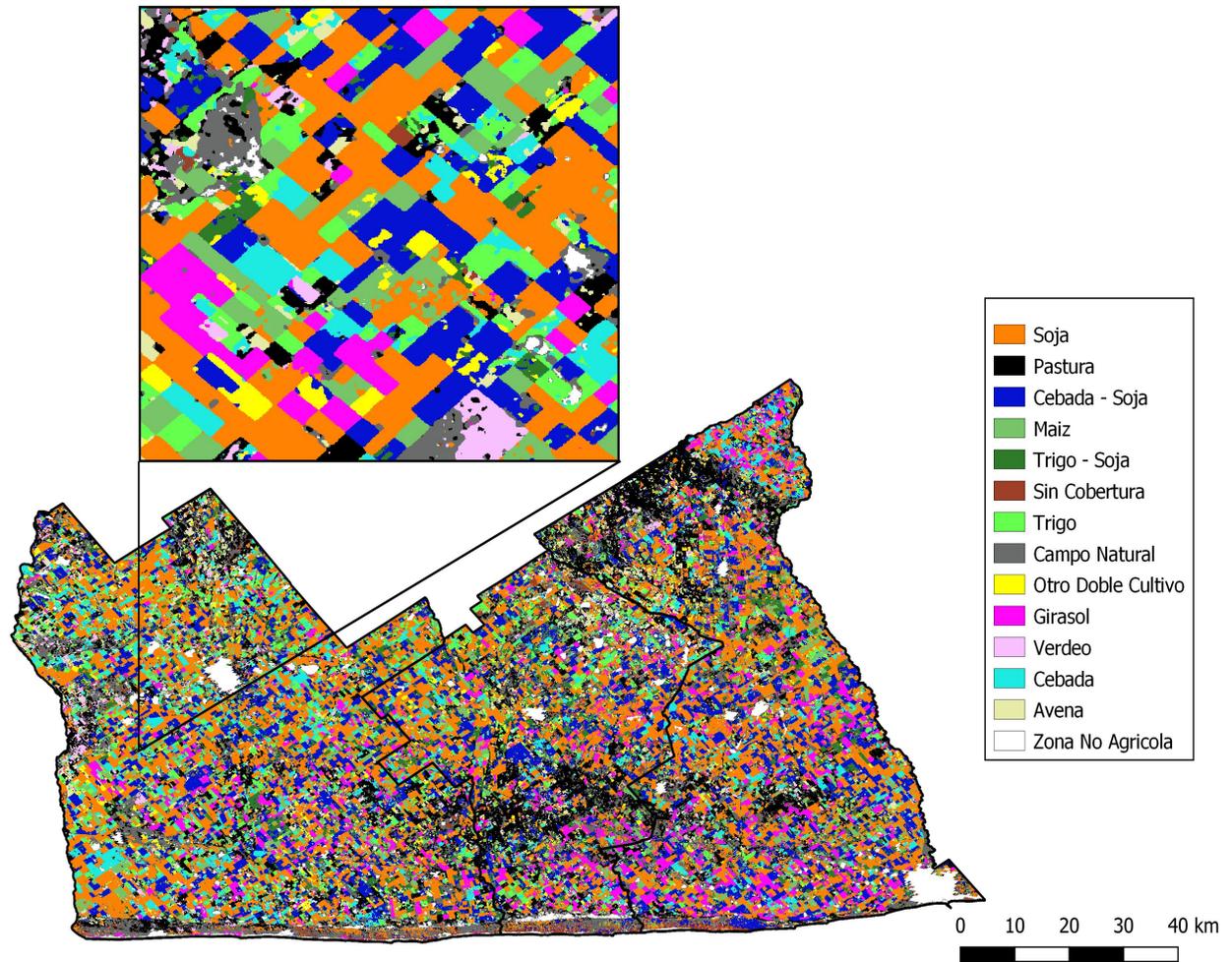


Figura 6: Mapa resultante de la clasificación para la Campaña Anual del muestreo balanceado para los departamentos de Tres Arroyos, San Cayetano y Necochea (en ese orden). Se muestra un detalle (zoom) de 14 km (Oeste- Este) x 12 km (Norte-Sur).

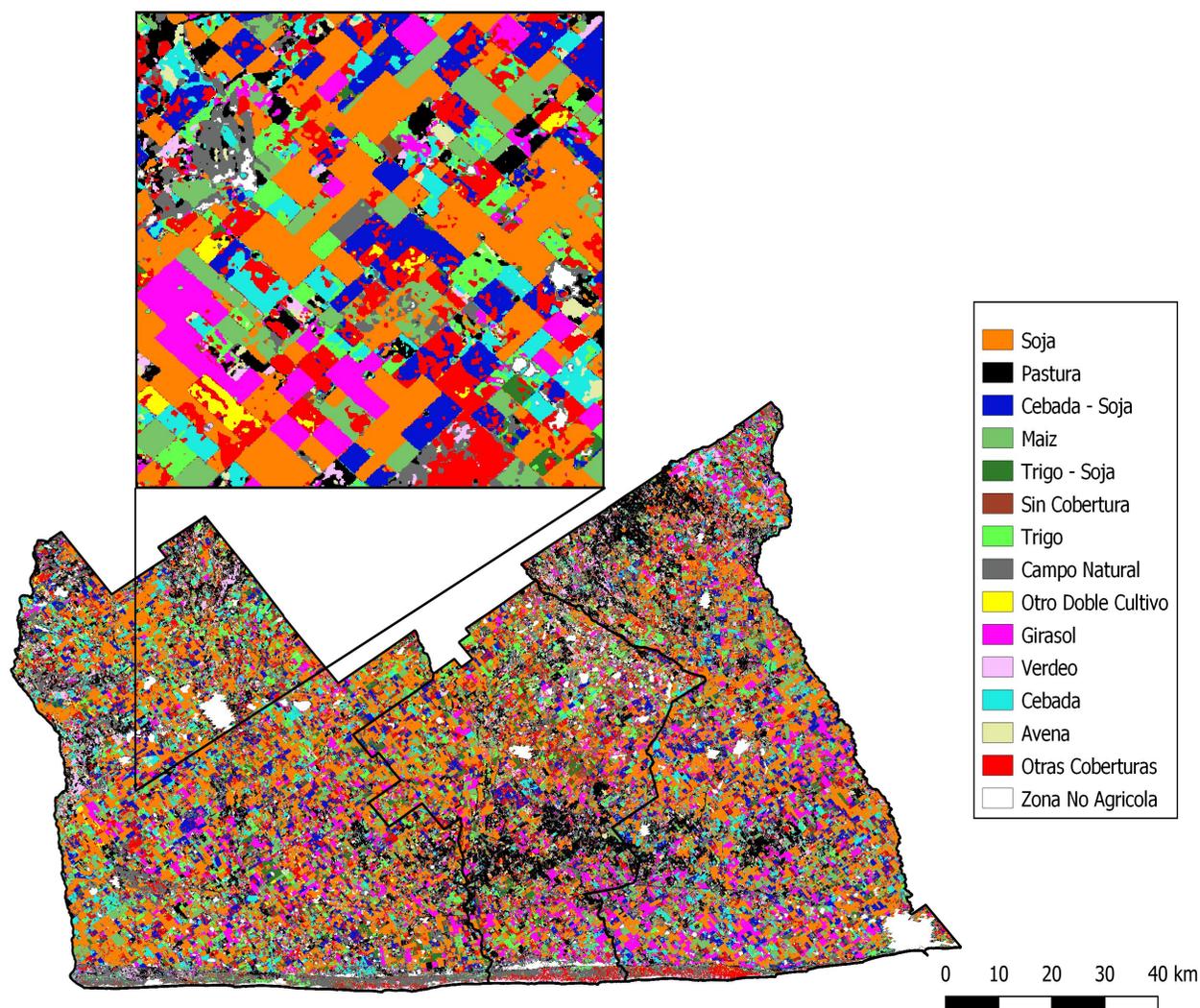


Figura 7: Mapa resultante de la suma de las clasificaciones para la Campaña de invierno y verano resultantes del muestreo balanceado. Se muestran de izquierda a derecha los departamentos de Tres Arroyos, San Cayetano y Necochea. Se muestra un detalle (zoom) de 14 km (Oeste- Este) x 12 km (Norte-Sur).

Clase	Anual			Suma			MINAGRI
	Area (ha)	LI	LS	Area (ha)	LI	LS	
Soja	288.248	279.553	296.943	248.638	240.128	257.148	566.400
Maiz	166.657	159.330	173.984	164.940	158.622	171.257	89.700
Sin Cobertura	158.599	152.469	164.729	111.238	105.876	116.600	
Otra Cobertura	79.959	74.785	85.133	81.507	77.693	85.320	
Verdeo	59.890	56.492	63.289	61.012	56.148	65.876	
Pastura	37.227	34.158	40.296	9.403	6.764	12.043	
Girasol	79.704	76.038	83.370	75.718	71.684	79.751	80.500
Campo Natural	120.637	114.272	127.002	138.607	132.990	144.225	
Trigo-Soja	54.542	49.730	59.354	82.066	78.648	85.485	
Trigo	88.541	84.636	92.446	94.155	90.225	98.085	120.400
Cebada	16.837	16.575	17.100	38.066	36.899	39.234	248.000
Cebada-Soja	86.316	82.669	89.963	65.275	61.286	69.264	
Avena	64.093	59.733	68.453	84.975	81.764	88.185	51.000
Otras Coberturas				44.925	40.033	49.818	

Tabla 7: Estimaciones de superficie para cada clase obtenidas con la metodología de Oloffson para ambas aproximaciones (Clasificación Anual y Suma de clasificaciones).

	Precisión general			
	Anual		Suma	
	85,66	67,49		
	Precisión del productor		Precisión del usuario	
	Anual	Suma	Anual	Suma
Soja	93,16	91,29	85,32	80,56
Pastura	96,21	75,65	93,18	73,55
Cebada-Soja	94,13	59,44	95,21	80,51
Maíz	87,83	84,57	91,13	81,82
Trigo-Soja	94,77	53,41	95,66	71,96
Sin Cobertura	89,21	51,99	91,18	23,73
Trigo	90,78	63,72	96,29	64,22
Campo Natural	86,12	70,41	93,55	77,09
Otro Doble Cultivo	89,51	10,44	93,44	98,04
Girasol	93,83	83,50	95,12	82,57
Verdeo	90,22	95,25	92,87	87,27
Cebada	91,52	63,78	95,31	52,18
Avena	91,13	37,00	96,70	74,25
Otras Coberturas		24,72		7,53

Tabla 8: Indicadores de precisión general, del productor y del usuario obtenidos por la metodología de Oloffson para la clasificación de Campaña Anual y la suma de las Campañas Invierno y Verano.

Para evaluar la capacidad predictiva del mapa obtenido en la clasificación anual, se compararon los valores de área estimada por esta clasificación con los valores de estimaciones agrícolas brindados por el Ministerio de Agricultura Ganadería y Pesca. Para la comparación con los datos ministeriales, los cuales tienen una menor resolución conceptual (Solo discrimina las clases Soja, Pastura, Girasol, Verdeo y Avena) hubo que unificar varias clases resultantes de la clasificación. Para esto se tuvieron en cuenta dentro de la superficie clasificada como cultivo de soja las clases: Soja, Cebada-Soja y Trigo-Soja. Algo similar se realizó para la estimación de la superficie de los cultivos de trigo y cebada, los cuales comprendieron a las clases Trigo y Trigo-Soja, y Cebada y Cebada-Soja, respectivamente. Se encontró que la mayoría de las estimaciones brindadas por el MINAGYP se encontraban fuera del intervalo de confianza estimado para dichos cultivos. El cultivo que presentó mayores diferencias fue el maíz, con una diferencia del 86% entre ambas aproximaciones. En este caso, el valor estimado mediante la metodología de Oloffson et al. (2014) resultó muy superior (166 mil ha) respecto a los datos públicos (90 mil ha aprox). Una diferencia de magnitud similar se observó para el cultivo de cebada, la cual se sobreestimó en un 58% aproximadamente (248 mil ha según el MINAGYP vs 103 mil ha estimadas mediante la metodología de Oloffson et al. (2014)). Esto se observó en menor medida para los cultivos de Soja, Avena y Trigo, los cuales difirieron en un 24, 25 y 18%, respectivamente, siendo sobreestimado el valor en el caso del cultivo de Soja y subestimado en los otros dos. Finalmente es importante resaltar el acierto en la estimación de superficie sembrada con Girasol, la cual cayó dentro del intervalo de confianza, con una desviación del 1% respecto a la media (Tabla 7).

Lo mencionado hasta aquí demuestra que la metodología de consulta a expertos suele diferir respecto a una estimación de superficie basadas en datos de sensores remotos; esto se suma a la ya mencionada ausencia de valores estadísticos presentados que den una idea de la incertidumbre con la cual se reportan estimaciones de superficie a escala poblacional. Esto plantea un cuestionamiento a esta metodología y lleva a plantearse la necesidad de modelos que adopten el aporte de datos satelitales. Fritz et. al (2019) mencionan distintos sistemas de monitoreo agrícolas y rescatan la importancia de los analistas (“analyst”) en sistemas híbridos que utilizan como input las estimaciones y recomendaciones de los expertos. En Argentina es

posible plantear un modelo de estas características teniendo en cuenta el relevamiento otorgado por el Método de Segmentos Aleatorios (MAGYP 2016), la disponibilidad de recursos capacitados para la metodología de consulta a expertos y la disponibilidad de información satelital abierta disponible.

4. Conclusiones

En este trabajo se plantearon distintas metodologías para estimar los usos y coberturas del suelo a partir de datos de campo provenientes de un muestreo estratificado. En este aspecto se observó que el uso de muestras balanceadas respecto a su presencia en el territorio para todos los cultivos que se deseaban estimar brindaban mejores resultados respecto al uso total de la información o aún de un muestreo que contemple la variabilidad.

Estos resultados obtenidos se exploraron con la intención de obtener distintas estimaciones y poder compararlas entre sí y con datos públicos provistos por instituciones gubernamentales. De estas estimaciones se rescata que las clases estimadas para la ventana temporal correspondiente a toda la campaña 2015/2016 ofrecieron precisiones y estimaciones de área de mejor calidad respecto a una aproximación por subcampaña que combina los resultados en un único mapa anual mediante la suma de ambas clasificaciones.

Finalmente, se rescata que los resultados obtenidos por esta metodología de trabajo, que combina la información satelital con relevamientos exhaustivos a campo (los cuales permiten caracterizar los usos y coberturas del suelo a escala de lote), permitirán avanzar sobre el análisis de diversos aspectos relacionados con el uso y la intensificación agrícola, a partir de identificar las secuencias o rotaciones de cultivos a lo largo de las campañas agrícolas. A su vez, brindaría información de interés para diversas instituciones públicas y privadas vinculadas con la planificación y la toma de decisiones a distintas escalas, tanto local, regional como nacional.

5. Bibliografía

Baeza. S. 2016. El uso/cobertura del suelo en Uruguay y los Pastizales del Rio de La Plata: Caracterización, análisis de sus cambios, a lo largo del tiempo e impactos sobre el funcionamiento ecosistémico. Tesis de Doctorado Ciencias Biológicas. PEDECIBA. Montevideo, Uruguay. 2016

Breiman L. 2001. Random Forests. *Machine Learning* 45, 5-32, pp 5-32. Public Academic Publishers. Netherlands

Cabrera, A. L.; 1976. Regiones fitogeográficas argentinas. Enciclopedia Argentina de Agricultura y Jardinería, Tomo II Fs. 1. Ed. ACME. Bs. As. Argentina. 1-85 pp

Cap E. J, Giancola S. I, Malach, V. 2010. Las limitaciones de las estadísticas ganaderas en los estudios de productividad en Argentina: Las encuestas a informantes calificados como fuente complementaria de datos. El caso de la provincia de San Luis. Instituto Nacional de Tecnología Agropecuaria. Trabajo de Investigación. Julio 2010

Cihlar J. 2000. Land cover mapping of large areas from satellites: status and research priorities. *Int. J. Remote Sensing*, 2000, Vol 21. Nro 6 & 7, 1093-1114.

Copati E., Humber M. L, Puricelli E., Sanchez A., Sahajpal R., Becker-Reshef I., O'Neill D. 2021. Winter crops coverages classification in the southeast of the Buenos Aires province, Argentina. Bolsa de Cereales de Buenos Aires.

Daily, G. C. 1997. Description and recording of vegetation upon a structural basis. *Ecology*, 32:172-229.

de Abelleira D, Banhero S, Verón S, Mosciaro J, Volante J. 2019. Mapa Nacional de Cultivos campaña 2018/2019. Colección 1. Versión 1. Informe. Instituto Nacional de Tecnología Agropecuaria (INTA), Argentina.

de Abelleira D, Banhero S, Verón S, Mosciaro J, Volante J. 2020. Mapa Nacional de Cultivos campaña 2019/2020. Colección 1. Versión 1. Informe. Instituto Nacional de Tecnología Agropecuaria (INTA), Argentina.

de Abelleira D, Banhero S, Verón S, Mosciaro J, Volante J. 2021. Mapa Nacional de Cultivos campaña 2020/2021. Colección 1. Versión 1. Informe. Instituto Nacional de Tecnología Agropecuaria (INTA), Argentina.

FAOSTAT 2013. Food and Agriculture Organization of the United Nations Statistical Databases 2013, Technical Report. Recopilado de Baeza 2016.

Fritz, T.; See, L.; Laso Bayas, J.C.; Waldner, F.; Jacques D.; Becker-Reshef, I.; Whitcraft, A.; Baruth, B.; Bonifacio, R.; Crutchfield, J.; Rembold, F.; Rojas, O.; Schucknecht, A.; Van der Velde, M.; Verdin, J.; Wu, B.; Yan, N.; You, L.; Gilliams, S.; Múcher, S.; Tetrault, R.; Moorthy, I.; McCallum, I. 2019. A comparison of global agricultural monitoring systems and current gaps, *Agricultural Systems*, Volume 168, Pages 258-272, <https://doi.org/10.1016/j.agsy.2018.05.010>.

Hansen, M.C.; Defries, R.; Townshend, J.R.G.; Sohlberg, R. 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*. 21: 1331-1364.

Hey, T. Tansley, S. Tolle, K. 2009. *The Fourth Paradigm: Data-Intensive Scientific Discovery*, 286 pp., Microsoft Res., Redmond, Wash.

IPCC (Intergovernmental Panel on Climate Change). 2000. *Special report on land use, land-use change, and forestry*. Cambridge University Press, Cambridge, UK.

Jensen, J.R., 1996, *Introduction to Digital Image Processing: A remote sensing perspective*, 2nd edn (Piscataway, NJ: Prentice Hall). Recopilado de Lu y Weng (2007).

Laboratorio de Analisis Regional y Teledetección (LART). 2005. *Estimación de Areas para la Provincia de Buenos Aires (2003-2004)*. Consultado de: <https://ced.agro.uba.ar/lart/areas.html>

Mcguire, A.D.; Sitch, S.; Clein, J.S.; Dargaville, R.; Esser, G.; Foley, J.; Heimann, M.; Joss, F.; Kaplan, J.; Kicklighter, D.W.; Meier, R.A.; Melillo, J.M.; Moore III, B.; Prentice, I.C.; Ramankutty, N.; Reichenau, T.; Schloss, A.; Tian, H.; Williams, L.J.; Wittenberg, Y. U. 2001. Carbon balance of the terrestrial biosphere in the twentieth century: Analyses of CO₂, climate and land use effects with four process-based ecosystem models. *Global Biogeochemical Cycles* 15: 183-206.

Leavy S, Allen B, Qüesta M, Piermatei C, Jones R. 2020. *Informes de Mercado, su uso en la Argentina*. Revista de investigaciones de la Facultad de Ciencias Agrarias. Nro 35. Universidad de Rosario UNR

Lu, D. & Weng, Q. 2007. A survey of image classification methods and techniques for improving classification performance, *International Journal of Remote Sensing*, 28:5, 823-870, DOI: 10.1080/01431160600746456

Matteucci S. D 2012. *Ecorregión Pampa. Ecorregiones y Complejos Ecosistémicos argentinos*. Cap. 12. Orientación Gráfica Editora. Buenos Aires 2012.

Mosciaro M., Dimuro V. 2009. Zonas Agroeconómicas Homogeneas Buenos Aires Sur. Estudios socioeconómicos de la sustentabilidad de los sistemas de producción y recursos naturales. Instituto Nacional de Tecnología Agropecuaria. Ciudad Autónoma de Buenos Aires 2009.

Ministerio de Agricultura Ganadería y Pesca (MAGYP). Circa 2016. Método de Segmentos Aleatorios: Metodología para la Estimación de la Superficie Sembrada con Cultivos Extensivos. Version V. Recuperado de https://www.magyp.gob.ar/sitio/areas/estimaciones/acerca_de/metodologia_archivos/000000_Metodo-de-segmentos-aleatorios-Version-5.pdf

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, 148, 42-57.

Oyarzabal M., Clavijo J., Oakley L., Biganzoli F., Tognetti P., Barberis I., Maturo H. M., Aragón R., Campanello P. I., Prado D., Oesterheld M., Leon R. J. C. 2018. Unidades de vegetación de la Argentina. *Revista Ecología Austral* 28: 040-063. Asociación Argentina de Ecología. Abril 2018.

Paruelo, J. 2014. Clasificación de Datos Espectrales. *Percepción Remota y Sistemas de Información Geográfica: Sus Aplicaciones en Agronomía y Ciencias Ambientales*, pp 161-185. Ed. Hemisferio Sur, 2014

Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., and Skakun S. 2017. Exploring Google Earth Engine Platform for Big Data Processing: Classification of Multi-Temporal Satellite Imagery for Crop Mapping. *Front. Earth Sci.* 5:17. doi: 10.3389/feart.2017.00017

Toutin, T., 2004, Geometric processing of remote sensing images: models, algorithms and methods. *International Journal of Remote Sensing*, 25, pp. 1893–1924. Recopilado de Lu y Weng (2007).

6. Anexo I: Matrices de confusión para las distintas clasificaciones

	Soja	Pastura	Cebada-Soja	maiz	trigo soja	Sin cobertura	trigo	campo natural	otro dc	girasol	verdeo	cebada	avena	Total Classified
Soja	592	3	10	24	4	4	6	8	10	12	0	6	6	685
Pastura	10	518	13	13	4	10	10	57	21	9	0	7	25	697
Cebada-Soja	12	8	596	5	19	1	9	5	25	0	0	20	18	718
maiz	19	1	3	623	0	2	4	4	3	3	0	2	1	665
trigo soja	9	0	17	3	591	0	8	0	16	0	0	3	1	648
Sin cobertura	14	3	0	1	1	667	3	3	1	0	0	1	1	695
trigo	1	4	2	1	7	3	599	0	12	0	0	7	2	638
campo natural	9	75	15	6	9	22	9	574	10	5	0	11	22	767
otro dc	11	1	2	3	0	1	5	5	552	0	0	2	1	583
girasol	11	2	3	1	2	0	4	1	10	631	0	1	0	666
verdeo	2	17	0	1	0	2	1	2	0	0	653	1	2	681
cebada	0	4	6	3	5	0	16	1	4	2	0	626	2	669
avena	0	16	1	4	4	2	2	4	1	3	0	7	549	593
Total Reference	690	652	668	688	646	714	676	664	665	665	653	694	630	6096

Tabla 1: Matriz de confusión para la clasificación anual.

	Soja	Pastura	Cebada- Soja	Maiz	Trigo- Soja	Sin Cobertura	Trigo	Campo Natural	Otro DC	Girasol	Verdeo	Cebada	Avena	Otras Coberturas	Total Classified
Soja	609	23	26	12	20	1	1	14	11	11	0	2	7	19	756
Pastura	8	809	7	11	2	7	23	131	15	6	6	18	35	22	1100
Cebada- Soja	4	1	347	0	25	2	1	8	2	1	0	9	14	17	431
Maiz	36	11	2	630	5	2	2	15	36	4	0	0	6	21	770
Trigo-Soja	6	1	43	0	195	0	13	0	5	0	0	1	3	4	271
Sin Cobertura	9	2	0	2	1	14	5	0	0	5	0	2	9	10	59
Trigo	1	4	6	1	37	2	368	20	5	1	3	83	16	26	573
Campo Natural	18	106	10	12	1	0	5	821	21	5	1	11	25	29	1065
Otro DC	0	0	2	1	0	0	3	0	702	3	0	0	0	5	716
Girasol	18	21	6	6	16	3	18	5	13	663	0	1	11	22	803
Verdeo	8	54	2	2	0	12	0	32	2	3	1138	5	12	34	1304
Cebada	4	32	40	2	4	0	64	23	8	4	0	251	40	9	481
Avena	2	61	2	1	1	5	12	36	2	2	0	9	447	22	602
Otras Coberturas	33	13	94	16	40	6	27	10	301	35	2	24	136	60	797
Total Reference	756		587	696	347	54	542	1115	1123	743	1150	416	761	300	9728

Tabla 2: Matriz de Confusión para la clasificación resultante de la suma de campañas.

	Soja	Maiz	Girasol	Pastura	Campo Natural	Otro Cultivo	Verdeo	Sin Cobertura	Total Classified
Soja	2881	150	10	10	70	40	51	114	3326
Maiz	148	3273	2	8	20	22	18	96	3587
Girasol	82	31	3616	8	16	25	17	85	3880
Pastura	74	17	14	3597	19	10	18	46	3795
Campo Natural	79	26	3	3	3510	14	17	33	3685
Verdeo	128	37	12	2	17	3489	22	82	3789
Otro Cultivo	118	44	4	3	10	10	3514	65	3768
Sin Cobertura	106	50	17	5	13	32	16	3106	3345
Total Reference	3616	3628	3678	3636	3675	3642	3673	3627	29175

Tabla 3: Matriz de confusión para la clasificación supervisada de la campaña de verano 2015/2016.

	Trigo	Cebada	Avena	Pastura	Campo Natural	Otro Cultivo	Verdeo	Sin Cobertura	Total Classified
Trigo	3392	29	14	19	29	15	15	72	3585
Cebada	5	3442	1	2	3	12	11	44	3520
Avena	9	10	3542	5	11	5	7	9	3598
Pastura	6	4	2	3548	3	1	4	26	3594
Campo Natural	15	6	8	22	3415	24	13	66	3569
Otro Cultivo	12	17	24	6	16	3410	7	44	3536
Verdeo	18	9	19	16	22	19	3425	42	3570
Sin Cobertura	0	0	0	0	0	0	0	2809	2809
Total Reference	3457	3517	3610	3618	3499	3486	3482	3112	27781

Tabla 4: Matriz de confusión para la clasificación supervisada de la campaña de invierno 2015/2016.