

# FACULTAD DE CIENCIAS LICENCIATURA EN GEOLOGÍA TELEDETECCIÓN

## **INFORME FINAL**

Rubens Merlino

## **Docentes:**

Dra. Virginia Fernandez Arq. Virginia Pedemonte Zerpa



## **CONTENIDO**

#### 1. Introducción

- a. Características de las imágenes ópticas
- b. Particularidades de la misión Landsat
- c. Particularidades de la misión Sentinel

## 2. Objetivos

- a. General
- b. Específicos
- 3. Área de estudio
  - a. Localización
  - b. Descripción general

#### 4. Correcciones

- a. Corrección atmosférica
- b. Geométrica
- c. Procesamientos necesarios para mejorar la definición de clases
- 5. Otros procesamientos
- 6. Clasificación
  - a. Definición de las clases según la imagen
  - b. Clasificación no supervisada
- 7. Matriz de confusión
- 8. Conclusiones
  - a. Conclusiones sobre los resultados obtenidos
  - b. Aciertos y Desaciertos
  - c. Posibles mejoras
- 9. Bibliografía



## 1. INTRODUCCIÓN

## a. Características de las imágenes ópticas

Las imágenes ópticas son aquellas generadas por la reflexión de las longitudes de onda provenientes de la energía solar (Chuvieco, 1995). Dada esta definición, las características presentes en estas imágenes, provienen de las diferentes longitudes de onda que reflejan los objetos, las cuales son registradas en los sensores bajo la denominación de bandas.

Cada misión particular provee imágenes con distintas características según los sensores que portan. Así, a grandes rasgos, podemos hablar de distintos tipos de resoluciones en los sensores (Frank 2017, Chivieco 1995)

- Resolución espacial: ésta refiere al tamaño del píxel de la imagen, siendo que a menor tamaño de píxel, mayor será la resolución espacial de la imagen.
- Resolución espectral: este valor indica el ancho y número de bandas presentes en la imagen. Siendo cada banda, como mencionamos anteriormente, un rango específico del espectro electromagnético.
- Resolución radiométrica: este parámetro indica la cantidad de valores posibles dentro de cada píxel.

De esta manera, de acuerdo a estos tipos de resolución y a otros parámetros, se generan las imágenes a las que luego se les pueden aplicar combinaciones de bandas, índices, filtros y clasificaciones para poder estudiar desde diferentes ópticas la cobertura presente en un área del planeta y su evolución en el tiempo.

#### b. Particularidades de la misión Landsat

El programa Landsat es una serie de misiones de observación de la Tierra por satélite administrado conjuntamente por la NASA y el Servicio Geológico de EE.UU. (Frank 2017).

En particular, la misión Landsat 8, de la cual se extrae la imagen a procesar, contiene las siguientes características:

- Proyección UTM, zona 21.
- Datum: WGS84.
- Altitud del satélite 705 km.
- Resolución radiométrica: 16 Bits.
- Resolución espectral: multiespectral.
- Resolución espacial para las bandas:
  - o 30m OLI bandas multiespectral 1-7,9.
  - o 15m OLI banda pancromática 8.
  - o 100m TIRS bandas 10-11.
- Resolución temporal o tiempo de revisita: Heliosincrónico de 16 días.



#### c. Particularidades de la misión Sentinel

Sentinel-2 es una misión de observación terrestre desarrollada por la Agencia Espacial Europea (ESA) dentro del programa Copérnico, para desarrollar observaciones del planeta Tierra para dar servicios como el seguimiento de la evolución de los bosques, los cambios en la corteza terrestre y la gestión de los desastres naturales. Está compuesto por dos satélites idénticos: Sentinel-2A y Sentinel-2B.

La misión Sentinel-2 tiene las siguientes características:

- Imagen multiespectral, 13 bandas en el Espectro visible, en el infrarrojo cercano e infrarrojos de onda corta además del espectro electromagnético
- Cobertura global sistemática de las capas de tierra de 56° S a 84° N.
- Resolución temporal o tiempo de revisita: Heliosincrónico de 5 días.
- Resolución espacial de 10m, 20m y 60m según las bandas.

#### 2. OBJETIVOS

#### a. General

El objetivo general fue aprender a utilizar diversas herramientas que nos permitan poder realizar estudios de teledetección a través de la manipulación de diversas imágenes provenientes de distintos tipos de misiones. Para ello se utilizaron imágenes de las misiones Landsat (NASA) y Sentinel (The European Space Agency), las cuales fueron procesadas mediante el uso de índices, combinaciones de bandas, filtros y clasificaciones.

Como objetivos generales podemos resumir:

- Conocer las misiones Landsat y Sentinel-2a
- Comprender los conceptos de reflectancia y sus resultados en imágenes satelitales.
- Descargar imágenes y aplicar distintos filtros de búsqueda.
- Introducirnos en el uso del software SoPI (Software de Procesamiento de Imágenes)
- Introducirnos en el uso del software SNAP (Sentinel Application Platform)
- Desarrollar los conceptos de índices, bandas espectrales y combinaciones de las mismas para visualizaciones con diferentes objetivos.
- Conocer los conceptos de clasificación automática tanto supervisada como no supervisada y sus diferentes procesamientos para medir su precisión y efectividad para clasificar los diferentes tipos de uso de suelo.
- Introducirnos en el uso del software QGIS (Quantum GIS) para ampliar el procesamiento de imágenes tipo Raster previamente analizadas en SNAP o SoPI.

#### b. Específicos

Como objetivos específicos se intentó a través de las imágenes descargadas obtener una visualización clara de los afloramientos rocosos vinculados a la zona de estudio (I) y a través de algoritmos de clasificación precisar su distribución areal y cuantificar su extensión en la zona estudiada (II).



## 3. ÁREA DE ESTUDIO

#### a. Localización

La zona a estudio se encuentra entre los departamentos de Lavalleja, Treinta y Tres y rocha, de la República Oriental del Uruguay. (Fig. 1)

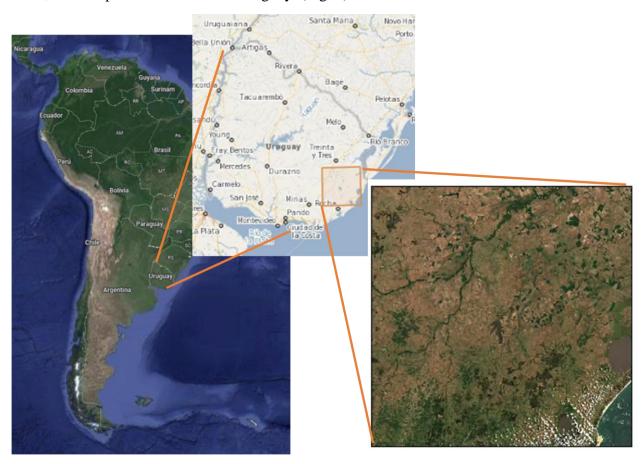


Fig 1: Ubicación del área a estudio.

## b. Descripción general

El estudio se focalizó en el sector NE del llamado lineamiento Santa Lucía – Aiguá – Merín (SaLAM), el cual está vinculado a eventos tectónicos acaecidos en el período cretácico (De Sant Ana *et al. 2004*) (fig. 2). En este sector, a través de imágenes satelitales y cartas topográficas pueden identificarse claramente afloramientos pertenecientes a la formación Arequita y formación Valle Chico ambas inferidas según la memoria explicativa de la carta geológica del Uruguay (Preciozzi *et al. 1985*) disponible en la Dirección Nacional de Minería y Geología (DINAMIGE).

Los afloramientos mencionados se encuentran a las riberas del río Cebollatí, y su distribución se ve intercalada con sectores de suelo desnudo y herbáceo. Si bien a simple vista en imágenes satelitales pueden identificarse claramente, la intercalación con las otras coberturas del suelo y su valor de reflectancia similar a valores de suelo desnudo *a priori* representan un desafío mayor para su cuantificación por algoritmos de clasificación.



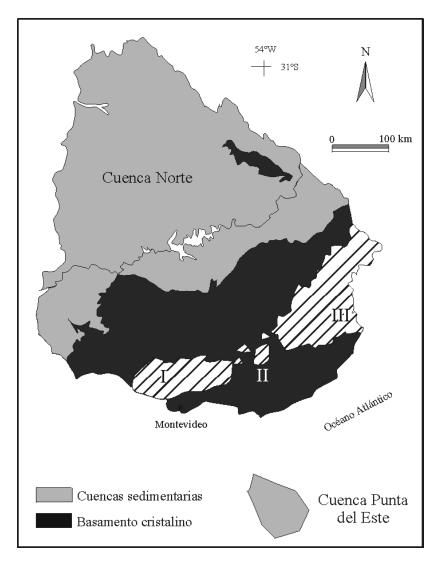


Fig 2: Mapa de las cuencas sedimentarias del Uruguay, en III se encuentra el área de estudio.

#### Las imágenes utilizadas fueron:

#### Imagen Sentinel-2a:

• Id: S2A\_MSIL2A\_20220112T133231\_N0301\_R081\_T21HYC\_20220112T160117

• **Fecha**: 2022-01-12

#### Imagen Landsat 8:

• **Id**: LC08\_L1TP\_223083\_20211122\_20211130\_01\_T1

• **Fecha**: 2021-11-30

Path: 223Row: 83



#### 4. CORRECCIONES

#### a. Corrección atmosférica

Se procedió a realizar el cálculo de reflexión con corrección atmosférica TOA. De acuerdo a los datos provistos por USGS (United State Geological Survey) para los productos de Landsat nivel 1, el cálculo de reflexión con corrección atmosférica TOA se realiza mediante la siguiente ecuación:

$$\rho \lambda' = M\rho * Qcal + A\rho$$

#### Donde:

 $\rho \lambda'$  = reflectancia al tope de atmósfera.

 $M\rho$ = factor multiplicativo específico para cada banda. Obtenido de la metadata de la imagen.

 $A\rho$  = valor específico aditivo para cada banda. Obtenido de la metadata de la imagen.

*Qcal* = Valor cuantizado del pixel

#### Ecuación para TOA:

$$\rho\lambda = \rho\lambda'\cos(\theta SZ) = \rho\lambda'\sin(\theta SE)\rho\lambda = \rho\lambda'\cos[f(\theta SZ)] = \rho\lambda'\sin[f(\theta SE)]$$

#### Donde:

 $\rho\lambda$ = TOA planetary reflectance

 $\theta SE$  = Elevación local del sol.

 $\theta SZ$  = Angulo local del zenith



**Fig 3:** Selección de las bandas 2 a la 7 y aplicación de realce "*Ecualización de Histograma*" (*software SoPI*).



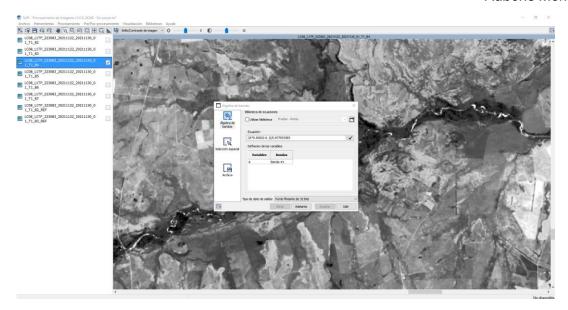


Fig 4: Aplicación de corrección atmosférica (software SoPI).

De acuerdo a los datos suministrados por la metadata del archivo se utilizó la siguiente ecuación de corrección atmosférica:

#### (A\*0.00002-0.1)/0.877953585

Siendo A el valor de cada píxel de la capa



**Fig 5:** Detalle del afloramiento perteneciente a la Fm Valle Chico en la imagen de la banda N° 2 con realce "*Ecualización de Histograma*" *Izq:* sin aplicación de la corrección atmosférica TOA. *Der:* con aplicación de la corrección atmosférica TOA (*software SoPI*).

Como puede observarse, la aplicación de la corrección atmosférica TOA realza varios sectores del afloramiento.



#### b. Geométrica

Se procedió a realizar el apilado de bandas y la reproyección al sistema UTM, zona 21S.



Fig 6: Izq: Apilado de bandas. Der: Reproyección del raster a UTM zone 21S (software SoPI).

## c. Procesamientos necesarios para mejorar la definición de clases

Para la mejora de la visualización se aplicaron índices y combinaciones de bandas, a continuación se muestra el álgebra de bandas realizada para la aplicación del índice NDVI al sector del afloramiento de la Fm Valle Chico

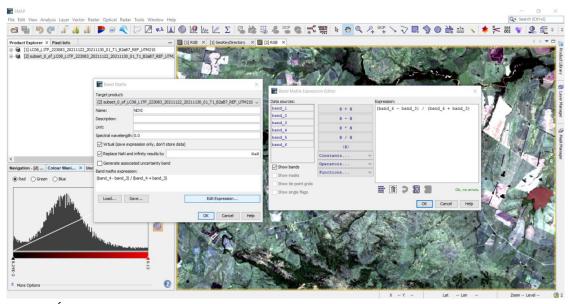
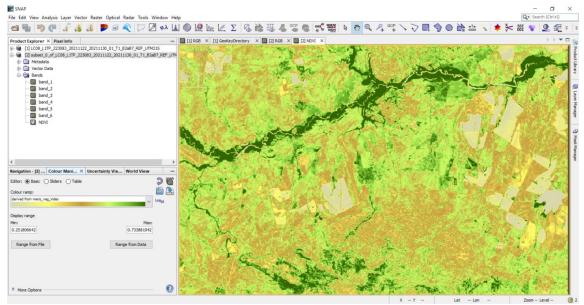


Fig 7: Índice NDVI: Landsat 8 (banda 5 - banda 4) / (banda 5 + banda 4). (software SNAP).

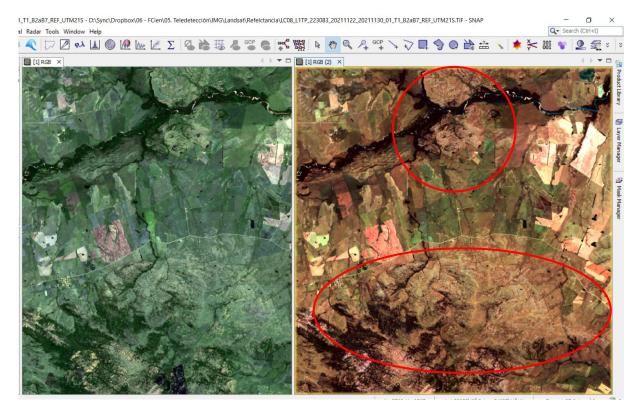




**Fig** 8: Aplicación de color map: *meris\_veg\_index* Que indica en los colores verde más fuertes la vegetación vigorosa. (*software SNAP*).

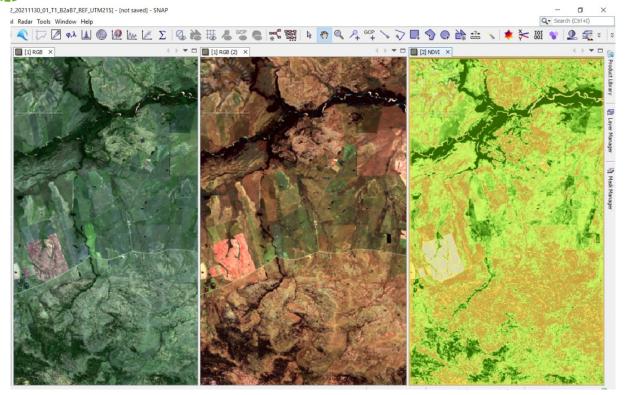
Como puede apreciarse en la fig 8, la aplicación del índice NDVI, debido a las características que ofrece de realce de vegetación y a las características de la zona, no dan buen resultado para realzar las zonas del afloramiento.

Debido a lo anterior se optó por realizar una combinación de bandas particular, la 742 (RGB) que tiene como objetivo el realce de estructuras gelógicas. A continuación se presentan los resultados.



**Fig 9:** Izq: imagen con color natural, der: aplicación de la combinación de bandas 742 (RGB) para destaque de afloramientos y estructuras geológicas. (*software SNAP*).





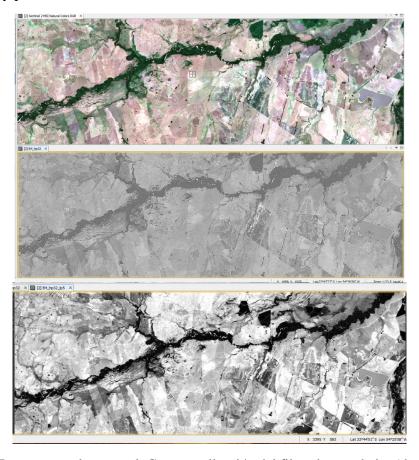
**Fig 10: Izq**: imagen con color natural; **centro**: aplicación de la combinación de bandas 742 (RGB) para destaque de afloramientos y estructuras geológicas; **der**: índice NDVI con color map *meris\_veg\_index*. (*software SNAP*).

Como puede observarse en la figura 9 y 10, la aplicación de la combinación de bandas 742, permite destacar considerablemente los macizos rocosos aflorantes marcados con círculos rojos en la fig. 9. Asimismo, comparándola con la imagen de la derecha de la fig.9 puede observarse claramente como la aplicación del índice NDVI destaca sobremanera la vegetación del monte nativo a las riberas del río Cebollatí.



## 5. OTROS PROCESAMIENTOS

Continuando con el objetivo de resaltar los afloramientos rocosos, se aplicaron los filtros de paso bajo y paso alto.



**Fig 11: Arriba:** Imagen en color natural. Centro: aplicación del filtro de paso bajo. Abajo aplicación del filtro de paso alto. (*software SNAP*).

Como se puede apreciar en la fig 11, la aplicación del filtro de paso alto, reforzó los contornos y enfatizó los rasgos lineales de los afloramientos (Chuvieco, 1995).



## 6. CLASIFICACIÓN

## a. Definición de las clases según la imagen

Para la imagen extraída de la misión Sentinel 2A, para el sector de interés, se definieron siete clases de acuerdo a la cobertura presente:

Clase:
Herbáceo
Afloramientos
Suelo desnudo
Forestación
Monte nativo
Cultivos
Agua

Estas clases serán cuantificadas por medio de algoritmos de clasificación no supervisada (K-Means) y de clasificación supervisada (Random Forest y Maximum Likelihood)

Previo a la aplicación de los algoritmos de clasificación, se realizó un subconjunto de la imagen Sentinel, conteniendo únicamente las bandas (*B2*, *B3*, *B4*, *B5*, *B6*, *B7*, *B8*, *B8A*, *B11 Y B12*) y a este subconjunto se le aplicó un remuestreo a 10m tomando como parámetro la banda B2.

## b. Clasificación no supervisada

Para la aplicación de la clasificación no supervisada se empleó el algoritmo K-Means, en una primera instancia con 9 clusters y 10 iteraciones, pero al ver resultados muy magros, se intentó con una clasificación basada en 15 clusters y 30 interaciones<sup>1</sup>, está última resulto mucho mejor clasificada y en base a ella se procedió a la reclasificación de las 15 clases a las 7 descritas anteriormente.

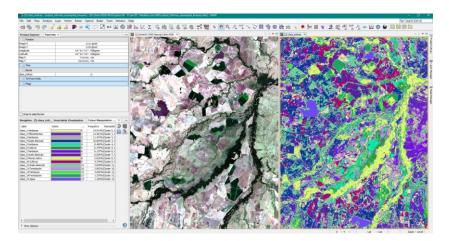


Fig 12: Resultado de la clasificación a través del algoritmo no supervisado K-Means.  $(software SNAP)^2$ .

<sup>&</sup>lt;sup>1</sup> En ambos casos se utilizó el parámetro que venía por defecto de "random seed = 31415", y se seleccionaron todas las bandas.

<sup>&</sup>lt;sup>2</sup> El procesamiento de este algoritmo llevó 3hs, 15m, 38.099s, para un equipo con procesador Intel Core i5 2400, 8Gb de memoria RAM y disco HDD.



Como se indicó anteriormente, el algoritmo obtuvo 15 clases, las cuales luego de haber sido identificadas, fueron reagrupadas según su misma categoría a través del software QGIS, mediante su herramienta "reclasificación por tabla". Obteniendo el siguiente resultado:

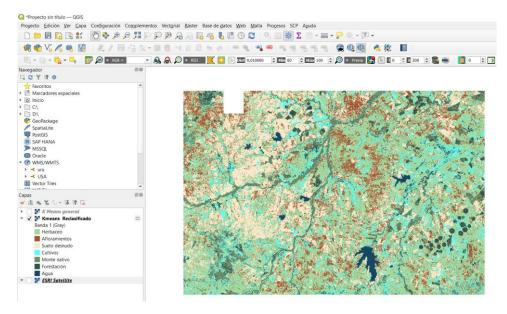


Fig 13: Resultado de la reclasificación a las 7 bandas establecidas en el punto 6.a. (software QGIS).

En la siguiente tabla se muestran los porcentajes establecidos según los distintos tipos de cobertura:

Clase	% de cobertura
Herbáceo	44,0790
Suelo desnudo	20,4530
Afloramientos	12,2610
Cultivos	11,1540
Monte nativo	5,8100
Forestación	5,0780
Agua	1,1650

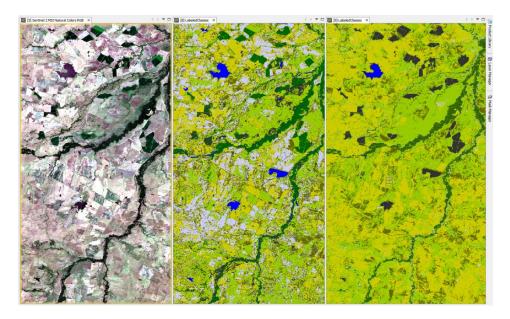
## c. Clasificación supervisada

Para poder aplicar los algoritmos de clasificación supervisada, primero se reproyectó la imagen a proyección geográfica (WGS84). Posteriormente para la fase de entrenamiento se desarrollaron 70 polígonos, 10 polígonos por cada clase. A continuación se muestran los resultados de los algoritmos de clasificación supervisada utilizados.





**Fig 14:** Detalle de algunos polígonos creados para la fase de entrenamiento de los algoritmos (*software SNAP*).



**Fig 15: Izq**: imagen con color natural; **centro**: resultado de las clases generadas por el algoritmo Random Forest; **der**: resultado de las clases generadas por el algoritmo Maximum Likelihood. (*software SNAP*).



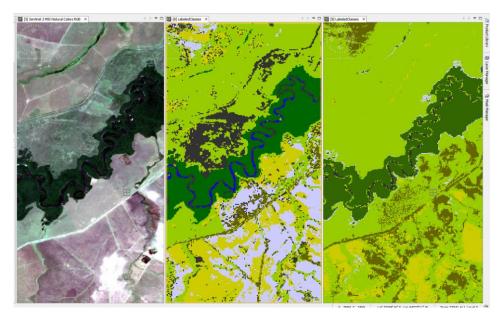


Fig 16: Detalle de una sección del río observando las clsase Izq: imagen con color natural; centro: resultado de las clases generadas por el algoritmo Random Forest; der: resultado de las clases generadas por el algoritmo Maximum Likelihood. (software SNAP).

Los parámetros utilizados para la creación de los algoritmos y sus resultados fueron los siguientes:

Algoritmo: Random Forests

Muestras: 10000

Número de árboles = 10

**Tiempo de procesamiento<sup>3</sup>**: 262,633 minutos (17890 B/s 5963 Pixels/s)

Clase	% de cobertura				
Herbáceo	40,2710				
Suelo desnudo	18,2280				
Cultivos	14,6860				
Afloramientos	9,8800				
Agua	7,3150				
Monte nativo	4,8130				
Forestación	4,8080				

Algoritmo: Maximum Likelihood

Muestras: 5000

Tiempo de procesamiento<sup>4</sup>: 50,166 minutos (109978 B/s 36659 Pixels/s)

Clase	% de cobertura
Herbáceo	43,4010
Suelo desnudo	34,0090
Afloramientos	13,6660
Monte nativo	3,7257
Forestación	3,0860
Cultivos	1,8310
Agua	0,7500

<sup>&</sup>lt;sup>3</sup> Ver nota al pie #2. <sup>4</sup> Ver nota al pie #2.



## 7. MATRIZ DE CONFUSIÓN

Para analizar los resultados de los algoritmos de clasificación previamente se reproyectaron las imágenes a UTM zona 21S y luego se les realizó una matriz de confusión mediante el plugin *AcATaMa* del software QGIS.

Para cada algoritmo se seleccionaron 15 muestras por clase y luego se procedió a clasificar las muestras.

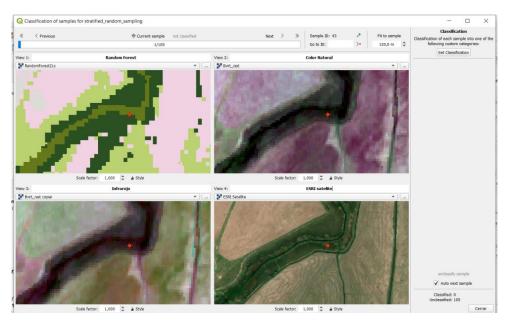


Fig 17: Detalle del proceso de testeo de la clasificación. (software QGIS).

Luego del testeo se obtuvieron las siguientes matrices de confusión:

RANDOM FOREST									
	Forestación	Cultivos	Monte Nativo	Herbáceo	Afloramiento	Suelo desnudo	TOTAL	PRECISIÓN	ÁREA DE LA CLASE (m²)
Forestación	8	0	5	2	0	0	15	53,33%	26.897.141.779.716
Cultivos	0	9	0	5	0	1	15	60,00%	74.995.471.493.574
Monte Nativo	4	0	10	1	0	0	15	66,67%	23.143.419.535.419
Herbáceo	0	4	0	11	0	0	15	73,33%	198.429.170.922.363
Afloramientos	0	3	0	7	3	2	15	20,00%	49.598.499.893.949
Suelo desnudo	0	1	0	3	0	11	15	73,33%	88.874.899.693.626
TOTAL	12	17	15	29	3	14	90		4.619.386.033.186.470
PRECISIÓN	66,67%	52,94%	66,67%	37,93%	100,00%	78,57%	PRECISIÓN:	57,78%	

Tabla 1. Matriz de confusión para el algoritmo de clasificación Random Forest.



4/10										
MAXIMUM LIKELIHOOD										
	Forestación	Cultivos	Monte Nativo	Herbáceo	Afloramiento	Suelo desnudo	TOTAL	PRECISIÓN	ÁREA DE LA CLASE (m²)	
Forestación	11	0	0	0	0	0	11	100,00%	14.723.789.951.878	
Cultivos	0	0	0	0	0	1	1	0,00%	8.679.041.117.338	
Monte Nativo	1	0	4	0	0	0	5	80,00%	15.024.500.807.499	
Herbáceo	0	2	4	26	1	0	33	78,79%	203.282.858.212.324	
Afloramientos	0	2	0	8	2	2	14	42,86%	64.595.345.852.221	
Suelo desnudo	0	1	0	6	1	18	26	69,23%	159.860.122.524.351	
TOTAL	12	5	8	40	4	21	90		4.661.656.584.656.110	
PRECISIÓN	91,67%	0,00%	50,00%	65,00%	50,00%	85,71%	PRECISIÓN:	67,78%		

Tabla 2. Matriz de confusión para el algoritmo de clasificación Maximum Likelihood.

A su vez cabe aclarar que no figura la clase Agua, debido a que, como se puede ver en la fig 18, la reproyección efectuada en el software SNAP genera un marco que es identificado como la clase 0 (Agua), la cual a su vez luego de la reproyección realizada en QGIS se amplía y cómo la toma automática de muestras selecciona alguno de estos puntos, seguramente dé algún error interno que, al generar la matriz de confusión, elimina la clase agua de la misma.

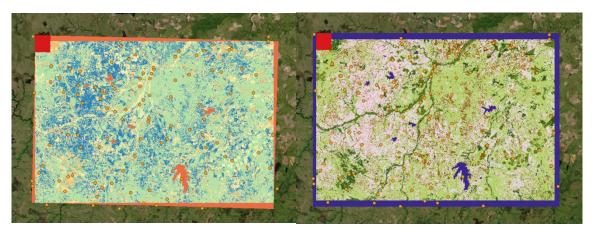


Fig 18: Detalle del error de margen y su asimilación a la clase agua. (software QGIS).

Asimismo, se intentó generar la matriz de confusión para el algoritmo de clasificación no supervisada K-Means sin resultados ya que el plugin del software QGIS dio el siguiente error "AcATaMa: Error, thematic raster must be byte or integer as data type".



#### 8. CONCLUSIONES

#### a. Conclusiones sobre los resultados obtenidos

Los resultados obtenidos mediante los algoritmos de clasificación se expresan en la siguiente tabla donde se especifica el porcentaje de uso de suelo estimado por cada algoritmo, así como el promedio de estos porcentajes, la diferencia entre el porcentaje calculado y el promedio para cada algoritmo y por último la suma de diferencias como indicador de variabilidad del porcentaje calculado para cada uso del suelo:

		LIKELIHOOD		RANDOM FOREST		K-	MEANS	
Uso Suelo	Promedio	% Uso	dif Promedio	% Uso	dif Promedio	% Uso	dif Promedio	Suma diferencias
Herbáceo	42,58	43,40	0,82	40,27	-2,31	44,08	1,50	4,63
Suelo desnudo	24,23	34,01	9,78	18,23	-6,00	20,45	-3,78	19,56
Afloramientos	11,94	13,67	1,73	9,88	-2,06	12,26	0,33	4,11
Monte nativo	4,78	3,73	-1,06	4,81	0,03	5,81	1,03	2,11
Forestación	4,32	3,09	-1,24	4,81	0,48	5,08	0,75	2,48
Cultivos	9,22	1,83	-7,39	14,69	5,46	11,15	1,93	14,79
Agua	3,08	0,75	-2,33	7,32	4,24	1,17	-1,91	8,48
_	Precisión:	67,78%		57,78%		No o	calculada	

Tabla 3. Resultados generales.

Dados los resultados obtenidos, podemos decir que debido a la baja precisión calculada para cada algoritmo de clasificación, estos datos resultan poco concluyentes. Como primera hipótesis sobre esta baja precisión, se puede inferir que dadas las características de la imagen y el período del año en que fueron extraídas, resulta muy difícil diferenciar, incluso observando la imagen satelital, qué tipo de cobertura corresponde a herbáceo, cuál a cultivos y cuál a suelo desnudo, ya que en esta zona para esta época del año los cultivos pueden verse poco frondosos o estar en una etapa de cosecha y generar falsos positivos para la clase de suelo desnudo y falsos negativos para la clase cultivos. De igual modo suelen confundirse, debido a sus valores de reflectancia, el suelo desnudo y los afloramientos rocosos que eran el objetivo principal de esta investigación.

De esta manera, siguiendo los objetivos, podemos decir que (I) mediante la combinación de bandas y la aplicación de filtros, pueden realzarse considerablemente los afloramientos rocosos permitiendo una mejor identificación y la posibilidad de creación de cartas de afloramientos de cierta precisión. Por otro lado (II) no se puede determinar con cierto grado respetable de precisión el valor cuantitativo de los afloramientos presentes en la zona.

#### **b.** Aciertos y Desaciertos

Se tomó como acertada la decisión de seleccionar únicamente la siete clases de cobertura mencionadas en el punto 6.a ya que esas clases resultan suficientes para caracterizar la zona. Asimismo se considera correcta la utilización de los algoritmos de clasificación mencionados ya que su eficiencia es reconocida y están presentes en muchos trabajos en la actualidad, formando parte del estado del arte en la materia.



Por otro lado podemos mencionar como desaciertos, la utilización de una imagen con un área tan extensa, pudiéndose expresar con mejor diferencia los afloramientos en áreas más reducidas.

#### c. Posibles mejoras

Se entiende que para obtener mejores resultados se podrían realizar los mismos procedimientos para imágenes de distintas épocas del año para determinar con mayor precisión las clases de cobertura.

Asimismo se considera que trabajar con áreas más reducidas puede ayudar a mejorar la precisión de la clasificación.

## 9. BIBLIOGRAFÍA

CHUVIECO, E. (1995). Fundamentos de Teledetección. Ed. Rial Ariel.

FRANK, N. (2017) Teoría - Teledetección espacial, Landsat, Sentinel-2, Aster L1T y Modis. Primera edición. Geomática Ambiental SRL

DE SANTA ANA, H. GOSO AGUILAR, C. MONTAÑO, J. PIÑEIRO, G. MUZIO, R. ROSSELLO, E. PEREA, D. UCHA, N. 2004 Cuencas sedimentarias del Uruguay. Segunda Edición.

PRECIOZZI, F. SPOTURNO, J. HEINZEN, W. ROSSI, P. 1985, Memoria explicativa de la Carta geológica del Uruguay. DINAMIGE.

Atlas de Cobertura del Suelo del Uruguay. 2011