

UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS AGRONOMICAS
ESCUELA DE PREGRADO

MEMORIA DE TÍTULO

**PREDICCIÓN DE LA RIQUEZA DE PLANTAS VASCULARES Y CARBONO
AÉREO EN TURBERAS ANTROPOGÉNICAS: COMPARACIÓN DE DOS
MÉTODOS DE MODELACIÓN CON TRES SENSORES REMOTOS DE USO
LIBRE**

IVÁN CASTILLO RIFFART

SANTIAGO, CHILE

2017

UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS AGRONÓMICAS
ESCUELA DE PREGRADO

MEMORIA DE TÍTULO

**PREDICCIÓN DE LA RIQUEZA DE PLANTAS VASCULARES Y CARBONO
AEREO EN TURBERAS ANTROPOGÉNICAS: COMPARACIÓN DE DOS
MÉTODOS DE MODELACIÓN CON TRES SENSORES REMOTOS DE USO
LIBRE**

**PREDICTING VASCULAR PLANT RICHNESS AND ABOVEGROUND CARBON
IN ANTHROPOGENIC PEATLANDS: COMPARISON OF MODELING
METHODS WITH FREELY AVAILABLE REMOTE SENSING DATA**

IVÁN CASTILLO RIFFART

Santiago, Chile

2017

UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS AGRONÓMICAS
ESCUELA DE PREGRADO

**PREDICCIÓN DE LA RIQUEZA DE PLANTAS VASCULARES Y
CARBONO AEREO EN TURBERAS ANTROPOGÉNICAS:
COMPARACIÓN DE DOS MÉTODOS DE MODELACIÓN CON TRES
SENSORES REMOTOS DE USO LIBRE**

Memoria para optar al título profesional de:
Ingeniero en Recursos Naturales Renovables

IVÁN CASTILLO RIFFART

PROFESORES GUÍA

Calificaciones

Mauricio Galleguillos T.
Ingeniero Agrónomo, M.S., Ph.D.

7,0

Jorge Pérez Q.
Ingeniero Agrónomo, M.S., Ph.D.

7,0

PROFESORES EVALUADORES

Jaime Hernández P.
Ingeniero Forestal, PhD.

6,2

Ian Homer B.
Ingeniero Agrónomo, Ph.D.

6,7

SANTIAGO, CHILE

2017

AGRADECIMIENTOS

Primero que todo agradezco al proyecto CONICYT/FONDECYT 1130935 “Balance de gases de efecto invernadero en un bosque maduro y una turbera secundaria en Nordpatagonia” por el financiamiento que hizo posible esta memoria.

Segundo agradezco a mis profesores guía Jorge y Mauricio por todo el apoyo entregado, la buena disposición y la paciencia en todo el proceso, hicieron este proceso mucho más agradable y fue bkn conocerlos en otro contexto más personal, son unos grandes.

Tercero a Lopa y Julian por la ayuda entregada en los códigos y en todas las dudas que se me presentaron al respecto, además de los consejos y sobre todo la buena onda. Agradecimientos especiales a la Chica por ser mi sombra en los muestreos y por toda la ayuda entregada en terreno y laboratorio (la mejor ahijada de todas), así como a la javi y la maca que también estuvieron ahí cuando las necesite.

Al Laboratorio SAP por la buena disposición y facilitar los hornos de secado. A la Eme, Aira, Rodrigo, Wlady y Carla por la ayuda en la identificación de especies. A Juanito por prestar las herramientas y por las conversaciones del desayuno. Al equipo del LECs por acompañarme en todo el proceso, ojala sigua siendo un laboratorio con un grupo humano tan diverso y buena onda, espero seguir un tiempo trabajando con ustedes.

Finalmente gracias a mi familia, especialmente a Paty (madre) por todo el apoyo incondicional y por darme un techo donde vivir y alimentarme.

Y gracias a todos los que sienten que debieran estar aquí y no están (todos los amigos encontrados en el camino), pero la verdad es que en la tesis no me ayudaron en nada, aunque sí en muchas otras cosas que tiene mucha más importancia en la vida... gracias totales por estar siempre presentes y hacer de la etapa universitaria una de las mejores (hasta ahora...).

ÍNDICE

RESUMEN	1
ABSTRACT	2
1. INTRODUCCIÓN	3
1.1. OBJETIVOS	6
1.1.1. <i>Objetivo general</i>	6
1.1.2. <i>Objetivos específicos</i>	6
2. MATERIALES Y MÉTODOS	7
2.1. ÁREA DE ESTUDIO	7
2.2. MATERIALES Y MÉTODOS	8
2.2.1 <i>Determinación de los puntos de muestreo</i>	8
2.2.2 <i>Obtención de variables espectrales</i>	8
2.2.3 <i>Selección de variables</i>	11
2.2.4 <i>Modelos estadísticos</i>	11
2.2.5 <i>Validación de los modelos y mapa predictivo de especies</i>	12
3. RESULTADOS	14
3.1. MODELOS PREDICTIVOS	14
3.1.1 <i>Riqueza de plantas vasculares</i>	14
3.1.2 <i>Carbono aéreo</i>	18
4. DISCUSIÓN	21
4.1. RIQUEZA Y CARBONO AÉREO.....	21
4.2. VARIABLES PREDICTIVAS SELECCIONADAS Y SUS IMPLICANCIAS ECOLÓGICAS	22
4.2.1 <i>Riqueza de plantas vasculares</i>	24
4.2.2 <i>Carbono aéreo</i>	24
4.3. COMPARACIÓN DE SENSORES SATELITALES.....	25
4.3.1 <i>Riqueza de plantas vasculares</i>	26
4.3.1 <i>Carbono aéreo</i>	27
4.4. COMPARACIÓN DE MODELOS.....	28
5. CONCLUSIONES	30
8. BIBLIOGRAFÍA	31
8. APÉNDICE	39

RESUMEN

Las turberas son ecosistemas de gran relevancia debido a su capacidad de proveer bienes y servicios a los seres humanos. Las turberas antropogénicas se originan tras la quema o tala de bosque en suelos poco drenados, permitiendo el asentamiento de musgo del género *Sphagnum*. El carbono (C) aéreo en general no es estimado en estos ecosistemas debido a que el reservorio subterráneo es más relevante, aun cuando en turberas antropogénicas existe un importante crecimiento de especies vasculares sobre el sustrato, cuyo estudio ha sido escasamente abordado con sensores remotos. En el presente estudio se realizó una predicción de la riqueza de plantas vasculares y C aéreo en tres turberas antropogénicas de la Isla de Chiloé, Chile, utilizando datos satelitales de los sensores de uso libre OLI, ASTER y MSI, mediante los métodos de modelación predictiva: no paramétrica random forest (RF) y paramétrica generalized linear model (GLM). Se realizaron 15 mediciones por turbera ($n = 45$) en intervalos regulares de 60 m, con cuadrantes de muestreo de 2×2 m para medir riqueza y de $0,5 \times 0,5$ m para medir C aéreo. Se utilizaron como variables predictoras las bandas espectrales, índices vegetacionales y variables texturales (Grey level Co-occurrence Matrix). Adicionalmente se probó una variable categórica como diferenciador de las condiciones estructurales de la vegetación predominante entre las turberas. Las variables más importantes fueron identificadas usando el método de selección de atributos recursive feature elimination (RFE). Se obtuvieron buenas predicciones de riqueza con todos los sensores tanto para GLM (R^2 : 0,53, 0,49, 0,60 para OLI, ASTER y MSI, respectivamente) como para RF (R^2 : 0,52, 0,54, 0,58 para OLI, ASTER y MSI, respectivamente) con RRMSE menores a 20%. Por su parte, en las predicciones de C aéreo los ajustes fueron bastante más bajos tanto con GLM (R^2 : 0,17, 0,13, 0,08 para OLI, ASTER y MSI respectivamente) como con RF (R^2 : 0,2, 0,10, 0,34 para OLI, ASTER y MSI respectivamente) con RRMSE mayores a 66%. La comparación entre los métodos de modelación en general no presentó diferencias significativas, salvo en ASTER al modelar riqueza y MSI en C aéreo, donde RF fue más preciso. La variable categórica mejoró los resultados en los modelos de riqueza, mientras que en los de C aéreo no apareció dentro del ranking de RFE, quedando fuera de la modelación. El sensor MSI resultó ser el más preciso con ambos métodos de predicción, mostrando que la resolución espectral de la zona del infrarrojo cercano es fundamental para estudiar la vegetación en ecosistemas altamente variables, como las turberas antropogénicas.

Palabras claves: Turberas antropogénicas, Riqueza de plantas, OLI, ASTER, MSI, Random forest, Generalized linear models

ABSTRACT

Peatlands are ecosystems of great relevance due to its ability to provide services and goods to mankind. Anthropogenic peatlands emerge after the burn or forest cut in poorly drained soils, allowing the establishment of moss of the genus *Sphagnum*. In general, aboveground carbon (C) is not estimated in these ecosystems since the underground reservoir is more relevant. Nevertheless, in anthropogenic peatlands there is an important growth of vascular species on the aboveground substrate, whose study has been scarcely addressed with remote sensing. In the present study, a prediction of vascular plant richness and aboveground C in three anthropogenic peatlands of the Island of Chiloé, Chile, was done. Satellite data from three freely available sensors (OLI, ASTER and MSI) were used to provide predictors in order to test two predictive modeling methods, non-parametric random forest (RF) and parametric generalized linear models (GLM). Fifteen measurements were made in peatland ($n = 45$) at regular intervals every 60 m, with sampling squares of 2×2 m to measure richness and 0.5×0.5 m to measure aboveground C. Spectral bands, vegetation indices and textural variables (Gray level Co-occurrence Matrix) were used as predictors. Additionally, a categorical variable was used to include the structural conditions of the predominant vegetation among peatlands. The predictor selection was computed with a recursive feature elimination (RFE) method. Good predictions of richness were obtained with all the sensors for both GLM (R^2 : 0.53, 0.49, 0.60 for OLI, ASTER and MSI respectively) and for RF (R^2 : 0.52, 0.54, 0.58 for OLI, ASTER and MSI respectively) with RRMSE $< 20\%$, whereas for aboveground C predictions, the coefficient of correlations were much lower for GLM (R^2 : 0.17, 0.13, 0.08 for OLI, ASTER and MSI respectively) and for RF (R^2 : 0.2, 0.10, 0.34 for OLI, ASTER and MSI respectively) with RRMSE $> 66\%$. The comparison between the modeling methods in general did not present significant statistical differences, except in ASTER when modeling richness and MSI in aboveground C, where RF was more accurate. The categorical variable improved the results in the richness models, while in the aboveground C modelling it was not relevant. The MSI sensor proved to be the most accurate in both methods of prediction, showing that spectral resolution of the near infrared zone is fundamental to study vegetation in highly variable ecosystems such as anthropogenic peatlands.

1. INTRODUCCIÓN

Las turberas son un tipo de humedal en que se produce y deposita la materia orgánica más rápido de lo que se descompone, llevando a su acumulación y formación de turba (Gore, 1983). Estos ecosistemas ofrecen numerosos servicios, destacando su influencia sobre el ciclo hidrológico, al actuar como fuente de almacenamiento y purificación de agua (Bullock y Acreman, 2003) y sobre el ciclo de carbono (C), al ser uno de los mayores sumideros terrestres a nivel mundial (Zedler y Kercher, 2005; Parish et al., 2008). Adicionalmente, estos ecosistemas albergan especies con fuerte singularidad, al estar adaptadas a estados de inundación donde priman los procesos anaeróbicos (Keddy, 2010). Las turberas antropogénicas en Chiloé tienen su origen en la tala o quema del bosque nativo en lugares poco drenados, lo que ocasiona una sucesión vegetacional donde destaca la presencia de musgo del género *Sphagnum* que cubre gran parte de la superficie (Díaz et al., 2007). Este es objeto de manejo debido a su valor como sustrato para mejorar la calidad de los suelos para viveros, cultivos de hortalizas y plantas ornamentales. En este sentido, el estudio de la dinámica vegetacional y sus procesos es fundamental para preservar las funciones y servicios que estos ecosistemas ofrecen a los seres humanos (Duffy, 2009).

La percepción remota representa una herramienta útil para establecer estrategias de conservación y de manejo apropiadas (Turner et al., 2014), cuya ventaja radica en su capacidad de realizar observaciones sinópticas, espacialmente continuas y frecuentes, resultando en un amplio volumen de datos a varias escalas espaciales y temporales (de Araujo-Barbosa et al., 2014). Dichas características son particularmente útiles en ecosistemas de humedales, debido a que permite caracterizar la dinámica espacio-temporal de la vegetación en áreas de difícil acceso donde los muestreos de campo son complejos (Lee y Lunetta, 1996), ofreciendo un método práctico y económico para estimar parámetros tanto bioquímicos como biofísicos de la vegetación de humedal, permitiendo enfocar los esfuerzos de muestreo (Adam et al., 2010). Existen numerosos trabajos que describen técnicas de monitoreo de la vegetación en humedales, donde se ha hecho uso de sensores activos como “Radio Detection and Ranging” (RADAR) (Evans et al., 2014) o “Light Detection and Ranging” (LiDAR) (Luo et al., 2015), pasivos como los multiespectrales (Proisy et al., 2007; Cole et al., 2014) o hiperespectrales (Hestir et al., 2008), o bien combinando ambos tipos (Aslan et al., 2016). Una revisión detallada del tema es abordado por Adam et al. (2010), en el cual presenta las aplicaciones de la percepción remota en la discriminación y mapeo de la vegetación en humedales.

Entre los tipos de datos satelitales disponibles se destacan aquellos de uso libre, los cuales son accesibles a todo el público y están exentos de cobro. Estos son de particular importancia para aplicaciones donde no hay un interés económico directo, entre estos, “Landsat Legacy Project” es la familia de satélites líder en el estudio de la vegetación, principalmente por poseer un registro temporal muy extenso (desde 1972 a la fecha), documentándose varias aplicaciones en el campo de la predicción de la diversidad de plantas (Savage et al., 2015) y de biomasa aérea (Tan et al., 2003), que es una aproximación comúnmente usada para estudiar el ciclo del carbono (Lu, 2006). El satélite

más reciente de esta familia es Landsat 8, el cual se encuentra en órbita desde febrero de 2013 y cuenta con el sensor “Operational Land Imager” (OLI) que cuenta con una resolución espacial de 30 m y temporal de 16 días.

El sensor “Advanced Spaceborne Thermal Emission and Reflection” (ASTER) es otra alternativa que ha sido utilizada para estudiar la vegetación y su diversidad (Mutowo y Murwira, 2012) y estimar la biomasa aérea (Muukkonen y Heiskanen, 2007), aunque los ejemplos son mucho menos frecuentes que con Landsat, ya que se encuentra en órbita desde el año 1999 y sus productos fueron liberados de costo recientemente (abril de 2016), con lo que se proyecta que su utilización vaya en aumento de ahora en adelante. Por otro lado, en junio de 2015 se lanzó la misión Sentinel 2A, pensada especialmente para estudiar la vegetación, al llevar a bordo el sensor “Multi Spectral Instrument” (MSI) que cuenta con cinco bandas que cubren la zona del espectro infrarrojo cercano y dos en el infrarrojo medio, zona del espectro electromagnético más idónea para discriminar y estudiar las coberturas vegetales (Chuvieco, 2008). Actualmente son pocos los estudios que han utilizado estas imágenes para monitorear la vegetación (Fernandez-Manso et al., 2016) dado lo reciente de su puesta en órbita, encontrando un único caso en la discriminación de especies (Laurin et al., 2016) y en la estimación de biomasa aérea (Sibanda et al., 2016), proyectándose un aumento importante dadas sus características.

No obstante, los numerosos y variados estudios vinculados al uso de la percepción remota para caracterizar la vegetación en humedales, estos son menos frecuentes cuando se considera la diversidad de plantas usando modelos empíricos y aún menos frecuentes para ecosistemas de turberas, donde solo se reporta el trabajo de Cabezas et al. (2016), quienes estimaron la riqueza de plantas vasculares en una turbera antropogénica de la Isla grande de Chiloé, Chile, utilizando modelación predictiva basada en imágenes Landsat 8.

La estimación de la riqueza o diversidad alfa presenta características y desafíos que deben ser tomados en cuenta. Rocchini et al. (2010) afirma que una de las áreas significativas de refinamiento en las estimaciones es el proceso de creación de los modelos empíricos, donde los estudios que estiman riqueza mediante teledetección usualmente utilizan regresiones univariantes simples (o lineales) (Oldeland et al., 2010) y regresiones múltiples (Ceballos et al., 2015), mientras que otros estudios utilizan modelos más complejos como “partial least square” (PLS) (Feilhauer y Schmidtlein, 2009), “path modeling” (Lopatin et al., 2015), “generalized additive models” (GAM) (Fava et al., 2010) y métodos de “machine learning” como las redes neuronales (NN) (Foody and Cutler, 2003) o “random forest” (RF) (Cabezas et al., 2016) entre otros.

Se ha planteado con anterioridad que la aplicación directa de índices vegetacionales y modelos de regresiones simples no son capaces de utilizar completamente la información procedente de datos de teledetección (Foody y Cutler, 2006; Lopatin et al., 2016), por lo cual el uso de modelos no paramétricos y “machine learning” se han implementado en mayor medida en la última década (Rocchini et al., 2010). Se suele pensar que este tipo de modelos son mejores que los modelos lineales en cualquier tipo de estudio. Sin embargo Lopatin et al. (2016) demostraron que tal afirmación no es necesariamente cierta en cuanto a modelos de riqueza de plantas vasculares. Lopatin et al. (2016) realizaron un análisis comparativo demostrando que “genalized linear models” (GLM) se comportaban de mejor

forma que RF en un bosque complejo de Chile central usando datos LiDAR, debido a que RF no fue capaz de representar correctamente la asimetría característica de los “count data” (e.g. número de especies), los cuales son discretos y limitados a valores positivos. Sin embargo, es preciso verificar si estas tendencias son generalizables a otros ecosistemas.

Por otro lado, la predicción del contenido de carbono presenta un alto grado de incertidumbre (Lu, 2006), y una manera útil para proceder en su estimación es desarrollando aproximaciones a partir del estudio de la biomasa, dado que el C representa una fracción de ésta (Dube y Mutanga, 2016). En ese sentido, el C aéreo generalmente no es estimado en ecosistemas de turbera, debido a que el reservorio subterráneo es el de mayor importancia (Gore, 1983). A pesar de esto, las turberas antropogénicas poseen la particularidad de presentar un importante crecimiento de plantas vasculares sobre el sustrato de musgo *Sphagnum*, el cual puede ser interpretado como un indicador de sucesión, aportando información sobre el estado en que se encuentra el ecosistema.

La percepción remota se ha convertido en la herramienta principal para la estimación de la biomasa en la vegetación verde (Lu, 2006) y en una importante herramienta para estimar los reservorios terrestres de C (Kelsey y Neff, 2014), siendo abordada desde distintos enfoques predictivos con distintos tipos de sensores remotos, no habiendo consenso respecto a cuál es la mejor práctica para estimar esta variable (Fassnacht et al., 2014a). No obstante, según Fassnacht et al. (2014a) la mayor cantidad de estudios han aplicado diferentes tipos de modelos lineales (Morel et al., 2012), seguido de modelos no paramétricos basados en aproximaciones de vecinos más cercanos (“nearest neighbor”) (Karjalainen et al., 2012) o utilizando método de ML como “support vector machine” (SVM) (Zhao et al., 2011) y RF (Latifi y Koch, 2012), siendo estos últimos lo que han predominado en los últimos años, por ser más flexibles y ofrecer mejores rendimientos que los modelos lineales en escenarios donde el número de puntos de muestreo son limitados.

Es así como el estudio de la biomasa aérea ha sido ampliamente desarrollado desde la modelación predictiva con sensores remotos, los cuales se han dedicado principalmente al estudio de ecosistemas boscosos (Latifi y Koch, 2012), habiendo algunos ejemplos desarrollados en humedales de tipo manglar (Proisy et al., 2007) y bosques de turba pantanosa (Manuri et al., 2017), encontrando un único caso de estudio aplicado en turberas antropogénicas, el cual fue desarrollado por Cabezas et al. (2016) en la Isla de Chiloé, Chile, donde se estimó el contenido de C aéreo mediante datos derivados de Landsat 8 y Pleiades.

El objetivo de este estudio fue predecir la riqueza de plantas vasculares y el reservorio de C aéreo en tres turberas antropogénicas de Chiloé, donde se compararon dos métodos de modelación mediante datos de tres sensores de uso libre. Se realizará una comparación de la exactitud de la predicción obtenida mediante los métodos de modelación random forest y generalized linear models, utilizando los sensores OLI, ASTER y MSI a bordo de las plataformas satelitales Landsat 8, Terra y Sentinel 2A respectivamente.

1.1. Objetivos

1.1.1. Objetivo general

Predecir la riqueza de plantas vasculares y el contenido de carbono aéreo en tres turberas antropogénicas, comparando dos métodos de modelación mediante datos de tres sensores satelitales de uso libre.

1.1.2. Objetivos específicos

- Cuantificar la riqueza de plantas vasculares y carbono aéreo
- Generar y comparar los modelos predictivos

2. MATERIALES Y MÉTODOS

2.1. Área de estudio

El área de estudio está ubicada en la zona norte de la Isla Grande de Chiloé, en la comuna de Ancud, Región de Los Lagos, Chile ($41^{\circ}52'40''\text{S}$, $73^{\circ}40'00''\text{W}$) (Figura 1). En específico corresponde a tres turberas antropogénicas, ubicadas en la Estación Biológica Senda Darwin (T2) y sus alrededores, cuyas extensiones son de 10,5, 5,6 y 6,2 ha para T1, T2 y T3, respectivamente (Figura 1). Estas turberas tienen en común que en la actualidad no presentan perturbaciones antrópicas ni presiones externas significativas, permitiendo que especies de plantas vasculares se desarrollen sobre el sustrato de musgo *Sphagnum*.

El clima es templado con fuerte influencia oceánica, con presencia de un período seco en los meses de verano. La precipitación anual es de 2.110 mm con una temperatura media anual de 10°C .

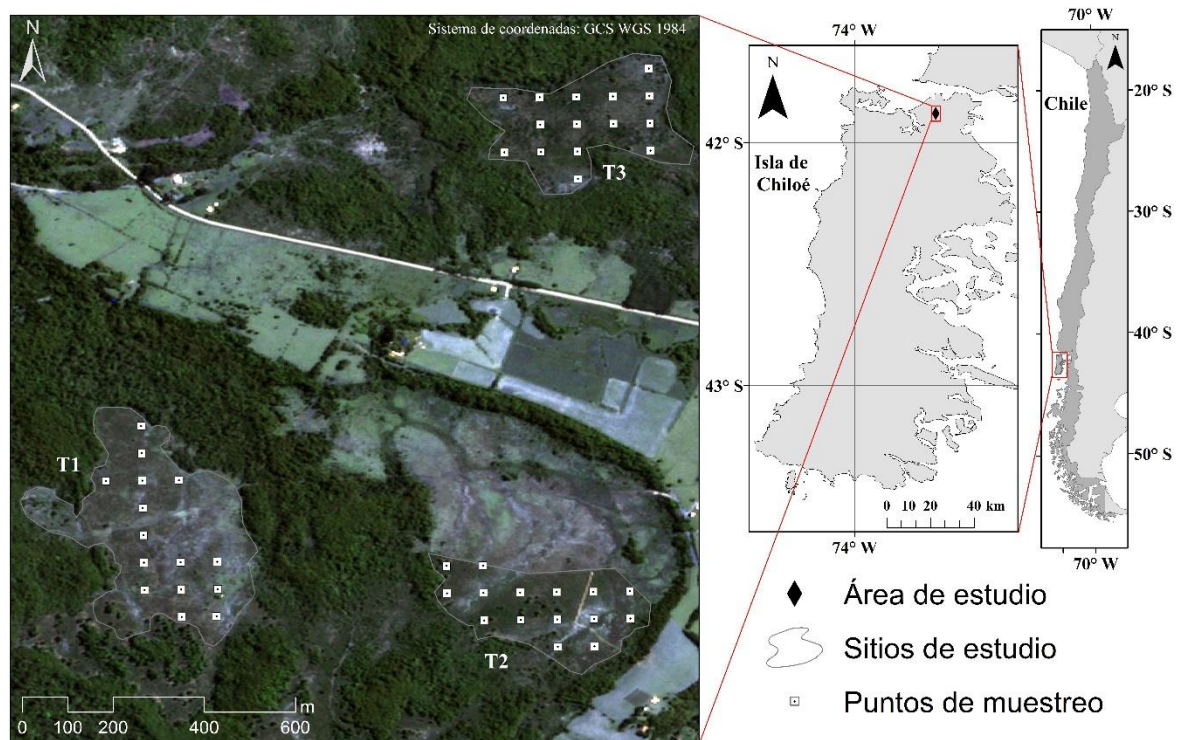


Figura 1. Área de estudio. Se muestran las turberas estudiadas (T1, T2 y T3).

2.2 Materiales y Métodos

2.2.1 Determinación de los puntos de muestreo

Se realizó un muestreo sistemático de la riqueza de plantas vasculares (expresada en número de especies) y biomasa aérea, para representar adecuadamente la variabilidad del ecosistema, ya que según Ulanowski y Branfireun (2013), las turberas presentan gran heterogeneidad espacial. Para esto, se estableció una grilla regular con 15 puntos de muestreo separados cada 60 m en cada turbera, dando un total de 45 muestras (Figura 1). Para medir la riqueza de plantas vasculares se utilizaron cuadrantes de muestreo de 2×2 m en cada punto (Díaz et al., 2008; Cabezas et al., 2015), en los cuales se identificaron las especies de plantas vasculares presentes. Las especies que no se lograron identificar en terreno fueron herborizadas e identificadas en laboratorio mediante guías de campo y ayuda de expertos en flora.

Para medir la biomasa y el contenido de C aéreo se realizó una cosecha total de la biomasa en 0,25 m² (cuadrantes de 0,5×0,5 m) en el centro de cada punto de muestreo (Bubier et al., 2006), diferenciando entre especies herbáceas y arbustivas. La biomasa total fue pesada en terreno, para luego obtener submuestras de peso conocido, las cuales fueron secadas en un horno a 70 °C durante 72 horas, obteniendo la masa seca y el porcentaje de humedad de las muestras. Así se obtuvo la biomasa seca de cada cuadrante, que luego fue extrapolado a kg m⁻². Finalmente, se utilizaron los porcentajes de C de la biomasa reportados por Cabezas et al. (2015), siendo 51,2±0,3% para las arbustivas y 48,9±0,3% para las herbáceas, datos que fueron obtenidos de la turbera ubicada en la Estación Biológica Senda Darwin (T2).

Los muestreos se realizaron en dos fechas, el primero en enero de 2014 y abarcó a la turbera T2, y el segundo en enero de 2016, abarcando las turberas T1 y T3 (Figura 1). Cabe señalar que las tres turberas no presentan manejo observables, siendo la turbera T2 una zona de conservación, aldaño a un sector con manejo productivo (Figura 1).

Adicionalmente, se midió la firma espectral de la vegetación en los 15 puntos de la turbera T2 utilizando un espectroradiómetro (ASD Fieldspec 4 JRTM, ASD Inc., Boulder, CO), para ayudar a explicar la relación entre la firma espectral de la vegetación en ecosistemas de turbera y las variables que fueron seleccionadas en los modelos predictivos.

2.2.2 Obtención de variables espectrales

Las variables predictivas fueron derivadas de imágenes satelitales de tres sensores de uso libre: sensor OLI a bordo de Landsat 8, sensor ASTER a bordo de Terra y sensor MSI a bordo de Sentinel 2A, cuyas especificaciones se presentan en el Cuadro 1. Las imágenes de OLI y ASTER corresponden a las fechas más cercanas a las mediciones en terreno, mientras que con el sensor MSI solo se usó una imagen tomada durante el segundo

muestreo, dada su puesta en órbita en junio de 2015 (posterior a la fecha de la primera toma de datos en terreno) (Cuadro 2).

Cuadro 1. Características de los sensores y sus bandas utilizadas en este estudio

Sensor	Bandas		Resolución			
			Espectral (μm)	Espacial	Temporal	Radiométrica
OLI	Blue	B2	0,450-0,510	30 m	16 días	12 bits
	Green	B3	0,530-0,590			
	Red	B4	0,640-0,670			
	NIR	B5	0,850-0,880			
	SWIR 1	B6	1,570-1,650			
	SWIR 2	B7	2,110-2,290			
	ASTER	Green	B1			
Red		B2	0,630-0,690			
NIR		B3N	0,780-0,860			
MSI	Blue	B2	0,457-0,522	10 m	10 días	12 bits
	Green	B3	0,542-0,578			
	Red	B4	0,650-0,680			
	NIR	B8	0,785-0,900			
	NIR	B5	0,698-0,713	20 m		
	NIR	B6	0,733-0,748			
	NIR	B7	0,773-0,793			
	NIR	B8a	0,855-0,875			
	SWIR 1	B11	1,565-1,655			
	SWIR 2	B12	2,100-2,280			

Cuadro 2. Productos utilizados por sensor y su fecha de adquisición

Características	Sensor		
	OLI	ASTER	MSI
Producto utilizado	Climate Data Record	7XT	Level-1C
Nivel de procesamiento	Reflectancia superficial	Reflectancia superficial	Reflectancia sobre la atmosfera
Fecha de adquisición	24/12/2013 (T1 y T3) 06/01/2016 (T2)	23/12/2013 (T1 y T3) 31/01/2016 (T2)	05/01/2016 (T1, T2 y T3)

Las variables predictoras utilizadas fueron la reflectancia de la zona del espectro electromagnético del visible (0.4-0.7 μm) e infrarrojo cercano (NIR: 0.7-1.3 μm) de todos los sensores, además del infrarrojo de onda corta (SWIR: 1.3-8 μm) en OLI y MSI, dado que el sensor ASTER no posee bandas en esta zona del espectro de las fechas estudiadas. A

partir de estas variables se obtuvieron los índices vegetacionales presentados en el Cuadro 3. También se utilizaron variables texturales de las bandas y los índices vegetacionales, siendo estas la mediana y la desviación estándar y las métricas GLCM (grey-level co-occurrence matrix), que corresponden a la media, la varianza, la homogeneidad, el contraste, la correlación, “second moment”, la entropía y la disimilitud. Estas variables fueron incorporadas al procesamiento ya que según Rocchini et al., (2010), la heterogeneidad espectral puede predecir mejor la diversidad de especies, siendo también utilizadas en el modelamiento de C aéreo, dado que estudios previos utilizaron estas variables para ayudar a mejorar los modelos de C (Kelsey y Neff, 2014; Cabezas et al., 2015). Todas las variables texturales fueron calculadas considerando una ventana móvil de 3×3 píxeles.

Cuadro 3. Índices calculados desde los satélites OLI, ASTER y MSI

Índices	Referencia	OLI	ASTER	MSI
All Normalized Difference Index (ND)	Mutanga et al. (2012)	x	x	x
Angle at NIR (ANIR)	Khanna et al. (2007)	x		
Difference Vegetation Index (DVI)	Ling et al. (2014)	x		
Enhanced Vegetation Index (EVI)	Huete et al. (2002)	x		x
Green Normalized Difference Vegetation Index (GNDVI)	Gitelson et al. (1996)	x	x	x
Modified Simple Ratio (MSR)	Haboudane et al. (2004)	x	x	x
Modified Soil Adjusted Vegetation Index 2 (MSAVI2)	Qi et al. (1994)	x	x	
Normalized Difference Vegetation Index (NDVI)	Silleos et al. (2006)	x	x	x
Renormalized Difference Vegetation Index (RDVI)	Haboudane et al. (2004)	x		
Shortwave angle slope index (SASI)	Khanna et al. (2007)	x		
Simple Ratio (SR)	Ling et al. (2014)	x	x	x
Soil Adjusted Vegetation Index (SAVI)	Huete et al. (1997)	x	x	x
Tasseled Cap Transformation (greenness, brightness y wetness)	Baig et al. (2014)	x		x ¹
Transformed Vegetation Index (TVI)	Silleos et al. (2006)	x	x	x ²
Tree Band Vegetation Index (3BSI-T)	Tian et al. (2013)		x	x ²
Tree-Band-Index (3BSI-W)	Wang et al. (2012)		x	
Triangular Vegetation Index (TriVI)	Broge y Leblanc (2000)	x		

¹Se utilizaron los mismos parámetros que para Landsat 8, además se calculó usando las diferentes bandas del NIR.

²Este índice fue calculado selectivamente con las bandas del NIR y Rojo para evitar sobrecargar el procedimiento con muchos predictores.

Además se evaluó la incorporación de una variable categórica como diferenciador entre turberas (de ahora en adelante Id.turbera) en los modelos de riqueza y C aéreo, fundamentado en que podría existir una diferencia entre estas variables dependiendo de la

turbera en que se encuentren. En el caso de la riqueza, esta diferencia se evaluó mediante el índice de Jaccard (Jaccard, 1902), que toma valores entre 0 y 1, donde 1 indica que ambos sitios comparten las mismas especies. Este índice arrojó valores de 0,33, 0,46 y 0,42 entre los pares de turberas T1-T2, T1-T3 y T2-T3 respectivamente, los cuales se consideraron lo suficientemente bajos para establecer que las tres turberas son diferenciables entre ellas.

En el caso del C aéreo, la diferencia se evaluó mediante el test de Wilcoxon (Wilcoxon, 1945), el cual arrojó valores de P de 0,008, 0,016 y 0,566 entre los pares de turberas T1-T2, T1-T3 y T2-T3 respectivamente, indicando que dos turberas (T2-T3) pertenecen a la misma población de datos. No obstante, esta variable fue incluida en el procesamiento posterior para evaluar si con ella era posible mejorar los modelos predictivos de C aéreo.

Con lo anterior se obtuvieron tres conjuntos de predictores, dando un total de 360 para el sensor OLI, 148 para ASTER y 1640 para MSI. Los valores de los puntos de muestreo fueron extraídos usando interpolación bilineal (Cabezas et al., 2016).

2.2.3 Selección de variables

Para seleccionar las variables más explicativas se utilizó el algoritmo “Recursive Feature Elimination” (RFE) (Guyon et al., 2002), que opera en base a un proceso de iteraciones en el cual se va eliminando una variable en cada paso basado en la importancia de las variables, obtenida con el algoritmo RF (Breiman, 2001). El algoritmo se implementó utilizando “Leave-one-out cross-validation” (LOOCV) para obtener la relación entre la raíz del error cuadrático medio (RMSE) y el número de predictores, así como también para seleccionar las variables más importantes en cada paso. Se buscó el modelo con menor RMSE y menor número de variables. Además se siguió lo dispuesto por Hair et al. (2009), quienes establecen que el número de observaciones no debe ser menor a 5 por cada predictor, e idealmente debe encontrarse entre 15 y 20.

Con el primer subconjunto obtenido, se evaluó la correlación entre estas variables aplicando el coeficiente de correlación de Pearson (R), seleccionando los pares de variables con un R absoluto $\geq |0.6|$ y eliminando una a una la menos importante de acuerdo al ranking de RFE. Este proceso se realizó diferenciadamente para las variables de cada sensor, puesto que se obtuvieron modelos predictivos con cada uno de ellos para poder compararlos posteriormente.

2.2.4 Modelos estadísticos

La riqueza de especies y el contenido de C aéreo fue modelada usando dos enfoques: RF y GLM.

Random forest (Breiman, 2001) es un método de regresión no paramétrico y estocástico que opera en base a mecanismos de aprendizaje no supervisado. RF ha sido catalogado como un método eficiente para desarrollar predicciones, especialmente cuando el número de observaciones es significativamente menor al número de predictores (Svetnik et al., 2003). Este método requiere el ajuste de dos parámetros: 1) n_{tree} , que define el número de árboles del modelo, los cuales son generados con un conjunto de datos de entrenamiento,

seleccionados de forma aleatoria (con reemplazo), de igual tamaño que el conjunto original, los datos que no pertenecen a dicho conjunto son usados para la validación; y 2) mtry, que define el número óptimo de variables utilizadas en cada nodo del árbol, las cuales son seleccionadas aleatoriamente. El ntree usado fue de 500 (Fassnacht et al., 2014b; Lopatín et al., 2016), mientras que el mtry fue ajustado para cada modelo. Todos los procesos estadísticos fueron realizados en el software R 3.2.5 (paquete randomForest (Liaw y Wiener, 2002) y Caret (Kuhn, 2016)).

Por otro lado, GLM es un método de regresión paramétrico y determinístico, que se diferencia de las regresiones lineales ordinarias por ser más flexible, ya que permite incorporar la distribución de los residuos al proceso a través de una función de enlace. Según Lopatin et al. (2016), el procesamiento de datos con GLM puede ser subdividido en tres pasos:

i) Identificar una familia de modelos apropiada capaz de manejar las propiedades estadísticas de los datos a predecir. Para esto se compararon los gráficos de cuantiles de los residuos normalizados de varios modelos, calculados en primera instancia con la variable más importante según RFE, con las siguientes familias propias de distribución de “count data”: Poisson, Quasi-Poisson y binomial negativa, además se probaron las distribuciones Gamma y Gaussiana con y sin la función de enlace log-link.

ii) Seleccionar el subconjunto de variables predictoras. Se escogieron las cuatro variables más importantes de acuerdo al ranking de RFE, una vez eliminadas las que presentaran una alta colinealidad (ver ítem 2.2.3).

iii) Selección del modelo final. El número de variables predictivas se determinó calculando varios modelos mediante todas las combinaciones posibles entre las cuatro variables más importantes y comparándolos con el criterio de información de Akaike (Posada y Buckley, 2004). En el caso de los modelos de C se transformó la variable de respuesta a su logaritmo natural para cumplir con los supuestos de los modelos lineales.

Para comprobar los supuestos de los GLM se aplicó el test de Shapiro-Wilk (Shapiro y Wilk, 1965) a los residuos de las regresiones obtenidas a partir de los datos de cada sensor y el test de Moran para evaluar la autocorrelación espacial de los mismos (Cliff y Ord, 1972), descartando aquellos modelos que no cumplieran con alguno de estos supuestos.

2.2.5 Validación de los modelos y mapa predictivo de especies

Para la comparación y validación se siguió lo realizado por Lopatín et al. (2016), donde los mejores RF y GLM de cada sensor fueron incorporados individualmente en un “bootstrap” (Efron, 1979) con 500 iteraciones. En estas iteraciones los datos fueron divididos en dos subconjuntos, uno de entrenamiento y uno de prueba, con un promedio de 37.8% (~17 muestras) de los datos para la validación independiente (Fassnacht et al., 2014b). Los modelos construidos con RF y GLM fueron comparados en base a la diferencia del coeficiente de determinación (R^2), RMSE relativo (RMSE dividido en la media de las observaciones, de ahora en adelante RRMSE) y bias o sesgo, calculado como uno menos la pendiente de la regresión de los valores predichos y los observados, sin el intercepto

Finalmente se calculó un mapa predictivo de especies a partir del mejor modelo obtenido. Las predicciones del modelo estadístico se realizaron exclusivamente dentro de los límites de las turberas para evitar predicciones fuera del rango del modelo. Adicionalmente se calculó un mapa del coeficiente de variación (CV, dados en %) de los valores de riqueza y C predichos, calculados en cada pixel a partir de las 500 iteraciones del “bootstrapping”.

3. RESULTADOS

Las mediciones de Riqueza y C aéreo por punto de muestreo se presentan en el Apéndice 1, cuya síntesis se muestra en el Cuadro 4, en el cual se puede observar las diferencias que existen entre las tres turberas respecto de las variables evaluadas. Se puede observar que T1 es la turbera que cuenta con mayor número de especies vasculares, seguida de T3 y T2, situación que es inversa en el caso del C aéreo, es decir que a mayor riqueza, menor es el C aéreo en la vegetación vascular.

Cuadro 4. Estadísticas descriptivas de la riqueza (n° de especies) y el C aéreo (kg m⁻²).

	Mínimo	Máximo	Media	Mediana	DE	CV
Riqueza						
Total	5	14	9.42	10	2.58	27%
T1	7	14	11.31	11.5	1.8	16%
T2	5	10	6.93	7	1.28	18%
T3	6	13	9.73	10	2.05	21%
Carbono						
Total	0.05	3.65	0.57	0.41	0.6	105%
T1	0.05	0.96	0.33	0.25	0.18	55%
T2	0.16	3.65	0.86	0.54	0.9	105%
T3	0.15	1.46	0.57	0.54	0.37	65%

(DE = desviación estándar, CV = coeficiente de variación, N=15 en cada turbera).

3.1 Modelos predictivos

3.1.1 Riqueza de plantas vasculares

I. Generación de modelos predictivos

En el Cuadro 5 se presentan las variables utilizadas en los modelos RF y GLM en la modelación de riqueza. La mayoría de los predictores seleccionados corresponden a variables texturales y todos ellos tienen la influencia del infrarrojo. Además, la incorporación de la variable Id.turbera ayudó a mejorar las predicciones en la mayoría de los modelos, salvo en OLI con RF.

En cuanto a los GLM finales, se utilizó la distribución del error de Poisson con la función log-link, la cual resultó ser ligeramente mejor a las demás distribuciones probadas en este estudio.

Cuadro 5. Variables seleccionadas en los modelos finales de riqueza. Se marca con una “x” las variables usadas en cada modelo, ordenadas por sensor y de acuerdo a su importancia obtenida con RFE.

Sensor	Variables	Métricas texturales	GLM	RF
OLI	Id.turbera	Media	x	
	NDVI	Contraste		x
	B5		x	x
	ND entre B2 y B7		x	x
ASTER	Id.turbera		x	x
	SR entre B2 y B3N	Varianza		x
	$3BSI-W = (B1 - B2 + 2 B3N) / (B1 + B2 - 2 B3N)$	Second Moment	x	x
	B3N)			
MSI	Id.turbera		x	x
	$3BSI-T = (B7 - B6 - B8a) / (B7 + B6 + B8a)$	Homogeneidad		x
	SR entre B6 y B8a	Media	x	
	B8a		x	x

El significado de las variables se presenta en el Cuadro 3.

II. Procedimiento de validación

Al aplicar la estrategia de validación mediante la prueba de “bootstrap”, los GLM presentaron una leve tendencia a ser mejores que RF. Esta tendencia se dio en los modelos con los sensores OLI y MSI, los que presentaron mejores ajustes de R^2 y menor error y sesgo (Figura 2 y Cuadro 6). No obstante, las diferencias no resultaron ser significativas entre ellos. Por otro lado, los modelos con el sensor ASTER resultaron ser los únicos significativamente diferentes, donde RF mostró mayor precisión en cuanto al R^2 y RRMSE que GLM (Cuadro 6). Respecto al sesgo, presentó una leve tendencia a sobreestimar los valores en todos los modelos (entre 0,01 y 0,04), no presentando diferencias significativas en ninguno de los casos.

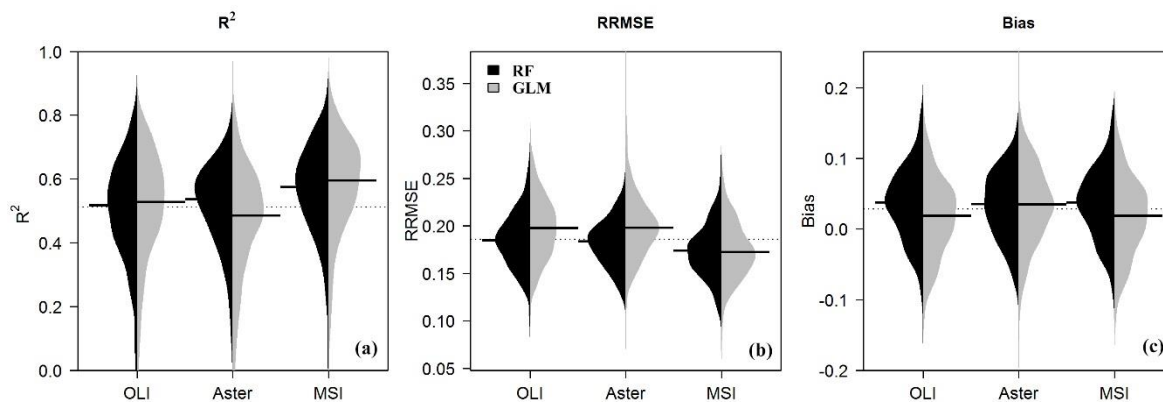


Figura 2. Precisión de los modelos de riqueza utilizando RF (negro) y GLM (gris) en términos de R^2 (a), RRMSE (b) y bias (c). Los Beanplots muestran la distribución de los resultados de las 500 iteraciones. Las líneas negras horizontales indican la mediana de las distribuciones.

Pese a lo anterior, los modelos generados con el sensor MSI obtuvieron la mayor precisión tanto con GLM como con RF, con R^2 de 0,6 y 0,58, RRMSE de 0,17 y 0,17 y Bias de 0,01 y 0,03 respectivamente. Los valores de R^2 , RRMSE y sesgo de todos los modelos se muestran en el Cuadro 6.

Cuadro 6. Precisión de los modelos predictivos de riqueza. Se presenta el valor de la mediana de las 500 iteraciones.

Sensor	R^2			RRMSE			Bias		
	GLM	RF		GLM	RF		GLM	RF	
OLI	0,53	0,52	-	0,20	0,18	-	0,01	0,04	-
ASTER	0,49	0,54	*	0,20	0,17	*	0,03	0,02	-
MSI	0,60	0,58	-	0,17	0,17	-	0,01	0,03	-

* Resultados significativamente diferentes ($p < 0,05$).

III. Mapa predictivo

En la Figura 3 se presenta el mapa predictivo de riqueza de plantas vasculares y el CV asociado, calculado mediante el método GLM con el sensor MSI, que resultó ser el mejor modelo entre los evaluados en este estudio. La distribución de la riqueza sigue los patrones descritos en el Cuadro 4, estimando mayor riqueza en la turbera T1 y menor riqueza en la turbera T2, aun cuando en algunos bordes de ambas turberas se puede ver que el modelo tiende a arrojar valores más altos. La turbera T2 por su parte presenta una distribución más homogénea de la riqueza, que podría estar asociado a la menor variabilidad espacial de sus coberturas. El mapa de CV muestra que la riqueza es predicha relativamente con baja variación, cuyo valor máximo es de 15%, encontrándose en las zonas donde la riqueza es mayor, sin embargo en gran parte de las turberas el error es cercano al 9%, lo que da cuenta de un modelo consistente, resultados que concuerdan con el RRMSE de las predicciones, el cual es de 17%, que es considerado un error bajo.

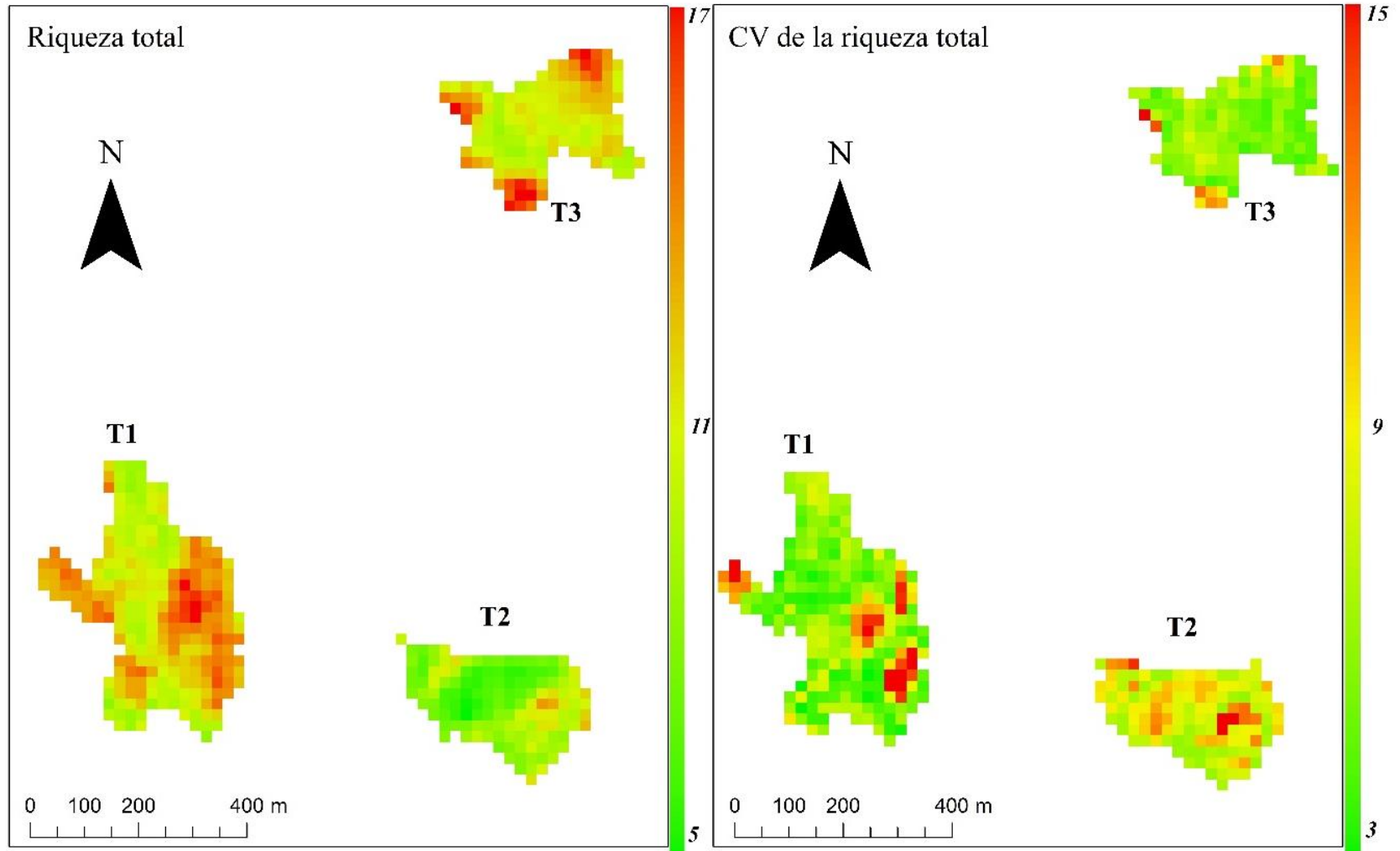


Figura 3. Mapa predictivo de riqueza de especies (en número de especies) y su coeficiente de variación (CV en %).

3.1.2 Carbono aéreo

I. Generación de modelos predictivos

En el Cuadro 7 se presentan las variables utilizadas en los modelos RF y GLM en la modelación de C aéreo, destacando la poca influencia que poseen los predictores texturales sobre esta variable, a diferencia de la riqueza. Nuevamente la zona del espectro infrarrojo aparece como la zona más idónea en las predicciones, al estar incluidas en once de los doce predictores finales.

Respecto a los GLM, la distribución normal (gaussiana) resultó ser levemente mejor a las demás distribuciones probadas, la que corresponde a un modelo lineal tradicional.

Cuadro 7. Variables seleccionadas en los modelos finales de C aéreo. Se marca con una “x” las variables usadas en cada modelo, ordenadas por sensor y de acuerdo a su importancia obtenida con RFE.

Sensor	Variables	Métricas texturales	GLM	RF
OLI	ND entre las Bandas 2 y 4		x	x
	ND entre las Bandas 4 y 7		x	x
	ND entre las Bandas 6 y 7	Entropía		x
	ND entre las Bandas 5 y 6	Media		x
ASTER	$BSI-T = (B1 - B2 - B3N) / (B1 + B2 + B3N)$	Media		x
	$BSI-T = (B3N - B1 - B2) / (B3N + B1 + B2)$	Desviación estándar	x	x
	$BSI-W = (B1 - B2 + 2 B3N) / (B1 + B2 - 2 B3N)$			x
	$BSI-T = (B1 - B2 - B3N) / (B1 + B2 + B3N)$		x	
MSI	SR entre Banda 2 y 5			x
	SR entre Banda 8a y 8		x	x
	$BSI-T = (B4 - B3 - B8) / (B4 + B3 + B8)$		x	
	ND entre las Bandas 6 y 7			x

II. Procedimiento de validación

Al aplicar la validación mediante la prueba de “bootstrap”, todos los modelos presentaron bajos ajustes de R^2 a la vez que el error y sesgo fueron altos (Figura 4 y Cuadro 8). No obstante el método de modelación RF mostró mayor precisión que GLM, aun cuando estas diferencias fueron significativas en solo dos casos, en el R^2 de MSI y RRMSE de ASTER y MSI. Por otro lado, los modelos RF con los sensores ASTER y MSI presentaron una distribución bimodal del error a diferencia de los GLM donde la distribución sigue una tendencia más parecida a la normal.

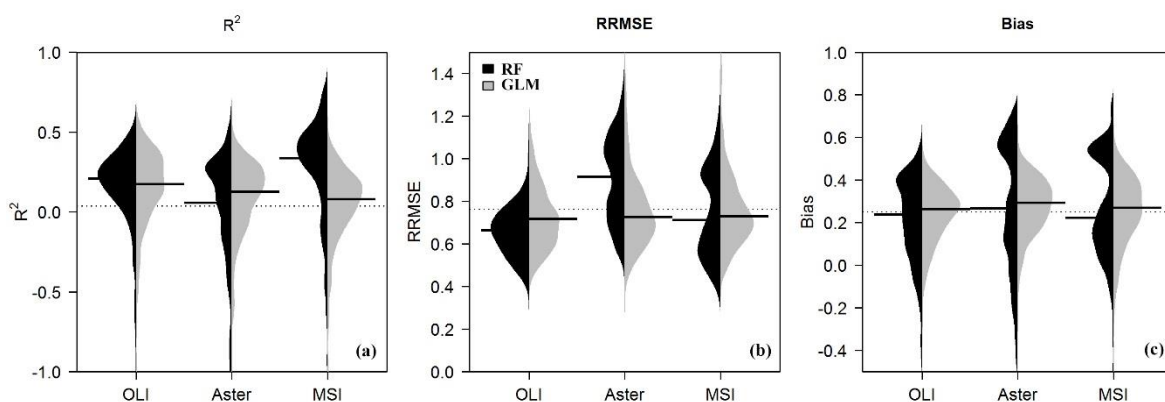


Figura 4. Precisión de los modelos de C aéreo utilizando RF (negro) y GLM (gris) en términos de R^2 (a), RRMSE (b) y bias (c). Los Beanplots muestran la distribución de los resultados de las 500 iteraciones. Las líneas negras horizontales indican la mediana de las distribuciones.

En el Cuadro 8 se observa que el mejor modelo es generado con el sensor MSI y el método RF ($R^2 = 0,34$) seguido de OLI con el mismo método ($R^2 = 0,20$), mientras que el mejor GLM solo alcanzó un valor de R^2 de 0,17. Respecto al RRMSE, en todos los modelos fue $\geq 60\%$ lo cual indica un alto error asociado a las modelaciones. Por su parte, el sesgo de los modelos fue $\geq 0,2$ en todos los casos, con una tendencia a sobreestimar los valores.

Cuadro 8. Precisión de los modelos predictivos de C aéreo. Se presenta el valor de la mediana de las 500 iteraciones.

Sensor	R^2			RRMSE			Bias		
	LM	RF		LM	RF		LM	RF	
OLI	0,17	0,20	-	0.66	0.70	-	0,23	0,25	-
ASTER	0,13	0,10	-	0.94	0.72	*	0,31	0,31	-
MSI	0,08	0,34	*	0.67	0.74	*	0,23	0,28	-

* Resultados significativamente diferentes ($p < 0,05$).

III. Mapa predictivo

En la Figura 5 se muestra el mapa predictivo de carbono aéreo y el CV asociado, calculado mediante el método RF con el sensor MSI, que resultó ser el mejor modelo de los evaluados. En el mapa predictivo se puede observar que las tres turberas en general presentan un contenido de carbono aéreo similar, con valores entre 0,22 y 0,84 kg m^{-2} en la mayor parte de sus superficies, donde los valores cercanos o mayores a 1 kg m^{-2} son escasos y se encuentran distribuidos en el espacio de forma aleatoria. El mapa de CV muestra que la variación en las predicciones de C es alta, alcanzando un valor máximo de 50%, aun cuando la mayoría de los valores se encuentran cercanos a 30%, valor que de todas formas indica un alto grado de variabilidad en las predicciones. Lo anterior se condice con el bajo R^2 del modelo (0,34) y el alto RRMSE asociado (74%), siendo un modelo predictivo poco consistente.

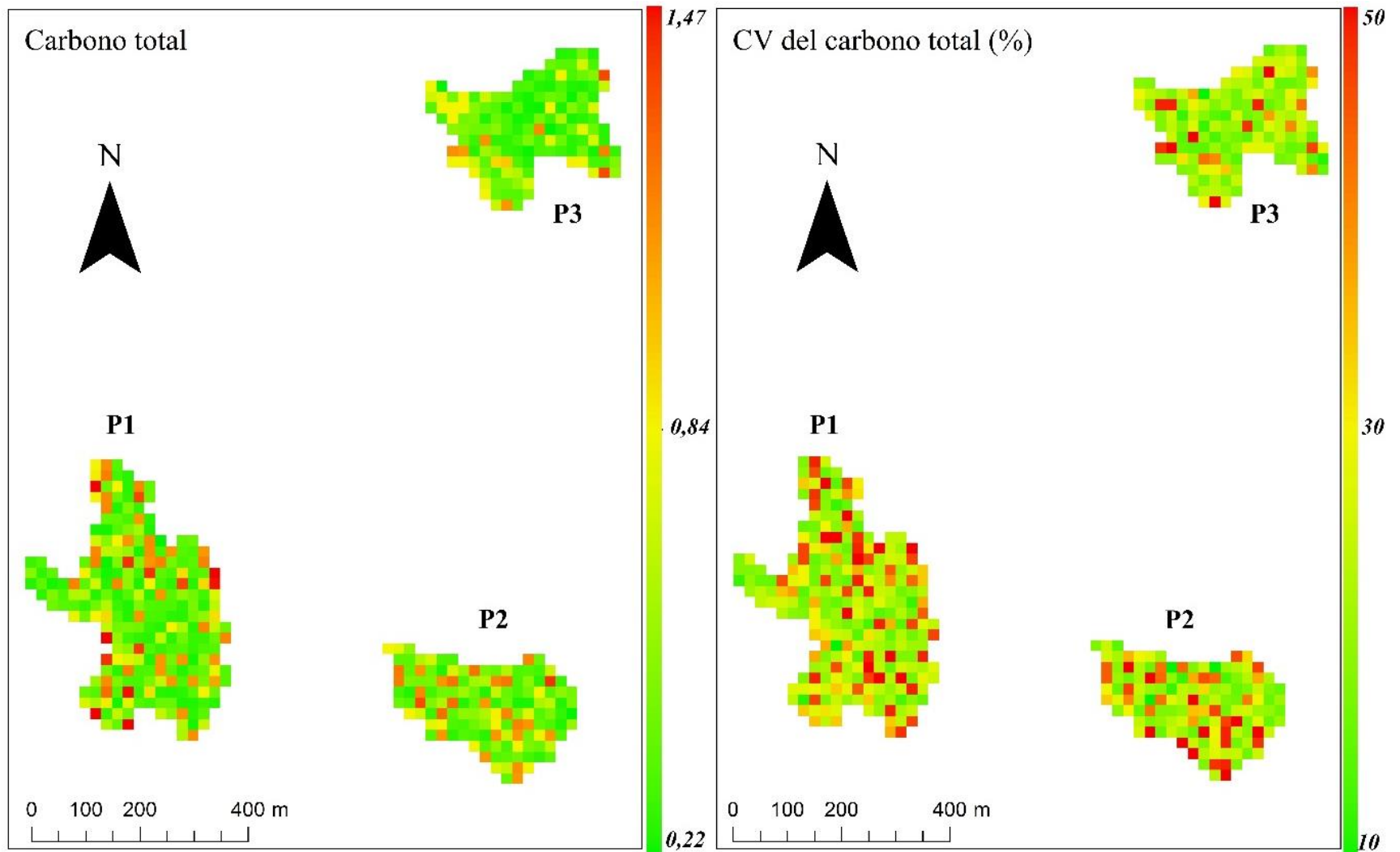


Figura 5. Mapa predictivo de C aéreo (kg m^{-2}) y su coeficiente de variación (CV en %).

4. DISCUSIÓN

4.1. Riqueza y carbono aéreo

La tendencia de la riqueza y el C aéreo a comportarse de manera inversa podría estar relacionado a que la turbera con mayor riqueza (T1) presenta una cobertura dominada por matorral abierto con especies como *Baccharis patagonica* y *Myrteola numularia* dispuestas de manera irregular en el espacio (Figura 6A), permitiendo que una cantidad importante de especies herbáceas crezcan sobre el sustrato de musgo, las cuales poseen una baja biomasa y consecuentemente bajo contenido de C aéreo. Por su parte, la turbera T2 cuenta con una cobertura de matorral denso en gran parte de su superficie dominada por *B. patagonica*, acompañada por *M. numularia*, *Gaultheria insana* y *G. mucronata*, limitando el crecimiento herbáceo bajo su dosel (Figura 6B), razón por la cual el contenido de C aéreo es mayor a las otras dos turberas, a la vez que la riqueza es la menor de las tres. Finalmente en la turbera T3 aparecen individuos arbóreos con mayor regularidad (Figura 6C), destacando el Ciprés de las Guaitecas (*Pilgerodendron uviferum*), los que interactúan con las coberturas de matorral abierto y semidenso adyacentes con especies como *B. patagonica* y *Philesia magellanica*, lo que permite que otras especies se desarrollen bajo su dosel, siendo un estado intermedio entre las turberas T1 y T2 respecto a la riqueza de plantas vasculares y el contenido de C aéreo.

No obstante, la relación inversa entre riqueza y C aéreo no es tan clara entre las turberas T2 y T3, donde si bien la riqueza de plantas mostró diferencias en su composición y número de especies, en el caso del contenido de C aéreo no se pudo establecer tal diferencia ($P = 0.566$ en la prueba de Wilcoxon), siendo necesario realizar otras investigaciones que se enfoquen en específico en las sucesiones de estos ecosistemas.

Un meta análisis realizado por Wang (2007) obtuvo resultados similares en distintos ecosistemas (no incluyendo turberas), encontrando que las poblaciones de especies de plantas vasculares contaban con mayor biomasa que las comunidades, habiendo una relación inversa entre el número de especies (o riqueza) y la biomasa, concordando con estudios posteriores que llegaron a los mismos resultados (Potter and Woodall, 2014). Sin embargo, Wu et al. (2014) obtuvieron resultados contrapuestos en ecosistemas de tipo forestal, donde la relación entre la riqueza de especies y biomasa fue positiva, aun cuando el aumento de la biomasa estaba más relacionado a otros factores ambientales. En la misma línea Schmid et al. (2009), indican que los efectos de la biodiversidad sobre las propiedades funcionales (entre las que figura la biomasa) difiere entre tipos de ecosistemas, pero no entre ecosistemas acuáticos o terrestres, no habiendo consenso sobre la relación entre la biodiversidad y la biomasa.

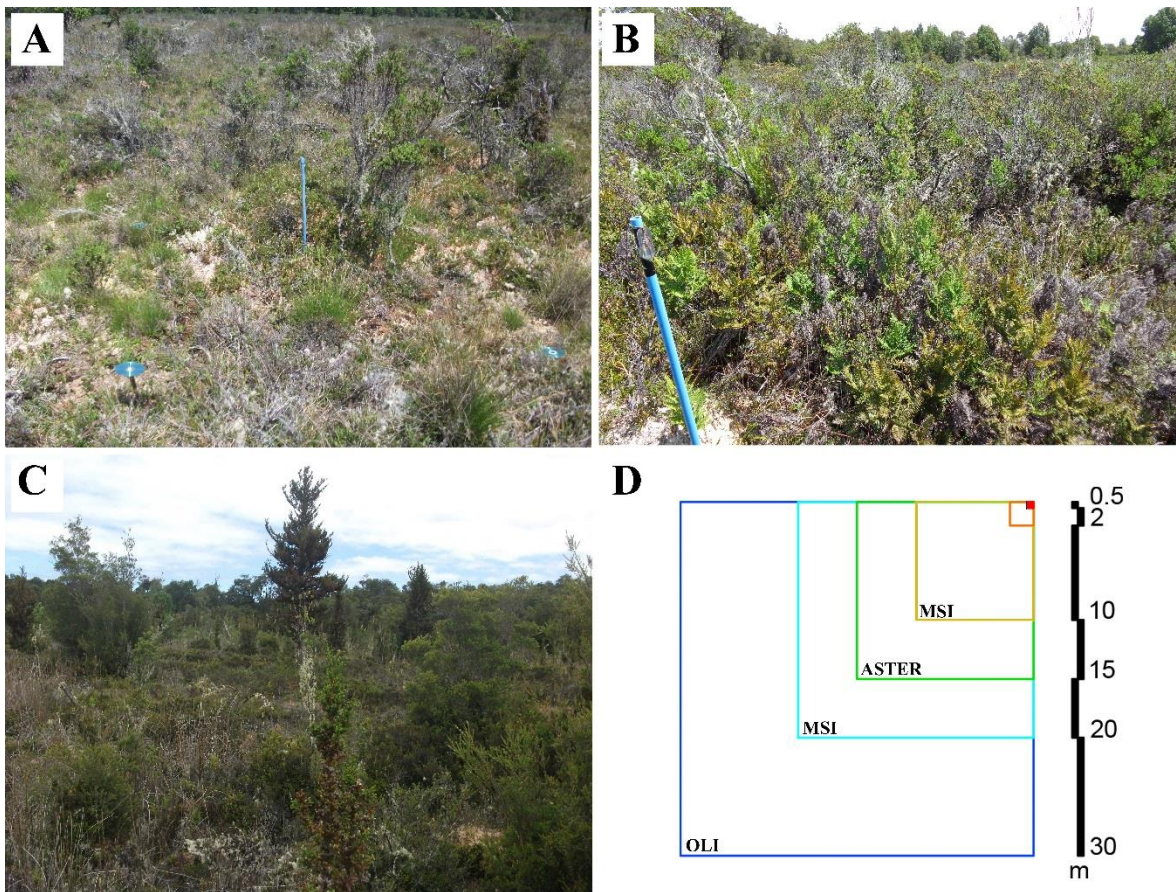


Figura 6. Fotos de las turberas desde T1 (A), T2 (B) y T3 (C). El panel D muestra la relación entre la superficie de los puntos de muestreo (C aéreo en rojo y riqueza en naranja) y los píxeles de las imágenes satelitales.

4.2. Variables predictivas seleccionadas y sus implicancias ecológicas

De las 20 variables espectrales utilizadas en los 12 modelos evaluados, 19 poseen la influencia del espectro infrarrojo, ya sea de forma directa mediante la utilización de una banda o métrica GLCM o a partir de su influencia dentro de un índice espectral. Lo anterior según Chuvieco (2008) se relaciona a que la zona del espectro electromagnético del infrarrojo es la más idónea para estudiar la vegetación, teoría que ha sido respaldada por diversos estudios que a partir de imágenes hiperespectrales han seleccionado y utilizado esta parte del espectro para discriminar entre especies de plantas (Schmidt and Skidmore, 2003; Vaiphasa et al., 2007; Adam y Mutanga, 2009). Al descomponer esta zona del espectro se obtiene que en el NIR se produce una alta reflectancia debido a la estructura interna de la hoja y sus procesos físicos, mientras que en el SWIR se produce una mayor absorción debido al contenido de moléculas de agua, por lo que se relaciona a la humedad de las coberturas vegetales (Kumar et al., 2001). Por otra parte el espectro visible es menos idóneo debido a que esta zona está dominada por el efecto de absorción de la clorofila y

otros pigmentos que son los responsables del color de las plantas, el cual generalmente no posee tanta variabilidad en las superficies vegetales en comparación al infrarrojo (Kumar et al., 2001). En la Figura 7 se observa que la zona del NIR es la que presenta mayor variabilidad en ecosistemas de turberas, seguido por el SWIR y el Rojo.

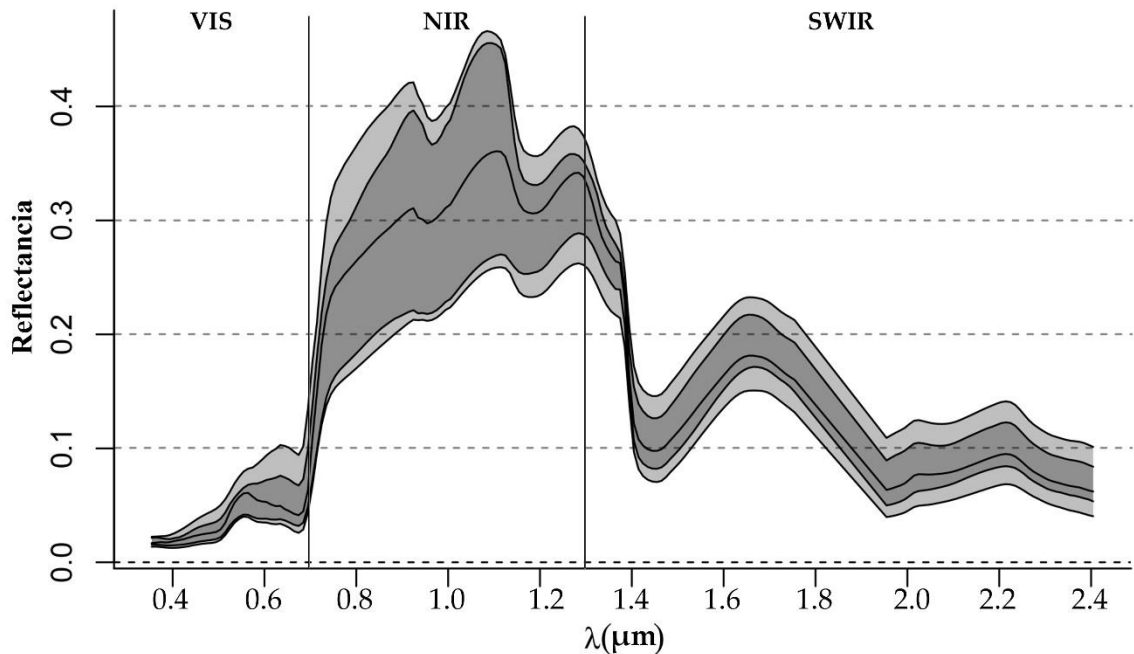


Figura 7. Firma espectral de la vegetación obtenida por espectroradiometría de campo. Se muestra la variación de las firmas espectrales de la vegetación de las 15 parcelas de muestreo de la turbera T2, representando con las líneas los percentiles 5, 25, 50, 75 y 95%. VIS = visible, NIR = infrarrojo cercano y SWIR = infrarrojo de onda corta.

Algunos autores exponen que la vegetación de los humedales y sus propiedades son más difíciles de estudiar y detectar que en los ecosistemas terrestres, debido a que la vegetación herbácea en los humedales presenta gran variabilidad espacial y espectral debido a los gradientes ambientales, que producen marcados ecotonos entre las unidades vegetacionales adyacentes (Zomer et al., 2008; Adam y Mutanga, 2009) y además la reflectancia del dosel de la vegetación de humedal suele ser muy similar entre coberturas, por lo que se confunden con el suelo subyacente, régimen hidrológico y el vapor atmosférico (Lin and Liqian, 2006). Esta combinación puede afectar la clasificación óptica en humedales que se encuentran inundados, resultando en una disminución de la reflectancia, especialmente en los canales del NIR y SWIR donde la absorción de agua es mayor (Silva et al., 2008). No obstante, la intensidad de dicho efecto puede estar determinado por la densidad vegetal o la estructura de la canopia (Jakubauskas et al., 2000). Sin embargo, las turberas de este estudio no se encontraban en condiciones de inundación a la hora de hacer el muestreo de campo ni de la toma de imágenes satelitales, por lo que este efecto podría no ser determinante ya que las coberturas se comportan más parecido a vegetación terrestre

(Figura 7), en la cual este efecto no representa un problema significativo (Jakubauskas et al., 2000; Adam et al., 2010).

4.2.1. Riqueza de plantas vasculares

Según Hobbs (1989), la percepción remota es eficaz para detectar tipos estructurales, así como la cantidad y el estado de la cubierta vegetal, pero es menos capaz de detectar diferentes especies o variación en la composición de comunidades vegetales a gran escala. Sin embargo existen estudios donde se modeló la riqueza de especies de plantas vasculares mediante sensores remotos a escala local alcanzaron buenos ajustes, con R^2 de hasta 0,79 usando imágenes QuickBird con 2.4 m de resolución espacial (Viedma et al., 2012), 0,64, usando imágenes ASTER (Mutowo and Murwira, 2012) y 0,6, usando Landsat 8 (Cabezas et al., 2016), valores similares a los obtenidos en este estudio.

Por otro lado, seis de las ocho variables espectrales seleccionadas para predecir la riqueza resultaron ser variables texturales, concordando con lo reportado por Stein et al. (2014), quienes establecen que el factor más importante relacionado a la riqueza de plantas es la heterogeneidad espectral. En este sentido Rocchini et al. (2010) encontraron que la heterogeneidad espectral puede predecir de mejor manera la diversidad biológica, fundamentado en la hipótesis de la variación espectral (SVH) (Palmer et al., 2002), que indica que a mayor heterogeneidad espectral habrá una correlación positiva con la riqueza de especies, dado que zonas ambientalmente más heterogéneas tienen la capacidad de albergar más nichos ecológicos, con mayor probabilidad de encontrar un número más alto de especies. Esta teoría ha sido exitosamente probada en varias taxas, entre las que figuran las plantas vasculares (Gould, 2000; Foody and Coutler, 2006; Levin et al., 2007; Cabezas et al., 2016) y es respaldada por los resultados de este estudio.

Por su parte, la incorporación de la variable Id.turbera mejoró en la mayoría de los casos las predicciones de riqueza, lo que podría estar asociado a la disimilitud que presentan las turberas entre sí en cuanto a su estructura horizontal, entendida como la distribución de las especies en el espacio, pudiéndose interpretar como distintos estados sucesionales (Diaz et al., 2007), cuyo clímax en este caso, sería un ecosistema boscoso.

4.2.2. Carbono aéreo

En el caso de los modelos de C aéreo, únicamente cuatro de las doce variables predictivas resultaron ser texturales, muy por debajo de los modelos de riqueza, donde el mejor modelo de este estudio no utilizó estas variables en sus predicciones. Esta situación es contraria a lo descrito por Eckert (2012) quien realizó una comparación entre variables espectrales y texturales para predecir la biomasa y C aéreo en un bosque tropical mediante datos de alta resolución WorldView-2, obteniendo que las variables texturales lograban estimaciones más precisas que las variables espectrales, con R^2 de 0,58 v/s 0,41 respectivamente. La diferencia obtenida en este estudio podría estar atribuida a las resoluciones espaciales de los sensores usados en este estudio y los datos de WorldView-2 (2 m) o bien al tipo de ecosistema estudiado, no permitiendo homologar dichos resultados a este estudio. En un estudio similar, Lu (2005) obtuvo que tanto las variables espectrales como sus métricas

texturales eran similarmente precisas para estimar la biomasa aérea usando datos Landsat-TM (30 m) en un bosque amazónico, con R^2 de 0,74 v/s 0,70 respectivamente, donde el mejor modelo presentado resultaba de la combinación de ambas variables (R^2 de 0,78). Por otro lado, Mutanga et al. (2012) y Dube y Mutanga (2016) desarrollaron modelos predictivos de C y biomasa aérea obteniendo resultados similares sin utilizar variables texturales (R^2 de 0,79 y 0,66 respectivamente), indicando que el número de bandas representativas de la zona del NIR es lo más relevante en este tipo de modelos, lo que es respaldado por Kelsey y Neff (2014).

De lo anterior se desprende que la selección de las mejores variables predictivas depende en gran medida de las condiciones locales del estudio, como el tipo de ecosistema y su características específicas, no habiendo consenso sobre cuáles son las más adecuadas para realizar estimaciones de C o biomasa aérea. Otro factor que toma relevancia es el tipo de sensores remotos usados y sus características espaciales y espectrales, y de acuerdo con Lu (2005), de los algoritmos de selección utilizados. Con todo lo anterior, los resultados de este estudio indican que las variables espectrales permiten representar más adecuadamente el C aéreo en ecosistemas de turbera antropogénica, aun cuando el uso de variables texturales puede ayudar en algunos casos a mejorar dichas predicciones.

Respecto a la incorporación de la variable Id.turbera, esta no apareció en el ranking RFE como una de las variables más importantes, quedando fuera de la modelación de C aéreo, lo que puede estar relacionado a los resultados obtenidos en el test de Wilcoxon, donde no se encontró que existiera diferencias significativas entre todos los pares de turberas.

4.3. Comparación de sensores satelitales

Al comparar la modelación de riqueza y C aéreo mediante los distintos sensores, MSI a 20 m resultó ser el de mejores ajustes y menor error, seguidos por OLI y ASTER. Esta diferencia puede ser atribuida a la resolución espectral de los sensores, ya que el espectro infrarrojo en los modelos aparece como el factor más relevante en las predicciones y el sensor MSI es el que posee mejor representatividad en esta zona, al contar con cuatro bandas entre los 0,698 y 0,875 μm más una que va entre los 0,785-0,900 μm (Cuadro 1).

Junto con lo anterior, es necesario mencionar que un factor que pudo influir en los resultados finales, es el hecho de haber utilizado una sola imagen con el sensor MSI que fuera representativa de dos fechas de muestreo con un intervalo de dos años de diferencia. Esto pudo haber generado un error adicional, ya que idealmente la adquisición de imágenes deben ser al mismo tiempo que los muestreos de campo. Sin embargo, si el muestreo de campo se realiza en condiciones estacionales o de crecimiento similares hará que dicho error sea menor (Bradley, 2014). Dado lo anterior, los muestreos se realizaron en la misma fecha del año, reduciendo el efecto estacional; además ambos años presentaron similares niveles de precipitación y temperaturas medias mensuales (Figura 8). Otro factor a considerar desde el punto de vista de la variación que presentó la vegetación ambos años, es que las tres turberas estudiadas no presentan grandes perturbaciones externas, por lo que las alteraciones que estas puedan tener van a depender principalmente de las condiciones

atmosféricas e hidrológicas. Un estudio de Nagendra et al. (2010) utilizó la misma metodología de muestreo, el cual se realizó en varias campañas entre los años 2003 y 2005, utilizando una sola imagen del año 2000, obteniendo buenos resultados pese al desfase temporal entre el muestreo de campo y la adquisición de imágenes.

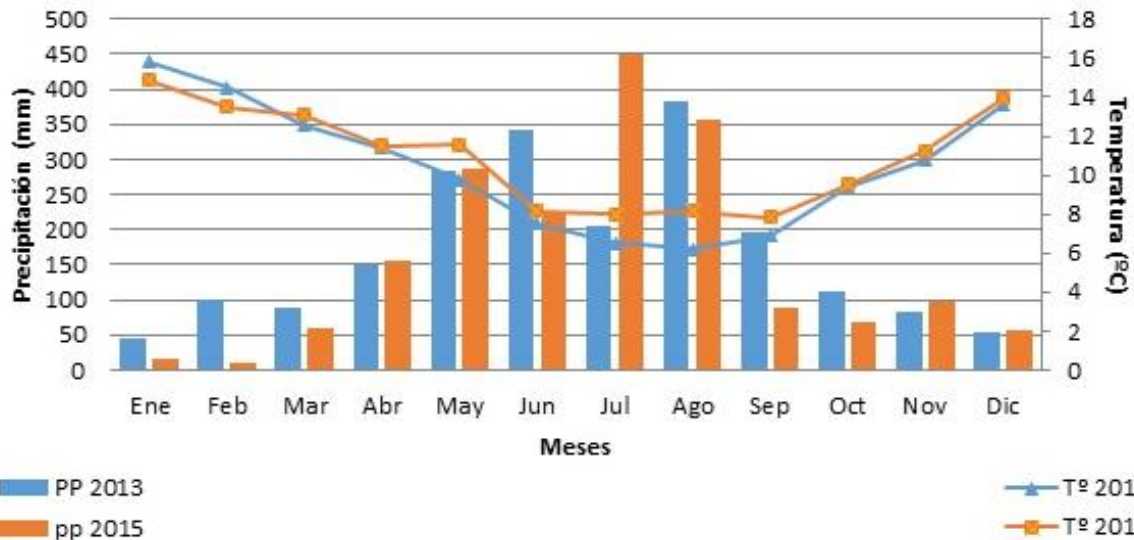


Figura 8. Precipitación (pp) y temperatura (T°) media mensual de la estación meteorológica ubicada en la EBSD.

4.3.1. Riqueza de plantas vasculares

Lo anterior va en la misma línea que los resultados obtenidos por Adam y Mutanga (2009), quienes realizaron una discriminación entre especies en un humedal pantanoso usando datos hiperespectrales, cuyos resultados indicaron que las bandas entre los 0,745 y 0,989 μm eran las más idóneas para discriminar entre especies, similar a lo reportado por Vaiphasa et al. (2007), en un ecosistema de manglar. En ese sentido, Feilhauer et al. (2013) indican que a mayor cobertura en los canales del NIR y SWIR, mejor será la discriminación de la composición florística, resultados que son respaldados por este estudio.

Respecto a las variables texturales, la resolución espacial de la imagen es importante para su análisis, ya que dependiendo del objeto de estudio se pueden identificar individuos o una agregación de ellos (Bradley, 2014). Según Rocchini et al. (2016), idealmente el tamaño del pixel debe ser como mínimo del mismo tamaño que las unidades de muestreo, al menos cuando se calcula la heterogeneidad espectral para estimar la diversidad local de especies, como en este estudio, indicando además que una baja resolución espacial puede afectar la captura de la heterogeneidad real debido a que ocurre un proceso de suavizado que dificulte la detección de patrones a una escala más fina. No obstante, cuando el tamaño del pixel usado tiene dimensiones muy reducidas (1-5 m), la sombra puede crear una alta heterogeneidad espacial que conduzca a generar más ruido en vez de aportar información relevante (Nagendra y Rocchini, 2008), teoría que es respaldada por Cabezas et al. (2016) en una turbera antropogénica.

Es así como en este estudio se utilizaron ventanas móviles de 3×3 píxeles para calcular las variables texturales, independiente de la resolución espacial de cada sensor, por lo que el efecto de suavizado difiere entre los sensores, pudiendo afectar en el cálculo de la heterogeneidad real del ecosistema. No obstante, a las tres escalas utilizadas (~10-30 m) se consiguieron precisiones similares en los modelos, lo que podría ser atribuido a que los sensores con menor resolución espacial poseen mayor resolución espectral en la zona del NIR y SWIR, generando una compensación entre ambas resoluciones, o bien podría deberse a que la variación de la riqueza en el espacio no es tan abrupta como para influir en los resultados.

4.3.1. Carbono aéreo

Según Kelsey y Neff (2014) la diferencia entre el tamaño de los puntos de muestreo y la resolución espacial del pixel utilizado incorpora un factor de error que puede ser importante, debido que el punto de muestreo no necesariamente va a representar la agregación de C dentro del pixel (ver Figura 6D), situación que en este estudio es bastante notoria, debido a la alta heterogeneidad que presentan las tres turberas analizadas, en las cuales existen cambios abruptos en las coberturas, encontrando superficies de matorral, helechos, juncos, herbáceas y musgo agregadas en espacios reducidos, menores a los tamaños analizados en este estudio, además de las diferencias existentes dentro de las mismas coberturas vegetales.

Lo anterior puede ser una de las razones de que los modelos generados alcanzaron ajustes tan bajos, a la vez que los errores y sesgos fueron altos, debido a que el porcentaje de la superficie del pixel cubierta por el cuadrante de muestreo es menor a uno, porcentaje que además no necesariamente es representativo de la superficie total de estos. Siguiendo lo anterior, se podría inferir que con las imágenes de mayor resolución espacial debieran obtenerse las mejores predicciones, sin embargo este no es el caso, siendo el sensor MSI a 20 m con el que se llegó a mejores ajustes, lo que presumiblemente está relacionado a que la variabilidad del ecosistema en 100 m² (pixel de 10×10 m) no es muy diferente a la variabilidad en 900 m² (pixel de 30×30 m), o a que la resolución espectral de los sensores remotos es más relevante que la espacial, o bien podría darse ambas situaciones a la vez.

Otra razón que pudiera haber afectado en el bajo ajuste de los modelos es el tipo de datos usados, ya que según Zolkos et al. (2013), los datos hiperespectrales entregan un nivel de información muy bajo en comparación a los datos LiDAR en la modelación de biomasa, lo que es respaldado por Fassnacht et al. (2014) quienes compararon modelos generados con datos LiDAR e hiperespectrales, obteniendo bajos rendimientos en los modelos mediante los datos hiperespectrales, con R² de 0,20 con modelos lineales y 0,35 con RF aproximadamente, similares a los mejores modelos obtenidos en este estudio.

4.4. Comparación de modelos

En la generación de los modelos predictivos el número de variables utilizado se encontró entre dos y cuatro, dentro del rango recomendado por Hair et al. (2009), las que de acuerdo al número de observaciones de este estudio ($n=45$) debían encontrarse alrededor de tres, con una tendencia de los GLM a requerir menos variables que RF.

Los mejores GLM y los mejores RF en general no resultaron ser significativamente diferentes en ninguno de los parámetros evaluados, salvo en ASTER en los modelos de riqueza, donde RF obtuvo mayor precisión y menor error y en MSI en los modelos de C donde se presentó la misma situación. Pese a lo anterior, en los modelos de riqueza los GLM resultaron ser ligeramente superiores a RF (Cuadro 6), situación contraria a lo encontrado en los modelos de C, donde RF se comportó mejor en todos los casos (Cuadro 8).

La leve superioridad de los GLM encontrada en los modelos de riqueza está presumiblemente conectada a las características estadísticas de los “count data”, que en general poseen una distribución no simétrica del error (Lopatin et al., 2016). Dado lo anterior, la distribución de Poisson resultó ser en este caso la más adecuada para predecir la riqueza, concordando con lo establecido por otros autores (Ciampi, 1991; Cook and Lawless, 2007), aun cuando la distribución binomial también se considera útil y ha sido utilizada (Lopatin et al., 2016). Lo anterior fue bastante diferente en los modelos de riqueza generados a partir del sensor ASTER, donde RF fue significativamente superior, con mayor R^2 y menor RRMSE (Cuadro 6), situación que también se dio en todos los modelos de C aéreo, donde el rendimiento de los GLM fueron más bajos que RF. En ese sentido, Fassnacht et al. (2014) compararon diferentes métodos de modelación de biomasa, entre los que figuraban los modelos lineales y RF, obteniendo los mejores resultados mediante RF, por ser un algoritmo con mayor adaptabilidad y flexibilidad a los datos de entrada, mientras que con los modelos lineales se obtuvieron los peores rendimientos, lo que fue asociado a que la relación entre los datos hiperespectrales y la biomasa observada probablemente no son lineales y por lo tanto los modelos entregan bajos ajustes, situación que se dio en este estudio.

Con lo anterior, los métodos de ML, como RF, son conocidos por poseer mayor precisión y se prefieren por sobre los métodos paramétricos, sin embargo sus predicciones de “black-box” son difíciles de interpretar, a diferencia de los GLM donde las relaciones son lineales y su interpretación es más directa (Song et al., 2013), ventaja que es necesario tomar en cuenta en los casos donde el objetivo es encontrar las relaciones físicas entre los predictores y las variables de respuesta. Otro problema con RF es que la aplicación de submuestras en el algoritmo puede resultar en una varianza considerable de las estimaciones, más aún cuando el número de muestras es pequeño (Latifi et al., 2012), lo que puede volverse un problema mayor si consideramos el proceso de validación mediante la prueba de bootstrap, donde son seleccionados subconjuntos que reducen aún más los datos de entrada del modelo. Lo anterior puede ser una de las razones que explique el comportamiento bimodal del RRMSE y sesgo de los modelos RF en las predicciones de C aéreo (ver Figura 4), ya

que la reducción de los datos de entrada pueden volver más sensible al modelo, llevándolo a comportarse de manera variable dependiendo el set de datos seleccionado en la validación.

Finalmente cabe destacar que si bien la superficie de las turberas estudiadas son reducidas (superficie total de 22,3 ha), estas demostraron presentar características de disimilitud entre sí que las vuelven un objeto representativo de estudio de las turberas sin manejo productivo presentes en la zona. Sin embargo los modelos generados en este estudio no son necesariamente representativos de todas las turberas de la zona, debido a que según Cabezas et al. (2015), las turberas destinadas a la conservación presentan características significativamente diferentes en cuanto a los reservorios de C aéreo y a la riqueza de especies vasculares respecto a las turberas manejadas productivamente, siendo necesario realizar otro tipo de estudios que evalúen el nivel de ajuste del modelo en ese tipo de turberas.

5. CONCLUSIONES

La riqueza y el C aéreo pueden ser considerados como indicadores de estado y dar una aproximación sobre las características sucesionales en el que se encuentran las turberas antropogénicas.

La comparación entre RF y GLM mostró que ambos métodos alcanzan resultados similares en cuanto a la precisión de los modelos y su error asociado, al no presentar diferencias significativas en la mayoría de los casos evaluados. Es así como los resultados indican que ambos métodos son útiles en la modelación de riqueza de especies vasculares en turberas antropogénicas, a la vez que para modelar el C aéreo, presentaron en general resultados insatisfactorios, siendo necesario ampliar este estudio o probar nuevas metodologías.

La comparación entre sensores mostró que con MSI se obtienen mejores resultados que con OLI y ASTER. Esto presumiblemente se debe a la resolución espectral del sensor, que posee mayor detalle en la zona del espectro infrarrojo cercano, donde la variabilidad espectral es mayor en este tipo de ecosistema, permitiendo detectar con mayor precisión las diferentes coberturas vegetacionales y sus propiedades. Por su parte, la resolución espacial no mostró ser un factor determinante en la modelación a las escalas dadas, al obtenerse los mejores resultados a una resolución intermedia entre los sensores probados. Es así como este estudio mostró el potencial de los sensores de uso libre en el modelamiento y predicción de variables vegetacionales en ecosistemas de turbera.

Las variables texturales ayudan a mejorar las predicciones, aun cuando no necesariamente son las más explicativas de la varianza de los modelos, siendo más útiles en los modelos de riqueza que en los de C aéreo. La incorporación de una variable categórica que diferenciara entre las distintas turberas ayudó a mejorar las predicciones de riqueza, ya que las condiciones ambientales de las turberas evaluadas presentan características de disimilitud que fue necesario incorporar en el análisis, situación que no se dio en el C aéreo.

Finalmente los hallazgos de este estudio apoyan el potencial de los datos de la percepción remota para modelar y mapear la distribución de la riqueza de plantas vasculares en turberas antropogénicas, ampliando el conocimiento sobre este campo de la modelación predictiva sobre un ecosistema que ha sido escasamente abordado desde esta perspectiva. Sin embargo, en el caso del C aéreo los resultados no fueron tan satisfactorios en términos de precisión y error, lo que presumiblemente está conectado a la alta variabilidad espacial que poseen las turberas estudiadas, dificultando su modelación a las resoluciones espaciales probadas.

8. BIBLIOGRAFÍA

- Adam, E. and O. Mutanga. 2009. Spectral discrimination of papyrus vegetation (*Cyperus papyrus* L.) in swamp wetlands using field spectrometry. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64:612–620.
- Adam, E.; O. Mutanga and D. Rugege. 2010. Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: A review. *Wetlands Ecology and Management*, 18(3): 281–296.
- Aslan, A.; A. Rahman; M. Warren and S. Robeson. 2016. Mapping spatial distribution and biomass of coastal wetland vegetation in Indonesian Papua by combining active and passive remotely. *Remote Sensing of Environment*, 183: 65–81.
- Baig, M.H.A.; L. Zhang; T. Shuai and Q. Tong. 2014. Derivation of a tasseled cap transformation based on Landsat 8 at-satellite reflectance. *Remote Sensing Letters*, 5(5): 423–431.
- Bradley, B.A. 2014. Remote detection of invasive plants: a review of spectral, textural and phenological approaches. *Biological Invasions*, 16: 1411–1425.
- Breiman, L. 2001. Random forests. *Machine Learning*, 45: 5–32.
- Broge, N. and E. Leblanc. 2000. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment*, 76: 156–172.
- Bubier, J.L.; T.R. Moore y G. Crosby. 2006. Fine-scale vegetation distribution in a cool temperate peatland. *Canadian Journal of Botany*, 84(6): 910–923.
- Bullock, A. and M. Acreman. 2003. The role of wetlands in the hydrological cycle. *Hydrology and Earth System Sciences*, 7: 358–389.
- Cabezas, J.; M. Galleguillos and J.F. Perez–Quezada. 2016. Predicting vascular plant richness in a heterogeneous wetland using spectral and textural features and a random forest algorithm. *IEEE Geoscience and Remote Sensing Letters*, 13: 646–650.
- Cabezas, J.; M. Galleguillos; A. Valdés; J.P. Fuentes; C. Pérez and J.F. Perez–Quezada. 2015. Evaluation of impacts of management in an anthropogenic peatland using field and remote sensing data. *Ecosphere*, 6(12): 282.
- Ceballos, A.; J. Hernández; P. Corvalán and M. Galleguillos. 2015. Comparison of airborne LiDAR and satellite hyperspectral remote sensing to estimate vascular plant richness in deciduous mediterranean forests of central Chile. *Remote Sensing*, 7: 2692–2714.
- Chuvieco, E. 2008. Teledetección ambiental: La observación de la tierra desde el espacio. 3 ed. Barcelona: Ariel S.A. 594p.

Ciampi, A. 1991. Generalized regression trees. *Computer Statistics Data Analysis*, 12: 57–78.

Cliff, A. y K. Ord. 1972. Testing for spatial autocorrelation among regression residuals. *Geographical Analysis*, 4(3): 267–284.

Cole, B.; J. McMorro and M. Evans. 2014. Spectral monitoring of moorland plant phenology to identify a temporal window for hyperspectral remote sensing of peatland. *ISPRS Journal of Photogrammetry and Remote Sensing*, 90: 49–58.

Cook, R.J. and J.F. Lawless. 2007. The statistical analysis of recurrent events. New York: Springer. 404p.

De Araujo Barbosa, C.C.; P.M. Atkinson and J.A. Dearing. 2015. Remote sensing of ecosystem services: A systematic review. *Ecological Indicators*, 52: 430-443.

Díaz, M.F.; J. Larrain; G. Zegers and C. Tapia. 2008. Caracterización florística e hidrológica de turberas de la Isla Grande de Chiloé, Chile. *Revista Chilena de Hististoria Natatural*, 81: 455–468.

Díaz, M.F.; S. Bigelow and J.J. Armesto. 2007. Alteration of the hydrologic cycle due to forest clearing and its consequences for rainforest succession. *Forest Ecology and Management*, 244: 32–40.

Dube, T. and O. Mutanga. 2016. The impact of integrating WorldView-2 sensor and environmental variables in estimating plantation forest species aboveground biomass and carbon stocks in uMgeni Catchment, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing*, 119: 415–425.

Duffy, E. 2009. Why biodiversity is important to the functioning of real-world ecosystems. *Frontiers in Ecology and the Environment*, 7: 437–444.

Eckert, S. 2012. Improved forest biomass and carbon estimations using texture measures from WorldView-2 satellite data. *Remote Sensing*, 4: 810–829.

Efron, B. 1979. Bootstrap methods: Another look at the Jackknife. *Annals of Statistics*, 7: 1–26.

Evans, T.L.; M. Costa; W.M. Tomas and A.R. Camilo. 2014. Large-scale habitat mapping of the Brazilian Pantanal wetland: A synthetic aperture radar approach. *Remote Sensing of Environment*, 155: 89–108.

Fassnacht, F.; C. Neumann; M. Förster; H. Buddenbaum; A. Ghosh; A. Clasen; P.K. Joshi and B. Koch. 2014b. Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three central European test sites. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*, 7: 2547–2561.

Fassnacht, F.; F. Hartig; H. Latifi; C. Berger; J. Hernandez; P. Corvalan and B. Koch. 2014a. Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sensing of Environment*, 154: 102–114.

Fava, F.; G. Parolo; R. Colombo; F. Gusmeroli; G. Della Marianna; A. Monteiro and S. Bocchi. 2010. Fine-scale assessment of hay meadow productivity and plant diversity in the European Alps using field spectrometric data. *Agriculture, Ecosystems & Environment*, 137: 151–157.

Feilhauer, H. and S. Schmidtlein. 2009. Mapping continuous fields of forest alpha and beta diversity. *Applied Vegetation Science*, 12: 429–439.

Feilhauer, H.; F. Thonfeld; U. Faude; K. He; D. Rocchini and S. Schmidtlein. 2013. Assessing floristic composition with multispectral sensors: A comparison based on monotemporal and multiseasonal field spectra. *International Journal of Applied Earth Observation and Geoinformation*, 21: 218–229.

Fernandez-Manso, A.; O. Fernandez-Manso and C. Quintano. 2016. SENTINEL-2A red-edge spectral indices suitability for discriminating burn severity. *International Journal of Applied Earth Observation and Geoinformation*, 50: 170–175.

Foody, G.M. and M.E.J. Cutler. 2003. Tree biodiversity in protected and logged Bornean tropical rain forests and its measurement by satellite remote sensing. *Journal of Biogeography*, 30: 1053–1066.

Foody, G.M. and M.E.J. Cutler. 2006. Mapping the species richness and composition of tropical forests from remotely sensed data with neural networks. *Ecological Modelling*, 195: 37–42.

Gitelson, A.A.; Y.J. Kaufman and M.N. Merzljak. 1996. Use of a green channel in remote sensing of global vegetation from EOS–MODIS. *Remote Sensing of Environment*, 55(3): 289–298.

Gore, A.J.P. 1983. Ecosystems of the world. Mires: Swamp, bog, fen and moor 4A, General studies. Amsterdam, The Netherlands: Elsevier. 440p.

Gould, W. 2000. Remote sensing of vegetation, plant species richness, and regional biodiversity hot spots. *Ecological Applications*, 10: 1861–1870.

Guyon, I.; J. Weston and S. Barnhill. 2002. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46: 389–422.

Haboudane, D.; J.R. Miller; E. Pattey; P.J. Zarco-Tejada and I.B. Strachan. 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 90: 337–352.

Hair, J.F.J.; W.C. Black; B.J. Babin and R.E. Anderson. 2009. Multivariate data analysis: A global perspective. 7 ed. Upper Saddle River: Prentice Hall. 800p.

Hestir, E.L.; S. Khanna; M.E. Andrew; M.J. Santos; J.H. Viers; J.A. Greenberg; S.S. Rajapakse and S.L. Ustin. 2008. Identification of invasive vegetation using hyperspectral remote sensing in the California Delta ecosystem. *Remote Sensing of Environment*, 112: 4034–4047.

Hobbs, R.J. 1989. Remote sensing and temporal dynamics of vegetation (cap. 9, pp. 203–219). En: Hobbs, R.J. and H.A. Mooney (ed.). *Remote Sensing of Biosphere Functioning*. New York: Springer-Verlag. 306p.

Huete, A.; H.Q. Liu; K. Batchily and W. van Leeuwen. 1997. A Comparison of vegetation indices over a global set of TM images for EOS–MODIS. *Remote Sensing of Environment*, 59: 440–451.

Huete, A.; K. Didan; T. Miura; E. Rodriguez; X. Gao and L. Ferreira. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83: 195–213.

Jaccard, P. 1902. Lois de distribution florale. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 38: 67–130.

Jakubauskas, M.; K. Kindscher; A. Fraser; D. Debinski and K.P. Price. 2000. Close-range remote sensing of aquatic macrophyte vegetation cover. *International Journal of Remote Sensing*, 21(8): 3533–3538.

Karjalainen, M.; V. Kankare; M. Vastaranta; M. Holopainen and J. Hyypä. 2012. Prediction of plot-level forest variables using TerraSAR-X stereo SAR data. *Remote Sensing of Environment*, 117: 338–347.

Keddy, P.A. 2010. *Wetland ecology: Principles and conservation*. 2 ed. Ney York: Cambridge University Press. 497p.

Kelsey, K. and J. Neff. 2014. Estimates of aboveground biomass from texture analysis of Landsat imagery. *Remote Sensing*, 6:6407–6422.

Khanna, S.; A. Palacios-Orueta; M.L. Whiting; S.L. Ustin; D. Riaño and J. Litago. 2007. Development of angle indexes for soil moisture estimation, dry matter detection and land–cover discrimination. *Remote Sensing of Environment*, 109 (2): 154–165.

Kuhn, M. 2016. *Caret: Classification and Regression Training*. R package version 6.0–71.

Kumar, L.; K.S. Schmidt; S. Dury and A.K. Skidmore. 2001. Imaging Spectrometry and vegetation Science. (cap. 5, pp. 111–155). En: Van Der Meer, F.D. and S.M. De Jong. (ed.). *Imaging spectrometry: Basic principles and prospective applications*. The Netherlands: Springer. 403p.

Latifi, H. and B. Koch. 2012. Evaluation of most similar neighbour and random forest methods for imputing forest inventory variables using data from target and auxiliary stands. *International Journal of Remote Sensing*, 33(21): 6668–6694.

Latifi, H.; F. Fassnacht and B. Koch. 2012. Forest structure modeling with combined airborne hyperspectral and LiDAR data. *Remote Sensing of Environment*, 121: 10–25.

Laurin, G.; N. Puletti; W. Hawthorne; V. Liesenberg; P. Corona; D. Papale; Q. Chen and R. Valentini. 2016. Discrimination of tropical forest types, dominant species, and mapping of functional guilds by hyperspectral and simulated multispectral Sentinel-2 data. *Remote Sensing of Environment*, 176: 163–176.

Lee, K.H. and R.S. Lunetta. 1996. Wetland detection methods. (cap. 18, pp. 249–284). En: Lyon J.G. and J. McCarthy (ed.). Wetland and environmental application of GIS. New York: Lewis Publishers.

Levin, N.; A. Shmida; O. Levanoni; H. Tamari and S. Kark. 2007. Predicting mountain plant richness and rarity from space using satellite-derived vegetation indices. *Diversity and Distributions*, 13: 692–703.

Liaw, A. and M. Wiener. 2002. Classification and regression by randomForest. R News, 2, 18–22.

Lin, Y. and Z. Lique. 2006. Identification of the spectral characteristics of submerged plant *Vallisneria spiralis*. *Acta Ecologica Sinica*, 26:1005–1011.

Ling, C.; H. Sun; H. Zhang; H. Lin; H. Ju and H. Liu. 2014. Study on above-ground biomass estimation of East Dong Ting lake wetland based on WorldView-2 data. *2014 Third International Workshop on Earth Observation and Remote Sensing Applications (EORSA)* Jun. 2014: 428–432.

Lopatin, J.; K. Dolos; H.J. Hernández; M. Galleguillos and F.E. Fassnacht. 2016. Comparison generalized linear model and random forest to model vascular plant species richness using LiDAR data in a natural forest in central Chile. *Remote Sensing of Environment*, 173: 200–210.

Lopatin, J.; M. Galleguillos; F.E. Fassnacht; A. Ceballos and J. Hernández. 2015. Using a multistructural object-based LiDAR approach to estimate vascular plant richness in Mediterranean forests with complex structure. *IEEE Geoscience and Remote Sensing Letters*, 12(5): 1008–1012.

Lu, D. 2006. The potential and challenge of remote sensing-based biomass estimation. *International Journal of Remote Sensing*, 27(7): 1297–1328.

Luo, S.; C. Wang; F. Pan; X. Xi; G. Li and S. Nie. 2015. Estimation of wetland vegetation height and leaf area index using airborne laser scanning data. *Ecological Indicator*, 48: 550–559.

Manuri, S.; H.E. Andersen; R. McGaughey and C. Brack. 2017. Assessing the influence of return density on estimation of lidar-based aboveground biomass in tropical peat swamp forests of Kalimantan, Indonesia. *International Journal of Applied Earth Observation and Geoinformation*, 56: 24–35.

Morel, A.C.; J.B. Fisher and Y. Malhi. 2012. Evaluating the potential to monitor aboveground biomass in forest and oil palm in Sabah, Malaysia, for 2000–2008 with Landsat ETM plus and ALOS-PALSAR. *International Journal of Remote Sensing*, 33: 3614–3639.

Mutanga, O.; E. Adam and M.A. Cho. 2012. High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. *International Journal of Applied Earth Observation and Geoinformation*, 18: 399–406.

Mutowo, G. and A. Murwira. 2012. Relationship between remotely sensed variables and tree species diversity in savanna woodlands of Southern Africa. *International Journal of Remote Sensing*, 33(20): 6378–6402.

Muukkonen, P. and J. Heiskanen. 2007. Biomass estimation over a large area based on standwise forest inventory data and ASTER and MODIS satellite data: A possibility to verify carbon inventories. *Remote Sensing of Environment*, 107: 617–624.

Nagendra, H. and D. Rocchini. 2008. High resolution satellite imagery for tropical biodiversity studies: the devil is in the detail. *Biodiversity and Conservation*, 17: 3431–3442.

Nagendra, H.; D. Rocchini; R. Ghate; B. Sharme and S. Pareeth. 2010. Assessing plant diversity in a dry tropical forest: Comparing the utility of Landsat and Ikonos satellite images. *Remote Sensing*, 2: 478–496.

Oldeland, J.; D. Wesuls; D. Rocchini; M. Schmidt and N. Jurgens. 2010. Does using species abundance data improve estimates of species diversity from remotely sensed spectral heterogeneity?. *Ecological Indicator*, 10: 390–396.

Palmer, M.; G. Earls; B. Hoagland; W. White and T. Wohlgemuth. 2002. Quantitative tools for perfecting species lists. *Environmetrics* 13(2); 121–259.

Parish, F.; A. Sirin; D. Charman; H. Joosten; T. Minayeva; M. Silvius and L. Stringer. 2008. Assessment on peatlands, biodiversity and climate change: Main report. Wageningen, Paises Bajos: Global Environment Centre, Kuala Lumpur and Wetlands International, 179p.

Posada, D. and T.R. Buckley. 2004. Model selection and model averaging in phylogenetics: Advantages of Akaike information criterion and bayesian approaches over likelihood ratio tests. *Systematic Biology*, 53: 793–808.

Potter, K. and C. Woodall. 2014. Does biodiversity make a difference? Relationships between species richness, evolutionary diversity, and aboveground live tree biomass across U.S. forests. *Forest Ecology and Management*, 321: 117–129.

Proisy, C.; P. Coutron and F. Fromard. 2007. Predicting and mapping mangrove biomass from canopy grain analysis using Fourier-based textural ordination of IKONOS images. *Remote Sensing of Environment*, 109: 379–392.

Qi, J.; A. Chehbouni; A.R. Huete; Y.H. Kerr and S. Sorooshian. 1994. A modified soil adjusted vegetation index. *Remote Sensing of Environment*, 48(2): 119–126.

Rocchini, D.; Boyd, D.S.; Féret, J.B.; Foody, G.M.; He, K.S.; Lausch, A.; Nagendra, H.; Wegmann, M. and N. Pettorelli. 2016. Satellite remote sensing to monitor species diversity: potential and pitfalls. *Remote Sensing in Ecology and Conservation*, 2: 25–36.

Rocchini, D.; N. Balkenhol; G.A. Carter; G.M. Foody; T.W. Gillespie; K.S. He; S. Kark; N. Levin; K. Lucas; M. Luoto; H. Nagendra; J. Oldeland; C. Ricotta; J. Southworth and M. Neteler. 2010. Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. *Ecological Information*, 5(5): 318–329.

Savage, S.L.; R.L. Lawrence and J.R. Squires. 2015. Predicting relative species composition within mixed conifer forest pixels using zero-inflated models and Landsat imagery. *Remote Sensing of Environment*, 171: 326–336.

Schmid, B.; P. Balvanera; J. Godbold and M. Solan. 2009. Consequences of species loss for ecosystem functioning: Meta-analyses of data from biodiversity experiments. En: Naeem, S.; E. Bunker; A. Hector; M. Loreau and C. Perrings. (ed.). *Biodiversity, ecosystem functioning, and human wellbeing: an ecological and economic perspective*. Oxford Univ. Press, pp. 14–29.

Schmidt, K.S. and A.K. Skidmore. 2003. Spectral discrimination of vegetation types in a coastal wetland. *Remote Sensing of Environment*, 85 (1): 92–108.

Shapiro, S. S. y M. B. Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3): 591–611.

Sibanda, M.; O. Mutanga and R. Mathieu. 2016. Examining the potential of Sentinel-2 MSI spectral resolution in quantifying above ground biomass across different fertilizer treatments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 110: 55–65.

Silleos, N.G.; T.K. Alexandridis; I.Z. Gitas and K. Perakis. 2006. Vegetation indices: Advances made in biomass estimation and vegetation monitoring in the last 30 years. *Geocarto International*, 21(4): 21–28.

Silva, T.S.F.; M.P.F. Costa; J.M. Melack and E. Novo. 2008. Remote sensing of aquatic vegetation: Theory and applications. *Environment Monitoring and Assessment*, 140: 131–145.

Song, L.; P. Langfelder and S. Horvath. 2013. Random generalized linear model: a highly accurate and interpretable ensemble predictor. *BMC Bioinformatics*, 14: 1–22.

Stein, A.; K. Gerstner and H. Kreft. 2014. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecological Letters*, 17: 866–880.

Svetnik, V.; A. Liaw; C. Tong; J. Culberson; R.P. Sheridan and B.P. Feuston. 2003. Random forest: A classification and regression tool for compound classification and QSAR modeling. *Journal of Chemical Information and Computation Science*, 43(6): 1947–1958.

Tan Q.; Y. Shao; S. Yang and Q. Wei. 2003. Wetland vegetation biomass estimation using Landsat-7 ETM data. *IEEE Transaction on Geoscience and Remote Sensing*, 3(4): 2629–2631.

Tian, Y.C.; K.J. Gu; X. Chu; X. Yao; W.X. Cao and Y. Zhu. 2014. Comparison of different hyperspectral vegetation indices for canopy leaf nitrogen concentration estimation in rice. *Plant Soil*, 376: 193–109.

Turner, W. 2014. Sensing biodiversity. *Science*, 346: 301–302.

Ulanowski, T.A. and B.A. Branfireun. 2013. Small-scale variability in peatland pore-water biogeochemistry, Hudson Bay Lowland, Canada. *Science of Total Environment*, 454-455: 211–218.

Vaiphasa, C.; A.K. Skidmore; W.F. de Boer and T. Vaiphasa. 2007. A hyperspectral band selector for plant species discrimination. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62: 225–235.

Viedma, O.; I. Torres; B. Pérez and J.M. Moreno. 2012. Modeling plant species richness using reflectance and texture data derived from QuickBird in a recently burned area of Central Spain. *Remote Sensing of Environment*, 119: 208–221.

Wang, W.; X. Yao; X. Yao; Y. Tian; X. Liu; J. Ni; W. Cao and Y. Zhu. 2012. Estimating leaf nitrogen concentration with three-band vegetation indices in rice and wheat. *Field Crops Research*, 129: 90–98.

Wang, X. 2007. Effects of species richness and elevated carbon dioxide on biomass accumulation: a synthesis using meta-analysis. *Ecophysiology*, 152: 595–605.

Wilcoxon, F. 1945. Individual comparison by ranking method. *Biometrics Bulletin*, 1(6): 80–83.

Wu, X.; X. Wang; Z. Tang; Z. Shen; C. Zheng; X. Xia and J. Fang. The relationship between species richness and biomass changes from boreal to subtropical forests in China. *Ecography*, 37: 1–12.

Zedler, J.B. and S. Kercher. 2005. Wetland resources: Status, trends, ecosystem services, and restorability. *Annual Review of Environment and Resources*, 30: 39–74.

Zhao, K.; S. Popescu; X. Meng; Y. Pang and M. Agca. 2011. Characterizing forest canopy structure with lidar composite metrics and machine learning. *Remote Sensing of Environment*, 115: 1978–1996.

Zolkos, S.G.; S.J. Goetz and R. Dubayah. 2013. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sensing of Environment*, 128: 289–298.

Zomer, R.J.; A. Trabucco and S.L. Ustin. 2008. Building spectral libraries for wetlands land cover classification and hyperspectral remote sensing. *Journal of Environmental Management*, 90: 2170–2177.

8. APÉNDICE

Apéndice 1. Riqueza de especies vasculares (n° de especies) y C aéreo (km m⁻²) por punto de muestreo.

Turbera	Riqueza	C	Turbera	Riqueza	C	Turbera	Riqueza	C
T1	13	0,12	T2	7	0,66	T3	13	0,36
T1	11	0,31	T2	5	0,40	T3	11	1,15
T1	14	0,60	T2	8	0,47	T3	10	0,57
T1	13	0,41	T2	7	0,16	T3	9	1,46
T1	13	0,22	T2	6	1,21	T3	9	0,79
T1	12	0,27	T2	8	0,50	T3	8	0,48
T1	12	0,16	T2	7	0,18	T3	8	0,27
T1	7	0,17	T2	7	0,73	T3	9	0,69
T1	11	0,19	T2	6	0,18	T3	7	0,75
T1	11	0,05	T2	6	1,77	T3	10	0,63
T1	12	0,60	T2	8	3,65	T3	6	0,26
T1	10	0,12	T2	7	0,54	T3	12	0,21
T1	10	0,55	T2	10	0,29	T3	10	0,54
T1	14	0,28	T2	5	1,18	T3	11	0,20
T1	11	0,23	T2	7	0,96	T3	13	0,15

Apéndice 2. Listado de especies presentes en cada turbera

Especie	Hábito	T1	T2	T3
<i>Agrostis tenuis</i> L.	Herbácea			x
<i>Aster vahlii</i> (Gaudich.) Hooker et Arnott	Herbácea			x
<i>Avena</i> sp	Herbácea			x
<i>Baccharis patagonica</i> H. et A.	Arbustiva	x	x	x
<i>Baccharis sagittalis</i> (Less.) DC.	Herbácea			x
<i>Berberis microphylla</i> G. Forst.	Arbustiva		x	x
<i>Blechnum chiensis</i> (Kaulf.) Mett.	Herbácea	x		x
<i>Blechnum cordatum</i> (Desv.) Hieron.	Herbácea		x	
<i>Blechnum magellanicum</i> (Desv.) Mett.	Herbácea	x	x	x
<i>Blechnum pennamarina</i> (Poir.) Kuhn.	Herbácea	x	x	x
<i>Carex magellanica</i> Lam. Gaud.	Herbácea	x	x	x
<i>Carex eamatorrhyncha</i> É. Desv.	Herbácea			x
<i>Centella asiatica</i> (L.) Urb.	Herbácea		x	x
<i>Cyperaceae</i> sp	Herbácea			x
<i>Eleocharis pachycarpa</i> É. Dsv.	Herbácea	x		x
<i>Empetrum rubrum</i> Vahl ex Willd.	Arbustiva	x		x
<i>Gaultheria insana</i> (Molina) D.J. Middleton	Arbustiva		x	
<i>Gaultheria mucronata</i> (L. f.) Hook. & Arn.	Arbustiva	x	x	x
<i>Holcus lanatus</i> L.	Herbácea			x
<i>Hypochaeris radicata</i> L.	Herbácea			x
<i>Juncus planifolius</i> R. Br.	Herbácea	x	x	x
<i>Juncus procerus</i> E. Mey.	Herbácea	x	x	x
<i>Juncus stipulatus</i> Nees & Meyen	Herbácea	x	x	x
<i>Lotus pedunculatus</i> Cav.	Herbácea			x
<i>Luma apiculata</i> (DC.) Burret	Arbórea		x	
<i>Myrteola numularia</i> (Poir.) O. Berg	Arbustiva	x	x	x
<i>Nothofagus nitida</i> (Phil.) Krasser	Arbustiva		x	
<i>Oreobolus obtusangulus</i> Gaudich	Herbácea			
<i>Ovidia pillo-pillo</i> (Gay) meisn.	Arbustiva			
<i>Philesia magellanica</i> J.F. Gmel.	Herbácea	x	x	
<i>Pilgerodendron uviferum</i> (D. Don) Florin.	Arbórea	x		
<i>Poaceae</i> sp 1	Herbácea	x		x
<i>Poaceae</i> sp 2	Herbácea	x		
<i>Poaceae</i> sp 3	Herbácea	x		x
<i>Poaceae</i> sp 4	Herbácea			x
<i>Schizaea fiotulosa</i> Labill	Herbácea	x		
<i>Shoenus rhynchosporoides</i> (Steud.) Kuek.	Herbácea	x	x	x
<i>Sticherus cryptocarpus</i> (Hook.) Ching	Herbácea	x	x	x
<i>Tepualia stipularis</i> (Hook. & Arn.) Griseb.	Arbórea		x	
<i>Trifolium repens</i> L.	Herbácea			x
<i>Uncinia</i> sp	Herbácea	x	x	x
<i>Uncinia tenuis</i> Poepp. Ex Kunth	Herbácea		x	x

