

Metodología para la clasificación de cobertura del suelo en Argentina

Gonzalo Dieguez Gaviola¹, Evelyn Tengler², Augusto Van Der Ploeg³, Martín Fernando Ortiz⁴

^{1,2,3 y 4} Departamento de Aplicaciones Geoespaciales, parte de la Dirección de Información Geoespacial del Instituto Geográfico Nacional (IGN). Avenida Cabildo 381, CABA. {aploeg, gdieguez, etengler, mortiz}@ign.gob.ar

Resumen: El Instituto Geográfico Nacional de Argentina tiene la misión de entender la representación oficial del territorio nacional y elaborar marcos normativos para su realización, mediante la obtención de información geográfica precisa, oportuna y concisa. La disponibilidad de información satelital con amplia cobertura espacial y temporal, sumado a la acumulación de series de datos que permiten incorporar a la dimensión espacial y temporal, inducen a dirigir los esfuerzos en el desarrollo de metodologías que admiten una actualización permanente y el estudio de cambios cuali-cuantitativos en las coberturas y usos de suelo en el tiempo, especialmente debidos a intervenciones humanas, que tienen efectos significativos en el ambiente. En este contexto, se propone una metodología para establecer, de manera gráfica y cuantitativa, los distintos tipos de cobertura de la nación Argentina, para un análisis de su evolución espacial y temporal. Se dividió la nación en 6 regiones y para ello se definieron 12 categorías de cobertura y se procedió a utilizar el algoritmo Random Forest en la plataforma basada en la nube Google Earth Engine. Los resultados otorgan índices de precisión aproximados al 80% dando una idea de la calidad de la misma.

Palabras Clave: COBERTURA, PLANIFICACION, RANDOM FOREST, AMBIENTE, ANALISIS TEMPORAL, ANALISIS ESPACIAL

1. INTRODUCCIÓN

Conocer y caracterizar la cobertura de suelo, y eventualmente monitorear su cambio en el tiempo constituyen insumos fundamentales para múltiples aplicaciones, que van desde la planificación y gestión del territorio y sus recursos naturales, hasta la prevención e intervención en materia de Riesgo de Desastres [1]–[3].

El Comité de Expertos sobre la Gestión Mundial de la Información Geoespacial (UN-GGIM) es el órgano intergubernamental de Naciones Unidas encargado de debatir, mejorar y coordinar las actividades de gestión de la información geoespacial mundial con la participación de los Estados miembros y cuenta con su capítulo regional: UN-GGIM Américas. Su objetivo es maximizar el uso y los beneficios económicos, sociales y ambientales derivados de la información geoespacial, basado en el conocimiento y el intercambio de experiencias y tecnologías por parte de los Estados miembros [4], [5]. En este contexto, un grupo especializado del UN-GGIM realizó la identificación de una serie de 14 datos básicos y fundamentales para fortalecer la infraestructura de datos espaciales de un país, entre los cuales se encuentran las coberturas y usos de suelo.

Por su parte, el Instituto Geográfico Nacional tiene la misión de entender la representación oficial del territorio nacional y en la elaboración de los marcos normativos para su realización, mediante la obtención de información geográfica precisa, oportuna y concisa.

La gran disponibilidad de información satelital con amplia cobertura espacial y temporal, sumado a la acumulación de series de datos que permiten incorporar a la dimensión espacial y la dimensión temporal, inducen a dirigir los esfuerzos en el desarrollo de metodologías que admiten una actualización permanente, y el estudio de cambios cuali-cuantitativos en las coberturas y usos de suelo en el tiempo, especialmente debidos a intervenciones antrópicas, que tienen efectos significativos en el ambiente [2], [3].

Se propone aquí explorar una metodología para establecer, de manera gráfica y cuantitativa, los distintos tipos de cobertura de la nación Argentina, para un análisis de su evolución espacial y temporal, compatibilizando las coberturas de suelo que pueden ser discernidas con técnicas de teledetección, y las clases establecidas en el Catálogo de Objetos Geográficos institucional [6]

2. MATERIALES Y MÉTODOS

2 - 1. Área de estudio

El área de estudio seleccionada para implementar la metodología fue la República Argentina, exceptuando los territorios de la Antártida. En la figura 1 se observa el recorte del mapa ArgenMap (<https://mapa.ign.gob.ar>) con el área propiamente dicha. Argentina cuenta con un total de 18 ecorregiones, resultado de un conjunto amplio de características climáticas, topográficas, geológicas, biológicas, entre muchas otras que conforman un gran mosaico de coberturas [7].

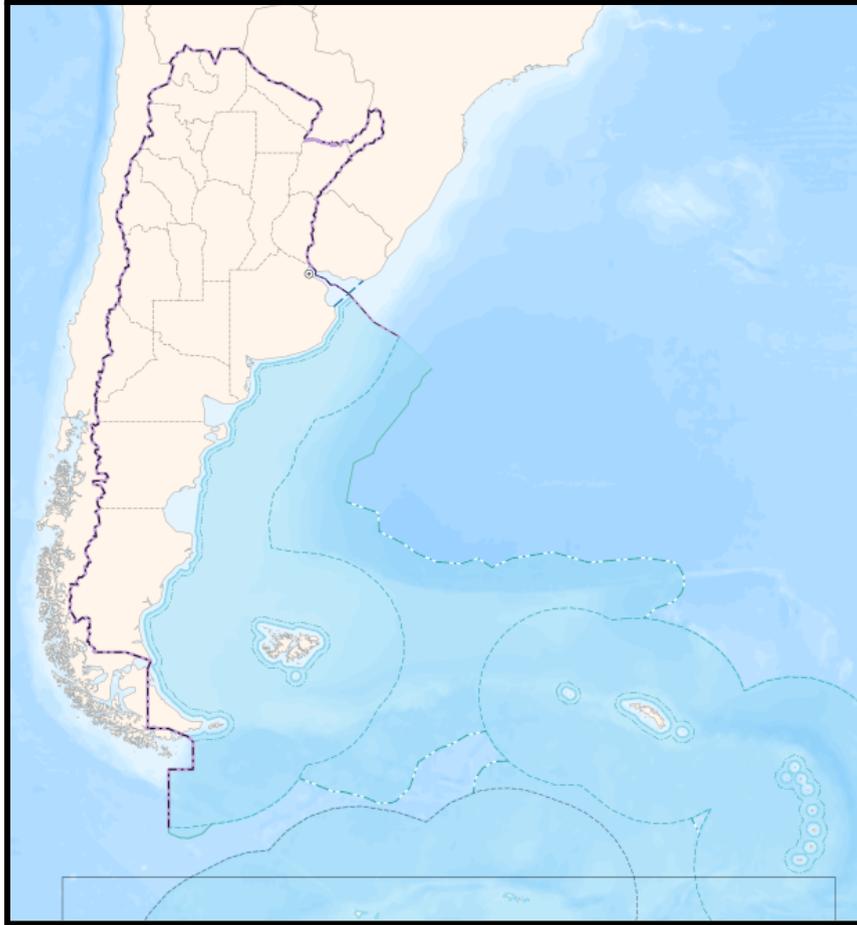


Figura 1: Área de estudio

Para lograr optimizar el tiempo de procesamiento se dividió a la totalidad del área en las 6 regiones de la Argentina: Centro, Patagonia, NOA, NEA, Cuyo e Islas del Atlántico Sur (Figura 2).



Figura 2: División en regiones de la Argentina, de la región patagónica por procesamiento se separó a las islas del atlántico sur

2 - 2. Categorías de cobertura

Las categorías fueron definidas previamente a partir del conocimiento del terreno, información bibliográfica e información geoespacial disponible. Se buscó compatibilizar las clases presentes en el Catálogo de Objetos institucionales [6], que fueron relevadas en cartografía antecedente, con las clases consideradas por organismos nacionales e internacionales. Del análisis de antecedentes se concluyó que los trabajos realizados por el Instituto Nacional de Tecnología Agropecuaria (INTA), la Agencia Espacial Europea (ESA), la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) y el mapa de las unidades de vegetación en la Argentina realizada por Oyarzabal et al. [8] son las más pertinentes a considerar por su similitud con los antes mencionados Objetos del Catálogo [9,10,11]

En base a todo lo antes mencionado, las categorías finales fueron:

1. Corriente de Agua
2. Nieve
3. Bosque, selva
4. Estepa Arbustiva
5. Tierra para cultivo
6. Afloramiento rocoso
7. Humedal
8. Salina

9. Médano, duna
10. Pastizal
11. Monte
12. Planta Urbana*

Es importante remarcar que la planta urbana no fue categorizada a través del algoritmo implementado, sino que se usó una máscara proveniente del shapefile proporcionado por la institución que se realizó a base de fotointerpretación por el departamento de Actualizaciones de Información Geoespacial perteneciente al Instituto Geográfico Nacional.

2 - 3. Imágenes satelitales

Se utilizaron las imágenes satelitales de la misión SENTINEL-2 MSI, con nivel de preprocesamiento 1-C y 2-A con una resolución espacial de 10m y 20m dependiendo la banda utilizada. Del mismo proveedor, se utilizó el producto SENTINEL-1 SAR GRD: escala log, con una resolución de 10m, que brinda imágenes del tipo Radar de Apertura Sintética (SAR), banda C. La colección de imágenes se utiliza directamente desde la plataforma Google Earth Engine, y son provistas por Copernicus Sentinel Data [12]. El periodo analizado fue para el año 2021. En determinadas áreas se trabajó con imágenes Sentinel con nivel de procesamiento 2-A con un filtro de nubosidad específico, dado que esto nos permitió resolver inconvenientes debido a falta de imágenes o clasificaciones inconsistentes en dichas zonas.

Adicionalmente, se utilizó el Modelo Digital de Elevación (DEM por sus siglas en inglés) provisto por el instituto [13] con una resolución de 30m para proveer al clasificador de información auxiliar de altitud.

2 - 4. Metodología

Todo el procesamiento fue realizado en la nube, en la plataforma Google Earth Engine. Una vez seleccionados los productos y periodo de procesamiento, se realizó un filtrado de los metadatos para las imágenes Sentinel 1C que permite obtener pixeles con un porcentaje de nubosidad menor al 30%, ya que la presencia de nubes en imágenes que no son radar tienen impacto en el cálculo de índices y datos de reflectancia. A partir de los pixeles seleccionados, se calculó la media aritmética de los pixeles en el período considerado, tal como se observa en la figura 3 en el uso del algoritmo *ee.Reducer* de la plataforma de Google:

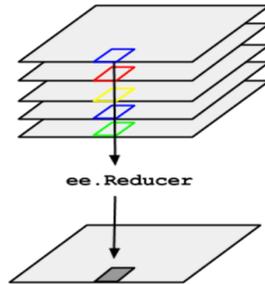


Figura 3: Esquema de procesamiento, donde `ee.Reducer()` calcula la media

Para las imágenes Sentinel 2A que se utilizaron para zonas con poca disponibilidad de imágenes, se utilizó el filtro de probabilidad de nubes, algoritmo propio de Copernicus. Estas imágenes se utilizaron para enmascarar algunas zonas que no se pudieron clasificar con las imágenes Sentinel 1C. El filtro utiliza la banda provista por Copernicus y se procede a enmascarar las áreas con una alta probabilidad de nubes.

Posteriormente se seleccionaron las bandas espectrales a utilizar en la clasificación. Para el producto SENTINEL 2 se utilizaron las siguientes bandas:

- Banda 2: **Azul** (0,49 μm)
- Banda 3: **Verde** (0,56 μm)
- Banda 4 : **Rojo** (0,65 μm)
- Banda 5: **Borde del rojo** (0,70 μm)
- Banda 8: Infrarrojo cercano (**NIR** por sus siglas en inglés) (0,84 μm)
- Banda 11: Infrarrojo onda corta (**SWIR** por sus siglas en inglés) (1,61 μm)

Para las imágenes SAR se utilizaron la banda de co-polarización simple, transmisión vertical/recepción vertical. Esta fue seleccionada debido a su fuerte interacción con volúmenes, ya que la vegetación, especialmente arbórea despolariza la onda, generando una respuesta de mayor intensidad [14].

Los índices calculados fueron seleccionados las categorías definidas previamente:

$$\diamond \text{ NDVI} = (\text{NIR} - \text{Rojo}) / (\text{NIR} + \text{Rojo})$$

El NDVI (*Normalized Difference Vegetation Index*) está altamente asociado con el contenido de vegetación, permitiendo su identificación gracias a su fuerte respuesta en NIR, y absorción en el Rojo [15] [16]. Por lo tanto, los valores altos de este índice estarán asociados a las categorías de bosque y cultivo, mientras que los bajos a zonas de afloramiento rocoso, arenal y estepa.

$$\diamond \text{ NDWI} = (\text{Verde} - \text{NIR}) / (\text{Verde} + \text{NIR})$$

El Índice de Agua de Diferencia Normalizada (NDWI) se utiliza para resaltar el aspecto de las masas de agua en una imagen de satélite. Para ello, se reduce considerablemente la reflectancia del suelo y la vegetación, lo que permite que dichas masas de agua “destaquen” en la imagen [17].

$$\diamond \text{ BSI} = [(SWIR + \text{Rojo}) - (NIR + \text{Azul})] / (SWIR + \text{Verde}) + (NIR + \text{Azul})]$$

El BSI (*Bare Soil Index*) utiliza las bandas de infrarrojos de onda corta y las bandas espectrales rojas para cuantificar la composición mineral del suelo, mientras que las bandas azules y las bandas espectrales de infrarrojos cercanos se utilizan para mejorar la presencia de vegetación [18]. Este índice estará asociado a la detección de las categorías de estepa y afloramiento rocoso principalmente.

$$\diamond \text{ EVI2} = 2,4 * (NIR - \text{Rojo}) / (NIR + \text{Rojo} + 1)$$

Se desarrolló una variación de 2 bandas de EVI (EVI2) para sensores sin banda azul, como MODIS y ASTER. Se ha informado que EVI2 se corresponde bien con el EVI original y se espera que contribuya a un registro de índice de vegetación continuo multisensor. Sin embargo, antes de usar EVI2 para describir los gradientes de la superficie, esta variable debe probarse para determinar su capacidad para representar la diversidad de la cobertura del suelo sin perder información importante sobre el paisaje [19].

$$\diamond \text{ NDSI} = (\text{Verde} - SWIR) / (\text{Verde} + SWIR)$$

El Índice Diferencial Normalizado de Nieve (NDSI) es un indicador numérico que muestra la cobertura de nieve en áreas terrestres. Las bandas espectrales de infrarrojos verdes y de onda corta (SWIR) se utilizan dentro de esta fórmula para trazar el mapa de la cubierta de nieve. Dado que la nieve absorbe la mayor parte de la radiación incidente en el SWIR mientras que las nubes no lo hacen, esto permite a NDSI distinguir la nieve de las nubes. Esta fórmula se utiliza comúnmente en la aplicación de mapas de cobertura de nieve y hielo, así como en el monitoreo de glaciares [20].

$$\diamond \text{ SAVI} = NIR - \text{Rojo} / (NIR + \text{Rojo} + 0,5) * (1 + 0,5)$$

El Índice de Vegetación Ajustado al Suelo (SAVI) se utiliza para corregir el NDVI por la influencia del brillo del suelo en áreas donde la cobertura vegetativa es baja. El SAVI derivado de la reflectancia de la superficie del Landsat se calcula como una relación entre los valores R y NIR con un factor de corrección de la luminosidad del suelo (L) definido en 0.5 para acomodar la mayoría de los tipos de cobertura del suelo [21].

$$\diamond \text{ NDMI} = NIR - SWIR / NIR + SWIR$$

El Índice de Diferencia Normalizada de Humedad (NDMI) se utiliza para determinar el contenido de agua de la vegetación. Se calcula como una relación entre los valores NIR y SWIR de forma tradicional [22].

Para cada uno de estos índices, se calculó la media, el máximo, la mínima y el rango del período analizado. Además, se incorporó el modelo digital de elevación para proveer al clasificador de información sobre el relieve de la provincia. Toda esta información se agrupó en una imagen stack, donde cada banda de los productos, índices e información auxiliar se incorporó como una banda independiente.

2 - 5. Puntos de muestreo

A partir de la fotointerpretación de imágenes satelitales provistas por Google, así también como las imágenes SENTINEL-2 en falso color compuesto, se realizó la elección de un variado número de píxeles para cada una de las categorías, tratando de incorporar la heterogeneidad dentro de las clases como espacialmente dentro del área de estudio. Por otro lado, se utilizó la plataforma Collect Earth, donde se incorporó el shape de cartas topográficas de las distintas regiones y estableciendo un muestreo aleatorio dentro de cada una. Luego se exportó como .csv para ser editado y georeferenciado en QGIS. También se incorporó un muestreo no aleatorio y focalizado en las categorías con mayor representación en el territorio para obtener una mayor equitatividad de las mismas. Las zonas como Cuyo y la provincia de Misiones poseen un mayor número de muestras dado a la incorporación de puntos de muestreo de trabajos previos realizados en esas áreas.

A los efectos de garantizar un mínimo de puntos que resulten en una clasificación robusta y confiable, se identificaron como mínimo $10x$ píxeles, siendo x la cantidad de bandas utilizadas para la clasificación [23]. Se intentó distribuir los puntos de manera aleatoria siempre que sea posible, para reducir el riesgo de autocorrelación [24].

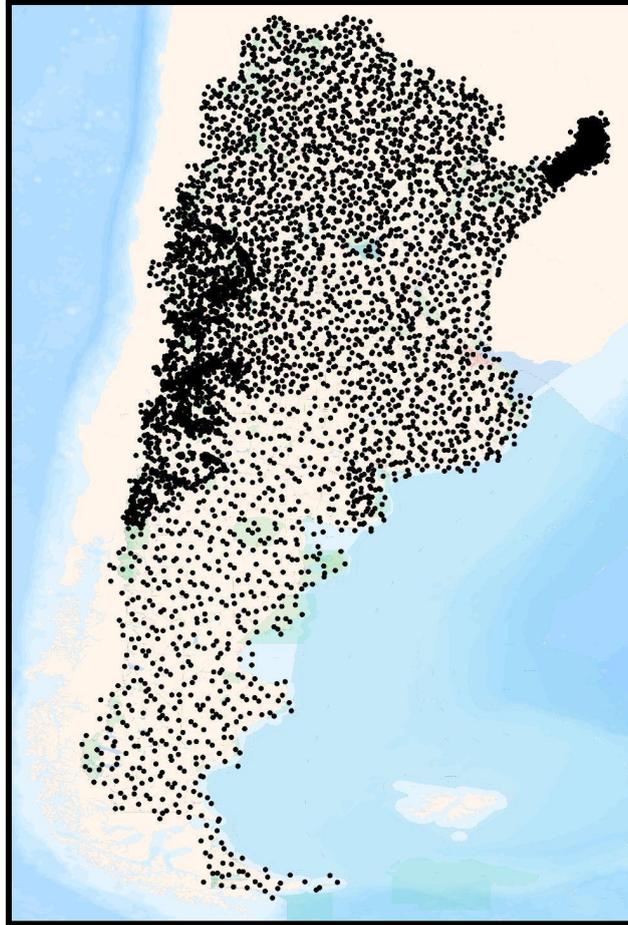


Figura 4: Distribución de los puntos de muestreo en el área de estudio

Del total de puntos muestreados, se utilizó el 70% para entrenamiento del algoritmo, y el 30% para validación. Todas las clases deben tener una buena proporción de datos en el grupo de entrenamiento y validación para subsanar los problemas de la división entrenamiento-validación, utilizada en la validación cruzada. Se pueden hacer sucesivas iteraciones, donde la muestra se divide aleatoriamente, y comparar cómo resultan los parámetros luego de esa serie de repeticiones (es decir, se dividen en 70/30 de manera aleatoria, pero se procesa varias veces el algoritmo, así cada división reparte los datos de manera diferente). La plataforma GEE reparte el 70% y 30% en cada iteración de manera aleatoria.

2 - 6. Algoritmo Random Forest

El Random Forest es un método de clasificación supervisada que ejecuta varios algoritmos de árboles de decisiones. Para clasificar un nuevo objeto basado en atributos (bandas en este caso), cada árbol de decisión da una clasificación y finalmente la decisión con mayor "votos" es la predicción del algoritmo. Uno de los beneficios que más llama la atención es el poder de manejar grandes cantidades de datos con mayor dimensionalidad. Puede manejar miles de variables de entrada e identificar las variables más significativas, por lo que se considera uno de los métodos de reducción de dimensionalidad. Además el modelo muestra la importancia de la variable, que puede ser una característica muy útil a la hora de analizar las bandas del stack que se usó en la clasificación.

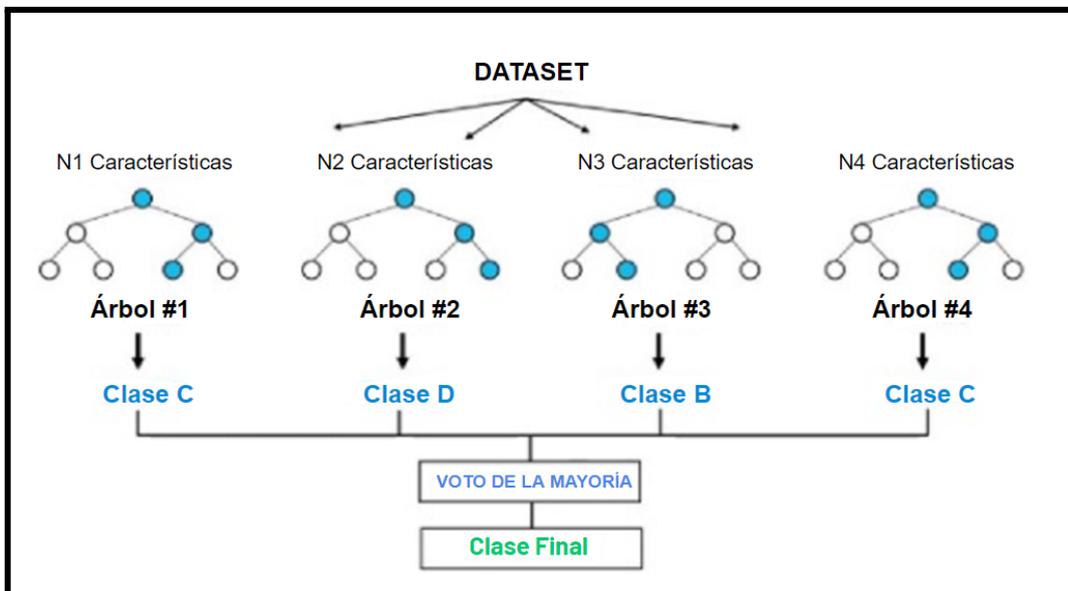


Figura 5: Representación de cómo funciona el algoritmo de clasificación supervisada Random Forest

Se utilizó esta metodología para realizar la clasificación supervisada ya que ofrece una buena respuesta con muestras reducidas, aunque si los puntos de muestreo son escasos puede tener problemas con la identificación de clases de menor peso relativo [25]. Entonces, entre las características deseables del clasificador RF podemos nombrar [26]:

- Es relativamente robusto a valores atípicos y al ruido.
- Brinda información útil sobre estimaciones internas del error, la fuerza, correlación e importancia de variables.
- Es simple y fácilmente paralelizado.

- Es capaz de tomar numerosas variables de entrada y ejecutar clasificaciones rápidas, sin ser severamente afectado por el "curse of dimensionality" [27].

2 - 7. Análisis de la clasificación

Finalmente, con la zona de estudio ya clasificada, se analizó la precisión del modelo sobre los puntos de validación, a partir de los índices de precisión general, precisión del usuario, precisión del productor y el índice Kappa. A su vez, se realizó la matriz de confusión para evidenciar las precisiones del clasificador en cada categoría. Cabe destacar que estos índices de precisión y matriz son propios de cada región.

Clasificación	Observaciones a campo			Suma	Precisión del usuario
	Agua	Vegetación	Suelo desnudo		
Clase 1	245	0	8	253	0,97
Clase 2	23	356	5	384	0,93
Clase 4	0	6	156	162	0,96
Suma	268	362	169	799	
Precisión del productor	91,42	98,34	92,31		0,95

Figura 6: Ejemplo de Matriz de Confusión, donde se observa los distintos tipos de precisión, de errores y las diferencias entre las clases analizadas.

2 - 8. Análisis espacial y temporal

Desde la plataforma Google Earth es posible realizar el cálculo en km² de las distintas categorías dentro de las regiones, por lo que permite un análisis cuantitativo de la evolución de las mismas a lo largo del tiempo, ya que el algoritmo no cambia, solo el insumo de las imágenes SENTINEL para cada año correspondiente. Dado el lanzamiento del producto SENTINEL en 2015, es posible hacer uso de estas imágenes desde este período hasta la actualidad, proveyendo información útil en la detección de cambios.

3. RESULTADOS

Como resultados se obtuvo un producto que representa la cobertura del suelo del territorio nacional para las 12 categorías propuestas anteriormente. El tamaño del píxel es de 30m y se realizaron en las proyecciones EPSG:4326 y

EPSG:3857. Se adecuaron los estilos a las categorías correspondientes y se realizaron los procesos de post-producción (compresión del archivo y creación de pirámides).

Se obtuvieron las matrices y resultados de precisión desde el plug-in de QGIS "Semi-Automatic Classification". Para utilizarlo, se procesaron las muestras de puntos con un buffer de 30 metros, aproximadamente el tamaño de un píxel de clasificación, para transformarlas a polígonos. Luego el plugin toma como inputs la clasificación y las muestras para corroborar la verdad terrestre con la brindada por el clasificador.

3 - 1. Centro

- Precisión general (%) = 80
- Índice Kappa = 0.6428

Matriz de Confusión		Verdad Terrestre											Precisión	
		Corriente de agua	Bosque_selva	Estepa Arbustiva	Cultivo	Afloramiento rocoso	Humedal	Salina	Arenal	Pastizal	Monte	Planta urbana	Total	P. Usuario
C l a s i f i c a c i ó n	Corriente de agua	940	0	0	0	0	2	5	0	51	0	0	998	0,94
	Bosque_selva	3	179	3	8	0	4	0	0	6	23	0	226	0,79
	Estepa Arbustiva	0	3	221	22	10	10	22	2	6	26	0	322	0,69
	Cultivo	67	17	10	1224	0	34	5	0	132	12	0	1501	0,82
	Afloramiento rocoso	0	0	0	0	106	0	0	0	0	0	0	106	1,00
	Humedal	2	15	5	21	3	627	7	0	92	7	0	779	0,80
	Salina	0	0	0	0	0	11	255	0	5	0	0	271	0,84
	Arenal	0	0	0	4	0	0	4	233	0	0	0	241	0,97
	Pastizal	5	11	16	194	3	70	3	12	664	11	0	989	0,67
	Monte	0	56	26	4	0	7	0	0	4	202	0	299	0,68
	Planta urbana	54	7	0	7	0	0	0	11	8	0	0	87	0,00
Total	1071	288	281	1484	122	765	301	258	968	281	0	5819	-	
Precisión														
P. Productor	0,88	0,62	0,79	0,82	0,87	0,82	0,85	0,90	0,69	0,72	0,00	-	0,80	

3 - 2. Cuyo

- Precisión general (%) = 75
- Índice Kappa = 0.5536

Matriz de Confusión		Verdad terrestre											Precisión		
		Corriente de agua	Nieve	Bosque_selva	Estepa arbustiva	Tierra para cultivo	Afloramiento rocoso	Humedal	Salina	Médano, duna	Pastizal	Monte	Planta Urbana*	Total	P. Usuario
C l a s i f i c a c i ó n	Corriente de agua	325	0	0	0	0	0	0	0	0	88	0	0	413	0,79
	Nieve	0	554	0	0	0	20	0	0	0	0	0	0	574	0,97
	Bosque_selva	0	0	350	46	0	21	22	0	0	3	20	0	482	0,76
	Estepa arbustiva	66	3	139	2778	147	250	274	92	696	23	100	0	4568	0,61
	Tierra para cultivo	0	0	2	8	429	0	0	0	0	0	3	0	442	0,97
	Afloramiento rocoso	33	64	15	63	0	832	21	0	2	0	0	0	1030	0,81
	Humedal	5	0	7	3	14	0	389	27	0	1	0	0	446	0,87
	Salina	0	0	0	15	0	0	16	695	0	0	0	0	726	0,96
	Médano, duna	0	0	0	17	0	0	0	2	682	0	0	0	701	0,97
	Pastizal	3	0	0	0	0	0	0	0	0	47	0	0	50	0,94
	Monte	0	0	21	40	0	0	0	0	0	0	277	0	338	0,82
Planta Urbana*	0	0	3	4	9	0	0	0	0	0	0	0	16	0,00	
Total	432	621	537	2974	599	1123	722	816	1380	162	400	0	9766	-	
Precisión															
P. Productor	0,75	0,89	0,65	0,93	0,72	0,74	0,54	0,85	0,49	0,29	0,69	0,00	0,75		

3 - 3. NOA

- Precisión general (%) = 77
- Índice Kappa = 0.6323

Matriz de Confusión		Verdad terrestre											Precisión		
		Corriente de agua	Nieve	Bosque, selva	Estepa arbustiva	Tierra para cultivo	Afloramiento rocoso	Humedal	Salina	Médano, duna	Pastizal	Monte	Planta Urbana*	Total	P. Usuario
C l a s i f i c a c i ó n	Corriente de agua	992	0	0	2	0	0	9	0	0	14	0	0	1017	0,98
	Nieve	0	121	0	0	0	0	0	0	0	0	0	0	121	0,99
	Bosque, selva	272	0	915	78	26	4	37	0	0	8	79	0	1419	0,64
	Estepa arbustiva	0	3	5	416	7	50	60	8	23	0	9	0	581	0,72
	Tierra para cultivo	8	0	43	41	323	13	17	0	0	10	10	0	465	0,69
	Afloramiento rocoso	0	10	6	111	0	494	42	10	13	4	4	0	694	0,71
	Humedal	0	0	0	13	12	0	160	6	0	4	0	0	195	0,82
	Salina	4	0	0	3	0	0	2	233	2	0	0	0	244	0,95
	Médano, duna	0	0	0	0	0	0	0	0	8	0	0	0	8	1,00
	Pastizal	0	0	0	0	0	0	0	0	0	16	0	0	16	1,00
	Monte	0	0	18	14	3	0	4	0	0	0	195	0	234	0,83
	Planta Urbana*	3	0	3	0	0	0	0	0	0	0	0	0	6	0,00
Precisión	Total	1279	134	990	678	371	569	331	262	46	56	297	0	5013	-
P.Productor		0,78	0,90	0,92	0,61	0,87	0,87	0,48	0,89	0,17	0,29	0,66	0,00	-	0,77

3 - 4. Patagonia

- Precisión general (%) = 85
- Índice Kappa = 0.5256

Matriz de confusión		Verdad terrestre											Precisión		
		Corriente de agua	Nieve	Bosque, selva	Estepa arbustiva	Tierra para cultivo	Afloramiento rocoso	Humedal	Salina	Médano, duna	Pastizal	Monte	Planta Urbana*	Total	P. Usuario
C l a s i f i c a c i ó n	Corriente de agua	2324	0	0	0	0	0	0	0	0	96	0	0	2428	0,96
	Nieve	0	190	2	3	0	0	23	7	0	0	0	0	225	0,84
	Bosque, selva	1	0	769	9	21	3	12	0	0	142	10	0	967	0,80
	Estepa arbustiva	24	0	48	1673	47	161	56	164	3	61	37	0	2274	0,74
	Tierra para cultivo	0	0	0	2	302	0	6	0	0	27	2	0	339	0,89
	Afloramiento rocoso	0	0	1	8	0	215	0	6	0	0	0	0	230	0,93
	Humedal	0	0	6	14	21	0	322	0	0	32	0	0	395	0,82
	Salina	34	0	0	3	0	0	0	485	2	2	0	0	526	0,92
	Médano, duna	0	0	0	0	0	0	0	2	8	0	0	0	329	0,98
	Pastizal	22	0	4	10	0	2	8	0	0	178	0	0	224	0,78
	Monte	0	0	0	0	3	0	0	0	0	0	31	0	34	0,91
	Planta Urbana*	0	0	0	0	4	0	0	0	0	0	0	0	4	0,00
Precisión	Total	2405	190	830	1722	398	415	411	657	329	538	80	0	7975	-
P.Productor		0,97	1,00	0,93	0,97	0,76	0,52	0,78	0,74	0,88	0,33	0,39	0,00	-	0,85

3 - 5. NEA

- Precisión general (%) = 89
- Índice Kappa = 0.6411

Matriz de Confusión		Verdad terrestre								Precisión	
		Corriente de agua	Bosque, selva	Estepa arbustiva	Tierra para cultivo	Humedal	Salina	Pastizal	Monte	Total	P. Usuario
C l a s i f i c a c i ó n	Corriente de agua	522	1	0	0	0	0	47	0	1017	0,51
	Bosque, selva	16	2703	8	119	61	0	308	43	1419	1,90
	Estepa arbustiva	0	0	0	0	0	0	0	0	581	0,00
	Tierra para cultivo	0	15	6	187	12	0	14	7	465	0,40
	Humedal	13	46	11	11	346	0	49	15	195	1,77
	Salina	0	0	0	0	0	0	0	0	244	0,00
	Pastizal	0	184	0	182	50	0	591	41	16	36,94
	Monte	0	2	38	3	6	0	11	122	234	0,52
Precisión	Total	1279	990	678	371	331	262	56	297	5013	-
P.Productor		0,41	2,73	0,00	0,50	1,05	0,00	10,55	0,41	-	0,89

3 - 6. Atlántica

- Precisión general (%) = 97
- Índice Kappa = 0.8623

Matriz de Confusión		Verdad terrestre									Precisión	
		Corriente de agua	Bosque, selva	Estepa arbustiva	Tierra para cultivo	Humedal	Salina	Pastizal	Monte	Total	P. Usuario	
C l a s i f i c a c i ó n	Corriente de agua	522	1	0	0	0	0	47	0	1017	0,51	
	Bosque, selva	16	2703	8	119	61	0	308	43	1419	1,90	
	Estepa arbustiva	0	0	0	0	0	0	0	0	581	0,00	
	Tierra para cultivo	0	15	6	187	12	0	14	7	465	0,40	
	Humedal	13	46	11	11	346	0	49	15	195	1,77	
	Salina	0	0	0	0	0	0	0	0	244	0,00	
	Pastizal	0	184	0	182	50	0	591	41	16	36,94	
	Monte	0	2	38	3	6	0	11	122	234	0,52	
Precisión	Total	1279	990	678	371	331	262	56	297	5013	-	
P. Productor		0,41	2,73	0,00	0,50	1,05	0,00	10,55	0,41	-	0,89	

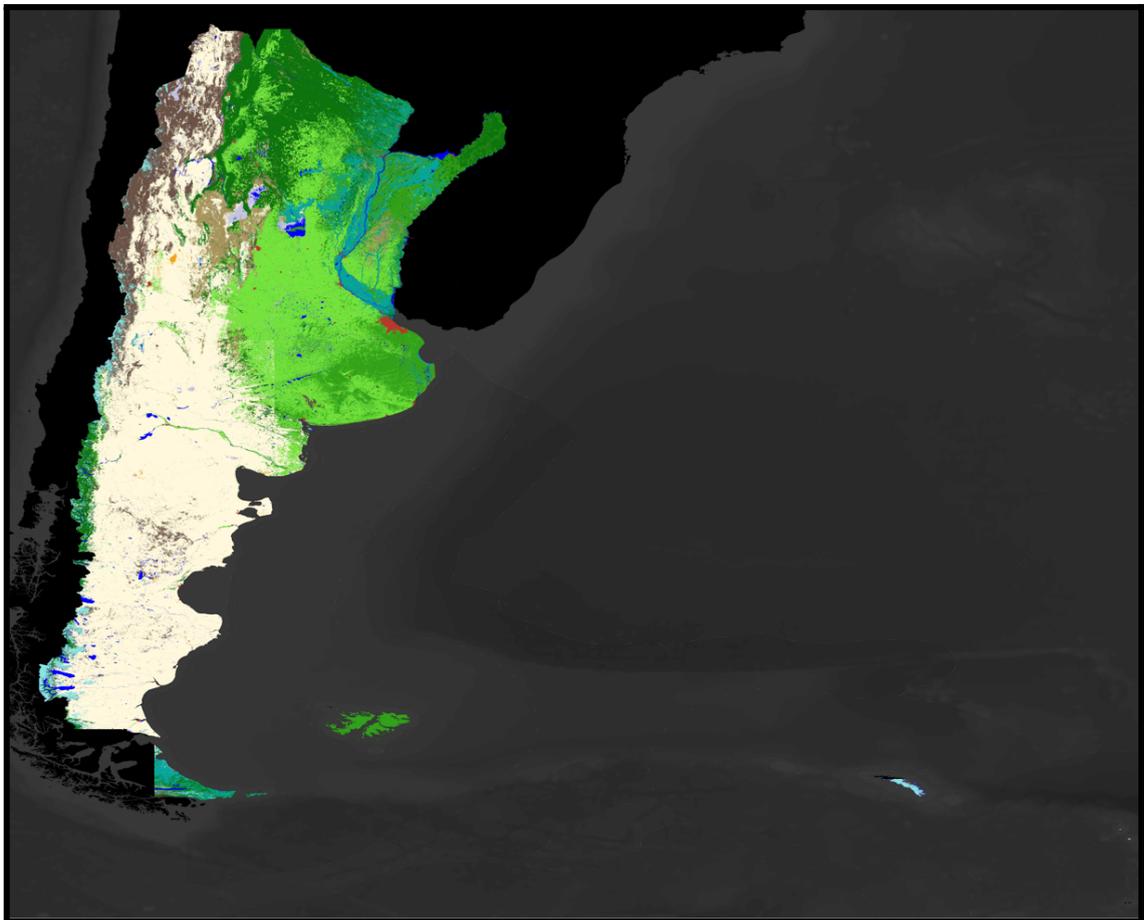


Figura 7: Resultado final de cobertura del suelo de la República Argentina

4. CONCLUSIONES

La detección de cobertura de suelo para la República Argentina constituye un

insumo nacional de calidad para el análisis temporal y espacial de las mismas. El contar con un producto nacional y de esta escala, brinda la posibilidad de generar una inmensa cantidad de análisis o subproductos posteriores tales como ordenamientos territoriales, gestión de riesgos, gestión ambiental, entre muchos otros.

Durante el trabajo se detectaron ciertas oportunidades de mejora, tales como, mejorar el tamaño de píxel para trabajar en menores escalas, la toma de muestras para la clasificación, sub-categorías en las clases y la unión entre regiones para que la continuidad sea más fluida.

5. REFERENCIAS

- [1] S. Talukdar et al., "Land-use land-cover classification by machine learning classifiers for satellite observations-A review," *Remote Sens.*, vol. 12, no. 7, 2020, doi: 10.3390/rs12071135.
- [2] Y. A. Liou, A. K. Nguyen, and M. H. Li, "Assessing spatiotemporal eco-environmental vulnerability by Landsat data," *Ecol. Indic.*, vol. 80, no. May, pp. 52–65, 2017, doi: 10.1016/j.ecolind.2017.04.055.
- [3] K. A. Nguyen and Y. A. Liou, "Global mapping of eco-environmental vulnerability from human and nature disturbances," *Sci. Total Environ.*, vol. 664, pp. 995–1004, 2019, doi: 10.1016/j.scitotenv.2019.01.407.
- [4] UN-GGIM, UN-GGIM: Américas. Estatutos, Comité Regional de Naciones Unidas Sobre Gestión Global de la Información Geoespacial para las Américas (UN-GGIM: Américas). 2019, pp. 1–12.
- [5] Instituto Geográfico Nacional, "DETERMINACIÓN DE LA SUPERFICIE CORRESPONDIENTE AL TERRITORIO CONTINENTAL, ANTÁRTICO E INSULAR DE LA REPÚBLICA ARGENTINA," 2022. [Online]. Available: <https://www.ign.gob.ar/NuestrasActividades/InformacionGeoespacial/UnidadesTerritoriales/DocumentacionTecnica/>.
- [6] Instituto Geográfico Nacional, "Descripción del Diccionario de Conceptos y Catálogo de Objetos Geográficos del Instituto Geográfico Nacional República Argentina," 2017.
- [7] Waller, Tomas & Micucci, Patricio & Draque, Juan & Barros, Mariano & Estavillo, Candelaria. (2010). CONSERVACIÓN DE LA BOA AMPALAGUA (*Boa constrictor occidentalis*) EN LA REPÚBLICA ARGENTINA.
- [8] Oyarzabal, M., Clavijo, J., Oakley, L., Biganzoli, F., Tognetti, P., Barberis, I., Maturo, H. M., Aragón, R., Campanello, P. I., Prado, D., Oesterheld, M., & León, R. J. (2018). Unidades de vegetación de la Argentina. *Ecología Austral*, 28(1), 040–063. <https://doi.org/10.25260/EA.18.28.1.0.399>
- [9] Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsensbazar, N.E., Xu, P., Ramoino, F., Arino, O., 2022. ESA WorldCover 10 m 2021 v200. <https://doi.org/10.5281/zenodo.7254221>
- [10] LAND COVER CLASSIFICATION SYSTEM. (s/f). Fao.org. Recuperado el 23 de febrero de 2023, de <https://www.fao.org/3/x0596e/x0596e00.htm>
- [11] Volante, José & Collado, Alfredo & Ferreyra, Eduardo & López, Carlos &

- Navarro Rau, Maria & Pezzola, Alejandro & Puentes, María & Olmedo, Guillermo & al., (2009). Monitoreo de la cobertura y el uso del suelo a partir de sensores remotos.
- [12] Ruiz-Ramos, J., Marino, A., Boardman, C., & Suarez, J. (2020). Continuous Forest Monitoring Using Cumulative Sums of Sentinel-1 Timeseries. *Remote Sensing*, 12(18), 3061. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/rs12183061>
- [13] MDE-Ar | Instituto Geográfico Nacional. (s. f.). <https://www.ign.gov.ar/category/tem%C3%A1tica/geodesia/mde-ar>
- [14] M. Shimada, T. Itoh, T. Motooka, M. Watanabe, and T. Shiraishi, "New global forest / non-forest maps from ALOS PALSAR data (2007 – 2010) Remote Sensing of Environment New global forest / non-forest maps from ALOS PALSAR data (2007 – 2010)," *Remote Sens. Environ.*, no. February 2018, 2014, doi: 10.1016/j.rse.2014.04.014.
- [15] L. Pasolli et al., "Estimation of soil moisture in mountain areas using SVR technique applied to multiscale active radar images at C-band," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 8, no. 1, pp. 262–283, 2015, doi: 10.1109/JSTARS.2014.2378795.
- [16] K. J. Raj and S. Sivasathya, "SVM and Random Forest Classification of Satellite Image with NDVI as an Additional Attribute to the Dataset," doi: 10.1007/978-81-322-1771-8.
- [17] S. K. McFEETERS (1996) The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features, *International Journal of Remote Sensing*, 17:7, 1425-1432, DOI: 10.1080/01431169608948714
- [18] M. Castelo-cabay, J. A. Piedra-fernandez, and R. Ayala, "Deep learning for land use and land cover classification from the Ecuadorian Paramo .," 2022, doi: 10.1080/17538947.2022.2088872.
- [19] Mondal, P. (2011). Quantifying surface gradients with a 2-band Enhanced Vegetation Index (EVI2). *Ecological Indicators*, 11(3), 918–924. <https://doi.org/10.1016/j.ecolind.2010.10.006>
- [20] Raster calculator. (s/f). [Blumarblegeo.com](https://www.blumarblegeo.com). Recuperado el 23 de febrero de 2023, de https://www.blumarblegeo.com/knowledgebase/global-mapper-19/Raster_Calculator.htm
- [21] Landsat Soil Adjusted Vegetation Index. (s/f). [Usgs.gov](https://www.usgs.gov). Recuperado el 23 de febrero de 2023, de <https://www.usgs.gov/landsat-missions/landsat-soil-adjusted-vegetation-index>
- [22] Normalized Difference Moisture Index. (s/f). [Usgs.gov](https://www.usgs.gov) recuperado el 28 de febrero de 2023, de [Normalized Difference Moisture Index | U.S. Geological Survey \(usgs.gov\)](https://www.usgs.gov/landsat-missions/landsat-soil-adjusted-vegetation-index)
- [23] D. A. Griffith, "Effective Geographic Sample Size in the Presence of Spatial Autocorrelation," vol. 95, no. December 2004, pp. 740–760, 2005.
- [24] L. C. L. High-resolution, C. A. Ramezan, T. A. Warner, A. E. Maxwell, and B. S. Price, "Effects of Training Set Size on Supervised Machine-Learning Remotely Sensed Data," *Remote Sens.*, pp. 1–27, 2021.
- [25] J. Bialas, T. Oommen, and T. C. Havens, "Int J Appl Earth Obs Geoinformation Optimal segmentation of high spatial resolution images for the classification of buildings using random forests," *Int J Appl Earth Obs Geoinf.*, vol. 82, no. June, p.

- 101895, 2019, doi: 10.1016/j.jag.2019.06.005.
- [26] Breiman, L. (2001). Random Forest. *Machine learning*, 45(1), 5–32. <https://doi.org/10.1023/a:1010933404324>
- [27] Rodriguez-Galiano, V. F., Chica-Olmo, M., Abarca-Hernandez, F., Atkinson, P. M., & Jeganathan, C. (2012). Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sensing of Environment*, 121, 93–107. <https://doi.org/10.1016/j.rse.2011.12.003>