

# Synergy of Landsat, climate and LiDAR data for aboveground biomass mapping in medium-stature tropical forests of the Yucatan Peninsula, Mexico

## Sinergia de datos espectrales Landsat, climáticos y LiDAR para el mapeo de biomasa aérea en selvas medianas de la península de Yucatán, México

Alma D. Ortiz-Reyes<sup>1</sup>; José R. Valdez-Lazalde<sup>2\*</sup>; Gregorio Ángeles-Pérez<sup>2</sup>; Héctor M. De los Santos-Posadas<sup>2</sup>; Laura Schneider<sup>3</sup>; Carlos A. Aguirre-Salado<sup>4</sup>; Alicia Peduzzi<sup>5</sup>

<sup>1</sup>Centro Nacional de Investigación Disciplinaria en Conservación y Mejoramiento de Ecosistemas Forestales (CENID-COMEF). Av. Progreso núm. 5, Barrio de Santa Catarina. C. P. 04110. Coyoacán, Ciudad de México, México.

<sup>2</sup>Colegio de Postgraduados, Postgrado en Ciencias Forestales. Carretera México-Texcoco km 36.5, Montecillo, Texcoco, México. C. P. 56230. México

<sup>3</sup>Rutgers University, Department of Geography. 54 Joyce Kilmer Avenue, Blvd. Piscataway, NJ 08854, USA.

<sup>4</sup>Universidad Autónoma de San Luis Potosí, Facultad de Ingeniería. Dr. Manuel Nava núm. 8, Zona Universitaria. C. P. 78000. San Luis Potosí, México.

<sup>5</sup>Woods Hole Research Center. 149 Woods Hole Road, 02540-1644, Falmouth, MA, USA.

\*Corresponding author: valdez@colpos.mx; tel.: +52 595 952 0200 ext. 1482.

### Abstract

**Introduction:** Tropical forests represent complex and dynamic ecosystems that cover extensive areas, hence the importance of determining biomass content and representing spatial variability.

**Objective:** Estimating and mapping aboveground biomass and its associated uncertainty for medium-stature semi-evergreen (SMSP) and semi-deciduous (SMSC) tropical forests of the Yucatan Peninsula.

**Materials and methods:** Aboveground biomass was estimated as a function of explanatory variables taken from Landsat images and climatic variables, using the random Forest algorithm. Aboveground biomass was mapped from previous biomass estimates for stripes of the territory with the presence of LiDAR (Light Detection And Ranging) and field data. Uncertainty at the pixel level was estimated as the coefficient of variation.

**Results and discussion:** A combination of climatic and spectral variables showed acceptable capacity to estimate biomass in the medium-stature semi-evergreen and semi-deciduous tropical forest with an explained variance of 50 % and RMSE (root mean squared error) of 34.2 Mg·ha<sup>-1</sup> and 26.2 Mg·ha<sup>-1</sup>, respectively, prevailing climate variables. SMSP biomass ranged from 4.0 to 185.7 Mg·ha<sup>-1</sup> and SMSC ranged from 11.7 to 117 Mg·ha<sup>-1</sup>. The lowest values of uncertainty were recorded for the medium-stature semi-evergreen tropical forest, being higher in areas with lower amounts of aboveground biomass.

**Conclusion:** Aboveground biomass was estimated and mapped by the combined use of auxiliary variables with an acceptable accuracy, against uncertainty of predictions, which represents an opportunity for future improvement.

### Resumen

**Introducción:** Los bosques tropicales constituyen ecosistemas complejos y dinámicos que cubren áreas extensas, de ahí la importancia de determinar su contenido de biomasa y representar su variabilidad espacial.

**Objetivo:** Estimar y mapear la biomasa aérea y su incertidumbre asociada en selvas medianas subperennifolia (SMSP) y subcaducifolia (SMSC) de la península de Yucatán.

**Materiales y métodos:** La biomasa aérea se estimó en función de variables explicativas obtenidas de imágenes Landsat y variables climáticas, mediante el algoritmo random Forest. La biomasa aérea se mapeó a partir de estimaciones previas de biomasa para franjas del territorio con presencia de datos LiDAR (Light Detection And Ranging) y datos de campo. La incertidumbre a nivel de pixel se estimó como el coeficiente de variación.

**Resultados y discusión:** Una combinación de variables climáticas y espectrales mostraron capacidad aceptable para estimar la biomasa en la selva mediana subperennifolia y mediana subcaducifolia con una varianza explicada de 50 % y RMSE (raíz del error cuadrático medio) de 34.2 Mg·ha<sup>-1</sup> y 26.2 Mg·ha<sup>-1</sup>, respectivamente, prevaleciendo las variables climáticas. La biomasa de la SMSP varió entre 4.0 y 185.7 Mg·ha<sup>-1</sup> y la de la SMSC osciló entre 11.7 y 117 Mg·ha<sup>-1</sup>. Los valores más bajos de incertidumbre se registraron en la selva mediana subperennifolia, siendo mayores en zonas con cantidades menores de biomasa aérea.

**Conclusión:** La biomasa aérea se estimó y mapeó mediante el uso combinado de las variables auxiliares con una precisión aceptable, contrario a la incertidumbre de las predicciones, lo que representa una oportunidad de mejora futura.

**Keywords:** tropical forest; satellite images; vegetation indices; random Forest, uncertainty.

**Palabras clave:** bosque tropical; imágenes satelitales; índices de vegetación; random Forest, incertidumbre.

Please cite this article as follows (APA 6): 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. (2021). Synergy of Landsat, climate and LiDAR data for aboveground biomass mapping in medium-stature tropical forests of the Yucatan Peninsula, Mexico. *Revista Chapingo Serie Ciencias Forestales y del Ambiente*, 27(3), 383-400. doi: 10.5154/r.rchscfa.2020.08.050

Received: August 14, 2020 / Accepted: July 1, 2021.



<https://revistas.chapingo.mx/forestales>

## Introduction

Tropical forests represent 45 % of the Earth's forest area and have high carbon storage capacity (Food and Agriculture Organization of the United Nations [FAO], 2020; FAO & UNEP, 2020). Unfortunately, in recent decades, these ecosystems have recorded high rates of deforestation and degradation (FAO, 2020) becoming sources of carbon dioxide into the atmosphere. Therefore, monitoring biomass (carbon) inventories and changes in the area of these forests is important for the planning of policies that contribute to the mitigation of negative effects of climate change (Dupuy-Rada, Hernández-Stefanoni, Hernández-Juárez, Tun-Dzul, & May-Pat, 2012; Houghton, Byers, & Nassikas, 2015).

Biomass can be monitored by collecting field data on a large number of sampling units, which represents a heavy investment in time and cost (Wulder et al., 2012); however, as relationships between field-measured biomass density and data from different types of remote sensors have been identified, it has been possible to reduce the number of field samples without sacrificing accuracy (Saatchi et al., 2011).

Information on vegetation type and conditions is provided by indices estimated from spectral values of satellite images (Foody et al., 2001). NDVI (normalized difference vegetation index) and SAVI (soil-adjusted vegetation index) have been the most widely used for modeling aboveground biomass, quantification of tree structure and composition of tropical forests (Foody et al., 2001; Ghosh & Behera, 2018). Other indices (brightness, greenness and wetness) estimated by Tasseled Cap transformation are able to distinguish phenological changes and key attributes in different forest types and conditions (Deo et al., 2016). LiDAR (Light Detection And Ranging) active sensors are considered a suitable technology for the study of forest structure, as they provide detail and spatial accuracy. LiDAR-derived metrics tend to be highly correlated with aboveground biomass measurements; therefore, methodologies that take advantage of this relationship are more beneficial when incorporating satellite image products, which cover the study area perfectly and provide information in areas of difficult access (Wulder et al., 2012).

LiDAR technology has restrictions of use for monitoring large areas, due to its high cost and the large amount of data for storage and processing. The viable option is to obtain data in strategically established stripes and rely on satellite image data to estimate variables of interest at a regional scale (Saatchi et al., 2011; Wilkes et al., 2015); furthermore, it is known that the lowest biomass estimation errors are achieved when using LiDAR-derived data and auxiliary variables from optical

## Introducción

Los bosques tropicales representan 45 % del área forestal de la Tierra y tienen una capacidad elevada de almacenamiento de carbono (Food and Agriculture Organization of the United Nations [FAO], 2020; FAO & UNEP, 2020). Desafortunadamente, en las últimas décadas, estos ecosistemas han registrado tasas altas de deforestación y degradación (FAO, 2020) convirtiéndolos en emisores de bióxido de carbono a la atmósfera. Por tanto, el monitoreo de los inventarios de biomasa (carbono) y de los cambios de superficie de estos bosques es importante para la planificación de políticas que contribuyan a la mitigación de los efectos negativos del cambio climático (Dupuy-Rada, Hernández-Stefanoni, Hernández-Juárez, Tun-Dzul, & May-Pat, 2012; Houghton, Byers, & Nassikas, 2015).

La biomasa puede monitorearse mediante la colecta de datos en campo en un número grande de unidades de muestreo, lo que representa una inversión fuerte en tiempo y costos (Wulder et al., 2012); sin embargo, a medida que se han identificado las relaciones entre la densidad de biomasa medida y los datos de sensores remotos de diversa índole, ha sido posible reducir el número de muestras sin sacrificar la precisión (Saatchi et al., 2011).

La información sobre el tipo y las condiciones de vegetación se obtiene a través de índices calculados a partir de los valores espectrales de las imágenes satelitales (Foody et al., 2001). El NDVI (índice de vegetación de diferencia normalizada) y el SAVI (índice de vegetación ajustado al suelo) han sido los más utilizados para la modelación de biomasa aérea, cuantificación de la estructura arbórea y composición de bosques tropicales (Foody et al., 2001; Ghosh & Behera, 2018). Otros índices (brillo, verdor y humedad) obtenidos mediante la transformación Tasseled Cap son capaces de distinguir cambios fenológicos y atributos clave en varios tipos y condiciones del bosque (Deo et al., 2016). Por su parte, los sensores activos tipo LiDAR (Light Detection And Ranging) son considerados una tecnología adecuada para el estudio de la estructura del bosque, ya que brindan detalle y precisión espacial. Las métricas derivadas de LiDAR tienden a estar altamente correlacionadas con las observaciones de biomasa aérea; por lo tanto, las metodologías que aprovechan dicha relación son más beneficiadas cuando incorporan productos de imágenes satelitales, los cuales cubren el área de estudio completamente y proporcionan información en áreas de acceso difícil (Wulder et al., 2012).

La tecnología LiDAR presenta restricciones de uso para el monitoreo de áreas extensas, debido a su alto costo y a la gran cantidad de datos para el almacenamiento y procesamiento. La opción factible es obtener datos

images and radar data (Zolkos, Goetz, & Dubayah, 2013). Accuracy of estimates depends on factors such as landscape heterogeneity, density of sampled data, and the remote sensors used. Hence the importance of estimating and expressing, spatially, the uncertainty of estimates at the pixel level (Barbosa, Broadbent, & Bitencourt, 2014).

Based on the above, this study aimed to identify a combination of spectral and climate variables to estimate aboveground biomass for two types of medium-stature tropical forest in the Yucatan Peninsula; to evaluate the behavior of the models fitted with the random Forest algorithm; and to map the aboveground biomass and its associated uncertainty at the pixel level. Estimates of aboveground biomass and its associated uncertainty, expressed spatially, can contribute to the management of policies to mitigate the effects of climate change in tropical forests.

## Materials and Methods

### Study area

The study area includes the medium-stature semi-deciduous (SMSC) and semi-evergreen (SMSP) tropical forests of the Yucatan Peninsula, Mexico (Figure 1). It covers approximately 77 000 km<sup>2</sup>, where 28 000 km<sup>2</sup> correspond to SMSC and 49 000 km<sup>2</sup> to SMSP (Instituto Nacional de Estadística y Geografía [INEGI], 2013).

SMSC has a warm sub-humid ( $Aw_1$ ) climate with rain in summer (May-October) and a dry season (November-April). Mean annual temperature is 26 °C (Dupuy-Rada et al., 2012). SMSP has a warm sub-humid climate with mean annual precipitation of 950 mm (July-October) and mean annual temperature of 22 °C (Aryal, De Jong, Ochoa-Gaona, Esparza-Olguin, & Mendoza-Vega, 2014). In most of the Yucatan peninsula, precipitation gradient is dry to the northwest (600 mm) and wetter to the southeast (1 400 mm) (Martínez & Galindo, 2002).

### Aboveground biomass data

This study used aboveground biomass estimates reported by Ortiz-Reyes et al. (2019), corresponding to raster files. These authors employed an area-based approach for biomass estimation by linking field measurements (data from the Inventario Nacional Forestal y de Suelos de México, 2009-2014 remeasurement) with LiDAR metrics, using the random Forest algorithm. Estimates from such protocol, frequently, maintain and even exceed operational accuracy standards than those obtained through traditional inventory, with an acceptable level of bias (White et al., 2013). Each pixel (20 m x 20 m) containing an estimated aboveground biomass value about the stripes were considered as

en franjas establecidas estratégicamente y apoyarse en datos de imágenes satelitales para la estimación de las variables de interés a una escala regional (Saatchi et al., 2011; Wilkes et al., 2015); además, se reconoce que los errores de estimación de biomasa más bajos son aquellos que usan datos derivados de LiDAR y variables auxiliares de imágenes ópticas y datos de radar (Zolkos, Goetz, & Dubayah, 2013). La precisión de las estimaciones depende de factores como la heterogeneidad del paisaje, la densidad de datos muestreados, así como de los sensores remotos empleados. De ahí la importancia en calcular y expresar, espacialmente, la incertidumbre de las estimaciones a nivel de pixel (Barbosa, Broadbent, & Bitencourt, 2014).

Con base en lo anterior, en este trabajo se planteó identificar una combinación de variables espectrales y climáticas para la estimación de la biomasa aérea en dos tipos de selva mediana en la península de Yucatán; evaluar el comportamiento de los modelos ajustados con el algoritmo random Forest; y mapear la biomasa aérea y su incertidumbre asociada a nivel de pixel. Las estimaciones de biomasa aérea y su incertidumbre asociada, expresadas espacialmente, pueden contribuir a la gestión de políticas de mitigación de los efectos del cambio climático en bosques tropicales.

## Materiales y métodos

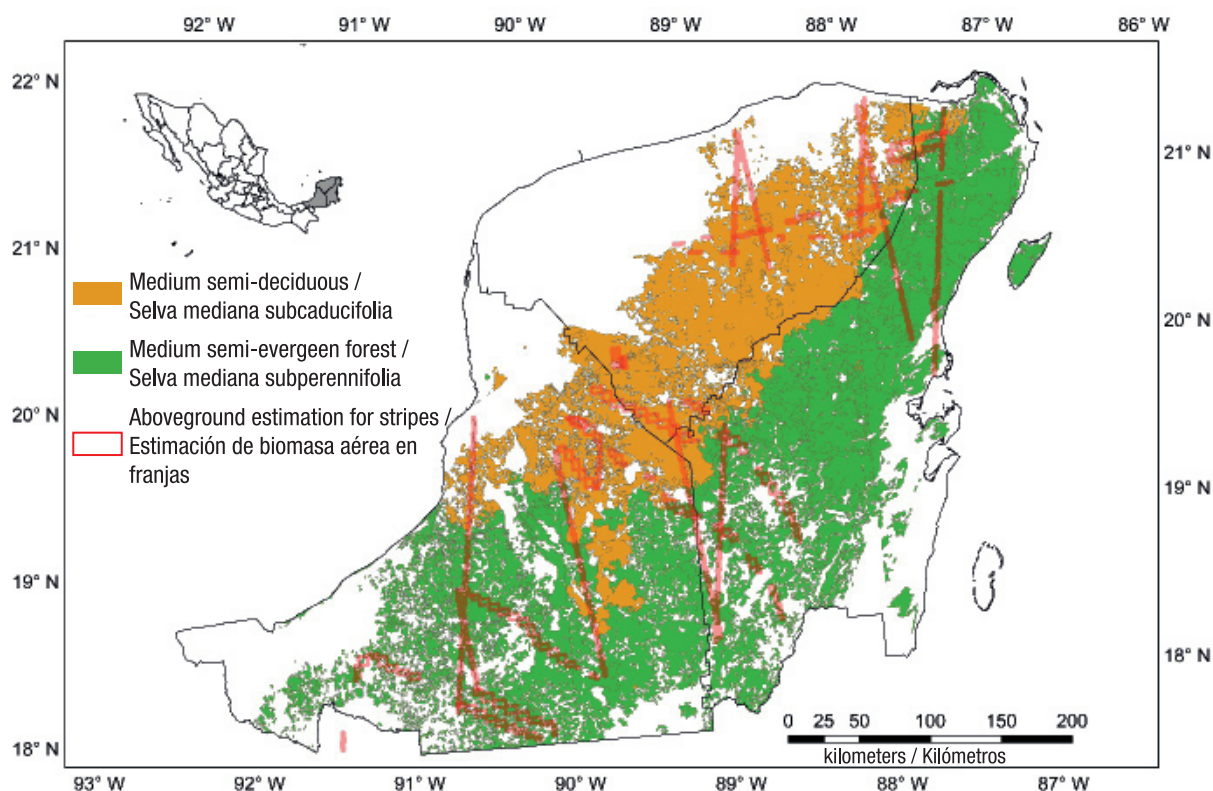
### Área de estudio

El área comprende las selvas medianas subcaducifolia (SMSC) y subperennifolia (SMSP) de la península de Yucatán, México (Figura 1). La superficie aproximada del área bajo análisis es de 77 000 km<sup>2</sup>, donde 28 000 km<sup>2</sup> corresponden a SMSC y 49 000 km<sup>2</sup> a la SMSP (Instituto Nacional de Estadística y Geografía [INEGI], 2013).

En la SMSC, el clima es cálido subhúmedo ( $Aw_1$ ) con lluvia en verano (mayo-octubre) y una estación seca (noviembre-abril). La temperatura media anual es 26 °C (Dupuy-Rada et al., 2012). En la SMSP, el clima es cálido subhúmedo con precipitación media anual de 950 mm (julio-octubre) y temperatura media anual de 22 °C (Aryal, De Jong, Ochoa-Gaona, Esparza-Olguin, & Mendoza-Vega, 2014). En la mayor parte de la península de Yucatán, el gradiente de precipitación es seco hacia el noroeste (600 mm) y más húmedo hacia el sureste (1 400 mm) (Martínez & Galindo, 2002).

### Datos de biomasa aérea

Se emplearon las estimaciones de biomasa aérea reportadas por Ortiz-Reyes et al. (2019), correspondientes a archivos ráster. Estos autores emplearon un enfoque basado en el área para la



**Figure 1. Study area: Medium-stature semi-deciduous and semi-evergreen tropical forests—series V, INEGI, 2013— of the Yucatan Peninsula, Mexico. Red stripes indicate the location of biomass data estimated from LiDAR data (Ortiz-Reyes et al., 2019).**

**Figura 1. Área de estudio: selvas mediana subcaducifolia y subperennifolia –serie V, INEGI, 2013– de la península de Yucatán, México. Las franjas rojas indican la ubicación de datos de biomasa estimados a partir de datos LiDAR (Ortiz-Reyes et al., 2019).**

“LiDAR plots”. These estimates increase the distribution and sample size of local data and are similar to field plot estimates (Wulder et al., 2012). Subsequently, the strip pixels were resampled to a spatial resolution of 30 m using the nearest neighbor method to match pixels from Landsat images (Cracknell, 1998).

#### Landsat images and climate data

Six Landsat 8 images taken by the OLI (Operational Land Imager) sensor, processed at the surface reflectance level (Vermote, Justice, Clavarié, & Franch, 2016) were downloaded from the United States Geological Survey database (USGS, 2017).

The image search period was one year (April 1, 2013 up to April 30, 2014) to establish closeness between the conditions of previous aboveground biomass estimates from LiDAR data with Landsat images. Those images with cloudiness less than 21 %, corresponding to the winter season, were downloaded (Table 1). A cloud mask was applied to each scene using the Pixel QA (Quality Assessment) filter (Vermote et al., 2016). The procedure was performed with the QGIS software

estimación de la biomasa mediante la vinculación de mediciones en campo (datos del Inventario Nacional Forestal y de Suelos de México, remediación 2009-2014) con métricas LiDAR, a través del algoritmo random Forest. Las estimaciones de dicho protocolo, frecuentemente, mantienen e incluso superan los estándares de precisión operacional que las obtenidas mediante el inventario tradicional, con un nivel de sesgo aceptable (White et al., 2013). Cada pixel (20 m x 20 m) que contenía un valor de biomasa aérea estimada sobre las franjas se consideraron como “parcelas LiDAR”. Estas aumentan la distribución y el tamaño de muestra de los datos locales y son similares a las estimaciones de las parcelas en campo (Wulder et al., 2012). Posteriormente, los píxeles de las franjas se remuestrearon a una resolución espacial de 30 m empleando el método del vecino más cercano, para hacerlos coincidentes con los píxeles de las imágenes Landsat (Cracknell, 1998).

#### Datos de imágenes Landsat y climáticos

Seis imágenes Landsat 8 obtenidas por el sensor OLI (Operational Land Imager) se procesaron a nivel de

**Table 1. Characteristics of Landsat 8 scenes processed at the ground reflectance level for the estimation of aboveground biomass for medium-stature tropical forests of the Yucatan Peninsula.**

**Cuadro 1. Características de las escenas Landsat 8 procesadas a nivel de reflectancia en superficie para la estimación de biomasa aérea en las selvas medianas de la península de Yucatán.**

| Landsat scene identifier /<br>Identificador de la escena Landsat | WRS Path | WRS Row | Scene cloud cover (%) /<br>Cobertura de nubes<br>de la escena (%) | Acquisition date /<br>Fecha de adquisición |
|------------------------------------------------------------------|----------|---------|-------------------------------------------------------------------|--------------------------------------------|
| LC80190452014046LGN01SR                                          | 19       | 45      | 2.08                                                              | February 15, 2014 /<br>febrero 15, 2014    |
| LC80190462014046LGN01SR                                          | 19       | 46      | 1.97                                                              | February 15, 2014 /<br>febrero 15, 2014    |
| LC80190472014046LGN01SR                                          | 19       | 47      | 10.79                                                             | February 15, 2014 /<br>febrero 15, 2014    |
| LC80200452014005LGN01SR                                          | 20       | 45      | 0.54                                                              | January 5, 2014 /<br>Enero 5, 2014         |
| LC80200462014021LGN01SR                                          | 20       | 46      | 8.30                                                              | January 5, 2014 /<br>Enero 5, 2014         |
| LC80200472014005LGN01SR                                          | 20       | 47      | 20.90                                                             | January 5, 2014 /<br>Enero 5, 2014         |

version 3.6 Noosa (QGIS, 2019), using CloudMasking plugin. Cloud and shadow areas were excluded from subsequent analyses.

Later, NDVI, MSAVI (modified soil-adjusted vegetation index), SAVI and EVI (enhanced vegetation index) spectral indices were created with the preprocessed images. Brightness, greenness and wetness indices were estimated using the Tasseled Cap transformation to take advantage of information from more bands, using coefficients for Landsat products with surface reflectance reported by Crist (1985). The above was estimated using the raster package in R (R Development Core Team, 2013). Spectral bands alone (2 to 7) were also used as independent variables in the estimation of aboveground biomass.

Climate information was taken from the WorldClim (2017) database which has monthly average, minimum and maximum temperature, and precipitation for the period 1970 to 2000. Average monthly temperature and monthly precipitation data were downloaded for January, February, November, and December, in addition to mean annual temperature (°C) and annual precipitation (mm), biologically significant variables (Fick & Hijmans, 2017). All these variables had ~1 km<sup>2</sup> resolution so they were resampled to 30 m, using the nearest neighbor method, to match them with the other variables. These variables were chosen because of their proven relevance in other forest parameter estimation studies (Ahmed, Franklin, Wulder, & White, 2015; Wilkes et al., 2015). The list of predictor variables processed is shown in Table 2.

reflectancia en superficie (Vermote, Justice, Clavarié, & Franch, 2016). Las imágenes se descargaron en la base de datos del Servicio Geológico de Estados Unidos (USGS, 2017).

El periodo de búsqueda en las imágenes fue un año (1 de abril del 2013 hasta el 30 de abril del 2014) para establecer cercanía entre las condiciones de las estimaciones previas de biomasa aérea de los datos LiDAR con las imágenes Landsat. Se descargaron aquellas imágenes con nubosidad menor del 21 %, correspondientes a la época de invierno (Cuadro 1). Una máscara de nubes se aplicó a cada escena empleando el filtro Pixel QA (Quality Assessment) (Vermote et al., 2016). El procedimiento se realizó en el programa QGIS versión 3.6 Noosa (QGIS, 2019), empleando el complemento CloudMasking. Las áreas de nubes y sombras se excluyeron de los análisis subsiguientes.

Posteriormente, los índices espectrales NDVI, MSAVI (índice de vegetación ajustado al suelo modificado), SAVI y EVI (índice de vegetación mejorado) se crearon con las imágenes preprocesadas. Los índices de brillo, verdor y humedad se calcularon mediante la transformación Tasseled Cap para aprovechar la información de más bandas, empleando los coeficientes para productos Landsat con reflectancia en superficie señalados por Crist (1985). Lo anterior se calculó utilizando el paquete *raster* en R (R Development Core Team, 2013). También se emplearon las bandas espectrales por sí solas (2 a 7) como variables independientes en la estimación de biomasa aérea.

Table 2. Predictor variables (spectral and climate variables) used for modeling of aboveground biomass. Variables correspond to raster files. Cuadro 2. Variables predictor (espectrales y climáticas) empleadas en el modelamiento de la biomasa aérea. Las variables corresponden a archivos ráster.

| Variable (abbreviation)/<br>Variable (abreviación) | Features/Formula/<br>Características/Fórmula                                                                                                                                      | Trait/<br>Propiedad                                                                                                                                                                                |
|----------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Band 2 / Banda 2<br>(B2fc)                         | B2 blue (λ: 0.452 - 0.512 μm) / B2 Azul (λ: 0.452 - 0.512 μm)                                                                                                                     | Differentiates vegetation soil and deciduous<br>coniferous vegetation (USGS, 2019) /<br>Diferencia el suelo de la vegetación y la<br>vegetación decidua de coníferas (USGS, 2019)                  |
| Band 3 / Banda 3<br>(B3fc)                         | B3 green (λ: 0.533 - 0.590 μm) / B3 Verde (λ: 0.533 - 0.590 μm)                                                                                                                   | Evaluates plant vigor (USGS, 2019) /<br>Evalúa el vigor de la planta (USGS, 2019)                                                                                                                  |
| Band 4 / Banda 4<br>(B4fc)                         | B4 red (λ: 0.636 - 0.673 μm) / B4 Rojo (λ: 0.636 - 0.673 μm)                                                                                                                      | Discriminates vegetation slopes (USGS, 2019) /<br>Discrimina las pendientes de vegetación (USGS, 2019)                                                                                             |
| Band 5 / Banda 5<br>(B5fc)                         | B5 Near infrared (λ: 0.851 - 0.879 μm) / B5 Infrarrojo cercano (λ: 0.851 - 0.879 μm)                                                                                              | Emphasizes moisture conditions of plants<br>and soils (Young et al., 2017) /<br>Enfatiza condiciones de humedad de<br>plantas y suelos (Young et al., 2017)                                        |
| Band 6 / Banda 6<br>(B6fc)                         | B6 Shortwave infrared 1 (λ: 1.566 - 1.651 μm) /<br>B6 Infrarrojo de onda corta 1 (λ: 1.566 - 1.651 μm)                                                                            | Emphasizes moisture conditions of plants<br>and soils (Young et al., 2017) /<br>Enfatiza condiciones de humedad de<br>plantas y suelos (Young et al., 2017)                                        |
| Band 7 / Banda 7<br>(B7fc)                         | B7 Shortwave infrared 2 (λ: 2.107 - 2.294 μm) /<br>B7 Infrarrojo de onda corta 2 (λ: 2.107 - 2.294 μm)                                                                            | Enhances soil and vegetation moisture<br>content (USGS, 2019) /<br>Mejora el contenido de humedad del<br>suelo y la vegetación (USGS, 2019)                                                        |
| NDVI<br>(bNDVIfc)                                  | $NDVI = \frac{(Band\ 5 - Band\ 4)}{Band\ 5 + Band\ 4}$ $NDVI = \frac{(Banda\ 5 - Banda\ 4)}{Banda\ 5 + Banda\ 4}$                                                                 | Sensitive to photosynthetic activity<br>(Ghosh & Behera, 2018) /<br>Sensible a la actividad fotosintética<br>(Ghosh & Behera, 2018)                                                                |
| MSAVI<br>(bMSAVIfc)                                | $MSAVI = [2 * Band\ 5 + 1 - \sqrt{(2 * Band\ 5 + 1)^2 - 8 * (Band\ 5 - Band\ 4)}] / 2$ $MSAVI = [2 * Banda\ 5 + 1 - \sqrt{(2 * Banda\ 5 + 1)^2 - 8 * (Banda\ 5 - Banda\ 4)}] / 2$ | Sensitive to the amount of vegetation (Qi et al., 1994) /<br>Sensible a la cantidad de vegetación (Qi et al., 1994)                                                                                |
| SAVI<br>(bSAVIfc)                                  | $SAVI = \left( \frac{Band\ 5 - Band\ 4}{Band\ 5 + Band\ 4 + 0.05} \right)$ $SAVI = \left( \frac{Banda\ 5 - Banda\ 4}{Banda\ 5 + Banda\ 4 + 0.05} \right)$                         | Highly<br>correlated with vegetation cover dynamics<br>(Gao, Huete, Ni, & Miura, 2000) /<br>Altamente<br>correlacionado con la dinámica de la cobertura<br>vegetal (Gao, Huete, Ni, & Miura, 2000) |

**Table 2. Predictor variables (spectral and climate variables) used for modeling of aboveground biomass. Variables correspond to raster files. Cuadro 2. Variables predictor (espectrales y climáticas) empleadas en el modelamiento de la biomasa aérea. Las variables corresponden a archivos ráster.**

| Variable (abbreviation)/<br>Variable (abreviación)                | Features/Formulas/<br>Características/Fórmula                                                                                                                                                                                                                                                                                                           | Trait/<br>Propiedad                                                                                                                                                                                                                                                     |
|-------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| EVI<br>(bEVI/c)                                                   | $EVI = 2.5 * \left( \frac{\text{Band 5} - \text{Band 4}}{\text{Band 5} + 6 * \text{Band 4} - 7.5 * \text{Band 2} + 1} \right)$ $EVI = 2.5 * \left( \frac{\text{Banda 5} - \text{Banda 4}}{\text{Banda 5} + 6 * \text{Banda 4} - 7.5 * \text{Banda 2} + 1} \right)$                                                                                      | Sensitive to canopy structural variations (Gao et al., 2000; Vieilledent et al., 2016)/<br>Sensible a variaciones estructurales del dosel (Gao et al., 2000; Vieilledent et al., 2016)                                                                                  |
| TCB<br>(brighVal)                                                 | $TCB = 0.2043 * \text{Band 2} + 0.4158 * \text{Band 3} + 0.5524 * \text{Band 4} + 0.5741 * \text{Band 5} + 0.3124 * \text{Band 6} + 0.2303 * \text{Band 7}$ $TCB = 0.2043 * \text{Banda 2} + 0.4158 * \text{Banda 3} + 0.5524 * \text{Banda 4} + 0.5741 * \text{Banda 5} + 0.3124 * \text{Banda 6} + 0.2303 * \text{Banda 7}$                           | Sensitive to ground brightness (Crist, 1985)/<br>Sensible al brillo de la superficie (Crist, 1985)                                                                                                                                                                      |
| TCG/TCV<br>(GreenVal)                                             | $TCG = (-0.1603) * \text{Band 2} + (-0.2819) * \text{Band 3} + (-0.4934) * \text{Band 4} + 0.7940 * \text{Band 5} (-0.0002) * \text{Band 6} + (-0.1446) * \text{Band 7}$ $TCV = (-0.1603) * \text{Banda 2} + (-0.2819) * \text{Banda 3} + (-0.4934) * \text{Banda 4} + 0.7940 * \text{Banda 5} (-0.0002) * \text{Banda 6} + (-0.1446) * \text{Banda 7}$ | Sensitive to greenness of vegetation (Crist, 1985)/<br>Sensible al verdor de la vegetación (Crist, 1985)                                                                                                                                                                |
| TCW/TCV<br>(WetVal)                                               | $TCW = 0.0315 * \text{Band 2} + 0.2021 * \text{Band 3} + 0.3102 * \text{Band 4} + 0.1594 * \text{Band 5} + (-0.6806) * \text{Band 6} + (-0.6109) * \text{Band 7}$ $TCH = 0.0315 * \text{Banda 2} + 0.2021 * \text{Banda 3} + 0.3102 * \text{Banda 4} + 0.1594 * \text{Banda 5} + (-0.6806) * \text{Banda 6} + (-0.6109) * \text{Banda 7}$               | Sensitive to moisture content of vegetation (Crist, 1985)/<br>Sensible al contenido de humedad de la vegetación (Crist, 1985)                                                                                                                                           |
| Mean annual temperature/<br>Temperatura media<br>anual (Var_Biot) | Data from 1970 to 2000<br>°C at 30" spatial resolution<br>(~1 km <sup>2</sup> )/<br>Datos desde 1970 – 2000<br>°C a 30" de resolución espacial<br>(~1 km <sup>2</sup> )                                                                                                                                                                                 | Influence vegetation growth and mortality<br>processes (Álvarez-Dávila et al., 2017)/<br>Influye en los procesos de crecimiento y mortalidad de la vegetación (Álvarez-Dávila et al., 2017)                                                                             |
| Annual precipitation/<br>Precipitación anual<br>(Var_Bio12)       | Data from 1970 to 2000<br>mm at 30" spatial resolution (~1 km <sup>2</sup> )/<br>Datos desde 1970 – 2000<br>mm a 30" de resolución espacial (~1 km <sup>2</sup> )                                                                                                                                                                                       | Positive relationship with biomass.<br>Influence vegetation growth and mortality<br>processes (Álvarez-Dávila et al., 2017)/<br>Relación positiva con la biomasa.<br>Influye en los procesos de crecimiento y mortalidad de la vegetación (Álvarez-Dávila et al., 2017) |

Table 2. Predictor variables (spectral and climate variables) used for modeling of aboveground biomass. Variables correspond to raster files. Cuadro 2. Variables predictor (espectrales y climáticas) empleadas en el modelamiento de la biomasa aérea. Las variables corresponden a archivos ráster.

| Variable (abbreviation) /<br>Variable (abreviación)                                                                                                                                                                                                                              | Features/Formula /<br>Características/Fórmula                                            | Trait /<br>Propiedad                                                                                                                                                              |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Average temperature for<br>January (TemAv_M01),<br>February (TemAv_M02),<br>November (TemAv_M11) and December<br>(TemAv_M12) °C/<br>Temperatura promedio<br>de los meses de<br>enero (TemAv_M01),<br>febrero (TemAv_M02),<br>noviembre (TemAv_M11) y<br>diciembre (TemAv_M12) °C | °C at 30" spatial resolution<br>(~1 km²)/<br>°C a 30" de resolución espacial<br>(~1 km²) | Influence activation of growth processes<br>in plants (Fick & Hijmans, 2017)/<br>Influye en la activación de los procesos de<br>crecimiento en las plantas (Fick & Hijmans, 2017) |
| Average precipitation for<br>January (Prec_M01),<br>February (Prec_M02),<br>November (Prec_M11) and<br>December (Prec_M12)/<br>Precipitación promedio<br>de los meses de<br>enero (Prec_M01),<br>febrero (Prec_M02),<br>noviembre (Prec_M11) y<br>diciembre (Prec_M12)           | mm at 30" spatial resolution (~1 km²)/<br>mm a 30" de resolución espacial (~1 km²)       | Influence activation of growth processes<br>in plants (Fick & Hijmans, 2017)/<br>Influye en la activación de los procesos de<br>crecimiento en las plantas (Fick & Hijmans, 2017) |

NDVI: normalized difference vegetation index; SAVI: soil adjusted vegetation index; MSAVI: modified soil adjusted vegetation index; EVI: enhanced vegetation index; TCB: Tasseled Cap brightness, TCG: Tasseled Cap greenness, TCW: Tasseled Cap wetness.  
NDVI: índice de vegetación de diferencia normalizada; SAVI: índice de vegetación ajustado al suelo; MSAVI: índice de vegetación ajustado al suelo modificado; EVI: índice de vegetación mejorado; TCB: Tasseled Cap brillo, TCV: Tasseled Cap verdor, TCH: Tasseled Cap humedad.

## Aboveground biomass estimate using the random Forest algorithm

From aboveground biomass data previously estimated by Ortiz-Reyes et al. (2019) in transects with LiDAR data (more than 300 000 pixels for each vegetation type), a sample of 5 000 pixels per vegetation type was randomly selected without replacement to fit two models and estimate the biomass for the entire area of interest. The sample of 5 000 pixels represented the values of the dependent variable (aboveground biomass). Climate data and spectral data recorded in Landsat images, corresponding to the same pixels of the random sample, represented the independent variables.

The random Forest algorithm of R (R Development Core Team, 2013) builds a set of decision trees from training data, which are internally validated to generate a prediction of the response variable given the predictors (Cutler, Cutler, & Stevens, 2012). The algorithm is easy to apply and capable of processing large databases efficiently, as an option in regional studies (Asner & Mascaro, 2014). Final predictor variables were selected regarding the influence that each one represented on the mean squared error (MSE) of the fitted model. Sequentially, the algorithm evaluated the performance of the model for each vegetation type based on RMSE (root mean square error), number of predictor variables and percentage of variance explained.

## Mapping aboveground biomass in two types of tropical forest

Aboveground biomass maps were created in the R raster package (R Development Core Team, 2013). The maps were produced with the previously generated model using the raster files corresponding to the spectral and climate variables chosen by the model as relevant for predicting aboveground biomass.

## Quantification of uncertainty

Uncertainty refers to the level of ignorance of the true value of a parameter or variable of interest due to multiple factors and can be quantified with common statistical estimators such as standard deviation, coefficient of variation (CV) or by an interval with a preset confidence level (Global Observation of Forest and Land Cover Dynamics [GOF-C-GOLD]). This study evaluated and mapped uncertainty of aboveground biomass predictions by CV associated with the estimates generated at the pixel level. Estimations were performed using the ModelMap package of R (Freeman, Frescino, & Moisen, 2018).

La información climática se obtuvo de la base de datos del WorldClim (2017) que dispone de temperatura mensual promedio, mínima y máxima, y precipitación para el periodo 1970 a 2000. Se descargaron los datos de temperatura mensual promedio y precipitación mensual de enero, febrero, noviembre y diciembre, además de la temperatura media anual (°C) y la precipitación anual (mm), variables biológicamente significativas (Fick & Hijmans, 2017). Todas estas variables tenían resolución de ~1 km<sup>2</sup> por lo que fueron remuestreadas a 30 m, usando el método del vecino más cercano, para hacerlas coincidentes con las demás variables. Estas variables se eligieron por su relevancia probada en otros estudios de estimación de parámetros forestales (Ahmed, Franklin, Wulder, & White, 2015; Wilkes et al., 2015). La lista de variables predictoras procesadas se muestra en el Cuadro 2.

## Estimación de biomasa aérea mediante el algoritmo random Forest

A partir de los datos de biomasa aérea, estimados previamente por Ortiz-Reyes et al. (2019) en transectos con datos LiDAR (más de 300 000 píxeles para cada tipo de vegetación), se seleccionó aleatoriamente y sin remplazo una muestra de 5 000 píxeles por tipo de vegetación para ajustar sendos modelos y estimar la biomasa para toda el área de interés. La muestra de 5 000 píxeles representó los valores de la variable dependiente (biomasa aérea). Los datos climáticos y los datos espectrales registrados en las imágenes Landsat, correspondientes a los mismos píxeles de la muestra aleatoria, representaron a las variables independientes.

El algoritmo random Forest de R (R Development Core Team, 2013) construye un conjunto de árboles de decisión a partir de datos de entrenamiento, mismos que son validados internamente para generar una predicción de la variable de respuesta dados los predictores (Cutler, Cutler, & Stevens, 2012). El algoritmo es de aplicación fácil y con capacidad de procesamiento de grandes bases de datos de manera eficiente, lo que permite ser una opción en estudios regionales (Asner & Mascaro, 2014). Las variables predictoras definitivas se seleccionaron considerando la influencia que cada una representó sobre el error cuadrático medio (MSE) del modelo ajustado. De manera iterativa, el algoritmo evaluó el rendimiento del modelo para cada tipo de vegetación con base en la RMSE (raíz del error cuadrático medio), el número de variables predictoras y el porcentaje de varianza explicado.

## Mapeo de biomasa aérea en los dos tipos de selva mediana

Los mapas de biomasa aérea se crearon en el paquete *raster* de R (R Development Core Team,

Results and Discussion

Models for estimating aboveground biomass in two tropical forest types

A separate model was fitted to estimate aboveground biomass in each medium-stature tropical forest type using the random Forest algorithm. The main parameters are shown in Table 3.

Predictive ability of models is within the range reported in other studies for tropical forests ( $R^2 = 0.50-0.92$ ). Those studies used similar data and approaches to this research; for example, Lu et al. (2012) estimated aboveground biomass in the Amazon basin using a multiple regression model and differentiated mature ( $R^2 = 0.50$ ) and secondary successional ( $R^2 = 0.76$ ) forests. The authors point out that aboveground biomass estimation using Landsat images is site-dependent, due to variation in phenology, vegetation type and structure. In contrast, Basuki, Skidmore, Hussin, and Van Duren (2013) used images taken by a synthetic aperture radar (SAR) and Landsat ETM+ imagery for aboveground biomass estimation in tropical forests under management in Indonesia. By regression models, these authors explained 75 % of the variance ( $RMSE = 78.9 \text{ Mg}\cdot\text{ha}^{-1}$ ), while in a tropical forest in Malaysia, Phua et al. (2017) attributed 63 % of the variance to LiDAR metrics and 18 % to Landsat 8 green band texture variables ( $RMSE = 112.15 \text{ Mg}\cdot\text{ha}^{-1}$ ). Meanwhile, Ghosh and Behera (2018) estimated the aboveground biomass of two species grown in a tropical forest in India with SAR data and Sentinel-2A imagery; the explained variance was 60 % and 71 % ( $RMSE = 79.45 \text{ Mg}\cdot\text{ha}^{-1}$ ;  $105.02 \text{ Mg}\cdot\text{ha}^{-1}$ ) with random Forest and Gradient Boosting autonomous learning techniques, respectively. At the regional scale,

2013). Los mapas se produjeron con el modelo generado previamente utilizando los archivos ráster correspondientes a las variables espectrales y climáticas elegidas por el modelo como relevantes para predecir la biomasa aérea.

Cuantificación de incertidumbre

La incertidumbre refiere el nivel de desconocimiento del valor verdadero de un parámetro o de una variable de interés debido a factores múltiples y se puede cuantificar con estimadores estadísticos comunes como la desviación estándar, el coeficiente de variación (CV) o mediante un intervalo con un nivel de confianza preestablecido (Global Observation of Forest and Land Cover Dynamics [GOFD-GOLD]). Este estudio evaluó y mapeó la incertidumbre de las predicciones de biomasa aérea a través del CV asociado a las estimaciones generadas a nivel de pixel. El cálculo se realizó con el paquete ModelMap de R (Freeman, Frescino, & Moisen, 2018).

Resultados y discusión

Modelos para estimar biomasa aérea en dos tipos de selva mediana

Se ajustó un modelo por separado para la estimación de la biomasa aérea en cada tipo de selva mediante el algoritmo random Forest. Los parámetros principales se indican en el Cuadro 3.

La capacidad predictiva de los modelos está dentro del intervalo reportado en otros estudios para bosques tropicales ( $R^2 = 0.50-0.92$ ). En dichos estudios se emplean datos y enfoques similares a los de esta investigación; por ejemplo, Lu et al. (2012) estimaron

Table 3. Relevant parameters of the random Forest models for aboveground biomass estimation per type of medium-stature tropical forest in the Yucatan Peninsula.

Cuadro 3. Parámetros relevantes de los modelos random Forest para la estimación de biomasa aérea por tipo de selva mediana en la península de Yucatán.

| Parameters/<br>Parámetros                                         | Semi-evergreen tropical forest/<br>Selva mediana subperennifolia | Semi-deciduous tropical forest/<br>Selva mediana subcaducifolia |
|-------------------------------------------------------------------|------------------------------------------------------------------|-----------------------------------------------------------------|
| $R^2$                                                             | 0.50                                                             | 0.50                                                            |
| r                                                                 | 0.71                                                             | 0.70                                                            |
| RMSE ( $\text{Mg}\cdot\text{ha}^{-1}$ )                           | 34.10                                                            | 26.20                                                           |
| Number of predictor variables/<br>Número de variables predictoras | 12                                                               | 15                                                              |

$R^2$ : coefficient of determination, r: correlation between measured vs. predicted aboveground biomass data, RMSE: root mean squared error.  
 $R^2$ : coeficiente de determinación, r: correlación entre los datos de biomasa aérea observados vs. predichos, RMSE: raíz del error cuadrático medio.

Asner and Mascaro (2014) estimated aboveground carbon density in 14 tropical ecoregions in five countries and, by fitting nonlinear maximum likelihood models, explained 92.3 % of the variance (RMSE = 17.12 Mg C·ha<sup>-1</sup>).

In the previous cases, prediction method yield was superior to that obtained in this study; however, the errors obtained were also high (RMSE between 78.9 Mg·ha<sup>-1</sup> and 112.15 Mg·ha<sup>-1</sup>) compared to those of this study (RMSE = 34.1 Mg·ha<sup>-1</sup> and 26.2 Mg·ha<sup>-1</sup> for SMSP and SMSC, respectively). However, if results are compared with the regional study of Asner and Mascaro (2014), the reported error is similar in terms of aboveground biomass.

Regarding the variables selected by the random Forest algorithm (Figure 2), results showed that the green band (B3fc) and annual precipitation (Var\_Bio12) were the most important in models for estimating aboveground biomass in SMSP and SMSC, respectively. Climate variables prevailed in predicting aboveground biomass in both medium-stature tropical forest types, pointing to an influence of climatic processes on estimated aboveground biomass. This influence has been recognized in several studies on biomass estimation in tropical vegetation (Álvarez-Dávila et al., 2017; Vieilledent et al., 2016).

Regarding precipitation, the main constraint of dry forests is water in the soil, which could suggest relevance of annual precipitation (Var\_Bio12) in the SMSC model, while monthly precipitation averages (Prec\_M01, Prec\_M02) remained in the modeling of both vegetation types. Cao et al. (2015) mention that the growth of this forest type not only varies with age, soil type or land use background, but also with precipitation.

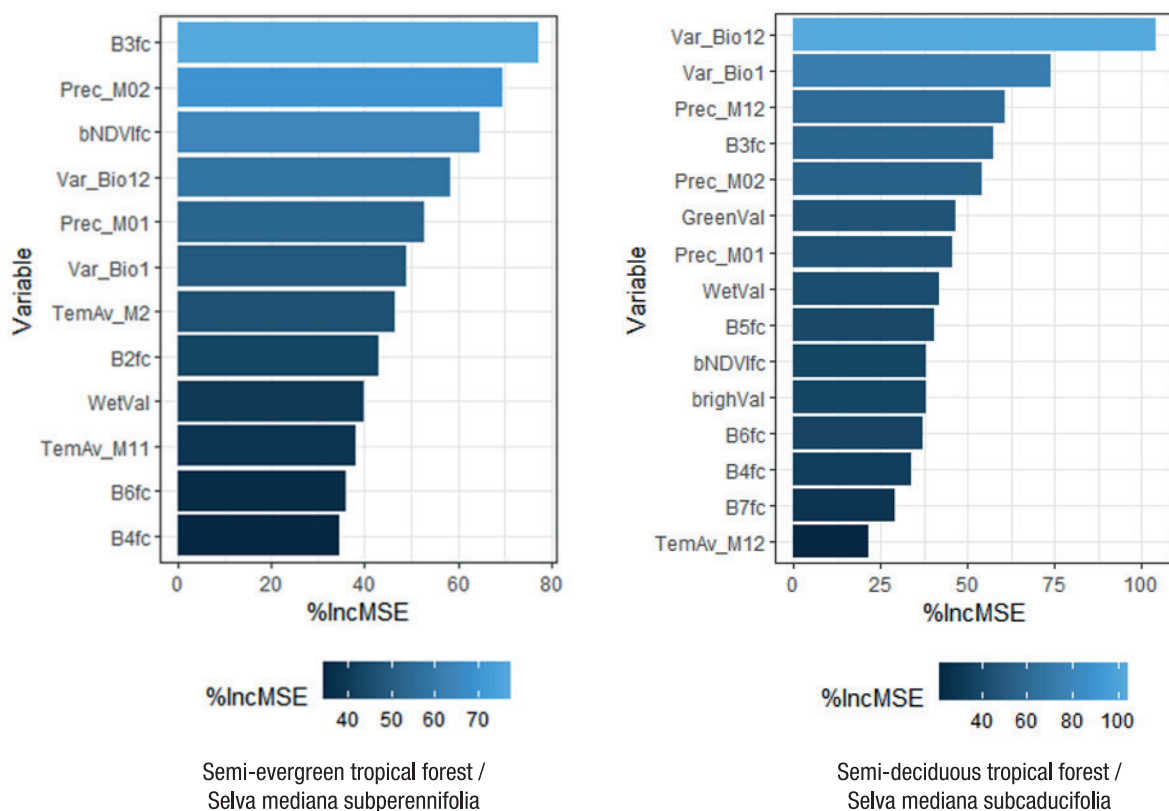
The fact that precipitation and temperature remained as relevant variables in the models could be an indication of the relationship between amount of available water and its interaction with temperature to influence biomass growth processes. In such a case, both precipitation and temperature would be having superior control over aboveground biomass density in tropical forests, because both vary regionally and are scale-dependent (Álvarez-Dávila et al., 2017; White & Hood, 2004). Saatchi et al. (2011) report that spatial variability of aboveground biomass depends on climate, natural and human-induced disturbance and recovery processes, soil type and variations in topography. Martínez and Galindo (2002) mentioned that high spatial and temporal variability of precipitation, geological substrate and scarce development of the soil were decisive factors in the distribution of vegetation in an area with similar characteristics to the one evaluated in this study.

la biomasa aérea en la cuenca amazónica mediante un modelo de regresión múltiple y diferenciaron bosques maduros ( $R^2 = 0.50$ ) y en sucesión secundaria ( $R^2 = 0.76$ ). Los autores señalan que la estimación de biomasa aérea usando imágenes Landsat es dependiente del sitio, debido a la variación en la fenología, el tipo y estructura de la vegetación. En contraste, Basuki, Skidmore, Hussin, y Van Duren (2013) usaron imágenes obtenidas por un radar de apertura sintética (SAR, por sus siglas en inglés) e imágenes Landsat ETM+ para la estimación de la biomasa aérea en bosques tropicales bajo manejo en Indonesia. A través de modelos de regresión, dichos autores explicaron 75 % de la varianza (RMSE = 78.9 Mg·ha<sup>-1</sup>), mientras que en un bosque tropical de Malasia, Phua et al. (2017) adjudicaron 63 % de la varianza a las métricas LiDAR y 18 % a las variables de textura de la banda verde de Landsat 8 (RMSE = 112.15 Mg·ha<sup>-1</sup>). Por su parte, Ghosh y Behera (2018) estimaron la biomasa aérea de dos especies plantadas en un bosque tropical en India con datos SAR e imágenes Sentinel-2A; la varianza explicada fue 60 % y 71 % (RMSE = 79.45 Mg·ha<sup>-1</sup>; 105.02 Mg·ha<sup>-1</sup>) con técnicas de aprendizaje autónomo random Forest y Gradient Boosting, respectivamente. A escala regional, Asner y Mascaro (2014) estimaron la densidad de carbono aéreo en 14 ecorregiones tropicales de cinco países y, mediante el ajuste de modelos de probabilidad máxima no lineal, explicaron 92.3 % de la varianza (RMSE = 17.12 Mg C·ha<sup>-1</sup>).

En los casos anteriores, el rendimiento de los métodos de predicción fue superior al obtenido en el presente trabajo; sin embargo, los errores obtenidos también fueron altos (RMSE entre 78.9 Mg·ha<sup>-1</sup> y 112.15 Mg·ha<sup>-1</sup>) en comparación con los de este estudio (RMSE = 34.1 Mg·ha<sup>-1</sup> y 26.2 Mg·ha<sup>-1</sup> para la SMSP y SMSC, respectivamente). No obstante, si los resultados se comparan con el estudio regional de Asner y Mascaro (2014), el error reportado es similar en términos de biomasa aérea.

Con relación a las variables seleccionadas por el algoritmo random Forest (Figura 2), los resultados mostraron que la banda verde (B3fc) y la precipitación anual (Var\_Bio12) fueron las más importantes en los modelos para la estimación de biomasa aérea en la SMSP y SMSC, respectivamente. Las variables climáticas prevalecieron para predecir la biomasa aérea en ambos tipos de selvas medianas, apuntando a una influencia de los procesos climáticos sobre la biomasa aérea estimada. Dicha influencia ha sido reconocida en varios trabajos de estimación de biomasa en vegetación tropical (Álvarez-Dávila et al., 2017; Vieilledent et al., 2016).

Respecto a la precipitación, la principal limitante de las selvas secas es el agua en el suelo, lo cual podría sugerir



**Figure 2. Relative importance of variables selected by the random Forest model in the medium-stature semi-evergreen tropical forest (12 variables) and the medium-stature semi-deciduous tropical forest (15 variables), for the estimation of aboveground biomass.** B3fc: reflectance of Band 3 – green from Landsat 8 OLI sensor, Prec\_M02: precipitation month 2 (February) (mm); bNDVIfc: normalized difference vegetation index; Var\_Bio12: annual precipitation (mm); Prec\_M01: precipitation month 1 (January, mm); Var\_Bio1: mean annual temperature (°C); TemAv\_M2: mean temperature of month 02 (February, °C); B2fc: reflectance of Band 2 - blue from Landsat 8 OLI sensor; WetVal: wetness in Tasseled Cap transformation; TemAv\_M11: mean temperature of month 11 (November, °C); B6fc: reflectance of Band 6 - shortwave infrared 1 from Landsat 8 OLI sensor; B4fc: reflectance of Band 4 – red from Landsat 8 OLI sensor; Prec\_M12: precipitation of month 12 (December, mm); GreenVal: greenness in the Tasseled Cap transform; B5fc: reflectance of Band 5 - near infrared from Landsat 8 OLI sensor; brighVal: brightness in Tasseled Cap transform; B7fc: reflectance of Band 7 - shortwave infrared 2 from Landsat 8 OLI sensor; TeamAv\_12: average temperature of month 12 (December, °C). IncMSE %: percent increase in mean square error.

**Figura 2. Importancia relativa de las variables seleccionadas por el modelo random Forest en la selva mediana subperennifolia (12 variables) y la selva mediana subcaducifolia (15 variables), para la estimación de biomasa aérea.** B3fc: reflectancia de la Banda 3 - verde del sensor Landsat 8 OLI; Prec\_M02: precipitación del mes 2 (febrero) (mm); bNDVIfc: Índice de vegetación de diferencia normalizada; Var\_Bio12: precipitación anual (mm); Prec\_M01: precipitación del mes 1 (enero, mm); Var\_Bio1: temperatura media anual (°C); TemAv\_M2: temperatura promedio del mes 02 (febrero, °C); B2fc: reflectancia de la Banda 2 - azul del sensor Landsat 8 OLI; WetVal: humedad en la transformación Tasseled Cap; TemAv\_M11: temperatura promedio del mes 11 (noviembre, °C); B6fc: reflectancia de la Banda 6 – infrarrojo de onda corta 1 del sensor Landsat 8 OLI; B4fc: reflectancia de la Banda 4 – roja del sensor Landsat 8 OLI; Prec\_M12: precipitación del mes 12 (diciembre, mm); GreenVal: verdor en la transformación Tasseled Cap; B5fc: reflectancia de la Banda 5 - infrarrojo cercano del sensor Landsat 8 OLI; brighVal: brillo en la transformación Tasseled Cap; B7fc: reflectancia de la Banda 7 – infrarrojo de onda corta 2 del sensor Landsat 8 OLI; TeamAv\_12: temperatura promedio del mes 12 (diciembre, °C). % IncMSE: incremento porcentual del error cuadrático medio.

For biomass prediction in the SMSC, the random Forest algorithm selected the same or similar spectral variables that have shown good predictive capacity in forests with similar conditions. Freitas, Mello, and Cruz (2005) report that NDVI is a good indicator of aboveground biomass for dry and deciduous tropical forests. Of the spectral indices, NDVI was the only one that was maintained for both models, the rest of the indices were removed because they did not contribute to the yield of models. The blue, green and infrared bands were similar components to those reported in the study of Foody et al. (2001), who indicate the importance of regarding all useful sensor bands and not only the indices dependent on the red band. The green, red and infrared bands were maintained as explanatory variables in both vegetation types, highlighting the green band (B3fc). Such relevance was also reported by Foody et al. (2001) and Phua et al. (2017).

### Mapping aboveground biomass for two types of medium-stature tropical forest

Figure 3 shows the spatial variation of aboveground biomass in the two types of medium-stature tropical forest. SMSP biomass ranged from 4.0 to 185.7 Mg·ha<sup>-1</sup> (mean  $\bar{y}$  = 85.2; standard deviation  $s$  = 23.2), an interval that is within the results reported by other authors for the same type of vegetation. Aryal et al. (2014) reported values from 11.72 to 99.56 Mg C·ha<sup>-1</sup> for four-year-old secondary forests ( $s$  = 4.92) and for mature forests ( $s$  = 20.83), which is similar to the interval reported in this study. Recently, Hernández-Stefanoni et al. (2020) reported 127.5 Mg·ha<sup>-1</sup> of average aboveground biomass and a CV lower than 40 %.

For SMSC, aboveground biomass ranged from 11.7 to 117 Mg·ha<sup>-1</sup> ( $\bar{y}$  = 51.1;  $s$  = 17.5). This value is within the range reported by Dupuy-Rada et al. (2012) for dry tropical forests of the Yucatan Peninsula ( $\bar{y}$  = 56 Mg·ha<sup>-1</sup>). Similarity could be due to the fact that aboveground biomass data of the two studies come from mosaics of forest fragments at different successional ages and spatial arrangement. For this vegetation type, Hernández-Stefanoni et al. (2020) reported 100.4 Mg·ha<sup>-1</sup> average aboveground biomass and Dai et al. (2014) estimated 5.0 to 115.0 Mg C·ha<sup>-1</sup> with  $\bar{y}$  = 56.6 Mg C·ha<sup>-1</sup>.

On the other hand, there are estimates that report higher amounts of aboveground biomass than that reported in the present study; e.g., Hernández-Stefanoni et al. (2014) reported mean biomass values of 109.71 Mg·ha<sup>-1</sup> and 376.77 Mg·ha<sup>-1</sup> for SMSC and SMSP, respectively, when they used field sampling plots of 1 000 m<sup>2</sup>. These same authors reported mean biomass values of 147.2 and 270.2 Mg·ha<sup>-1</sup> for SMSC and SMSP, respectively, when using 400 m<sup>2</sup> field sampling plots in the same study area. This shows the complexity of

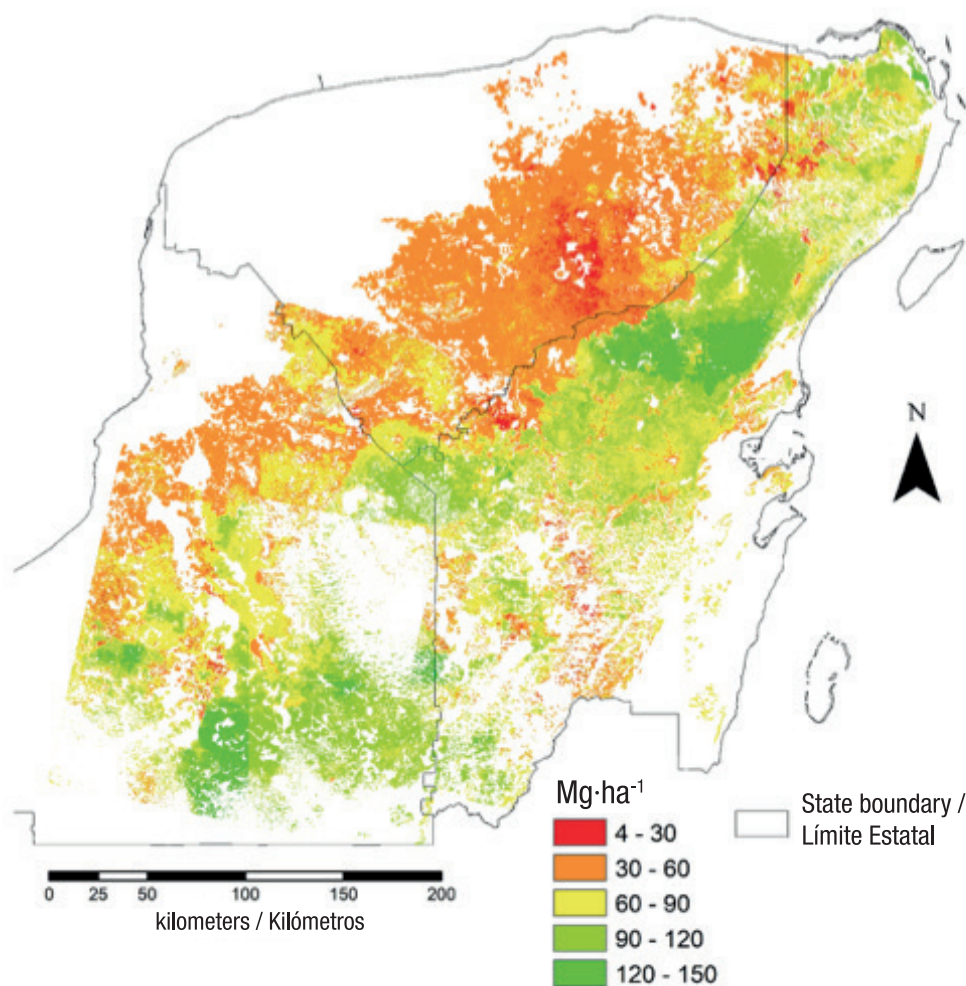
la relevancia de la precipitación anual (Var\_Bio12) en el modelo de la SMSC, mientras que los promedios mensuales de precipitación (Prec\_M01, Prec\_M02) permanecieron en la modelación de ambos tipos de vegetación. En este sentido, Cao et al. (2015) mencionan que el crecimiento de este tipo de bosques no solo varía con la edad, el tipo de suelo o el antecedente del uso de la tierra, sino también con la precipitación.

El hecho de que la precipitación y la temperatura permanecieran como variables relevantes en los modelos generados podría ser un indicio de la relación entre la cantidad de agua disponible y su interacción con la temperatura para influir sobre los procesos de crecimiento en biomasa. En tal caso, tanto la precipitación como la temperatura estarían ejerciendo un control superior sobre la densidad de biomasa aérea en los bosques tropicales, puesto que ambos varían regionalmente y son dependientes de la escala (Álvarez-Dávila et al., 2017; White & Hood, 2004). Saatchi et al. (2011) señalan que la variabilidad espacial de la biomasa aérea depende del clima, de procesos de perturbación y recuperación provocados de forma natural y por el hombre, del tipo de suelo y variaciones en la topografía. Por su parte, Martínez y Galindo (2002) mencionan que la alta variabilidad espacial y temporal de la precipitación, el sustrato geológico y el desarrollo escaso del suelo fueron factores decisivos en la distribución de la vegetación en un área con características similares a la evaluada en este estudio.

Para la predicción de biomasa en la SMSC, el algoritmo random Forest seleccionó las mismas variables espectrales o similares que han mostrado buena capacidad predictiva en bosques de condiciones semejantes. Freitas, Mello, y Cruz (2005) señalan que el NDVI es un buen indicador de la biomasa aérea en bosques secos y caducifolios. De los índices espectrales, el NDVI fue el único que se mantuvo para ambos modelos, el resto de índices se eliminaron ya que no contribuyeron al rendimiento de los modelos. Las bandas azul, verde e infrarroja fueron componentes similares a los reportados en el trabajo de Foody et al. (2001), quienes señalan la importancia de considerar todas las bandas útiles del sensor y no solo los índices dependientes de la banda roja. Las bandas verde, roja e infrarroja se mantuvieron como variables explicativas en ambos tipos de vegetación, destacando la banda verde (B3fc). Dicha relevancia también fue reportada por Foody et al. (2001) y Phua et al. (2017).

### Mapeo de biomasa aérea en los dos tipos de selva mediana

La Figura 3 muestra la variación espacial de la biomasa aérea en los dos tipos de selva mediana. La biomasa de la SMSP varió entre 4.0 y 185.7 Mg·ha<sup>-1</sup>



**Figure 3. Spatial distribution of average aboveground biomass ( $\text{Mg}\cdot\text{ha}^{-1}$ ) in medium-stature semi-evergreen and semi-deciduous tropical forests of the Yucatan Peninsula, Mexico.**

**Figura 3. Distribución espacial de la biomasa aérea promedio ( $\text{Mg}\cdot\text{ha}^{-1}$ ) en las selvas mediana subperennifolia y subcaducifolia de la península de Yucatán, México.**

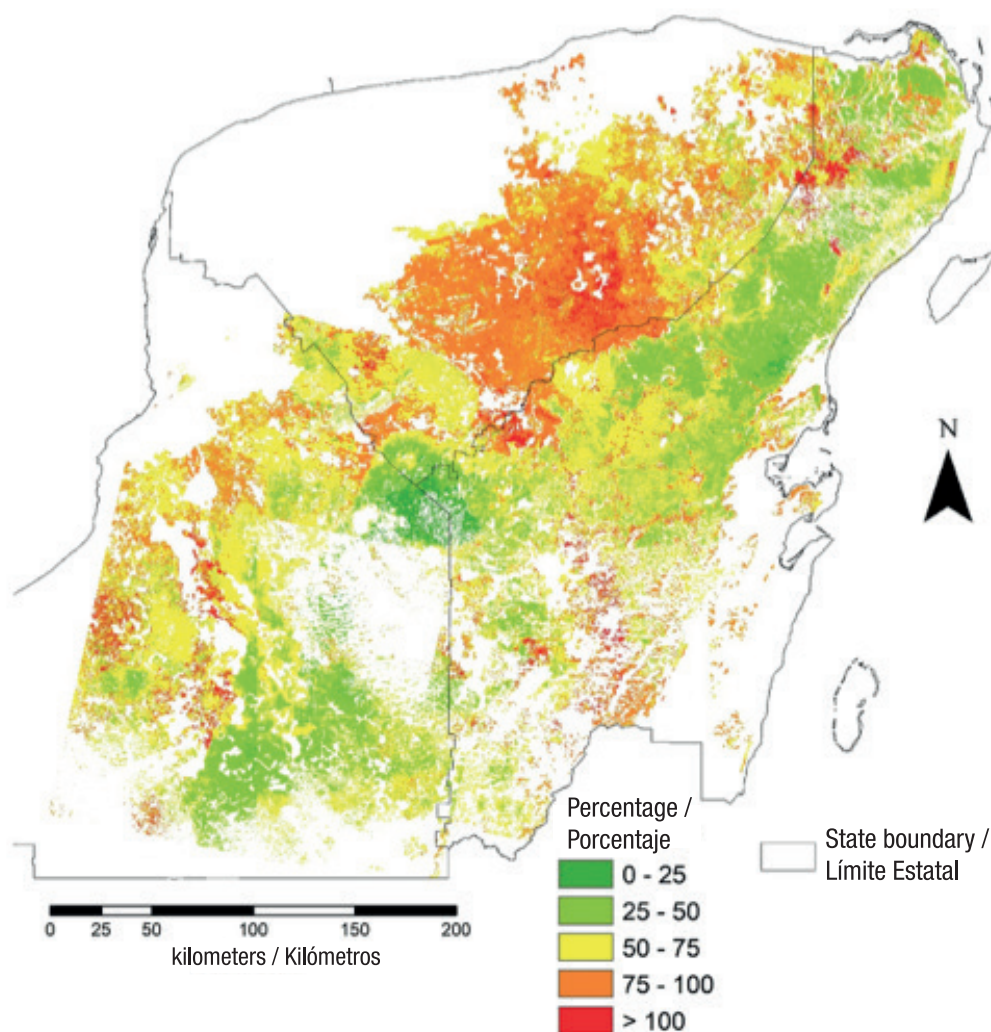
comparing results between studies of similar purpose, but using different methods or analysis approaches, specially when the size of the areas under analysis is uneven and landscape elements are contrasting as a result of spatially haphazard successional states due to anthropogenic activities and natural disturbances (Aryal et al., 2014; Dupuy-Rada et al., 2012).

#### **Spatial uncertainty of aboveground biomass predictions**

Spatial variability of prediction uncertainty was expressed as the percentage of CV (Figure 4). For SMSP, the CV ranged mostly between 25 and 75 % and was of higher magnitude for SMSC (75 to more than 100 %), particularly in areas with low average aboveground biomass and adjacent to others where information was removed, due to the hiding process to exclude cells containing clouds. These CV values, in general, are

(media  $[\bar{y}] = 85.2$ ; desviación estándar  $[s] = 23.2$ ), intervalo que se encuentra dentro de los resultados reportados por otros autores para el mismo tipo de vegetación. Aryal et al. (2014) reportaron valores de 11.72 a 99.56  $\text{Mg C}\cdot\text{ha}^{-1}$  en bosques secundarios de cuatro años ( $s = 4.92$ ) y en bosques maduros ( $s = 20.83$ ), lo cual es similar al intervalo reportado en este estudio. Recientemente, Hernández-Stefanoni et al. (2020) reportaron 127.5  $\text{Mg}\cdot\text{ha}^{-1}$  de biomasa aérea promedio y un CV menor de 40 %.

En el caso de la SMSC, la biomasa aérea osciló entre 11.7 y 117  $\text{Mg}\cdot\text{ha}^{-1}$  ( $\bar{y} = 51.1$ ;  $s = 17.5$ ). Este valor se encuentra dentro del intervalo reportado por Dupuy-Rada et al. (2012) para bosques tropicales secos de la península de Yucatán ( $\bar{y} = 56 \text{ Mg}\cdot\text{ha}^{-1}$ ). La similitud podría deberse a que los datos de biomasa aérea de los dos estudios provienen de mosaicos de fragmentos de bosque en edades sucesionales y arreglo espacial diferentes.



**Figure 4. Spatial distribution of uncertainty (% coefficient of variation) of aboveground biomass in the medium-stature semi-evergreen and semi-deciduous tropical forests of the Yucatan Peninsula, Mexico.**

**Figura 4. Distribución espacial de la incertidumbre (% coeficiente de variación) de biomasa aérea en las selvas medianas subperennifolia y subcaducifolia de la península de Yucatán, México.**

higher than those reported by Hernández-Stefanoni et al. (2020) for the same vegetation types (0 to 75 %, but mostly below 60 %); however, it is important to highlight the contrast in the size of the area analyzed in both studies. The aforementioned authors analyzed 3 600 km<sup>2</sup> of each vegetation type, while in this study 28 000 km<sup>2</sup> of SMSC and 49 000 km<sup>2</sup> of SMSP were analyzed, therefore, it is to be expected that variability is greater.

Like most of the scarce research, this study used the CV to report spatial variation of uncertainty in biomass estimation. This highlights the importance of assessing uncertainty per component to identify which component contributes the most error to estimates. For example, two components that probably affected the results of this study are temporal discordance between field and remote sensing data, and

Para este tipo de vegetación, Hernández-Stefanoni et al. (2020) reportaron 100.4 Mg·ha<sup>-1</sup> de biomasa aérea promedio y Dai et al. (2014) estimaron 5.0 a 115.0 Mg C·ha<sup>-1</sup> con  $\bar{y} = 56.6$  Mg C·ha<sup>-1</sup>.

Por otra parte, existen estimaciones que reportan cantidades mayores de biomasa aérea que la reportada en el presente estudio; e.g., Hernández-Stefanoni et al. (2014) reportaron valores medios de biomasa de 109.71 Mg·ha<sup>-1</sup> y 376.77 Mg·ha<sup>-1</sup> para la SMSC y SMSP, respectivamente, cuando utilizaron parcelas de muestreo en campo de 1 000 m<sup>2</sup>. Estos mismos autores reportaron valores medios de biomasa de 147.2 y 270.2 Mg·ha<sup>-1</sup> para la SMSC y SMSP, respectivamente, cuando utilizaron parcelas de muestreo en campo de 400 m<sup>2</sup> en la misma área de estudio. Esto muestra la complejidad de la comparación de resultados entre estudios de propósito similar, pero que utilizan métodos

the lack of a priori spatial planning of remote sensing data collection. Another component responsible for measured uncertainty is the model previously fitted by Ortiz-Reyes et al. (2019) to estimate aboveground biomass in stripes, whose data were used in this study as a starting point to fit a larger model by the random Forest algorithm. Therefore, the use of approaches that correct for the errors that each component adds is a pending task that could improve the precision of aboveground biomass estimates.

The analysis performed provides a record of the current biomass quantification effort and offers points of comparison on the road to improving uncertainty quantification methodologies in complex forest ecosystems. On the other hand, results represent an attempt to standardize the reports of spatial variation of uncertainty as an important part of forest aboveground biomass estimation.

## Conclusions

Structural variability of medium-stature semi-deciduous (SMSC) and semi-evergreen (SMSP) tropical forests of the Yucatan Peninsula was collected by training data from stripes, which impacted the performance of models for predicting aboveground biomass in both vegetation types. Models provided a continuous map detailing spatial distribution of aboveground biomass at the pixel level for SMSC and SMSP. This distribution was explained in greater proportion by precipitation and temperature. The error of predictions, expressed as the coefficient of variation, allowed spatially explicit visualization of uncertainty associated with aboveground biomass estimation at 30 m resolution. Both the methodology and the results of this study are acceptable regarding the available elements and represent a contribution towards the development of more effective methods for estimating aboveground biomass at the regional level.

*End of English version*

## References / Referencias

- Ahmed, O. S., Franklin, S. E., Wulder, M. A., & White, J. C. (2015). Characterizing stand-level forest canopy cover and height using Landsat time series, samples of airborne LiDAR, and the Random Forest algorithm. *ISPRS Journal of Photogrammetry and Remote Sensing*, 101, 89–101. doi: 10.1016/j.isprsjprs.2014.11.007
- Álvarez-Dávila, E., Cayuela, L., González-Caro, S., Aldana, A. M., Stevenson, P. R., Phillips, O., ... Rey-Benayas, J. M. (2017). Forest biomass density across large climate gradients in northern South America is related to water availability or enfoques de análisis diferentes, particularmente cuando la dimensión de las áreas bajo análisis es desigual y los elementos paisajísticos son contrastantes como resultado de estados sucesionales espacialmente azarosos, a causa de actividades antrópicas y disturbios naturales (Aryal et al., 2014; Dupuy-Rada et al., 2012).
- Incertidumbre espacial de las predicciones de biomasa aérea**
- La variabilidad espacial de la incertidumbre de las predicciones se expresó como el porcentaje de CV (Figura 4). Para la SMSP, el CV osciló mayoritariamente entre 25 y 75 % y fue de magnitud mayor para la SMSC (75 a más de 100 %), particularmente en áreas con biomasa aérea promedio baja y aledañas a otras donde se eliminó información, debido al proceso de enmascaramiento para excluir las celdas que contenían nubes. Estos valores de CV, en general, son mayores que los reportados por Hernández-Stefanoni et al. (2020) para los mismos tipos de vegetación (0 a 75 %, pero mayoritariamente por debajo de 60 %); sin embargo, es importante resaltar el contraste en el tamaño de la superficie analizada en ambos estudios. Los autores antes mencionados analizaron 3 600 km<sup>2</sup> de cada tipo de vegetación, mientras que en este estudio se analizaron 28 000 km<sup>2</sup> de SMSC y 49 000 km<sup>2</sup> de SMSP, por tanto, es de esperar que la variabilidad sea mayor.
- Como la mayoría de las escasas investigaciones, este trabajo utilizó el CV para reportar la variación espacial de la incertidumbre en la estimación de biomasa. Lo anterior pone de manifiesto la importancia de la evaluación de la incertidumbre por componente, para identificar cual aporta mayor error sobre las estimaciones. Por ejemplo, dos componentes que probablemente afectaron los resultados de este trabajo es la discordancia temporal entre los datos en campo y los de sensores remotos, y la no planificación espacial *a priori* de la colecta de datos con sensores remotos. Otro componente responsable de la incertidumbre observada es el modelo ajustado previamente por Ortiz-Reyes et al. (2019) para estimar la biomasa aérea en franjas, cuyos datos se utilizaron en este estudio como punto de partida para ajustar un modelo de mayor envergadura a través del algoritmo random Forest. Por lo tanto, el uso de enfoques que corrijan los errores que cada componente agrega es una tarea pendiente que podría mejorar la precisión de las estimaciones de biomasa aérea.
- El análisis realizado deja constancia del esfuerzo actual de cuantificación de biomasa y ofrece puntos de comparación en la ruta de mejora de las metodologías de cuantificación de incertidumbre en ecosistemas forestales complejos. Por otro lado, los resultados representan un intento de estandarización de los

- but not with temperature. *PLoS ONE*, 12(3), e0171072. doi: 10.1371/journal.pone.0171072
- Aryal, D. R., De Jong, B. H. J., Ochoa-Gaona, S., Esparza-Olguin, L., & Mendoza-Vega, J. (2014). Carbon stocks and changes in tropical secondary forests of southern Mexico. *Agriculture, Ecosystems & Environment*, 195, 220–230. doi: 10.1016/j.agee.2014.06.005
- Asner, G. P., & Mascaro, J. (2014). Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. *Remote Sensing of Environment*, 140, 614–624. doi: 10.1016/j.rse.2013.09.023
- 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*, Article ID 715796. doi: 10.1155/2014/715796
- Basuki, T. M., Skidmore, A. K., Hussin, Y. A., & Van Duren, I. (2013). Estimating tropical forest biomass more accurately by integrating ALOS PALSAR and Landsat-7 ETM+ data. *International Journal of Remote Sensing*, 34(13), 4871–4888. doi: 10.1080/01431161.2013.777486
- Cao, S., Yu, Q., Sanchez-Azofeifa, A., Feng, J., Rivard, B., & Gu, Z. (2015). Mapping tropical dry forest succession using multiple criteria spectral mixture analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 109, 1–29. doi: 10.1016/j.isprsjprs.2015.08.009
- Cracknell, A. P. (1998). Synergy in remote sensing-what's in a pixel? *International Journal of Remote Sensing*, 19(11), 2025–2047. doi: 10.1080/014311698214848
- Crist, E. P. (1985). A TM tasseled cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, 17(3), 301–306. doi: 10.1016/0034-4257(85)90102-6
- Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random forests. In C. Zhang, & Y. Ma (Eds.), *Ensemble machine learning: Methods and applications* (pp. 157–175). New York, USA: Springer.
- Dai, Z., Birdsey, R. A., Johnson, K. D., Dupuy, J. M., Hernandez-Stefanoni, J. L., & Richardson, K. (2014). Modeling Carbon stocks in a secondary tropical dry forest in the Yucatan Peninsula, Mexico. *Water, Air, & Soil Pollution*, 225, Article 1925. doi: 10.1007/s11270-014-1925-x
- Deo, R. K., Russell, M. B., Domke, G. M., Woodall, C. W., Falkowski, M. J., & Cohen, W. B. (2016). Using landsat time-series and LiDAR to inform aboveground forest biomass baselines in Northern Minnesota, USA. *Canadian Journal of Remote Sensing*, 43(1), 28–47. doi:10.1080/07038992.2017.1259556
- Dupuy-Rada, J., Hernández-Stefanoni, J., Hernández-Juárez, R., Tun-Dzul, F., & May-Pat, F. (2012). Efectos del cambio de uso del suelo en la biomasa y diversidad de plantas leñosas en un paisaje de bosque tropical seco en Yucatán. *Investigación Ambiental Ciencia y Política Pública*, 4(2), 130–140. Retrieved from [https://cicy.repositorioinstitucional.mx/jspui/bitstream/1003/1242/1/id26022\\_Dupuy\\_Juan.pdf](https://cicy.repositorioinstitucional.mx/jspui/bitstream/1003/1242/1/id26022_Dupuy_Juan.pdf)
- Food and Agriculture Organization of the United Nations (FAO). (2020). *Global forest resources assessment 2020: Main report*. Rome: Author. doi: 10.4060/ca9825en
- FAO & UNEP. (2020). *The state of the world's forests 2020. Forests, biodiversity and people*. Rome: Author. doi: 10.4060/ca8642en
- reportes de la variación espacial de incertidumbre como parte importante de la estimación de biomasa aérea forestal.
- ## Conclusiones
- La variabilidad estructural de las selvas medianas subcaducifolia (SMSC) y subperennifolia (SMSP) de la península de Yucatán se capturó en buena medida por los datos de entrenamiento provenientes de franjas, lo que repercutió en el desempeño de los modelos para la predicción de biomasa aérea en ambos tipos de vegetación. Los modelos generados permitieron la obtención de un mapa continuo que detalla la distribución espacial de biomasa aérea a nivel de pixel en la SMSC y SMSP. Dicha distribución se explicó en mayor proporción por la precipitación y la temperatura. El error de las predicciones, expresado como el coeficiente de variación, permitió la visualización espacialmente explícita de la incertidumbre asociada a la estimación de biomasa aérea a una resolución de 30 m. Tanto la metodología empleada como los resultados de este estudio son aceptables considerando los elementos disponibles y representan un aporte hacia el desarrollo de métodos más efectivos de estimación de biomasa aérea a nivel regional.
- ## Fin de la versión en español
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315. doi: 10.1002/joc.5086
- Foody, G. M., Cutler, M. E., Mcmorrow, J., Pelz, D., Tangki, H., Boyd, D. S., & Douglas, I. (2001). Mapping the biomass of Bornean tropical rain forest from remotely sensed data. *Global Ecology and Biogeography*, 10(4), 379–387. doi: 10.1046/j.1466-822X.2001.00248.x
- Freeman, E. A., Frescino, T. S., & Moisen, G. G. (2018). ModelMap: an R package for model creation and map production. Retrieved from <https://cran.r-project.org/web/packages/ModelMap/vignettes/VModelMap.pdf>
- Freitas, S. R., Mello, M. C. S., & Cruz, C. B. M. (2005). Relationships between forest structure and vegetation indices in Atlantic Rainforest. *Forest Ecology and Management*, 218(1–3), 353–362. doi: 10.1016/j.foreco.2005.08.036
- Gao, X., Huete, A. R., Ni, W., & Miura, T. (2000). Optical-biophysical relationships of vegetation spectra without background contamination. *Remote Sensing of Environment*, 74(3), 609–620.
- Ghosh, S. M., & Behera, M. D. (2018). Aboveground biomass estimation using multi-sensor data synergy and machine learning algorithms in a dense tropical forest. *Applied Geography*, 96, 29–40. doi: 10.1016/j.apgeog.2018.05.011
- Global Observation of Forest and Land Cover Dynamics (GOFC-GOLD). (2016). A sourcebook of methods and

- procedures for monitoring and reporting anthropogenic greenhouse gas emissions and removals associated with deforestation, gains and losses of carbon stocks in forests remaining forests, and forestation. GOF-C-GOLD Report version COP22-1. Retrieved from <http://www.monitoreoforestal.gob.mx/repositorioidigital/items/show/459>
- Hernández-Stefanoni, J., Dupuy, J., Johnson, K., Birdsey, R., Tun-Dzul, F., Peduzzi, A., ...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
- Hernández-Stefanoni, J. L., Castillo-Santiago, M. Á., Mas, J. F., Wheeler, C. E., Andres-Mauricio, J. A., Tun-Dzul, F., . . . Vaca, R. (2020). Improving aboveground biomass maps of tropical dry forests by integrating LiDAR, ALOS PALSAR, climate and field data. *Carbon Balance and Management*, 15(1), 1–17. Retrieved from <https://cbmjournals.biomedcentral.com/articles/10.1186/s13021-020-00151-6>
- Houghton, R. A., Byers, B., & Nassikas, A. A. (2015). A role for tropical forests in stabilizing atmospheric CO<sub>2</sub>. *Nature Climate Change*, 5(12), 1022–1023. doi:10.1038/nclimate2869
- Instituto Nacional de Estadística y Geografía (INEGI). (2013). Conjunto Nacional de Uso del Suelo y Vegetación a escala 1:250,000. Aguascalientes, México: Author.
- 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*, Article ID 436537. doi: 10.1155/2012/436537
- Martínez, E., & Galindo, L. C. (2002). La vegetación de Calakmul, Campeche, México: clasificación, descripción y distribución. *Boletín de la Sociedad Botánica de México*, 71, 7–32. doi: 10.17129/botsoci.1660
- 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
- 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
- Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, 48(2), 119–126. doi: 10.1016/0034-4257(94)90134-1
- QGIS (2019). QGIS Geographic Information System. Open Source Geospatial Foundation Project. (Version 3.6 Noosa). Retrieved from <https://qgis.org>
- R Development Core Team. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
- Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T., Salas, W., . . . Morel, A. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences*, 108(24), 9899–9904. doi: 10.1073/pnas.1019576108
- USGS. (2017). EROS science processing architecture on demand interface. Retrieved January, 2017, from <https://espa.cr.usgs.gov/ordering/new>
- USGS. (2019). What are the best Landsat spectral bands for use in my research? Retrieved January, 2017, from [https://www.usgs.gov/faqs/what-are-best-landsat-spectral-bands-use-my-research?qt-news\\_science\\_products=7#qt-news\\_science\\_products](https://www.usgs.gov/faqs/what-are-best-landsat-spectral-bands-use-my-research?qt-news_science_products=7#qt-news_science_products)
- Vermote, E., Justice, C., Claverie, M., & Franch, B. (2016). Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sensing of Environment*, 185, 46–56. doi: 10.1016/j.rse.2016.04.008
- Vieilledent, G., Gardi, O., Grinand, C., Burren, C., Andriamanjato, M., Camara, C., . . . Lines, E. (2016). Bioclimatic envelope models predict a decrease in tropical forest carbon stocks with climate change in Madagascar. *Journal of Ecology*, 104(3), 703–715. doi: 10.1111/1365-2745.12548
- White, D. A., & Hood, C. S. (2004). Vegetation patterns and environmental gradients in tropical dry forests of the northern Yucatan Peninsula. *Journal of Vegetation Science*, 15(2), 151–160. doi: 10.1111/j.1654-1103.2004.tb02250.x
- 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-Berelov, 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
- WorldClim (2017). Historical climate data. Retrieved January 2017 from <https://worldclim.org/data/worldclim21.html>
- Wulder, M. A., White, J. C., Bater, C. W., Coops, N. C., Hopkinson, C., & Chen, G. (2012). Lidar plots – A new large-area data collection option: Context, concepts, and case study. *Canadian Journal of Remote Sensing*, 38(5), 600–618. doi: 10.5589/m12-049
- Young, N. E., Anderson, R. S., Chignell, S. M., Vorster, A. G., Lawrence, R., & Evangelista, P. H. (2017). A survival guide to Landsat preprocessing. *Ecology*, 98(4), 920–932. doi: 10.1002/ecy.1730
- 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