

Clasificación de cultivos en la provincia de Buenos Aires mediante la utilización de imágenes SAR e imágenes ópticas

Por Ing. Agr. Luciano Elias Di Paolo

Presentado ante la Facultad de Matemática, Astronomía y Física y el Instituto de Altos Estudios Espaciales Mario Gulich como parte de los requerimientos para la obtención del grado de

MAGÍSTER EN APLICACIONES ESPACIALES DE ALERTA Y RESPUESTA TEMPRANA A EMERGENCIAS

UNIVERSIDAD NACIONAL DE CÓRDOBA
Julio, 2017

©IG - CONAE/UNC 2017
©FaMAF - UNC 2017

Director: Agrim. Mauro J. Giovanniello
Agencia de Recaudación de la Provincia de Buenos Aires
ARBA

Co-director: Lic. Mario Lanfri
Instituto Gulich-CONAE



Esta obra está bajo una [Licencia Creative Commons Atribucion-NonCommercial-SinDerivar 2.5 Argentina](https://creativecommons.org/licenses/by-nc-nd/2.5/arg/).

Resumen

La tesis de maestría presenta tres aplicaciones obtenidas a partir de información satelital que son de interés de la administración fiscal de la Provincia de Buenos Aires:

1. La detección remota de cultivos y estimación de su superficie cultivada,
2. La clasificación supervisada de cultivos a través de imágenes satelitales ópticas
3. La utilización de imágenes SAR (Radar de Apertura Sintética) para clasificar cultivos.

Se utilizaron series temporales de imágenes SAR Cosmo SkyMed, Sentinel-1 A y Landsat 8 – OLI, para clasificar de manera supervisada cultivos de interés en la Provincia de Buenos Aires. Se probaron distintas combinaciones de imágenes SAR y Landsat 8 para clasificar cultivos. Se utilizaron los clasificadores de Máxima verosimilitud, Árboles de decisión (DT), “Random Forest”, “Gradient Boosted Tree”, “Support Vector Machine”, “Neural Network” para clasificar imágenes SAR con el objetivo de confeccionar mapas de cultivos en tres zonas de la provincia de Buenos Aires. Se obtuvieron precisiones de entre 89% y 92% en todas las zonas de estudio. Las clasificaciones sobre imágenes SAR obtuvieron mejores precisiones con clasificadores no paramétricos en dos de tres casos. El clasificador “Random Forest” presentó el mejor desempeño. Por último, se ha propuesto una metodología de trabajo para incorporar imágenes SAR a los productos cartográficos de la agencia de Recaudación de la provincia de Buenos Aires.

Abstract

This Magister thesis is about three major applications obtained from satellite information of special interest for the Revenue Administration of Buenos Aires Province, Argentina.

1. Remote assessment of crops: area and type.
2. Supervised classification of crops using optical images
3. The use of SAR images for crop assessment.

We performed supervised classification of temporal series of satellite images: Cosmo SkyMed, Sentinel-1A, and Landsat 8-OLI over three sites in Buenos Aires Province. We combined both SAR and Landsat images to achieve crop classification. Maximum Likelihood, Decision Trees Classifier, Random Forest, Gradient Boosted Trees, Support Vector Machine and Neural Network classifiers were used to classify crops. Overall accuracies of 89% to 92% were achieved for each region of interest. Classification of SAR images performed better with non-parametric classifiers in two of three regions. “Random Forest” classifier performed the best. Finally, we propose a work frame for including SAR images to cartographic products produced by the Revenue Agency of Buenos Aires Province (ARBA).

Palabras Clave:

Clasificación de Cultivos, SAR, series temporales, Sentinel-1A, Cosmo SkyMed, Landsat 8.

Agradecimientos

A la Agencia de Recaudación de la Provincia de Buenos Aires, al Instituto de altos estudios Mario Gulich, a la Comisión Nacional de Actividades Espaciales y a la Agencia Espacial Italiana, que han brindado el apoyo y los recursos necesarios para llevar adelante este trabajo.

Mis agradecimientos también, a Juan Arrigoni, Juan Eder, Mauro Giovanniello y Martin Ripa de ARBA; y a Gabriel Platzeck y Mario Lanfri del Instituto Gulich, CONAE.

Contenido

Resumen.....	2
Palabras Clave:.....	2
Agradecimientos	3
Contenido	4
Índice de Figuras	6
Índice de Tablas.....	7
Capítulo 1. Introducción & Antecedentes.....	8
1.1. Motivación	8
1.2. Antecedentes	9
1.3. Revisión conceptual	10
1.3.1. Clasificadores	13
1.4. Zonas de estudio	16
1.5. Objetivos	17
1.6. Hipótesis.....	18
1.7. Niveles de procesamiento imágenes Cosmo SkyMed.....	18
Capítulo 2. Cultivo de Cebolla en el Sur de la provincia de Buenos Aires.....	19
2.1. Introducción	19
2.2. Materiales & Métodos	19
2.2.1. Datos de campo.....	20
2.2.2. Imágenes SAR disponibles.....	21
2.2.3. Pre-procesamiento.....	21
2.2.4. Procesamiento	23
2.3. Resultados	25
2.4. Discusión	27
2.5. Conclusiones.....	27
Capítulo 3. Cultivo de papa en el Sudeste de la provincia de Buenos Aires	29
3.1. Introducción	29
3.2. Materiales y Métodos	30
3.2.1. Datos de campo.....	30
3.2.2. Imágenes SAR disponibles.....	32
3.2.3. Pre-procesamiento.....	33
3.2.4. Procesamiento	35
3.3. Resultados	36
3.4. Discusión	41
3.5. Conclusiones.....	41

Capítulo 4. Cultivo de granos en el Norte de la provincia de Buenos Aires.....	42
4.1. Introducción	42
4.2. Materiales y Métodos	42
4.2.1. Datos de campo.....	43
4.2.2. Imágenes SAR disponibles.....	44
4.2.3. Pre-procesamiento.....	45
4.2.4. Procesamiento	47
4.3. Resultados	47
4.4. Discusión	51
4.5. Conclusiones.....	52
Capítulo 5. Cultivo de granos en el Oeste de la provincia de Buenos Aires.....	54
5.1. Introducción	54
5.2. Materiales y Métodos	54
5.2.1. Datos de campo.....	55
5.2.2. Imágenes SAR disponibles.....	56
5.2.3. Pre-procesamiento.....	57
5.2.4. Procesamiento	58
5.3. Resultados	59
5.4. Discusión	62
5.5. Conclusiones.....	63
Capítulo 6. Discusión y Conclusiones	65
Anexos.....	70
1. Propuesta del nuevo sistema de detección de cultivos de ARBA. Gráfico de roles y tareas.....	70
2. Gráficos de retrodispersión por cobertura.	72
Boxplots clases Capítulo 3. Sudeste Provincia de Buenos Aires.	72
Boxplots clases Capítulo 4. Norte Provincia de Buenos Aires.....	75
Boxplots clases Capítulo 5. Oeste Provincia de Buenos Aires.....	77
3. Parámetros de los grafos de pre-procesamiento de las imágenes SAR.....	80
1. Figura 3. Imágenes Cosmo Himage SLC. Producto: Imágenes sin filtrar, georreferenciadas, a su máxima resolución espacial posible en píxeles cuadrados.	80
2. Figura 12. Imágenes Cosmo Himage SLC. Producto: Imágenes filtradas y georreferenciadas.	80
3. Figura 20. Imágenes Cosmo Wide Region GEC. Producto: Imágenes filtradas.....	80
4. Figura 21. Imágenes Cosmo Wide Region DGM. Producto: Imágenes georreferenciadas y filtradas.	81

5. Figura 31. Imágenes Sentinel 1 IW. Producto: Imágenes georreferenciadas y filtradas.
81

Bibliografía 82

Índice de Figuras

Figura 1. Zonas de estudio.	17
Figura 2. Puntos de observación de cultivos tomados en el mes de Febrero del 2015.....	20
Figura 3. Grafo de procesos aplicados a las imágenes CSK-Himage con el programa Nest.....	22
Figura 4. Izquierda: imagen CSK-Himage (2m resolución).	23
Figura 5. Vista de parvas de cebolla en Google Earth.....	24
Figura 6. Mapa detalle del área relevada. En marrón las superficies con cultivo de cebolla detectadas a campo.	25
Figura 7. Parcelas con cultivo de cebolla detectadas de forma remota,	26
Figura 8. Cultivo de papa sobre camellones. 11 Diciembre 2014.....	29
Figura 9. Cultivo de papa listo para cosechar. 4 de Marzo 2015.	30
Figura 10. Polígonos de clasificación o entrenamiento y de control.	31
Figura 11. Imagen Composición Color (R:12 Nov HV, G:11 Ene HV, B:22 Mar HV). Imágenes Cosmos Himage del sudeste de la provincia de Buenos Aires. Detalle.....	32
Figura 12. Grafo de procesos aplicados a las imágenes CSK-Himage con el programa Nest. Los parámetros se pueden ver en el Anexo 3.2.	34
Figura 13. Máxima Precisión alcanzada por las clasificaciones supervisadas sobre los distintos sets de imágenes utilizadas.....	36
Figura 14. Precisión Global en porcentaje de los clasificadores no paramétricos y de máxima verosimilitud sobre set #1 (cinco imágenes Cosmo Himage)..	36
Figura 15. Máxima precisión global de los clasificadores sobre set de índices texturales calculados sobre imágenes Cosmo Himage (set #6).	38
Figura 16. Clasificaciones del sudeste de Buenos Aires. Errores de Omisión y Comisión	39
Figura 17. Parcelas clasificadas con cultivo de papa detectadas de forma remota.....	40
Figura 18. Parcelas relevadas a campo en el norte de la Provincia de Buenos Aires durante la campaña agrícola 2014-2015.	43
Figura 19. Polígonos de muestreo sobre imagen compuesta RGB:11 Noviembre 2014 HV, 10 de Enero 2015 HV, 20 Marzo 2015 HV.....	44
Figura 20. Grafo de pre procesamiento de imágenes nivel "GEC" ..	46
Figura 21. Grafo de pre procesamiento de imágenes "DGM" ..	46
Figura 22. Clasificaciones zona Núcleo Prov. Buenos Aires. Sets.....	48
Figura 23. Clasificaciones zona Núcleo Prov. Buenos Aires. Clasificadores ..	48
Figura 24. Clasificaciones zona Núcleo Prov. Buenos Aires. Errores de Omisión y Comisión	50
Figura 25. Retrodispersión de los cultivos de Soja y Maíz para el mes de Enero en polarización HH y HV	50
Figura 26. Retrodispersión de los cultivos de Soja y Maíz para el mes de Noviembre en polarización HV y VV	51
Figura 27. Clasificaciones de cultivos de grano Norte de la Provincia de Buenos Aires.	52

Figura 28. Producto cartográfico. Clasificación de cultivos y la superficie calculada por cobertura.....	53
Figura 29. Polígonos que representan los lotes registrados a campo con las distintas coberturas de la "verdad de campo"	55
Figura 30. Polígonos de coberturas registradas a campo sobre imágenes multitemporales Sentinel1. Imágenes Compuestas RGB: 3 Noviembre 2014 VV, 7 Febrero 2015 VV, 26 Mayo 2015 VV.	56
Figura 31. Grafo de pre procesamiento de las imágenes Sentinel1.	58
Figura 32. Clasificaciones zona Oeste de Buenos Aires.Sets.....	59
Figura 33. Clasificaciones zona Oeste de Prov. Buenos Aires.Clasificadores.....	60
Figura 34. Clasificaciones zona Oeste de Prov. Buenos Aires. Errores de Omisión y Comisión.	62
Figura 35. Clasificaciones de cultivos en el Oeste de la Provincia de Buenos Aires. Campaña 2014-2015	63
Figura 36. Producto Cartográfico. Sentinel.1A.....	64

Índice de Tablas

Tabla 1. Imágenes CSK-Himage utilizadas.....	21
Tabla 2. Imágenes CSK-Himage recibidas del sudeste de la provincia de Buenos Aires.	32
Tabla 3. Errores de Omisión y Comisión de la clase "papa" para los tres mejores clasificadores.	39
Tabla 4. Imágenes Wide Region del norte de la provincia de Buenos Aires.	44
Tabla 5. Imágenes Sentinel-1 utilizadas para las clasificaciones supervisadas.	57
Tabla 6. Comparación de las superficies cultivadas estimadas mediante clasificaciones supervisadas.....	67

Capítulo 1. Introducción & Antecedentes

1.1. Motivación

El departamento de Tecnología de Imágenes de la Agencia de Recaudación de la Provincia de Buenos Aires, República Argentina, (ARBA) tiene por misión contribuir a la recaudación de impuestos provinciales de manera masiva, ayudar en la gestión de emergencias, cooperar con otros organismos públicos, y mejorar la gestión de los datos espaciales en todo el territorio de la Provincia de Buenos Aires. La tecnología desarrollada e implementada por el departamento ha sido compartida con distintos organismos, municipios, ministerios provinciales y nacionales. A su vez, la Agencia posee desde el año 2006 un convenio de provisión de imágenes satelitales y capacitación con la Comisión Nacional de Actividades Espaciales (CONAE). El citado convenio ha permitido la capacitación de agentes de ARBA en el Instituto Gulich-CONAE. La presente tesis de maestría está enmarcada en esa línea de cooperación ARBA-CONAE y pretende contribuir al desarrollo de tecnologías y procedimientos para mejorar la administración de la Provincia.

La actividad agropecuaria es de gran importancia para las economías de la Provincia de Buenos Aires y la República Argentina. En el territorio provincial, la agricultura se encuentra gravada con el impuesto a los ingresos brutos (IIBB) y constituye un aporte de relevancia a las cuentas provinciales. A su vez, existen otros impuestos provinciales asociados a la actividad agropecuaria como ser el impuesto a los sellos y el impuesto inmobiliario rural. Estos impuestos contribuyen a financiar la educación pública, el sistema de salud provincial, el mantenimiento del orden público y las instituciones provinciales. En el presente trabajo se intentará mejorar el sistema de detección satelital utilizado en el departamento de Tecnología de Imágenes de ARBA (“Manuales de detección remota de cultivos”, sistema “Índice Verde”, departamento Tecnología de Imágenes, Agencia de Recaudación de Buenos Aires, 2007) en tres sistemas de producción agrícola (cultivos de grano, cultivo de papa, cultivo de cebolla) de cuatro zonas de la Provincia de Buenos Aires (norte, sudeste y oeste de la provincia, y en el valle inferior del Rio Colorado) utilizando imágenes SAR y ópticas disponibles para la agencia. En la actualidad, el sistema de detección mencionado se basa en la clasificación de cultivos mediante el uso de una serie temporal de imágenes de índice verde normalizado extraídas de imágenes multiespectrales Landsat 7 y Landsat 8. Este proceso es altamente dependiente de las condiciones de nubosidad al momento de la toma de las imágenes. Las imágenes SAR permitirían salvar esta limitación contribuyendo a aumentar la solidez del sistema de detección remota.

La identificación de las parcelas fiscales donde se producen granos, papa, y cebolla, su superficie cultivada y la estimación de su producción permite el control fiscal de dichas actividades, pero además contribuiría al control del trabajo informal y las condiciones de vida del trabajador rural. Es por esto, que la presente tesis de maestría estará enfocada a la mejora en la detección y cuantificación de la producción de dichas actividades con la finalidad de mejorar los ingresos de la provincia de manera equitativa y presentar herramientas sólidas para el control de las condiciones laborales de los trabajadores rurales de la provincia.

1.2. Antecedentes

En el año 2011, el Área Rural del departamento de Tecnología de Imágenes (T.I.) de ARBA, presentó dos proyectos en el Anuncio de Oportunidad de la misión SAOCOM. Uno de ellos fue el primer intento de clasificación de cultivos utilizado por parte de la Agencia basado en imágenes SAR en banda X. El proyecto intitulado “Detección de cultivos de papa en la Provincia de Buenos Aires mediante imágenes Radar”, fue el primer contacto con las imágenes SAR por parte de la Agencia. El proyecto arrojó resultados positivos tanto en lo formativo para el grupo de trabajo como en los resultados de las clasificaciones supervisadas realizadas. Utilizando una serie de imágenes SAR Cosmo SkyMed, modo Wide Region, banda X, polarización VV, ángulo de incidencia: 40°, nivel de procesamiento GEC (por sus siglas en inglés, Geocoded ellipsoid Corrected), entre Noviembre de 2011 y Febrero de 2012 (6 imágenes CSK Wide Region), se logró una exactitud global de 73%, contra un 81% utilizando solamente imágenes L5 y L7 (4 imágenes), y 95% al combinar imágenes SAR y Landsat. Estos resultados fueron prometedores ya que mostraron que el uso de las imágenes SAR mejoró la clasificación supervisada realizada solamente con imágenes ópticas Landsat. A su vez, se encontraron ciertas complicaciones al utilizar imágenes Cosmo SkyMed en el nivel GEC, las cuales mostraron leves diferencias entre sí en su georreferencia, obligando a realizar una corrección. Esta primitiva investigación generó también otros interrogantes: ¿Sería mejor pre-procesar las imágenes SAR desde un nivel más bajo?, ¿Es VV la mejor polarización?, y por último, ¿Es el ángulo de incidencia determinante? Las respuestas a estas preguntas darían como resultado una configuración óptima del sensor para realizar la mejor clasificación posible utilizando imágenes SAR.

En el primer semestre del año 2014, el departamento de Tecnología de Imágenes de ARBA, realizó una investigación utilizando imágenes Cosmo SkyMed sobre un área del Sudeste de la provincia. En la misma, buscó responder algunas de esas preguntas. Para ello se requirieron imágenes Cosmo SkyMed Himage con distinto ángulo de incidencia en polarización HH y HV. El trabajo realizado dio como resultado dos productos, un informe intitulado “Multi-incidence angle SAR Images for monitoring agricultural practices in the wet pampas, Buenos Aires province” y una clasificación supervisada de los cultivos de papa del sudeste provincial utilizando imágenes SAR Cosmo SkyMed e imágenes Landsat 8. La investigación plasmada en el informe intentaba extraer de un set de imágenes con distinto ángulo de incidencia un índice de biomasa de cultivos, si bien no se llegó a ese resultado, se encontraron otras aplicaciones al uso de sets de imágenes multi-angulares. Una de las conclusiones del trabajo fue: encontrar más confiable el pedido de series temporales que el pedido de imágenes con distinto ángulo. La adquisición de imágenes de distinto ángulo entra en compromiso con el área de interés. Además, al solicitar imágenes con distinto ángulo se corre el riesgo de que entre una adquisición y la subsiguiente pasen varios días o semanas incrementándose el efecto temporal sobre los valores digitales de las imágenes. Todo esto fundó la idea de que la obtención de una serie temporal de imágenes era más sólida que el uso de distintos ángulos para la clasificación de cultivos, aunque para algunas aplicaciones el uso de un set multi-angular es preferible. En cuanto a las clasificaciones supervisadas realizadas (tipo : máxima verosimilitud), se obtuvo un porcentaje de exactitud global del 88% utilizando cuatro imágenes CSK de distinto ángulo de incidencia y tres imágenes L8, mientras que utilizando solamente L8 se obtuvo un porcentaje de 85%. Si bien a nivel global la diferencia es pequeña, cuando se analizan los errores de clasificación por cultivo, se observa por ejemplo que el error de comisión de la clasificación del cultivo de papa se reduce del 27% al 16 % por el agregado de las imágenes SAR. Esto permite suponer que la utilización de imágenes SAR ayuda a discriminar mejor ciertos cultivos donde por las condiciones de su sistema de producción generan un efecto en la retrodispersión que contribuye a su discriminación de

manera más efectiva. En el caso del cultivo de papa se supone que dicho efecto es producto de la contribución del riego y/o del suelo arado sobre el cual se desarrolla el cultivo.

Durante las investigaciones llevadas a cabo en el departamento de T.I. de ARBA, se comprobó empíricamente que el mejor nivel de procesamiento para el rápido y correcto despacho de productos es el nivel DGM, el cual permite mediante software libre (NEST, de la Agencia Espacial Europea) georreferenciar y corregir por terreno de manera más precisa. Las imágenes obtenidas se encontraban mejor georreferenciadas que aquellas obtenidas de los productos GEC o GTC (Geocoded Terrain Corrected).

1.3. Revisión conceptual

Los sensores activos, en general, y los sistemas SAR (por sus siglas en inglés, Synthetic Aperture Radar) en particular, disponen de su propia energía para sensar la superficie terrestre. La antena SAR transmite un pulso de energía electromagnética perpendicular a la dirección de avance del vehículo que transporta la antena (avión, satélite). El haz de ondas electromagnéticas impacta en un objeto o superficie, la energía interactúa con el objeto o superficie, la onda electromagnética es dispersada, y una parte de esa energía dispersada vuelve en dirección a la antena SAR que la transmitió. Esa misma antena registra la potencia recibida o retrodispersada. La variable “retrodispersión” puede ser expresada como la relación entre la energía recibida y la energía emitida, y se representa con la letra sigma (σ). Los sistemas de radar de observación terrestre SAR se caracterizan por la frecuencia de la onda electromagnética que transmiten y su capacidad para emitir y recibir en una o varias polarizaciones. Las frecuencias más utilizadas son: banda X: 8-12.5 GHz, con 2.4 a 3.8 cm de longitud de onda (ejemplo: Cosmo SkyMed); banda C: 4 a 8 GHz, con 3.8 a 7.5 cm de longitud de onda (ejemplo: Sentinel 1A y Sentinel 1B); banda L: 1 a 2 GHz, con 15 a 30 cm de longitud de onda (ejemplo: ALOS). La polarización de la onda electromagnética hace referencia al plano en el que oscilan las ondas del campo eléctrico con respecto a la dirección de propagación de la onda electromagnética. Si el campo eléctrico oscila en el plano vertical entonces la onda posee polarización vertical, si el campo eléctrico oscila en el plano horizontal entonces la onda posee polarización horizontal. Es importante destacar que los sistemas SAR pueden emitir y recibir pulsos electromagnéticos tanto en polarización vertical como en polarización horizontal. Las configuraciones más comunes son: HH: transmisión horizontal, recepción horizontal; VV: transmisión vertical, recepción vertical; HV: transmisión horizontal, recepción vertical; VH: transmisión vertical, recepción horizontal. La interacción de la onda transmitida con un objeto/cobertura terrestre puede transformar la polarización de la onda retrodispersada. Otro aspecto a tener en cuenta en los sistemas SAR es el ángulo de incidencia del pulso electromagnético. El ángulo formado por la dirección del pulso transmitido por la antena y la vertical del terreno donde incide la onda es llamado ángulo de incidencia, (Paruelo, Di Bella, Milkovic, Verón, de Abelleira, capítulo 5. 2014).

Los satélites SAR de observación terrestre presentan distintas configuraciones variables que pueden ser optimizadas según el objetivo del relevamiento o área de interés: Polarización, ángulo de incidencia, resolución espacial y resolución temporal. A esta lista puede agregarse el modo de adquisición : full-polarimétrico, multi-polarimétrico, o de polarización simple, (Kasischke, Melack, and Craig Dobson 1997). La diferencia entre full-polarimétrico y multi-polarimétrico reside en que el primero preserva la coherencia de fase entre los distintos canales permitiendo recrear posteriormente polarizaciones circulares o elípticas. La capacidad de adquirir imágenes full-polarimétricas sería de importancia, según Kasischke, para cuantificación de biomasa y para estimación de humedad de suelo. Sin embargo, en relación a

esta última aplicación el autor menciona un estudio donde se estimó humedad de suelo a partir de imágenes multi-polarimétricas HH/VV (Dubois, Van Zyl, and Engman 1995). Dubois (1995), logra estimar la humedad del suelo, tanto con sensores aerotransportados como satelitales, en banda L, pero solo para zonas sin vegetación o vegetación dispersa, el procedimiento requiere el enmascarado de píxeles con valores de NDVI superiores a 0.4. La generación de la máscara de vegetación puede lograrse, según el trabajo citado, utilizando el ratio HV/VV como indicador de vegetación. La relación que presenta entre el índice HV/VV y el NDVI es cercana a una lineal a partir de valores de NDVI de 0,2.

En el trabajo de Kasischke, citado anteriormente, el autor clasifica las aplicaciones de las imágenes SAR en ecología terrestre en cuatro grupos: 1) Clasificación del terreno (Land Cover), y 2) Cuantificación de biomasa 3) Monitoreo de inundaciones 4) Monitoreo de procesos dinámicos. El mencionado trabajo resulta interesante ya que presenta una tabla con las configuraciones óptimas que debería poseer un sistema SAR para cada tipo de aplicación. Para el mapeo de vegetación, concluye que la configuración deseable del sistema SAR sería multifrecuencia (por lo menos dos frecuencias), multipolarización o full-polarimétrico, ángulos de incidencia altos (40°-60°) y bajos (20°-40°), alta resolución espacial, y alta frecuencia temporal de muestreo (sobre todo para sistemas de una sola polarización). Esto significa, que el sistema SAR satelital ideal requerido para el mapeo de vegetación sería muy complejo, en la actualidad inexistente, ya que a la fecha no existe un sistema SAR satelital multifrecuencia. Ferrazoli (2002), analiza una serie de trabajos sobre la materia y propone la frecuencia y polarización más adecuadas para monitorear una serie de cultivos. Las conclusiones de su trabajo son importantes. Para los cultivos de papa, maíz y remolacha azucarera, soja y girasol la configuración ideal del sistema SAR sería banda L, polarización HV, con ángulos de incidencia altos; para el cultivo de colza canola, banda C, polarización HV; para los cultivos de trigo y cebada las bandas C y X mostraron atenuación de la señal retrodispersada en la polarización VV con ángulos de incidencia bajos (20°-30°) y en banda L mejoró la sensibilidad a la densidad del cultivo; para el cultivo de arroz, el mecanismo de backscatter varió durante el ciclo para bandas C y L, aumentando el efecto de “doble rebote” en la retrodispersión medida con el crecimiento del cultivo. Ferrazoli (2002) afirma que un incremento de la densidad de los tallos, que generalmente se encuentra asociado a un decremento en el diámetro de los mismos (ej. trigo, cebada, arroz, colza), lleva a un incremento en la frecuencia conveniente para su monitoreo: bandas C y X. Para cultivos de diámetro de tallo mayores como ser los cultivos de papa y maíz, frecuencias más bajas (banda L) serían más convenientes. Otra conclusión interesante, es que la polarización HV sería útil para cultivos con muchas ramificaciones, como ser el maíz, la papa y la colza, y para cultivos con estructuras verticales dominantes como en los cereales de invierno, el efecto de la atenuación y el doble rebote en la polarización VV aportaría información útil. El efecto de atenuación en cultivos de invierno (trigo, cebada) se ha comprobado también en banda X, polarización VV por el departamento de T.I. de ARBA. Kim (2013) encontró que para el cultivo de soja, la combinación banda L, en polarización HH es la que mejor se relaciona con el índice de área foliar (en inglés, LAI: Leaf area index) y el contenido de agua del cultivo. La cultura utilizada en el tratamiento del suelo también puede influenciar los niveles de retrodispersión obtenidos de coberturas terrestres agrícolas, en el caso de la utilización de siembra directa con rastrojos en superficie, el mecanismo de retrodispersión en volumen fue mayor con el aumento del volumen de rastrojo (H. McNairn et al. 2009).

Las variables agronómicas más importantes de los cultivos que pueden ser monitoreadas con sistemas SAR son: 1) el contenido de agua de los cultivos, 2) el índice de área foliar, 3) la biomasa, (Ferrazoli P. 2002; Kim et al. 2013), estas variables son importantes porque son inputs para modelos de estimación de rendimiento (Steduto et al. 2009).

La clasificación de cobertura de suelo, tema que trata la presente tesis, suele realizarse mediante clasificaciones supervisadas digitales, usualmente utilizando el clasificador de “Máxima Verosimilitud” (Kasischke, Melack, and Craig Dobson 1997).

Una clasificación supervisada de imágenes satelitales es un proceso digital por el cual se extrae, a partir de una imagen multibanda, un archivo raster de una única banda cuyos valores digitales son etiquetas que se corresponden con las clases con las que se entrenó al clasificador digital (Chuvieco E. S. 2006). Los clasificadores digitales del tipo supervisados son algoritmos que permiten clasificar imágenes a partir de un conjunto de datos de “entrenamiento”, los cuales al ser tomados del terreno pueden ser denominados “verdad de campo”. Estos datos de entrenamiento son comúnmente áreas identificadas dentro del espacio de la imagen a clasificar mediante “áreas de interés”. Las áreas de interés pueden obtenerse mediante interpretación visual de la imagen o mediante campañas a campo de recolección de datos. Las áreas de interés podrán ser utilizadas luego en el cálculo de errores y precisión del mapa. Los clasificadores supervisados pueden dividirse en dos grupos, los paramétricos y los no paramétricos. Los paramétricos son aquellos que utilizan parámetros estadísticos para estimar los límites entre clases, ejemplo de este grupo son: “clasificador de mínima distancia”, “clasificador de paralelepípedos”, y el más utilizado: “clasificador de máxima probabilidad o verosimilitud”. Dentro de los clasificadores “no paramétricos” más utilizados encontramos las “Neural Network Classifier” (Clasificador de Red Neuronal) y “Support Vector Machine” (Máquinas de Soporte Vectorial).

En la bibliografía se pueden encontrar abundantes ejemplos tanto de clasificadores paramétricos como no paramétricos siendo utilizados para clasificar digitalmente coberturas terrestres. Más adelante se presenta una descripción de los algoritmos de clasificación utilizados en la presente tesis. Ambos tipos de clasificadores han sido utilizados para clasificar imágenes SAR, (Lee, Grunes, and Pottier 2001; Chen et al. 1996; Del Frate et al. 2003).

Las primeras clasificaciones de imágenes SAR se hicieron utilizando imágenes adquiridas por radares aerotransportados, (Lee, Grunes, and Pottier 2001; Chen et al. 1996; Del Frate et al. 2003), los mismos se realizaban sobre áreas de cultivo en condiciones conocidas y controladas. Con el paso del tiempo las imágenes SAR producidas desde satélites permitieron realizar clasificaciones de cultivos de superficies mayores en el orden de 25 km² a 700 km² sobre superficies no completamente conocidas y con una mayor heterogeneidad espacial (Satalino et al. 2011; Tian et al. 2010; Lohmann et al. 2009; Jin, Mountrakis, and Stehman 2014). Entre los primeros estudios citados y los últimos se observan grandes diferencias en las precisiones globales alcanzadas en las clasificaciones (90-95% vs. 70-90%). La heterogeneidad del territorio a clasificar y la densidad de muestras de entrenamiento son determinantes en la precisión global alcanzada y en los errores de omisión y comisión de cada clase por separado. Es de esperar que a mayor superficie a clasificar y menor conocimiento de las coberturas que se encuentran en ese territorio, la precisión disminuya independientemente del clasificador utilizado o del tipo de imagen utilizada.

En general los trabajos que utilizaron mayor cantidad de información satelital, ya sea por mayor número de polarizaciones o por utilizar más de una frecuencia han mostrado mejores precisiones en las clasificaciones (Chen et al. 1996; Lee, Grunes, and Pottier 2001; Del Frate et al. 2003; Jin, Mountrakis, and Stehman 2014; Loosvelt, Peters, Skriver, Lievens, et al. 2012; Loosvelt, Peters, Skriver, De Baets, et al. 2012).

En cuanto al tipo de clasificador no se observan diferencias en las precisiones alcanzadas por clasificadores paramétricos (máxima probabilidad) y no paramétricos (Red Neuronal, Random Forest, Support Vector Machine) (Lohmann et al. 2009; Satalino et al. 2011; Jin, Mountrakis, and Stehman 2014; Chen et al. 1996). Esto presumiblemente, se deba al efecto de normalización de los datos ocurrido al filtrar las imágenes SAR. Los valores de intensidad de las

imágenes SAR sin filtrar siguen una distribución gamma. Es muy probable que los clasificadores paramétricos (que requieren que los valores a clasificar sigan una distribución normal) presenten una menor precisión global en la clasificación de una imagen SAR en comparación con los clasificadores no paramétricos, dado que estos últimos no requieren que los valores a clasificar presenten distribución normal.

Finalmente, cabe destacar que dentro de las bandas SAR típicas más utilizadas X, C y L, todas ellas han presentado buenos desempeños en clasificaciones de cultivos (Tian et al. 2010; Skriver et al. 2011).

En la presente tesis, se utilizará el concepto de “clasificación de cultivos” para referir a una clasificación digital de coberturas terrestres donde las clases predominantes son los cultivos. El resultado de la clasificación de cultivos se verá reflejada en un mapa. El tipo de clasificación utilizada en esta tesis será del tipo supervisada, es decir, se utilizarán muestras de entrenamiento para “entrenar” a un algoritmo, el cual servirá posteriormente para la predicción de una clase. El clasificador generará en una primera etapa una función que vincule los datos de entrada y salida (etapa de entrenamiento). En una segunda etapa el algoritmo clasificador utilizará esa función para predecir valores de salida a partir de un nuevo universo de valores de entrada (etapa de clasificación). Las muestras de entrenamiento, surgen de un dato de campo, o “verdad de campo”, ese dato de campo posee una componente geográfica y una etiqueta. De esa ubicación geográfica se extrae un conjunto de valores digitales del set de imágenes satelitales utilizadas, y se asocia a ese conjunto de valores la etiqueta o clase registrada en la “verdad de campo”. Esa clase se corresponde con una cobertura terrestre observada en el terreno. De manera que una “muestra de entrenamiento” puede ser caracterizada como un conjunto de vectores, donde cada uno acumula una serie de valores obtenidos a partir de un set de imágenes satelitales y una etiqueta que identifica cada vector con una cobertura del terreno. Para llevar a cabo un control de las clasificaciones, se utiliza un conjunto de vectores llamado “muestra de control” compuesta por vectores independientes de la “muestra de entrenamiento”. En la presente tesis, el control de las clasificaciones se llevará a cabo con los siguientes indicadores: Precisión global de la clasificación, error de comisión, y el error de omisión que surgen de una matriz de confusión. La matriz de confusión es una tabla de doble entrada, $N \times N$, donde N es la cantidad de clases definidas por el usuario, las columnas corresponden a los valores observados y los de las filas a los predichos por la clasificación.

La precisión global de la clasificación se define como el porcentaje de píxeles bien clasificados en relación a la cantidad de píxeles total clasificados.

Los errores de omisión y comisión se calculan siempre sobre una clase específica.

Error de omisión: suponiendo dos clases A y B, se define el error de omisión de la clase A, como el porcentaje de píxeles que pertenecen a la clase A según la muestra de control pero fueron clasificados en la clase B.

Error de comisión: suponiendo dos clases A y B, se define el error de comisión de la clase A, como el porcentaje de píxeles que fueron clasificados como A pero según la muestra de control pertenecen a la clase B.

1.3.1. Clasificadores

Los clasificadores presentados a continuación y utilizados en el procesamiento de los capítulos 3, 4, y 5 son rutinas de "machine learning" o máquinas de entrenamiento supervisadas. El objetivo de las rutinas es "entrenar" a una máquina digital o algoritmo la relación funcional entre una variable de entrada (input) y una variable de salida (output). Cuando la máquina es

capaz de predecir una variable de salida cualitativa, entonces produce una "clasificación". Cuando la máquina predice valores cuantitativos el resultado es llamado "regresión". Normalmente, las máquinas de entrenamiento supervisadas, funcionan en dos fases, la primera de entrenamiento, en la cual se genera un modelo predictivo, y la segunda de clasificación de un conjunto de vectores de entrada. El operador, deberá posteriormente calcular la precisión y los errores de la clasificación mediante la utilización de matrices de confusión.

El primer clasificador es el más simple: "Decisión Tree" o "Árbol de decisión", el segundo y el tercer algoritmo ("Random Forest" y "Gradient Boosted Trees") son combinaciones de muchos árboles de decisión generados de distinta forma. El cuarto y el quinto son algoritmos no paramétricos que no utilizan árboles de decisión. El último de los algoritmos es el de "Máxima verosimilitud", representa el algoritmo estándar de clasificación de coberturas, y es uno de los más utilizados en la actualidad.

1. "Decision tree" (DT) o "Árbol de decisión auto entrenado"

El clasificador DT es un clasificador que genera un árbol de decisión de tipo binario de manera recursiva. El clasificador DT consta de un nodo "raíz" o nodo inicial que contiene todo el conjunto de muestras de entrenamiento. El nodo "raíz" se divide en dos utilizando una regla o criterio aplicado sobre alguna de las variables que componen cada vector de la muestra de entrenamiento. Las "hojas" del árbol, donde el nodo tiene una sola salida, son las clases finales de la clasificación. Los datos de entrenamiento son divididos utilizando un criterio, por ejemplo para el caso de clasificaciones se suele utilizar el criterio de "pureza" de gini. El índice de "pureza" de gini es una medida de cuan seguido un elemento elegido al azar del set de entrenamiento sería incorrectamente clasificado si este fuera clasificado al azar de acuerdo con la distribución de clases de todo el conjunto.

La construcción del árbol se detiene si:

1. la profundidad del árbol llega a un límite máximo pre-determinado.
2. La cantidad de muestras de entrenamiento es estadísticamente no representativa como para partir el nodo.
3. Todas las muestras de entrenamiento en un nodo pertenecen a la misma clase.
4. La mejor opción de corte no mejora la clasificación al azar de una muestra.

(Open Source Computer Vision Library, n.d.; Waske and Braun 2009)

2. "Random Forest"

El algoritmo "Random Forest" es un algoritmo de clasificación que utiliza una colección de árboles de decisión, la cual es llamada "forest" o bosque. La clasificación funciona de la siguiente manera: el clasificador genera una colección de árboles de decisión autoentrenados (DT) a partir de las muestras de entrenamiento (ver clasificador "Árboles de decisión"). Todos los árboles son entrenados con los mismos parámetros pero con distintas muestras. Cada muestra contiene igual cantidad de vectores que la muestra de entrenamiento original, pero son elegidos al azar, es decir algunas muestras tendrán vectores repetidos y vectores ausentes. En cada nodo de cada árbol, no todas las variables son utilizadas para dividir el conjunto, sino solo algunas seleccionadas al azar. Luego de generado el "bosque" o conjunto de árboles, el algoritmo utiliza la salida de cada árbol para clasificar el conjunto de valores de entrada. Un

vector de entrada es clasificado tantas veces como árboles tenga el “bosque”, ese vector es clasificado finalmente con la clase donde ha sido clasificado más veces (criterio de votación simple). (Open Source Computer Vision Library, n.d.; Waske and Braun 2009)

3. “Gradient Boosted Trees Classifier” (GBT)

El clasificador “Gradient Boosted tree” es un tipo de clasificador Boost que combina muchos “Árboles de decisión autoentrenados” (DT), cada clasificador DT por separado es considerado un “clasificador débil”. Un “clasificador” débil es un algoritmo de clasificación que arroja un resultado levemente superior al obtenido por azar, por lo tanto, puede ser muy simple y consumir pocos recursos computacionales. El algoritmo “GBT” genera de manera secuencial una cantidad predeterminada de árboles de decisión. La palabra “Gradient” (Gradiente en castellano) hace referencia a un algoritmo de “gradiente descendente” o de búsqueda del mínimo en una función de manera iterativa. Esa función para el caso del algoritmo GBT es llamada función “Loss” (L) o función de costo. La función “Loss” es una función cuyo valor de salida aumenta cuando empeora el desempeño de la clasificación. Cada árbol es generado utilizando información generada por el árbol anterior, minimizando la función Loss, es decir disminuyendo el error. El clasificador GBT genera una familia de árboles (M) por cada clase (k). Para predecir a qué clase pertenece una determinada muestra, la muestra recorre todos los árboles de todas las clases, para cada clase (k), calcula la suma de los resultados obtenidos en cada árbol, multiplicados por un valor de regularización “v”: $f(x) = f_0 + v \cdot \sum_{i=1}^M T_i(x)$ (siendo f_0 : el árbol inicial, $T(x)$ el resultado de cada árbol para una clase, el coeficiente v permite disminuir el riesgo de “sobre entrenamiento” o “over-fitting”). Finalmente la muestra es clasificada como la clase que mayor valor obtuvo ($\arg \max_{i=1 \dots k} (f_i(x))$). (Open Source Computer Vision Library, n.d.; Friedman, 1999)

4. Neural Network

El algoritmo de redes neuronales o redes neuronales artificiales utiliza una capa de entrada o “input”, una o más capas “ocultas” y una capa de resultado o “output”. Cada capa se define por la cantidad de “neuronas” que contiene y una función de activación. La cantidad de “neuronas” del “input” es igual en tamaño al vector de entrenamiento. La cantidad de “neuronas” de la capa “output” es igual a la cantidad de “clases” que se desee clasificar el conjunto de datos. La cantidad de capas “ocultas” y la cantidad de “neuronas” por capa oculta, es definida por el operador humano (arquitectura de la red).

En la fase de entrenamiento el algoritmo calcula un peso para cada neurona del vector “input” de manera iterativa hasta encontrar la solución que mejor ajuste al conjunto de datos de entrenamiento proporcionados. La suma de los pesos de cada neurona de la capa input pasa a cada neurona de la capa oculta. Mediante una función de activación (gaussiana o sigmoidea) la señal pasa a una u otra neurona de la capa oculta subsiguiente y finalmente a la capa output.

Una vez entrenada la red, el algoritmo puede clasificar un nuevo set de datos de la siguiente manera:

1. Toma el vector “input”
2. Pasa los valores de la capa “input” a la primera capa oculta.
3. Calcula los valores de salida de la capa oculta utilizando la función de activación y los pesos (ajustados en la etapa de entrenamiento)

4. Pasa los valores a la siguiente capa oculta o a la capa "output", otorgando una clase a cada vector de entrada.

(Open Source Computer Vision Library, n.d.)

5. “Support Vector Machine”(SVM) (máquinas de vectores de soporte)

El algoritmo SVM fue originalmente pensado como clasificador binario. Luego el algoritmo se extendió para poder ser utilizado en problemas de regresión y clasificación para más de dos clases. El algoritmo encuentra el límite entre clases utilizando planos de soporte generados a partir del set de datos de entrenamiento. El límite entre dos clases se encontrará equidistante a los dos planos generados por los vectores más cercanos a los de la clase vecina (vectores de soporte).

El algoritmo “Lib Support Vector Machine” (Chang and Lin 2011), resuelve el problema de multi-clases con la estrategia de “un voto contra un voto”. Cuando se utiliza para resolver clasificaciones de más de dos clases, el algoritmo divide el problema en $P = n(n-1)/2$ (siendo n el número de clases, y P el número de subproblemas), los clasificadores binarios se combinan con el criterio de votación simple (Löw et al. 2013; Open Source Computer Vision Library, n.d.).

6. Máxima Verosimilitud

Máxima Verosimilitud es un clasificador paramétrico, es decir utiliza parámetros estadísticos y asume una distribución normal de los valores digitales de los píxeles de una clase. El algoritmo calcula la probabilidad para cada pixel de pertenecer a cada clase utilizando los parámetros estadísticos (media y varianza), a partir de una muestra de entrenamiento. Luego en la etapa de clasificación, cada pixel es clasificado a la clase que tiene mayor probabilidad de pertenecer. Se puede definir un límite de probabilidad por debajo del cual el pixel es clasificado como “sin clase” o “clase nula”, es decir que no pertenece a ninguna clase. (Bustos, Frery, Lamfri, Scavuzzo. 2004)

1.4. Zonas de estudio

Los principales cultivos de granos (cereales y oleaginosas) de la Provincia de Buenos Aires son: la soja, el maíz, el trigo y el girasol. Dichos cultivos se desarrollan en casi todo el territorio de la provincia pero presentan rendimientos más altos en la llamada “Zona Núcleo”, la misma comprende grandes zonas de los partidos de Leandro N. Alem, General Viamonte, Bragado, General Arenales, Junin, Alberti, Rojas, Chivilcoy, Chacabuco, Colon, Salto, San Nicolás, Ramallo, San Pedro, Baradero, San Antonio de Areco, Exaltación de la Cruz, Capitan Sarmiento, San Andrés de Giles, Pergamino, Arrecifes, y Carmen de Areco, en la Provincia de de Buenos Aires. El cultivo de papa se produce principalmente en los partidos de Gral. Pueyrredon, Gral. Alvear, y Balcarce, que pertenecen a la zona Sudeste de la provincia. En tanto que el cultivo de cebolla se encuentra localizado casi exclusivamente en dos partidos: Villarino y Patagones, en la zona bajo riego del cauce inferior del Rio Colorado, ver Figura 1. Los tres sistemas de agricultura son de importancia económica para la provincia dado que sus productos son apreciados comercialmente tanto en el mercado interno como externo. A su vez, los dos últimos (cultivo de papa y cebolla) requieren de gran inversión de capital y mano de obra para su desarrollo, y presentan altos rindes por hectárea (en el orden de los 20 a 60 tn/ha promedio). La detección

temprana de estos cultivos permite la fiscalización no solo de la superficie producida por parcela fiscal sino también el control por parte de la autoridad competente de las condiciones de trabajo de los trabajadores rurales aplicados a las tareas de cosecha y embolsado del producto.

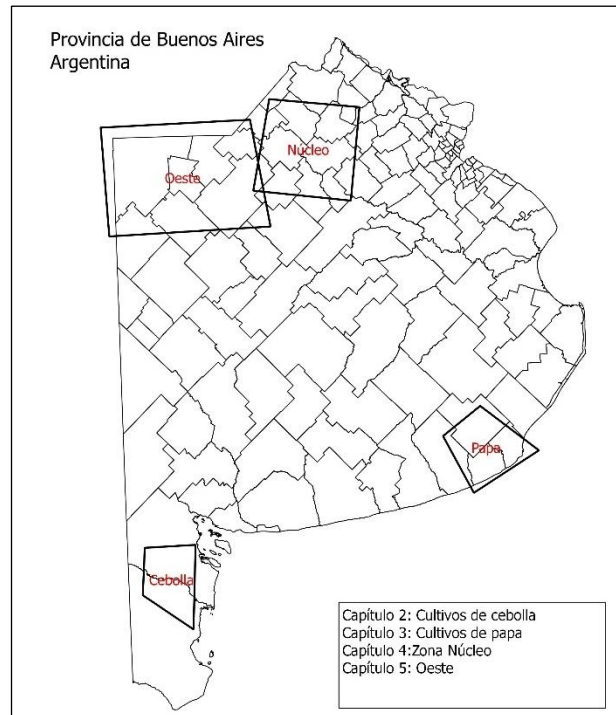


Figura 1. Zonas de estudio.

1.5. Objetivos

El objetivo general de la presente tesis es proporcionar herramientas metodológicas al departamento de Tecnología de Imágenes de la Agencia de Recaudación de la Provincia de Buenos Aires con el fin de incorporar imágenes SAR y métodos estadísticos de clasificación de cultivos, al sistema de detección de cultivos. Si es posible, intentar encontrar la mejor metodología de procesamiento, según estándares actuales (2015), para clasificar cada cultivo en términos de repetitividad, calidad, solidez y rapidez en la entrega de los productos derivados.

Objetivos particulares:

- 1- Generar una clasificación de cultivo de cebolla del sur de la provincia de Buenos Aires y estimar la superficie cultivada.
- 2- Generar una clasificación de cultivo de papa en el Sudeste de la provincia de Buenos Aires, utilizando imágenes SAR Cosmo SkyMed y estimar la superficie cultivada.
- 3- Generar una clasificación de cultivos de granos del norte y noroeste de la provincia de Buenos Aires, utilizando una serie temporal de imágenes SAR Cosmo SkyMed e imágenes SAR Sentinel-1, y estimar la superficie cultivada.
- 4- Comparar clasificaciones supervisadas generadas a partir de imágenes SAR e imágenes Multiespectrales.

1.6. Hipótesis

1. La primera hipótesis es que se pueden clasificar cultivos utilizando series temporales de imágenes SAR, sea banda X como en banda C.
2. La segunda hipótesis es que la clasificación supervisada de cultivos utilizando sets de imágenes SAR multitemporales es igual o mejor que una clasificación supervisada utilizando imágenes multiespectrales Landsat 8.
3. La tercera hipótesis es que clasificar imágenes SAR en forma no paramétrica es preferible a clasificarlas en forma paramétrica, basando dicha preferencia en los errores y la precisión global del mapa.

1.7. Niveles de procesamiento imágenes Cosmo SkyMed

Las imágenes satelitales en general son distribuidas con un nivel de procesamiento realizado por el productor de la imagen. Los niveles de procesamiento hacen referencia a la cantidad de procesos realizados por el productor de la imagen. En general, a mayor cantidad de procesos, más alto el nivel. A continuación se detallan algunos de los niveles con los que se solicitaron las imágenes Cosmo SkyMed.

SLC: "Single Look Complex". Nivel 1 A, datos focalizados calibrados radiométricamente. Contiene los datos de fase y amplitud.

DGM: "Detected Ground Multilook". Nivel 1 B, datos focalizados calibrados radiométricamente. Contiene únicamente datos de amplitud. Los datos de amplitud se presentan filtrados ("de-speckle").

GEC: "Geocoded Ellipsoid Corrected". Nivel 1 C. Presenta los datos de amplitud, calibrados, filtrados y proyectados a un elipsoide de referencia.

GTC: "Geocoded Terrain Corrected". Nivel 1D. Presenta los datos de amplitud, calibrados, filtrados y proyectados a un modelo digital de elevación de referencia.

Capítulo 2. Cultivo de Cebolla en el Sur de la provincia de Buenos Aires

2.1. Introducción

El cultivo de cebolla es un cultivo muy importante en el sur de la provincia de Buenos Aires. Constituye la actividad económica rural más importante de los partidos de Villarino y Patagones. La misma, es una actividad estacional (primavera, verano) y demanda gran cantidad de mano de obra. Es de destacar que la mano de obra utilizada a menudo se encuentra precarizada. Esto determina que la fiscalización de esta actividad impacte al mismo tiempo en la fiscalización del trabajo informal que muchas veces se encuentra asociado a la misma.

El sitio de estudio es el valle de riego del curso inferior del Río Colorado (ver Figura 1). El mismo fue desarrollado mediante canales sobre-elevados que distribuyen agua de riego a las distintas parcelas. El riego es administrado por una corporación llamada “Corfo” (Corporación de Fomento del valle del Río Colorado), y representa el área bajo riego más importante de la provincia de Buenos Aires. Las precipitaciones de la zona son escasas e irregulares por lo cual la mayoría de los años son insuficientes para mantener cultivos en secano. El riego se constituye como un factor determinante de la producción. Las áreas bajo riego tienen un régimen especial de valuación de la tierra según el catastro bonaerense dado que las mismas poseen una capacidad productiva diferencial con respecto a las parcelas que no poseen riego por canal.

El objetivo de este capítulo es lograr una clasificación de las parcelas con cultivos de cebolla, utilizando una “verdad de campo” y una serie temporal de imágenes Cosmo Himage.

El cultivo de cebolla se siembra en los meses de Agosto-Noviembre y se cosecha de Enero-Abril. La cosecha es casi siempre manual y su objetivo es la confección de “parvas” de cebolla para que los bulbos se sequen en el campo y puedan ser posteriormente embolsados. Las parvas permanecen un tiempo variable en el campo según las necesidades comerciales del productor. El 50% del costo de producción corresponde a mano de obra. La incidencia de las labores manuales en todo el ciclo de producción es muy alto en comparación a otros cultivos. Se ha detectado que gran parte de esa mano de obra es de carácter informal, llegando inclusive a encontrarse trabajo infantil (La Nueva Provincia 2014).

2.2. Materiales & Métodos

Los materiales utilizados para generar una clasificación de cultivos de cebolla fueron una “verdad de campo” conformada por puntos de observación de cultivos georreferenciados y una serie temporal de imágenes Cosmo SkyMed modo Himage y nivel de procesamiento Single Look Complex, provistas por CONAE. La “verdad de campo” es un registro georreferenciado, en este caso, de coberturas de terreno.

La “verdad de campo” se generó en un viaje de campaña desarrollado en el área de estudio durante la semana del 23 al 27 del mes de Febrero del año 2015. Durante la misma se registraron la ubicación y el estadio del cultivo encontrado en las parcelas observadas sobre un recorrido predeterminado (Figura 2). Se observaron cultivos de cebolla en diversos estadios: fase de crecimiento herbáceo, bulbificación, y cosecha. Sin embargo, la mayoría de los cultivos observados ya se encontraban cosechados, y la producción almacenada a campo en forma de “parvas”.

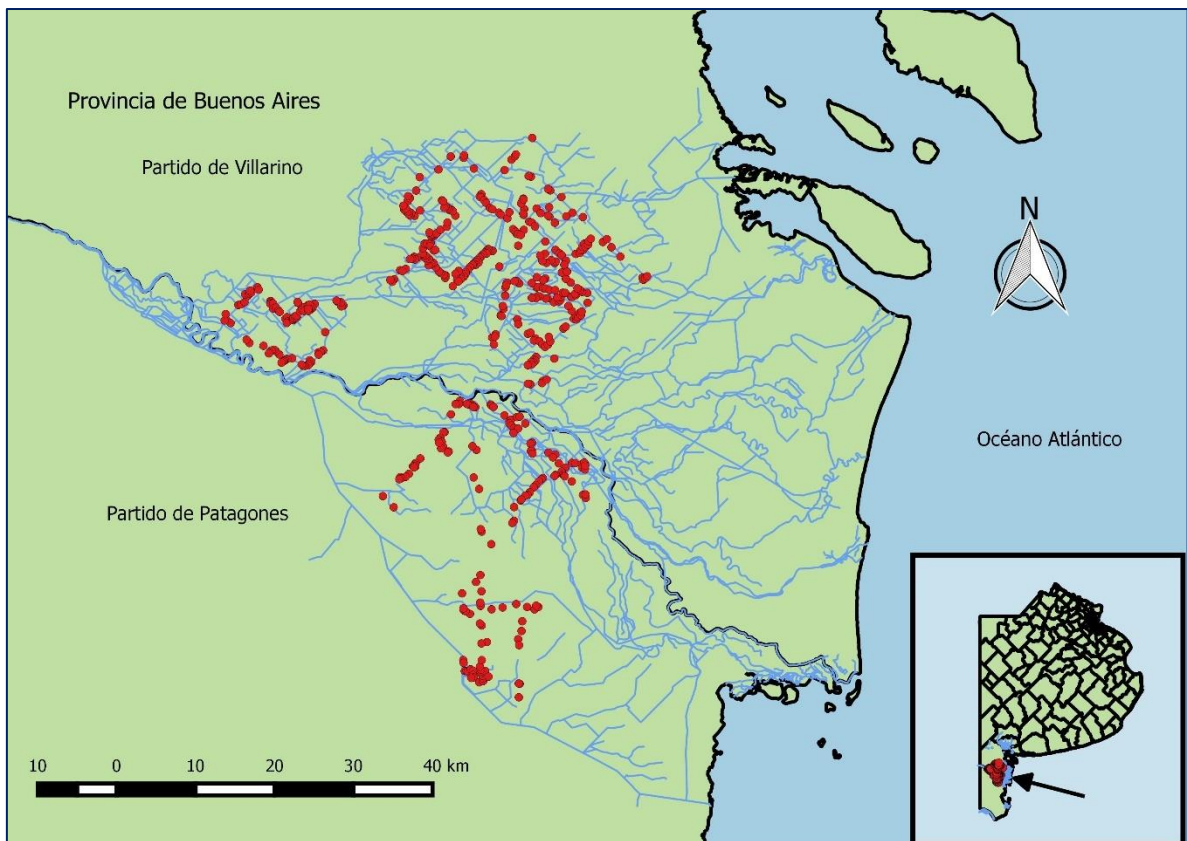


Figura 2. Puntos de observación de cultivos tomados en el mes de Febrero del 2015.

2.2.1. Datos de campo

El registro de los datos en un sistema de información geográfica se logró mediante el uso de una Tablet PC marca Lenovo con sistema operativo Windows 8, y utilizando el programa ArcPad de ESRI. En paralelo se tomaron fotografías con una cámara digital, asociando el número de foto al punto georreferenciado cargado al sistema. Además se probó tomar fotografías georreferenciadas directamente desde un teléfono celular “Smart” con sistema operativo Android 4.2. El sistema operativo del celular permite guardar las coordenadas obtenidas del sistema GPS/Glonass directamente al archivo digital de cada fotografía. De esta manera utilizando un programa que lea la metadatos del archivo digital se puede reconstruir la ubicación donde fue tomada la fotografía e integrarla a un sistema de información geográfica.

La toma de datos a campo, tanto en este capítulo como en los subsiguientes, representó un esfuerzo importante. En el caso del presente capítulo implicó cinco jornadas completas de trabajo a campo, transitando caminos rurales en mal estado. El trabajo consistió en recorrer caminos rurales de la zona de estudio en un vehículo con chofer provisto por la Agencia de Recaudación de la Provincia de Buenos Aires, junto al tesista y dos colaboradores (Ing. Agr. Jorge Morgan e Ing. Agr. Gabriel Ferrero). Durante el recorrido se utilizó una Tablet PC para navegar mapas con imágenes ópticas y SAR de referencia. El trabajo de la recolección de puntos consistió en detenerse a registrar en el sistema de información geográfica móvil los cultivos y coberturas agrícolas o ganaderas. En total se registraron 503 puntos de observación. Las tareas de observación de cultivos, determinación, y registro de datos en el sistema fueron alternándose entre los integrantes del grupo de trabajo.

2.2.2. Imágenes SAR disponibles

Se solicitó una serie de imágenes Cosmo SkyMed Himage (Tabla 1) con distinto ángulo de incidencia y polarización de manera de completar un set multitemporal con variedad de polarizaciones y ángulos con el fin de aumentar la diferencia capturada en la clasificación entre las distintas coberturas presentes en el sitio. Las imágenes obtenidas fueron las que se muestran en la Tabla 1.

IMAGEN	MODO	NIVEL DE PROCESAMIENTO	POLARIZACIÓN	FECHA	ÁNGULO DE INCIDENCIA	ORBITA	MIRADA
1	Himage	SLC	HV	11/11/2014	51.15-51.98	ASCENDENTE	Derecha
2	Himage	SLC	HV	11/11/2014	20.05-23.5	ASCENDENTE	Derecha
3	Himage	SLC	HV	12/11/2014	41.79-43.62	ASCENDENTE	Derecha
4	Himage	SLC	HV	20/11/2014	41.79-43.63	ASCENDENTE	Derecha
5	Himage	SLC	HH	08/01/2015	41.79-43.64	ASCENDENTE	Derecha
6	Himage	SLC	HV	15/01/2015	41.79-43.65	ASCENDENTE	Derecha
7	Himage	SLC	HH	24/01/2015	41.79-43.66	ASCENDENTE	Derecha
8	Himage	SLC	HH	23/03/2015	29.27-31.96	ASCENDENTE	Derecha

Tabla 1. Imágenes CSK-Himage utilizadas. Sur de la prov. de Buenos Aires. Los satélites de la constelación CSK pueden obtener imágenes en sus órbitas descendentes y ascendentes, apuntando tanto a la derecha como a la izquierda de la dirección de avance del satélite.

Las imágenes Cosmo SkyMed fueron obtenidas a través de la Comisión Nacional de Actividades Espaciales (CONAE).

2.2.3. Pre-procesamiento

Para poder utilizar las imágenes SAR tanto en clasificaciones digitales como en interpretaciones visuales, es preciso realizar un procesamiento previo o pre-procesamiento. Las imágenes fueron pre-procesadas con el programa libre NEST de la Agencia Espacial Europea. A continuación se grafica los procesos realizados a las imágenes receptadas en formato "Hierarchical Data format 5", (".h5").

El grafo utilizado (Figura 3) realiza los siguientes procesos a cada imagen: primero lee la imagen, luego la calibra a valores de intensidad Sigma cero (Formula 1), a continuación corrige la imagen utilizando un modelo digital de elevación (SRTM de 90m, por sus siglas en inglés: Shuttle Radar Topography Mission, cortesía del USGS) que descarga de la web y finalmente graba la imagen resultante en el formato BEAM, nativo del programa NEST. De esta manera obtenemos imágenes georreferenciadas por orbita y terreno en su máxima resolución (2 metros por pixel). Los parámetros utilizados para el pre-procesamiento se describen en el Anexo 3.

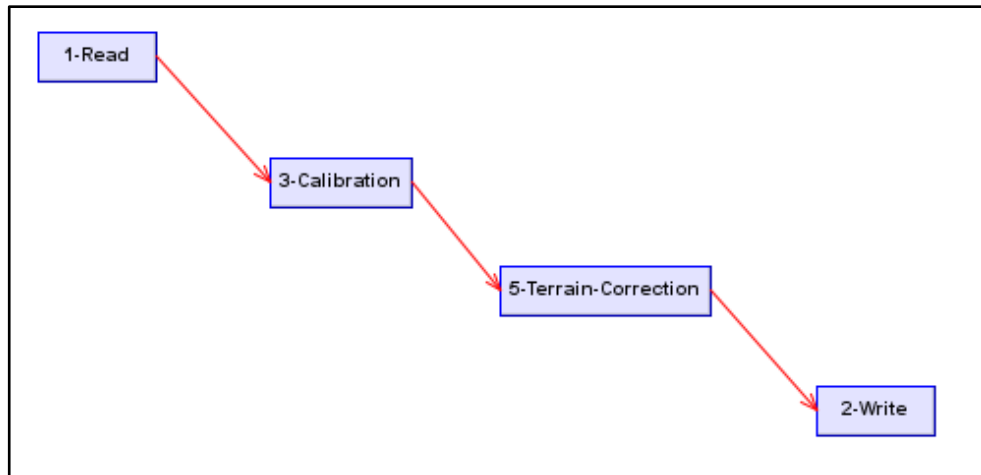


Figura 3. Grafo de procesos aplicados a las imágenes CSK-Himage con el programa Nest. Los parámetros se pueden ver en el Anexo 3.1.

Fórmula 1. Calibración de las imágenes CSK:

$$\sigma^0(i,j) = Amplitud(i,j)^2 \times Ftot$$

$$Ftot = R_{ref}^{2*Rexp} \times \sin(\alpha_{ref}) \times \frac{1}{F^2} \times \frac{1}{K}$$

Las constantes de la fórmula se encuentran en la metadata de las imágenes. El programa NEST 5.1 resuelve automáticamente la fórmula extrayendo las constantes de la metadata.

Ref: "Reference Slant Range"

Rexp: "Reference Slant Range Exponent"

α ref: "Reference Incidence Angle"

K: "Calibration Constant"

F: "Rescaling Factor"

La fórmula anterior permite el cálculo de la retrodispersión sensada por el radar como el coeficiente de retrodispersión sigma cero. La retrodispersión es la relación entre la energía dispersada por una superficie hacia la fuente (la antena radar) y la energía incidente sobre esa misma superficie. Los coeficientes de retrodispersión se calculan mediante la normalización de la retrodispersión por un área estándar. El área estándar para el caso del coeficiente de retrodispersión sigma cero es el área sobre el terreno donde impacta el haz electromagnético producido por la antena SAR ("Ground range"). Sigma cero es una variable utilizada para evaluar las condiciones de retrodispersión de una superficie, minimiza el efecto de la topografía sobre los valores de retrodispersión y permite comparar imágenes de distintas fechas.

2.2.4. Procesamiento

Al observar las imágenes del mes de Enero se encontró un patrón en la imagen que refleja de manera inequívoca y sin ambigüedades la presencia del cultivo de cebolla. El patrón identificado con las parvas de cebolla en campo, es el de hileras de puntos brillantes espaciadas de manera regular sobre un fondo oscuro (Figura 4), correspondiéndose a las parvas de cebolla sobre suelo desnudo (sin malezas en su entorno). Esto produjo un cambio de enfoque en el problema y se decidió realizar una clasificación visual de las parcelas identificando aquellas donde se observa el cultivo apilado en parvas listas para el embolsado. El proceso por el cual se identificaron las parcelas consistió en componer un proyecto en el programa Qgis2.1 superponiendo las imágenes SAR Cosmo SkyMed de las fechas de Enero y Marzo. Sobre ellas se volcaron los puntos de observación relevados a campo. Una vez identificado el patrón de textura e intensidad de la imagen que se corresponde con las parvas de cebolla a campo, se relevaron las imágenes SAR identificando todos los lotes con el mismo patrón. Para ello, se generó un nuevo vector de polígonos, y se dibujó un nuevo polígono sobre cada área que presentara el patrón descrito para las parvas de cebolla sobre el suelo desnudo. El mismo patrón de hileras sobre un fondo sin vegetación puede observarse en el programa Google Earth (figura 5).

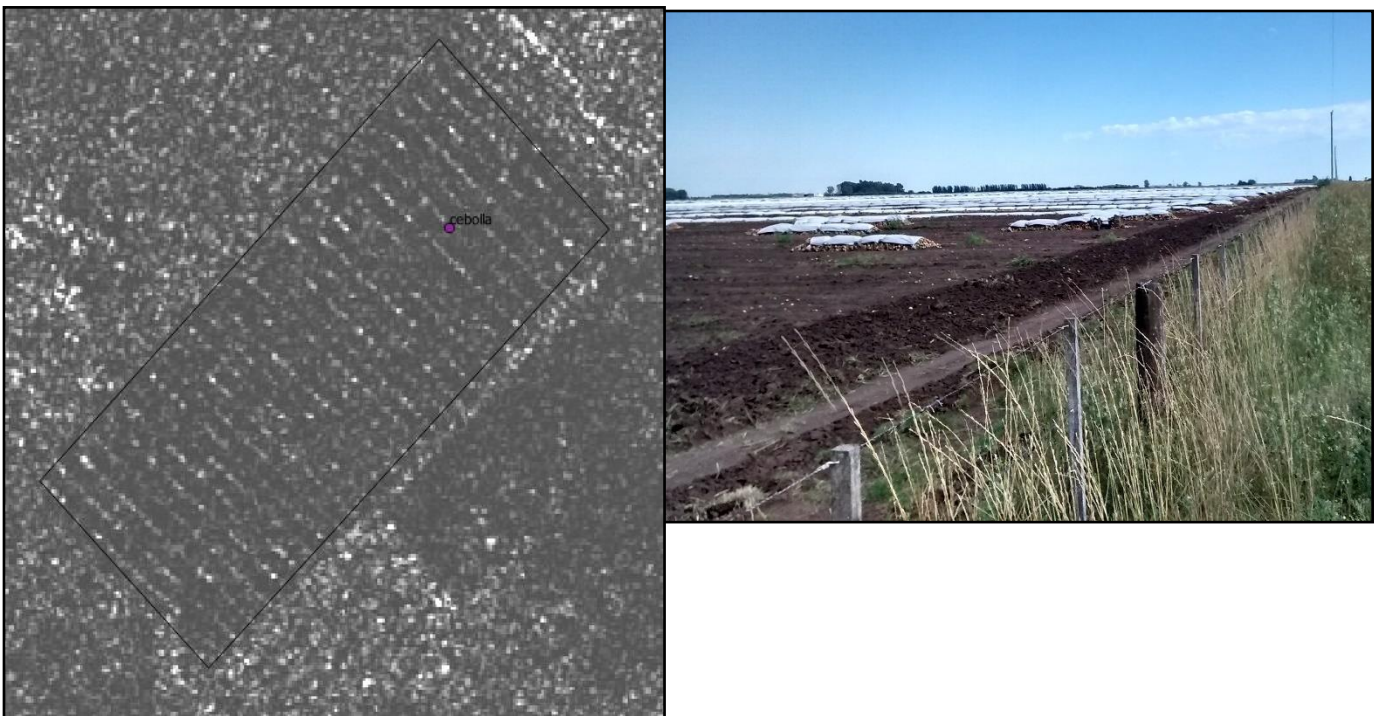


Figura 4. Izquierda: imagen CSK-Himage (2m resolución). Fecha: 23 de Enero. Derecha .Fotografía parvas de cebolla.



Figura 5. Vista de parvas de cebolla en Google Earth.

2.3. Resultados

La clasificación visual de las imágenes del 24 de Enero y del 23 de Marzo donde se podían detectar las parvas de cebolla en el campo arrojó resultados positivos (Figura 6).

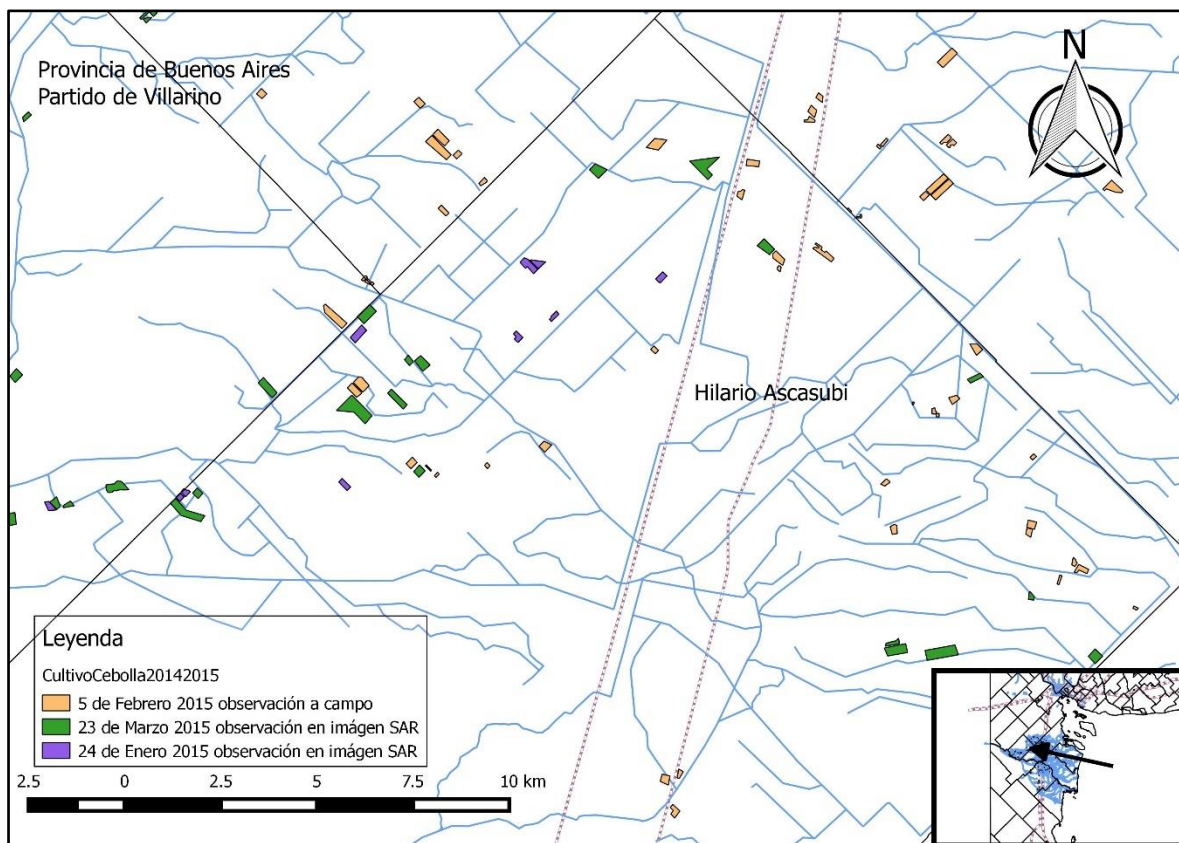


Figura 6. Mapa detalle del área relevada. En marrón las superficies con cultivo de cebolla detectadas a campo. En verde las relevadas mediante imágenes SAR el 23 de Marzo de 2015 y en violeta las relevadas mediante imágenes SAR con imagen del 24 de Enero de 2015.

La superficie detectada total para la campaña 2014-2015 fue de **1385 ha** de cultivo de cebolla. El número de parcelas fiscales con cultivo fue de **269** (Figura 7).

Las imágenes que más información aportaron fueron las del mes de Enero y Marzo. La polarización simple HH con ángulo de incidencia cercano a 40° mostró mejor calidad que las imágenes de polarización cruzada HV.

No todas las parcelas detectadas a campo, fueron necesariamente detectadas en las imágenes, solo aquellos lotes que fueron cosechados antes del 23 de Marzo, fecha de la imagen con que se contaba posterior a la campaña de reconocimiento a campo.

Deteccion cultivos de cebolla campaña 2014-2015

Superficie cultivada detectada de cebolla 1385 ha

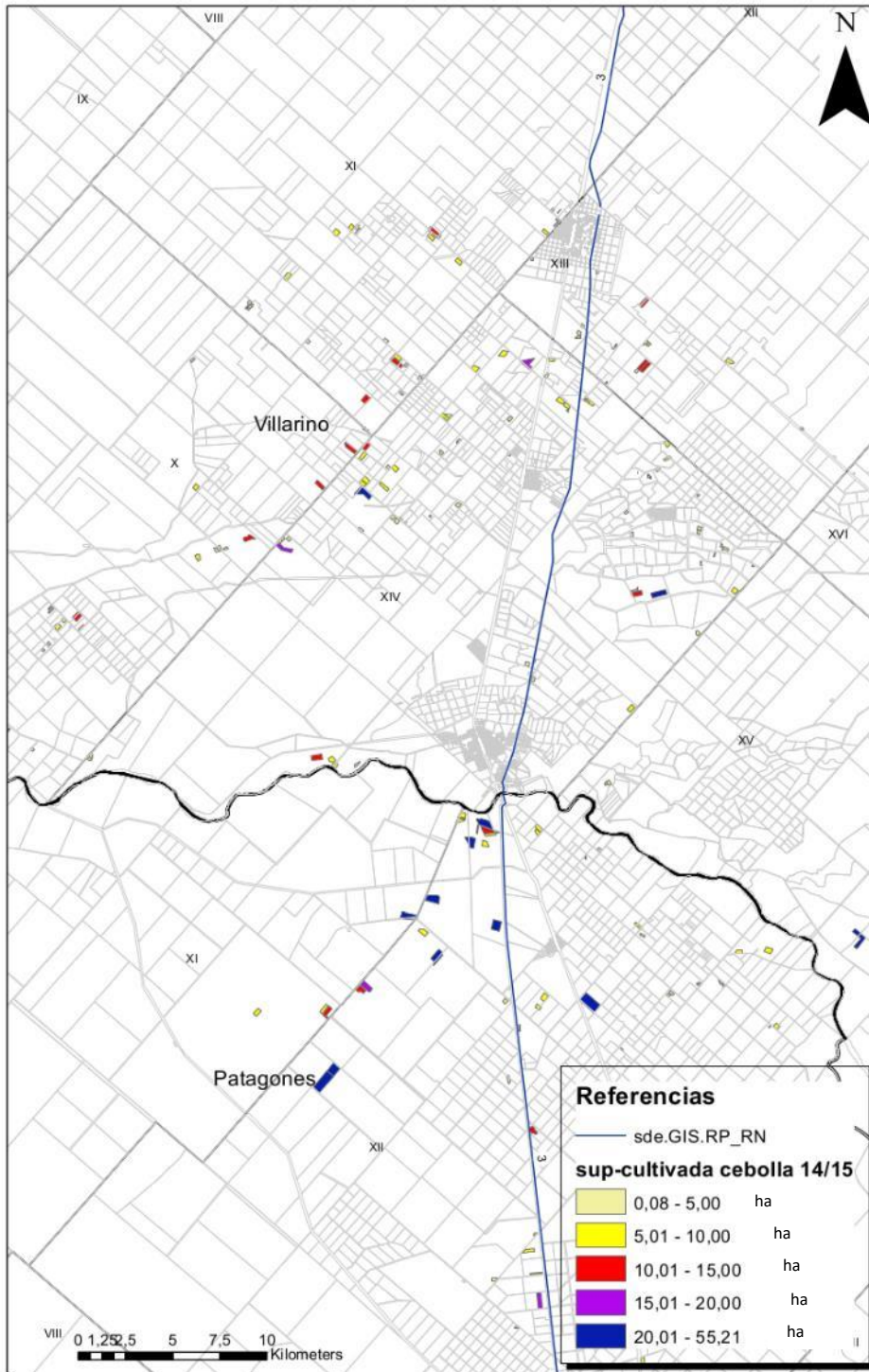


Figura 7. Parcelas con cultivo de cebolla detectadas de forma remota, mediante las imágenes Cosmo SkyMed de las fechas 24 de Enero de 2015 y 23 de Marzo de 2015, y las observadas a campo en la fecha 5 de Febrero de 2015.

2.4. Discusión

La clasificación visual de las imágenes mostró ser un método novedoso por la utilización de imágenes SAR para la detección de este tipo de objetos (no se ha encontrado esta aplicación en la bibliografía revisada). Es un proceso no automático dada la necesidad de tener un clasificador humano para la obtención del producto. Si bien la tendencia es la automatización de los procesos, no siempre es conveniente hacerlo. La calidad del proceso de clasificación visual en este caso ha demostrado un desempeño cualitativamente alto, dada las características del objeto observado (las parvas de cebolla sobre suelo desnudo). Lo más importante en este proceso es que se logre el contraste entre las parvas y el suelo circundante a las mismas. Para ello es imprescindible que la fecha de la imagen coincida con la presencia de la parva y que no exista presencia de maleza en torno a la misma. La aparición de malezas circundantes sucede aproximadamente un mes luego de la confección de la parva ya que la fertilidad y la humedad del suelo promueven el crecimiento de malezas post-cosecha. Cumplida la condición temporal de la imagen es importante mencionar que cualitativamente se mostró superior la polarización simple HH por sobre la polarización cruzada HV.

La clasificación visual arrojó un producto cartográfico que fue utilizado para la fiscalización por parte de la ANSES (Administración Nacional de Seguridad Social) y el Ministerio de trabajo de la Nación para el control del trabajo rural informal y el trabajo infantil en el sitio de estudio. El mismo consta de la ubicación georreferenciada de las parcelas donde se detectó el cultivo de cebolla de manera inequívoca de manera que los organismos citados puedan acceder a las mismas y realizar los controles pertinentes sobre el personal de los establecimientos hortícolas individualizados. El producto se observa en la figura 7.

Debido a que no se poseían al momento de realizar la detección visual de la cantidad óptima de imágenes, la superficie estimada ha sido subestimada con respecto a la superficie real.

No se ha probado una clasificación automática de los cultivos de cebolla debido a la multiplicidad de estadios distintos y condiciones de humedad de suelo (por riego) que presentan los cultivos de cebolla en un mismo momento de la campaña agrícola, y la escasa cantidad de imágenes SAR. Es muy probable que un algoritmo clasificador supervisado sufriera un “subentrenamiento”, es decir, no contara con la cantidad de información necesaria para poder separar la clase “cebolla” de otras clases con comportamiento radiométrico temporal similar.

2.5. Conclusiones

Si bien la necesidad de procesar de manera rápida las imágenes para la obtención del producto citado anteriormente impidió el desarrollo de un método automático de detección de “parvas” de cebolla, queda como tema de estudio a futuro. Para ello se podría explorar la utilización de filtros de Fourier para la detección de la frecuencia o las frecuencias con las que se disponen las “parvas” sobre el terreno. Otro enfoque para la clasificación automática de lotes con cultivo de cebolla podría ser la clasificación a partir de una segmentación previa de una imagen y posterior cálculo de estadísticas espaciales sobre los vectores producto de la segmentación.

Es importante destacar que a partir de esta tesis se puede reproducir el uso de imágenes CSK-Himage para la detección de parcelas con cultivo de cebolla y la fiscalización de trabajo informal y trabajo infantil sobre las mismas. La versatilidad de la metodología propuesta radica en la reducción del costo de realizar campañas de relevamiento en campo y la posibilidad de “observar” parcelas inaccesibles por medio de la red vial actual (ver Figura 6). Se espera que la próxima campaña agrícola 2015-2016 se realice la detección de parcelas con cultivo de cebolla

utilizando imágenes CSK-Himage sin la necesidad de realizar un control a campo de las mismas. La metodología propuesta reemplazaría la observación de las parvas de cebolla a campo.

El pedido de imágenes debería realizarse para la totalidad de la superficie del valle inferior del Rio Colorado con una cobertura a fines de Enero y otra a fines de Febrero de manera de cubrir la mayor parte del período de cosecha de cebolla del sur de la provincia de Buenos Aires. Este producto podría realizarse para otras zonas productoras de cebolla del país como el valle del Rio Negro, Santiago del Estero y San Juan.

Capítulo 3. Cultivo de papa en el Sudeste de la provincia de Buenos Aires

3.1. Introducción

El cultivo de papa en la provincia de Buenos Aires es de gran importancia económica. La producción provincial de papa se concentra en los partidos de Tandil, Balcarce, Gral. Pueyrredón, y Gral. Alvarado y tiene su destino tanto en el mercado interno como en la exportación. Además es una producción primaria de gran valor agregado que a su vez sufre a industrias alimenticias de la zona.

El cultivo se desarrolla durante los meses de Septiembre a Marzo. Su cultivo requiere gran cantidad de insumos, riego y labores mecánicas. Las labores culturales previas a la plantación del tubérculo incluyen refinamiento de la tierra y confección de camellones de aproximadamente un metro por treinta a cuarenta centímetros de alto (ver Figura 8 y Figura 9). Durante el desarrollo del cultivo se utiliza riego por aspersión mediante equipos de avance frontal o “pivot” central mayormente.



Figura 8. Cultivo de papa sobre camellones. 11 Diciembre 2014.



Figura 9. Cultivo de papa listo para cosechar. 4 de Marzo 2015.

El objetivo de este capítulo es generar una clasificación supervisada de los cultivos presentes en el sitio de estudio, estimar la superficie implantada, e identificar especialmente las parcelas con cultivo de papa.

3.2. Materiales y Métodos

La clasificación de las parcelas con cultivos se llevó a cabo mediante la utilización de sets multitemporales de imágenes Cosmo-SkyMed modo Himage, procesadas a partir del nivel SLC (Single Look Complex) (Tabla 2). Las clasificaciones fueron de carácter estadístico supervisado para lo cual se realizó una recolección de datos de campo en los meses de Diciembre y Marzo. Los programas utilizados para el entrenamiento y clasificación fueron Envi 4.8 y Monteverdi 2 de la CNES (Centre National d'Études Spatiales). Para el pre-procesamiento de las imágenes Cosmo se utilizó el programa NEST 5.1 de la ESA (Agencia Espacial Europea). Adicionalmente, se clasificaron sets de imágenes Landsat 8 y sets combinados de imágenes Cosmo Himage e imágenes Landsat 8. Finalmente, se probó una clasificación utilizando un set compuesto de índices texturales (ver sección 3.2.3) extraídos a partir de las imágenes Cosmo SkyMed Himage con el programa Sentinel-1 toolbox también de la ESA.

3.2.1. Datos de campo

Los datos de campo utilizados para el entrenamiento de los clasificadores supervisados, se obtuvieron mediante dos campañas de recolección de datos. La primera se llevó a cabo el once de Diciembre de 2014 y la segunda el cuatro de Marzo de 2015 (ver Figura 10). Durante la primera campaña de recolección de puntos de observación se registraron 243 puntos, mientras que en la segunda campaña se recolectaron 157 puntos de observación, totalizando entre ambas campañas 400 puntos de observación. De esta manera se cubrió el ciclo de cultivo de las distintas fechas de plantación presentes en el sudeste de la provincia de Buenos Aires. La mayoría del recorrido se desarrolló sobre caminos de tierra y ripio. Los datos se almacenaron en una Tablet PC sobre el programa licenciado de ESRI, ArcPad®. A su vez se tomaron fotografías y se incluyeron en los datos registrados de cada parcela.

Durante las campañas de recolección de datos se registraron los cultivos observados en cada parcela sobre un recorrido predeterminado en la zona de estudio. Posteriormente en

gabinete se dibujó sobre las imágenes SAR los polígonos correspondientes a las parcelas de cultivo identificadas a campo (ver Figura 11). Si bien se relevaron diez clases, las clases “objetivo”, es decir las de interés en el presente estudio, fueron las correspondientes a los cultivos. Las capas restantes representan el resto de las coberturas posibles en el sitio de estudio.

Las coberturas relevadas fueron agrupadas en diez clases:

1. Cultivo de papa
2. Bajos anegables
3. Cultivo de Soja de primera
4. Cereal de Invierno/Soja de segunda
5. Cultivo de Maíz
6. Cultivo de Girasol
7. Pastizales/Pasturas
8. Urbano
9. Forestación
10. Agua

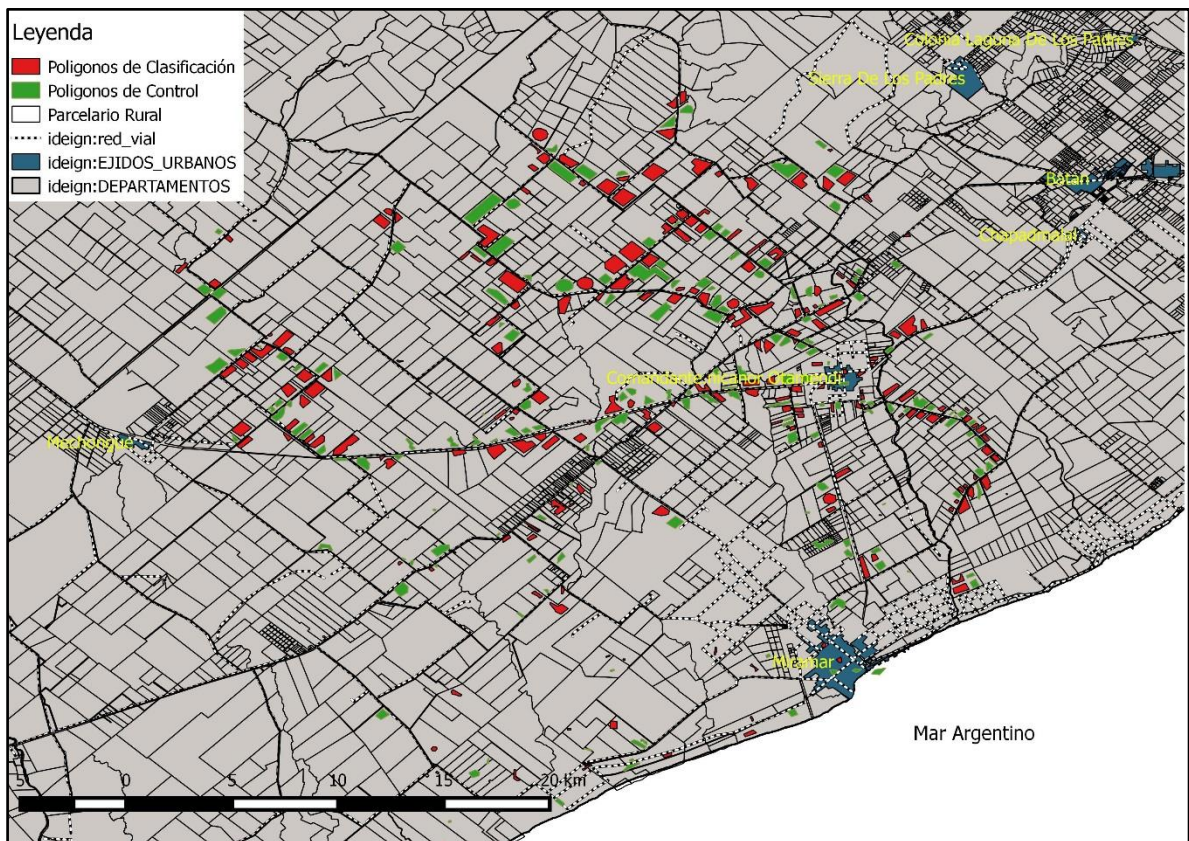


Figura 10. Polígonos de clasificación o entrenamiento y de control o testeo registrados a partir de dos viajes de recolección de datos de campo (Diciembre 2014 y Marzo 2015) en el sudeste de la provincia de Buenos Aires.

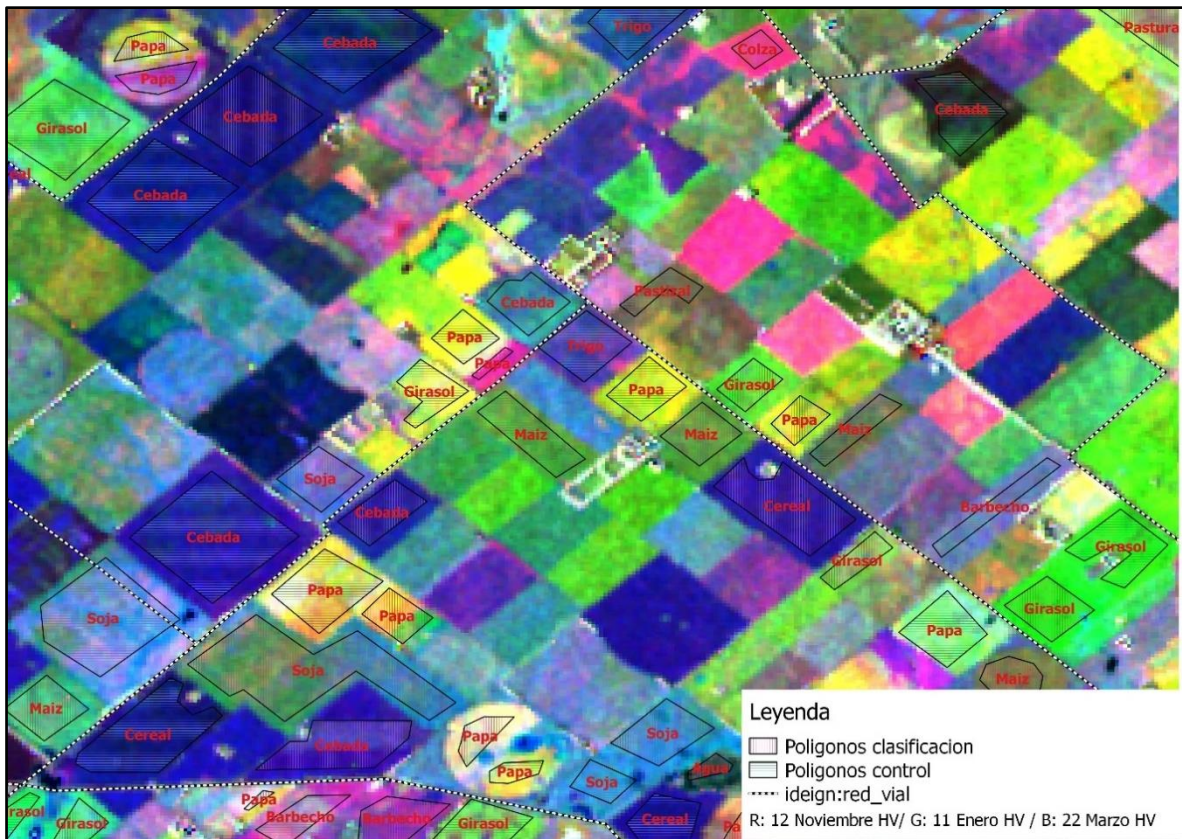


Figura 11. Imagen Composición Color (R:12 Nov HV,G:11 Ene HV, B:22 Mar HV). Imágenes Cosmos Himage del sudeste de la provincia de Buenos Aires. Detalle. Superpuesto: cultivos identificados a campo.

Los índices texturales calculados (Página 40) para cada imagen fueron concatenados y almacenados en un solo archivo raster multibanda. Los algoritmos de cálculo de los índices pueden encontrarse en (Joao Viane Soares, Camilo Daleles Rennó, and Antonio Roberto Formaggio 1997).

3.2.2. Imágenes SAR disponibles

IMAGEN	Modo	NIVEL DE PROCESAMIENTO	Polarizacion	FECHA	ÁNGULO DE INCIDENCIA	ORBITA	MIRADA
1	Himage	SLC	HV	12/11/2014	39.34-41.39	ASCENDENTE	DERECHA
2	Himage	SLC	HV	24/11/2014	39.34-41.40	ASCENDENTE	DERECHA
3	Himage	SLC	HV	11/01/2015	39.34-41.41	ASCENDENTE	DERECHA
4	Himage	SLC	HH	12/01/2015	39.34-41.42	ASCENDENTE	DERECHA
5	Himage	SLC	HV	22/03/2015	32.43-34.83	ASCENDENTE	DERECHA

Tabla 2. Imágenes CSK-Himage recibidas del sudeste de la provincia de Buenos Aires.

Las imágenes Cosmo SkyMed solicitadas no fueron adquiridas en su totalidad. En su lugar se obtuvieron cinco imágenes distribuidas a lo largo de la estación de crecimiento de las variedades de papa utilizadas en la zona de estudio. Dichas imágenes se receptaron en nivel de procesamiento SLC (por sus siglas en inglés, Single Look Complex) (Tabla 2). El ángulo de incidencia fue cercano a los 40°, excepto para la última imagen del mes de Marzo, cuyo ángulo de incidencia fue de 33°. La órbita para todas las imágenes fue ascendente y la mirada del haz

electromagnético fue hacia la derecha. Las imágenes Cosmo SkyMed fueron obtenidas a través de la Comisión Nacional de Actividades Espaciales (CONAE).

Se utilizaron imágenes Landsat 8 del path: 224, Row: 86 de las fechas: 26 de Noviembre de 2014, 12 de Diciembre de 2014, y 14 de Febrero de 2015, todas ellas no presentaban nubosidad. De las mismas se utilizaron las bandas de la 2 a la 7 del sensor OLI. Se utilizaron esas bandas por ser las correspondientes a las porciones del espectro electromagnético visible (bandas 2, 3, 4), infrarojo cercano (banda 5) e infrarojo medio (bandas 6 y 7). Esas porciones del espectro aportan información valiosa en lo que respecta a la clasificación de cultivos. Las imágenes fueron cortesía de la Oficina de Registros Geológicos de los Estados Unidos (USGS).

Las fechas de adquisición de las imágenes no coinciden con las fechas de las recolecciones de datos de campo. Esto no impide la utilización de los datos de campo, dado que el tipo de cobertura (ej. Soja, maíz, papa, etc) no varía dentro del rango de fechas de adquisición de las imágenes. Por lo tanto, no es necesario que el dato de campo sea tomado al mismo tiempo que la imagen.

3.2.3. Pre-procesamiento

El pre-procesamiento de las imágenes Cosmo tuvo por objetivos: hacer la imagen radar lo más parecido posible a una imagen óptica en cuanto a las características de homogeneidad radiométrica (disminuyendo el ruido "Speckle"), y georreferenciar las imágenes de manera de poder co-registrarlas, superponerlas en un SIG (sistema de información geográfica), y poder realizar clasificaciones estadísticas (Figura 12). Para ello se debió calibrar a Intensidad Sigma Cero (Fórmula 1) los valores digitales de las imágenes, georreferenciarlas con datos orbitales y mediante correcciones geométricas confeccionadas utilizando un modelo digital de elevación (SRTM de 90m, por sus siglas en inglés: Shuttle Radar Topography Mission, cortesía del USGS), y finalmente reducir el ruido "Speckle" producto de la generación misma de la imagen SAR. La calibración a sigma cero de todas las imágenes Cosmo, se realizó con el fin de hacerlas comparables entre sí. La reducción del ruido "Speckle" se realizó en dos pasos, primeramente utilizando un proceso conocido como "Multilooking" y en segundo lugar un filtrado de Lee posterior a la corrección geométrica. El pre-procesamiento se realizó mediante el programa NEST 5.1 de la Agencia Espacial Europea. Los parámetros utilizados para el pre-procesamiento se describen en el Anexo 3.

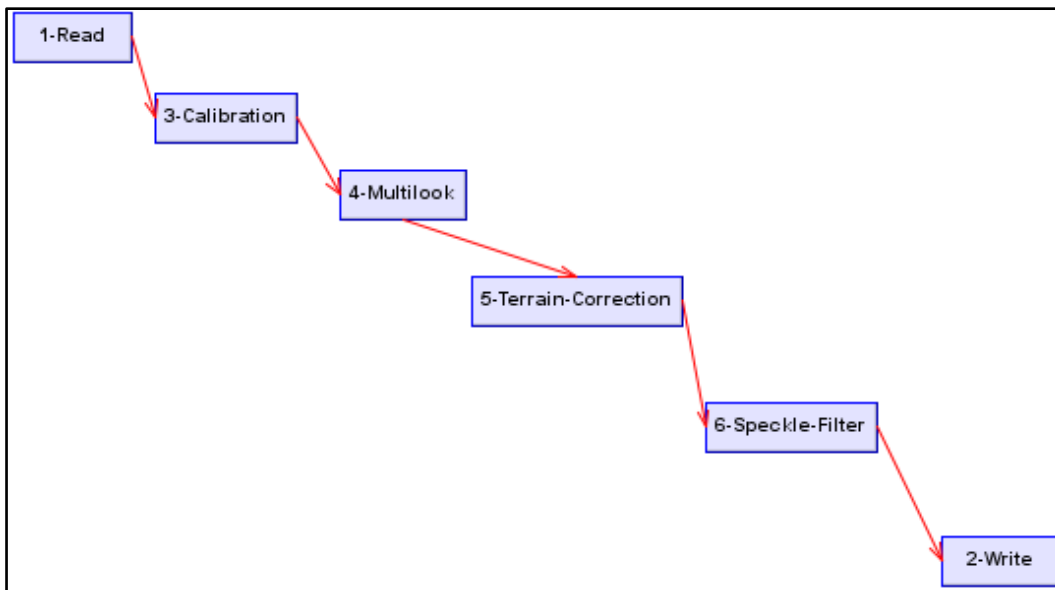


Figura 12. Grafo de procesos aplicados a las imágenes CSK-Himage con el programa Nest. Los parámetros se pueden ver en el Anexo 3.2.

Las bandas 2 a 7 del sensor OLI perteneciente al satélite Landsat 8 fueron calibradas a reflectividad y concatenadas en un solo archivo multibanda.

Finalmente mediante el programa Sentinel-toolbox de la Agencia Espacial Europea se calcularon índices texturales de las imágenes Cosmo sobre una ventana de 7x7 píxeles, lo que equivale aproximadamente a una ventana de 42 metros x 42 metros, dado que el píxel de las imágenes Cosmo luego del filtrado es de 6x6 metros. El uso de una ventana de 7x7 píxeles (49 píxeles) obedece al hecho de poder estimar los índices texturales sobre un número de datos lo suficientemente alto como para poder representar la textura de una superficie pequeña en relación al tamaño promedio de un lote para la región de estudio (50 ha), de manera de no degradar los límites de los lotes agrícolas. Los índices calculados fueron:

- Máxima Probabilidad (Max)
- Homogeneidad
- Varianza
- Media
- Entropía
- Energía
- Disimilaridad
- Segundo Momento Angular (ASM)

Las fórmulas matemáticas de los índices texturales puede encontrarse en Joao Vianeí Soares et col. 1997.

3.2.4. Procesamiento

El procesamiento consistió en realizar clasificaciones supervisadas de seis sets de datos confeccionados a partir del pre-procesamiento. Las clasificaciones supervisadas fueron realizadas de forma no automática por un operador. Los sets clasificados fueron:

1. Cinco imágenes Cosmo SkyMed Himage de distintas fechas, polarizaciones HH y HV. (5 bandas)(Tabla 2)
2. Tres imágenes Landsat 8 OLI (bandas 2 a 7) sin nubes. (18 bandas)
3. Los dos sets anteriores concatenados (CSK+L8).(23 bandas)
4. Una imagen de Febrero de 2015 Landsat OLI (bandas 2 a 7) sin nubes. (6 bandas)
5. Una imagen de Febrero de 2015 Landsat OLI (bandas 2 a 7) más las cinco imágenes CSK. (11 bandas)
6. Los ocho índices texturales de cada una de las cinco imágenes CSK concatenados (8x5 bandas=40 bandas).

Los sets se conformaron con los siguientes objetivos:

- A. Los sets uno y dos fueron elegidos para comparar las clasificaciones de imágenes SAR vs. Imágenes ópticas L8,
- B. Los sets número tres a cinco permitirían indagar el efecto combinado de los datos SAR y ópticos sobre la precisión de la clasificación efectuada,
- C. El set número seis, se realizará con el propósito de probar los índices texturales en la clasificación de datos SAR.

Los seis sets fueron clasificados mediante el algoritmo “Máxima verosimilitud” implementado en el programa ENVI 4.8®. No se le asignó una probabilidad diferencial a ninguna capa, y el factor de re-escalado permaneció igual a 1 dado que los valores no necesitaban ser re-escalados. Adicionalmente el set número uno (cinco imágenes CSK de distintas fechas corregistradas y concatenadas) fue clasificado mediante clasificadores no paramétricos implementados en el programa Monteverdi2-Orfeo Tool Box® de la Agencia espacial francesa (CNES). La descripción de los algoritmos clasificadores encuentra en el capítulo 1.3.1. La implementación de los algoritmos esta descrita en “The Orfeo Toolbox Software Guide” (OTB development team 2014) y una descripción más detallada puede encontrarse en (Open Source Computer Vision Library, n.d.). Los clasificadores no paramétricos utilizados fueron:

1. “Decision tree”
2. “Random Forest”
3. “Gradient Boosted tree”
4. “Lib Support Vector Machine”
5. “Neural Network”

A partir de los índices texturales calculados sobre las imágenes SAR, se realizaron clasificaciones de máxima verosimilitud probando distintas combinaciones de índices con el fin de ver la performance de los mismos.

Una vez realizadas las clasificaciones se calculó la precisión global de cada clasificación, y los errores de omisión y comisión para cada clase. El set de datos de entrenamiento fue del 50% del total de datos relevados mientras que el otro 50% de los datos fue utilizado para el cálculo de la precisión y los errores. En la Figura 10 se pueden observar ambos conjuntos de datos de campo.

3.3. Resultados

Las clasificaciones, por tipo de set clasificado, arrojaron una precisión global de entre un 89% (set #2, set #3, set #5, máxima verosimilitud) y un 72% (set #6, máxima verosimilitud). La precisión global más alta se alcanzó utilizando todas las imágenes Cosmo Himage disponibles en combinación con una o más imágenes Landsat 8 y en el set de tres imágenes Landsat (sets #5, #3, #2, máxima verosimilitud, precisión de 89%) (ver figura 13). Tanto los sets de tres imágenes Landsat 8 (set #2) como el set de una imagen Landsat más cinco imágenes Cosmo Himage (set #5), dieron resultados similares.

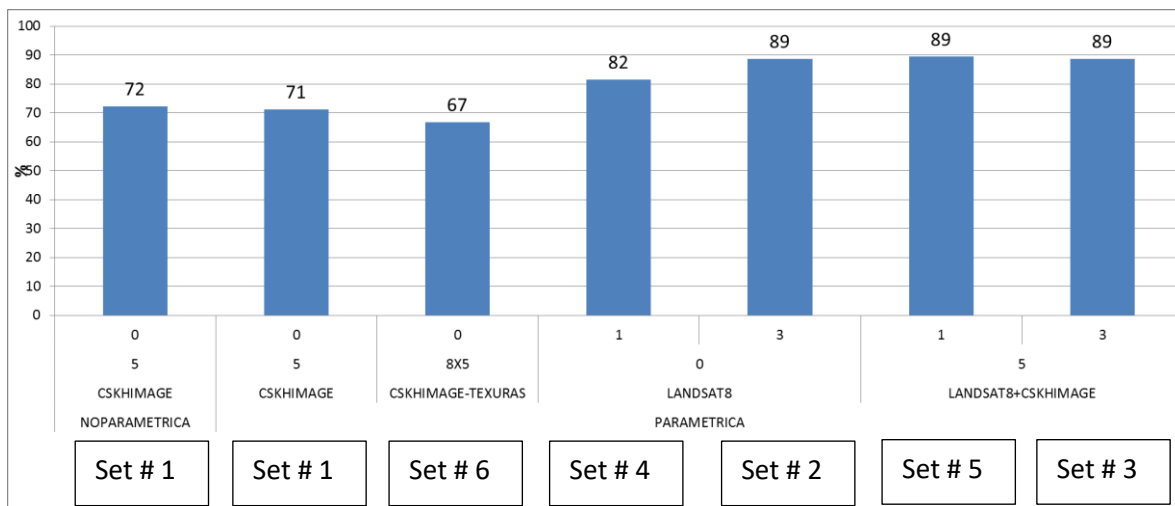


Figura 13. Máxima Precisión alcanzada por las clasificaciones supervisadas sobre los distintos sets de imágenes utilizadas.

En cuanto a los resultados por clasificador, se observa una leve superioridad del clasificador no paramétrico "Gradient Boosted Tree" sobre el clasificador "Máxima verosimilitud", sin embargo la diferencia entre ambos es despreciable. El peor desempeño lo presentó el clasificador "Decision Tree", el cual generó un único árbol de decisión (ver Árbol de Decisión (DT) 1).

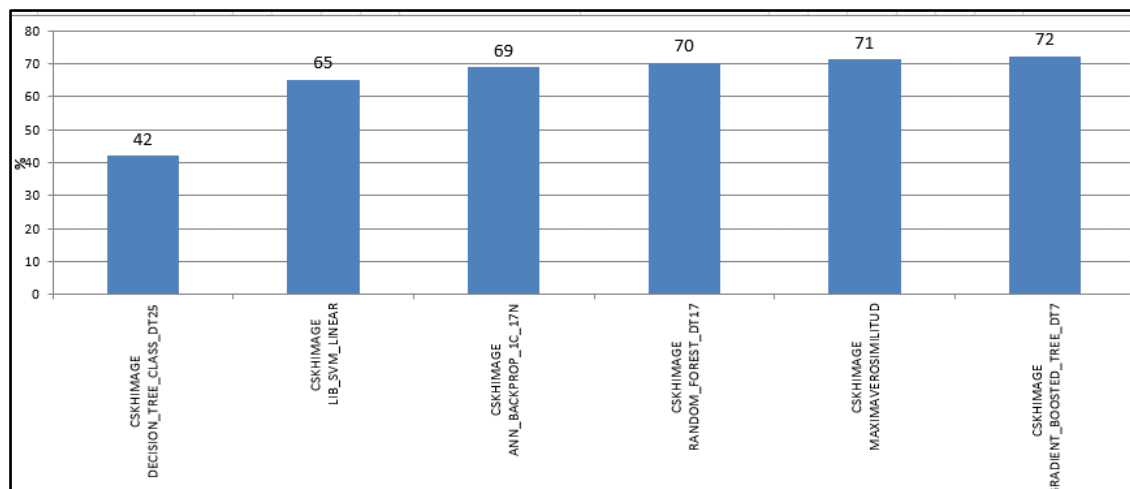
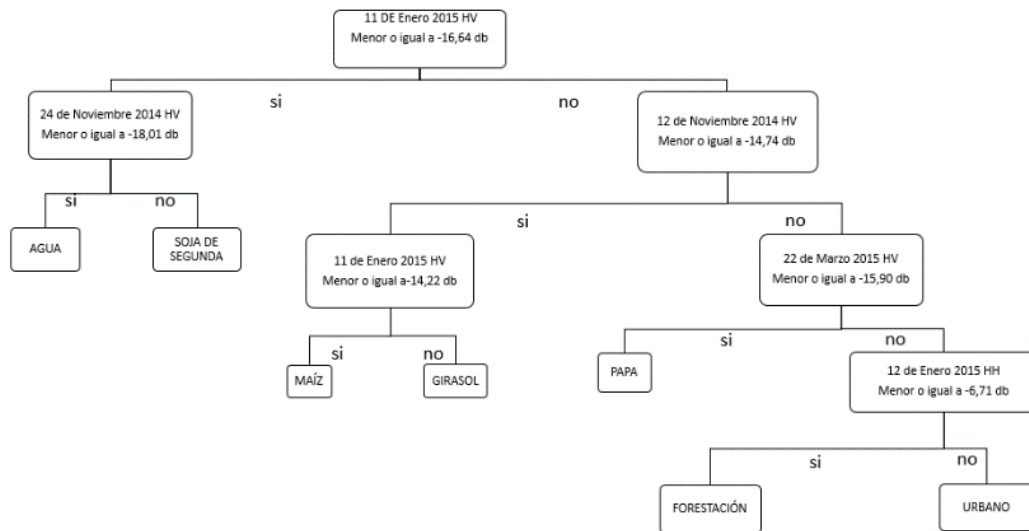


Figura 14. Precisión Global en porcentaje de los clasificadores no paramétricos y de máxima verosimilitud sobre set #1 (cinco imágenes Cosmo Himage). De izquierda a derecha: "Decision tree", profundidad de

árbol: 25; "Support Vector Machine", basado en la librería LibSVM, kernel: lineal; "Artificial Neural Network", backpropagation, 1 capa oculta, 17 neuronas; "; "Random Forest", profundidad de árbol: 17; "Máxima Verosimilitud" algoritmo implementado en programa ENVI ;"Gradient Boosted Tree", profundidad de árbol:7. La cantidad de pixeles de muestra con los que se entrenó a los clasificadores fue de 2022.

El clasificador DT, presentó errores muy altos, esto se debe en parte por que no pudo incluir en el árbol generado a tres de las diez clases de entrenamiento ("Bajos anegables", "Pastizales", "Soja de primera"). Sin Embargo, es interesante observar el modelo generado por el clasificador (ver Árbol de Decisión (DT) 1). El modelo es sencillo y permite observar la relación existente entre las clases. En primer término, el modelo divide las clases "Agua" y "Cereal de Invierno/Soja de segunda" del resto de las clases. En una segunda división separa las clases "Maíz" y "Girasol" de las clases con retrodispersión más altas (ver Anexo 2): "Cultivo de Papa", "Urbano", y "Forestación".



Árbol de Decisión (DT) 1. Árbol de decisión generado por el clasificador DT. No fueron clasificadas todas las clases. Se encuentran ausentes las clases: Bajos Anegables, Soja de primera, y Pastizales. Cada nodo presenta la fecha de la imagen y su correspondiente polarización, debajo se muestra el límite tomado para dividir el nodo en dos. Los valores de intensidad se encuentran expresados en decibeles. El árbol se desarrolla con respuesta afirmativa a la condición en cada nodo hacia la izquierda.

Finalmente, los índices texturales clasificados arrojaron distinto resultado según la combinación utilizada. La mejor combinación de índices se logró con alguna combinación de la Media, la Homogeneidad , la Energía y la Entropía (72% de precisión global) (ver Figura 15).

Clasificando unicamente la Media y la Entropía se alcanzó prácticamente el mismo resultado que utilizando los cuatro índices mencionados anteriormente.

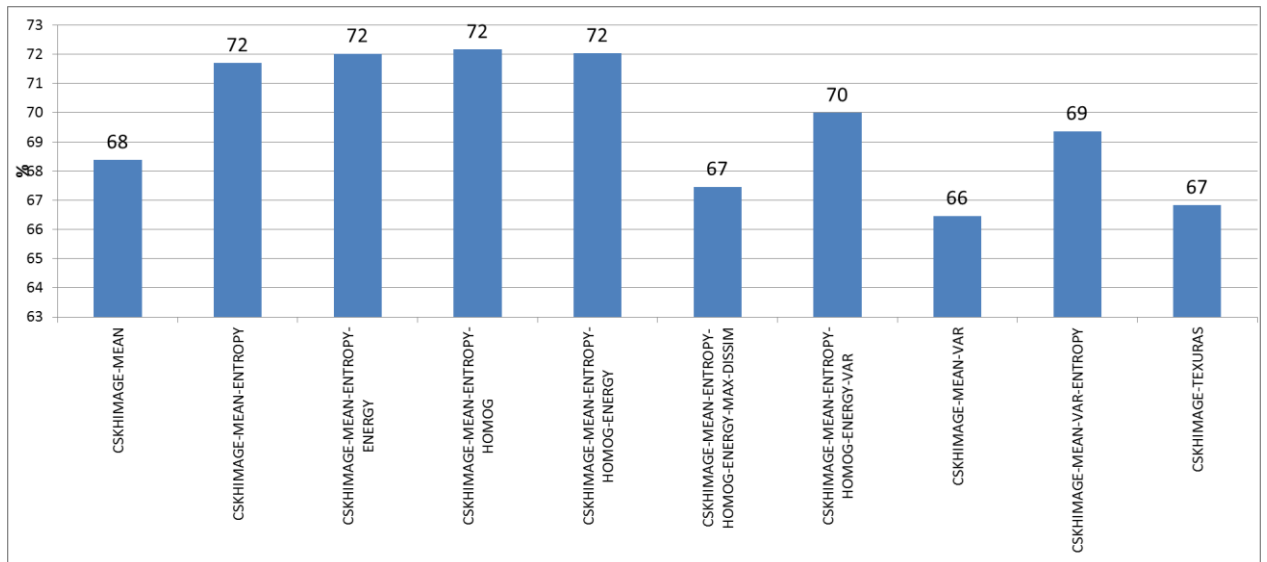


Figura 15. Máxima precisión global de los clasificadores sobre set de índices texturales calculados sobre imágenes Cosmo Himage (set #6).

En cuanto a los errores por clase de cultivo se observa una disminución marcada en los errores de comisión y omisión al pasar de la clasificación supervisada de una imagen Landsat 8 (bandas 2 a 7, set #4) a la clasificación de una imagen Landsat 8 más cinco imágenes Cosmo Himage (set #5). Luego el agregado de imágenes Landsat 8 no modifica sustancialmente los errores calculados (ver Figura 16). La clasificación de tres imágenes Landsat 8 sin nubes arrojó el mismo nivel de error que la clasificación de las mismas tres imágenes más cinco imágenes Cosmo Himage (ver Figura 16). Además la variabilidad de los errores fue mayor en el set #1 y en el set #4, lo que estaría influenciado por la menor a cantidad de información contenida en esos sets, (5 y 6 bandas vs. 11 bandas y más).

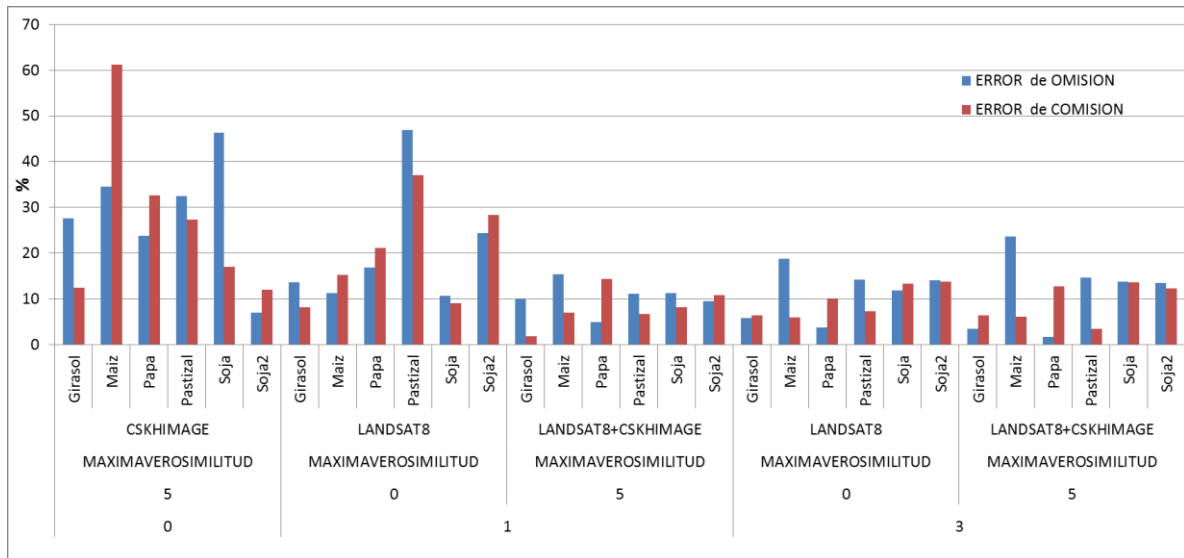


Figura 16. Clasificaciones del sudeste de Buenos Aires. Errores de Omisión y Comisión en las coberturas de cultivos, para las clasificaciones de máxima verosimilitud, sobre los sets #1, #4, #5, #2, #3. Los números de la primera fila indican el número de imágenes Cosmo, y los de la segunda el número de imágenes Landsat 8.

La tabla 3 muestra los errores obtenidos para la clase “papa” en los tres mejores clasificadores (ver Figura 14) de las clasificaciones realizadas sobre el set #1 (cinco imágenes Cosmo SkyMed Himage). Se observa una diferencia a favor del clasificador no paramétrico “Gradient Boosted Tree”.

Clase: Papa	Error de Comisión	Error de Omisión
Gradient Boosted Tree	0,2	0,21
Máxima Verosimilitud	0,33	0,24
Random Forest	0,34	0,17

Tabla 3. Errores de Omisión y Comisión de la clase “papa” para los tres mejores clasificadores. Los errores corresponden a las clasificaciones realizadas sobre el set #1 (cinco imágenes Cosmo SkyMED, Himage)

Deteccion cultivos de papa campaña 2014-2015

Superficie cultivada detectada de papa 9162 ha

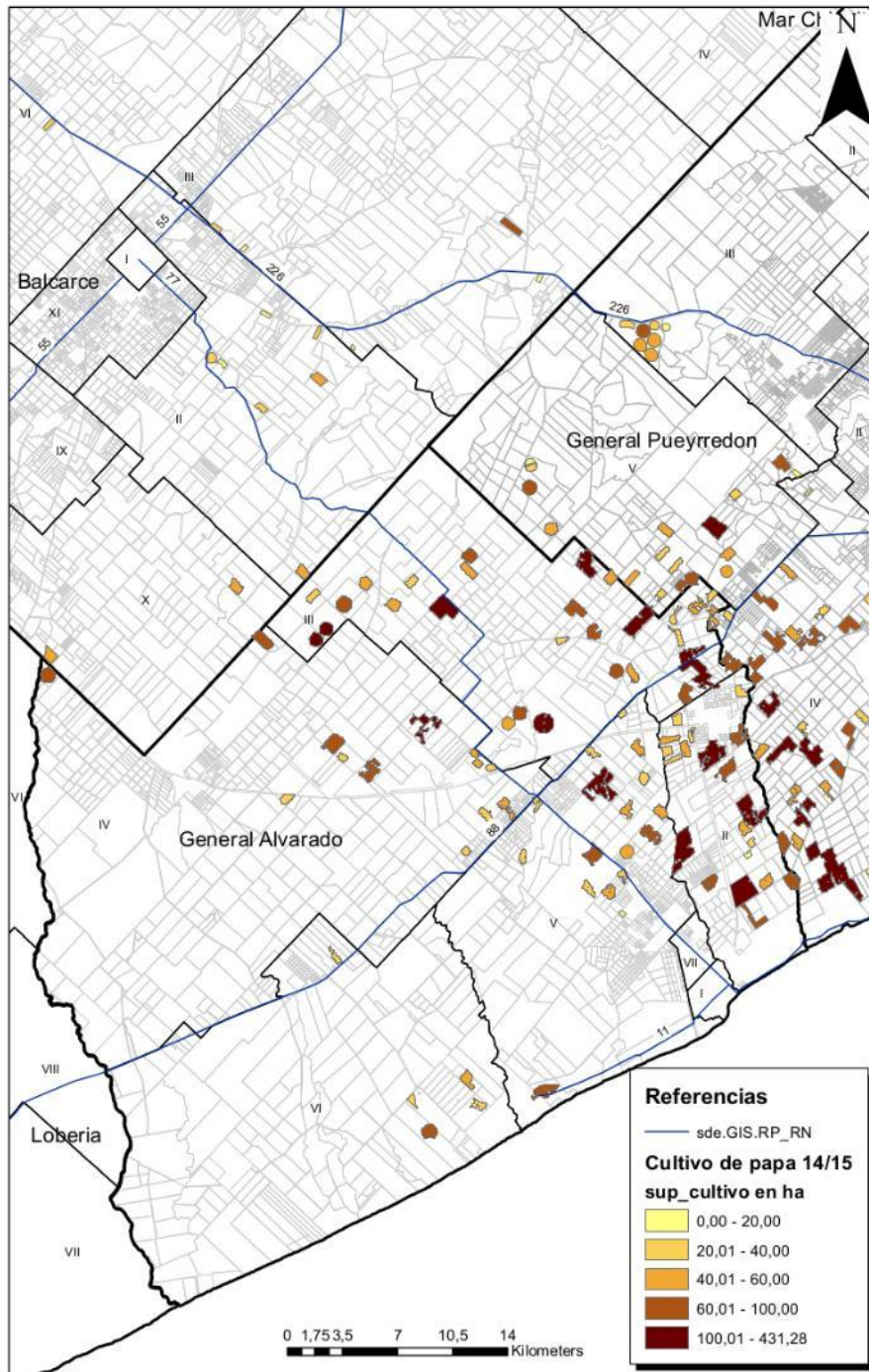


Figura 17. Parcelas clasificadas con cultivo de papa detectadas de forma remota.

3.4. Discusión

La clasificación de cultivos de secano (soja, maíz, girasol) y bajo riego (cultivo de papa) utilizando imágenes SAR arroja resultados aceptables. En este capítulo se observan resultados similares a los obtenidos en investigaciones anteriores realizados en el departamento de T.I. de ARBA, y los encontrados en la bibliografía (Satalino et al. 2011; Skriver et al. 2011; Loosvelt, Peters, Skriver, Lievens, et al. 2012; Lee, Grunes, and Pottier 2001). Es importante destacar que el agregado de una sola imagen Landsat 8 al set multitemporal de imágenes CSK Himage mejoró la precisión global de la clasificación en un 10%, no siendo superado por el agregado de dos imágenes Landsat 8 adicionales. Los resultados en términos de precisión y errores sugieren que el solo uso de hasta cinco imágenes SAR Cosmo Himage no sería suficiente para una correcta clasificación de cultivos en el sudeste bonaerense. También nos sugiere que ante una cantidad escasa de imágenes ópticas Landsat, el agregado de imágenes SAR puede mejorar sustancialmente la clasificación como se observa en las Figura 13 y Figura 16. En particular el cultivo de papa no llegó a ser clasificado correctamente utilizando solamente imágenes SAR banda x (Cosmo) dado que los errores de omisión y comisión se encontraron cercanos al 20% (en el set #1) (ver Tabla 3). No obstante, en el futuro, debería probarse nuevamente la clasificación de cultivos de papa utilizando imágenes SAR en banda C Sentinel-1 doble polarización, (a la fecha de realización de esta tesis no se encontraba disponible).

En cuanto al uso de clasificadores no paramétricos, no mejoraron significativamente la precisión global de la clasificación (Figura 14), sin embargo, el clasificador “Gradient Boosted Tree” sí mejoró en particular la clasificación del cultivo de papa (Tabla 3). Tampoco mejoró la clasificación el uso de índices texturales (Figura 15). Es destacable, que los índices que mejor se comportaron fueron la Entropía y la Media. Sin embargo, no se justifica su utilización dado que no incrementaron la precisión global del mapa con respecto al uso de las imágenes SAR en valores de intensidad.

3.5. Conclusiones

1. A los fines prácticos, el uso de imágenes SAR para la clasificación de cultivo de papa en el sudeste bonaerense se justifica en campañas donde la cantidad o calidad de imágenes ópticas Landsat es baja. En caso contrario su utilización no garantizaría una mejoría en la clasificación.
2. El uso de clasificadores no paramétricos sobre imágenes SAR no mejoró significativamente la precisión global comparándolos con el clasificador paramétrico de máxima verosimilitud.
3. El clasificador no paramétrico “Gradient Boosted Tree” de profundidad de árbol igual a 7 y 200 iteraciones presentó errores más bajos en la clasificación de cultivos de papa.

Finalmente, el presente capítulo arrojó un producto cartográfico que fue utilizado por la Agencia de Recaudación de Buenos Aires (ARBA), en sus tareas de fiscalización masiva del sector agropecuario (Figura 17).

Capítulo 4. Cultivo de granos en el Norte de la provincia de Buenos Aires

4.1. Introducción

La detección de parcelas cultivadas en la provincia de Buenos Aires es parte de un proceso de fiscalización masiva para el recupero del impuesto al ingreso bruto percibido por el fisco provincial. En la actualidad dicho proceso es llevado a cabo por el departamento de Tecnología de Imágenes de la Agencia de Recaudación de Buenos Aires. Este capítulo, así como el próximo intentarán aportar herramientas para mejorar la metodología de clasificación de cultivos actualmente vigente, utilizando información de satélites SAR.

El cultivo de grano más importante en la provincia de Buenos Aires por su volumen y valor actuales, es el cultivo de soja, el cual puede desarrollarse en un ciclo largo con siembras en Octubre a Noviembre y o bien en un ciclo corto o llamado “de segunda”, el cual se desarrolla de Enero a Abril-Mayo. Normalmente la soja de ciclo corto viene precedida por un cultivo de invierno como el trigo o la cebada. Le sigue en importancia el cultivo de maíz, el cual tiene fuerte presencia en el norte de la provincia, en el sitio de estudio del presente capítulo. Las áreas de pastizales y pasturas dedicadas a la ganadería también están presentes en la zona y se encuentran normalmente circunscriptas a zonas bajas, depresiones, y llanuras de inundación.

4.2. Materiales y Métodos

La clasificación de cultivos de grano en el norte de la Provincia de Buenos Aires se llevó a cabo mediante la utilización de sets multitemporales de imágenes Cosmo-SkyMed modo Wide Region. Las clasificaciones fueron de carácter estadístico supervisado para lo cual se realizó una recolección de datos de campo en los meses de Noviembre a Mayo. Los programas utilizados para el entrenamiento y clasificación fueron el Envi 4.8 y el Monteverdi 2 de la CNES (Centre National d'Études Spatiales). Para el preprocesamiento de las imágenes Cosmo se utilizó el programa NEST 5.1 de la ESA (Agencia Espacial Europea). Adicionalmente, se clasificaron sets de imágenes Landsat 8 y sets combinados de imágenes Cosmo Himage e imágenes Landsat 8.

4.2.1. Datos de campo

Los datos de campo utilizados para el entrenamiento de los clasificadores supervisados, se obtuvieron mediante una campaña de recolección de datos llevada a cabo en los días 29 de Octubre de 2014, 25 de Noviembre de 2014, 19 de Marzo de 2015, 7 al 9 de Abril de 2015, y 14 y 15 de Abril de 2015. En total se registraron 700 puntos de observación (ver Figura 18). De esta manera se cubrió el ciclo de cultivo de las distintas especies relevadas. Los datos se almacenaron en una Tablet PC sobre el programa licenciado de ESRI, ArcPad®. A su vez se tomaron fotografías y se incluyeron en los datos registrados de cada parcela.

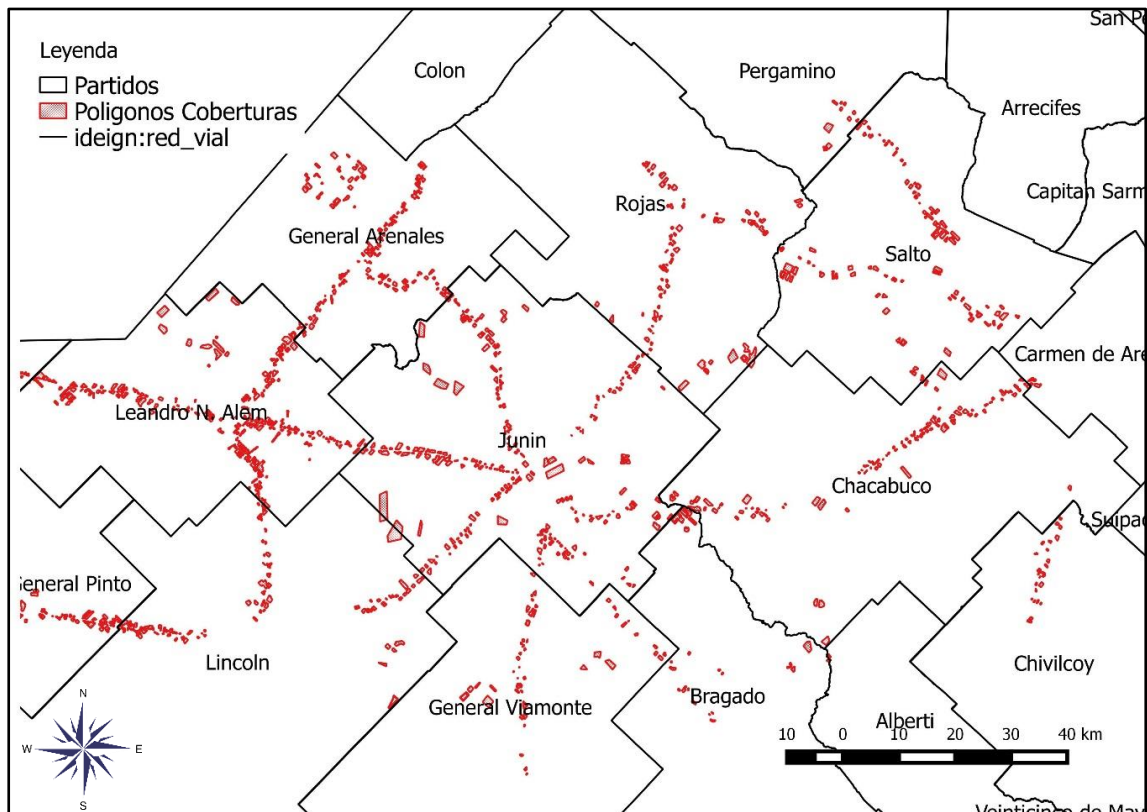


Figura 18. Parcelas relevadas a campo en el norte de la Provincia de Buenos Aires durante la campaña agrícola 2014-2015.

Durante la campaña de recolección de datos se registraron los cultivos observados en cada parcela sobre un recorrido predeterminado en la zona de estudio. Posteriormente en gabinete se dibujó sobre las imágenes SAR los polígonos correspondientes a las parcelas de cultivo identificadas a campo (ver Figura 19).

Las coberturas relevadas fueron agrupadas en ocho clases:

1. Cultivo de Soja de primera
2. Cereal de Invierno/Soja de segunda
3. Cultivo de Maíz
4. Pastizales/Pasturas
5. Urbano
6. Forestación
7. Agua
8. Juncal

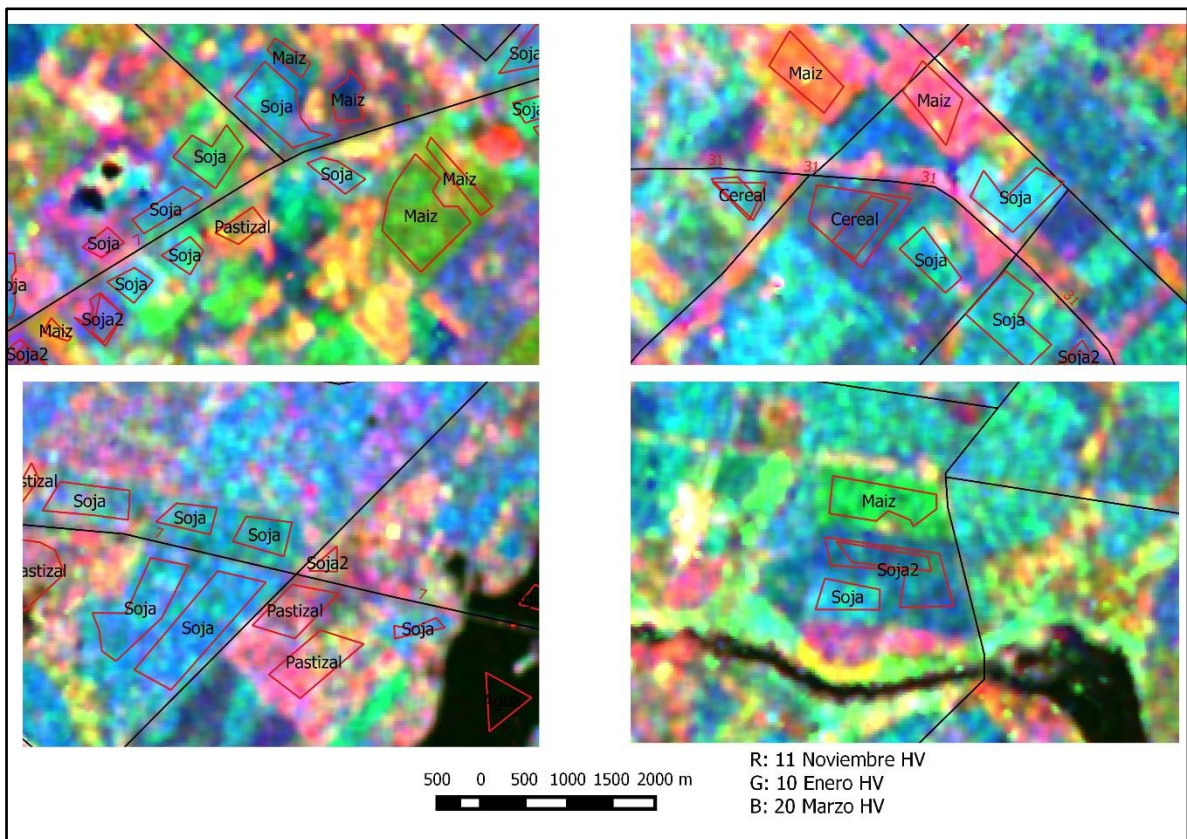


Figura 19. Polígonos de muestreo sobre imagen compuesta RGB:11 Noviembre 2014 HV, 10 de Enero 2015 HV, 20 Marzo 2015 HV.

4.2.2. Imágenes SAR disponibles

IMAGEN	MODO	NIVEL DE PROCESAMIENTO	POLARIZACIÓN	FECHA	ÁNGULO DE INCIDENCIA	ORBITA	MIRADA
1	Himage	GEC	HV	11/11/2014	36.8-43.5	ASCENDENTE	Derecha
2	Himage	GEC	VV	12/11/2014	30.12-38.67	ASCENDENTE	Derecha
3	Himage	DGM	HV	10/01/2015	36.8-43.5	ASCENDENTE	Derecha
4	Himage	DGM	HH	14/01/2015	36.8-43.5	ASCENDENTE	Derecha
5	Himage	DGM	HV	20/03/2015	30.12-38.67	ASCENDENTE	Derecha

Tabla 4. Imágenes Wide Region del norte de la provincia de Buenos Aires.

Las clasificaciones de cultivos se realizaron sobre sets de imágenes SAR, e imágenes SAR más imágenes ópticas. Las clasificaciones de imágenes SAR se realizaron sobre un set

multitemporal de cinco imágenes Cosmo SkyMed en modo Wide Region. Las primeras dos imágenes Cosmo receptoras fueron solicitadas en nivel de procesamiento GEC (por sus siglas en inglés, Geocoded ellipsoid Corrected, es decir corregidas geométricamente por datos de órbita del satélite). Las tres subsiguientes fueron solicitadas en nivel de procesamiento DGM (por sus siglas en inglés, Detected Ground Multilook, las cuales no fueron corregidas geométricamente por el productor de las imágenes) (ver tabla 3). Para poder trabajar con ambos tipos de imágenes se debieron corregir luego de pre-procesarlas. Las polarizaciones elegidas fueron HV a lo largo de toda la campaña, VV al inicio de la campaña por el efecto de atenuación que presentan los cereales de invierno al final de sus ciclo reproductivo, y HH en el mes de enero para explotar el efecto de retrodispersión diferencial de los cultivos de verano la polarización HH (Floyd Henderson and Anthony Lewis 1998).

Las imágenes Landsat 8 utilizadas fueron del path 226 row 84, fechas: 7 de Octubre de 2014, 16 de Marzo de 2015, y 01 de Abril de 2015. Estas imágenes fueron seleccionadas por no presentar nubosidad. De las mismas se utilizaron las bandas de la 2 a la 7 del sensor OLI. Las imágenes fueron cortesía del servicio Geológico de los Estados Unidos (USGS).

4.2.3. Pre-procesamiento

El pre-procesamiento de las imágenes Cosmo Wide Region, resumidos en las figuras 20 y 21, tuvo por objetivo: hacer la imagen radar lo más parecido posible a una imagen óptica en cuanto a las características de homogeneidad radiométrica (ruido "Speckle"), y georreferenciar las imágenes de manera de poder coregistrarlas, superponerlas en un SIG (sistema de información geográfica), y realizar las clasificaciones. Para ello se debió calibrar a Intensidad Sigma Cero los valores digitales de las imágenes, georreferenciarlas con datos orbitales y un modelo digital de elevación (SRTM de 90m, por sus siglas en inglés: Shuttle Radar Topography Mission, cortesía del USGS), y finalmente reducir el ruido "Speckle" producto de la generación misma de la imagen SAR. La reducción del ruido "Speckle" se realizó en dos pasos, primeramente utilizando un proceso conocido como "Multilooking" y en segundo lugar un filtrado de Lee posterior a la corrección geométrica. El preprocesamiento se realizó mediante el programa NEST 5.1 de la Agencia Espacial Europea. Para el caso de las imágenes que llegaron en nivel de procesamiento "GEC" (Tabla 4), no fue posible la georreferenciación por terreno, (ya que el programa no permite esa función sobre ese nivel de procesamiento dado que la imagen se encuentra georreferenciada), y luego de filtrarlas, se coregistraron a las imágenes "DGM" siendo estas últimas las de referencia, dado que han sido corregidas por terreno por parte del generador del producto. La calibración de las imágenes se llevó a cabo mediante la fórmula 1 de calibración descrita en el Capítulo 2.

Las bandas 2 a 7 del sensor OLI perteneciente al satélite Landsat 8 fueron calibradas a reflectividad y concatenadas en un solo archivo multibanda.

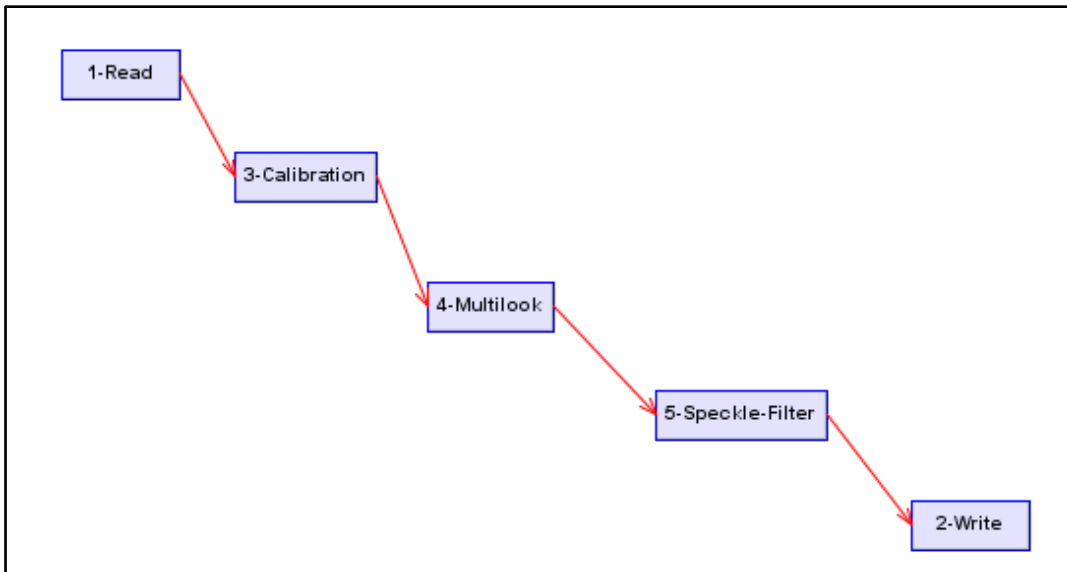


Figura 20. Grafo de pre procesamiento de imágenes nivel "GEC". Se observan los procesos utilizados para normalizar radiométricamente los valores digitales de las imágenes. Los parámetros se pueden ver en el Anexo 3.3.

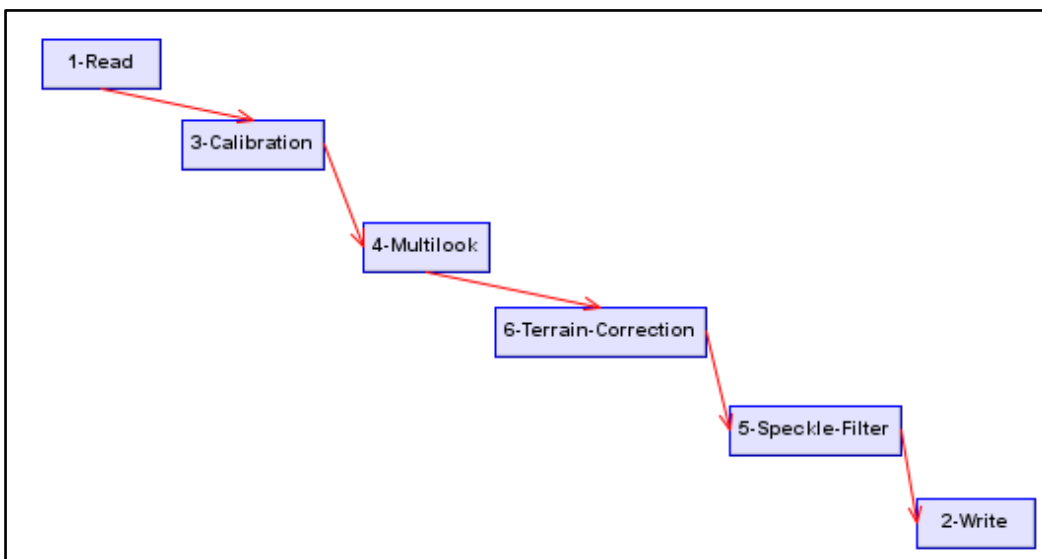


Figura 21. Grafo de pre procesamiento de imágenes "DGM". Se observan los procesos utilizados para corregir geoméricamente, y normalizar los valores digitales de las imágenes. Los parámetros se pueden ver en el Anexo 3.4.

4.2.4. Procesamiento

El procesamiento consistió en realizar clasificaciones supervisadas de cinco sets de datos confeccionados a partir del pre-procesamiento. Las clasificaciones supervisadas fueron realizadas de forma no automática por un operador. Los sets se eligieron de manera de poder comparar la conveniencia de utilizar las imágenes SAR con imágenes ópticas y en cuanto se mejora la clasificación utilizando una imagen óptica o tres imágenes ópticas sumadas a las imágenes SAR. Los sets clasificados fueron:

1. Cinco imágenes Cosmo SkyMed Wide Region de distintas fechas.(5 bandas)(ver Tabla 4)
2. Tres imágenes Landsat 8 OLI (bandas 2 a 7) sin nubes. (18 bandas)
3. Los sets 1 y 2 concatenados (CSK+L8). (23 bandas)
4. Una imagen de Marzo de 2015 Landsat OLI (bandas 2 a 7) sin nubes. (6 bandas)
5. Una imagen de Marzo de 2015 Landsat OLI (bandas 2 a 7) más las cinco imágenes CSK.(11 bandas)

Los sets se conformaron con los siguientes objetivos:

- a. Los sets uno y dos fueron elegidos para comparar las clasificaciones de imágenes SAR vs Imágenes ópticas L8.
- b. Los sets número tres a cinco permitirían indagar el efecto combinado de los datos SAR y ópticos sobre la precisión de la clasificación efectuada.

Los cinco sets fueron clasificados mediante el algoritmo “Máxima verosimilitud” implementado en el programa ENVI 4.8[®]. Adicionalmente el set número uno (cinco imágenes CSK de distintas fechas corregistradas y concatenadas) fue clasificado mediante clasificadores no paramétricos implementados en el programa Monteverdi2-Orfeo Tool Box[®] de la Agencia espacial francesa (CNES). La descripción de los algoritmos clasificadores encuentra en el capítulo 1.3.1. La implementación de los algoritmos esta descrita en “The Orfeo Toolbox Software Guide” (OTB development team 2014) y una descripción más detallada puede encontrarse en (Open Source Computer Vision Library, n.d.). Los clasificadores no paramétricos utilizados fueron:

1. “Decision tree”
2. “Random Forest”
3. “Gradient Boosted tree”
4. “Lib Support Vector Machine”
5. “Neural Network”

Una vez realizadas las clasificaciones se calculó la precisión global de cada clasificación, y los errores de omisión y comisión para cada clase. El set de datos de entrenamiento fue del 50% del total de datos relevados mientras que el otro 50% de los dato fue utilizado para el cálculo de la precisión y los errores.

4.3. Resultados

La precisión global de las clasificaciones fue similar para todos los sets clasificados menos para el set compuesto de una sola imagen Landsat 8 del mes de Marzo de 2015. El set de cinco imágenes Cosmo Wide Region clasificado por medio del algoritmo de máxima verosimilitud presentó un desempeño menor, con una precisión global de 85%, que el set de

tres imágenes Landsat 8 (91%). La combinación de imágenes Landsat 8 y Cosmo Wide Region en una concatenación simple no presentó una precisión significativamente mayor que el set de tres imágenes Landsat 8 (91% vs. 92%). En cambio, se observa una mejoría en la precisión de la clasificación cuando solo se cuenta con una imagen Landsat 8 (69%) y se la combina con cinco imágenes Cosmo Wide Region (90%), mejorando un 20% (ver Figura 22).

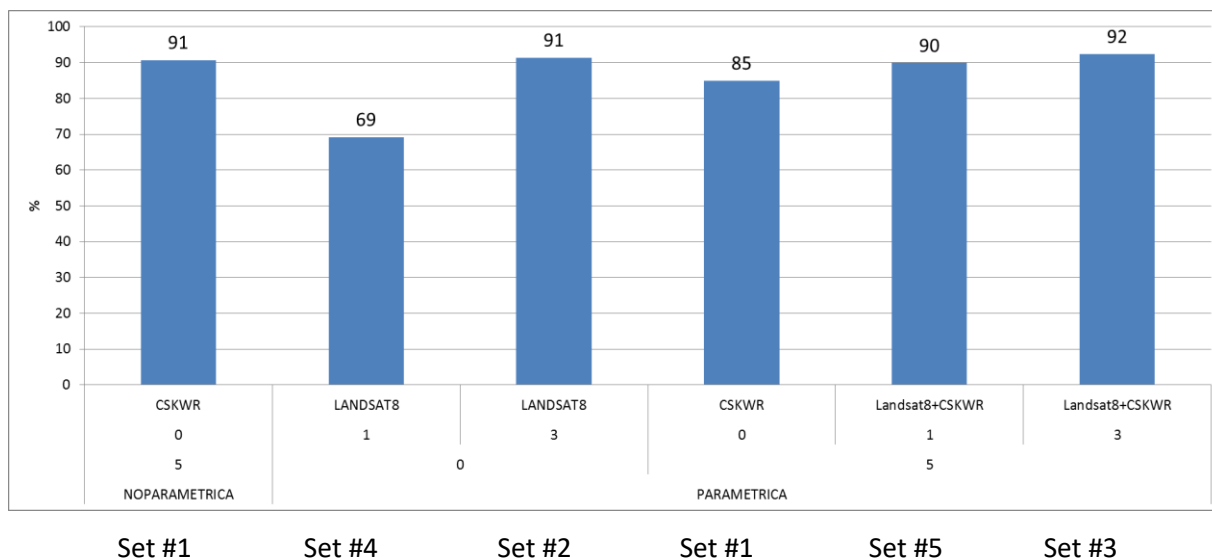


Figura 22. Clasificaciones zona Núcleo Prov. Buenos Aires. Máxima precisión global alcanzada por los distintos sets de imágenes utilizados, en función del tipo de clasificador utilizado (paramétrico y no paramétrico). La primera fila de números corresponde a la cantidad de imágenes Landsat 8. La segunda a la cantidad de imágenes Cosmo.

En cuanto a los clasificadores no paramétricos testeados, casi todos superaron al clasificador paramétrico de “Máxima Verosimilitud” del programa Envi 4.8, con excepción del clasificador “Decision Tree” (ver Figura 23).

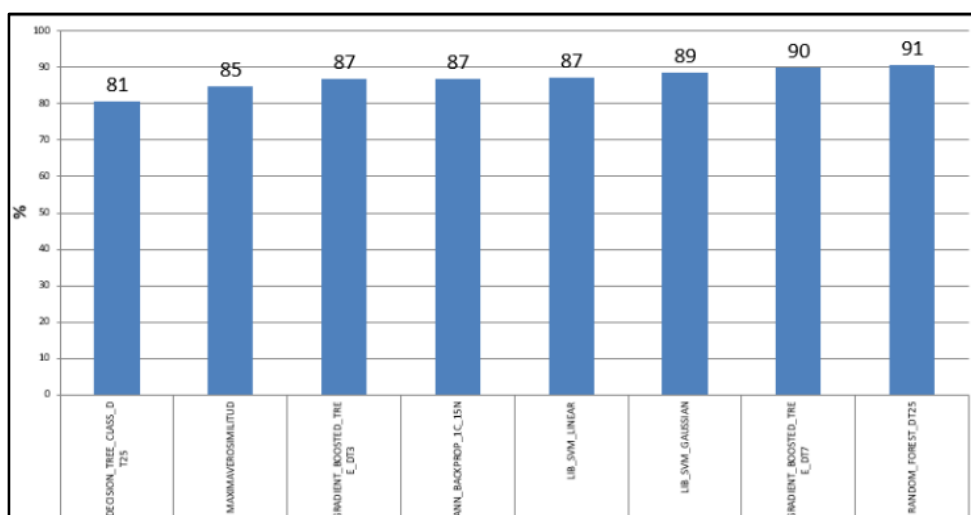
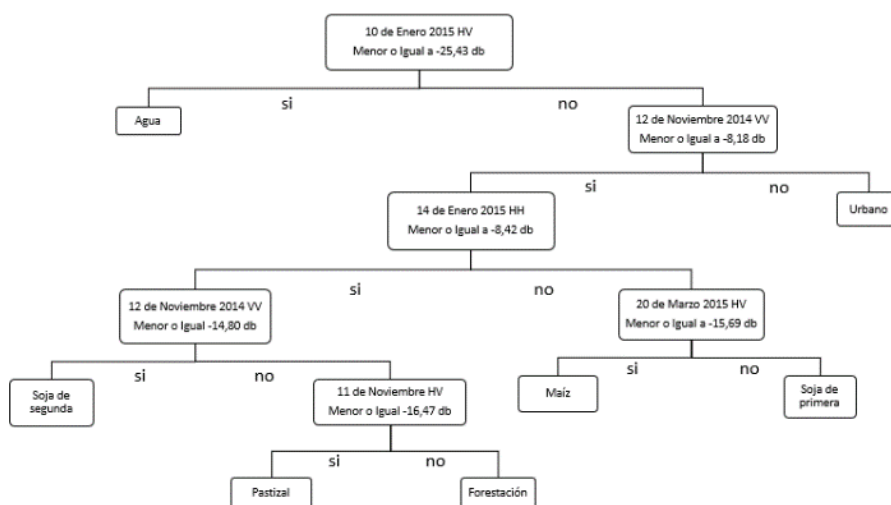


Figura 23. Clasificaciones zona Núcleo Prov. Buenos Aires. Precisión global de las clasificaciones con el programa Monteverdi2-Orfeo Toolbox y el programa ENVI (máxima verosimilitud). De izquierda a derecha: “Decision tree”, profundidad de árbol: 25; “Máxima verosimilitud” (implementado en el programa ENVI 4.8); “Gradient Boosted Tree”, profundidad de árbol: 3; “Artificial Neural Network”, backpropagation, 1 capa oculta, 15 neuronas; “Support Vector Machine”, basado en la librería LibSVM, kernel: gaussiano; “Support Vector Machine”, basado en la librería LibSVM, kernel: lineal; “Gradient Boosted

Tree”, profundidad de árbol:7; ”Random Forest”, profundidad de árbol: 25. La cantidad de píxeles muestra con los que se entrenó a los clasificadores fue de 4092 píxeles.

El clasificador DT generó un árbol de decisión que puede observarse a continuación (ver Árbol de Decisión (DT) 2). El árbol generado es un modelo sencillo y permite observar de forma clara la relación entre las clases. Su desempeño global en términos de precisión fue del 81%. El clasificador “Random Forest” (un combinación de 200 árboles de decisión) alcanzó una precisión global solamente 10 % superior al árbol de decisión generado por el algoritmo DT. En primer término, el árbol generado a partir del clasificador DT, separa la clase “Agua” del resto de las clases. Seguidamente separa la clase “Urbano” y continua sucesivamente con las clases de cultivos : “Soja de segunda”, “Maíz”, “Soja de primera”. Finalmente separa los “Pastizales” de las “Forestaciones”.



Árbol de Decisión (DT) 2. Árbol de decisión. En cada nodo se observa la imagen utilizada y el límite para dividir el nodo. El valor del límite se encuentra expresado en decibeles.

En la Figura 24 se observan los errores de comisión y de omisión de las principales coberturas. En la misma se pueden comparar los errores de la clasificación no paramétrica “random forest” con la realizada mediante el algoritmo de “Máxima verosimilitud”. Las diferencias a favor del algoritmo no paramétrico son importantes. A la derecha, podemos observar que casi no hay diferencias entre los errores de la clasificación realizada sobre las imágenes Landsat 8 y la realizada sobre imágenes Landsat 8 y Cosmo SkyMed. Sin embargo, puede verse una mínima mejora en la clasificación de las clases “Maíz” y “Pastizal” al agregar las imágenes Cosmo SkyMed.

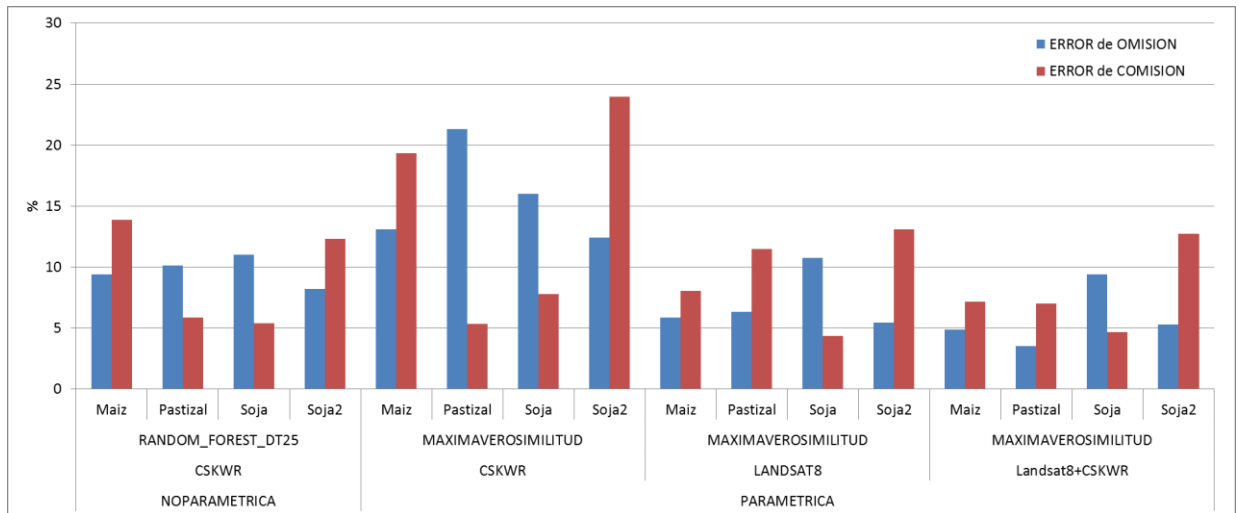


Figura 24. Clasificaciones zona Núcleo Prov. Buenos Aires. Errores de Omisión y Comisión para las principales coberturas clasificadas en la mejor clasificación no paramétrica (random forest), y en las de máxima verosimilitud para los sets de imágenes Cosmo y Landsat 8, y la combinación de ambos.

En la Figura 25 y la Figura 26 se muestra la retrodispersión sensada sobre las coberturas de Maíz y Soja de primera para los meses de Enero y Noviembre. Las mismas están indicando poca diferencia entre las polarizaciones HV y HH para el mes de Enero, y entre las polarizaciones VV y HV para el mes de Noviembre. La poca diferenciación en la retrodispersión entre ambos cultivos se ve reflejada en la matriz de confusión. La poca diferenciación entre ambos cultivos para esa variable es el mayor aporte a los errores de omisión y comisión de los mismos (Soja de primera y Maíz). De hecho, viendo el árbol de decisión generado por el clasificador "DT" observamos que la variable que divide ambos cultivos es la intensidad en la fecha "20 de Marzo".

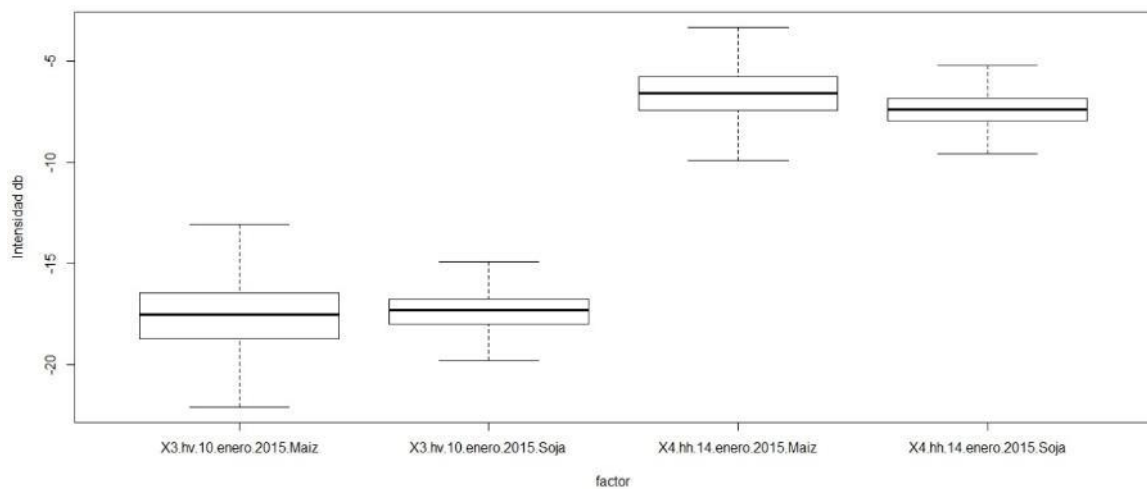


Figura 25. Retrodispersión de los cultivos de Soja y Maíz para el mes de Enero en polarización HH y HV

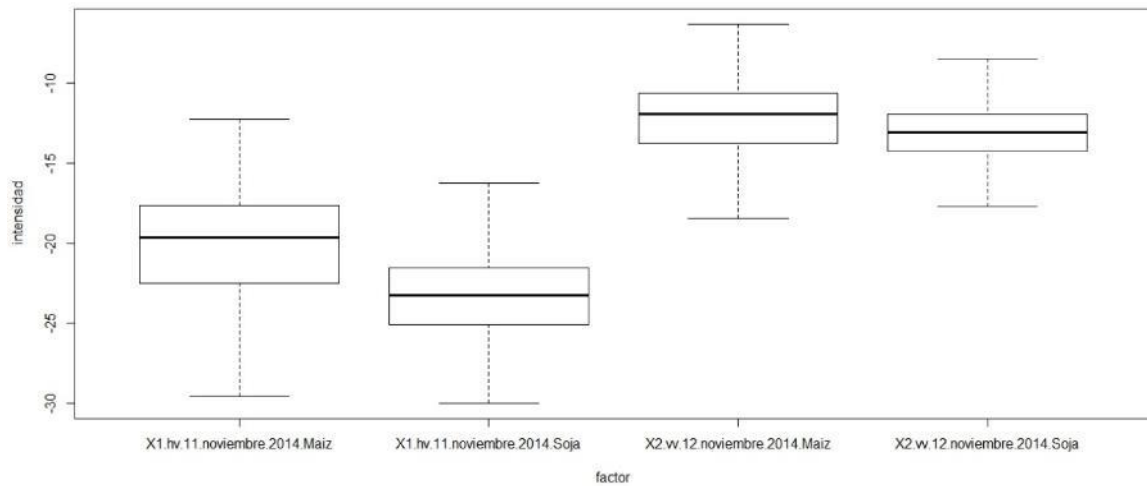


Figura 26. Retrodispersión de los cultivos de Soja y Maíz para el mes de Noviembre en polarización HV y VV

4.4. Discusión

Los resultados de precisión global de las clasificaciones obtenidos en este capítulo coinciden con lo encontrado en la bibliografía (Heather McNairn et al. 2009; Waske and Braun 2009). Las precisiones alcanzadas en las clasificaciones de granos de imágenes SAR Cosmo Wide Region pueden considerarse muy buenas dado que superan el 80% (Joao Viane Soares, Camilo Daleles Rennó, and Antonio Roberto Formaggio 1997). Los clasificadores no paramétricos como “random forest”, implementados en el programa Monteverdi2-OTB del CNES (Centre National d’Études Spatiales), presentaron errores más bajos que un clasificador paramétrico de “máxima verosimilitud” (Figura 22 y Figura 23). En este capítulo se arriba a la misma conclusión que en el capítulo tres: la utilización de imágenes SAR para clasificar cultivos se justifica solamente en el caso de baja cantidad de imágenes ópticas (Figura 22).

La calidad visual, para los fines de esta tesis, es una variable cualitativa subjetiva que toma en cuenta: la precisión en la demarcación de los límites entre lotes y la cantidad de pequeños polígonos de distintas clases inmersos en polígonos más grandes. El clasificador de máxima verosimilitud presentó la más alta calidad visual como se observa en la Figura 27. En la misma, puede apreciarse que no hay diferencias visuales significativas entre las clasificaciones conseguidas con los clasificadores de “máxima verosimilitud” y “random forest” para el set de imágenes SAR de cinco imágenes Cosmo Wide Region. La calidad visual de las clasificaciones de imágenes SAR podría mejorarse con un filtro post clasificación.

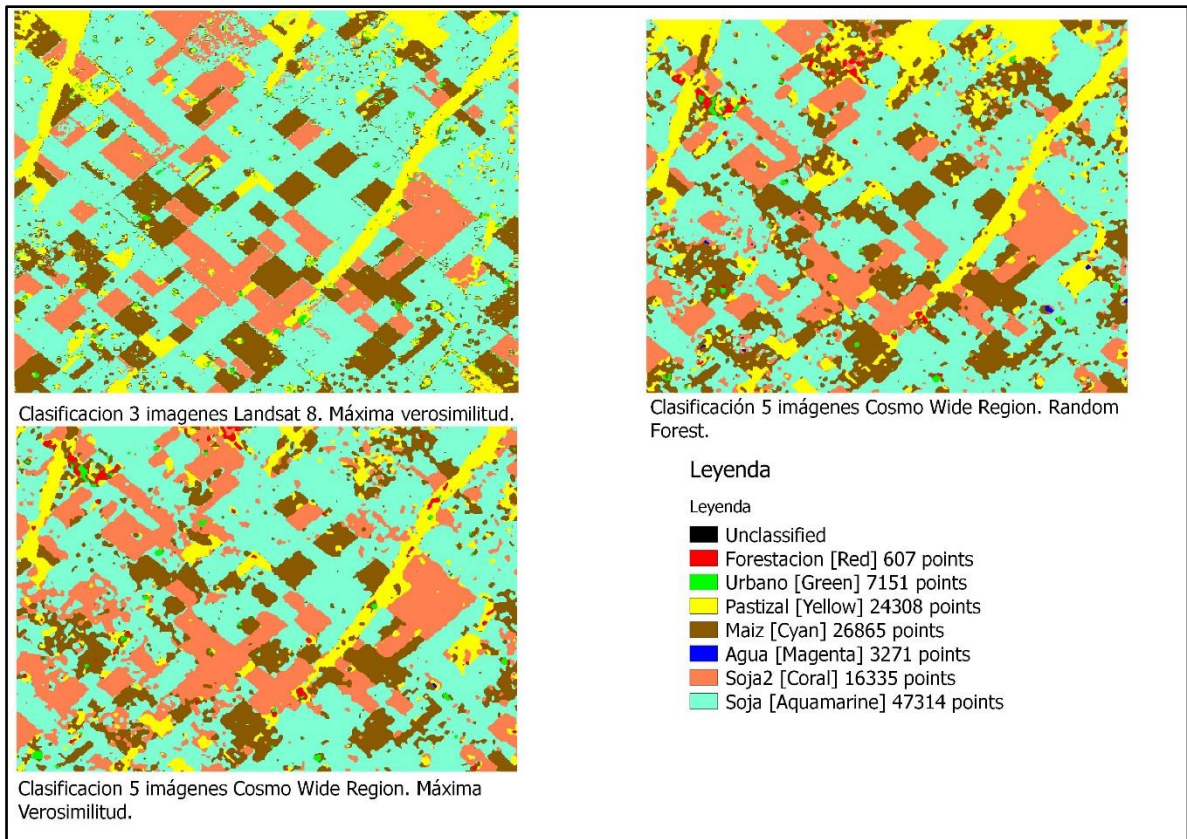


Figura 27. Clasificaciones de cultivos de grano Norte de la Provincia de Buenos Aires.

4.5. Conclusiones

Las conclusiones del presente capítulo se detallan a continuación:

1. La primera conclusión que arroja este capítulo es que es válido utilizar un set multi-temporal de por lo menos cinco imágenes SAR Cosmo Wide Region y distintas polarizaciones para clasificar cultivos con una precisión aceptable (superior al 80%).
2. En segundo lugar, si se opta por utilizar imágenes SAR es preferible utilizar un clasificador no paramétrico. Dentro de los clasificadores no paramétricos se destaca el clasificador “Random Forest”.
3. También es importante destacar que en caso de contar con imágenes ópticas multiespectrales del tipo Landsat 8, su clasificación arrojará mejores resultados desde la calidad visual del mapa obtenido que la clasificación de un set multi-temporal de imágenes SAR (ver Figura 27).
4. Es importante considerar a las imágenes SAR para cubrir posibles faltantes de imágenes ópticas multiespectrales para la implementación de un sistema de clasificación de cultivos preciso y confiable.
5. Por último, no se observan diferencias entre utilizar una configuración del sensor en polarización cruzada HV o copolar HH o VV, en las fechas de Noviembre y Enero, para diferenciar los principales cultivos de la zona, Maíz y Soja (ver Figura 25 y Figura 26).

El presente capítulo sentó las bases para el diseño de una propuesta de mejora al sistema de detección de cultivos utilizado en la actualidad por el departamento de Tecnología

de Imágenes de A.R.B.A (Agencia de Recaudación de la Provincia de Buenos Aires) (ver Anexo 1). Además, se generó un producto cartográfico de coberturas de cultivos en los partidos comprendidos dentro del área de interés (ver Figura 28).

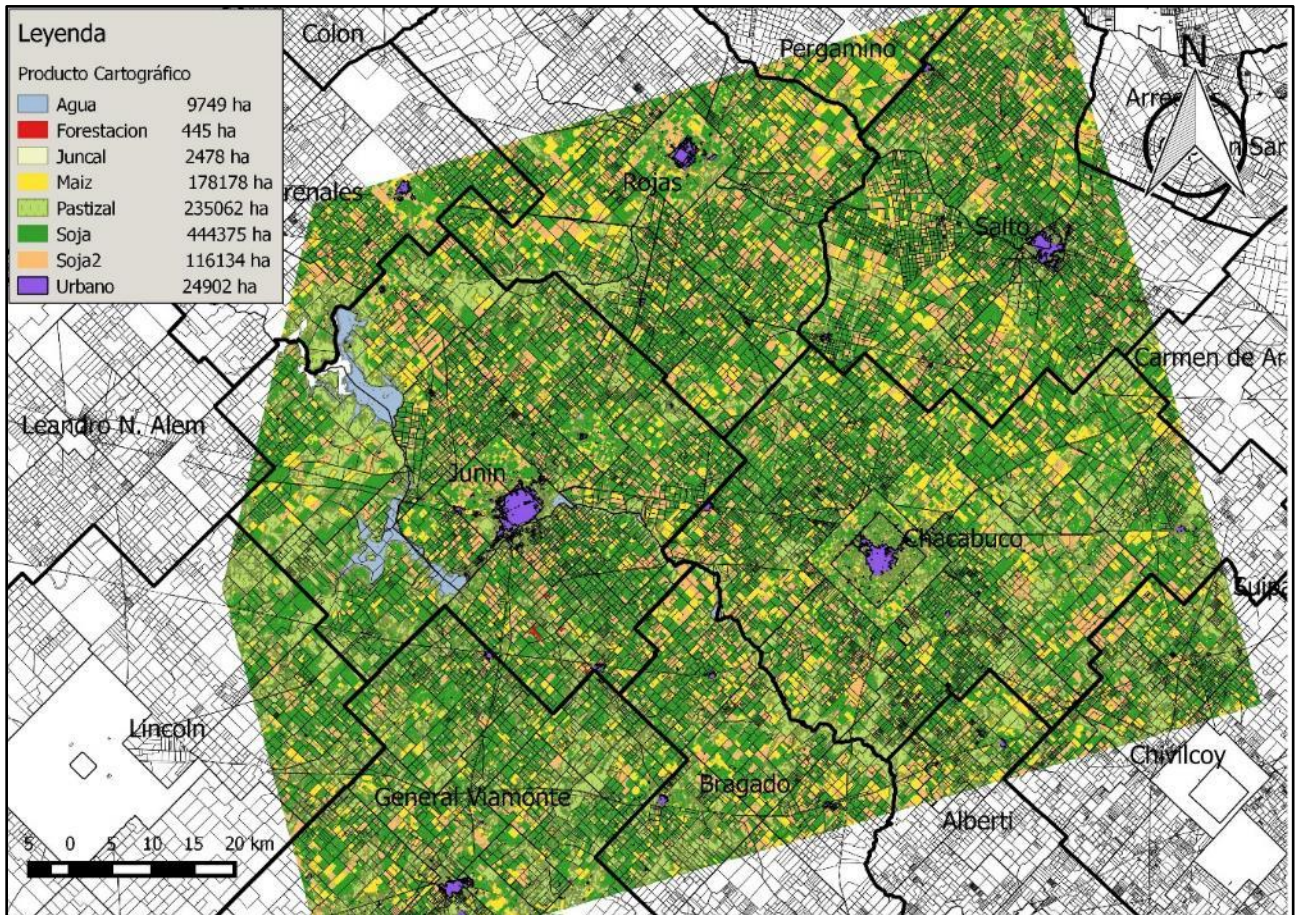


Figura 28. Producto cartográfico. Clasificación de cultivos y la superficie calculada por cobertura. La clasificación corresponde al algoritmo de Máxima Verosimilitud sobre el set de cinco imágenes Cosmo Wide Region y tres imágenes Landsat 8.

Capítulo 5. Cultivo de granos en el Oeste de la provincia de Buenos Aires

5.1. Introducción

El capítulo cinco presentará clasificaciones de cultivos de granos realizadas sobre el Oeste de la provincia de Buenos Aires utilizando las imágenes del nuevo sensor SAR banda C, 5.405 GHz de frecuencia, a bordo del Sentinel-1 A de la Agencia Espacial Europea (ESA).

El sitio de estudio sobre el cual se realizaron las clasificaciones comprende aproximadamente 18 mil kilómetros cuadrados dentro de la provincia de Buenos Aires (ver Figura 1).

Las imágenes Sentinel-1 son imágenes SAR en banda C. Las mismas se encuentran disponibles a través del sitio web de la Agencia Espacial Europea (ESA). El sensor radar genera imágenes de sectores del planeta pre definidas pero también permite programar adquisiciones especiales en función de las necesidades y prioridades de la Agencia Espacial Europea. Uno de los motivos de adquisición de imágenes por fuera de la programación es la ocurrencia de desastres naturales como inundaciones y terremotos. En la programación normal del sensor, el Noroeste de la provincia de Buenos Aires ha sido observado de manera recurrente desde la puesta en funcionamiento del satélite en Agosto de 2014. Las adquisiciones fueron realizadas en polarización VV, en modo “Interferometric Wide” y órbita ascendente. Las ventajas de utilizar este modo es la gran cantidad de imágenes disponibles, dado que a la fecha es el modo que más se ha utilizado sobre la provincia de Buenos Aires, la resolución espacial (5mx20m de pixel), y la cobertura espacial de cada escena (100km x 250 km). Estas características hacen de las imágenes Sentinel 1 IW ideales para su utilización en procesos de fiscalización masiva actividades económicas dispersas en amplios territorios.

Las principales coberturas vegetales de esta región son la soja, el maíz, el girasol, soja de segunda/cereales de invierno, pastizales, y juncuales.

Los objetivos de este capítulo son:

- Estimar la precisión global de una clasificación de cultivos sobre un set de imágenes Sentinel-1 de una sola polarización
- Comparar distintos clasificadores
- Comparar la clasificación de un set de imágenes SAR Sentinel-1 banda C y la clasificación de un set de imágenes multispectrales Landsat 8.

5.2. Materiales y Métodos

Las clasificaciones supervisadas se realizaron en base a distintos sets imágenes Sentinel-1 polarización simple VV, e imágenes Landsat8, y una “verdad de campo” registrada durante la campaña agrícola 2014-2015 en el oeste de la provincia de Buenos Aires. Los programas utilizados fueron el Envi 4.8, y el programa Monteverdi2-Orfeo toolbox del CNES (Centre National d’Études Spatiales)(OTB development team 2014). El pre-procesamiento de las imágenes SAR se realizó con el programa Sentinel1-toolbox 1.1 de la Agencia Espacial Europea. Finalmente, se calcularon la precisión global y los errores de cada clase en cada clasificación. Adicionalmente se utilizaron imágenes Landsat 8 para probar clasificaciones alternativas del sitio de estudio. Las imágenes L8 utilizadas corresponden al path 227, row: 084; de las fechas: 15 de Noviembre de 2014, 2 de Enero de 2015 y 7 de Marzo de 2015.

5.2.1. Datos de campo

Los datos de campo utilizados para el entrenamiento de los clasificadores, se obtuvieron mediante una campaña de recolección de datos, llevada a cabo en los días 12 de Noviembre de 2014, 25 de Noviembre de 2014, 7 al 9 de Abril de 2015, y 21 de Abril de 2015. Durante las campañas de reconocimiento de cultivos a campo se registraron 950 puntos de observación (ver Figura 29). Los datos se almacenaron en una Tablet PC sobre el programa licenciado de ESRI, ArcPad®. A su vez se tomaron fotografías y se incluyeron en los datos registrados de cada parcela relevada.

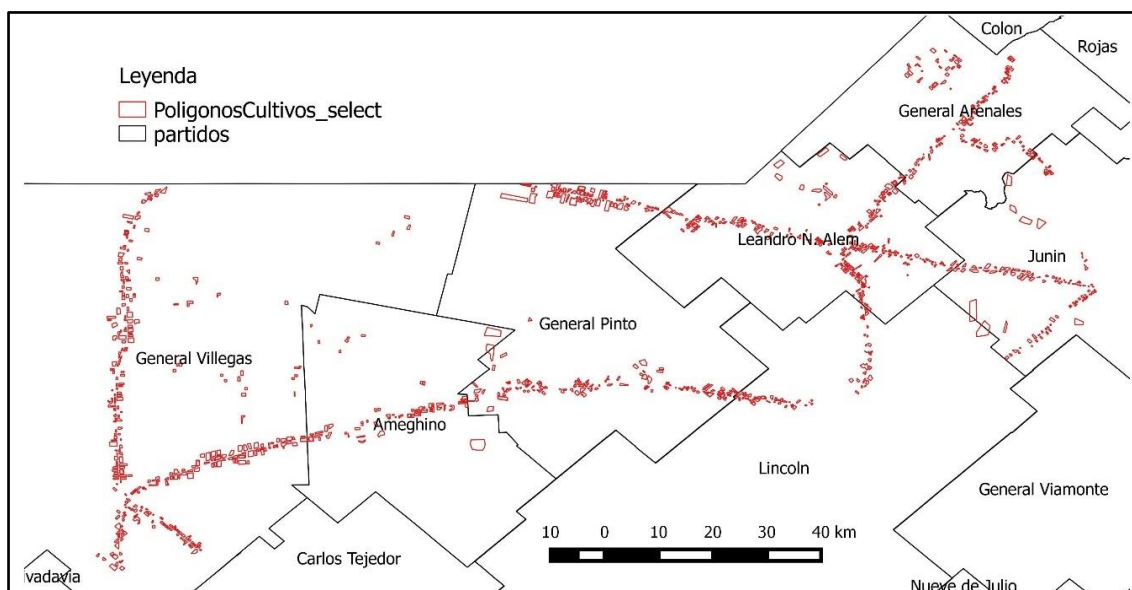


Figura 29. Polígonos que representan los lotes registrados a campo con las distintas coberturas de la "verdad de campo".

Durante la campaña de recolección de datos se registraron las coberturas observadas en cada parcela sobre un recorrido predeterminado en la zona de estudio. Posteriormente en gabinete se dibujó sobre las imágenes SAR los polígonos correspondientes a las parcelas de cultivo identificadas a campo o a las coberturas naturales (ver Figura 30). A diferencia de las clasificaciones del capítulo 4, en este sitio el cultivo de girasol es más frecuente y fue incluido en las coberturas clasificadas.

Las coberturas relevadas fueron agrupadas en ocho clases:

1. Cultivo de Soja de primera
2. Cereal de Invierno/Soja de segunda
3. Cultivo de Maíz
4. Cultivo de Girasol
5. Pastizales/Pasturas
6. Urbano
7. Agua
8. Juncal

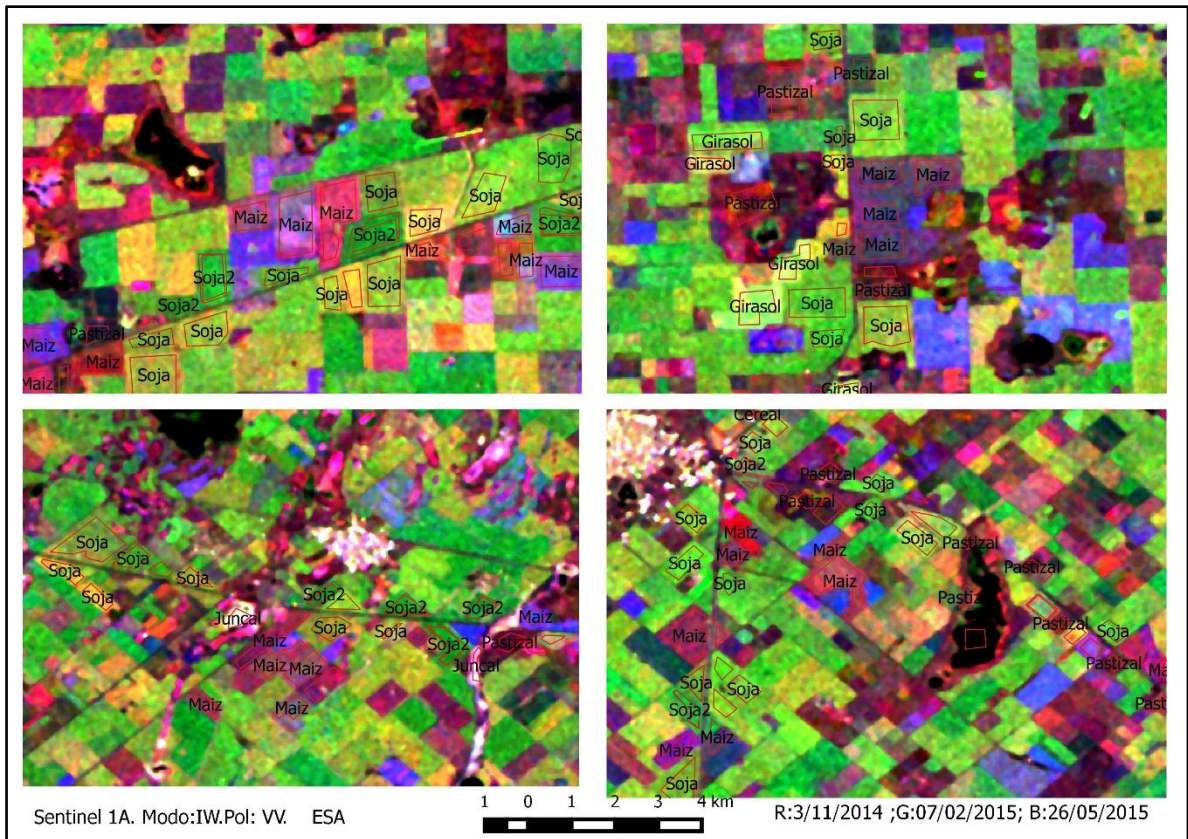


Figura 30. Polígonos de coberturas registradas a campo sobre imágenes multitemporales Sentinel1. Imágenes Compuestas RGB: 3 Noviembre 2014 VV, 7 Febrero 2015 VV, 26 Mayo 2015 VV.

5.2.2. Imágenes SAR disponibles

Las clasificaciones de imágenes SAR Sentinel 1 A se realizaron sobre un set de imágenes multitemporal del Oeste de la provincia de Buenos Aires (ver tabla 4). Las imágenes correspondientes al satélite mencionado no fueron programadas en función de los objetivos específicos de este capítulo, sino que solamente se tomaron del servidor de imágenes scihub.esa.int, perteneciente a la ESA. Las mismas fueron programadas en base a los objetivos de la Agencia Espacial Europea y en base a una programación fija de adquisición de imágenes de la superficie terrestre. Dicha programación puede encontrarse en el sitio web de la misión (<https://sentinel.esa.int/web/sentinel/missions/sentinel-1/observation-scenario>, (ESA, n.d.)). Las imágenes utilizadas fueron descargadas en nivel de procesamiento “Ground Range Detected”, (nivel 1). Este nivel de procesamiento incluye, datos focalizados, en amplitud, y proyectado a usando un modelo de elipsoide global.

IMAGEN	MODO	NIVEL DE PROCESAMIENTO	POLARIZACIÓN	FECHA	ÁNGULO DE INCIDENCIA	ORBITA	MIRADA
1	IW	GRD	VV	03/11/2014	29-40	ASCENDENTE	DERECHA
2	IW	GRD	VV	27/11/2014	29-40	ASCENDENTE	DERECHA
3	IW	GRD	VV	21/12/2014	29-40	ASCENDENTE	DERECHA
4	IW	GRD	VV	14/01/2015	29-40	ASCENDENTE	DERECHA
5	IW	GRD	VV	07/02/2015	29-40	ASCENDENTE	DERECHA
6	IW	GRD	VV	15/03/2015	29-40	ASCENDENTE	DERECHA
7	IW	GRD	VV	02/05/2015	29-40	ASCENDENTE	DERECHA
8	IW	GRD	VV	26/05/2015	29-40	ASCENDENTE	DERECHA

Tabla 5. Imágenes Sentinel-1 utilizadas para las clasificaciones supervisadas.

5.2.3. Pre-procesamiento

El pre-procesamiento de las imágenes Sentinel, se llevó a cabo con el programa de la ESA dedicado a la misión llamado Sentinel-toolbox. Los objetivos del pre-procesamiento fueron: hacer la imagen radar lo más parecido posible a una imagen óptica en cuanto a las características de homogeneidad radiométrica (ruido "Speckle"), georreferenciar las imágenes de manera de poder coregistrarlas, superponerlas en un SIG (sistema de información geográfica), y realizar clasificaciones. Para ello se debió calibrar a Intensidad Sigma Cero los valores digitales de las imágenes, georreferenciarlas con datos orbitales y un modelo digital de elevación (SRTM de 90m, por sus siglas en inglés: Shuttle Radar Topography Mission, cortesía del USGS). La reducción del ruido "Speckle" se realizó mediante el algoritmo de "Multilooking", aumentando la precisión radiométrica pero reduciendo la resolución espacial (30m x 30m tamaño de pixel). Los parámetros utilizados para el pre-procesamiento se describen en el Anexo 3.5. La fórmula de calibración utilizada se observa en la Fórmula 2.

Fórmula 2. Calibración a intensidad:

$$\sigma^0(i) = \frac{(abs(DNi))^2}{A_{(i)}^2}$$

DN: "Digital Number"

A: Vector de calibración a sigma cero. Se encuentra en la metadata de la imagen.

(Luis Veci 2015)

La figura 31 muestra el orden de los procesos aplicados a las imágenes Sentinel1A.

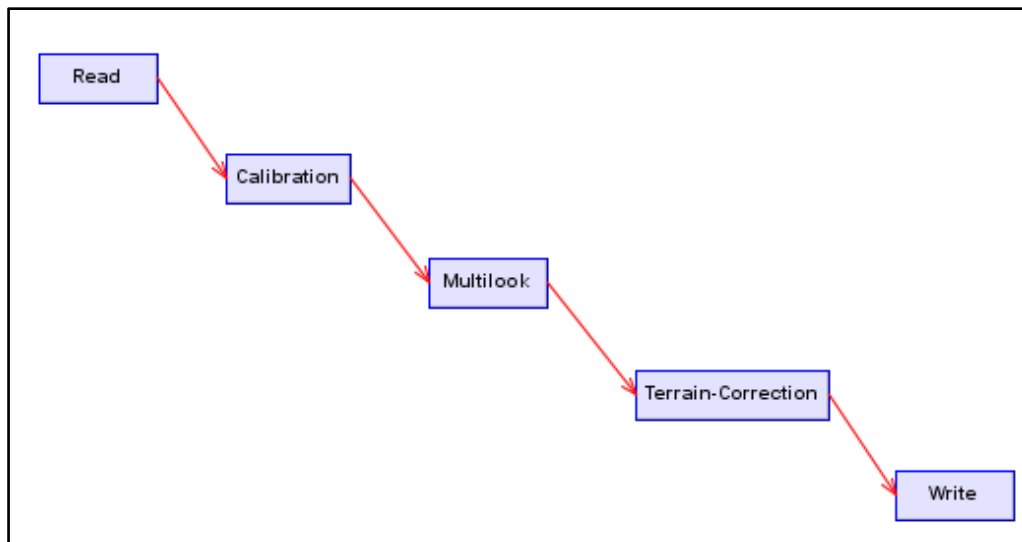


Figura 31. Grafo de pre procesamiento de las imágenes Sentinel1. Se observan los procesos utilizados para corregir geoméricamente y normalizar los valores digitales de las imágenes. Los parámetros se pueden ver en el Anexo 3.5.

Por otra parte, las bandas 2 a 7 del sensor OLI perteneciente al satélite Landsat 8, utilizadas para clasificar las coberturas mencionadas, fueron previamente calibradas a reflectividad y concatenadas en un solo archivo multibanda.

5.2.4. Procesamiento

El procesamiento consistió en realizar clasificaciones supervisadas de cinco sets de datos confeccionados a partir del pre-procesamiento. Las clasificaciones supervisadas fueron realizadas de forma no automática por un operador. Los sets clasificados fueron:

1. Ocho imágenes Sentinel-1A en polarización VV de distintas fechas.(8 bandas)
2. Tres imágenes Landsat 8, sensor OLI, de las bandas 2 a la 7. (18 bandas)
3. Los dos sets anteriores concatenados en una imagen multibanda. (26 bandas)
4. Seis imágenes Sentinel-1A en polarización VV correspondientes a fechas de adquisición de Diciembre a Mayo de 2015, con una imagen Landsat 8 del 15 de Noviembre de 2014.(12 bandas)
5. Tres imágenes Sentinel-1A de los meses de Diciembre 2014, Enero 2015 y Febrero 2015, con una imagen Landsat 8 del 15 de Noviembre 2014 y una imagen Landsat 8 del 7 de Marzo de 2015. (15 bandas)

Los sets se conformaron con los siguientes objetivos:

- a. Los sets uno y dos fueron elegidos para comparar las clasificaciones de imágenes SAR vs Imágenes ópticas L8.
- b. Los sets número tres a cinco permitirían indagar el efecto combinado de los datos SAR banda C y ópticos sobre la precisión de la clasificación efectuada.

Los cinco sets fueron clasificados mediante el algoritmo “Máxima verosimilitud” implementado en el programa ENVI 4.8®. Adicionalmente el set número 1 (8 imágenes

Sentinel1A de distintas fechas corregistradas y concatenadas) fue clasificado mediante clasificadores no paramétricos implementados en el programa Monteverdi2-Orfeo Tool Box® de la Agencia espacial francesa (CNES). La descripción de los algoritmos clasificadores encuentra en el capítulo 1.3.1. La implementación de los algoritmos esta descrita en “The Orfeo Toolbox Software Guide” (OTB development team 2014) y una descripción más detallada puede encontrarse en (Open Source Computer Vision Library, n.d.). Los clasificadores no paramétricos utilizados fueron:

1. “Decision tree”
2. “Random Forest”
3. “Gradient Boosted tree”
4. “Lib Support Vector Machine”
5. “Neural Network”

Una vez realizadas las clasificaciones se calculó la precisión global de cada clasificación, y los errores de omisión y comisión para cada clase. El set de datos de entrenamiento fue del 50% del total de datos relevados mientras que el otro 50% de los datos fue utilizado para el cálculo de la precisión y los errores.

5.3. Resultados

Los cinco sets clasificados superaron el umbral del 80% de precisión global, el cual indica un buen desempeño del combo set-clasificador-verdad de campo. La primera diferencia importante se observa entre la precisión del set de imágenes SAR Sentintel-1A y los sets que incluyen imágenes ópticas, presentando estos últimos mejores resultados. Luego el set de 8 imágenes SAR alcanzó una mayor precisión global al ser clasificado por un clasificador no paramétrico: 85%, *Random Forest* vs. 80% *Máxima Verosimilitud*, (ver Figura 32 y Figura 33).

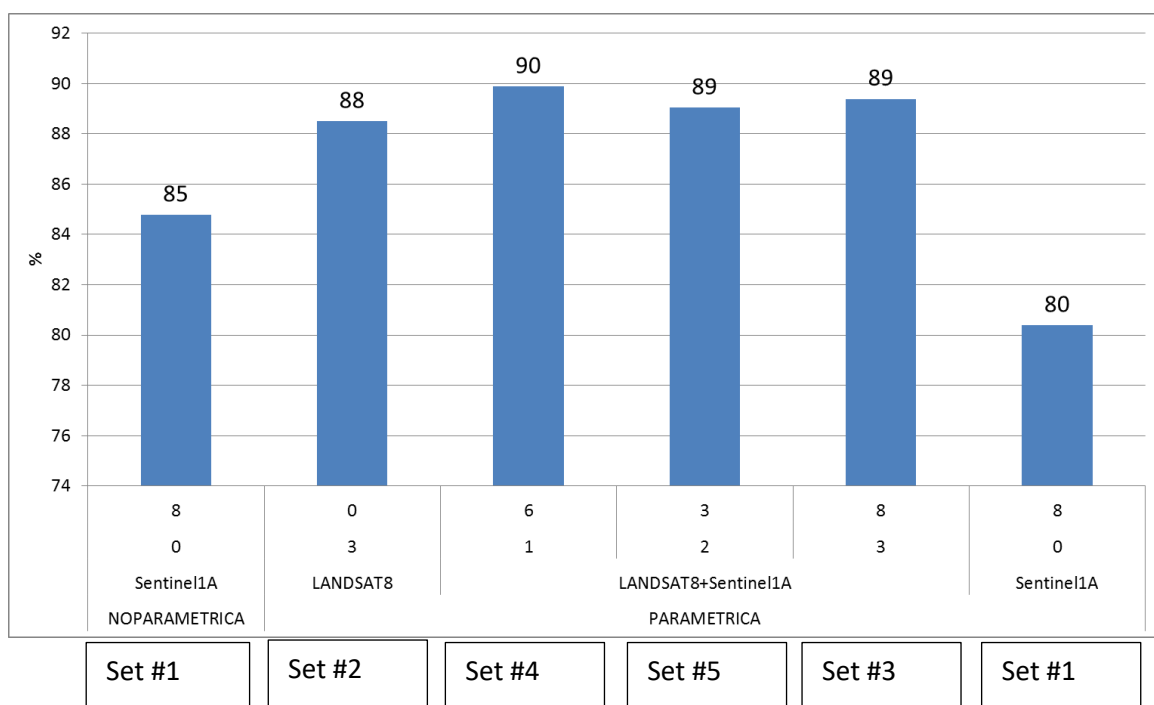


Figura 32. Clasificaciones zona Oeste de Buenos Aires. Precisiones globales alcanzadas por distintos clasificadores y sets de datos. La primera línea corresponde al número de imágenes SAR Sentinel 1, y la segunda al número de imágenes Landsat 8.

En cuanto a los clasificadores no paramétricos ensayados se observa un buen desempeño en casi todos, con excepción de las variantes de *Support Vector Machine* y *DT* (ver Figura 33).

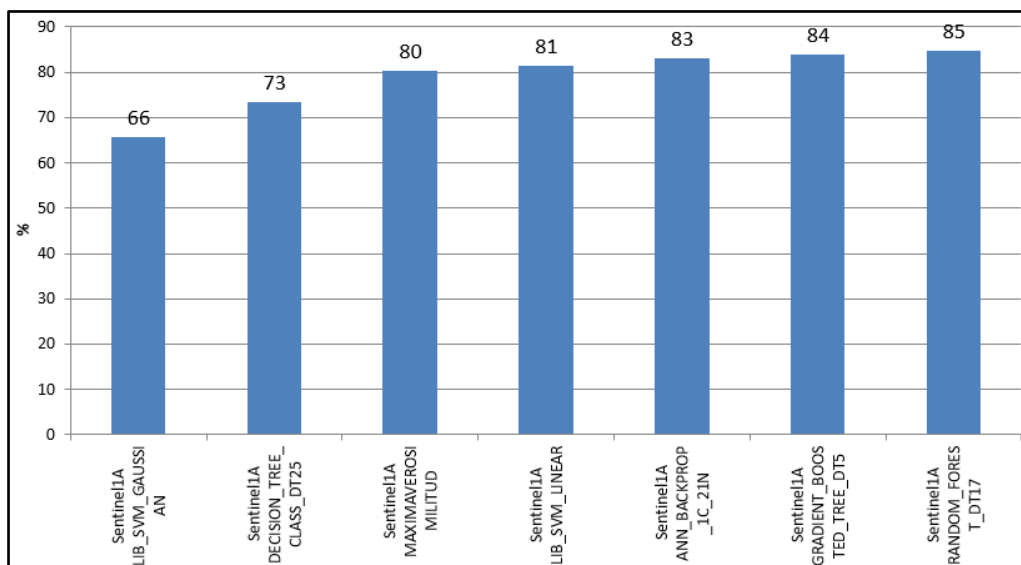
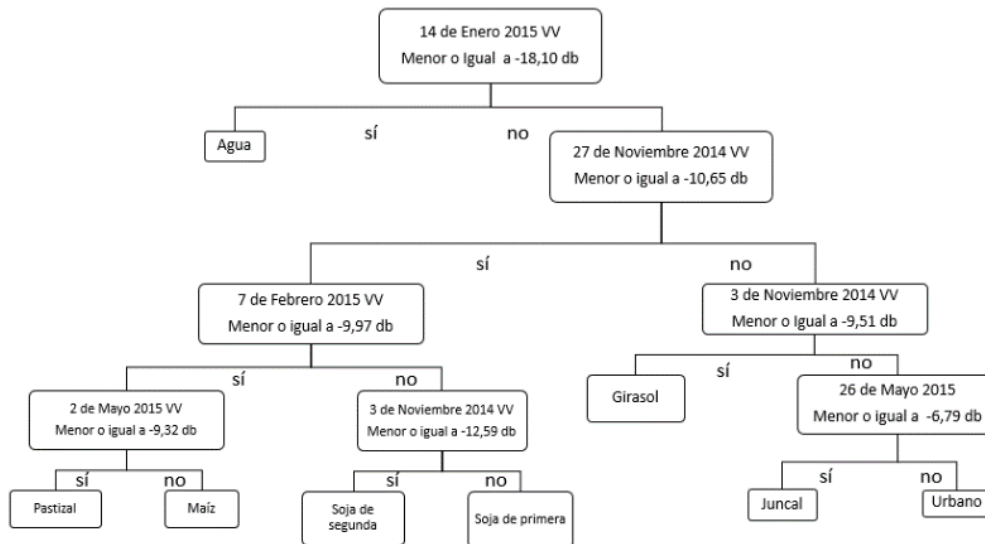


Figura 33. Clasificaciones zona Oeste de Prov. Buenos Aires. Precisión global alcanzada por distintos clasificadores para el set de datos de 8 imágenes sentinel1. De izquierda a derecha: "Support Vector Machine", basado en la librería LibSVM, kernel: gaussiano; "Decision tree", profundidad de árbol: 25; "Máxima Verosimilitud" implementado en el programa ENVI 4.8; "Support Vector Machine", basado en la librería LibSVM, kernel: lineal; "Artificial Neural Network", backpropagation, 1 capa oculta, 21 neuronas; "Gradient Boosted Tree", profundidad de árbol: 5; "Random Forest", profundidad de árbol: 17. La cantidad de píxeles muestra con los que se entrenó a los clasificadores fue de 11035 píxeles.

A continuación se presenta el árbol de decisión generado por el clasificador DT. El desempeño del mismo fue del 73% de precisión global, mientras que el mejor clasificador alcanzó una precisión global del 85% ("Random Forest", combinación de 200 árboles). El árbol muestra una primera división que separa la clase "Agua", del resto de las coberturas. En un nivel subsiguiente separa la clase "Girasol", y en el último nivel del árbol separa todas las otras clases.



Árbol de Decisión (DT) 3. Árbol generado por el algoritmo DT. Cada nodo muestra la imagen y el límite utilizado para dividir el nodo. El valor de intensidad está expresado en decibeles. No todas las imágenes fueron utilizadas.

En la Figura 34 se muestran los errores por clase para los sets #1(imágenes Sentinel1A), #2(imágenes L8) y #3(imágenes Sentinel 1A e imágenes L8). El agregado de imágenes SAR al set de imágenes L8 ocasiona la disminución de los errores de la clase “Girasol” de forma bien marcada. Esto se debería a que el Girasol presenta valores de retrodispersión muy distintos al resto de los cultivos, con valores que van entre -6 y -12 db, con un marcado incremento de Diciembre a Febrero. La retrodispersión más baja se observa en la clase “Agua” que corresponde a cuerpos de agua superficial, la misma se encuentra por debajo de los -16 db para todas las fechas analizadas. De las clases correspondientes a vegetación natural, pastizales y juncuales, vemos diferencias notables. La clase “pastizal”, presenta valores entre -9 y -15 db con una tendencia a mayor retrodispersión en Primavera y Otoño. Esto coincidiría con el ciclo de crecimiento de especies de gramíneas de clima templado. La clase “juncal” presenta valores altos de retrodispersión (para una cobertura vegetal) en todas las fechas analizadas (de Noviembre a Mayo), de entre -5 y -9 db. Lamentablemente el cultivo de Maíz presentó valores de retrodispersión similares a los presentados por la clase “Pastizal”. En cuanto a la clase “Soja de primera”, presentó valores de retrodispersión de -7 a -13 db con valores. La clase “Soja de segunda”, presentó una dinámica temporal similar a la clase “Soja de primera” pero con valores inferiores de retrodispersión entre Noviembre y Diciembre, coincidentes con las etapas finales del cultivo de invierno antecesor (Ver Anexo 2).

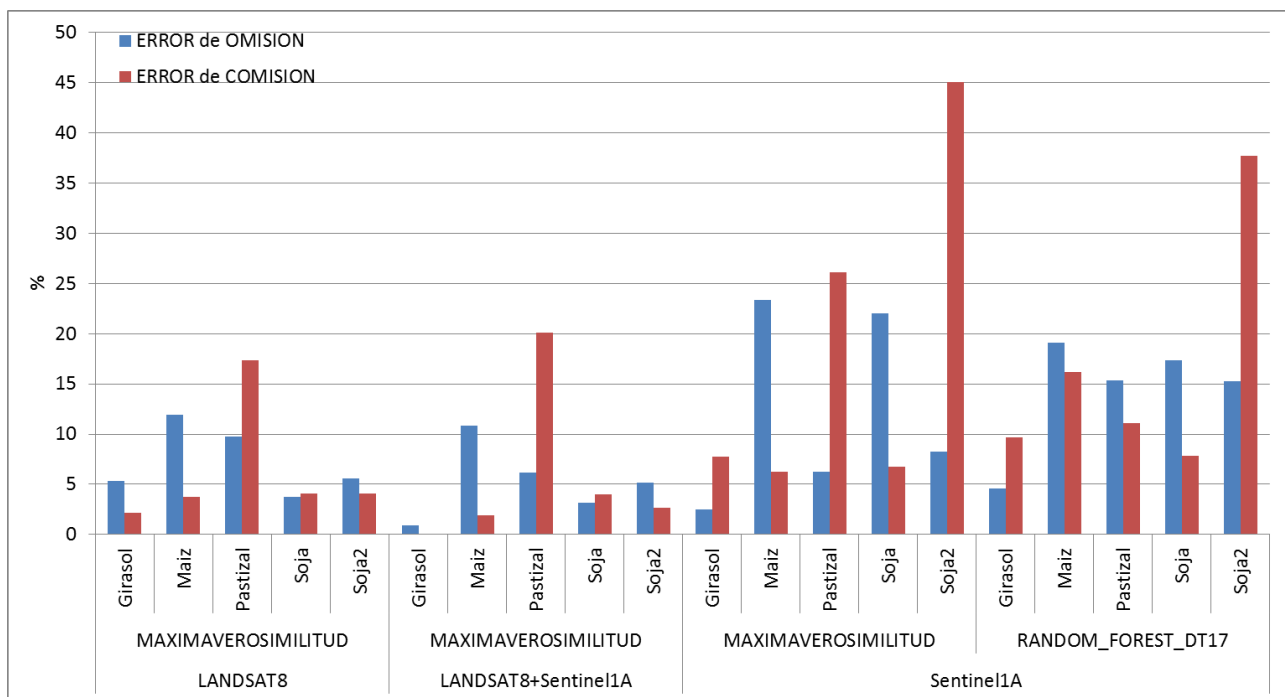


Figura 34. Clasificaciones zona Oeste de Prov. Buenos Aires. Errores de Omisión y Comisión para las principales coberturas clasificadas.

5.4. Discusión

En este capítulo se logró clasificar cultivos en una amplia zona del oeste de la Provincia de Buenos Aires, utilizando un set multi-temporal de imágenes SAR del nuevo satélite de la ESA Sentinel1A. Es importante destacar la mejoría en la precisión de la clasificación al utilizar un clasificador no paramétrico (ver Figura 32). La afirmación anterior coincide con la bibliografía revisada sobre el tema (Loosvelt, Peters, Skriver, Lievens, et al. 2012; Jin, Mountrakis, and Stehman 2014). Es importante también observar que la combinación de datos ópticos (OLI-L8), y SAR (Sentinel1A), no solo mejora la precisión y disminuyen los errores de la clasificación sino que además mejora la calidad visual de la clasificación como se observa en la Figura 35. Al respecto, se observa que el solo hecho de adicionar una imagen multibanda óptica L8 la precisión global se incrementa aproximadamente un 5%. En cuanto a la calidad visual de las clasificaciones es destacable el hecho de la “contaminación” de pixeles clasificados como “urbano” dentro de la clase “maíz”, en los sets de imágenes SAR (paneles superiores de la Figura 35). El agregado de imágenes ópticas reduce ese error. Por otro lado, el agregado de imágenes SAR a la clasificación de imágenes ópticas disminuye el error de la clase “Girasol”.

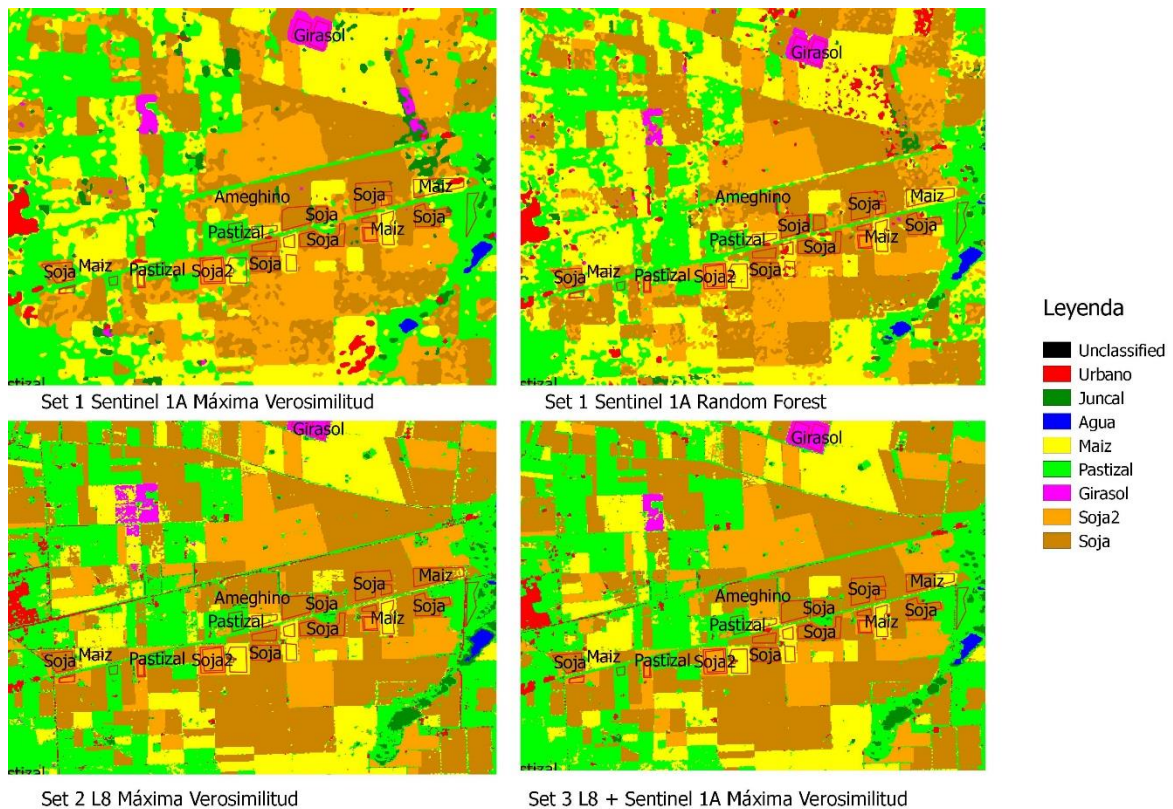


Figura 35. Clasificaciones de cultivos en el Oeste de la Provincia de Buenos Aires. Campaña 2014-2015

5.5. Conclusiones

1. Las imágenes Sentinel1A pueden ser utilizadas solas o combinadas con imágenes ópticas para generar clasificaciones de cultivos de calidad buena a muy buena (ver Figura 32 y Figura 33).
2. La clasificación de imágenes Sentinel 1A e imágenes Landsat 8 sin nubes (set #3) es superior a la clasificación de cada set de imágenes por separado.
3. La mayoría de los clasificadores (ej. "Random Forest") no paramétricos demostraron ser mejores para clasificar sets temporales de imágenes Sentinel 1A, en comparación con el clasificador paramétrico de Máxima Verosimilitud.

En la figura 36 se muestra el producto cartográfico extraído de la clasificación de cultivos por medio del algoritmo "Random Forest", sobre el set de ocho imágenes Sentinel 1A.

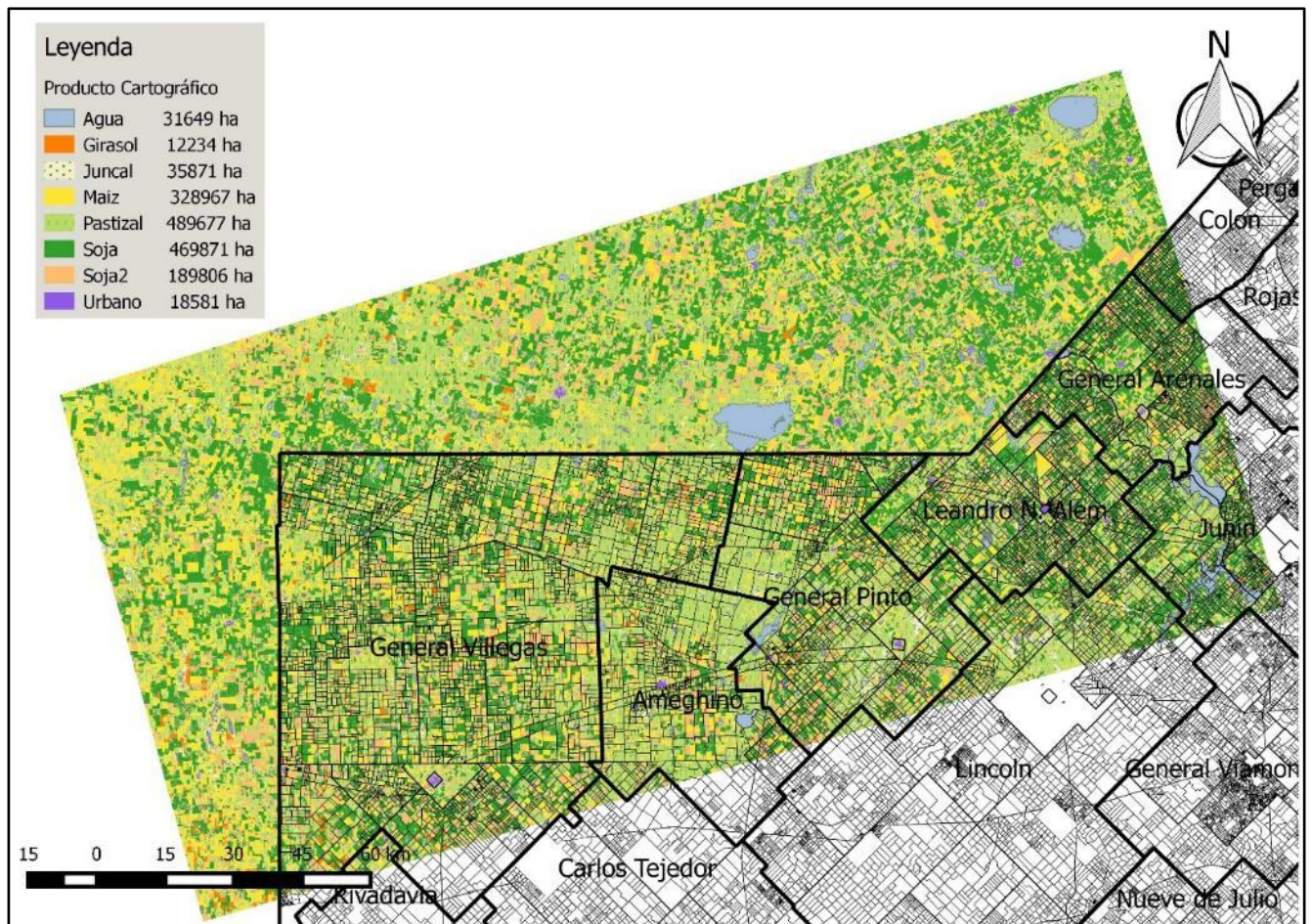


Figura 36. Producto Cartográfico. Clasificación cultivos y superficie calculada por cobertura, sobre el área de la provincia de Buenos Aires. La clasificación corresponde a la realizada con el algoritmo "Random Forest", sobre el set de ocho imágenes Sentinel 1A.

Capítulo 6. Discusión y Conclusiones

La retrodispersión observada para los distintos cultivos en las distintas zonas donde se trabajó presentan patrones que se repiten para las distintas polarizaciones. En el caso del cultivo de Girasol, la retrodispersión fue alta para los meses de Diciembre a Febrero en particular en las polarizaciones VV y HH (ver Anexo 2). Es posible que esto se deba al efecto sobre la retrodispersión por parte de la inflorescencia típica de la planta de Girasol. En el cultivo de papa la retrodispersión en polarización HV es más alta que en el resto de los cultivos en los meses de Noviembre a Enero. Esto debería asociarse al efecto del suelo arado sobre el cual se implante el cultivo, y al riego suplementario utilizado. En el caso del maíz, la retrodispersión en polarización HV acompaña el crecimiento del cultivo, yendo de menos a más a medida que avanza la campaña. En el caso de la soja y la soja de segunda, la retrodispersión es superior en la primera, en especial en los meses de Noviembre a Enero. El efecto del cultivo de cereal sobre la retrodispersión es la atenuación de la misma (Ferrazzoli P. 2002), es por esto que en los meses previos a la implantación de la soja de segunda (de Octubre a Diciembre) la retrodispersión es menor que en los lotes con cultivo soja de primera. En términos generales, analizando los árboles de decisión generados por el clasificador "DT", se puede afirmar que la clase más sencilla de separar es la clase "Agua", la cual corresponde a cuerpos de agua en superficie.

Si bien se han probado clasificaciones sobre sets de imágenes SAR en banda X (Cosmo SkyMed) como sobre sets de imágenes SAR en banda C (Sentinel 1), no sería recomendable su comparación dado que el número de imágenes de ambos sets, las fechas de adquisición, los ángulos de incidencia y sus polarizaciones no son las mismas. Para poder realizar una comparación solo de frecuencias utilizadas por uno y otro sistema SAR, se debería mantener la misma configuración de uno y otro sistema en términos de polarización, ángulo de incidencia, y fecha de adquisición, dado que todos esos factores alteran los valores de la retrodispersión recibida por la antena SAR.

El objetivo general planteado en el capítulo uno, el cual consistía en proporcionar herramientas metodológicas al departamento de Tecnología de Imágenes de la Agencia de Recaudación de la Provincia de Buenos Aires, con el fin de incorporar imágenes SAR y métodos estadísticos de clasificación de cultivos al sistema de detección de cultivos, fue alcanzado con éxito dado que se lograron clasificar exitosamente diferentes cultivos utilizando imágenes SAR en distintas áreas de la provincia y bajo distintos regímenes culturales. Este objetivo se consolidó mediante la propuesta de funcionamiento del área dedicada a la clasificación de cultivos del departamento (Anexo 1). El aporte de la propuesta desarrollada en el primer Anexo consiste en definir roles y tareas específicas de los agentes intervinientes, de manera de garantizar la obtención de cartografía y estimaciones de superficie cultivada, lo más precisas posibles, para los niveles de tecnología disponibles para la Agencia en la actualidad (2015).

Los cuatro capítulos centrales de la tesis contribuyen con el proceso de clasificación de distintos cultivos en distintas áreas agro-económicas de la provincia de Buenos Aires. El capítulo dos, presenta una nueva forma de relevar lotes con cultivo de cebolla de manera remota, es importante destacar el ahorro monetario y logístico que implica la utilización de imágenes SAR de alta resolución para este fin, con respecto a la alternativa de inspeccionar las parcelas de forma presencial a campo. El capítulo tres, propone la utilización de imágenes SAR únicamente en el caso de disponer con escaso número de imágenes ópticas y contribuyó a la propuesta de mejora explicada en el Anexo 1. El capítulo cuatro, aporta datos para expandir la clasificación de cultivos sobre la base de la utilización de imágenes SAR Wide Region sobre amplias zonas de la provincia en combinación con datos ópticos. El capítulo cinco, introduce la utilización de un nuevo sensor, el primero de la serie Sentinel, ampliando de manera muy significativa el área clasificada. En este sentido es importante destacar que la alta frecuencia y calidad del sistema Sentinel 1A, el cual se verá potenciado con la puesta en órbita del satélite Sentinel 1B, permitiría

explorar la utilización del dato de “fase” en la clasificación de cultivos. Por ejemplo, calculando la diferencia de fase entre las bandas HH y HV a partir de imágenes SLC (Single Look Complex) (Lee, Grunes, and Pottier 2001; Loosvelt, Peters, Skriver, De Baets, et al. 2012) o la utilización de índices polarimétricos. Adicionalmente, la posibilidad que brinda el sistema SAR Sentinel 1 permitiría el cálculo de la correlación interferométrica entre un par de imágenes SAR, la cual se encuentra en relación con el volumen de vegetación cultivada (Wegmuller and Werner 1997).

Los objetivos particulares de la tesis consistieron en lograr una estimación de superficie cultivada y detección de parcelas cultivadas por cultivo y por área agroeconómica, también fueron alcanzados, generando productos cartográficos de buena calidad. La utilización de imágenes SAR para la producción de cartografía de coberturas de terreno presenta ventajas y desventajas. Entre las desventajas destacables de esta tecnología vemos en primer lugar la necesidad de programación de la adquisición (ej: Cosmo SkyMed), lo cual supone un proceso en el cual no se tiene certeza absoluta de si la imagen será tomada dado que el pedido compite con usuarios de otras aplicaciones más importantes como ser el militar o en emergencias. En segundo lugar, las imágenes SAR, requieren un pre-procesamiento adicional a las imágenes ópticas, tanto para reducir el efecto Speckle como para georreferenciarlas. En tercer lugar, la precisión radiométrica de las imágenes debe ser evaluada para cada una en particular dado que la misma se ve afectada por las condiciones atmosféricas, por ejemplo, en el caso de fuertes tormentas en el momento de la adquisición, se genera un efecto de atenuación de la retrodispersión. Dentro de las bondades de las imágenes SAR, encontramos su capacidad para “observar” la superficie terrestre por debajo de la capa nubosa, o en ausencia de iluminación solar. En la actualidad, han proliferado las misiones espaciales SAR poniendo cada vez más imágenes a disposición de los usuarios, por ejemplo, la misión Sentinel 1A, el cual será acompañado por otro satélite gemelo, el Sentinel 1B, o el satélite de observación terrestre argentino SAR banda L, SAOCOM, previsto para el año 2018.

Las imágenes Cosmo Himage presentan como ventaja la posibilidad de adquirir imágenes a resolución de hasta dos metros, permitiendo la detección de objetos sobre la superficie terrestre como se utilizó en el capítulo dos. Si bien el modo Himage permite tanto la adquisición en polarización cruzada como co-polar, es preferible esta última por presentar mayor nitidez en los bordes de la imagen. Las imágenes Himage son recomendables para la detección de objetos territoriales de tamaño medio de 2 a 15 metros y para clasificación de coberturas de terreno en superficies de 20 a 40 km².

Las imágenes Cosmo Wide Región muestran como ventaja principal su amplio ancho de barrido (150 km). En este modo de adquisición es preferible ordenar el nivel DGM para poder georreferenciar de manera más precisa la imagen, si no se requiere del dato de fase de la imagen.

Las nuevas imágenes del satélite Sentinel 1A, poseen la gran ventaja de estar programadas a priori y de no necesitar hacer un pedido explícito para su adquisición. Si bien en la actualidad no cubren la totalidad de la superficie terrestre con una revisita periódica, es probable que con la puesta en órbita de su satélite gemelo el Sentinel 1B se logre la cobertura completa del globo con imágenes SAR banda C de manera repetitiva y constante. Las imágenes ensayadas han dado buenos resultados y presentan todas las características para ser incluidas en procesos de producción de información territorial dada su disponibilidad y revisita (para las áreas programadas por la agencia productora).

En los tres casos donde se probaron los clasificadores no paramétricos vs. el clasificador de máxima verosimilitud (paramétrico) sobre sets de imágenes SAR, las diferencias en la precisión global de la clasificación fueron siempre favorables a alguno de los clasificadores no paramétricos testeados, destacándose el llamado “Random Forest”. Esta diferencia a favor del no paramétrico se explicaría por la característica de los datos de intensidad SAR, ya que

presentan una distribución gama. Si bien lo anterior es cierto, los filtros aplicados para reducir el efecto “Speckle” han normalizado parcialmente la distribución de los valores de retrodispersión de las coberturas, dando resultados análogos con el clasificador paramétrico de Máxima Verosimilitud.

A continuación se muestra una comparación de los resultados obtenidos con los publicados por el Ministerio de Agroindustria de la Nación (en http://www.siiia.gob.ar/sst_pcias/estima/estima.php).

El método utilizado por el citado ministerio se puede encontrar en:

http://www.agroindustria.gob.ar/site/agricultura/estimaciones_agricolas/03=estimaciones%20agricolas/01-metodolog%C3%ADa/Metodo_de_segmentos_aleatorios_Version5.pdf?version5.

Partido		Junín		
Superficie cultivada (ha)		<i>Maíz</i>	<i>Soja</i>	
Ministerio Agroindustria Nación (1)		18000	134500	
Clasificación Cosmo+L8 “Máxima Verosimilitud” (2)		31756	114523	
Diferencia (2/1) (%)		76 %	-15 %	
Partido		General Pinto		
Superficie cultivada (ha)		<i>Maíz</i>	<i>Soja</i>	<i>Girasol</i>
Ministerio Agroindustria Nación (1)		30000	87100	900
Clasificación Sentinel 1 "Random Forest" (2)		42862	89911	1086
Diferencia (2/1) (%)		43 %	3 %	21 %
Partido		Leandro N. Alem		
Superficie cultivada (ha)		<i>Maíz</i>	<i>Soja</i>	<i>Girasol</i>
Ministerio de Agroindustria Nación (1)		22000	83050	700
Clasificación Sentinel 1 "Random Forest" (2)		36476	75079	866
Diferencia (2/1) (%)		66 %	-10 %	24 %

Tabla 6. Comparación de las superficies cultivadas estimadas mediante clasificaciones supervisadas (Capítulos 4 y 5), y estimaciones para la misma campaña agrícola realizadas por el Ministerio de Agroindustria de la Nación.

En la tabla 5 se observan las estimaciones de superficie cultivadas en la presente tesis y por el Ministerio de Agroindustria de la Nación. Se pudieron comparar únicamente los partidos clasificados en toda su superficie. El primero corresponde al partido de Junín, clasificado en el capítulo cuatro, y los partidos de General Pinto y Leandro N. Alem, clasificados en el capítulo cinco. La diferencia mayor se observa en el cultivo de maíz para los tres partidos. Para los cultivos de Girasol y Soja las diferencias absolutas se encuentran entre el 24 % y el 3 %. Las diferencias mayores encontradas en las estimaciones del cultivo de maíz podría deberse a que en la clase “maíz” no se ha discriminado el maíz con destino a consumo animal de la planta en pie o los cultivos de maíz con destino a silo de maíz para alimentación de ganado. El hecho que ambos métodos, encuentren valores similares en por lo menos dos de tres cultivos, es un hecho destacable. Es importante mencionar que no se puede asegurar cuál de los dos métodos es superior en términos de precisión, ya que no se encuentran publicados los datos de precisión de las estimaciones del Ministerio de Agroindustria. Sin embargo, sería recomendable, mejorar la

cantidad de puntos de observación de cultivos o “verdades de campo” para las futuras clasificaciones supervisadas. Además podría discriminarse para la clase “maíz”, el uso que se le va a dar al cultivo, para ello es importante realizar los recorridos más de una vez durante la campaña agrícola.

En conclusión, la primera hipótesis no fue rechazada: se han clasificado cultivos sobre la base de imágenes SAR de distintos sensores (Cosmo SKyMed, Sentinel1A) y de distintos modos de adquisición (Wide Region, Himage). La segunda hipótesis se rechazó: las clasificaciones de cultivos sobre imágenes SAR no fueron mejores en por lo menos dos de tres casos. Con respecto a la tercera hipótesis, en los tres casos se ha demostrado superior por lo menos uno de los clasificadores no paramétrico testeados frente al popular clasificador de máxima verosimilitud. No todos los clasificadores no paramétricos se han mostrado mejores que el clasificador paramétrico de máxima verosimilitud. Tuvieron especial buen desempeño aquellos clasificadores no paramétricos que combinaron árboles de decisión autoentrenados: “Random Forest” y “Gradient Boosted Trees”. El clasificador “DT”, si bien tuvo un menor desempeño en términos de precisión global, genera modelos (árboles) de fácil interpretación.

Sería deseable en el futuro que las investigaciones sobre clasificaciones de coberturas terrestres expliciten la extensión del área de la clasificación y si es posible cuantificar la heterogeneidad del paisaje o de las coberturas objetivo. Específicamente, en clasificaciones de imágenes SAR es importante aclarar el tipo de filtrado que sufren los datos, ya que el mismo modificará la precisión global del mapa. En el caso de las clasificaciones supervisadas es igualmente importante explicitar el tamaño de las muestras de entrenamiento y testeo, y analizar los errores de la clasificación a la luz de esos parámetros.

En futuras investigaciones sobre clasificaciones de coberturas de cultivo, sería interesante explorar la capacidad de clasificar más tipos de cultivos o variedades y con mejor resolución de manera de poder estimar mejor la producción de cultivos intensivos. En este sentido, y tomando en cuenta los resultados del capítulo dos, podría abrirse una línea de investigación que centre sus esfuerzos en detectar patrones espaciales vinculados a actividades económicas de manera automática. La detección de cultivos de forma remota no solo es importante para investigaciones del tipo fiscal impositivo y gestión de emergencias agropecuarias sino también para otras áreas del sector público como ser la fiscalización del trabajo informal, catastro económico, y estudios agroeconómicos a gran escala, que tengan como objetivo relacionar el tamaño de una explotación y las actividades agropecuarias que se desarrollan en la misma.

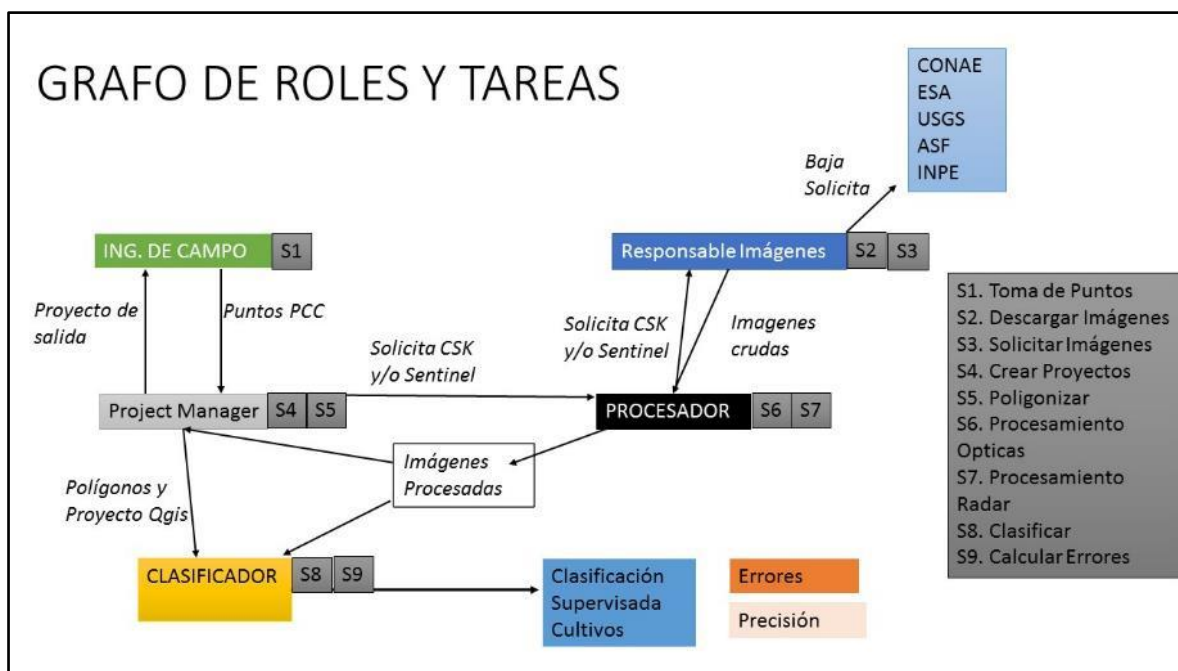
Otras líneas de investigación que podrían mejorar el sistema de detección remota de cultivos es clasificar previamente las zonas no agrícolas y realizar las clasificaciones exclusivamente sobre las áreas agrícolas. Para esto podría explorarse la utilización de modelos de elevación digital e imágenes ópticas para la determinación de áreas bajas y planicies de inundación que son utilizadas exclusivamente con fines ganaderos. En esta línea de trabajo también podría avanzarse con clasificaciones basadas en objetos, es decir, previa segmentación de una imagen, y cálculo de variables para cada vector resultante.

Finalmente, es importante destacar que la disponibilidad de imágenes satelitales tanto ópticas como de origen radar ha ido en aumento en los últimos años. A la histórica serie de satélites Landsat (gestionado por la NASA y el USGS) se ha sumado la serie de satélites de observación terrestre Sentinel del programa Copernicus de la ESA (Agencia Espacial Europea). Además, otras agencias han contribuido a aumentar la oferta de imágenes, como ser las imágenes del sensor Aster de la NASA/JAXA (National Aeronautics and Space Administration/Japan Aerospace Exploration Agency) y las imágenes CBERS 4 del INPE (Instituto Nacional de Pesquisas Espaciais). Esto implica que la cantidad de información de origen espacial disponible para el sector público es mayor. Por lo tanto, la cantidad de productos extraídos de

información de origen satelital podría ser eventualmente mayor o de mejor calidad. Tanto el sector académico como el sector público estarían en condiciones de realizar mayores esfuerzos para que esos productos impacten de manera efectiva en una mejor gestión del territorio y en una mejor administración pública, brindando más y mejores servicios a los habitantes.

Anexos

1. Propuesta del nuevo sistema de detección de cultivos de ARBA. Gráfico de roles y tareas.



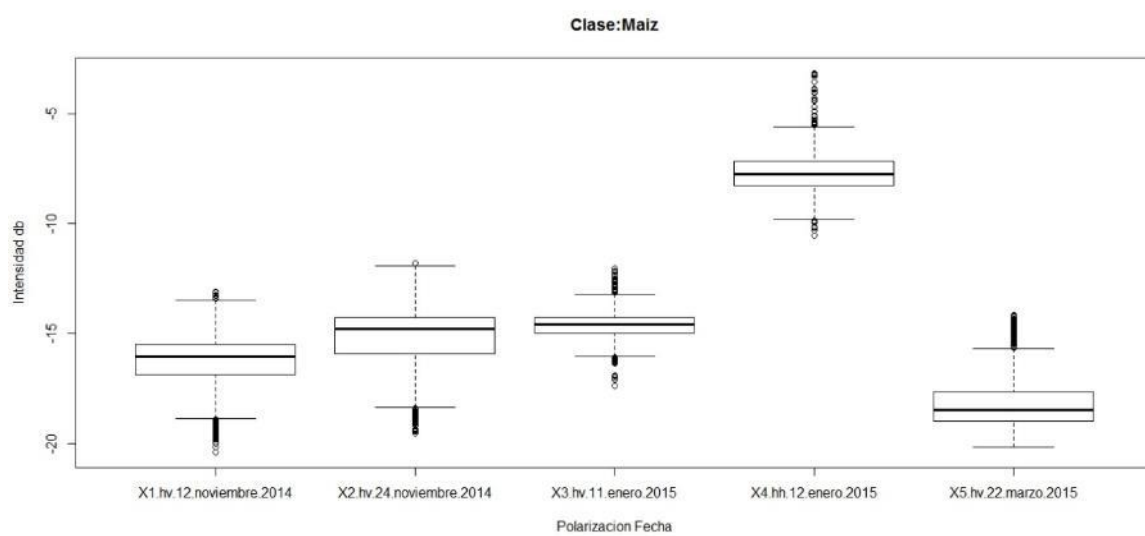
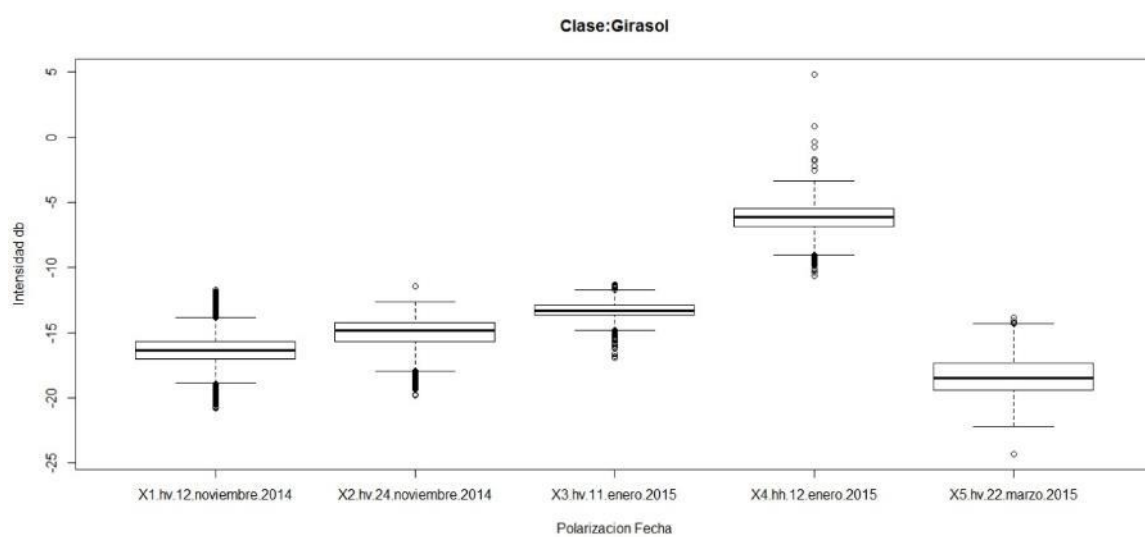
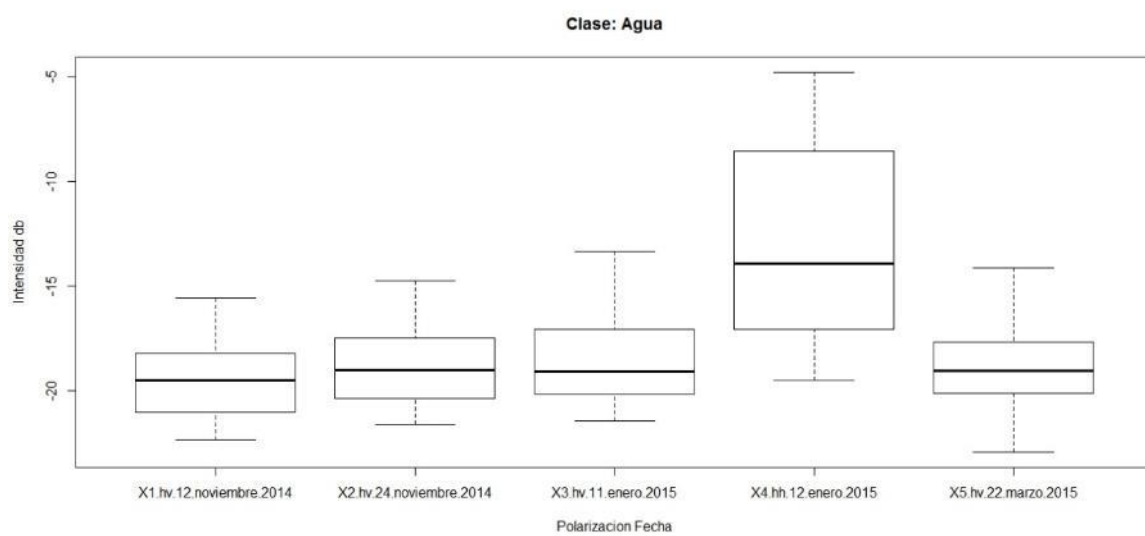
La figura precedente presenta una propuesta organizativa para el departamento de Tecnología de Imágenes de ARBA, pero podría utilizarse para cualquier dependencia pública o privada que deseara utilizar información de imágenes satelitales ópticas o SAR.

El sistema de detección, es un sistema que tiene como "inputs" los datos obtenidos a campo e imágenes satelitales de distintos orígenes, y como "outputs", una clasificación de terreno con los cultivos identificados, los errores de la clasificación y la precisión global de la misma. En el grafo anterior se identifican con la letra 's' los subprocesos. El S1 es la toma de datos en forma de puntos georreferenciados y etiquetados de los cultivos observados a campo. Esta tarea es llevada a cabo por un agrónomo. El subproceso S1 requiere de un proyecto SIG para volcar los datos de campo, el mismo es confeccionado por el "Project Manager" y suministrado en una Tablet PC. El Project Manager será el encargado también de programar los recorridos que utilizará el responsable de campo para la toma de puntos georreferenciados sobre el SIG. El programa SIG puede suministrarse tanto en sistema operativo licenciado Windows o en Android. La confección del proyecto incluye: el parcelario rural, las rutas y caminos rurales, y las últimas imágenes satelitales disponibles para poder identificar rápidamente posibles cultivos a lo largo del camino prefijado. La segunda tarea del Project Manager será llevar un registro de la calidad de las imágenes recibidas y de solicitar de ser necesario imágenes SAR al "procesador" de imágenes. El responsable de imágenes se ocupará de los subprocesos S2 y S3, que son la descarga y solicitud de imágenes. Es importante que estos subprocesos se concentren en una persona dado el gran número de imágenes e información disponibles en la actualidad (2015). Las imágenes descargadas serán puestas a disposición del "procesador", quien será el encargado de procesar las imágenes tanto ópticas como radar. Las imágenes procesadas serán puestas a disposición del Project Manager y del Clasificador. El Project Manager a su vez deberá tomar los puntos de campo georreferenciados, chequear su validez (que se encuentren dentro del rango visual de quien tomó el punto y que el mismo esté

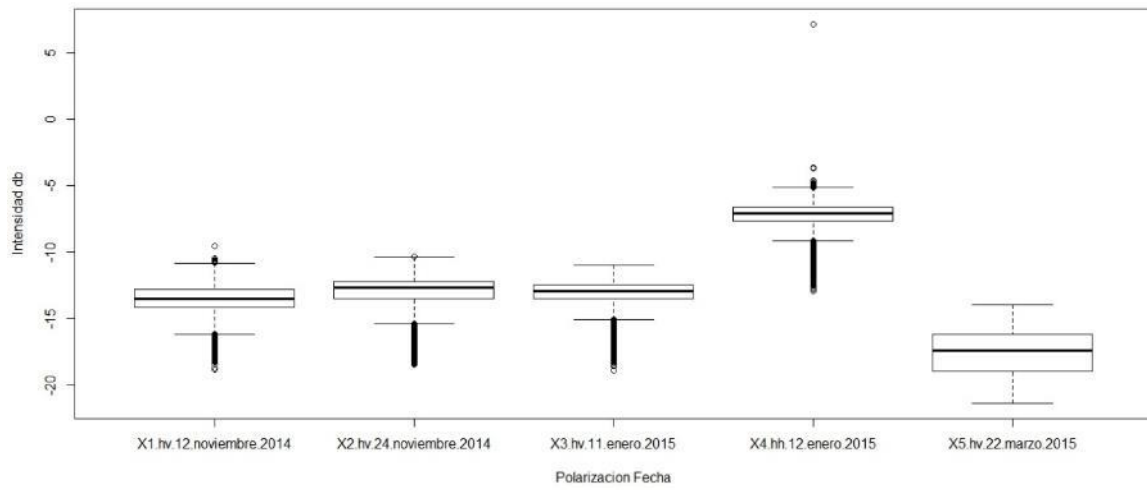
dentro de un lote agrícola), descartar los puntos dudosos, y dibujar los polígonos correspondientes a los lotes observados. Finalmente deberá dividir el conjunto de polígonos en dos grupos, uno de clasificación y otro de control. El Clasificador tomará los polígonos de clasificación y las imágenes procesadas y producirá la clasificación supervisada. Posteriormente calculará el error de cada clase, el error global y la precisión de la clasificación. La clasificación de cultivos tiene como fin la fiscalización de las declaraciones juradas de ingresos brutos: del cruce de lo declarado con lo detectado se puede fiscalizar el cobro debido del impuesto y recuperar lo que se declaró por debajo de lo producido realmente. Sin embargo, en una situación de emergencia climática, las clasificaciones también ayudarán a tomar decisiones relacionadas a exenciones fiscales y subsidios a los damnificados.

2. Gráficos de retrodispersión por cobertura.

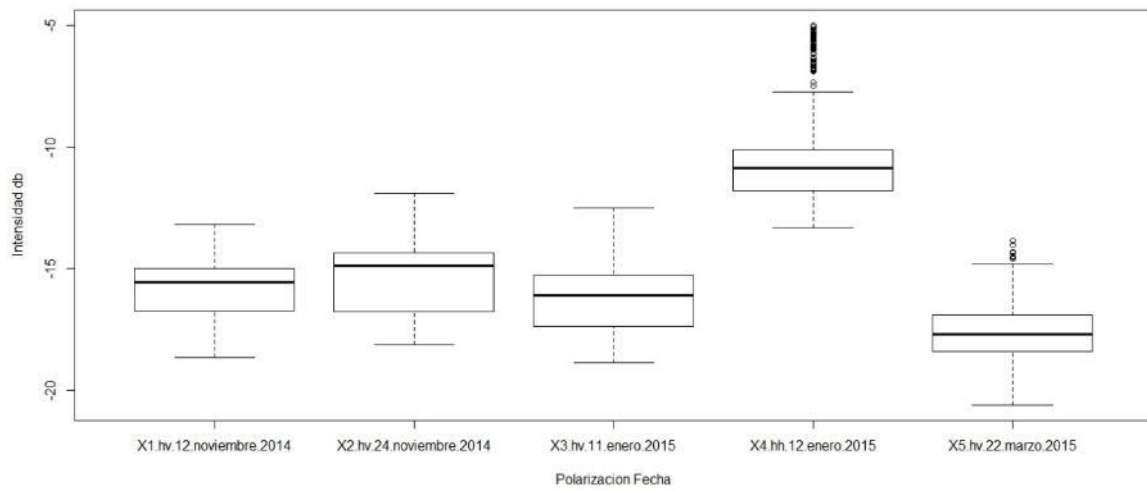
Boxplots clases Capítulo 3. Sudeste Provincia de Buenos Aires.



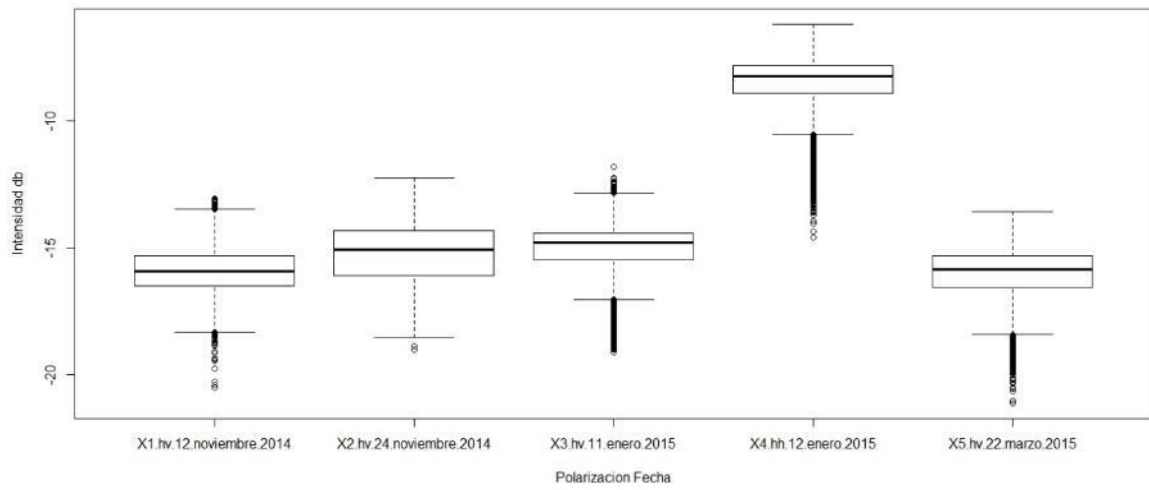
Clase:Papa



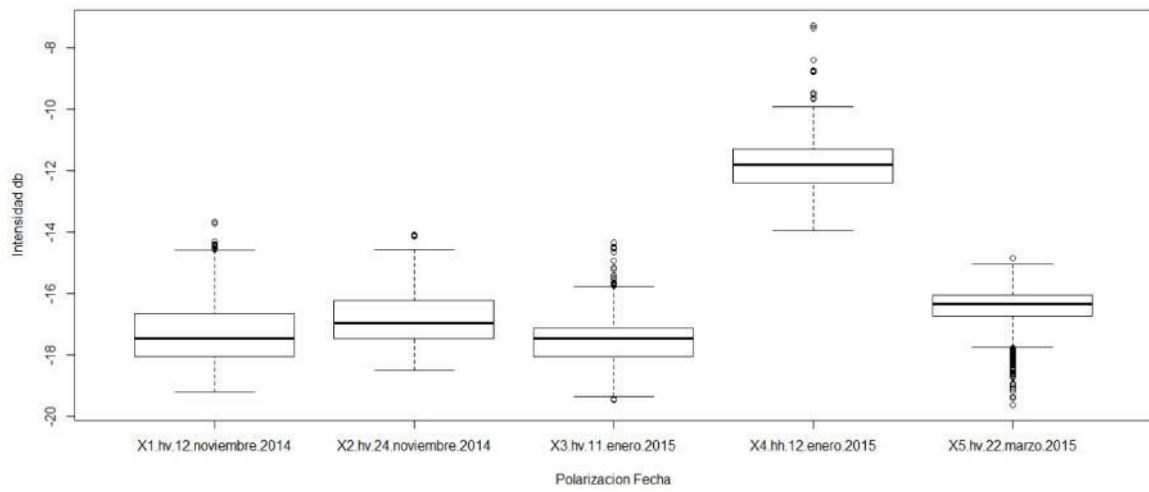
Clase:Pastizal



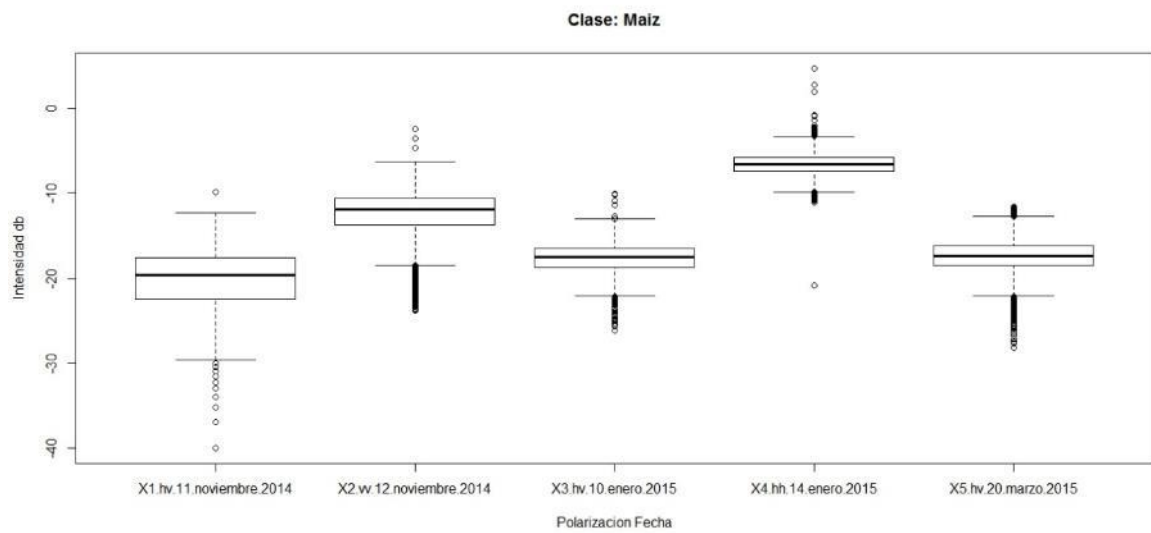
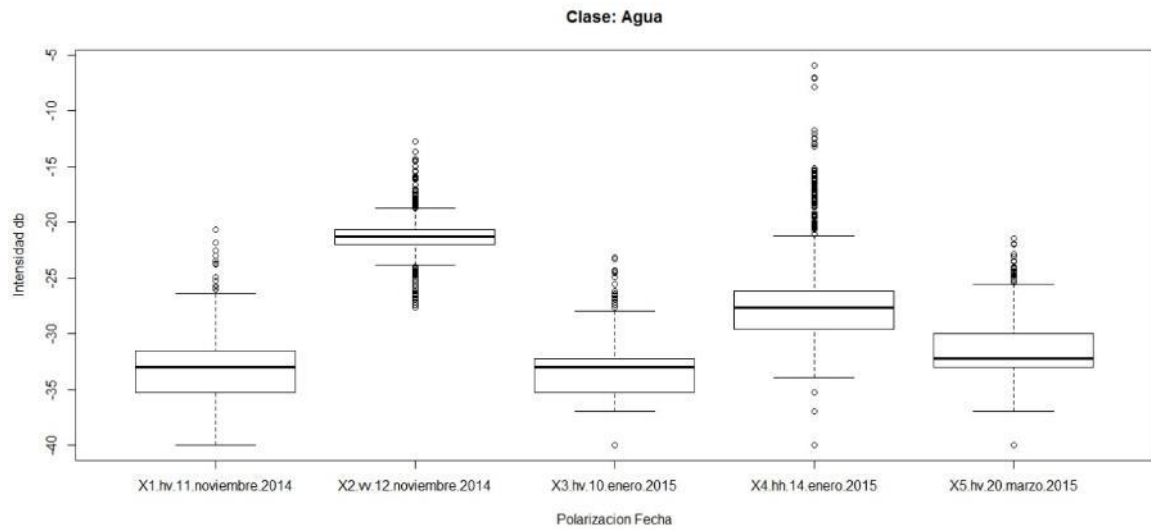
Clase: Soja de primera



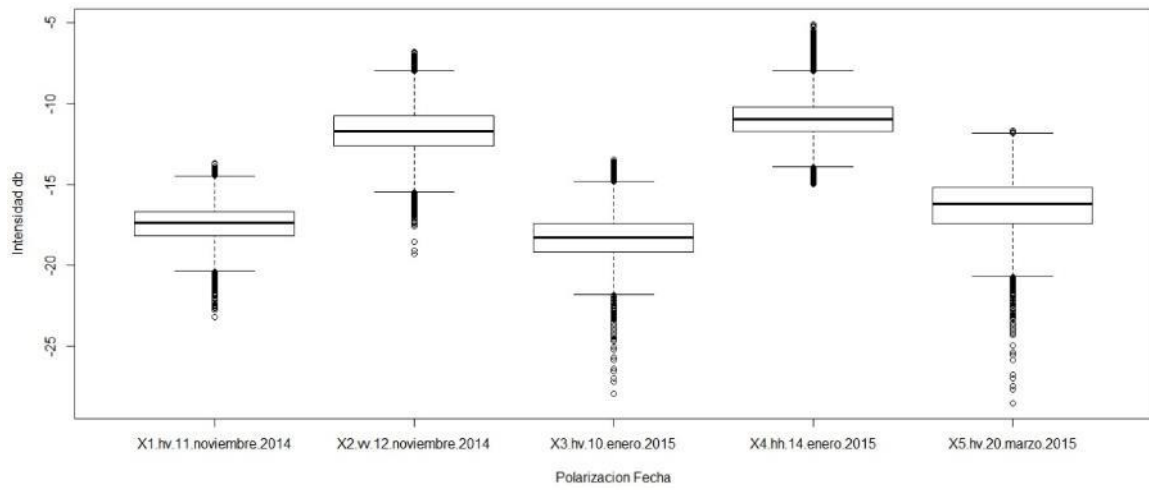
Clase: Soja de segunda



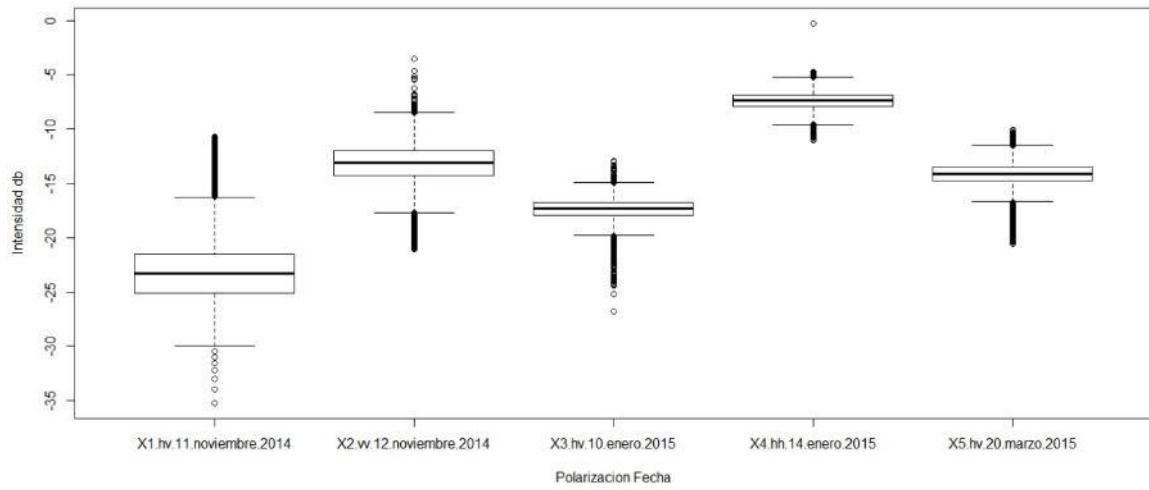
Boxplots clases Capítulo 4. Norte Provincia de Buenos Aires.



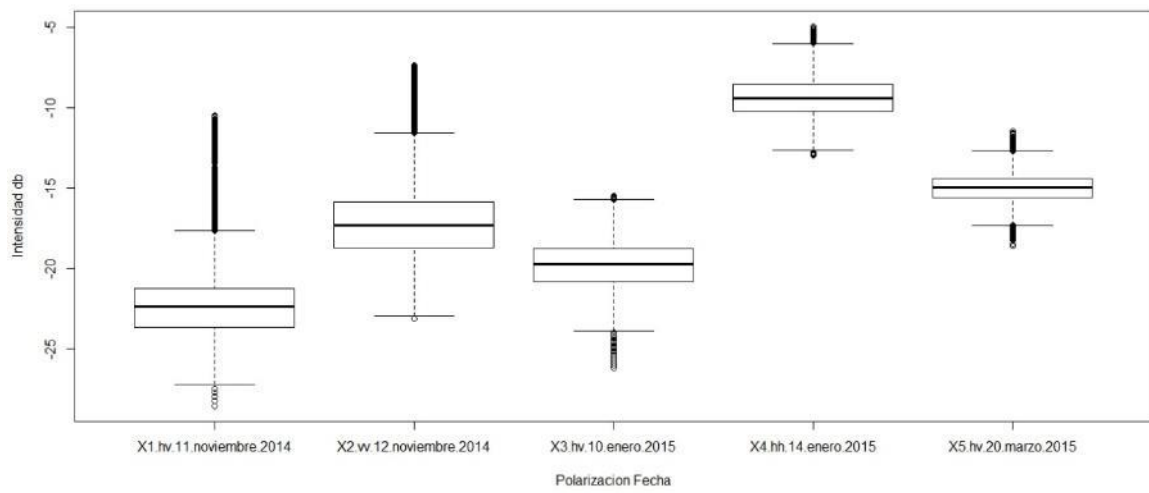
Clase: Pastizal



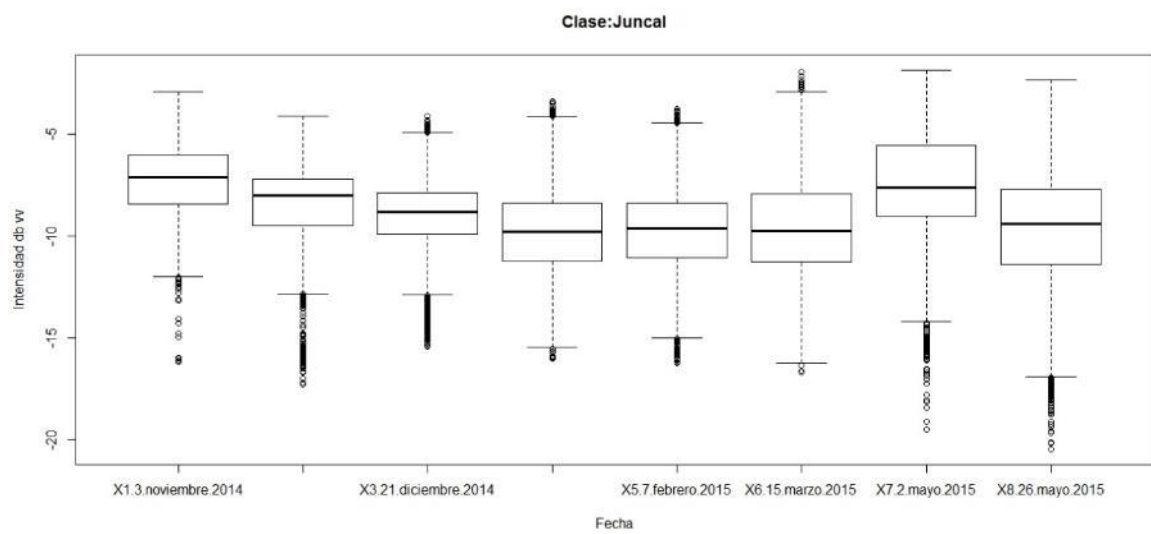
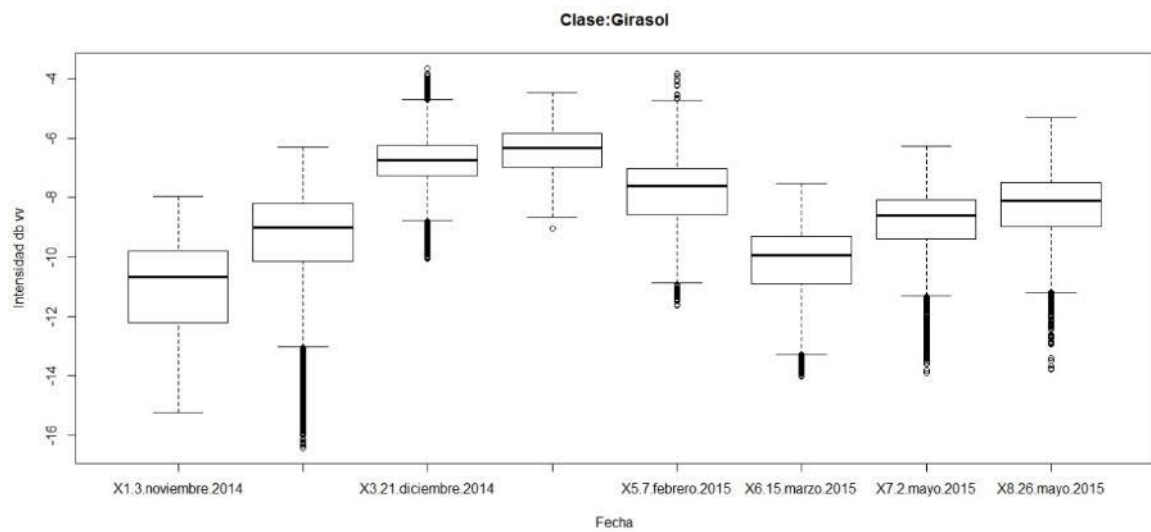
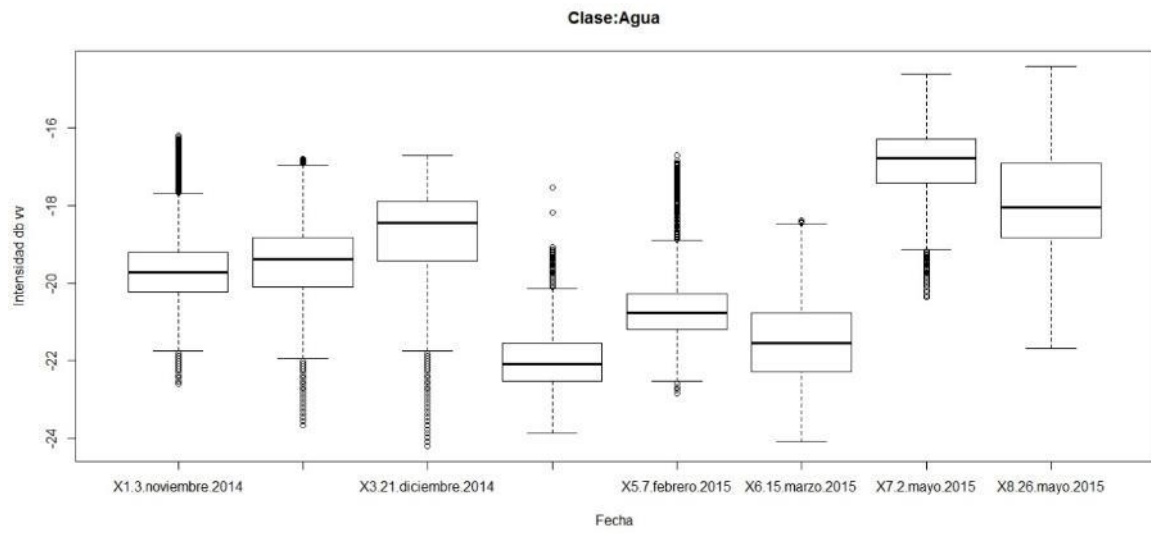
Clase: Soja de primera



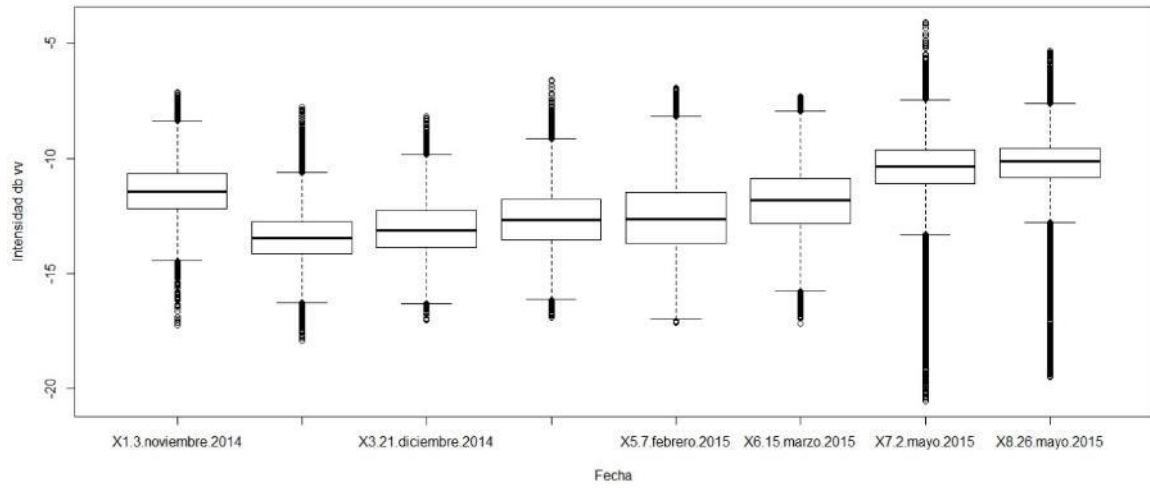
Clase: Soja de segunda



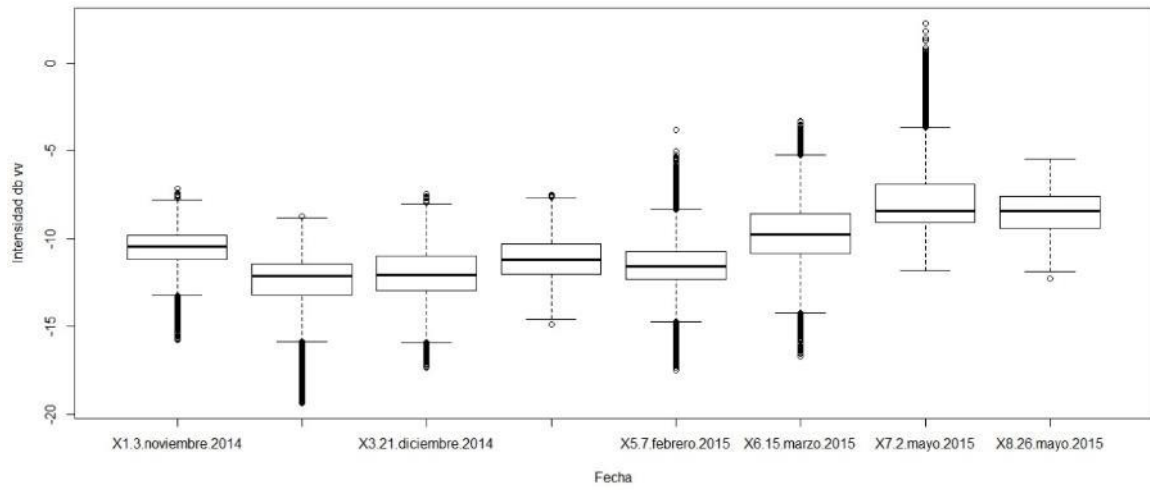
Boxplots clases Capítulo 5. Oeste Provincia de Buenos Aires.



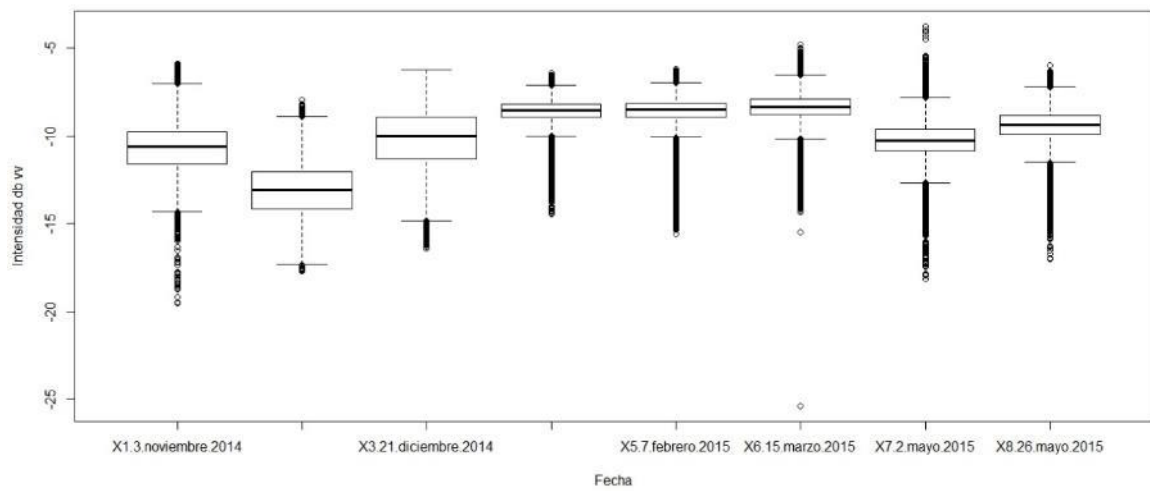
Clase:Pastizal



Clase:Maiz



Clase:Soja



3. Parámetros de los grafos de pre-procesamiento de las imágenes SAR.

1. Figura 3. Imágenes Cosmo Himage SLC. Producto: Imágenes sin filtrar, georreferenciadas, a su máxima resolución espacial posible en pixeles cuadrados.

Calibration: Source band: "Intensity". Se calibra a intensidad, y la salida son valores de retrodispersión calibrados.

Terrain Correction: Source banda: "Sigma0". Digital Elevation Model: "SRTM 3sec (autodownload)". Dem Resampling:"Cubic Convolution". Image Resampling:"Cubic Convolution". Map Projection: "WGS84(DD)". El resamplio elegido es Cubic Convolution para minimizer el efecto del resamplio sobre la distribución de los datos y los parámetros de la media y el desvío estándar.

2. Figura 12. Imágenes Cosmo Himage SLC. Producto: Imágenes filtradas y georreferenciadas.

Calibration: Source band: "Intensity". Se calibra a intensidad, y la salida son valores de retrodispersión calibrados.

Multilook: Source band: "Sigma0". GR Square Pixel. Number of Range looks:"3". Number of Azimuth looks:"3".El pixel resultante tendrá cerca de 6.5 m de lado, dependiendo del ángulo de incidencia de la imagen. Este es un proceso que filtra la imagen reduciendo el ruido "Speckle". Los números de looks elegidos permiten filtrar sin degradar demasiado la resolución espacial.

Terrain Correction: Source banda: "Sigma0". Digital Elevation Model: "SRTM 3sec (autodownload)". Dem Resampling:"Cubic Convolution". Image Resampling:"Cubic Convolution". Map Projection: "WGS84(DD)". El resamplio elegido es Cubic Convolution para minimizer el efecto del resamplio sobre la distribución de los datos y los parámetros de la media y el desvío estándar.

Speckle Filter: Source band: "Sigma0". Filter:"Lee". Filter Size X: "3". Filter Size Y: "3". Estimate equivalent number of looks. Este proceso permite filtrar más aún la imagen, reduciendo el ruido Speckle.

3. Figura 20. Imágenes Cosmo Wide Region GEC. Producto: Imágenes filtradas.

Calibration: Source band: "Intensity". Se calibra a intensidad, y la salida son valores de retrodispersión calibrados.

Multilook: Source band: "Sigma0". GR Square Pixel. Number of Range looks:"2". Number of Azimuth looks:"2".El pixel resultante tendrá cerca de 30 m de lado, dependiendo del ángulo de incidencia de la imagen. Este es un proceso que filtra la imagen reduciendo el ruido "Speckle". Los números de looks elegidos permiten filtrar sin degradar demasiado la resolución espacial.

Speckle Filter: Source band: "Sigma0". Filter:"Lee". Filter Size X: "5". Filter Size Y: "5". Estimate equivalent number of looks. Este proceso permite filtrar más aún la imagen, reduciendo el ruido Speckle.

4. Figura 21. Imágenes Cosmo Wide Region DGM. Producto: Imágenes georreferenciadas y filtradas.

Calibration: Source band: "Intensity". Se calibra a intensidad, y la salida son valores de retrodispersión calibrados.

Multilook: Source band: "Sigma0". GR Square Pixel. Number of Range looks:"2". Number of Azimuth looks:"2".El pixel resultante tendrá cerca de 30 m de lado, dependiendo del ángulo de incidencia de la imagen. Este es un proceso que filtra la imagen reduciendo el ruido "Speckle". Los números de looks elegidos permiten filtrar sin degradar demasiado la resolución espacial.

Terrain Correction: Source banda: "Sigma0". Digital Elevation Model: "SRTM 3sec (autodownload)". Dem Resampling:"Cubic Convolution". Image Resampling:"Cubic Convolution". Map Projection: "WGS84(DD)". El resamplio elegido es Cubic Convolution para minimizer el efecto del resamplio sobre la distribución de los datos y los parámetros de la media y el desvío estándar.

Speckle Filter: Source band: "Sigma0". Filter:"Lee". Filter Size X: "5". Filter Size Y: "5". Estimate equivalent number of looks. Este proceso permite filtrar más aún la imagen, reduciendo el ruido Speckle.

5. Figura 31. Imágenes Sentinel 1 IW. Producto: Imágenes georreferenciadas y filtradas.

Calibration: Source band: "Intensity". Se calibra a intensidad, y la salida son valores de retrodispersión calibrados.

Multilook: Source band: "Sigma0". GR Square Pixel. Number of Range looks:"3". Number of Azimuth looks:"3".El pixel resultante tendrá cerca de 30 m de lado, dependiendo del ángulo de incidencia de la imagen. Este es un proceso que filtra la imagen reduciendo el ruido "Speckle". Los números de looks elegidos permiten filtrar sin degradar demasiado la resolución espacial.

Terrain Correction: Source banda: "Sigma0". Digital Elevation Model: "SRTM 3sec (autodownload)". Dem Resampling:"Cubic Convolution". Image Resampling:"Cubic Convolution". Map Projection: "WGS84(DD)". El resamplio elegido es Cubic Convolution para minimizer el efecto del resamplio sobre la distribución de los datos y los parámetros de la media y el desvío estándar.

Bibliografía

- Bustos, Oscar H., Frery, Alejandro, Lamfri, Mario, and Scavuzzo, Marcelo. 2004. *Técnicas Estadísticas de Teledetección Espacial*.
- Chang, Chih-Chung, and Chih-Jen Lin. 2011. "LIBSVM: A Library for Support Vector Machines." *ACM Transactions on Intelligent Systems and Technology (TIST)* 2 (3): 27.
- Chen, K. S., W. P. Huang, D. H. Tsay, and F. Amar. 1996. "Classification of Multifrequency Polarimetric SAR Imagery Using a Dynamic Learning Neural Network." *Geoscience and Remote Sensing, IEEE Transactions on* 34 (3): 814–20.
- Chuvieco E. S. 2006. *Teledetección Ambiental*. Editorial Ariel.
- Dubois, Pascale C., Jakob Van Zyl, and Ted Engman. 1995. "Measuring Soil Moisture with Imaging Radars." *Geoscience and Remote Sensing, IEEE Transactions on* 33 (4): 915–26.
- ESA. n.d. "Sentinel-1 Observation Scenario." <https://sentinel.esa.int/web/sentinel/missions/sentinel-1/observation-scenario>.
- Ferrazzoli P. 2002. "Sar for Agriculture: Advances, Problems and Prospects."
- Floyd Henderson, and Anthony Lewis. 1998. *Principles and Applications of Imaging Radar*. Tercera edición. Vol. II. American Society for Photogrammetry and Remote Sensing.
- Frate, F. Del, G. Schiavon, D. Solimini, M. Borgeaud, D.H. Hoekman, and M.A.M. Vissers. 2003. "Crop Classification Using Multiconfiguration c-Band Sar Data." *IEEE Transactions on Geoscience and Remote Sensing* 41 (7): 1611–19. doi:10.1109/TGRS.2003.813530.
- Friedman, Jerome H. 1999. "Greedy Function Approximation: A Gradient Boosting Machine."
- Jin, Huiran, Giorgos Mountrakis, and Stephen V. Stehman. 2014. "Assessing Integration of Intensity, Polarimetric Scattering, Interferometric Coherence and Spatial Texture Metrics in PALSAR-Derived Land Cover Classification." *ISPRS Journal of Photogrammetry and Remote Sensing* 98 (December): 70–84. doi:10.1016/j.isprsjprs.2014.09.017.
- Joao Viane Soares, Camilo Daleles Rennó, Camilo Daleles Rennó, and Antonio Roberto Formaggio. 1997. "An Investigation of the Selection of Texture Features for Crop Discrimination Using SAR Imagery." *Remote Sensing of Environment* 59: 234–47.
- Kasischke, Eric S., John M. Melack, and M. Craig Dobson. 1997. "The Use of Imaging Radars for Ecological Applications—a Review." *Remote Sensing of Environment* 59 (2): 141–56.
- Kim, Yihyun, Thomas Jackson, Rajat Bindlish, Hoonyol Lee, and Sukyoung Hong. 2013. "Monitoring Soybean Growth Using L-, C-, and X-Band Scatterometer Data." *International Journal of Remote Sensing* 34 (11): 4069–82. doi:10.1080/01431161.2013.772309.
- La Nueva Provincia. 2014. "Pedro Luro: Atacan a Inspectores Que Controlaban Trabajo Informal." *La Nueva Provincia*, March 12.
- Lee, Jong-Sen, Mitchell R. Grunes, and Eric Pottier. 2001. "Quantitative Comparison of Classification Capability: Fully Polarimetric versus Dual and Single-Polarization SAR." *Geoscience and Remote Sensing, IEEE Transactions on* 39 (11): 2343–51.
- Lohmann, P., U. Soergel, M. Tavakkoli, D. Farghaly, and others. 2009. "Multi-Temporal Classification for Crop Discrimination Using TerraSAR-X Spotlight Images." *Proceedings IntArchPhRS* 38.
- Loosvelt, Lien, Jan Peters, Henning Skriver, Bernard De Baets, and Niko E. C. Verhoest. 2012. "Impact of Reducing Polarimetric SAR Input on the Uncertainty of Crop Classifications Based on the Random Forests Algorithm." *IEEE Transactions on Geoscience and Remote Sensing* 50 (10): 4185–4200. doi:10.1109/TGRS.2012.2189012.
- Loosvelt, Lien, Jan Peters, Henning Skriver, Hans Lievens, Frieke M.B. Van Coillie, Bernard De Baets, and Niko E.C. Verhoest. 2012. "Random Forests as a Tool for Estimating Uncertainty at Pixel-Level in SAR Image Classification." *International Journal of Applied*

- Earth Observation and Geoinformation* 19 (October): 173–84. doi:10.1016/j.jag.2012.05.011.
- Löw, F., U. Michel, S. Dech, and C. Conrad. 2013. “Impact of Feature Selection on the Accuracy and Spatial Uncertainty of per-Field Crop Classification Using Support Vector Machines.” *ISPRS Journal of Photogrammetry and Remote Sensing* 85 (November): 102–19. doi:10.1016/j.isprsjprs.2013.08.007.
- Luis Veci. 2015. “S1TBX SAR Basics Tutorial.pdf”. ESA.
- McNairn, H., Jiali Shang, Xianfeng Jiao, and C. Champagne. 2009. “The Contribution of ALOS PALSAR Multipolarization and Polarimetric Data to Crop Classification.” *IEEE Transactions on Geoscience and Remote Sensing* 47 (12): 3981–92. doi:10.1109/TGRS.2009.2026052.
- McNairn, Heather, Catherine Champagne, Jiali Shang, Delmar Holmstrom, and Gordon Reichert. 2009. “Integration of Optical and Synthetic Aperture Radar (SAR) Imagery for Delivering Operational Annual Crop Inventories.” *ISPRS Journal of Photogrammetry and Remote Sensing* 64 (5): 434–49. doi:10.1016/j.isprsjprs.2008.07.006.
- Open Source Computer Vision Library. n.d. “Machine Learning.” *OpenCV*. <http://docs.opencv.org/2.4/modules/ml/doc/ml.html>.
- OTB development team. 2014. “The Orfeo Toolbox Software Guide”. CNES. <http://www.orfeo-toolbox.org>.
- Paruelo, Di Bella, Milkovic, Verón, and de Abelleyra. 2014. *Percepción Remota Y Sistemas de Información Geográfica*. 1a ed. Ciudad Autonoma de Buenos Aires: Editorial Hemisferio Sur S.A.
- Satalino, Giuseppe, Donato Impedovo, Anna Balenzano, and Francesco Mattia. 2011. “Land Cover Classification by Using Multi-Temporal COSMO-SkyMed Data.” In *Analysis of Multi-Temporal Remote Sensing Images (Multi-Temp)*, 2011 6th International Workshop on the, 17–20. IEEE.
- Skriver, H, F Mattia, G Satalino, A Balenzano, V R N Pauwels, N E C Verhoest, and M Davidson. 2011. “Crop Classification Using Short-Revisit Multitemporal SAR Data.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 4 (2): 423–31. doi:10.1109/JSTARS.2011.2106198.
- Steduto, Pasquale, Theodore C. Hsiao, Dirk Raes, and Elias Fereres. 2009. “AquaCrop—The FAO Crop Model to Simulate Yield Response to Water: I. Concepts and Underlying Principles.” *Agronomy Journal* 101 (3): 426. doi:10.2134/agronj2008.0139s.
- Tian, Xin, Erxue Chen, Zengyuan Li, Z. Bob Su, Feilong Ling, Lina Bai, and Fengyu Wang. 2010. “Comparison of Crop Classification Capabilities of Spaceborne Multi-Parameter SAR Data.” In *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, 359–62. IEEE.
- Waske, Björn, and Matthias Braun. 2009. “Classifier Ensembles for Land Cover Mapping Using Multitemporal SAR Imagery.” *ISPRS Journal of Photogrammetry and Remote Sensing* 64 (5): 450–57. doi:10.1016/j.isprsjprs.2009.01.003.
- Wegmuller, Urs, and Charles Werner. 1997. “Retrieval of Vegetation Parameters with SAR Interferometry.” *Geoscience and Remote Sensing, IEEE Transactions on* 35 (1): 18–24.