



Inventario forestal asistido por LIDAR: efecto de la densidad de retornos y el diseño de muestreo sobre la precisión

LIDAR-assisted forest inventory: effect of return density and sampling design on accuracy

Bernardo Galeote-Leyva¹, José René Valdez-Lazalde^{2*}, Gregorio Ángeles-Pérez²,
Héctor Manuel de los Santos-Posadas² y Juan Manuel Romero-Padilla³

¹ Instituto Tecnológico Superior de Venustiano Carranza, Ciudad Lázaro Cárdenas, Puebla, México.

² Colegio de Postgraduados, Campus Montecillo, Posgrado en Ciencias Forestales, Montecillo, Estado de México, México.

³ Colegio de Postgraduados, Campus Montecillo, Posgrado en Estadística, Montecillo, Estado de México, México.

* Autor de correspondencia. valdez@colpos.mx

RESUMEN

El uso combinado de datos de campo y sensores remotos para la realización de inventarios forestales es un tema de interés actual. Uno de los retos importantes para su aplicación práctica consiste en optimizar/minimizar el volumen de datos a utilizar para lograr estimaciones aceptables. En este estudio se analizó el efecto del diseño de muestreo y la densidad de retornos LIDAR sobre la precisión del área basal (AB), el volumen maderable (V) y la biomasa (B), además de estimadores de muestreo asistidos por modelos aditivos generalizados (GAM) y el algoritmo *random forest* (RF) para un bosque bajo manejo ubicado Zacualtipán, Hidalgo. Se dispuso de 96 sitios de muestreo en campo (400 m²), tres diseños de muestreo LIDAR y 10 densidades de retornos. Se analizaron los estimadores en dos fases y dos etapas para estimar el inventario total. Los modelos GAM demostraron ser eficientes en la estimación (0.76 a 0.92 de R²) de las variables forestales a escala de transecto LIDAR. El algoritmo RF mostró bondades de ajuste aceptables (0.71 a 0.79 de R²) para estimar las variables a escala de área de estudio. Los estimadores asistidos por regresión presentaron una buena precisión teniendo un error menor a 6% en el inventario de las variables evaluadas. Se demostró que las muestras por transectos de datos LIDAR son una alternativa viable para la estimación de variables de interés forestal en predios bajo manejo.

PALABRAS CLAVE: área basal, biomasa, GAM, *random forests*, *RapidEye*, volumen maderable.

ABSTRACT

The combined use of field data and remote sensing to carry out forest inventories is a topic of current interest. One of the important challenges for its practical application is to optimize/minimize the volume of data to be used to achieve acceptable estimates. In this study, we analyzed the effect of the sampling design and density of LIDAR returns on the accuracy of basal area (AB), timber volume (V), and biomass (B), in addition to sampling estimators assisted by generalized additive models (GAM) and the Random Forest (RF) algorithm for a forest under management located in Zacualtipán, Hidalgo. There were 96 field sampling sites (400 m²), three LIDAR sampling designs, and 10 return densities. Two-phase, two-stage estimators were analyzed to estimate the total inventory. The GAM models proved to be efficient in estimating (0.76 to 0.92 of R²) forest variables at the LIDAR transect level. The RF algorithm showed acceptable goodness of fit (0.71 to 0.79 R²) for estimating variables at the study area level. The regression-assisted estimators showed good accuracy with an error of less than 6% in the inventory of the evaluated variables. It was demonstrated that transect sampling of LIDAR data is a viable alternative for the estimation of variables of forest interest in managed properties.

KEYWORDS: basal area, biomass, GAM, *random forests*, *RapidEye*, timber volume.

INTRODUCCIÓN

El inventario del volumen maderable o la biomasa contenidos en los bosques es considerado un proceso básico en su gestión, permite medir y monitorear a escala local los cambios que suceden en este tipo de ecosistemas (Urbazaev et al., 2018). Su realización implica la medición de árboles ubicados en unidades de muestreo de dimensión fija, comúnmente distribuidas en una red sistemática sobre un área determinada (Grafström y Ringvall, 2013). En las últimas décadas se ha impulsado el uso combinado de datos medidos en campo y datos colectados mediante sensores remotos para la realización de este tipo de inventarios (Grafström y Ringvall, 2013; Crespo-Peremarch et al., 2018).

La tecnología LIDAR (Light Detection and Ranging) aerotransportada ha demostrado utilidad para estimar variables biométricas del bosque con precisión aceptable a partir del registro de retornos (rebotes) de un haz de luz previamente emitido por un equipo especializado (Ortiz-Reyes et al., 2015; Marchi et al., 2018). Desafortunadamente, la adquisición de datos de LIDAR es costosa en comparación con datos provenientes de sensores ópticos o de radar. El costo está influenciado por el tiempo de vuelo y por la densidad de retornos por metro cuadrado que se desea registrar, los cuales a su vez varían con la velocidad, la altitud y el ancho de barrido del vuelo (Hansen Hofstad et al., 2015).

En años recientes se han estudiado estrategias para abaratar el costo de usar datos LIDAR en la estimación de inventarios forestales; particularmente la reducción de la densidad de retornos por metro cuadrado y el uso de transectos en lugar de datos de cobertura completa (Almeida et al., 2019; Ortiz-Reyes et al., 2019). Evaluaciones del efecto de la densidad de retornos en la precisión de la estimación de variables como la altura, diámetro cuadrático, área basal, volumen y biomasa, indican que las densidades óptimas varían de 0.5 m^{-2} a 15 m^{-2} en plantaciones, bosques tropicales (Hansen Hofstad et al., 2015; Silva et al., 2017; Almeida et al., 2019) y en bosques templados puros o mixtos (Treit et al., 2012; Singh et al., 2015). Respecto al uso de datos LIDAR para solo una

fracción del área a inventariar (a manera de una muestra), y a partir de esta expandir las estimaciones mediante modelos matemáticos alimentados con datos de otras imágenes de costo menor, se han publicado resultados prometedores (Ortiz-Reyes et al., 2019). Aunque el método tiene limitaciones debido a la trayectoria lineal de la plataforma que porta el instrumento (avión) (Wulder et al., 2012), éste es actualmente un procedimiento común en la selección de las muestras de campo, además de optimista para la estimación de variables de interés en el inventario (Mahoney et al., 2018; Ver et al., 2018). De acuerdo con Chen et al. (2012) y Wulder et al. (2012) la adquisición de una muestra de datos LIDAR en transectos, en combinación con un marco de muestreo apropiado, genera estadísticos precisos y de bajo costo para áreas extensas.

Los transectos LIDAR pueden incorporarse a un diseño de muestreo en dos etapas en el que las unidades de muestreo primarias (UMP) son los transectos y las secundarias (UMS) son parcelas medidas en campo (Gregoire et al., 2011; Wulder et al., 2012; Strimbu et al., 2017). Mediante un estimador complejo en dos etapas, se puede estimar la media general de variables de interés. Para este propósito regularmente se han utilizado métodos paramétricos como la regresión lineal (Strimbu et al., 2017) y no lineal (Saarela et al., 2015), además de modelos semiparamétricos o no paramétricos (McRoberts et al., 2011; Kangas et al., 2016).

Aunque el uso de estimadores complejos representa una solución para la estimación del inventario de variables de interés forestal a través de transectos LIDAR, paralelamente se ha propuesto la extracción de métricas derivadas de sensores ópticos para combinarlos con datos LIDAR y con ello estimar parámetros forestales a escala regional (Chen et al., 2012) y nacional. En México se han realizado estudios a escala nacional, combinando imágenes satelitales con datos LIDAR colectados en transectos para la estimación de biomasa aérea (Nelson et al., 2017; Urbazaev et al., 2018). A escala regional este enfoque se ha explorado para la estimación de biomasa aérea, carbono, riqueza y diversidad de especies en la península de Yucatán (Hernández-Stefanoni et al., 2014; Ochoa-Franco et al.,



2019; George-Chacon et al., 2019; Ortiz-Reyes et al., 2019). Si bien esto demuestra el interés nacional por fomentar el uso de estas herramientas de teledetección y enfoques de muestreo para la estimación de variables forestales, su uso es todavía limitado, particularmente para la estimación de inventarios operativos.

Este estudio tuvo como propósito analizar de manera conjunta el efecto del diseño de muestreo LIDAR y la densidad de retornos por metro cuadrado sobre la precisión del inventario de área basal (AB), volumen de madera (V) y biomasa (B) en un bosque templado bajo manejo ubicado en Zacualtipán, Hidalgo.

OBJETIVOS

Los objetivos de este estudio fueron analizar el efecto de diferentes diseños de muestreo y densidades de retornos (puntos) por unidad de área (m^2) sobre la estimación del inventario de AB, V y B en áreas con datos LIDAR y combinar los resultados del mejor diseño encontrado para estimar y mapear el inventario total de AB, V y B en un predio forestal mediante estimadores de muestreo complejos asistidos por modelos aditivos generalizados (GAM).

MATERIALES Y MÉTODOS

Área de estudio

El estudio se realizó en el área forestal que conforma el Sitio de Monitoreo Intensivo de Carbono (SMIC) en Atopixco, localizado en el municipio de Zacualtipán, Hidalgo, entre las coordenadas $20^{\circ}40'17''$ y $20^{\circ}34'51''$ latitud norte, $98^{\circ}40'07''$ y $98^{\circ}34'22''$ longitud oeste (Fig. 1). Su fisiografía comprende parte de la Sierra Alta Hidalguense y del Eje Neovolcánico. La precipitación promedio anual es de 1601 mm, con una temperatura media anual de $17^{\circ}C$ y clima templado húmedo con lluvias todo el año (Instituto Nacional de Estadística y Geografía [Inegi], 2017). El SMIC tiene una superficie de 900 ha cubiertas mayoritariamente de bosque coetáneo. Lo conforman rodales dominados por *Pinus patula* con edad variable desde 0 años hasta 31 años y

algunas áreas sin intervención silvícola en los últimos 80 años (naturales) (Torres-Vivar et al., 2017).

Datos de campo: estimación de atributos forestales

Se establecieron 24 unidades de muestreo primarias (UMP) de inventario permanente con arreglo sistemático y 17 adicionales siguiendo un gradiente de años transcurridos a partir de la cosecha de los rodales para capturar la variabilidad del paisaje forestal en el área de estudio (Fig. 1). Cada UMP está conformada por subunidades o unidades de muestreo secundarias (UMS), en ellas se ubicó el centro mediante un GPS Garmin® ($\pm 3m$). Las UMS estuvieron dispuestas de forma espacial en forma de “Y” invertida como el diseño del Inventario Nacional Forestal y de Suelos (INFyS) (Comisión Nacional Forestal [Conafor], 2011). En total se dispuso de 164 UMS circulares de $400 m^2$, pero solo se utilizaron 96 UMS, dejando fuera del análisis aquellas que no se ubicaron dentro de los diseños propuestos. En los sitios se midió el diámetro normal (DN) y la altura total (H) de todos los árboles con $DN \geq 7.5$ cm. Con estos datos se estimó su área basal (AB; m^2 ; Conafor, 2011), su volumen maderable (V; m^3) y su biomasa (B; kg), utilizando las ecuaciones para especies latifoliadas y pinos presentadas por Cruz-Martinez (2007) y Soriano-Luna (2014), respectivamente.

Datos del sensor LIDAR

Los datos LIDAR fueron obtenidos en mayo del 2013 con el sensor Riegl-VQ480 a una altitud de 397 m, una frecuencia de 200 kHz, un ángulo de barrido de $\pm 15^{\circ}$ y una intensidad de retornos de $5 m^{-2}$ a $10 m^{-2}$. Los retornos registrados permitieron generar una “nube de puntos” tridimensional, la cual se constituyó como la base de datos LIDAR (Ortiz-Reyes et al., 2015). Además, se generó un modelo digital de elevación (MDE) de 1 m de resolución para normalizar las alturas de los árboles sobre el suelo. La figura 2 muestra los pasos metodológicos generales implementados en el presente estudio. En los párrafos siguientes se detalla el procedimiento.

Diseños de muestreo LIDAR

Se definieron tres diseños de muestreo a partir de la ubicación en campo de las parcelas del inventario. Dos diseños consistieron en transectos de datos LIDAR de 200 m de ancho y 2.8 km de largo ubicados con dirección este-oeste y norte-sur (Fig. 1a y 1b, respectivamente). El tercer diseño correspondió a muestras de 400 m² ubicadas en rodales de edad diferente (Fig. 1c), previendo que en el futuro será factible coleccionar este tipo de nubes de puntos LIDAR mediante sensores montados en drones. Los tres diseños de muestreo fueron recortados con la herramienta *PolyClipData* de la aplicación FUSION (McGaughey, 2018) de los datos LIDAR disponibles para el SMIC, utilizando el mismo número de UMS (96 sitios).

Reducción de la densidad de retornos LIDAR

En cada uno de los sitios de 400 m² la densidad de retornos LIDAR original se redujo para simular diferentes

características de adquisición de datos. Esta reducción se realizó con la función *lasfilterdecimate (random)* de la paquetería *lidR* (Roussel et al., 2020) del software R (R Team Core, 2019). Se implementaron dos niveles de densidad: baja (0.2 m⁻², 0.4 m⁻², 0.6 m⁻², 0.8 m⁻² y 1 m⁻² retornos) y alta (2 m⁻², 4 m⁻², 6 m⁻², 8 m⁻² y 10 m⁻² retornos).

Métricas LIDAR

Se calcularon 40 métricas LIDAR (McGaughey, 2018) a nivel de sitio (400 m²) y de transecto para las 10 densidades de retornos evaluadas. Se utilizó una altura de corte de 3 m con el propósito de eliminar la variación (ruido) generada por la vegetación del sotobosque. Los transectos LIDAR para cada densidad de retornos evaluada fueron convertidos a formato Raster con la herramienta *Csv2Grid* de FUSION, con tamaño de pixel de 20 m × 20 m (400 m²), similar al tamaño de los sitios medidos en campo.

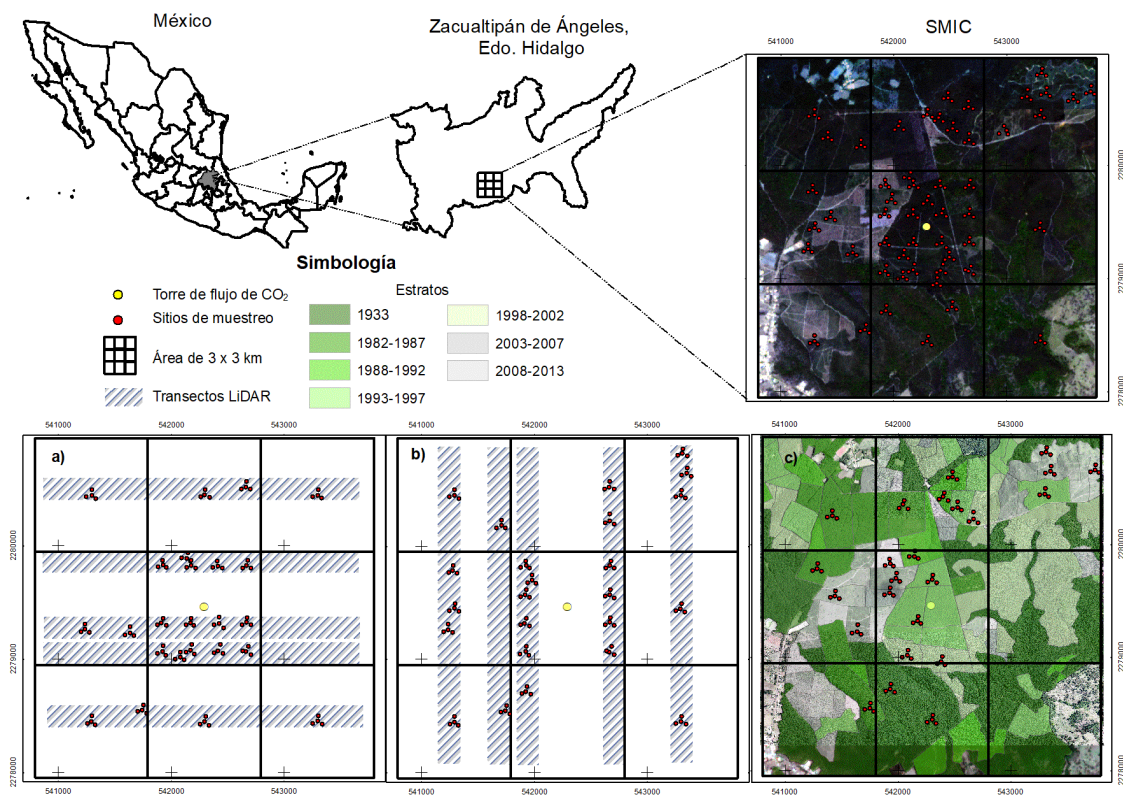


FIGURA 1. Mapa del Sitio de Monitoreo Intensivo de Carbono Atopixco y diseños de muestreo LIDAR.

a) transectos en dirección este-oeste; b) transectos en dirección norte-sur; y c) estratificado (estratos definidos con base en el año (periodo) de cosecha del rodal).

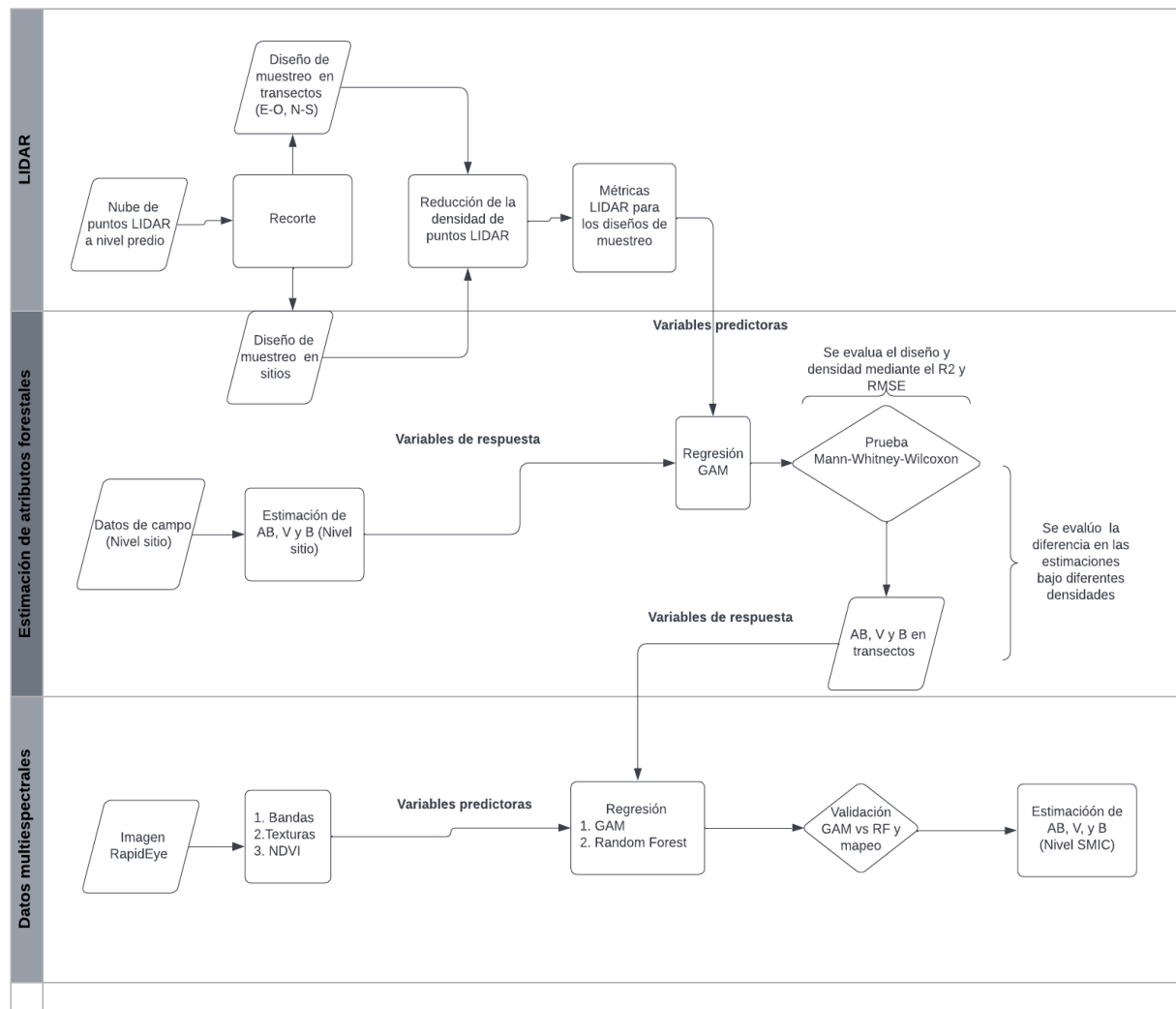


FIGURA 2. Flujo de trabajo utilizado para la modelización espacial de variables de inventario forestal mediante diferentes densidades de retornos de datos LIDAR y espectrales.

Datos multiespectrales: procesamiento de imágenes *RapidEye*

Se utilizaron dos imágenes *RapidEye* con nivel de procesamiento 3A tomadas el 23 de marzo y 13 de abril del año 2013. Estas imágenes tuvieron una resolución radiométrica de 12 bits y se reescalaron a un rango dinámico de 16 bits. La escala se realizó con un factor constante que convierte los números digitales (ND) del sensor en valores absolutos de radiancias (Planet Labs Inc, 2016). Posteriormente, se calculó la reflectancia en la parte superior de la atmósfera (TOA) conforme al procedimiento detallado por Planet Labs Inc. (2016). Las dos imágenes que cubren el área de

estudio fueron normalizadas mediante el procedimiento *Histogram matching* (Helmer y Rufenacht, 2005; Hong y Zhang, 2008). Este procedimiento se implementó en el software R con la función *histMatch* de la paquetería *RStoolbox* (Leutner et al., 2019).

A partir de los datos de reflectancia se calculó el índice de vegetación de diferencias normalizadas (NDVI) (Kim y Yeom, 2015) y 42 métricas de textura espectral (media, varianza, homogeneidad, contraste, disimilitud, entropía y segundo momento) y de coocurrencia de niveles de grises (GLCM) que consideran la relación espacial entre grupos de celdas adyacentes en una determinada ventana (15 × 15;

Ochoa-Franco et al., 2019). El proceso se aplicó a las cinco bandas (R, G, B, Red edge e Infrarrojo) y al NDVI mediante el software R con la paquetería GLCM (Zvoleff, 2020). Posteriormente, las métricas fueron remuestreadas, con el método del vecino más cercano, para empatar el tamaño de las celdas de las imágenes con la dimensión de los sitios de muestreo en campo (400 m²), *i.e.*, se transformó el tamaño de las celdas de 5 m × 5 m a 20 m × 20 m de resolución espacial.

Análisis estadístico: estimación de variables forestales a nivel transectos LIDAR y SMIC

Estimación de área basal, volumen y biomasa en los transectos LIDAR

Inicialmente se identificaron correlaciones relevantes entre AB, V y B calculadas a partir de datos de campo y las métricas LIDAR generadas a partir de la nube de puntos. Posteriormente se ajustaron modelos aditivos generalizados (GAM por sus siglas en inglés) para estimar, a nivel pixel (20 m × 20 m), valores de AB, V y B en el área muestreada mediante transectos LIDAR. Lo anterior con base en evidencia que señala que este tipo de modelos son más consistentes y de mejor precisión que otros modelos utilizados para estimar variables continuas (*e.g.* modelos lineales, no lineales, *random forest*) (Soriano-Luna et al., 2018). Los datos de AB, V y B colectados en campo se definieron como las variables dependientes y las métricas LIDAR representaron a los predictores. El ajuste de los modelos se realizó en la paquetería *mgcv* mediante la función *gam* del software R (Team R Core, 2019).

El efecto de la densidad de retornos sobre las estimaciones de AB, V y B a escala de sitio, se evaluó mediante el coeficiente de determinación (R²) y la raíz del error medio al cuadrado (RMSE) asociados a los modelos ajustados a partir de la validación cruzada repetida 30 veces de la muestra inicial (LOOCV). Estos valores fueron representados en diagramas de caja para comparar la variabilidad de las estimaciones y determinar la densidad de retornos LIDAR más adecuada para el mapeo de toda el

área de estudio (SMIC), y en la estimación del inventario total mediante estimadores asistidos por modelos. Para comprobar si las diferencias entre las densidades de retornos simuladas y la original eran significativas, se realizó una prueba de Mann-Whitney-Wilcoxon de dos caras.

Estimación de área basal, volumen y biomasa para el SMIC con imágenes RapidEye

Los valores de AB, V y B estimados previamente para los transectos se consideraron en esta fase como variables dependientes, y como nuevos predictores los valores registrados en las cinco bandas espectrales, las métricas de textura calculadas (Zvoleff, 2020), y los valores de NDVI. Se excluyeron del análisis los pixeles con valores atípicos de las estimaciones de AB, V y B, *i.e.*, aquellos con magnitud mayor al cuartil 3 por al menos 1.5 veces el rango intercuartil. Se eliminaron 6%, 2% y 3% de los datos para AB, V y B, respectivamente. Lo anterior para tener un mejor ajuste de los modelos a escala de área de estudio total (SMIC) y evitar sobrestimaciones.

Se comparó el uso de modelos GAM *versus* el algoritmo de aprendizaje automático *random forest* (RF), con el objetivo de identificar la mejor opción para determinar la distribución espacial de las variables de interés. Dado que el número de observaciones fue grande (7000 valores), 90% de los datos se usaron para el ajuste/entrenamiento de los modelos y 10% para la validación. Los modelos que resultaron con los mejores parámetros de bondad de ajuste (R² y RMSE) se utilizaron para mapear en la totalidad del área del SMIC, *i.e.*, las áreas no cubiertas por el sensor LIDAR. El ajuste de los modelos y el mapeo de las variables se realizó en el software R (Team R Core, 2019), paqueterías *mgcv* (ajuste GAM), *randomForest* y *raster* para la construcción de los mapas.

Estimadores muestrales e inventario total

Estimador en dos fases

El diseño en dos fases (Mandallaz, 2008) permitió asignar inicialmente una muestra grande n_1 a la primera fase (S_1),



representada por los datos de transectos LIDAR (6457 celdas de tamaño de 400 m²), distribuida uniformemente dentro del área forestal. En la segunda fase (S_2), se dispuso de los datos de inventario de las variables AB, V y B [$Y(x_2)$] tomados en campo mediante una muestra más pequeña n_2 (96 sitios de muestreo). Con base en las muestras n_1 y n_2 se ajustó un modelo de regresión GAM, identificado previamente como la mejor opción para relacionar los datos LIDAR y los datos de inventario. Para estimar la media de las variables AB, V y B, se utilizó el estimador por diferencia para el estimador en dos fases (Mandallaz et al., 2013; ecuación 1):

$$\bar{Y}_{Reg} = \frac{1}{n_1} \sum_{x \in S_1} \hat{Y}(x_1) + \frac{1}{n_2} \sum_{x \in S_2} (Y(x_2) - \hat{Y}(x_2)) \quad (1)$$

donde:

\bar{Y}_{Reg} : media estimada

n_1 : muestra intensiva en la primera fase (transectos)

n_2 : muestra de la segunda fase (datos de campo)

$\hat{Y}(x_1)$: valores estimados en los transectos mediante GAM con la muestra n_1

$\hat{Y}(x_2)$: valores estimados con la muestra n_2

$Y(x_2)$: valores observados en campo (n_2)

El estimador de varianza, de acuerdo con Mandallaz et al. (2013) y con Næsset et al. (2013), se calculó con la ecuación 2; y los residuales, dentro de esta ecuación, se calcularon con la ecuación 3.

$$Var(\bar{Y}_{Reg}) = Var(Y(x_2)) + \left(1 - \frac{n_2}{n_1}\right) \frac{1}{n_2} Var(\hat{R}(x_2)) \quad (2)$$

$$\hat{R} = Y(x_2) - \hat{Y}(x_2) \quad (3)$$

donde:

$Var(\bar{Y}_{Reg})$: varianza de la media estimada

\hat{R} : residuales

n_1 y n_2 se definieron anteriormente

La varianza de los valores observados en n_2 se calculó con la ecuación 4; la varianza de los residuales estimados en la muestra n_2 se calculó con la ecuación 5 y el promedio de los residuales con la ecuación 6.

$$Var(Y(x_2)) = \frac{\sum_{x \in n_2} (Y(x_2) - \bar{Y}(x_2))^2}{n_2(n_2 - 1)} \quad (4)$$

$$Var(\hat{R}(x_2)) = \frac{\sum_{x \in n_2} (\hat{R}(x_2) - \bar{R}(x_2))^2}{n_2(n_2 - 1)} \quad (5)$$

$$\bar{R}(x_2) = \frac{1}{n_2} \sum_{x \in n_2} Y(x_2) - \hat{Y}(x_2) \quad (6)$$

Estimador en dos etapas

El diseño en dos etapas parte del supuesto de una población U que está dividida en M conglomerados $\{U_1, \dots, U_M\}$ no traslapados del mismo tamaño. En la primera etapa de muestreo se elige una muestra de m conglomerados (UMP) que corresponden a un total M . La segunda etapa consiste en seleccionar una muestra de tamaño n_i (UMS) dentro de cada UMP seleccionada en la primera etapa (Mandallaz, 2008; Næsset et al., 2013). En el estudio se utilizaron los transectos LIDAR como UMP, donde el total de las celdas de 400 m² (UMS) se considera como de tamaño N_i (Gregoire et al., 2011).

El estimador de la media en dos etapas se calcula con la ecuación 7 (Næsset et al., 2013).

$$\bar{Y}_E \approx \frac{M}{m} \sum_{U_i \in S_1} \frac{N_i}{N} \frac{1}{n_i} \sum_{Y_i \in S_2} \hat{Y}(x_1) \quad (7)$$

donde:

\bar{Y}_E : media estimada en dos etapas

M : número total de conglomerados

m : muestra de conglomerados en la primera etapa

N : total de UMS

n : muestra de UMS seleccionada en la segunda etapa

$\hat{Y}(x_i)$: valores estimados de AB, V y B en los transectos (UMP)

El estimador de varianza total (ecuación 8; Gregoire et al., 2011) es la suma de la variación entre todas las muestras posibles de UMP para el diseño empleado en el muestreo

(ecuación 9) más la variación entre las muestras UMP dentro de cada UMP (ecuación 10).

$$Var(\hat{Y}_{uh}) \approx Var_I(\hat{Y}_{uh}) + Var_{II}(\hat{Y}_{uh}) \quad (8)$$

$$Var_I(\hat{Y}_{uh}) = M^2 \left(\frac{1}{m} - \frac{1}{M} \right) \sigma^2_{uh} \quad (9)$$

$$Var_{II}(\hat{Y}_{uh}) = \frac{M}{m} \sum_{p=1}^{M_u} N^2 \left(\frac{1}{n_p} - \frac{1}{N_p} \right) \sigma^2_{\epsilon_{uh}} \quad (10)$$

RESULTADOS Y DISCUSIÓN

Los valores promedio de AB, V y B (ha⁻¹) más altos, estimados con datos de campo, se obtuvieron en el diseño de muestreo dirección Este-Oeste seguido del diseño estratificado (Tabla 1). Del total de métricas LIDAR calculadas, la “altura media del dosel” (Elev_mean) y la “proporción de todos los retornos por encima de la media entre el total de primeros retornos” (Allretabmean_Tot1stret100) mostraron una correlación alta con el AB, V y B; alcanzando coeficientes de correlación de hasta 0.92 (Tabla 1). Estas dos métricas se utilizaron para predecir el AB, V y B en los transectos con datos LIDAR adicionando

como variables las coordenadas de longitud (x) y latitud (y), las cuales contribuyeron a los modelos GAM con un efecto espacial de suavizado bivariado.

La evaluación de los modelos GAM mediante validación cruzada (LOOCV) demostró que la reducción de la densidad de retornos a escala de sitio afectó la bondad de ajuste calculada, aumentando los valores de RMSE y disminuyendo el R² en los tres diseños de muestreo implementados en densidades menores o iguales a 1 retorno por metro cuadrado. Además, los promedios de R² calculados para los modelos de V, B y AB en las diferentes densidades de retornos oscilaron entre 0.86 y 0.90, 0.82 y 0.92, 0.76 y 0.83, respectivamente. El diseño estratificado (DE) presentó los valores promedio más bajos de RMSE (volumen = 37.01 m³ ha⁻¹, biomasa = 18.26 Mg ha⁻¹ y área basal = 4.97 m² ha⁻¹), seguido del diseño por transectos este-oeste (DH). Estos resultados demuestran que, si se utiliza el diseño estratificado, se captura la mayor variabilidad del atributo de interés. Por otro lado, de acuerdo con Wulder et al. (2012), los transectos LIDAR obtenidos de forma sistemática, operativamente resultan menos costosos en áreas grandes, sin embargo, se debe

TABLA 1. Valores promedio por hectárea de área basal, volumen total y biomasa aérea estimados a partir de datos de campo (n = 96) y su coeficiente de correlación (r) respecto a las métricas LIDAR relevantes.

Diseño	Variable	Mínimo	Máximo	Media	r	
					Elev_mean	Allretabmean_Tot1stret100
Horizontal	AB (m ² ha ⁻¹)	0.15	47.92	22.02	0.76	0.86
	V (m ³ ha ⁻¹)	0.31	393.05	167.16	0.85	0.84
	B (Mg ha ⁻¹)	0.16	281.77	90.82	0.81	0.81
Vertical	AB (m ² ha ⁻¹)	0.15	47.92	18.93	0.84	0.90
	V (m ³ ha ⁻¹)	0.31	438.80	143.92	0.88	0.91
	B (Mg ha ⁻¹)	0.16	281.77	76.00	0.86	0.86
Estratificado	AB (m ² ha ⁻¹)	0.27	35.84	19.66	0.87	0.90
	V (m ³ ha ⁻¹)	0.56	362.51	159.19	0.92	0.89
	B (Mg ha ⁻¹)	0.27	200.02	83.51	0.91	0.90

AB : área basal, V : volumen, B : biomasa aérea. Elev_mean es la altura media, Allretabmean_Tot1stret100 es proporción de todos los retornos por encima de la media entre el total de primeros retornos.



identificar el gradiente que logre capturar la mayor variación de la masa forestal dentro de cada línea de vuelo. Lo anterior puede explicar el hecho de que la variabilidad en los valores de RMSE y R^2 se mantuvo constantes y simétricos en estos diseños aun en densidades bajas de retornos por metro cuadrado (Fig. 3). Los resultados obtenidos concuerdan con lo señalado por Hansen Hofstad et al. (2015), quienes indicaron que las métricas derivadas de datos LIDAR aerotransportado son confiables cuando se construyen a partir de densidades mayores a 0.5 retornos por metro cuadrado. En el mismo sentido, Treitz et al. (2012) utilizaron una densidad de 0.5 retornos por metro cuadrado para el ajuste de modelos lineales múltiples y demostraron que fue suficiente para la estimación del inventario forestal a escala de sitio en Ontario, Canadá. Conviene mencionar que el ajuste de los modelos puede aumentar si se normaliza la nube de puntos (retornos) respecto al suelo a partir de un MDE de buena calidad; este proceso fue evaluado por Silva et al. (2017), quienes encontraron que las estimaciones de biomasa a escala de paisaje se subestimaron fuertemente a densidades por abajo de 0.8 retornos por metro cuadrado, principalmente en áreas con pendientes pronunciadas donde un MDE de baja calidad es más sensible en la normalización de las alturas. Al respecto, Hansen Hofstad et al. (2015) observaron que a partir de una densidad mayor a 1 retorno por metro cuadrado los efectos sobre la generación del MDE no fueron significativos. Este estudio demostró que una densidad promedio de 1 retorno por metro cuadrado puede considerarse como valor óptimo para escaneos LIDAR posteriores en el SMIC.

Las estimaciones de AB y V para los transectos del diseño DH fueron mapeados a nivel pixel con fines comparativos, mediante la aplicación de modelos GAM ajustados utilizando tres densidades de retornos: 0.8 m^{-2} , 1 m^{-2} y 10 m^{-2} (Fig. 4). Se ejemplifica este diseño por considerarlo una manera de uso operativo actual de un sensor LIDAR, además de que presentó bondades de ajuste adecuadas. Las estimaciones a partir de 0.8 m^{-2} y 1 m^{-2} retornos (densidades bajas) se contrastaron con las estimaciones a partir de 10 m^{-2} (densidad alta) mediante la

prueba estadística no paramétrica Mann-Whitney-Wilcoxon (Fig. 4), encontrándose que las estimaciones de AB fueron diferentes ($P < 0.5$), sin embargo, las estimaciones de V no fueron significativamente diferentes en términos de distribución y media ($P > 0.5$). Además, se encontró que las estimaciones derivadas de la densidad de 1 retorno por metro cuadrado fueron más conservadoras en comparación con las de 10 retornos por metro cuadrado.

Las estimaciones de AB, V y B generadas para el área de los transectos a partir de un modelo GAM y una densidad de 1 retorno por metro cuadrado (considerado el más adecuado), sirvieron como variables de respuesta para ajustar otros modelos GAM y RF para estimar y mapear las mismas variables para todo el SMIC. Se utilizaron como predictores los valores de las bandas espectrales, el NDVI y métricas de textura derivadas del sensor *RapidEye* que registraron un coeficiente de correlación de Pearson alto con las variables de interés (AB, V y B): texturas construidas a partir de la banda roja (B3) (media y varianza), azul (media) (B1_mean), verde (media) (B2_mean) y el NDVI (media y varianza) (ventana móvil de 15×15) (Tabla 2). El modelo GAM presentó menor valor del porcentaje de varianza explicada (R^2) y mayor RMSE en la estimación de AB, V y B, comparado con RF (Tabla 3). Sin embargo, todos los parámetros de suavizado que corresponden a las texturas del RGB y NDVI fueron altamente significativos ($P < 0.0001$).

El modelo RF resultante incluyó cinco bandas del sensor *Rapideye* y 42 texturas, además de las coordenadas longitud (x) y latitud (y). Este algoritmo logró explicar las variables dasométricas de interés por arriba de 70%, teniendo mejor rendimiento en la estimación de la B (79%, $RMSE = 23.67 \text{ Mg ha}^{-1}$) seguido del AB (73%, $RMSE = 5.69 \text{ m}^2 \text{ ha}^{-1}$) y V (71%, $RMSE = 46.95 \text{ m}^3 \text{ ha}^{-1}$).

RF mostró mejores bondades de ajuste al utilizar texturas derivadas del sensor *RapidEye*, lo que coincide con lo observado en otros estudios (Matasci et al., 2018; Zhu et al., 2020). Una ventaja de RF es que implementa una validación cruzada a partir de conjuntos de datos de entrenamiento y validación para evitar la inflación de la varianza causada por el uso de variables redundantes en la

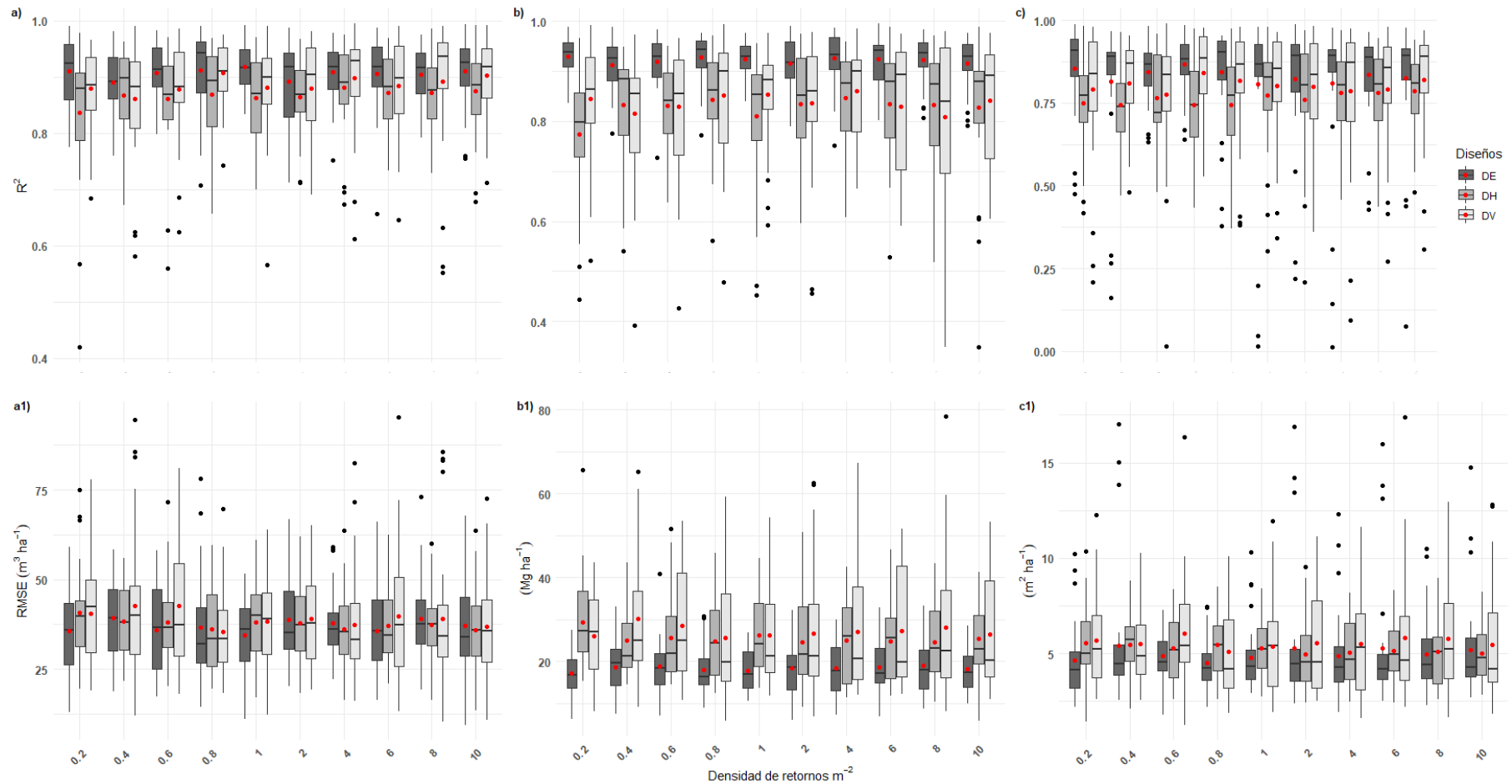


FIGURA 3. Efecto de la densidad de retornos LIDAR en la modelización de área basal (AB), volumen (V) y biomasa (B) para los sitios muestreados bajo los diferentes diseños (n=96); DH (este-oeste), DV (norte-sur); DE (estratificado).

R^2 y RMSE calculados mediante la validación cruzada repetida para; V (a-a1); B (b- b1) y AB (c-c1); los puntos negros representan puntos aberrantes.

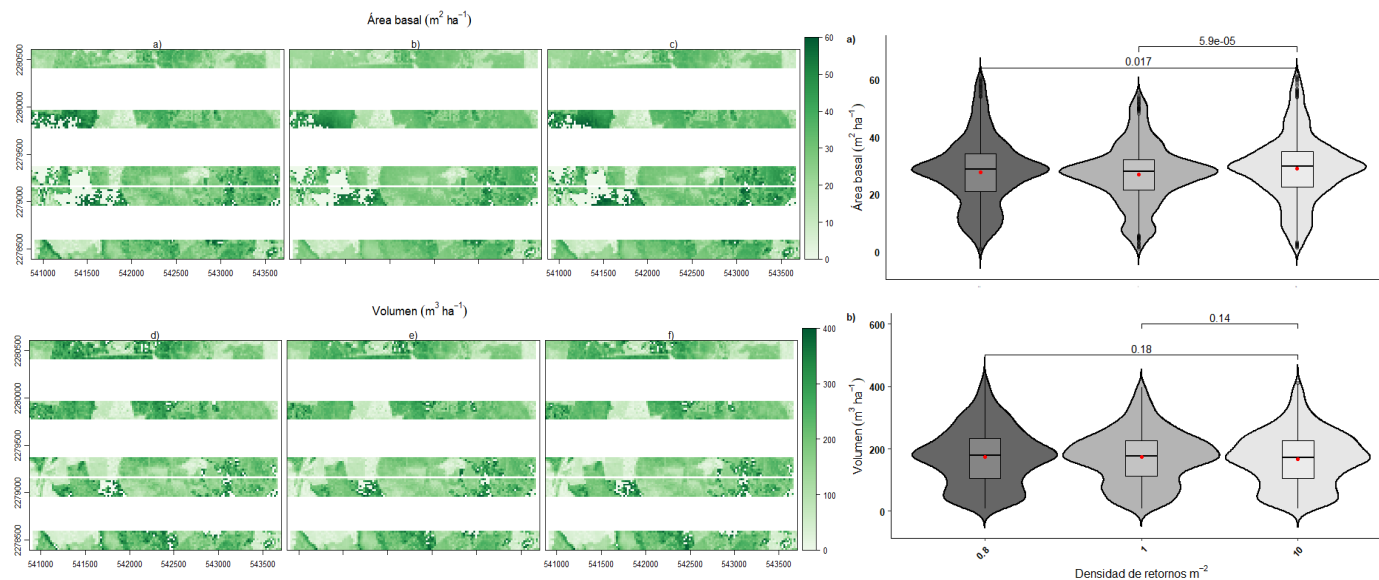


FIGURA 4. Estimación de área basal y volumen en transectos este-oeste a partir de variables LIDAR calculados a densidades de retornos diferentes (izquierda). a y d) 0.8 pulsos por metro cuadrado; b y e) 1 pulso por metro cuadrado; c y f) 10 pulsos por metro cuadrado. Prueba Mann-Whitney-Wilcoxon (0.6-10 y 1-10) (derecha).

TABLA 2. Variables (métricas) espectrales y de textura más correlacionadas (coeficiente de correlación -r) con las variables dasométricas.

RapidEye	Variables dasométricas		
	Área basal (m ²)	Volumen (m ³)	Biomasa (Mg)
B3_mean	-0.73	-0.57	-0.73
B3_var			-0.72
B1_mean	-0.72	-0.55	-0.72
B2_mean	-0.70	-0.53	-0.69
B3	-0.69	-0.54	-0.69
NDVI_var	0.63	0.48	0.65
NDVI_mean	0.63	0.48	0.65

B1: banda azul, B2: verde, B3: rojo, B4: borde rojo, B5: infrarrojo, texturas (15 × 15) _mean y _var: media y varianza.

TABLA 3. Valores de bondad de ajuste de los modelos aditivos generalizados y el algoritmo *random forest* evaluados para estimar área basal, volumen total y biomasa aérea en el Sitio de Monitoreo Intensivo de Carbono.

Modelo	Estructura del modelo	R ²	RMSE
GAM	AB=β ₀ + s(x,y)+s(B1_mean)+s(B2_mean)+s(B3_mean)+s(B3)+s(NDVI_mean)+s(NDVI_var)	0.61	6.85
RF	AB= s(x,y)+todas las métricas <i>RapidEye</i>	0.73	5.69
GAM	V= β ₀ +s(x,y)+s(B1_mean)+s(B2_mean)+s(B3_mean)+s(B3)+s(NDVI_mean)+s(NDVI_var)	0.61	54.15
RF	V= s(x,y)+todas las métricas <i>RapidEye</i>	0.71	46.95
GAM	B= β ₀ +s(x,y)+s(B1_mean)+s(B2_mean)+s(B3_mean)+s(B3)+s(NDVI_mean)+s(NDVI_var)	0.67	29.24
RF	B= s(x,y)+todas las métricas <i>RapidEye</i>	0.79	23.67

A: área basal (m²), V: volumen (m³), B: biomasa aérea (Mg). B1: banda azul, B2: verde, B3: rojo, B4: borde rojo, B5: infrarrojo, NDVI: índice de vegetación normalizado, texturas (15 × 15) _mean y _var: media y varianza, β₀ parámetro.

explicación de la variable de interés (Fassnacht et al., 2014). Además, se ha observado una mejoría en las estimaciones con el uso de texturas usando el método GLCM, con un tamaño de ventana de 15×15 (Zhu et al., 2020) o mayor dependiendo la variabilidad del dosel. A diferencia de las métricas comunes, las métricas de textura espectral pueden reducir los impactos de sombra y extraer más eficazmente la información de bosques maduros, con ello pueden aumentar las correlaciones entre la textura y la masa forestal madura (Lu y Batistella, 2005), como en el caso del presente estudio.

Los mapas de AB y el V correspondientes al área boscosa del SMIC (815.03 ha) muestran que ambos modelos (GAM y RF) estiman con tendencia similar (Fig. 5). Los datos indican que el modelo GAM estimó valores totales de AB y V de menor magnitud que RF (Tabla 4), pero valores absolutos más altos en ambas variables (Fig. 6). Conforme a lo esperado, los valores más bajos de AB y V corresponden a rodales jóvenes, cosechados en 2003-2013 y los valores máximos corresponden a los rodales maduros y a áreas sin intervención silvícola (Tabla 5).

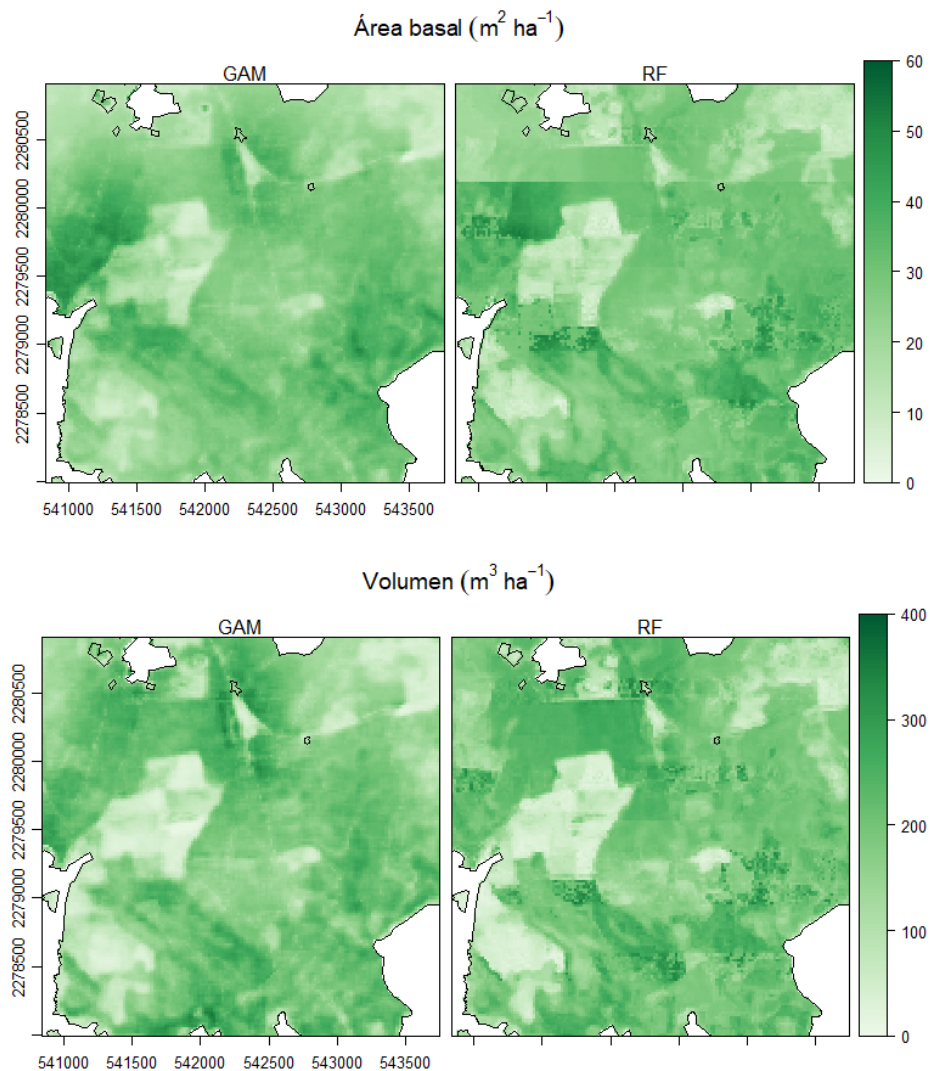


FIGURA 5. Variabilidad espacial del área basal y volumen de madera estimada a partir de modelos aditivos generalizados y el algoritmo *random forest* utilizando métricas derivadas del sensor *RapidEye*.



TABLA 4. Resumen de los valores de área basal y volumen total obtenidos mediante modelos aditivos generalizados y el algoritmo *random forest* en el Sitio de Monitoreo Intensivo de Carbono.

	Variable	N	Min	Max	Riq	Media	Sd	Inventario
GAM	AB (m ² ha ⁻¹)	20306	3.22	50.7	10.8	26.2	8.38	21 353.79
	Volumen (m ³ ha ⁻¹)	20306	7.39	468	86.3	173	64.9	141 000.19
RF	AB (m ² ha ⁻¹)	20306	3.2	53.1	10	27	8.4	22005.81
	Volumen (m ³ ha ⁻¹)	20306	13.3	354	76.9	178	64.8	145075.34

AB: área basal, V: volumen total, N : total de sitios de 400 m2, Min: valor mínimo estimado, Max: valor máximo estimado, Riq: rango intercuantil, Sd: desviación estándar.

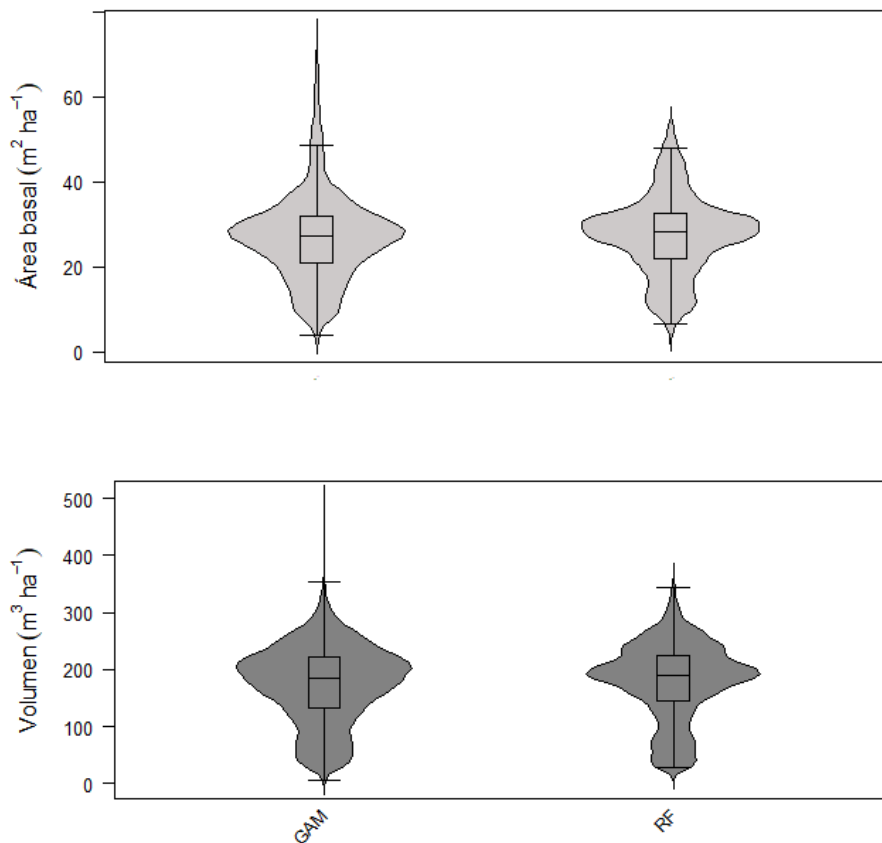


FIGURA 6. Diagrama de caja de las estimaciones de volumen de madera generadas a partir de modelos aditivos generalizados y el algoritmo *random forest* ajustados utilizando variables del sensor *RapidEye*.

La muestra de datos LIDAR (transectos) que representó 30% del total de la superficie del área de estudio, fue suficiente para estimar con precisión el AB, V y B, y para expandir las estimaciones a toda el área mediante la incorporación de datos espectrales *RapidEye*. Lo anterior

sugiere que la estimación del tamaño óptimo de la muestra requiere de estudios más detallados al respecto. Chen y Hay (2011) observaron que una muestra de 8.8% (un solo transecto) de la superficie total aumentó el rendimiento (Volumen: R²= 0.64 y RMSE= 156.9 m³ ha⁻¹) del algoritmo

de aprendizaje utilizado [regresión de soporte de vectores (SVR Support Vector Regression)] para predecir variables de interés. Estos autores diseñaron cuatro muestras de transectos LIDAR (8.8%, 17.6%, 26.4% y 35.2% del área total) y las combinaron con imágenes QuickBird para estimar la altura del dosel, biomasa y volumen en Vancouver Island, Canadá. En otro estudio, Chen et al.

(2012) evaluaron la proporción y dirección (Norte-Sur y Este-Oeste) de transectos LIDAR para estimar altura, biomasa y volumen en Quebec, Canadá y encontraron que un solo transecto que representó 7.6% del área total presentó mejores bondades de ajuste (volumen: $R^2 = 0.72$; $RMSE = 52.59 \text{ m}^3 \text{ ha}^{-1}$) en el algoritmo SVR.

TABLA 5. Estimador en dos etapas para la media e inventario total de área basal, volumen total y biomasa aérea obtenidos mediante los transectos estimados por modelos aditivos generalizados.

	<i>Variable</i>	<i>Estrato</i>	<i>Media</i>	<i>SE</i>	<i>IC-</i>	<i>IC+</i>	<i>Error %</i>	<i>Inventario</i>
Dos etapas	AB ($\text{m}^2 \text{ ha}^{-1}$)	Todos	25.93	0.75	23.84	28.01	5.80%	21 131.16
		1933	29.42	0.49	28.46	30.38	3%	8 086.09
		1982-1987	26.78	4.67	17.63	35.94	35%	2 550.64
		1988-1992	30.36	1.39	27.63	33.09	9%	1 654.68
		1993-1997	26.00	0.74	24.56	27.44	6%	2 605.09
		1998-2002	28.13	1.27	25.64	30.62	9%	4 051.57
		2003-2007	18.80	2.86	13.19	24.41	30%	1 285.31
		2008-2013	11.53	1.67	8.27	14.80	29%	897.79
	V ($\text{m}^3 \text{ ha}^{-1}$)	Todos	174.33	3.74	163.95	184.71	4.29%	142 085.96
		1933	213.75	3.10	207.68	219.82	3%	58 753.79
		1982-1987	210.37	10.53	189.73	231.02	10%	20 034.79
		1988-1992	228.85	7.08	214.97	242.73	6%	12 473.40
		1993-1997	170.86	5.10	160.87	180.85	6%	17 119.52
		1998-2002	165.70	10.37	145.37	186.03	13%	23 865.76
		2003-2007	80.82	8.80	63.58	98.07	22%	5 525.10
		2008-2013	55.42	7.50	40.73	70.11	27%	4 313.62
	B (Mg ha^{-1})	Todos	94.92	1.45	90.89	98.96	3.1%	77 366.12
		1933	119.47	1.79	115.95	122.98	3%	32 838.37
		1982-1987	112.60	5.38	102.05	123.15	10%	10 723.05
		1988-1992	126.48	1.58	123.39	129.57	2%	6 893.84
		1993-1997	92.26	5.05	82.37	102.16	11%	9 244.36
1998-2002		88.79	6.00	77.03	100.55	14%	12 788.16	
2003-2007		39.55	4.71	30.33	48.78	24%	2 703.99	
2008-2013		27.94	4.25	19.61	36.26	30%	2 174.35	

AB: área basal, V: volumen total, B: biomasa aérea, IC: intervalo de confianza a 95%.



Estimadores de muestreo e inventario total

Estimador en dos fases

En la primera fase se seleccionó una muestra de 6457 píxeles (400 m²), dentro del diseño DH, que fue utilizada como variables auxiliares (LIDAR). La segunda fase incorporó los datos de campo derivados de la muestra de 96 sitios medidos. El estimador de la media en dos fases asistido por el modelo GAM mostró un error menor a 5% (Tabla 6). Estos resultados corroboran la utilidad de este tipo de modelos señalada por otros autores. De acuerdo con Massey (2015) el proceso de estimación a partir de información auxiliar (LIDAR) y datos de inventario en la segunda fase ofrece una reducción del sesgo debido a que el estimador presenta una muestra grande en la primera fase, en lugar del uso de información de pared a pared (área SMIC). Estudios previos como el de Kangas et al. (2016) han utilizado el muestreo en dos fases asistido por un modelo GAM, demostrando cómo las técnicas de estimación no paramétricas incorporan resultados prometedores incluso en el caso de los modelos complejos (GAM) y diseños multifásicos.

Estimador en dos etapas

En este método los transectos LIDAR que contienen las estimaciones realizadas con el modelo GAM fueron consideradas como UMP y cada una de las celdas de 400 m² como UMS. Para cada UMP se asumió a las clases de edad como estratos para estimar una media general y por estrato con base en la ponderación (Gregoire et al., 2011).

Los resultados indican que el error estimado es bajo (4.29% para el volumen) y aumenta en los rodales más jóvenes o recién aprovechados (Tabla 5); esto se debe a que dentro de las UMP se encuentran menor número de celdas clasificadas para esas clases (Gregoire et al., 2011). Análisis previos para la misma área mencionan un error mayor para las áreas sin intervención silvícola argumentando que se establecieron pocos sitios en campo en ese estrato, lo que implica una mayor variación en la estimación de la media (Ortiz-Reyes et al., 2015; Torres-Vivar et al., 2017).

CONCLUSIONES

Los modelos GAM ajustados bajo tres diseños de muestreo para inventario forestal asistido por datos de un sensor LIDAR aerotransportado permitió identificar al diseño estratificado (DE) como el mejor (RMSE: V = 37.01 m³ ha⁻¹, B = 18.26 Mg ha⁻¹ y AB = 4.97 m² ha⁻¹), aun en densidades de retornos bajas. Aunque este diseño de muestreo no se utilizó para la estimación del inventario total del SMIC, los resultados alientan la posibilidad de su aplicación en inventarios forestales operativos mediante la captura de datos LIDAR a través de sensores montados en drones.

Se encontró que una densidad de datos LIDAR de 1 retorno por metro cuadrado fue suficiente para obtener estimaciones precisas de AB, V y B en bosques templados bajo manejo coetáneo, lo que permitirá considerar en futuros planes de vuelo para colecta de datos LIDAR una mayor altura de vuelo, con la consecuente reducción de tiempo y costo de escaneo.

TABLA 6. Estimador por regresión en dos fases para la media e inventario total de área basal, volumen total y biomasa aérea obtenidos mediante los transectos estimados por modelos aditivos generalizados.

Estimador	Variable	n ₁ :n ₂	Media	Se ext	IC-	IC+	Error %	Inventario
Dos fases	AB (m ² ha ⁻¹)		26.61	0.45	25.73	27.50	3.33%	21 689.52
	V (m ³ ha ⁻¹)	6 457:96	164.82	4.21	156.47	173.17	5.07%	134 333.05
	B (Mg ha ⁻¹)		91.12	2.15	86.84	95.40	4.69%	74 268.59

AB: área basal, V: volumen total, B: biomasa aérea, n₁ y n₂: tamaño de muestra en la primera y segunda fase, IC: intervalo de confianza a 95%.

Los modelos de regresión GAM y al algoritmo *Random Forest* demostraron propiedades importantes en la estimación de las variables de interés forestal tanto a escala de transectos como en áreas donde la información LIDAR no está presente, pero se cuenta con otro tipo de imágenes ópticas de resolución espacial fina.

Los estimadores de muestreo en dos fases y en dos etapas demostraron ser estimadores compatibles con las características de vuelo LIDAR mediante transectos, permitiendo alcanzar precisiones de muestreo menores o iguales a 6%.

RECONOCIMIENTOS

Los autores agradecen el financiamiento del Conacyt a través del proyecto "El papel de los bosques bajo gestión forestal comunitaria en la mitigación del cambio climático".

REFERENCIAS

- Almeida, D. R., Stark, S. C., Shao, G., Schiatti, J., Nelson, B. W., Silva, C. A., Gorgens, E. B., Valbuena, R., Papa, D. A., & Santin, P. H. B. (2019). Optimizing the remote detection of tropical rainforest structure with airborne lidar: Leaf area profile sensitivity to pulse density and spatial sampling. *Remote Sensing*, *11*(1), 92. <https://doi.org/10.3390/rs11010092>
- Chen, G., & Hay, G. (2011). A Support Vector Regression Approach to Estimate Forest Biophysical Parameters at the Object Level Using Airborne Lidar Transects and QuickBird Data. *American Society for Photogrammetry and Remote Sensing*, *77*(7), 733–741. <https://doi.org/10.14358/PERS.77.7.733>
- Chen, G., Hay, G., & St-Onge, B. (2012). A GEOBIA framework to estimate forest parameters from lidar transects, Quickbird imagery and machine learning: A case study in Quebec, Canada. *International Journal of Applied Earth Observation and Geoinformation*, *15*(1), 28–37. <https://doi.org/10.1016/j.jag.2011.05.010>
- Comisión Nacional Forestal [Conafor]. (2011). *Manual y procedimientos para el muestreo de campo*. Zapopan, Jalisco, México. Recuperado el 31 de 03 de 2020, de https://www.climateactionreserve.org/wp-content/uploads/2011/03/Sampling_Manual_Remuestreo-Conafor_INFyS.pdf
- Crespo-Peremarch, P., Ruiz, L., Balaguer-Beser, Á., & Estornell, J. (2018). Analyzing the role of pulse density and voxelization parameters on fullwaveform LiDAR-derived metrics. *ISPRS Journal of Photogrammetry and Remote Sensing*, *146*(1), 453–464. <https://doi.org/10.1016/j.isprsjprs.2018.10.012>
- Cruz-Martínez, Z. (2007). Sistema de Ecuaciones para estimación y partición de biomasa aérea en Atopixco, Zacualtipán, Hidalgo, México. *Tesis de Maestría*. Chapingo, Texcoco, Mexico.
- Fassnacht, F., Hartig, F., Latifi, h., Berger, C., Hernández, J., Corvalán, P., & Koch, B. (2014). Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sensing of Environment*, *154*(1)102-114. <https://doi.org/10.1016/j.rse.2014.07.028>
- George-Chacon, S., Dupuy, J. M., Peduzzi, A., & Hernandez-Stefanoni, J. (2019). Combining high resolution satellite imagery and lidar data to model woody species diversity of tropical dry forests. *Ecological Indicators*, *101*(1), 975-984. <https://doi.org/10.1016/j.ecolind.2019.02.015>
- Grafström, A., & Ringvall, H. (2013). Improving forest field inventories by using remote sensing data in novel sampling designs. *Canadian Journal of Forest Research*, *43*(11), 1015-1022. <https://doi.org/10.1139/cjfr-2013-0123>
- Gregoire, T. G., Ståhl, G., Næsset, E., Gobakken, T., Nelson, R., & Holm, S. (2011). Model-assisted estimation of biomass in a LiDAR sample survey in Hedmark County, Norway. *Canadian Journal of Forest Research*, *41*(1), 83–95. <https://doi.org/10.1139/X10-195>
- Hansen Hofstad, E., Gobakken, T., & Næsset, E. (2015). Effects of Pulse Density on Digital Terrain Models and Canopy Metrics Using Airborne Laser Scanning in a Tropical Rainforest. *Remote Sensing*, *7*(7), 8453-8468. <https://doi.org/10.3390/rs70708453>
- Helmer, E. H., & Ruefenacht, B. (2005). Cloud-free satellite image mosaics with regression trees and histogram matching. *Photogrammetric Engineering & Remote Sensing*, *71*(9), 1079-1089. <https://doi.org/10.14358/PERS.71.9.1079>
- Hernández-Stefanoni, J., Dupuy, J. M., Johnson, K. D., Birdsey, R., Tun-Dzul, F., Peduzzi, A., Caamal-Sosa, J. P., Sánchez-Santos, G., & López-Merlín, D. (2014). Improving Species Diversity and Biomass Estimates of Tropical Dry Forests Using Airborne LiDAR. *Remote Sensing*, *6*(6), 4741-4763. <https://doi.org/10.3390/rs6064741>
- Hong, G., & Zhang, Y. (2008). A comparative study on radiometric normalization using high resolution satellite images. *International Journal of Remote Sensing*, *29*(2), 425-438. <https://doi.org/10.1080/01431160601086019>
- Instituto Nacional de Estadística y Geografía [Inegi]. (2017). *Anuario estadístico y geográfico de Hidalgo 2017*. Instituto Nacional de Estadística: Inegi, México. Recuperado el 31 de 03 de 2020, de https://www.datatur.sectur.gob.mx/ITxEF_Docs/HGO_ANUARIO_PDF.pdf



- Roussel, J.-R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R. H., Sánchez Meador, A., Bourdon, J.-F., Boissieu, F., & Achim, A. (2020). lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. *Remote Sensing of Environment*, 251(1), 112061. <https://doi.org/10.1016/j.rse.2020.112061>
- Kim, H.-O., & Yeom, J.-M. (2015). Sensitivity of vegetation indices to spatial degradation of RapidEye imagery for paddy rice detection: a case study of South Korea. *GIScience & Remote Sensing*, 52(1), 1-17. <https://doi.org/10.1080/15481603.2014.1001666>
- Kangas, A., Myllymäki, M., Gobakken, T., & Næsset, E. (2016). Model-assisted forest inventory with parametric, semiparametric, and nonparametric models. *Canadian Journal of Forest Research*, 46(6), 855-868. <https://doi.org/10.1139/cjfr-2015-0504>
- Leutner, B., Horning, N., Schwalb-Willmann, J., & Hijmans, R. (2019). Package 'RStoolbox'. Tools for Remote Sensing Data Analysis. R Foundation for Statistical Computing, Version 0.1.
- Lu, D., & Batistella, M. (2005). Exploring TM Image Texture and its Relationships with Biomass Estimation in Rondônia, Brazilian Amazon. *Acta Amazonica*, 35(2), 249-257.
- Mahoney, C., Hall, J.R., Hopkinson, C., Filiatrault, M., Beaudoin, A., & Chen, Q. (2018). A Forest Attribute Mapping Framework: A Pilot Study in a Northern Boreal Forest, Northwest Territories, Canada. *Remote Sensing*, 10(1338), 1-30. <https://doi.org/10.3390/rs10091338>
- Mandallaz, D. (2008). *Sampling techniques for forest inventories*. Chapman & Hall/CRC.
- Mandallaz, D., Breschan, J. I., & Hil, A. (2013). New regression estimators in forest inventories with two-phase sampling and partially exhaustive information: a design-based Monte Carlo approach with applications to small-area estimation. *Canadian Journal Forest Research*, 43(11), 1023-1031. <https://doi.org/10.1139/cjfr-2013-0181>
- Marchi, N., Pirotti, F., & Lingua, E. (2018). Airborne and Terrestrial Laser Scanning Data for the Assessment of Standing and Lying Deadwood: Current Situation and New Perspectives. *Remote Sensing*, 10(1356), 1-21. <https://doi.org/10.3390/rs10091356>
- Massey, A. F. (2015). Multiphase estimation procedures for forest inventories under the design-based Monte Carlo approach. *Tesis Doctoral*. ETH-Zürich.
- Matasci, G., Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., Hobart, G. W., & Zald, H. S. (2018). Large-area mapping of Canadian boreal forest cover, height, biomass and other structural attributes using Landsat composites and lidar plots. *Remote Sensing of Environment*, 209(1), 90-106. <https://doi.org/10.1016/j.rse.2017.12.020>
- McGaughey, R. J. (11 de 06 de 2018). *FUSION / LDV: Software para análisis y visualización de datos LIDAR 3.80*. Obtenido de http://forsys.cfr.washington.edu/fusion/fusion_overview.html
- McRoberts, R., Magnussen, S., Tomppo, E., & Chirici, G. (2011). Parametric, bootstrap, and jackknife variance estimators for the k-Nearest Neighbors technique with illustrations using forest inventory and satellite image data. *Remote Sensing of Environment*, 115(12), 3165-3174. <https://doi.org/10.1016/j.rse.2011.07.002>
- Næsset, E., Gobakken, T., Bollandsås, O., Gregoire, T., Nelson, R., & Ståhl, G. (2013). Comparison of precision of biomass estimates in regional field sample surveys and airborne LiDAR-assisted surveys in Hedmark County, Norway. *Remote Sensing of Environment*, 130(1), 108-120. <https://doi.org/10.1016/j.rse.2012.11.010>
- Nelson, R., Margolis, H., Montesano, P., Sun, G., Cook, B., Corp, L., Andersen, H.-E., Jong, B. de, Paz Pellat, F., Fickel, T. Kauffman, J., & 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(1), 127-140. <https://doi.org/10.1016/j.rse.2016.10.038>
- Ochoa-Franco, A., Valdez-Lazalde, J. R., Ángeles-Pérez, G., de los Santos-Posadas, H. M., Hernández-Stefanoni, J., Valdez-Hernández, J. I., & Pérez-Rodríguez, P. (2019). Beta-Diversity Modeling and Mapping with LiDAR and Multispectral Sensors in a Semi-Evergreen Tropical Forest. *Forests*, 10(5), 419. <https://doi.org/10.3390/f10050419>
- Ortiz-Reyes, A. D., Valdez-Lazalde, J. R., Ángeles-Pérez, G., Santos-Posadas, H. M. de los, 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), 1-18. <https://doi.org/10.21829/myb.2019.2531872>
- Ortiz-Reyes, A. D., Valdez-Lazalde, J. R., De los Santos-Posadas, H. M., Ángeles-Pérez, G., Paz-Pellat, F., & 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. <https://doi.org/10.21829/myb.2015.213461>
- Planet Labs Inc. (2016). RapidEye Satellite imagery product specifications. Satellite imagery product specifications: Version 6.1.
- Saarela, S., Schnell, S., Grafström, A., Tuominen, S., Nordkvist, K., Hyyppä, J., Kangas, A., & Ståhl, G. (2015). Effects of sample size and model form on the accuracy of model-based estimators of growing stock volume. *Canadian Journal of Forest Research*, 45(11), 1524-1534. <https://doi.org/10.1139/cjfr-2015-0077>
- Silva, C. A., Hudak, A. T., Vierling, L. A., Klauber, C., Garcia, M., Ferraz, A., Keller, M., Ektel, J., & Saatchi, S. (2017). Impacts of

- Airborne Lidar Pulse Density on Estimating Biomass Stocks and Changes in a Selectively Logged Tropical Forest. *Remote Sensing*, 9(1068), 1-19. <https://doi.org/10.3390/rs9101068>
- Singh, K., Chen, G., McCarter, J., & Meentemeyer, R. (2015). Effects of LiDAR point density and landscape context on estimates of urban forest biomass. *ISPRS Journal of Photogrammetry and Remote Sensing*, 101(1), 310–322. <https://doi.org/10.1016/j.isprsjprs.2014.12.021>
- Soriano-Luna, M. Á. (2014). Estimación de biomasa y carbono en bosques manejados de Zacualtipán, Hidalgo. *Tesis de Maestría*. Colegio de Postgraduados, Montecillo, Texcoco, México.
- Soriano-Luna, M. d., Ángeles-Pérez, G., Guevara, M., Birdsey, R., Pan, Y., Vaquera-Huerta, H., Valdez-Lazalde, J. R., Johnson, K. D., & Vargas, R. (2018). Determinants of Above-Ground Biomass and Its Spatial Variability in a Temperate Forest Managed for Timber Production. *Forest*, 9(490), 1-20. <https://doi.org/10.3390/f9080490>
- Strimbu, V. F., Ene, L. T., Gobakken, T., Gregoire, T., Astrup, R., & Næsset, E. (2017). Post-stratified change estimation for large-area forest biomass using repeated ALS strip sampling. *Canadian Journal of Forest Research*, 47(6), 839–847. <https://doi.org/10.1139/cjfr-2017-0031>
- Team R Core. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Torres-Vivar, J. E., Valdez-Lazalde, J. R., Ángeles-Pérez, G., de los Santos-Posadas, H. M., & Aguirre-Salado, C. A. (2017). Inventario y mapeo de un bosque bajo manejo de pino con datos del sensor SPOT 6. *Revista Mexicana de Ciencias Forestales*, 8(39), 25-43.
- Treitz, P., Lim, K., Woods, M., Pitt, D., Nesbitt, D., & Etheridge, D. (2012). LiDAR Sampling Density for Forest Resource Inventories in Ontario, Canada. *Remote Sensing*, 4(4), 830-848. <https://doi.org/10.3390/rs4040830>
- Urbazaev, M., Thiel, C., Cremer, F., Dubayah, R., Migliavacca, M., Reichstein, M., & Schullius, 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(5), 1-20. <https://doi.org/10.1186/s13021-018-0093-5>
- Ver, N., Finley, A., Kershaw, J., & Weiskittel, A. (2018). Hierarchical Bayesian models for small area estimation of forest variables using LiDAR. *Remote Sensing of Environment*, 204, 287-295. <https://doi.org/10.1016/j.rse.2017.10.024>
- Wulder, M. A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N., Hilker, T. B., C. W., & Gobakken, T. (2012). Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment*, 121, 196–209. <https://doi.org/10.1016/j.rse.2012.02.001>
- Zhu, Y., Zhongke, F., Lu, J., & Jincheng, L. (2020). Estimation of Forest Biomass in Beijing (China) Using Multisource Remote Sensing and Forest Inventory Data. *Forests*, 11(2), 163. <https://doi.org/10.3390/f11020163>
- Zvoleff, A. (2020). *Calculate Textures from Grey-Level Co-Occurrence Matrices (GLCMs)*. Obtenido de <http://www.azvoleff.com/g lcm>.

Manuscrito recibido el 18 de febrero de 2021

Aceptado el 15 de julio de 2022

Publicado el 15 de diciembre de 2022

Este documento se debe citar como:

Galeote-Leyva, B., Valdez-Lazalde, J. R., Ángeles-Pérez, G., de los Santos-Posadas, H. M., & Romero-Padilla, J. M. (2022). Inventario forestal asistido por LIDAR: efecto de la densidad de retornos y el diseño de muestreo sobre la precisión. *Madera y Bosques*, 28(2), e2822330. <https://doi.org/10.21829/myb.2022.2822330>



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