



# Transectos de datos LiDAR: una estrategia de muestreo para estimar biomasa aérea en áreas forestales

## LiDAR data transects: a sampling strategy to estimate aboveground biomass in forest areas

Alma Delia Ortiz-Reyes<sup>1</sup>, José René Valdez-Lazalde<sup>\*</sup>, Gregorio Ángeles-Pérez<sup>1</sup>, Héctor M. De los Santos-Posadas<sup>1</sup>, Laura Schneider<sup>2</sup>, Carlos Arturo Aguirre-Salado<sup>3</sup> y Alicia Peduzzi<sup>4</sup>

<sup>1</sup> Colegio de Postgraduados. Postgrado en Ciencias Forestales. Montecillo, Texcoco, México.

<sup>3</sup> Universidad Autónoma de San Luis Potosí. Facultad de Ingeniería. San Luis Potosí, S.L.P., México.

<sup>4</sup> Woods Hole Research Center. Gilman Orday Campus. Falmouth, MA, USA.

<sup>2</sup> Rutgers University. Department of Geography. Piscataway, NJ, USA.

\* Autor de correspondencia: valdez@colpos.mx

### RESUMEN

La estimación y el mapeo de la biomasa aérea sobre áreas extensas puede realizarse haciendo uso de las herramientas que ofrece la percepción remota. El objetivo de este estudio fue estimar la biomasa aérea de dos tipos de selva mediana: subperennifolia (SMSP) y subcaducifolia (SMSC) en la península de Yucatán, México, empleando métricas generadas a partir de datos *Light Detection and Ranging* (LiDAR). Se usaron datos de 365 unidades de muestreo del Inventario Nacional Forestal y de Suelos (INFyS) de México para calibrar modelos de biomasa aérea usando regresión lineal múltiple y *Random Forest* (RF). Con estos modelos se mapeó la biomasa aérea sobre franjas de datos LiDAR. El modelo de regresión transformado logró explicar la varianza en un 62% (RMSE = 41.44 Mg ha<sup>-1</sup> para SMSP y 36.60 Mg ha<sup>-1</sup> para SMSC) para ambos tipos de vegetación. Los modelos generados a través de RF lograron explicar la varianza en un 57% (RMSE = 40.73 Mg ha<sup>-1</sup>) para la SMSP y solo de 52% (RMSE = 35.10 Mg Ha<sup>-1</sup>) para la SMSC. El desfase entre la toma de datos en campo y LiDAR, así como el error en la precisión de las coordenadas de los sitios de inventario, son factores reconocidos que influyeron en los resultados. A pesar de lo anterior, las estimaciones obtenidas podrían servir de base para estimar el inventario completo de biomasa en el área de estudio incorporando datos espectrales derivados de un sensor remoto que cubra la totalidad de esta.

PALABRAS CLAVE: franjas LiDAR; inventario forestal; muestreo; península de Yucatán.

### ABSTRACT

The estimation and mapping of aboveground biomass over large areas can be done using the remote sensing tools. The objective of this study was to estimate the aboveground biomass of two types of tropical forest: semi-evergreen (SETF) and semi-deciduous tropical forest (SDTF) in the Yucatan Peninsula, Mexico, using metrics obtained from LiDAR (Light Detection and Ranging) data. Data from 365 plots of the National Forest and Soils Inventory of Mexico were used to calibrate aboveground biomass models using multiple linear regression and Random Forest. These models were used for mapping aboveground biomass along LiDAR strips. The transformed regression model explained the variance by 62% (RMSE = 41.44 Mg ha<sup>-1</sup> for SETF & 36.60 Mg ha<sup>-1</sup> for SDTF) for both types of vegetation. The models of Random Forest explained the variance by 57% (RMSE = 40.73 Mg ha<sup>-1</sup>) for SETF and only 52% (RMSE = 35.10 Mg ha<sup>-1</sup>) by SDTF. The mismatch between the field data and LiDAR data, as well as the error in the precision of the coordinates of the inventory plots, were recognized as factors that influenced on the results. Despite the above, the estimates obtained could serve as a basis to estimate the complete biomass inventory in the study area by incorporating spectral data derived from a remote sensor that covers the entire area.

KEYWORDS: LiDAR strips; forest inventory; sampling; Yucatan Peninsula.

## INTRODUCCIÓN

Los bosques tropicales son los ecosistemas vegetales más biodiversos del mundo. Cuentan con aproximadamente la mitad de la biomasa viva de los ecosistemas terrestres (Houghton, Hall y Goetz, 2009) y juegan un papel importante en el ciclo de carbono global debido a su naturaleza de rápido crecimiento y acumulación de carbono atmosférico (Meyer *et al.*, 2013). Por lo tanto, la estimación precisa de su biomasa aérea es de gran interés para los gestores de estos recursos.

Las mediciones en campo y el muestreo destructivo son consideradas como la metodología más precisa para estimar biomasa (Zolkos, Goetz y Dubayah, 2013); sin embargo, su aplicación es impráctica cuando la región de interés es muy extensa o la frecuencia de evaluación es corta, debido a los costos asociados con la toma de datos. Estos problemas se magnifican en los bosques tropicales debido a la heterogeneidad del dosel, lo cerrado de los estratos y la irregularidad de los fustes (Xu *et al.*, 2016).

El uso de las herramientas de percepción remota ha incrementado la capacidad de realizar las estimaciones, especialmente en los bosques tropicales (Houghton *et al.*, 2009; Gonçalves *et al.*, 2017) que regularmente presentan un patrón de agregación y variabilidad alta en distancias espaciales cortas, lo que resulta en un mosaico complejo que es difícil de muestrear en su totalidad (Barbosa, Broadbent y Bitencourt, 2014).

Los datos provenientes de sensores Light Detection and Ranging (LiDAR), con capacidad de derivar información de la estructura del dosel en tres dimensiones, pueden usarse para generar información que, una vez calibrada y validada con datos de campo, se vuelve una alternativa para el mapeo continuo de la biomasa forestal (Hansen, Phillips, Dubayah, Goetz y Hofton, 2014). De acuerdo con Baltsavias (1999), el sistema LiDAR funciona bajo el principio de medir el tiempo de viaje entre el pulso emitido por el sensor láser y el pulso que recibe del objetivo. La medición de la distancia se obtiene al multiplicar este intervalo de tiempo por la velocidad de la luz y dividirlo entre dos. En su trayectoria, la señal láser proporciona

información de varios objetivos, penetrando a través del dosel, lo que resulta en productos como nubes de puntos de múltiples retornos o formas de onda completas que describen la estructura de la vegetación en tres dimensiones: x, y, z (Næsset y Økland, 2002).

A pesar de la utilidad de los datos LiDAR, no siempre es posible contar con una cobertura completa (de pared a pared) de datos para la zona de interés. Por lo anterior, en la mayoría de los estudios regionales LiDAR se ha usado como un mecanismo de muestreo en fases (Saarela *et al.*, 2015). El muestreo de LiDAR en franjas es una forma de obtener información completa, precisa y uniforme, tanto espacial como temporalmente, teniendo la ventaja (comparado con un inventario tradicional) de reducir el tiempo invertido y la intensidad del trabajo requerido. Para reducir la incertidumbre en las estimaciones de biomasa, es necesario procurar la realización de observaciones espacialmente continuas que sean lo suficientemente finas para capturar la variabilidad sobre una superficie (Hansen *et al.*, 2014; Saarela *et al.*, 2015).

Las estimaciones derivadas de LiDAR pueden ser un recurso clave en los sistemas de monitoreo forestal que se desarrollan en el contexto de los esfuerzos de mitigación del cambio climático como el mecanismo de Reducción de Emisiones por Deforestación y Degradación Forestal (REDD+) (Asner *et al.*, 2012), pues será posible el mapeo y monitoreo de la biomasa y carbono de los bosques tropicales en áreas geográficas grandes.

Existen diversos trabajos de estimación de biomasa aérea en bosques tropicales del mundo. Cada uno emplea diferentes enfoques metodológicos que van desde el modelado estadístico con técnicas paramétricas (Clark, Roberts, Ewel y Clark, 2011; Asner *et al.*, 2012) y no paramétricas (Mascaro *et al.*, 2014). Así mismo, utilizan diferentes datos de sensores ópticos pasivos como el *Moderate Resolution Imaging Spectroradiometer* (MODIS) (Saatchi, Houghton, Dos Santos Alvala, Soares y Yu, 2007; Baccini, Laporte, Goetz, Sun y Dong, 2008) o *Landsat* (Foody, Boyd y Cutler, 2003), sensores activos como LiDAR (Cao, Coops, Innes, Dai y She, 2014; Laurin *et al.*,



2016) y SAR (*Synthetic Aperture Radar*, por sus siglas en inglés) (Ningthoujam, Joshi y Roy, 2018) o una combinación de los anteriores (Wang y Qi, 2008; Phua *et al.*, 2017). Todos ellos tratando de encontrar un enfoque de estimación de biomasa que genere los mejores resultados.

En México, recientemente se han realizado estimaciones de biomasa a escala nacional. En ellos han utilizado datos del Inventario Nacional Forestal y de Suelos (INFiS) (de diferentes periodos de medición), escalas diferentes y técnicas de aprendizaje de máquinas variadas (Cartus *et al.*, 2014; Urbazaev *et al.*, 2018). Rodríguez-Veiga, Saatchi, Tansey y Balzter (2016) han combinado sensores ópticos activos y pasivos, y Nelson *et al.* (2017) utilizaron datos de dos tipos de LiDAR, satelital y aerotransportado para probar tres tipos de modelos paramétricos. La varianza explicada por los modelos en los trabajos anteriores va desde 31% hasta 68%.

Los estudios anteriormente señalados analizan puntualmente a la península de Yucatán como zona clave para centrar esfuerzos y conocer el estado actual de las reservas de carbono, dados los altos niveles de incertidumbre en las estimaciones, así como su importancia ecológica. Es una región contrastante con tasas altas de deforestación y un mosaico de rodales secundarios en diferentes estados sucesionales, producto de las rotaciones de cultivo, el abandono de tierras agrícolas o pastizales y los disturbios naturales como incendios o huracanes.

En este trabajo se muestra un enfoque de modelización basado en área, el cual vincula, a través de modelos matemáticos, datos medidos en campo con parámetros estadísticos espacialmente explícitos derivados de la distribución vertical de los datos LiDAR, es decir, los recortes de la nube de puntos LiDAR con los datos correspondientes a los sitios del INFiS (White *et al.*, 2013). Se empleó una base de datos espaciales y ecuaciones alométricas nuevas generadas para este tipo de selvas. Se ensayaron dos métodos estadísticos, uno paramétrico y otro no paramétrico para registrar el escalamiento de estimaciones de biomasa aérea de una escala local (datos de campo) a una segunda escala de muestreo (franjas). Las

estimaciones a escala de franja tienen como propósito aumentar la muestra de campo original y con ella facilitar estimaciones posteriores de biomasa a escala regional.

## OBJETIVOS

Este trabajo tuvo como objetivo estimar la biomasa aérea en franjas de selva mediana subperennifolia y subcaducifolia en la península de Yucatán, México. Se utilizaron transectos de datos LiDAR e información del INFiS. De manera particular el estudio se enfocó a 1) evaluar las métricas LiDAR (mejor correlacionadas con la biomasa aérea en campo) y su combinación para la estimación de biomasa aérea en dos tipos de selva mediana, 2) evaluar dos técnicas de modelado para estimar biomasa, un enfoque estadístico de regresión de mínimos cuadrados ordinarios y un enfoque de aprendizaje de máquinas, Random Forest y 3) mapear la distribución espacial de la biomasa aérea en las franjas con datos LiDAR. Lo anterior permitirá incrementar espacialmente la muestra de datos de biomasa y posteriormente realizar estimaciones en toda el área de estudio utilizando datos derivados de imágenes satelitales.

## MATERIALES Y MÉTODOS

### Área de estudio

Este estudio se llevó a cabo en dos tipos de selvas medianas: subperennifolia (SMSP) y subcaducifolia (SMSC) de la península de Yucatán, México, localizadas en la porción SE de la República Mexicana (17°00' y 21°45' N, 86°30' y 9°30' O). La superficie aproximada de la península es de 141 523 km<sup>2</sup>, equivalente a 7% del territorio nacional, de los cuales, 28 171 km<sup>2</sup> corresponden a SMSC y 48 699 km<sup>2</sup> a la SMSP (Instituto Nacional de Estadística y Geografía [Inegi], 2013) (Fig. 1). La selva mediana subcaducifolia se distribuye en la parte centro y norte de la península de Yucatán, en forma de franja. En la selva mediana subperennifolia, los componentes arbóreos pierden estacionalmente su follaje en un 25% a 50% durante la época de sequía. Se establece en regiones planas, ligeramente más secas y con drenaje rápido de la península (Pennington, 2005).

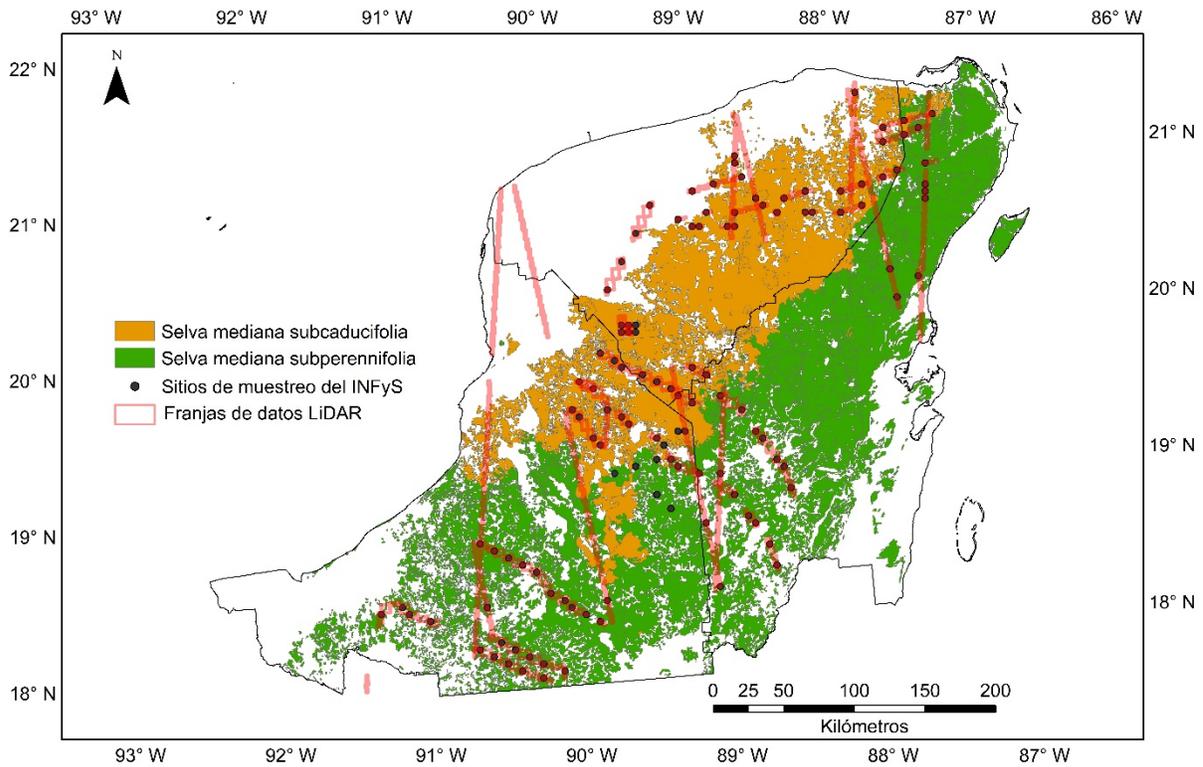


FIGURA 1. Distribución de las unidades de muestreo del Inventario Nacional y de Suelos (INFyS) de México, utilizados en el estudio y de las franjas de datos LiDAR, sobre los dos tipos de vegetación evaluados (Serie V, Inegi, 2013).

### Datos de campo

Los datos del INFyS fueron recolectados por la Comisión Nacional Forestal [Conafor] (2015), en el área de estudio durante el periodo 2009-2014 (Fig. 1) bajo un diseño de muestreo sistemático por conglomerados. Cada conglomerado está integrado por cuatro sitios rectangulares de 400 m<sup>2</sup> donde se midió y registró el arbolado cuyo diámetro normal a la altura de 1.3 m sobre la superficie del suelo, fue igual o mayor a 7.5 cm (Comisión Nacional Forestal [Conafor] 2011). Un total de 187 sitios coincidieron espacialmente con las franjas de datos LiDAR, en áreas que corresponden a la SMSC y 176 sitios en la SMSP. Las variables dasométricas que se consideraron para los árboles fueron el diámetro normal (DN), la altura total (AT), el nombre de la especie y la condición del individuo (vivo o muerto). El INFyS indicó que las coordenadas

geográficas de los sitios de muestreo tienen un error promedio de 5.9 m (Conafor, 2015).

Se encontraron superposiciones entre algunos sitios, por lo que se optó por reconstruir los polígonos tomando en cuenta las medidas de diseño para el establecimiento de las parcelas en campo (Conafor, 2011). En la tabla 1 se presenta un resumen de los datos dasométricos de los sitios de muestreo.

Los datos de campo sirvieron de base para calcular la biomasa aérea mediante ecuaciones alométricas. La biomasa de individuos con DN < 10 cm se estimó con ecuaciones para la especie cuando estuvo disponible (Cairns, Olmsted, Granados y Argaez, 2003; Puc-Kauil *et al.*, 2019), cuando la especie no contaba con una ecuación específica, se empleó la ecuación local de Puc-Kauil *et al.* (2019) para diámetros menores a 10 cm. Para individuos



con  $DN > 10$  cm se usaron las ecuaciones disponibles para cada especie (Cairns *et al.*, 2003), caso contrario, se investigó su valor de densidad de la madera y, tomando como referencia la base de datos de densidad de la madera global de Zanne *et al.* (2013), se aplicó la ecuación global de Urquiza-Haas, Dolman y Peres (2007). El valor de densidad de la madera mejora las predicciones de biomasa, debido a que las formas de crecimiento de las especies están en función de esta (Chave *et al.*, 2004). Para las especies que no se conoce el valor de la densidad de su madera y que además no tenían una ecuación específica se aplicó la ecuación global de Cairns *et al.* (2003) para diámetros mayores a 10 cm. En el caso de palmas y lianas, se usaron las ecuaciones propuestas por Chave *et al.* (2003).

### Datos LiDAR

Los datos LiDAR fueron adquiridos en 2013 por la *National Aeronautics & Space Administration* (NASA) mediante el sistema aerotransportado *Goddard's LiDAR, Hyperspectral & Thermal Imager* (G-LiHT). Este utilizó un escáner láser Riegl VQ-480, con una frecuencia de pulso de hasta 300 kHz. La velocidad de medición efectiva es de hasta 150 kHz a lo largo de una franja perpendicular a la dirección de vuelo de 60°. La huella del láser fue de 10 cm de diámetro, con un máximo de retornos por pulso de ocho y una densidad de seis pulsos por metro cuadrado en la superficie. La altitud de vuelo fue de 335 m (Cook *et al.*, 2013). Los datos se registraron a lo largo de franjas y clasificados en rutas para un mejor control de la información (Fig. 1). También se

utilizaron los modelos digitales del terreno disponibles para normalizar la nube de puntos a alturas por encima del terreno.

### Métricas LiDAR

Los datos LiDAR se procesaron en el software FUSION/LDV Versión 3.60+ (McGaughey, 2016) para extraer, a escala de unidad de muestreo, las métricas que posteriormente se utilizaron como variables predictoras de los modelos de biomasa aérea. Las métricas representan la distribución de altura de la vegetación y se dividen en tres grupos: 1) estadísticas que describen la distribución de las alturas correspondientes a los retornos del láser en una parcela o área determinada (p. ej. media, desviación estándar, moda), 2) percentiles de la distribución de la altura (P05, P10, P50, P70, P95) y 3) métricas de densidad que corresponden a la proporción de retornos del láser dentro de cierta porción de altura o por encima o debajo de cierta altura, conocida como altura de corte (Tabla 2). La descripción detallada de las métricas se encuentra en el manual del software FUSION (McGaughey, 2016). Se definió una altura de corte de 2 m para disminuir el ruido de los puntos próximos al suelo, causado por el sotobosque y fallas en el filtrado de los puntos del terreno. Se obtuvieron productos ráster por cada métrica calculada (Tabla 2) y corresponden a las franjas de datos LiDAR. El tamaño de pixel fue de 20 m, equivalente al tamaño de las parcelas en campo (400 m<sup>2</sup>).

TABLA 1. Estadísticas descriptivas de los 363 sitios de muestreo que coinciden espacialmente con las franjas de datos LiDAR.

Variable	<i>Selva mediana subperennifolia</i> (176)		<i>Selva mediana subcaducifolia</i> (187)	
	Media	Desviación estándar	Media	Desviación estándar
Área basal (m <sup>2</sup> ha <sup>-1</sup> )	16.27	9.32	14.14	8.48
Altura media (m)	9.24	2.18	8.91	1.52
Biomasa aérea (Mg ha <sup>-1</sup> )	89.38	62.26	71.50	50.84

TABLA 2. Resumen de las variables predictoras (métricas LiDAR) empleados en la estimación de biomasa aérea. Detalles de estas están disponible en McGaughey (2016).

<i>Tipo de variable predictorora</i>	<i>Descripción</i>	<i>Notación</i>
Estadísticas descriptivas de altura.	Altura media	elev_ave_2plus_20METERS
	Altura máxima	elev_max_2plus_20METERS
	Altura media generalizada cuadrática	elev_quadratic_mean_20METERS
	Altura media generalizada cúbica	elev_cubic_mean_20METERS
	Desviación estándar de la altura	elev_stddev_2plus_20METERS
Percentiles de altura (14)	Percentiles 05-99 de altura	elev_P05-P99_2plus_20METERS
Métricas de densidad	Tasa de relieve del dosel	elev_canopy_relief_ratio_20METERS
	Primeros retornos encima de la altura media	X1st_cnt_above_mean_20METERS
	% de los primeros retornos encima de la altura media	X1st_cover_above_mean_20METERS
	% de todos los retornos encima de 2 m	all_cover_above2_20METERS
	% de los primeros retornos encima de 2 m	X1st_cover_above2_20METERS
	% de todos los retornos encima de la altura media	all_cover_above_mean_20METERS
	(Todos los retornos encima de 2 m) / (Total de primeros retornos)*100	all_1st_cover_above2_20METERS
	(Todos los retornos encima de la altura media) / (Total de primeros retornos) * 100	all_1st_cover_above_mean_20METERS

Estimación de biomasa aérea basada en datos LiDAR

**Análisis de regresión**

Inicialmente se identificaron correlaciones de Pearson ( $r \geq 0.5$ ) entre los valores de las métricas LiDAR (Tabla 2) obtenidos de la nube de puntos correspondiente a los sitios de muestreo y las estimaciones de biomasa para los mismos sitios calculadas mediante ecuaciones alométricas. Posteriormente se probaron modelos lineales múltiples para estimar la biomasa aérea en función de las métricas LiDAR, empleando la función “lm” del paquete “stats” del software R (R Development Core Team, 2013). Sin embargo, se optó por construir los modelos con base en lo señalado por Ortiz-Reyes *et al.* (2015), en donde incorporaron características notorias al modelo, es decir, variables que describen tanto la estructura horizontal como vertical del dosel. La estructura horizontal considera la densidad del arbolado, mientras que la estructura vertical es

descrita por la distribución de alturas. Las métricas “primeros retornos por encima de la altura media” y el “percentil 95” son ejemplos de cada una de estas características, respectivamente.

Los primeros modelos no cumplían con los supuestos de la regresión, por lo que se consideró mejorarlos mediante la transformación Box-Cox a través de la librería MASS del software R, para conseguir una varianza constante y residuales distribuidos en forma normal. Adicionalmente se utilizó el factor de inflación de la varianza (FIV) como indicador para detectar problemas de colinealidad entre variables predictoras (Peduzzi, Wynne, Fox, Nelson y Thomas, 2012). Sin embargo, el hecho de usar solo dos métricas (correspondientes al plano vertical y horizontal) no significó gran problema de colinealidad entre ellas.



## Random Forest

*Random Forest* (RF) es una técnica avanzada de aprendizaje de máquinas utilizada para modelar asociaciones estadísticas complejas de datos de diferentes fuentes. Es un método de regresión (también de clasificación) basado en la agregación de un número grande de árboles de decisión. *Random Forest* es un conjunto de árboles construidos a partir de un conjunto de datos de entrenamiento y validados internamente para generar una predicción de la variable de respuesta dados los predictores. Cada árbol es construido a partir de una muestra *bootstrap* extraída con reemplazo del conjunto de datos original y las predicciones de los árboles resultantes son producto del promedio de todos ellos, en el caso de la regresión (Cutler, Cutler y Stevens, 2012).

El rendimiento de RF no se ve afectado por la colinealidad de las variables predictoras, ni por la falta de normalidad de la variable de respuesta (Wilkes *et al.*, 2015). Es de fácil aplicación y con capacidad de procesar eficientemente bases de datos grandes, lo que permite ser una opción en estudios regionales (Baccini *et al.*, 2012; Mascaro *et al.*, 2014).

Para reconocer las métricas LiDAR que aportan información significativa al modelo RF, es necesario medir la importancia de las variables predictoras identificando la influencia que tiene cada predictor sobre el error cuadrático medio (MSE, por sus siglas en inglés) del modelo. El algoritmo se describe ampliamente en Cutler *et al.* (2012). En términos prácticos, si la métrica LiDAR está contribuyendo al modelo y se permuta (o cambia) aleatoriamente por otro de los predictores en el grupo de prueba, entonces el MSE del modelo RF aumentará debido a que está perdiendo la información que aportaba dicha variable (Wilkes *et al.*, 2015).

Las variables predictoras o métricas LiDAR empleadas en RF para estimar la biomasa aérea en la SMSP y la SMSC fueron las que registraron una  $r \geq 0.5$ . Sin embargo, se eliminaron variables conforme la precisión del modelo mejoraba y se identificaba la influencia de cada variable predictora en el MSE. Se utilizó el paquete “randomForest” del software R.

Los modelos de biomasa fueron desarrollados usando el enfoque basado en área (White *et al.*, 2013). La precisión de los diferentes modelos fue evaluada en términos de la raíz del error cuadrático medio (RMSE por sus siglas en inglés) y RMSE relativo (%), así como el coeficiente de determinación ( $R^2$ ) entre las observaciones y predicciones.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

$$RMSE\% = 100 * \frac{RMSE}{\bar{y}}$$

Donde  $y_i$  es la biomasa observada,  $\hat{y}_i$  es la biomasa estimada,  $\bar{y}$  es la biomasa promedio observada, y  $n$  es el número de observaciones.

## Mapeo de la biomasa aérea en franjas con datos LiDAR

Los mapas de biomasa aérea para los dos tipos de selva se obtuvieron a través de la aplicación de ambos enfoques de regresión: regresión lineal múltiple y Random Forest. Se emplearon los productos ráster correspondientes a las métricas elegidas por el modelo de regresión lineal múltiple y el modelo RF, a lo largo de las franjas de datos LiDAR con un tamaño de pixel de 20 m. Se utilizó el paquete “raster” del software R.

## RESULTADOS

### Estimación de biomasa mediante modelo de regresión lineal

Del total de variables analizadas, 24 métricas LiDAR mostraron buena correlación con la biomasa aérea estimada en campo para la SMSP (valor de correlación de Pearson  $\geq 0.5$ ), siendo la altura media la métrica que tuvo la mayor correlación con la biomasa estimada para este tipo de vegetación ( $r=0.7$ ). En el caso de la SMSC, 23 métricas mostraron buena correlación y fue el percentil 90 la variable

con mayor correlación con la biomasa aérea estimada en campo.

Los resultados de los modelos de regresión lineal múltiple y transformado para la SMSP y SMSC se presentan en la tabla 3. En ambos casos se utilizaron solo dos métricas para predecir la biomasa aérea. La altura media y el porcentaje de los primeros retornos encima de 2 m para la SMSP. En el caso de la SMSC las métricas fueron el percentil 95 de la altura y el porcentaje de los primeros retornos encima de la altura media. Los valores más altos de R<sup>2</sup> se alcanzaron cuando se transformó a la biomasa aérea mediante el procedimiento Box Cox. Cuando se agregaron más variables explicativas a los modelos, se alcanzó una R<sup>2</sup> marginalmente más grande, sin embargo, esos modelos tendieron a sobre ajustar los datos

observados, resultando en valores altos de FIV entre las métricas explicativas de LiDAR similares.

### Estimación de biomasa mediante Random Forest

Los resultados de los modelos de RF para la SMSP y SMSC se presentan en la tabla 4. El número de variables predictoras empleadas en el modelo de SMSP fueron 20 métricas LiDAR, correspondientes a 13 percentiles de altura (P05 - P99), tres estadísticas descriptivas de la altura y cuatro tasas de retorno. Para el modelo de la SMSC fueron tres percentiles de altura (P75, P90 y P95), cuatro estadísticas descriptivas de la altura y seis tasas de retorno (Fig. 2).

TABLA 3. Modelos de regresión lineal múltiple que relacionan las métricas LiDAR con la biomasa aérea por tipo de vegetación.

<i>Tipo de vegetación</i>	<i>Modelo</i>	<i>R<sup>2</sup></i>	<i>RMSE</i>	<i>RMSE %</i>
SMSP	$B = -92.377 + \text{Altura media} * 15.09 + \% \text{ de los primeros retornos encima de } 2 \text{ m} * 0.728$	0.56	40.79	45.63
SMSP Box Cox	$B_t = -4.79733 + \text{Altura media} * 1.034 + \% \text{ de los primeros retornos encima de } 2 \text{ m} * 0.08617$	0.62	41.44	46.36
SMSC	$B = -110.435 + \text{Percentil } 95 \text{ de la altura} * 12.939 + \% \text{ de los primeros retornos encima de la altura media} * 0.573$	0.53	34.68	48.50
SMSC Box Cox	$B_t = -2.46 + \text{Percentil } 95 \text{ de la altura} * 0.567 + \% \text{ de los primeros retornos encima de la altura media} * 0.032$	0.62	36.50	51.05

\*\*\* B = Biomasa aérea, B<sub>t</sub> = Biomasa aérea transformada.

TABLA 4. Resultados de los modelos Random Forest para la estimación de biomasa aérea por tipo de vegetación.

	<i>selva mediana subperennifolia</i>	<i>selva mediana subcaducifolia</i>
% de varianza explicada	56.95	52.07
r	0.75	0.73
RMSE	40.73	35.10
# de variables predictoras	20	13

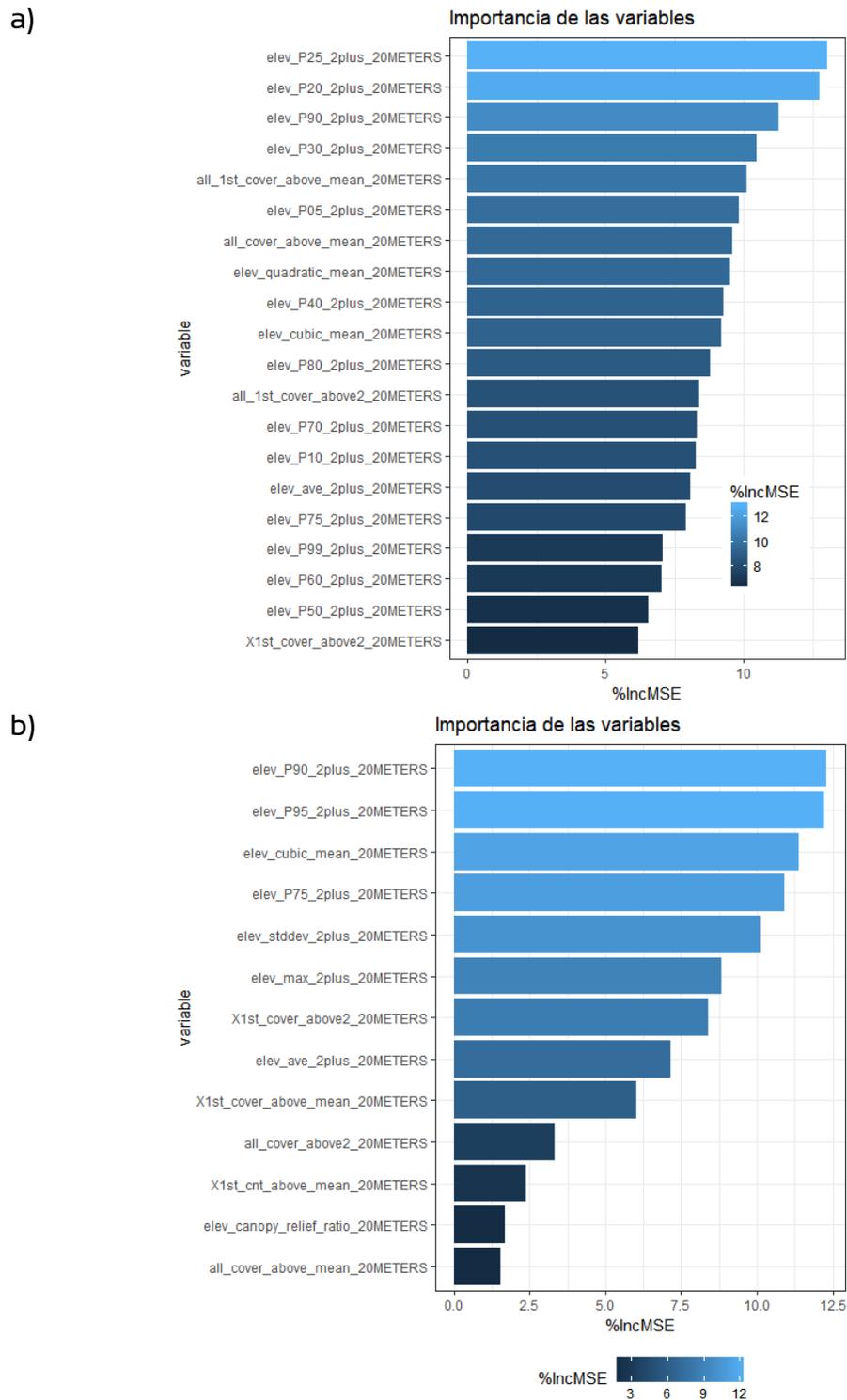


FIGURA 2. Importancia relativa de a) 20 variables seleccionadas para el modelo Random Forest en la selva mediana subperennifolia y b) de 13 variables en la selva mediana subcaducifolia. %IncMSE: Incremento porcentual del error cuadrático medio (*mean square error*).

La figura 3 muestra la relación entre la biomasa aérea predicha a partir de los modelos de regresión (a y b) y Random Forest (c y d) con la biomasa aérea de referencia calculada aplicando las ecuaciones alométricas a partir de las mediciones del INFyS. Los cuatro gráficos sugieren que los valores altos de biomasa aérea se subestimaron, mientras que los valores pequeños fueron sobreestimados.

### Mapas de biomasa aérea mediante los dos enfoques

En la figura 4 se presenta un acercamiento de los mapas de biomasa aérea en franjas generada mediante regresión lineal (a) y mediante RF (b) en la SMSC. c) un acercamiento de la biomasa aérea generada mediante regresión lineal y d) mediante RF para la SMSP.

### DISCUSIÓN

En el presente estudio, bajo el enfoque de regresión lineal, las métricas de LiDAR que más se correlacionaron con la biomasa aérea de la SMSC fueron los percentiles de altura 80, 90 y 95. El P90 y P95 son lo más parecido a contar con una estimación de la altura máxima del dosel (Goetz y Dubayah, 2011). Los resultados de este estudio son consistentes con otros de estimación de biomasa aérea en bosques tropicales. Cada una de estas métricas (P80, P90 y P95) se combinan con otras métricas de densidad o estadísticas de la distribución de altura para conformar los modelos generados por Clark *et al.* (2011) en Costa Rica, en manglares de Brasil (Rocha de Souza *et al.*, 2018), así como en bosques templados (Lefsky *et al.*, 2002).

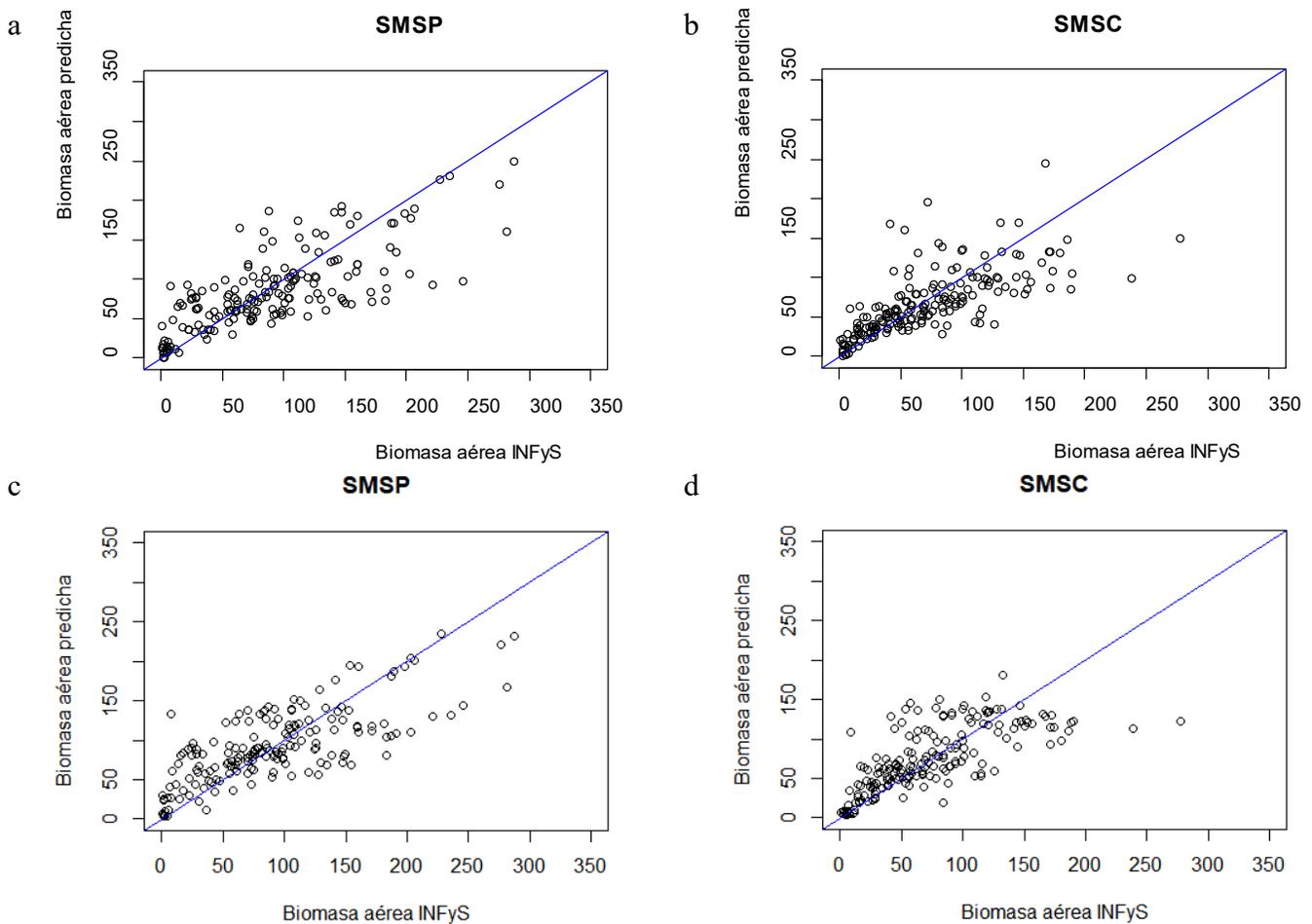


FIGURA 3. Predicción de la biomasa aérea usando los datos LiDAR en franjas: a) Biomasa aérea en la selva mediana subperennifolia y b) selva mediana subcaducifolia mediante regresión lineal y c) SMSP y d) SMSC mediante Random Forest. Se indica la línea 1:1.

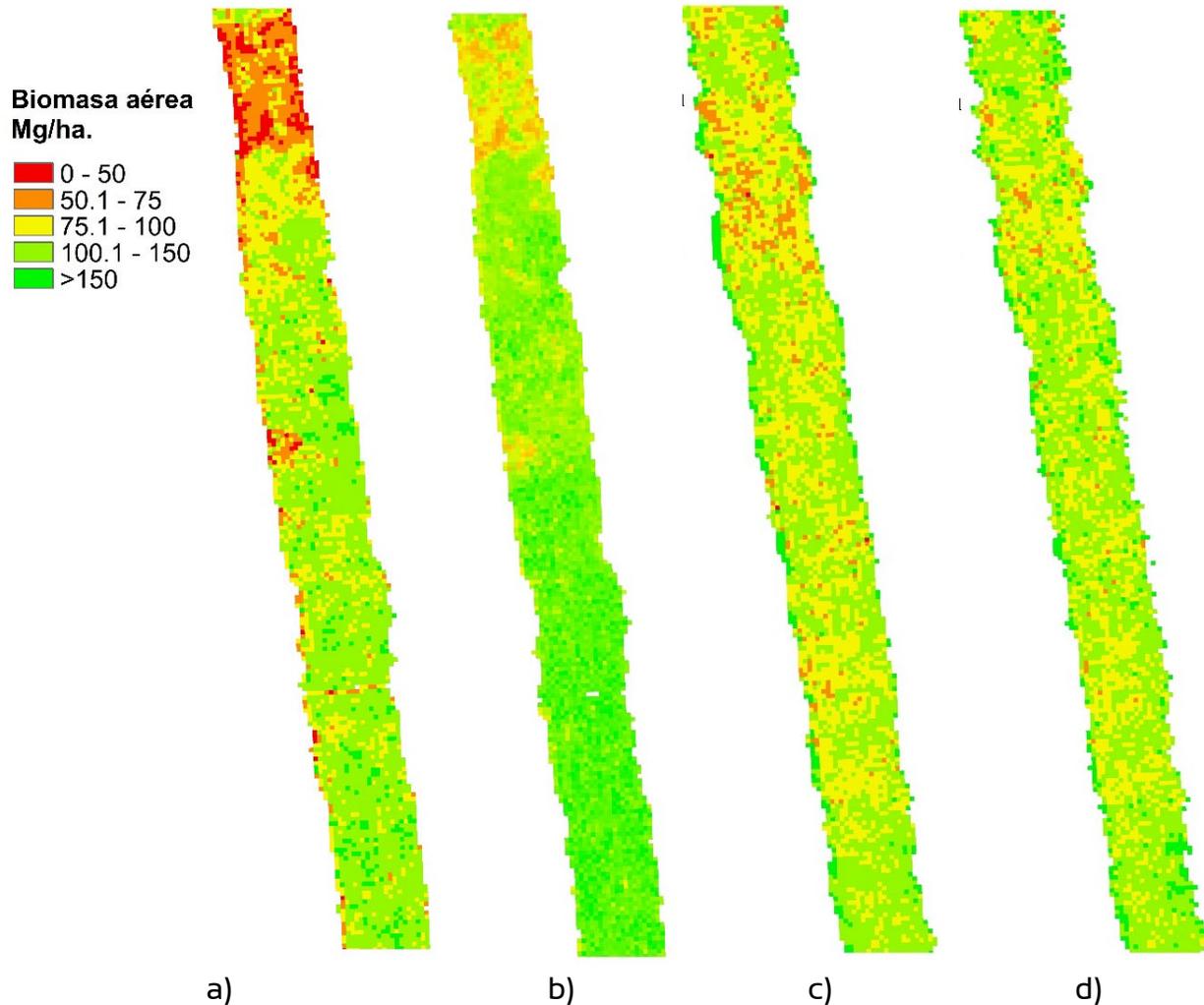


FIGURA 4. Mapas de biomasa aérea usando los datos LiDAR en franjas: a) biomasa aérea mediante regresión múltiple y b) mediante RF en la SMSC; c) biomasa aérea mediante regresión múltiple y d) mediante Random Forest en la SMSP.

Los percentiles P80 y P90 (9.6 m – 10.6 m) corresponden al patrón de altura promedio para este tipo de vegetación. Dupuy *et al.* (2012) encontraron que el patrón de altura (determinado principalmente por el arbolado adulto) se encuentra en un intervalo similar a estos percentiles. Hay que tener en cuenta que en el estudio mencionado se muestrearon individuos con DN > 5 cm y los datos del INFyS empleados en este trabajo registran los valores del DN > 7.5 cm.

La identificación de percentiles que corresponden a una medida de la altura total de los árboles (P90, P95) y que

resultan útiles para predecir la biomasa es completamente razonable. Esta variable (altura total de los árboles), al igual que el diámetro del fuste, consistentemente ha mostrado ser importante para predecir la biomasa de árboles individuales (Cairns *et al.*, 2003), consecuentemente, es entendible que contribuya de manera importante a predecir la biomasa a escala de parcela.

En el caso de la SMSP, la altura media y dos percentiles bajos (P20 y P25, 4.79 m – 5.32 m) fueron los que correlacionaron mejor con la biomasa de este tipo de selva, lo cual es consistente con lo indicado por Véga *et al.* (2015).

En el estudio de d'Oliveira, Reutebuch, McGaughey y Andersen (2012), realizado con el propósito de estimar la biomasa aérea en un bosque tropical en Brasil, los autores señalaron que el P25 y la varianza de la altura fueron las mejores variables explicativas de los modelos.

La alta correlación entre P20 y P25 con la biomasa aérea de la SMSP puede explicarse como la capacidad de la nube de puntos LiDAR de registrar adecuadamente la alta densidad (mediante los retornos) de este conjunto de árboles de altura menor en la SMSP. Echeverría, Arreola, Esparza, Morales y López (2014) y Zamora-Crescencio *et al.* (2017) estudiaron la composición y estructura de diferentes porciones de selva mediana subperennifolia en el estado de Campeche. En el primer estudio indicaron que 77.6% de los individuos tienen alturas menores a 5 m, mientras que la caracterización vertical del segundo estudio indicó que 42% de los individuos tenían alturas entre 2 m y 4.9 m. Es importante considerar que todo individuo con  $DN > 1$  cm fue medido. Respecto a la altura media (7.9 m), también registrada apropiadamente por la nube de puntos LiDAR, Read y Lawrence (2003) señalaron que las alturas promedio en la clase diamétrica de 5 cm - 9.9 cm se encuentran en el intervalo de 5 m a 8 m.

Por otro lado, en los dos tipos de selva estudiados, el percentil 95 y la altura media, además de dos métricas relacionadas a la cobertura del dosel (el porcentaje de los primeros retornos encima de 2 m y el porcentaje de los primeros retornos encima de la altura media) contribuyeron significativamente a predecir la biomasa mediante modelos de regresión lineal. La idea principal de usar dos métricas en cada modelo fue crear una relación simple que involucre la distribución de altura y la profundidad en que los retornos LiDAR penetran el arbolado, como lo señalan Rocha de Souza *et al.* (2018), además, es importante señalar que la adición de más variables en los modelos no incrementó significativamente la variación explicada.

Para el enfoque Random Forest, el percentil 90 (elev\_P90\_2plus\_20METERS) fue una métrica común entre las variables más importantes seleccionadas por el modelo. La tercera más importante para la SMSP y la

primera para la SMSC. En un estudio realizado por Li, Im, Quackenbush y Liu (2014), este percentil, así como el percentil 80, fueron las variables más importantes en los modelos evaluados para estimar la biomasa de bosques templados mediante técnicas de aprendizaje de máquinas.

El enfoque de regresión, usando variables transformadas, funcionó mejor que RF en cuanto a la varianza explicada, registrando 62% para ambos tipos de vegetación (RMSE = 41.44 Mg ha<sup>-1</sup> y 36.60 Mg ha<sup>-1</sup> para SMSP y SMSC respectivamente); mientras que con RF fue de 57% (40.73 Mg ha<sup>-1</sup>) para la SMSP y de 52% (35.10 Mg ha<sup>-1</sup>) para la SMSC. En general, los estudios realizados de estimación de biomasa a partir de datos LiDAR han señalado precisiones (RMSE) altamente variables que oscilan entre 20 Mg ha<sup>-1</sup> y 200 en escalas de parcelas de 30 m (900 m<sup>2</sup>) hasta 1 ha (Goetz y Dubayah, 2011).

No obstante, cuando se comparan los resultados de este trabajo con estudios similares de estimación de biomasa aérea en bosques tropicales existen diferencias notables. Por ejemplo, en un estudio realizado en la amazonia brasileña, d'Oliveira *et al.* (2012) lograron explicar la varianza hasta 72% (RMSE = 40.2 Mg ha<sup>-1</sup>); mientras que Véga *et al.* (2015) estimaron la biomasa aérea en un bosque tropical húmedo perenne en Western Ghats, India, con una  $R^2 = 0.96$  (RMSE<sub>vc</sub> = 28.83 Mg ha<sup>-1</sup>). Por otro lado, Meyer *et al.* (2013), utilizando sitios de muestreo del mismo tamaño que este estudio (0.04 ha), en un bosque tropical húmedo, hallaron una  $R^2$  de 0.19 (RMSE = 184.6 Mg ha<sup>-1</sup>) y 0.28 (RMSE = 173.8 Mg ha<sup>-1</sup>) para dos tipos de sensores LiDAR, respectivamente.

Hernández-Stefanoni *et al.* (2014) estimó la biomasa aérea en dos sitios ubicados en los mismos tipos de selva analizados en este trabajo empleando regresión múltiple. La varianza explicada para ambos sitios fue de 0.49 (RMSE<sub>vc</sub> = 37.4 Mg ha<sup>-1</sup> para el estudio situado en la SMSC y 60.2 Mg ha<sup>-1</sup> para el situado en la SMSP) con un tamaño de 0.04 ha. El RMSE obtenido en este trabajo a través del enfoque de regresión (36.50 Mg ha<sup>-1</sup>) fue similar al que obtuvieron Hernández *et al.* (2014) en la SMSC. Incluso la RMSE para la SMSP fue menor en este estudio (41.44 Mg ha<sup>-1</sup> con una



diferencia de 18.8 Mg ha<sup>-1</sup>) a pesar de que la varianza explicada no logró un valor mayor.

Por otro lado, Urbazaev *et al.* (2018) estimaron la biomasa en todo México mediante un modelo que explicó la varianza en un 68% (RMSE = 22.22 Mg ha<sup>-1</sup>) empleando simulaciones Monte Carlo y el algoritmo de aprendizaje de máquinas denominado Cubist. La varianza obtenida en su trabajo es parecida a la que se obtuvo en este estudio, la diferencia más grande es de 16%. Probablemente, la escala empleada de 1 ha (100 m tamaño del pixel) con respecto a los 400 m<sup>2</sup> (20 m) de este estudio, es un factor importante que explica estas diferencias. Los errores en la biomasa estimada con datos tridimensionales disminuyen cuando se utilizan parcelas más grandes debido al promedio espacial de los errores (Goetz y Dubayah, 2011; Hernández-Stefanoni *et al.*, 2014).

Además del promedio espacial, existen factores adicionales como el efecto de borde, en donde las copas de los árboles fuera del sitio se extienden dentro de la parcela (White *et al.* 2013). La nube de puntos LiDAR los detecta, pero en el inventario estos árboles no se registran. Dicho efecto de borde representa una discrepancia entre la cantidad de biomasa estimada y observada siendo más evidente en las parcelas pequeñas (Zolkos *et al.*, 2013).

Reiteradamente se han mencionado las ventajas de utilizar datos LiDAR para realizar las estimaciones de los parámetros forestales de interés, sin embargo, en este trabajo no se logró que los modelos explicaran la varianza más allá de 65%, lo que hace suponer que existen otros motivos que agregan error al modelo.

Uno de estos factores puede ser la poca precisión con la que se geolocalizaron los centros de los sitios del inventario, 5.9 m según el INFyS (Conafor, 2015), lo que resulta en una falta de coincidencia espacial con la nube de puntos LiDAR. De esta situación depende que árboles o partes de árboles se encuentran dentro del sitio y cuáles de ellos se miden o no (Mascaro, Detto, Asner y Muller-Landau, 2011; Gonçalves *et al.*, 2017).

En este sentido, estudios de propósito similar (Deo *et al.*, 2017; Rocha de Souza *et al.*, 2018) registran errores de

geoposicionamiento de los sitios de muestreo menores a 1 m. No obstante, White *et al.* (2013) recomiendan errores de geoposicionamiento de hasta 5 m en los sitios de muestreo en campo, además de otros métodos para disminuir el error, como registrar con el *global positioning system* (GPS) un mínimo de 500 puntos por parcela; aplicar corrección diferencial o incrementar el tamaño de parcela para asegurar un traslape suficiente.

Por otro lado, el uso de ecuaciones globales, que a pesar de ser desarrolladas en la zona puede ser otra fuente de error. Aunque en este estudio se utilizaron ecuaciones por especie, una gran cantidad de especies en los bosques tropicales no cuentan con un modelo específico. No obstante, se esperaría un buen resultado con el uso de las ecuaciones generales. Cuando se usan dichos modelos alométricos, la capacidad de predecir la biomasa aérea a partir de datos LiDAR se rige por la intensidad de la relación diámetro normal/altura a escala de individuo (Asner, 2009; Chen *et al.*, 2015). Sin embargo, en todo el proceso de estimación de biomasa existen más fuentes de error en el modelamiento de biomasa aérea: el tamaño de la muestra, el error en las mediciones en campo, la incertidumbre de muestreo relacionado al tamaño de la parcela, así como la representatividad de una red de parcelas en todo el paisaje que sin duda, están influyendo en la precisión de las estimaciones presentadas.

Otro factor importante que podría explicar los valores bajos de R<sup>2</sup> obtenidos en los modelos, es el desfase temporal entre la toma de datos del inventario y la recolección de datos LiDAR (Mascaro *et al.*, 2011); durante ese tiempo han ocurrido una serie de eventos que han modificado las condiciones de la vegetación, ya sea por mortalidad, deforestación, perturbación, etcétera (Goetz y Dubayah, 2011). La mayoría de los sitios se remidieron en 2009 y 2010 y el registro de datos LiDAR se llevó a cabo durante el año 2013. Los errores que se agregan en este punto podrían minimizarse al realizar la colecta de datos de campo y LiDAR simultáneamente (Andersen, Strunk y Temesgen, 2011; Woods *et al.*, 2011).

Es importante resaltar una constante en los modelos evaluados, los valores altos de biomasa aérea fueron subestimados, mientras que los valores bajos fueron sobreestimados. Con *Random Forest*, el proceso de promediado crea resultados sesgados hacia la media de la muestra, entonces los valores pequeños o grandes de biomasa aérea comúnmente son sub o sobreestimados (Xu *et al.*, 2016). En el caso de regresión múltiple, probablemente es consecuencia de no contar con suficiente información en los extremos. El valor de referencia fue calculado solo considerando los árboles medidos dentro del sitio, eliminando árboles pequeños y vegetación menor, sin embargo, los datos LiDAR consideran toda la información registrada dentro de la parcela.

A pesar de la limitada capacidad explicativa de los modelos generados, los resultados de este trabajo indican que la información obtenida es de alta calidad y es adecuada para mejorar las estimaciones de biomasa aérea a escala regional. Los errores registrados no exceden 50 Mg ha<sup>-1</sup>, aunque se esperaría que los errores obtenidos en este tipo de estudios y utilizando datos de percepción remota fueran próximos a 20 Mg ha<sup>-1</sup> o estar dentro del intervalo de 20% de las estimaciones obtenidas en campo (Zolkos *et al.*, 2013). No obstante, se espera una mejora al conjuntar esta información con datos auxiliares como los que proporcionan las imágenes satelitales (Lu *et al.*, 2012) u otros sensores activos (Rodríguez-Veiga *et al.*, 2016).

El proceso de obtener información extensa y de alta calidad de la estructura del bosque es difícil y costoso, principalmente en los ecosistemas tropicales, dadas sus características y factores estacionales adversos como inundaciones o huracanes durante la temporada de lluvias. Dichos factores pueden comprometer las mediciones recolectadas o la cantidad de sitios muestreados.

Aunque el problema tampoco se resolverá inmediatamente con los datos de LiDAR, debido a su limitada cobertura espacial, es necesario hacer uso de pequeñas muestras de datos que incrementan la precisión de la variable de interés evaluada, y a través de la conexión con los datos ópticos de satélite es posible realizar

estimaciones en superficies mayores sin tener que invertir el tiempo y el dinero que implica realizar el inventario forestal sobre esa misma gran superficie (Chi *et al.*, 2017).

## CONCLUSIONES

Este trabajo demostró que es posible realizar una estimación adecuada de biomasa aérea en bosques tropicales a través de una muestra de datos LiDAR y datos de inventario, a pesar de la complejidad y diversidad de este tipo de ecosistemas. Las métricas de altura y cobertura de LiDAR son las variables que correlacionaron mejor y contribuyeron, en gran medida, a estimar la biomasa aérea mediante los modelos empleados. Utilizando dos enfoques de regresión: lineal múltiple y *Random Forest*, fue posible estimar la biomasa aérea con resultados similares. El enfoque de regresión múltiple (transformado) logró explicar mejor la varianza que el método de *Random Forest*, no obstante, los valores de RMSE en ambos casos fueron los más altos (41.44 Mg ha<sup>-1</sup> y 36.50 Mg ha<sup>-1</sup> para la SMSP y SMSC, respectivamente). A partir de los modelos generados, se obtuvieron los mapas de biomasa aérea sobre las franjas de datos LiDAR, lo que permitió conocer su distribución espacial de forma continua y homogénea.

Este estudio pone en evidencia la importancia de corregir algunas deficiencias de diseño en la fase de campo que podrían mejorar la precisión de los modelos ajustados para predecir la biomasa aérea: el desfase entre la toma de datos en campo y el registro de datos LiDAR, la precisión en la geolocalización de las parcelas en campo, el uso de ecuaciones locales o adecuadas para estimar biomasa aérea y la necesidad de que la muestra incluya la totalidad de las diferentes condiciones de densidad existentes en el área de interés.

## REFERENCIAS

- Andersen, H.-E., Strunk, J., & Temesgen, H. (2011). Using airborne light detection and ranging as a sampling tool for estimating forest biomass resources in the Upper Tanana Valley of Interior Alaska. *Western Journal of Applied Forestry*, 26(4), 157-164. doi: 10.1093/wjaf/26.4.157
- Asner, G. P. (2009). Tropical forest carbon assessment: integrating satellite and airborne mapping approaches.



- Environmental Research Letters*, 4(3). doi: 10.1088/1748-9326/4/3/034009
- Asner, G. P., Mascaro, J., Muller-Landau, H. C., Vieilledent, G., Vaudry, R., Rasamoelina, M., & van Breugel, M. (2012). A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia*, 168(4), 1147-1160. doi: 10.1007/s00442-011-2165-z
- Baccini, Goetz, S., Walker, W., Laporte, N., Sun, M., Sulla-Menashe, D., & Friedl, M. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*, 2(3), 182. doi: 10.1038/nclimate1354
- Baccini, Laporte, N., Goetz, S., Sun, M., & Dong, H. (2008). A first map of tropical Africa's above-ground biomass derived from satellite imagery. *Environmental Research Letters*, 3(4), 045011. doi: 10.1088/1748-9326/3/4/045011
- Baltsavias, E. P. (1999). Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2), 199-214. doi: 10.1016/S0924-2716(99)00015-5
- Barbosa, J. M., Broadbent, E. N., & Bitencourt, M. D. (2014). Remote Sensing of Aboveground Biomass in Tropical Secondary Forests: A Review. *International Journal of Forestry Research*, 2014, 1-14. doi: 10.1155/2014/715796
- Cairns, M. A., Olmsted, I., Granados, J., & Argaez, J. (2003). Composition and aboveground tree biomass of a dry semi-evergreen forest on Mexico's Yucatan Peninsula. *Forest Ecology and Management*, 186(1-3), 125-132. doi: 10.1016/s0378-1127(03)00229-9
- Cao, L., Coops, N. C., Innes, J., Dai, J., & She, G. (2014). Mapping above-and below-ground biomass components in subtropical forests using small-footprint LiDAR. *Forests*, 5(6), 1356-1373. doi: 10.3390/f5061356
- Cartus, O., Kellndorfer, J., Walker, W., Franco, C., Bishop, J., Santos, L., & Fuentes, J. (2014). A National, Detailed Map of Forest Aboveground Carbon Stocks in Mexico. *Remote Sensing*, 6(6), 5559-5588. doi: 10.3390/rs6065559
- Chave, Condit, R., Aguilar, S., Hernandez, A., Lao, S., & Perez, R. (2004). Error propagation and scaling for tropical forest biomass estimates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1443), 409-420. doi: 10.1098/rstb.2003.1425
- Chave, Condit, R., Lao, S., Caspersen, J. P., Foster, R. B., & Hubbell, S. P. (2003). Spatial and temporal variation of biomass in a tropical forest: results from a large census plot in Panama. *Journal of Ecology*, 91(2), 240-252. doi: 10.1046/j.1365-2745.2003.00757.x
- Chen, Q., Lu, D., Keller, M., dos-Santos, M., Bolfe, E., Feng, Y., & Wang, C. (2015). Modeling and Mapping Agroforestry Aboveground Biomass in the Brazilian Amazon Using Airborne Lidar Data. *Remote Sensing*, 8(12). doi: 10.3390/rs8010021
- Chi, H., Sun, G., Huang, J., Li, R., Ren, X., Ni, W., & Fu, A. (2017). Estimation of Forest Aboveground Biomass in Changbai Mountain Region Using ICESat/GLAS and Landsat/TM Data. *Remote Sensing*, 9(12). doi: 10.3390/rs9070707
- Clark, M. L., Roberts, D. A., Ewel, J. J., & Clark, D. B. (2011). Estimation of tropical rain forest aboveground biomass with small-footprint lidar and hyperspectral sensors. *Remote Sensing of Environment*, 115(11), 2931-2942. doi: 10.1016/j.rse.2010.08.029
- Comisión Nacional Forestal [Conafor]. (2011). Manual y procedimientos para el muestreo de campo. Remuestreo 2012. Inventario Nacional Forestal y de Suelos. Zapopan, Jalisco, México: Conafor. Secretaría del Medio Ambiente y Recursos Naturales.
- Comisión Nacional Forestal [Conafor]. (2015). Base de datos del Inventario Nacional Forestal y de Suelos Remedición 2009-2014.
- Cook, B., Corp, L., Nelson, R., Middleton, E., Morton, D., McCorkel, J., & Montesano, P. (2013). NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. *Remote Sensing*, 5(12), 4045-4066. doi: 10.3390/rs5084045
- Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random forests. En C. Zhang, & Y. Ma (Eds.), *Ensemble machine learning: Methods and applications* (pp. 157-175). Nueva York, Estados Unidos: Springer.
- d'Oliveira, M. V., Reutebuch, S. E., McGaughey, R. J., & Andersen, H.-E. (2012). Estimating forest biomass and identifying low-intensity logging areas using airborne scanning lidar in Antimary State Forest, Acre State, Western Brazilian Amazon. *Remote Sensing of Environment*, 124, 479-491. doi: 10.1016/j.rse.2012.05.014
- Deo, R., Russell, M., Domke, G., Andersen, H.-E., Cohen, W., & Woodall, C. (2017). Evaluating Site-Specific and Generic Spatial Models of Aboveground Forest Biomass Based on Landsat Time-Series and LiDAR Strip Samples in the Eastern USA. *Remote Sensing*, 9(12). doi: 10.3390/rs9060598
- Dupuy, J. M., Hernández-Stefanoni, J. L., Hernández-Juárez, R. A., Tetetla-Rangel, E., López-Martínez, J. O., Leyequién-Abarca, E., & May-Pat, F. (2012). Patterns and correlates of tropical dry forest structure and composition in a highly

- replicated chronosequence in Yucatan, Mexico. *Biotropica*, 44(2), 151-162.
- Echeverría, F., Arreola, J., Esparza, L. G., Morales, V., y López, J. L. (2014). Análisis de la composición y estructura de la selva mediana subperennifolia del Ejido Pachuitz, Hopelchén, Campeche, México. *Foresta Veracruzana*, 16(2).
- Foody, G. M., Boyd, D. S., & Cutler, M. E. J. (2003). Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. *Remote Sensing of Environment*, 85(4), 463-474. doi: 10.1016/s0034-4257(03)00039-7
- Goetz, S. & Dubayah, R. (2011). Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. *Carbon Management*, 2(3), 231-244. doi: 10.4155/cmt.11.18
- Gonçalves, F., Treuhaft, R., Law, B., Almeida, A., Walker, W., Baccini, A., & Graça, P. (2017). Estimating Aboveground Biomass in Tropical Forests: Field Methods and Error Analysis for the Calibration of Remote Sensing Observations. *Remote Sensing*, 9(12). doi: 10.3390/rs9010047
- Hansen, A. J., Phillips, L. B., Dubayah, R., Goetz, S., & Hofton, M. (2014). Regional-scale application of lidar: Variation in forest canopy structure across the southeastern US. *Forest Ecology and Management*, 329, 214-226. doi: 10.1016/j.foreco.2014.06.009
- Hernández-Stefanoni, J., Dupuy, J., Johnson, K., Birdsey, R., Tun-Dzul, F., Peduzzi, A., y López-Merlín, D. (2014). Improving Species Diversity and Biomass Estimates of Tropical Dry Forests Using Airborne LiDAR. *Remote Sensing*, 6(12), 4741-4763. doi: 10.3390/rs6064741
- Houghton, R. A., Hall, F., & Goetz, S. J. (2009). Importance of biomass in the global carbon cycle. *Journal of Geophysical Research: Biogeosciences*, 114(G2), doi: 10.1029/2009jg000935
- Instituto Nacional de Estadística y Geografía [INEGI]. (2013). *Conjunto Nacional de Uso del Suelo y Vegetación a escala 1:250,000*. Aguascalientes, México: INEGI.
- Laurin, G. V., Puletti, N., Chen, Q., Corona, P., Papale, D., & Valentini, R. (2016). Above ground biomass and tree species richness estimation with airborne lidar in tropical Ghana forests. *International Journal of Applied Earth Observation and Geoinformation*, 52, 371-379. doi: 10.1016/j.jag.2016.07.008
- Lefsky, M. A., Cohen, W. B., Harding, D. J., Parker, G. G., Acker, S. A., & Gower, S. T. (2002). Lidar remote sensing of above-ground biomass in three biomes. *Global ecology and biogeography*, 11(5), 393-399. doi: 10.1046/j.1466-822x.2002.00303.x
- Li, M., Im, J., Quackenbush, L. J., & Liu, T. (2014). Forest Biomass and Carbon Stock Quantification Using Airborne LiDAR Data: A Case Study Over Huntington Wildlife Forest in the Adirondack Park. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(7), 3143-3156. doi: 10.1109/jstars.2014.2304642
- Lu, D., Chen, Q., Wang, G., Moran, E., Batistella, M., Zhang, M., & Saah, D. (2012). Aboveground Forest Biomass Estimation with Landsat and LiDAR Data and Uncertainty Analysis of the Estimates. *International Journal of Forestry Research*, 2012, 1-16. doi: 10.1155/2012/436537
- Mascaro, Asner, G. P., Knapp, D. E., Kennedy-Bowdoin, T., Martin, R. E., Anderson, C., & Chadwick, K. D. (2014). A tale of two "forests": random forest machine learning AIDS tropical forest carbon mapping. *PLoS One*, 9(1), 5993. doi: 10.1371/journal.pone.0085993
- Mascaro, Detto, M., Asner, G. P., & Muller-Landau, H. C. (2011). Evaluating uncertainty in mapping forest carbon with airborne LiDAR. *Remote Sensing of Environment*, 115(12), 3770-3774. doi: 10.1016/j.rse.2011.07.019
- McGaughey, R. J. (2016). *FUSION/LDV: software for LiDAR data analysis and visualization*. US Department of Agriculture, Forest Service, Pacific Northwest Research Station: Seattle, WA, USA.
- Meyer, V., Saatchi, S. S., Chave, J., Dalling, J. W., Bohlman, S., Fricker, G. A., & Hubbell, S. (2013). Detecting tropical forest biomass dynamics from repeated airborne lidar measurements. *Biogeosciences*, 10(8), 5421-5438. doi: 10.5194/bg-10-5421-2013
- Næsset, E. & Økland, T. (2002). Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment*, 79(1), 105-115. doi: 10.1016/S0034-4257(01)00243-7
- Nelson, R., Margolis, H., Montesano, P., Sun, G., Cook, B., Corp, L., & Prisley, S. (2017). Lidar-based estimates of aboveground biomass in the continental US and Mexico using ground, airborne, and satellite observations. *Remote Sensing of Environment*, 188, 127-140. doi: 10.1016/j.rse.2016.10.038
- Ningthoujam, R. K., Joshi, P. K., & Roy, P. S. (2018). Retrieval of forest biomass for tropical deciduous mixed forest using ALOS PALSAR mosaic imagery and field plot data. *International Journal of Applied Earth Observation and Geoinformation*, 69, 206-216. doi: 10.1016/j.jag.2018.03.007
- Ortiz-Reyes, A. D., Valdez-Lazalde, J. R., los Santos-Posadas, D., Héctor, M., Ángeles-Pérez, G., Paz-Pellat, F., y Martínez-Trinidad, T. (2015). Inventario y cartografía de variables del bosque con datos derivados de LiDAR: comparación de



- métodos. *Madera y Bosques*, 21(3), 111-128. doi: 10.21829/myb.2015.213461
- Peduzzi, A., Wynne, R. H., Fox, T. R., Nelson, R. F., & Thomas, V. A. (2012). Estimating leaf area index in intensively managed pine plantations using airborne laser scanner data. *Forest Ecology and Management*, 270, 54-65. doi: 10.1016/j.foreco.2011.12.048
- Pennington, T. D. (2005). *Árboles tropicales de México: manual para la identificación de las principales especies*. México, D.F.: UNAM.
- Phua, M.-H., Johari, S. A., Wong, O. C., Ioki, K., Mahali, M., Nilus, R., & Hashim, M. (2017). Synergistic use of Landsat 8 OLI image and airborne LiDAR data for above-ground biomass estimation in tropical lowland rainforests. *Forest Ecology and Management*, 406, 163-171. doi: 10.1016/j.foreco.2017.10.007
- Puc-Kaul, R., Ángeles-Pérez, G., Valdez-Lazalde, J. R., Reyes-Hernández, V. J., Dupuy-Rada, J. M., Schneider, L., García-Cuevas, X., Pérez-Rodríguez, P. (2019, aceptado para publicación). Species-specific biomass equations for small-size tree species in secondary tropical forests. *Tropical and Subtropical Agroecosystems*, 22(x): xx-xx.
- R Development Core Team. (2013). R: A language and environment for statistical computing, 2009-2017.
- Read, L. & Lawrence, D. (2003). Recovery of biomass following shifting cultivation in dry tropical forests of the Yucatan. *Ecological Applications*, 13(1), 85-97.
- Rocha de Souza, P. F., Kampel, M., Gomes Soares, M., Estrada, G., Bentz, C., & Vincent, G. (2018). Reducing Uncertainty in Mapping of Mangrove Aboveground Biomass Using Airborne Discrete Return Lidar Data. *Remote Sensing*, 10(4). doi: 10.3390/rs10040637
- Rodríguez-Veiga, P., Saatchi, S., Tansey, K., & Balzter, H. (2016). Magnitude, spatial distribution and uncertainty of forest biomass stocks in Mexico. *Remote Sensing of Environment*, 183, 265-281. doi: 10.1016/j.rse.2016.06.004
- Saarela, S., Grafström, A., Ståhl, G., Kangas, A., Holopainen, M., Tuominen, S., & Hyypä, J. (2015). Model-assisted estimation of growing stock volume using different combinations of LiDAR and Landsat data as auxiliary information. *Remote Sensing of Environment*, 158, 431-440. doi: 10.1016/j.rse.2014.11.020
- Saatchi, S. S., Houghton, R., Dos Santos Alvala, R., Soares, J. V., & Yu, Y. (2007). Distribution of aboveground live biomass in the Amazon basin. *Global Change Biology*, 13(4), 816-837. doi: 10.1111/j.1365-2486.2007.01323.x
- Urbazev, M., Thiel, C., Cremer, F., Dubayah, R., Migliavacca, M., Reichstein, M., & Schmullius, C. (2018). Estimation of forest aboveground biomass and uncertainties by integration of field measurements, airborne LiDAR, and SAR and optical satellite data in Mexico. *Carbon Balance and Management*, 13(1), 5. doi: 10.1186/s13021-018-0093-5
- Urquiza-Haas, T., Dolman, P. M., & Peres, C. A. (2007). Regional scale variation in forest structure and biomass in the Yucatan Peninsula, Mexico: Effects of forest disturbance. *Forest Ecology and Management*, 247(1-3), 80-90. doi: 10.1016/j.foreco.2007.04.015
- Véga, C., Vepakomma, U., Morel, J., Bader, J.-L., Rajashekar, G., Jha, C., & Dadhwal, V. (2015). Aboveground-Biomass Estimation of a Complex Tropical Forest in India Using Lidar. *Remote Sensing*, 7(12), 10607-10625. doi: 10.3390/rs70810607
- Wang, C. & Qi, J. (2008). Biophysical estimation in tropical forests using JERS-1 SAR and VNIR imagery. II. Aboveground woody biomass. *International Journal of Remote Sensing*, 29(23), 6827-6849. doi: 10.1080/01431160802270123
- White, J. C., Wulder, M., Vastaranta, M., Coops, N., Pitt, D., & Woods, M. (2013). The Utility of Image-Based Point Clouds for Forest Inventory: A Comparison with Airborne Laser Scanning. *Forests*, 4(3), 518-536. doi: 10.3390/f4030518
- Wilkes, P., Jones, S. D., Suarez, L., Mellor, A., Woodgate, W., Soto-Berelev, M., & Skidmore, A. K. (2015). Mapping forest canopy height across large areas by upscaling ALS estimates with freely available satellite data. *Remote Sensing*, 7(9), 12563-12587. doi: 10.3390/rs70912563
- Woods, M., Pitt, D., Penner, M., Lim, K., Nesbitt, D., Etheridge, D., & Treitz, P. (2011). Operational implementation of a LiDAR inventory in Boreal Ontario. *The Forestry Chronicle*, 87(4), 512-528.
- Xu, L., Saatchi, S. S., Yang, Y., Yu, Y., & White, L. (2016). Performance of non-parametric algorithms for spatial mapping of tropical forest structure. *Carbon Balance Manag*, 11(1), 18. doi: 10.1186/s13021-016-0062-9
- Zamora-Crescencio, P., Rico-Gray, V., Barrientos-Medina, R. C., Puc-Garrido, E. C., Villegas, P., Domínguez-Carrasco, M. d. R., y Gutiérrez-Báez, C. (2017). Estructura y composición florística de la selva mediana subperennifolia en Bethania, Campeche, México. *Polibotánica*, 43, 67-86.
- Zanne, A., Lopez-Gonzalez, G., Coomes, D., Ilic, J., Jansen, S., Lewis, S., & Chave, J. (2013). *Global wood density database 2009*.
- Zolkos, S. G., Goetz, S. J., & Dubayah, R. (2013). A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sensing of Environment*, 128, 289-298. doi: 10.1016/j.rse.2012.10.017

Manuscrito recibido el 1 de octubre de 2018

Aceptado el 19 de febrero de 2019

Publicado el 13 de noviembre de 2019

Este documento se debe citar como:

Ortiz-Reyes, A. D., Valdez-Lazalde, J. R., Ángeles-Pérez, G., De los Santos-Posadas, H. M., Schneider, L., Aguirre-Salado, C. A., & Peduzzi,

A. (2019). Transectos de datos LiDAR: una estrategia de muestreo para estimar biomasa aérea en áreas forestales. *Madera y Bosques*, 25(3), e2531872. doi: 10.21829/myb.2019.2531872



Madera y Bosques por Instituto de Ecología, A.C. se distribuye bajo una Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional.