

# Diseño de muestreo para la detección de Árboles Nativos y Otras Formaciones Boscosas fuera del Bosque Nativo.

---

María Eugenia Riaño

Instituto de Estadística, FCEA-UdelaR,  
Proyecto REDD+ Uruguay, MGAP- MVOTMA  
25 de noviembre de 2020



# ¿Qué es REDD+?

REDD+ es un marco respaldado por las Naciones Unidas que apunta a frenar el cambio climático al detener la destrucción de los bosques.

REDD significa “Reducción de Emisiones derivadas de la Deforestación y la Degradación de los bosques”; el símbolo + implica que en su implementación hay componentes de conservación, gestión sostenible de los bosques con participación de población local y aumento de las reservas forestales de carbono.

Equipo de trabajo: Lucía Bernardi, Juan Olivera, Ma. Laura García De Souza, Alejandra Boccardo y Diego Martino.

# Introducción

De acuerdo a la Ley Forestal N° 15.939:

*“Son bosques las asociaciones vegetales en las que predomina el arbolado de cualquier tamaño, explotado o no, y que estén en condiciones de producir madera u otros productos forestales o de ejercer alguna influencia en la conservación del suelo, en el régimen hidrológico o en el clima, o que proporcionen abrigo u otros beneficios de interés nacional”.*

Figura: Bosque Nativo Fluvial



Figura: Bosque Nativo de Quebrada



Figura: Bosque Nativo Serrano



Figura: Bosque Nativo de Parque



## ¿Qué son las OFBN y los ANFB?

Se utiliza el término OFBN para referirse al tipo de formación vegetal que puede estar incluida en el marco de la Ley Forestal N° 15.939, pero que no se encuentra contemplada dentro de la reglamentación e instructivos vigentes. A partir de estos últimos se excluye las formaciones boscosas de baja densidad o parches menores a 2500 m<sup>2</sup>.

Los ANFB se refiere al arbolado que crece de manera aislada en el paisaje de manera espontánea o plantada.



Figura: Otras Formaciones Boscosas Nativas - Sierras del Este



Figura: Otras Formaciones Boscosas Nativas - Cuenca Sedimentaria del Oeste



Figura: Árboles Nativos Fuera del Bosque - Cuesta Basáltica

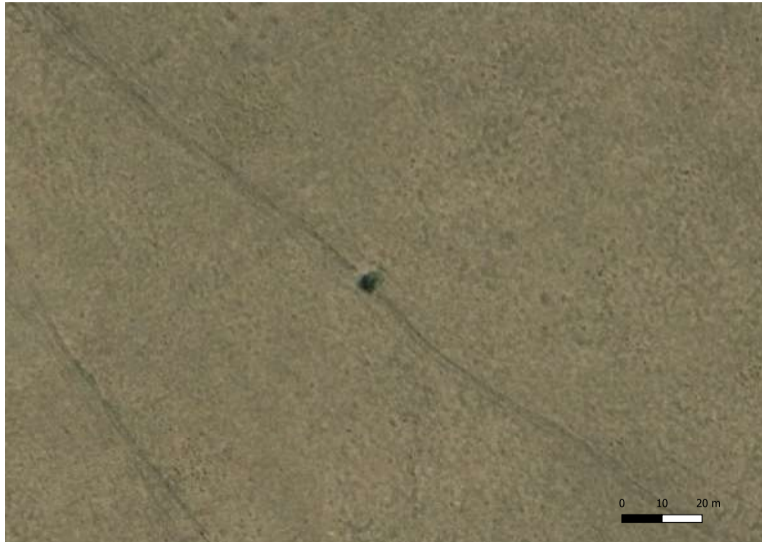
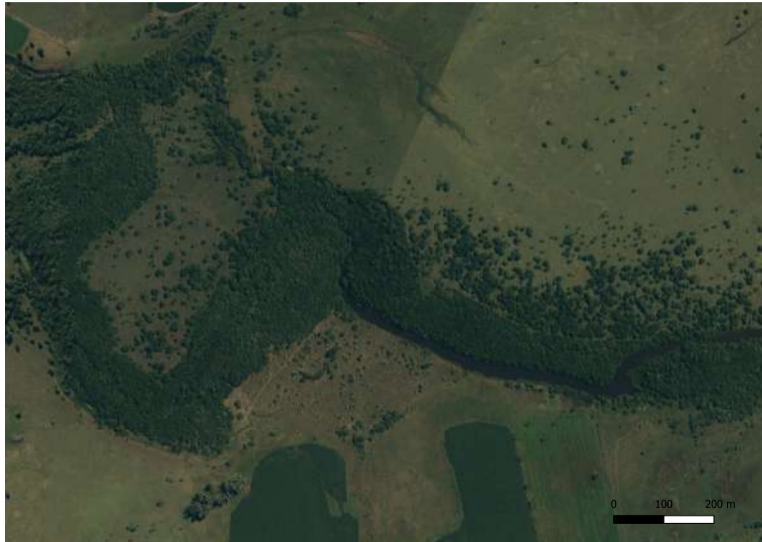


Figura: Gradiente de densidad de Bosque Nativo - Graven de Santa Lucía



## Objetivo

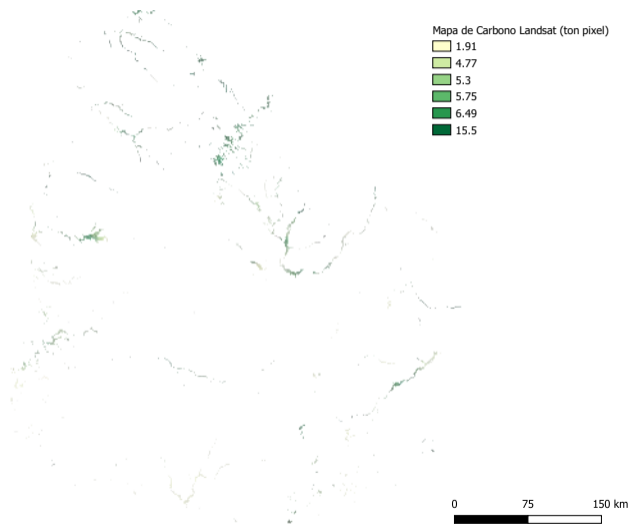
El objetivo principal es estimar el stock de carbono correspondiente a las áreas de ANFB y a OFBN.

El stock de carbono total en la biomasa viva del Bosque nativo se estima como:

**STOCK CARTOGRAFÍA + STOCK ÁREA OMITIDA - STOCK ÁREA DE SOBRECObERTURA**

- **STOCK CARTOGRAFÍA:** es el stock que se obtiene a partir del Mapa de Carbono (Riaño, Penengo y Bernardi, 2020).
- **STOCK DEL ÁREA OMITIDA:** corresponde al stock estimado para el área de omisión de la cartografía del Bosque Nativo.
- **STOCK DEL ÁREA DE SOBRECObERTURA:** corresponde al stock estimado para el área de sobrecobertura de la cartografía del Bosque Nativo.

Figura: Mapa de Carbono - LANDSAT 8



- STOCK CARTOGRAFÍA: 52.700.740 ton (error 3.7 %)
- STOCK DEL ÁREA OMITIDA: 14.339.385 (error 32 %)
- STOCK DEL ÁREA DE SOBRECUBERTURA: 5.741.025 (error 22.6 %)

STOCK TOTAL DE CARBONO: 61.299.100 ton (error 10.7 %)

# ¿Cuánto carbono hay en ANFB Y OFBN?



## ¿Qué diseño utilizar?

Los diseños muestrales utilizados en *sensoramiento remoto* son por lo general diseños Estratificados Simples o Sistemáticos. Son problemáticos si las clases se encuentran muy desbalanceadas.

### Ejemplo: Validación Cartográfica (Olofsson et al. 2014)

- Se estiman las áreas de Omisión y de Sobrecobertura de una Cartografía.
- Los estratos son las clases, “Bosque” y “No Bosque”.

## Bosque Nativo de Uruguay

993.600 ha. (error 5.7 %) en base a datos de Collect Earth.

5.6 % del territorio nacional (17.621.500 ha)

Se presume que el área de los ANFB y OFBN es menor que la del Bosque Nativo.



# ¿Cómo se construye la cartografía de Bosque Nativo?



# Clasificación Supervisada

En la construcción de la cartografía de Bosque Nativo (Bernardi et al., 2019) se aplica un modelo CART (Breiman et al., 1984) utilizando la plataforma Google Earth Engine (Gorelick et al., 2017) considerando las siguientes variables como explicativas:

## ■ Bandas Landsat/Sentinel

- Azul.
- Verde.
- Rojo.
- Infrarrojo Cercano (SWIR 1).
- Infrarrojo Medio (SWIR 2).

## ■ Índices de Vegetación

- EVI: *Enhanced Vegetation Index*.
- NDWI: *Normalized Difference Water Index*.
- MSAVI: *Modified soil-adjusted vegetation Index*.

## Diseño propuesto

1. Para cada pixel  $i$  se define la variable de respuesta  $y$  tal que:

$$y_i = \begin{cases} 1 & \text{si pertenece a la cartografía de BN} \\ 0 & \text{si no pertenece a la cartografía de BN} \end{cases}$$

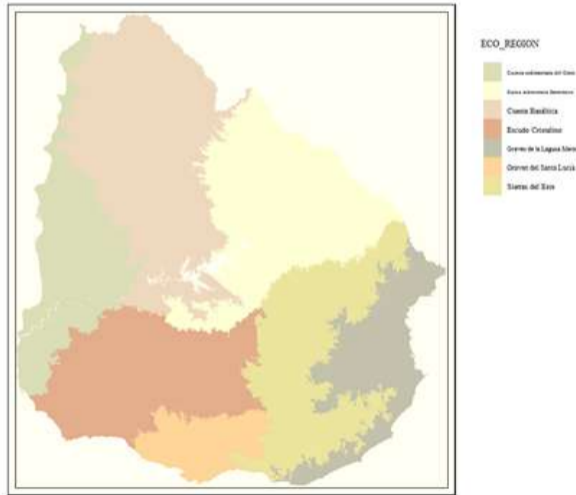
2. Utilizando las mismas variables explicativas y el mismo modelo para la clasificación se obtienen las **probabilidades de pertenecer a la cartografía de BN**.

$$\hat{p}_i = f(x_1, \dots, x_k)$$

# Diseño propuesto

3. Se estratifica de acuerdo a las probabilidades predichas de pertenecer a la cartografía BN en estratos de “alta” y “baja” probabilidad (en este caso además se estratifica por Eco-región (Brazeiro et al., 2012)).
4. Se selecciona una muestra aleatoria simple (puede ser también sistemática) en cada estrato asignando un mayor número de casos al estrato de alta probabilidad.

# Eco-Regiones del Uruguay





## Área de BN por Eco-Región (en ha)

Superficie de BN por Eco-región en ha		
Eco-Región	Área	% Superficie
Cuenca Sed.Oeste	141.635	6.2 %
Cuenca Sed. Gondwánica	157.009	5.3 %
Cuesta Basáltica	160.123	3.3 %
Escudo Cristalino	67.949	2.7 %
Graven Laguna Merín	60.989	3.8 %
Graven Santa Lucía	34.795	3.8 %
Sierras del Este	202.738	7.8 %

# Estimación del modelo CART

Se selecciona una muestra de entrenamiento de píxeles por Eco- Región y para cada una de ellas se estima un modelo CART. Se utiliza la librería *caret* (Kuhn, 2020) de R (R Core Team, 2017).

Los tamaños de las muestras de entrenamiento son de 16.000 casos para la categoría No Bosque y dependiendo de la Eco- Región, para la clase Bosque varía entre 4.000 y 10.000.

# Evaluación del desempeño: matriz de confusión

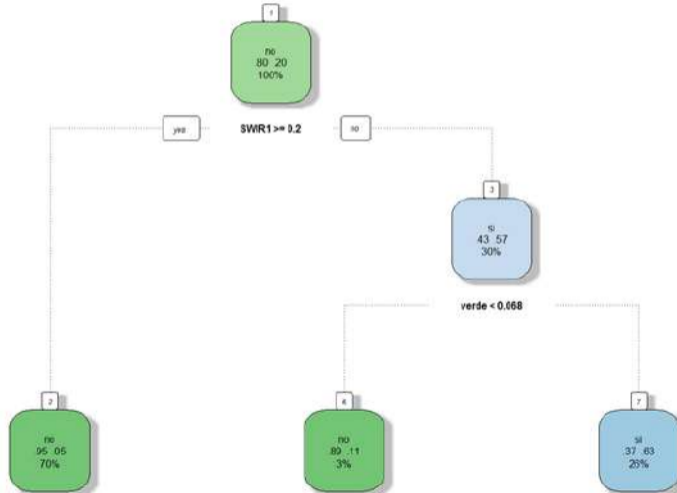
		Observado		
		Positivo	Negativo	
Predicción	Positivo	TP	FP	↔ VPP
	Negativo	FN	TN	↔ VPN

↕                      ↕  
Sensibilidad    Especificidad

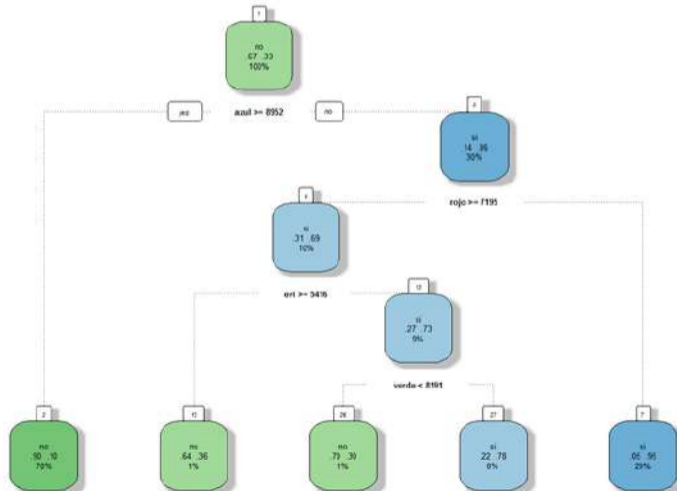
# Evaluación del desempeño

Indicadores de Desempeño por Eco- Región				
Eco-Región	Especificidad	Sensibilidad	VPN	VPP
Cuenca Sed.Oeste	87.7 %	79.3 %	89.6 %	76.2 %
Cuenca Sed. Gondwánica	90.8 %	84.7 %	92.3 %	82 %
Cuesta Basáltica	94.9 %	89.5 %	94.8 %	89.7 %
Escudo Cristalino	94.6 %	77.3 %	89.4 %	87.7 %
Graven Laguna Merín	92.8 %	81.3 %	90.7 %	85.2 %
Graven Santa Lucía	96 %	76.7 %	89.2 %	90.7 %
Sierras del Este	86.6 %	82.9 %	95.2 %	61 %

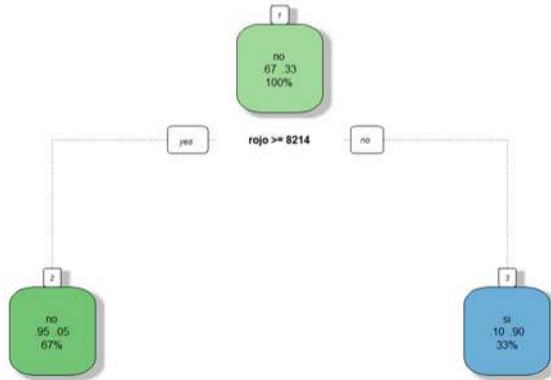
# Árbol de Clasificación para las Sierras del Este



# Árbol de Clasificación para el Graven de Santa Lucía



# Árbol de Clasificación para la Cuesta Basáltica



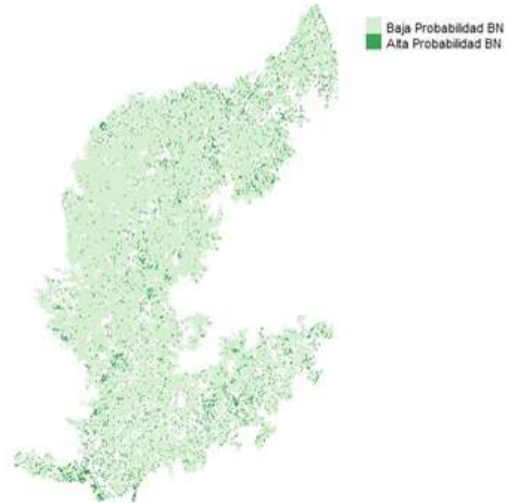
## Puntos de corte de los estratos de Alta y Baja probabilidad por Eco-Región

<b>Eco-Región</b>	<b>Punto de Corte</b>
Cuenca Sed. Oeste	0.77
Cuenca Sed. Gondwánica	0.82
Cuesta Basáltica	0.9
Escudo Cristalino	0.7
Graven Laguna Merín	0.84
Graven Santa Lucía	0.78
Sierras del Este	0.63

Se eliminan del marco las cartografías de BN y de Plantaciones Forestales.



# Marco - Estratos Sierras del Este



## Total de píxeles y peso relativo de los estratos

Eco-Región	Estrato Probabilidad	Píxeles	%
Cuenca Sed. Oeste	Alta	3,239,427	14.7
	Baja	18,758,412	85.3
Cuenca Sed. Gondwánica	Alta	2,363,134	7.7
	Baja	28,371,238	92.3
Cuesta Basáltica	Alta	1,497,551	3.4
	Baja	42,598,980	96.6
Escudo Cristalino	Alta	1,352,960	4.7
	Baja	27,750,862	95.3
Graven Laguna Merín	Alta	965,676	3.9
	Baja	19,885,721	96.1
Graven Santa Lucía	Alta	350,090	4.6
	Baja	8,680,424	95.4
Sierras del Este	Alta	2,353,747	8.2
	Baia	26.257.634	91.8

# Tamaño de muestra

El tamaño de muestra se estableció a partir de una muestra piloto en las Sierras del Este, para el estrato de Alta Probabilidad.

Se obtiene que con 600 casos el error de muestreo para las superficies de ANFB y OFBN es de un 20 %.

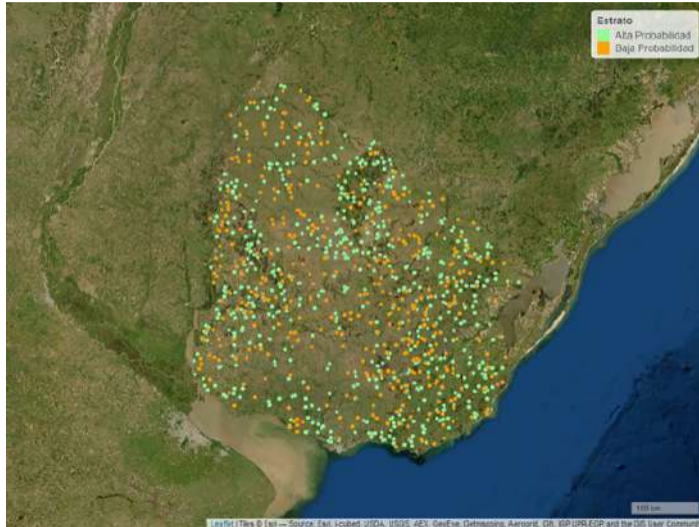
Para “completar” la muestra, se agregan 400 casos al estrato de Baja Probabilidad.

La asignación de la muestra entre Eco-regiones se realiza en forma proporcional a la superficie de BN en la región.

## Tamaños de muestra por estrato

Eco-Región	Estrato Probabilidad	Casos
Cuenca Sed. Oeste	Alta	104
	Baja	68
Cuenca Sed. Gondwánica	Alta	115
	Baja	76
Cuesta Basáltica	Alta	117
	Baja	78
Escudo Cristalino	Alta	50
	Baja	33
Graven Laguna Merín	Alta	44
	Baja	28
Graven Santa Lucía	Alta	26
	Baja	17
Sierras del Este	Alta	148
	Baja	98

# Puntos seleccionados



# Relevamiento

El relevamiento se realiza con las imágenes del IDE (Infraestructura de Datos Espaciales).

Se deben tener en cuenta los segmentos “base” de la cartografía para determinar si son casos de omisión.

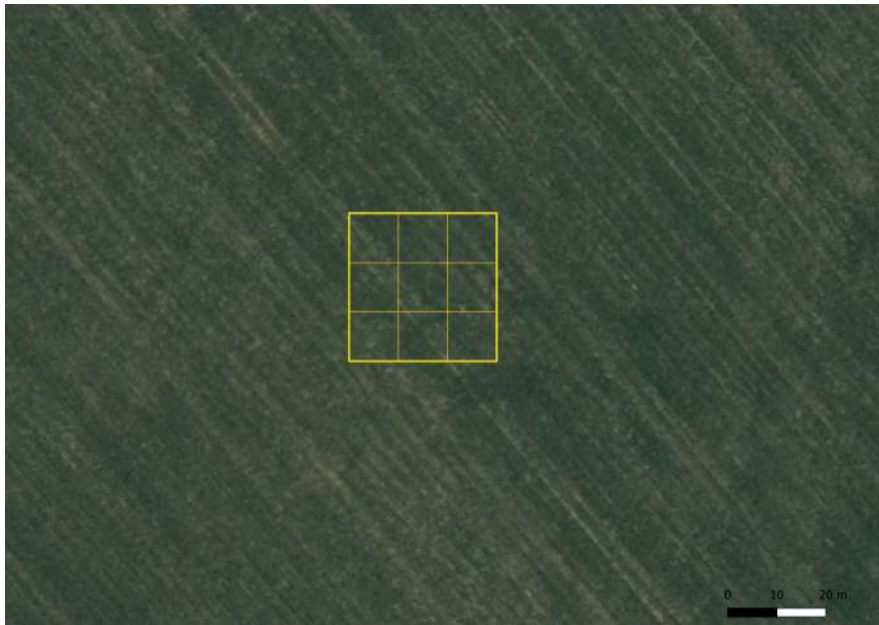
Para no duplicar en la estimación de carbono los ANFB y OFBN sólo se buscan en los segmentos no omitidos de BN.

## Etapas de relevamiento

1) Identificar los píxeles que no corresponden a la clase de interés.

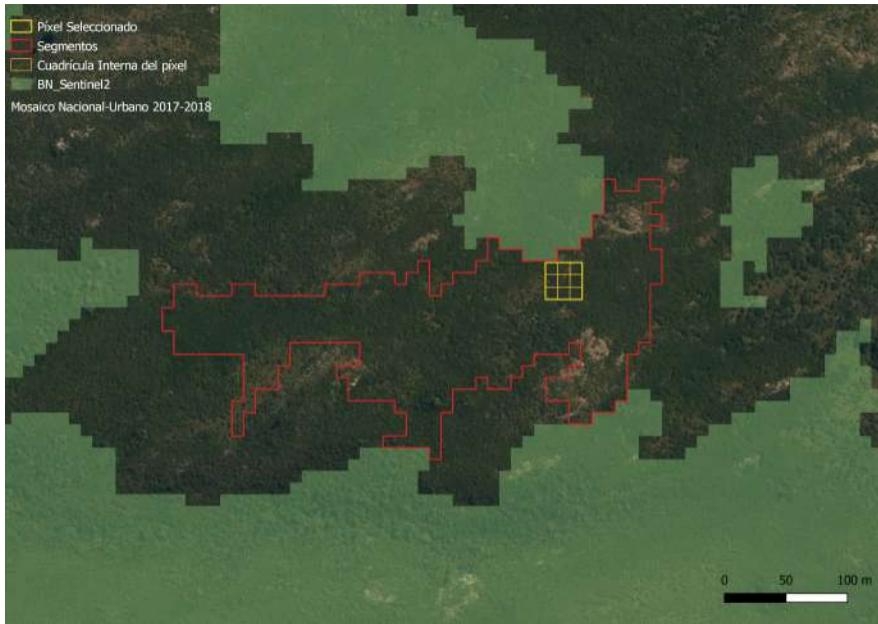






## Etapas de relevamiento

2) Clasificar los segmentos que intersectan con los píxeles no descartados en el paso anterior de acuerdo a la omisión.





## Etapas de relevamiento

3) Relevar la presencia de ANFB y OFBN en los píxeles que se encuentran en segmentos no omitidos de la cartografía.



# Resultados

		Estrato		
		Alta Probabilidad	Baja Probabilidad	Total
Clase	Bosque	238	50	288
	No Bosque	366	348	714
	Total	604	398	1002

Cuadro: Clasificación de los píxeles de acuerdo a la clase y a la estratificación.

$$\text{Bosque} = \text{ANFB} + \text{OFBN} + \text{BN}$$

# Resultados

- Un 82 % de la clase “Bosque” se encuentra en el estrato de Alta Probabilidad.
- En el estrato de Alta Probabilidad un 39 % es Bosque, mientras que para el estrato de Baja Probabilidad un 12 % corresponde a la clase Bosque



# Resultados

		Estrato		
		Alta Probabilidad	Baja Probabilidad	Total
Omisión	Si	84	7	91
	No	185	43	228
	Total	269	50	319

Cuadro: Clasificación de los segmentos de acuerdo a la Omisión

# Resultados

- Un 92 % de los segmentos omitidos se encuentran en el estrato de Alta Probabilidad.
- Como ejemplo de referencia, en la validación de la cartografía LANDSAT de BN se obtuvieron 22 casos de omisión en 1534 segmentos seleccionados, con este diseño se obtienen 91 en 319 de los segmentos analizados.

# Resultados

		Estrato		Total
		Alta Probabilidad	Baja Probabilidad	
Clase	ANFB+OFBN	81	23	104
	ANFB	52	18	70
	OFBN	29	5	34

Cuadro: Cantidad de píxeles con presencia de AFBN y/o AFBN por estrato

# Estimación de Varianza

Para la estimación de la varianza se asume que el muestreo es **con remplazo**, y que la estratificación es “fija”.

$$\hat{V}_0 = \sum_{h=1}^{h=H} \frac{1}{n_h(n_h - 1)} \sum_{s_h} (y_k/p_k - \hat{t}_{\pi h})^2$$

donde

- $p_k = \pi_k/n_h$
- $\hat{t}_{\pi h} = \sum_{s_h} y_k/\pi_k$
- $\pi_k$  es la probabilidad de inclusión del píxel  $k$  en la muestra.

# Estimación de Varianza

Las varianzas se obtienen con la librería *survey* de R (Lumley, 2020).

<b>Resultados del Análisis de Incertidumbre para Superficie de ANFB+OFBN en ha</b>			
Estimación	Inf. 95 %	Sup. 95 %	Error Relativo
57.743	36.645	78.874	36.5 %

# Stock de Carbono

Para el cálculo del stock de carbono en ANFB y OFBN se asigna el promedio de la Eco-región por ha a la superficie estimada.

<b>Resultados del Análisis de Incertidumbre para Stock de Carbono en ANFB+OFBN en ha</b>			
Estimación	Inf. 95 %	Sup. 95 %	Error Relativo
3.517.718	2.137.178	4.898.258	39.2 %

El carbono estimado en ANFB y OFBN representa entre un 3.5 y un 8 % del stock de carbono actual en BN a nivel nacional.

# Discusión

- El estrato de Baja Probabilidad se comporta como los diseños propuestos por Olofsson et al. (2014). Agrega variabilidad debido a la diferencia en los pesos muestrales con el estrato de Alta Probabilidad.
- Los ANFB tienen una mayor presencia en el estrato de Baja Probabilidad (agrega varianza).
- Las probabilidades obtenidas a partir del CART llevan a que no se pueda implementar un diseño proporcional al tamaño, son fijas por estrato.

# Líneas futuras

- Trabajar con los segmentos como unidades de muestreo.
- Utilizar otros modelos para clasificar.
- Estimar la varianza considerando la variabilidad que puede tener la muestra de entrenamiento del modelo a aplicar.



# ¿Preguntas?

# Bibliografía

- Bernardi, L., Boccoardo, A., Miguel, C., Olivera, J., Penengo, C., Rama, G. (2019). Metodología aplicada en la elaboración del mapa de cobertura del Bosque Nativo de Uruguay para el año 2016 con imágenes Sentinel 2. Proyecto REDD+, MGAP- MVOTMA.
- Brazeiro, A., Panario, D., Soutullo, A., Gutierrez, O., Segura, A. y Mai, P. (2012): Clasificación y delimitación de las eco-regiones de Uruguay. Informe Técnico. Convenio MGAP/PPR – Facultad de Ciencias/Vida Silvestre/ Sociedad Zoológica del Uruguay/CIEDUR. 40p.
- Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. (1984). Classification and regression trees: Wadsworth, Inc.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thaub, D., Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202: 18-2.7

# Bibliografía

- Kuhn, M. (2020). caret: Classification and Regression Training. R package version 6.0-86. <https://CRAN.R-project.org/package=caret>
- Lumley, T. (2020) "survey: analysis of complex survey samples". R package version 4.0.
- Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. Remote Sensing of Environment, 148:42-57.
- R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Riaño, M. E., Penengo, C., Bernardi, L. (2020). Mapa de Carbono en la Biomasa viva del Bosque Nativo de Uruguay. Proyecto REDD+, MGAP- MVOTMA.