



**Mapeo de la  
biomasa aérea  
de los bosques  
mediante datos  
de sensores remotos y**

**R**

**JOSÉ LUIS HERNÁNDEZ STEFANONI • MIGUEL ÁNGEL CASTILLO SANTIAGO  
JUAN ANDRÉS MAURICIO • JEAN-FRANÇOIS MAS  
FERNANDO TUN DZUL • JUAN MANUEL DUPUY RADA**







# Mapeo de la biomasa aérea de los bosques mediante datos de sensores remotos y R



Mapeo de la biomasa aérea de los bosques  
mediante datos de sensores remotos y  
**R**

José Luis Hernández Stefanoni  
Miguel Ángel Castillo Santiago  
Juan Andrés Mauricio  
Jean-François Mas  
Fernando Tun Dzul  
Juan Manuel Dupuy Rada



**EE**

**333.7516097267**

**M3**

*Mapeo de la biomasa aérea de los bosques mediante datos de sensores remotos y R* / José Luis Hernández Stefanoni, Miguel Ángel Castillo Santiago, Juan Andrés Mauricio, Jean-François Mas, Fernando Tun Dzul, Juan Manuel Dupuy Rada.- Chetumal, Quintana Roo, México : El Colegio de la Frontera Sur : Centro de Investigación Científica de Yucatán, A.C., 2021.

127 páginas : mapas ; 17x23 centímetros

ISBN: 978-607-8767-14-4 (ECOSUR)

ISBN: 978-607-7823-44-5 (CICY)

Bibliografía: páginas 121-126

1. Biomasa forestal, 2. Biomasa aérea, 3. R (Lenguaje de programación para computadora), 4. Sensores remotos, 5. Imágenes por satélites, 6. Inventarios forestales, 7. Felipe Carrillo Puerto (Quintana Roo, México), I. Hernández Stefanoni, José Luis (autor), II. Castillo Santiago, Miguel Ángel (autor), III. Andrés Mauricio, Juan (autor) , IV. Mas Causel, Jean-François (autor), V. Tun Dzul, Fernando (autor), VI. Dupuy Rada, Juan Manuel (autor)

D.A.R. 2021. *Mapeo de la biomasa aérea de los bosques mediante datos de sensores remotos y R*. José Luis Hernández Stefanoni, Miguel Ángel Castillo Santiago, Juan Andrés Mauricio, Jean-François Mas, Fernando Tun Dzul, Juan Manuel Dupuy Rada, El Colegio de la Frontera Sur (ECOSUR) y Centro de Investigación Científica de Yucatán, A.C. (CICY).

**Esta obra debe citarse de la siguiente forma:**

Hernández-Stefanoni, José Luis, Miguel Ángel Castillo Santiago, Juan Andrés Mauricio, Jean-François Mas, Fernando Tun Dzul, Juan Manuel Dupuy Rada. 2021. *Mapeo de la biomasa aérea de los bosques mediante datos de sensores remotos y R*. El Colegio de la Frontera Sur y Centro de Investigación Científica de Yucatán, A.C. México. 128 p.

Primera edición: Enero de 2021.

D. R. © El Colegio de la Frontera Sur  
Av. Centenario Km 5.5, C. P. 77014  
Chetumal, Quintana Roo, México  
[www.ecosur.mx](http://www.ecosur.mx)

D. R. © Centro de Investigación Científica de Yucatán, A. C.  
Calle 43 # 130 x 32 y 34, Col. Chuburná de Hidalgo, C. P. 97205  
Mérida, Yucatán, México  
[www.cicy.mx](http://www.cicy.mx)

Esta publicación fue sometida a un estricto proceso de arbitraje por pares, con base en los lineamientos establecidos por el Comité Editorial de El Colegio de la Frontera Sur

Impreso y hecho en México  
*Printed and made in Mexico*



# Contenido

Presentación .....	11
Propósitos del manual .....	11
Estructura del manual .....	12
Convenciones de escritura .....	12
Introducción .....	13
Datos de campo para la estimación de la biomasa .....	15
Datos de radar de apertura sintética .....	15
1. Estimación de la biomasa aérea del bosque .....	17
1.1 La biomasa del bosque.....	17
1.2 Estimación de la distribución espacial de la biomasa área.....	18
2. Configuración del entorno del <i>software</i> .....	23
2.1 Descargar e instalar R .....	23
2.2 Descargar e instalar RStudio .....	23
2.3 Cómo iniciar con R y RStudio.....	24
2.4 Paquetes de R.....	24
3. Cálculo de la biomasa aérea en campo .....	29
3.1 Área de estudio y unidades de observación.....	29
3.2 Creación de la base de datos .....	31
3.3 Correcciones taxonómicas .....	37
3.4 Extraer valores de densidad de la madera de bases de datos locales y globales .....	40
3.5 Cálculo de la biomasa a nivel de árbol individual .....	42
3.6 Cálculo de la biomasa y otros atributos de la vegetación por conglomerado.....	46
4. Preprocesamiento de las imágenes ALOS PALSAR .....	51
4.1 ¿Cómo descargar las imágenes de ALOS PALSAR?.....	51
4.2 Preparación de datos y descompresión de archivos .....	53
4.3 Creación de un mosaico con las escenas de ALOS PALSAR y recorte del área de estudio .....	57

4.4 Convertir valores de números digitales a retrodispersión en las polarizaciones HH y HV .....	61
4.5 Aplicación del filtro de Lee .....	63
4.6 Convertir las coordenadas de las imágenes preprocesadas, de geográficas a proyectadas.....	64
5. Procesamiento de las imágenes ALOS PALSAR .....	69
5.1 Configuración del directorio de trabajo y lectura de datos.....	69
5.2 Cálculo de NDBI .....	72
5.3 Cálculo de las medidas de textura en HH, HV y NDBI.....	73
5.4 Corrección de archivos.....	78
6. Preparación de bases de datos para la modelación .....	83
6.1 Crear un archivo espacial con la base de datos de campo.....	83
6.2 Extracción de los valores de las imágenes de textura con las coordenadas de puntos .....	86
6.3 Extracción de los valores de las polarizaciones HH, HV y NDBI .....	92
6.4 Conversión de datos espaciales a <i>dataframe</i> .....	98
7. Modelación y mapeo de la biomasa.....	99
7.1 Construcción del modelo de Random Forest para estimar la biomasa.....	99
7.2 Validación del modelo para estimar la biomasa .....	104
7.3 Crear un mapa tipo <i>raster</i> con la estimación de la biomasa.....	107
7.4 Impresión de mapas .....	116
Consideraciones finales.....	119
Referencias .....	121
Apéndice.....	127

# Presentación

EN ESTE APARTADO SE DESCRIBE EL propósito, la organización y la estructura de este manual, así como las convenciones de escritura utilizadas.

## Propósitos del manual

El propósito de este libro es servir de guía para llevar a cabo el mapeo de la biomasa aérea del bosque en grandes superficies, relacionando los datos de campo del Inventario Nacional Forestal u otras fuentes de datos con la información derivada de imágenes de satélite, particularmente de las imágenes del radar de apertura sintética ALOS PALSAR. Para ello, se utiliza el *software* libre R. La intención es que investigadores, estudiantes, técnicos forestales de organizaciones no gubernamentales y del gobierno, así como usuarios de la percepción remota en general, que tengan habilidades mínimas de programación para implementar los *scripts* para procesar la información de campo, analizar las imágenes de satélite y relacionar estos dos conjuntos de información, puedan obtener mapas con la distribución espacial de la biomasa aérea del bosque en sus áreas de interés.

El manual se enfoca en relacionar datos de biomasa del bosque estimados en campo, con la información de las imágenes de radar para producir modelos de estimación que permitan mapear la biomasa aérea del bosque. La metodología para el mapeo de la biomasa aérea del bosque utiliza *scripts* desarrollados en el lenguaje R, e incluye las siguientes actividades: primero se estima la biomasa en campo utilizando el diámetro, la altura y la densidad de la madera de los árboles con ecuaciones alométricas específicas; después, se ejecuta el preprocesamiento de las imágenes de radar y se obtienen las variables derivadas de estas imágenes que se asociarán con los datos de campo; en seguida, se ajustan modelos de predicción de la biomasa usando *Random Forest* y, por último, se aplican los modelos ajustados al conjunto de imágenes y se obtiene un mapa con la distribución espacial de la biomasa.

Este manual debe usarse en conjunto con los datos de campo del Inventario Nacional Forestal y de Suelos de un área ubicada en el municipio de Felipe Carrillo Puerto, en el estado de Quintana Roo, que se utiliza como ejemplo durante todo el

proceso. Además, se requieren los *scripts* empleados en las diferentes etapas de la estimación de la distribución espacial de la biomasa aérea del bosque. Ambos pueden ser descargados libremente como se describe en el Apéndice.

## Estructura del manual

Este manual consta de esta presentación, una introducción, siete capítulos, un apartado de referencias y un apéndice. Siguiendo a esta presentación, en la introducción se describe la importancia de los bosques como reservorios de carbono, se refieren algunos de los estudios más recientes sobre el mapeo de la distribución espacial de la biomasa aérea del bosque a nivel continental y nacional, las dificultades encontradas para el mapeo de la biomasa aérea, en particular la de los bosques tropicales, y las ventajas de utilizar imágenes de radar de apertura sintética, así como los procedimientos que podrían mejorar la precisión de la estimación de esta variable. El capítulo 1 describe de manera detallada los procedimientos para la estimación de la biomasa en campo, el procesamiento de las imágenes, el ajuste de modelos de predicción de la biomasa y el mapeo de la misma. En el capítulo 2 se expone de manera general la configuración del *software* R y las librerías utilizadas en los *scripts* desarrollados aquí. El capítulo 3 presenta los procedimientos para el cálculo de la biomasa en campo, utilizando datos dasométricos de árboles del Inventario Nacional Forestal y de Suelos y ecuaciones alométricas específicas. En los capítulos 4 y 5 se describen los procesos que deben realizarse con las imágenes de radar para obtener variables derivadas de dichas imágenes, que se asocien con la biomasa aérea. En el capítulo 6 se explican los procedimientos para crear una base de datos que contenga la biomasa estimada en campo y las variables derivadas de las imágenes de radar, la cual es utilizada en el capítulo 7 para ajustar modelos de predicción usando un método no paramétrico (*Random Forest*) y para mapear la distribución espacial de la biomasa aérea. Para finalizar se desarrolla un apartado con las conclusiones de este manual.

El apéndice de este manual describe la forma de descargar los datos de campo y los *scripts* utilizados para el procesamiento, así como la organización de los mismos.

## Convenciones de escritura

Para facilitar la lectura de este documento, el texto en general está escrito con un tipo de letra serif (Warnock Pro), las líneas de comando están escritas con una tipografía monospace (**Gemeli Mono**) en negritas y sobre un fondo gris. Los resultados arrojados por el *software* R tienen el mismo tipo de letra y tamaño que las líneas de comando (Gemeli Mono) en versión light, también sobre fondo gris. Por último, las gráficas y figuras que no son resultado de líneas de comando, están enumeradas y contienen títulos, mientras que el resto se presenta tal como se obtiene de R.

# Introducción

LOS BOSQUES TROPICALES SON RESERVORIOS DE carbono muy importantes que proporcionan bienes y servicios ecológicos a la sociedad. Estos bosques almacenan más carbono que cualquier otro bioma terrestre y ayudan a mitigar el cambio climático global (Houghton *et al.*, 2015; Pan *et al.*, 2011). Sin embargo, la pérdida de los bosques por intervenciones humanas, agravadas a veces por perturbaciones naturales, ha aumentado dramáticamente. La deforestación en los trópicos representa actualmente alrededor del 8% de las emisiones globales anuales de gases de efecto invernadero, aunque históricamente ha oscilado entre 15 y 25% (Fearnside y Laurance, 2004; Le Quéré *et al.*, 2013). La tala de bosques tropicales provoca, adicionalmente, disminuciones en la diversidad biológica, en las funciones ambientales y en los productos forestales (Fahrig, 2003). Por lo tanto, evaluar la biomasa aérea y los almacenes de carbono asociados a los bosques tropicales es fundamental para la mitigación del cambio climático y para mantener dicho almacén de carbono, así como para la conservación de la diversidad biológica.

Recientemente, se han publicado varios estudios que obtienen mapas de biomasa aérea o de densidad de carbono aéreo a escalas continental y nacional (Rodríguez-Veiga *et al.*, 2019; Urbazaev *et al.*, 2018; Rodríguez-Veiga *et al.*, 2016; Avitabile *et al.*, 2016; Cartus *et al.*, 2014; Bacchini *et al.*, 2012; Saatchi *et al.*, 2011). Sin embargo, dichos mapas a menudo presentan grandes incertidumbres, particularmente en áreas tropicales con una estructura vegetal compleja y altos valores de biomasa aérea (Rodríguez-Veiga *et al.*, 2019; Avitabile *et al.*, 2016). Por ejemplo, Rodríguez-Veiga *et al.* (2016) y Cartus *et al.* (2014) mapearon la biomasa aérea y la densidad de carbono respectivamente en México a nivel nacional. En estos estudios se reportan errores relativos más altos en los bosques tropicales en comparación con otros tipos de bosques en México.

Las imágenes de satélite utilizadas para mapear la distribución espacial de la biomasa aérea a escalas regionales o globales, utilizan resoluciones espaciales que varían desde una resolución alta, tales como Sentinel-2 y Spot, hasta aquellas que usan una resolución media, como Landsat y MODIS. Sin embargo, cada vez es más

frecuente utilizar datos de radar de apertura sintética para mapear la distribución espacial de la biomasa. Una de las principales ventajas de los sensores de radar, en comparación con los sensores ópticos, es su capacidad de penetrar la capa de nubes, lo cual es muy importante para las regiones tropicales, que tienen una capa de nubes persistente la mayor parte del año. Los datos de radar se pueden utilizar para predecir la biomasa aérea gracias a la facultad que tienen estos sensores de penetrar el dosel del bosque, permitiendo una mejor detección de tallos y ramas, que están asociados con la cantidad de biomasa aérea disponible en el bosque (Joshi *et al.*, 2017). El radar de la banda L es relativamente sensible a la biomasa forestal, ya que utiliza longitudes de onda que varían de 15 a 30 cm, lo que le permite una penetración profunda de la vegetación.

Sin embargo, el mapeo de la biomasa forestal utilizando datos de radar con la banda L tiene algunas limitaciones. Las relaciones entre la intensidad de la retrodispersión del radar y la biomasa aérea pueden saturarse en ecosistemas forestales densos y con estructura de vegetación compleja (Mermoz *et al.*, 2015), lo que también puede conducir a una subestimación de la biomasa aérea. Aunque el nivel de saturación se ve afectado por varios factores, como el tipo de vegetación, la complejidad del dosel y la topografía del sitio, entre otros, el punto de saturación generalmente está en alrededor de 150 Mg ha<sup>-1</sup> (Mermoz *et al.*, 2015; Morel *et al.*, 2011), aunque los valores de esta variable en algunos sitios de bosques tropicales pueden exceder incluso los 320 Mg ha<sup>-1</sup>. Para resolver el problema de saturación, se han aplicado varios métodos. Uno de ellos utiliza índices que combinan las polarizaciones SAR (HH y HV) para identificar la contribución del volumen de dispersión de diferentes polarizaciones (García *et al.*, 2018; Mitchard *et al.*, 2012; Almeida-Filho *et al.*, 2019); otro utiliza la textura de las imágenes de radar de la banda L, que han mostrado resultados satisfactorios al capturar variaciones en la estructura de la vegetación. Con una resolución espacial de 25 m, ALOS PALSAR puede capturar la variación de la estructura de la vegetación en áreas de bosque con diferentes edades de sucesión. Es decir, el sensor puede captar variaciones en parches de bosque de distintas edades, gracias a que tienen diferencias en la estructura de la vegetación, tales como la altura de los árboles, el tamaño de las copas y el número de tallos. Esta variación de la estructura del bosque está asociada con la cantidad de biomasa aérea (Huang *et al.*, 2019; Zhao *et al.*, 2016; Thapa *et al.*, 2015).

En este manual se explica la metodología a seguir para producir mapas con la distribución espacial de la biomasa aérea, la cual está basada en el uso de datos de campo relacionándolos con información de los datos de radar y sus texturas. Se incluye una guía paso a paso para el cálculo y la estimación de la biomasa aérea a par-

tir de parcelas de campo, así como el preprocesamiento y análisis de las imágenes de ALOS PALSAR. Esta guía incluye también la extracción de la información de las imágenes de radar y la creación de las bases de datos necesarias para construir los modelos de predicción de la biomasa usando Random Forest. Además, se presenta una guía técnica para la evaluación de los modelos construidos y el mapeo de la biomasa aérea. En este manual se desarrolla un ejemplo práctico usando datos de los bosques tropicales secos de la península de Yucatán; sin embargo, los métodos y programas de *software* que se explican son aplicables para mapear la biomasa aérea de cualquier tipo de vegetación usando cualquier otro tipo de imágenes de sensores remotos.

### **Datos de campo para la estimación de la biomasa**

El conocimiento de la distribución espacial de la biomasa aérea tiene como base la información de biomasa calculada a partir de datos de campo. Los datos de campo generalmente incluyen todos los diámetros de los árboles por encima de un umbral (que en el caso del Inventario Nacional Forestal y de Suelos de México es de 7.5 cm), así como las alturas de estos y la identidad de las especies. Este último dato es utilizado para inferir la densidad de la madera que se emplea en el cálculo de la biomasa usando ecuaciones alométricas. Los datos de campo para el cálculo de la biomasa pueden ser obtenidos del Inventario Nacional Forestal y de Suelos para cualquier parte de México.

El segundo Inventario Nacional Forestal y de Suelos (INFyS) se llevó a cabo desde 2009 hasta 2014. El diseño de muestreo de este inventario fue estratificado sistemático, es decir, se estableció una malla equidistante a lo largo y ancho del territorio nacional; sin embargo, la intensidad de muestreo varió entre tipos de vegetación (Conafor, 2013). En un primer grupo se incluyeron los bosques y las selvas con un distanciamiento de 5 x 5 km; en el segundo grupo se encuentran las zonas semiáridas y selvas bajas (10 x 10 km); por último, en el tercer grupo, las comunidades de zonas áridas (20 x 20 km). Las unidades de muestreo primarias (conglomerados) que se utilizaron para la información de campo están compuestas de cuatro parcelas o unidades de muestreo secundarias. Estas parcelas de 400 m<sup>2</sup> están colocadas en forma de “Y” invertida con respecto al norte y son de forma circular, representando de manera conceptual una unidad de muestreo de una hectárea (Conafor, 2013).

### **Datos de radar de apertura sintética**

Para mapear la distribución espacial de la biomasa aérea, se utilizaron datos de radar de la Agencia de Exploración Aeroespacial de Japón (JAXA), la cual proporciona

acceso libre al mosaico de imágenes ALOS PALSAR/ PALSAR -2. PALSAR (Phased Array type L-band Synthetic Aperture Radar, por sus siglas en inglés), el cual es obtenido por el satélite ALOS (Advanced Land Observing Satellite). El sensor PALSAR /PALSAR-2 es capaz de generar datos sobre la superficie de la Tierra durante el día y la noche sin importar las condiciones climáticas.

Los datos de ALOS PALSAR utilizados en este manual son del mosaico global PALSAR / PALSAR-2. El mosaico global fue creado por la JAXA sin fisuras en unidades de 10 x 10 grados de latitud y longitud. Las imágenes de este mosaico fueron previamente corregidas por las distorsiones geométricas de los sensores de radar de apertura sintética mediante una ortorrectificación. Además, se corrigieron los efectos topográficos sobre la imagen, usando una corrección de la pendiente. Estas correcciones fueron realizadas por la JAXA. El tamaño de un píxel es de 25 m<sup>2</sup>, mientras que el intervalo de tiempo de las imágenes en el mosaico es de un año. En el mosaico de imágenes se incluyen capas de datos con la siguiente información: una con la polarización HH y otra con la HV, mapa forestal / no forestal, fecha de observación y ángulo de incidencia local.



# 1. Estimación de la biomasa aérea del bosque

## 1.1 La biomasa del bosque

LA BIOMASA LIGNOCELULÓSICA, O BIOMASA DE las plantas leñosas, se refiere al tejido vegetal seco de las hojas, frutos, ramas, troncos y raíces de estas plantas. Se expresa generalmente en kilogramos o toneladas de materia seca. De acuerdo con su posición y características para su medición, se reconocen tres tipos de reservorios: biomasa (aérea y subterránea), materia orgánica muerta (madera muerta y mantillo) y suelo mineral.

La biomasa subterránea es uno de los reservorios más difíciles de medir, ya que implica conocer las dimensiones de la estructura radicular. En el caso de los árboles muy grandes, dicha estructura puede encontrarse a grandes profundidades. Solo en contados estudios se ha realizado una extracción y medición de las raíces de algunos árboles (Sorgona *et al.*, 2018), estos han sido utilizados para derivar relaciones cuantitativas entre la biomasa de las raíces y la de la parte aérea. En uno de los trabajos de este tipo realizado en la península de Yucatán, Cairns *et al.* (1997) destacan que la biomasa de las raíces puede representar entre el 20 y 30 % de la biomasa aérea.

La biomasa del material caído (a veces también referida como necromasa), usualmente solo representa una fracción menor de la biomasa total del sistema; sin embargo, en ocasiones su contribución puede ser significativamente alta; por ejemplo, en bosques templados bajo manejo forestal, la madera muerta puede representar hasta el 60 % de la biomasa del sistema (Galicia *et al.*, 2015). Evaluaciones realizadas después del paso del huracán Dean por la península de Yucatán muestran que, en algunas de las áreas más afectadas, hasta el 40 % de los árboles en pie fueron derribados por los vientos (Islebe *et al.*, 2009). Los métodos que se emplean para calcular la cantidad de material caído dependerán del tamaño de los fragmentos: en los más grandes puede ser más efectiva la medición de sus diámetros y longitudes, mientras que en los fragmentos de tamaño intermedio se pueden utilizar conteos a lo largo de transectos; por último, en los más pequeños, el secado y el pesado del material resulta una opción muy eficiente (Galicia *et al.*, 2015). La estimación de la cantidad

de biomasa en este reservorio es muy importante para calcular la acumulación de material combustible y evaluar los riesgos de ocurrencia de incendios forestales.

Por otra parte, la biomasa aérea es uno de los reservorios más estudiados. En los bosques tropicales, típicamente los árboles son los que aportan los mayores contenidos de biomasa, mientras que otras formas de vida, tales como las lianas o las palmas, solo contribuyen marginalmente a la biomasa total. Por ejemplo, Gehring *et al.* (2004) encontraron que las lianas aportaron entre 6.2 y 1.8 % de la biomasa aérea de bosques secundarios de la Amazonía; en las etapas sucesionales tempranas, la contribución de las lianas fue mayor y disminuyó con la edad. Un patrón similar se presenta con las palmas. En un estudio realizado en la península de Yucatán, estas aportaron entre 5 y 2 % de la biomasa de la vegetación secundaria en etapas sucesionales tempranas, y menos del 1 % en los bosques maduros (Read y Lawrence, 2003).

## 1.2 Estimación de la distribución espacial de la biomasa aérea

La estimación de la distribución espacial de la biomasa aérea del bosque es importante para cuantificar el papel que tienen los bosques como almacenes de carbono. Esta información es crucial para apoyar la conservación y el manejo de los bosques y mantener las reservas de carbono, debido a que el conocimiento sobre las existencias y flujos de carbono es esencial para comprender la condición actual y el curso que tendrá el ciclo del carbono ante los cambios en el uso del suelo y otros disturbios (Thomson *et al.*, 2010).

Los métodos para estimar la biomasa aérea de los bosques están basados en datos de campo; sin embargo, la estimación de la biomasa usando exclusivamente estos datos tiene dificultades logísticas, involucra costos altos y no describe la distribución de esta variable en el espacio. La percepción remota ofrece una fuente viable de información auxiliar para evaluar la biomasa aérea de los bosques, ya que brinda un medio económico para lograr una cobertura espacial completa de información en áreas grandes y con intervalos de tiempo regulares (Lu *et al.*, 2016). De igual manera, pueden utilizarse variables ambientales distribuidas de manera continua en el espacio para la estimación de la biomasa. Varios estudios han encontrado que la biomasa se ve afectada por variables climáticas; como ejemplo se ha demostrado que la variación de la temperatura y la precipitación influye en la biomasa (Keith *et al.*, 2009). Por lo tanto, se han utilizado variables climáticas (Hernández-Stefanoni *et al.*, 2020) además de variables topográficas (Cartus *et al.*, 2014) en conjunto con datos de imágenes de satélite para el mapeo de la biomasa aérea de los bosques.

Un método común para estimar la biomasa aérea es construir mapas de clases de la cobertura del suelo o tipos de vegetación a partir de imágenes de satélite. Estos

mapas se utilizan para predecir la biomasa aérea en función de los valores medios de esta variable dentro de cada tipo de vegetación. A pesar de su aplicabilidad general y la utilidad de este enfoque, se debe reconocer que tiene una simplificación excesiva, derivada de tener un valor medio único que predice todos los puntos no medidos dentro de cada tipo de vegetación, es decir, se ignora la variabilidad dentro de cada tipo de vegetación (Burrough, 2001). Un enfoque alternativo que aborda estas deficiencias es el uso de datos de imágenes de satélite para medir parámetros o extraer variables de percepción remota y construir modelos que relacionen la biomasa aérea estimada en campo con las variables extraídas de las imágenes, lo cual permite desarrollar mapas continuos de esta variable (Cartus *et al.*, 2014; Hernández-Stefanoni *et al.*, 2020; Rodríguez-Veiga *et al.*, 2019).

La estimación de la distribución espacial de la biomasa aérea del bosque incluye varias etapas, las cuales pueden observarse en la figura 1. En la primera etapa se realiza la estimación de la biomasa aérea en campo. La biomasa del bosque no se mide en campo, más bien, se estima, por lo que se requiere recolectar información de un número suficiente de parcelas, para medir el diámetro y la altura, así como identificar a nivel de especie cada uno de los árboles dentro de las parcelas. Adicionalmente, a la información se agrega la densidad de la madera para cada especie. Posteriormente, con estos datos y el uso de ecuaciones alométricas se estima la biomasa de cada árbol. Finalmente, la suma de la biomasa de todos los individuos dentro de una parcela brinda una estimación de la biomasa aérea por unidad de superficie de la parcela.

Las ecuaciones alométricas para estimar la biomasa a nivel de árbol se obtienen mediante muestreos destructivos y el ajuste de modelos de regresión lineales o no lineales. Estas ecuaciones estiman la biomasa total para árboles individuales y combinan el diámetro a la altura del pecho y la altura; además, en algunos se incluye la densidad de la madera como variable predictora (Urquiza-Haas *et al.*, 2007; Chave *et al.*, 2014; Ramírez *et al.*, 2019). Sin embargo, las relaciones entre la altura y el diámetro de los árboles varían considerablemente entre y dentro de las especies, así como en relación con el clima y la estructura de la vegetación (Banin *et al.*, 2012; Hulshof *et al.*, 2015). Por otro lado, la densidad de la madera rara vez se mide dentro de los inventarios forestales, a pesar de que la biomasa de cada árbol es directamente proporcional a la densidad de la madera, la cual, además, presenta variaciones entre especies de árboles y entre individuos de la misma especie (Bastin *et al.*, 2015). Por lo tanto, la precisión en la estimación de la biomasa de árboles individuales depende de las ecuaciones alométricas utilizadas (Molto *et al.*, 2013).

La segunda etapa para la estimación de la distribución espacial de la biomasa tiene que ver con el procesamiento de las imágenes de satélite. Aquí el objetivo es

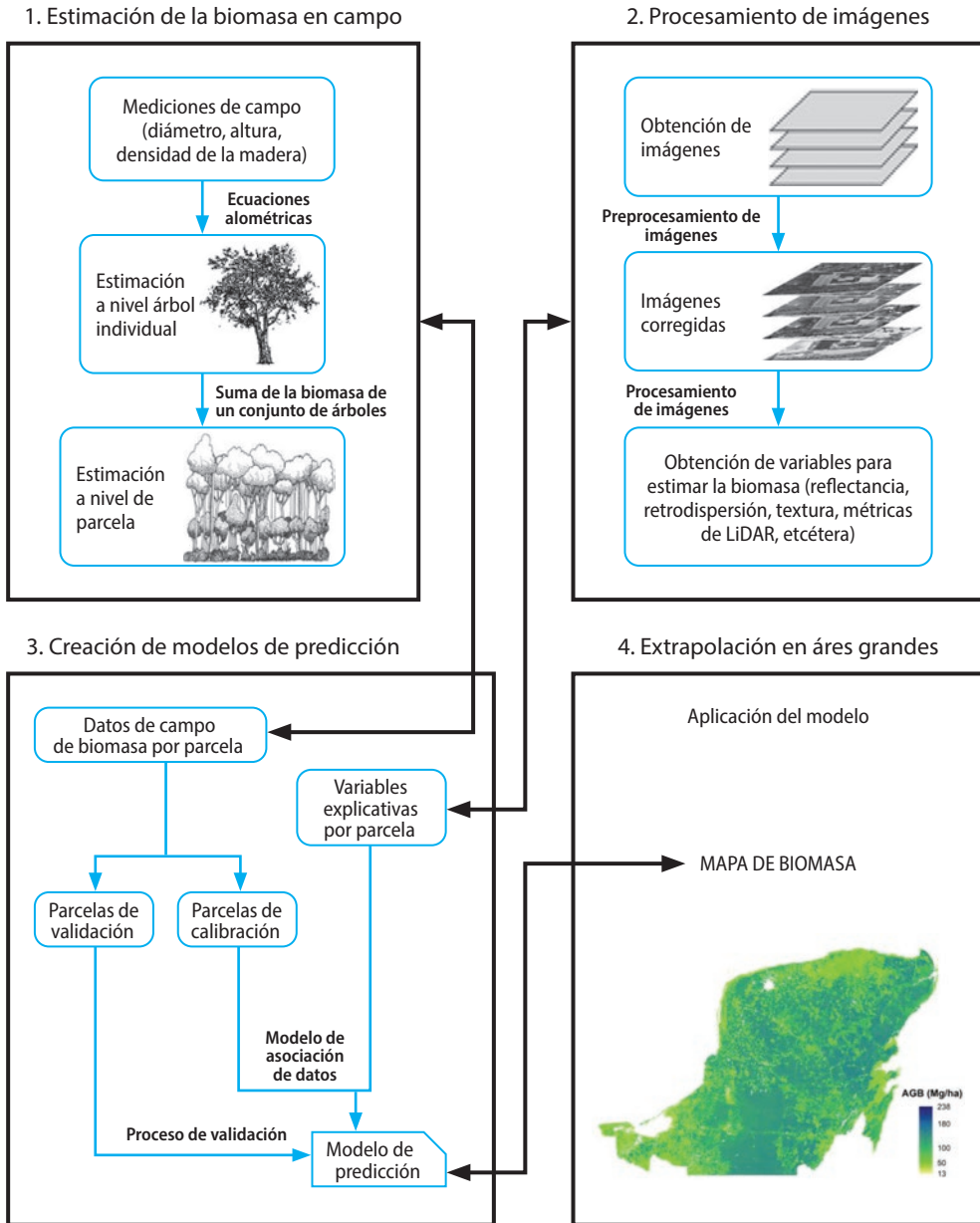
generar variables derivadas de los datos de percepción remota, que se relacionen con la biomasa aérea del bosque y que puedan ser utilizadas en la construcción de modelos para predecir la biomasa.

Los valores de reflectancia y los índices de vegetación derivados de imágenes ópticas de satélites han sido utilizados para estimar la biomasa en los bosques (Foody *et al.*, 2003). Diversos estudios muestran que se pueden obtener estimaciones de la biomasa, tanto para bosques templados, como tropicales, a escalas espaciales locales, regionales y globales (Saatchi *et al.*, 2011; Cartus *et al.*, 2014; Rodríguez-Veiga *et al.*, 2019; Hernández-Stefanoni *et al.*, 2020). Sin embargo, una limitación importante de las imágenes ópticas es que los índices de vegetación como el NDVI (“Normalized Difference Vegetation Index”) tienden a saturarse en áreas con vegetación densa (Lu *et al.*, 2012). Una manera de solucionar este problema es a través del uso de imágenes de alta resolución (resolución <10 m, tales como RapidEye, Quickbird e Iconos, entre otras). Este tipo de imágenes ofrece información más detallada de la estructura de la vegetación en comparación con imágenes de resolución media y alta (resolución >10 m como Spot, Sentinel 2, Landsat, etc.) y, a su vez, permite discriminar entre el dosel del bosque y las aperturas en este, así como distinguir el espacio entre árboles y otros atributos de la vegetación, lo que posibilita el uso de texturas de las imágenes para la estimación de la biomasa (Ploton *et al.*, 2017).

Los datos del radar de apertura sintética (SAR, por sus siglas en inglés) también se han utilizado con éxito para mapear la biomasa (Hernández-Stefanoni *et al.*, 2020). Los instrumentos SAR de longitud de onda moderada de 15 a 30 cm pueden penetrar el dosel del bosque interactuando con tallos y ramas, donde se almacena la mayor parte de la biomasa (Joshi *et al.*, 2017). Por lo tanto, la intensidad de la señal de retrodispersión del radar se relaciona con la biomasa del bosque; sin embargo, las relaciones entre la intensidad de la retrodispersión del radar y la biomasa también pueden saturarse (Joshi *et al.*, 2017). Por otro lado, los bosques exhiben una heterogeneidad a escala de paisaje, provocada por los cambios en el uso de la tierra, lo cual genera un mosaico de parches de bosques de diferentes edades. Estas variaciones en la estructura de la vegetación pueden caracterizarse usando la textura de las imágenes SAR de resolución media (resolución de 25 m) y con ello superar o minimizar los problemas de saturación (Huang *et al.*, 2019; Zhao *et al.*, 2016; Thapa *et al.*, 2015).

Por último, un enfoque muy eficaz para estimar la biomasa del bosque es mediante el uso de datos LiDAR (Light Detection and Ranging). LiDAR es un sensor activo que puede adquirir directamente la estructura vertical de la vegetación, debido a que tiene la capacidad de penetrar el dosel del bosque (Lefsky *et al.*, 2001). En consecuencia, LiDAR puede proporcionar una vista tridimensional de la estructura

**Figura 1.** Etapas para la estimación de la distribución espacial de la biomasa aérea de los bosques usando percepción remota



del bosque en una nube de puntos, lo que permite estimar la biomasa con mayor precisión que las imágenes ópticas (Lefsky *et al.*, 2001; Hernández-Stefanoni *et al.*, 2014). LiDAR permite obtener estimaciones muy precisas de diferentes atributos de la vegetación (altura, área basal y biomasa) usando métricas de altura y densidad de puntos (Hernández-Stefanoni *et al.*, 2014).

La tercera etapa tiene que ver con la construcción de modelos de predicción de la biomasa. Estos modelos se desarrollan pensando en relacionar la variable dependiente (es decir, la biomasa aérea) con una serie de variables independientes (derivadas de datos de las imágenes de satélite). Para ello, se pueden utilizar modelos de regresión lineal simple o múltiple; sin embargo, la biomasa suele estar relacionada de forma no lineal con las variables derivadas de las imágenes de satélite, de modo que también se han utilizado modelos no lineales. Por otra parte, como las relaciones entre la biomasa y las variables independientes son a menudo muy complejas, se han utilizado asimismo algoritmos no paramétricos para estimar la biomasa, entre los que se incluyen el vecino más cercano (K-NN), redes neuronales artificiales (ANN), bosque aleatorio (Random Forest), máquinas de soporte de vectores (SVM), entre otros. Un aspecto importante durante la construcción de los modelos de predicción de la biomasa consiste en evaluar la confiabilidad de los modelos y la precisión de las estimaciones de la biomasa. Para ello se requiere una evaluación de la precisión de la predicción de la biomasa a través de un proceso de validación.

En la última etapa, se utilizan los modelos de predicción construidos junto con las variables predictoras derivadas de las imágenes de satélite, para obtener la distribución espacial de la biomasa aérea estimada en el área de estudio. Los mapas con la distribución espacial de la biomasa se construyen al aplicar el modelo de predicción al conjunto de capas que representan las variables de las imágenes.

## 2. Configuración del entorno del *software*

PARA EL MAPEO DE LA BIOMASA aérea del bosque se requieren herramientas de cálculo, procesamiento, creación de modelos y mapeo de variables con *software* de código abierto. Las instrucciones de este capítulo guiarán al usuario a través de la instalación y configuración del *software* que se utilizará para los procedimientos de mapeo de la biomasa en una plataforma Microsoft Windows. Sin embargo, pueden encontrarse en la red instrucciones para otras plataformas (por ejemplo, Linux o MacOS).

Los programas que aquí se presentan fueron desarrollados en el lenguaje R, que es una plataforma para el análisis estadístico de datos tales como modelos lineales y no lineales, pruebas estadísticas, series de tiempo, clasificaciones, entre muchos otros. Además, contiene herramientas gráficas avanzadas y de gestión de datos. Por otro lado, miles de personas en el mundo colaboran para su desarrollo, lo cual hace que esta herramienta continuamente sea extendida con nuevas aplicaciones. Y no menos importante, R se puede obtener y distribuir de manera gratuita bajo una licencia pública general.

### 2.1 Descargar e instalar R

Los archivos e instrucciones para la instalación de R se pueden descargar desde la página web del Comprehensive R Archive Network (CRAN).

- a. Ir al enlace <https://cloud.r-project.org/index.html> para descargar e instalar R.
- b. Elegir un archivo de instalación para su plataforma.

### 2.2 Descargar e instalar RStudio

Debido a que R no tiene una interfaz gráfica para el usuario que facilite su utilización para los principiantes, en este manual se utilizó RStudio, en el cual algunas funciones de R son más fáciles de utilizar. RStudio incluye adicionalmente un editor de código

go, así como herramientas de depuración y de visualización. Para instalar RStudio se deben realizar las siguientes actividades:

- a. Ir al enlace <https://www.rstudio.com/products/rstudio/download/> para descargar e instalar la edición de código abierto de RStudio.
- b. En la página de descarga, seleccionar la opción RStudio Desktop, Open Source License.
- c. Elegir un archivo de instalación para su plataforma.
- d. Adicionalmente se debe instalar la versión más reciente de las herramientas de R, las cuales pueden obtenerse en <https://cran.rstudio.com/bin/windows/Rtools/>

### 2.3 Cómo iniciar con R y RStudio

- a. Manuales de R: <http://cran.r-project.org/manuals.html>
- b. Documentación de R: <http://cran.r-project.org/other-docs.html>
- c. Una guía rápida en R, Quick-R: <http://www.statmethods.net/index.html>
- d. Una guía rápida de RStudio, RStudio Tutorial – A Complete Guide for Novice Learners!: <https://data-flair.training/blogs/rstudio-tutorial/>
- e. Una guía rápida de R y RStudio en español se encuentra en el libro de Mas (2018), el cual puede ser descargado de <https://eujournal.org/files/journals/1/books/JeanFrancoisMas.pdf>

### 2.4 Paquetes de R

Al instalar R se obtiene una configuración del sistema básico, que incluye funcionalidades mínimas. Para poder extender las funcionalidades de R, es necesario instalar paquetes adicionales (*packages*).

Existen alrededor de 12 000 paquetes de R disponibles para su descarga. Información sobre los paquetes disponibles puede obtenerse directamente en CRAN, con la función `available.packages()`. Dicha función devuelve una matriz con información detallada sobre los paquetes disponibles actualmente en uno o más de los repositorios de R. Para instalar un paquete de R es necesario tener una conexión a internet, conectarse con un repositorio y utilizar la función `install.packages("nombre del paquete")`. Por ejemplo, `install.packages("BIOMASS")` instalará el paquete llamado **BIOMASS**.

Los paquetes de R son colecciones de funciones, datos, documentación y código compilado que, además, son fáciles de compartir con otros usuarios. A con-



tinuación, se presentan los paquetes utilizados en este manual que se relacionan directamente con el mapeo de la biomasa aérea de bosques usando datos de campo e imágenes de radar.

**openxlsx.** Este paquete se utiliza para la creación de archivos Excel (\*.xlsx), proporcionando una interfaz de alto nivel para escribir, diseñar y editar hojas de trabajo. En el Inventario Nacional Forestal existen diferentes variables de cadenas de texto que pueden contener comas, lo que dificulta la lectura de las variables en cuestión cuando se usan archivos de Excel separados por comas (\*.CSV), para evitar esto fue que se optó por usar el paquete **openxlsx**.

**BIOMASS.** Este paquete contiene funciones para estimar la biomasa y su incertidumbre en los bosques tropicales. Las funciones permiten: 1) recuperar y corregir la taxonomía de las especies de plantas; 2) estimar la densidad de la madera y su incertidumbre; 3) construir modelos diámetro-altura; 4) modificar datos individuales, y 5) estimar la biomasa aérea y la incertidumbre asociada a esta estimación (Réjou-Méchain *et al.*, 2017).

**rgdal.** Este paquete proporciona enlaces a la Biblioteca de datos ‘geoespaciales’ (‘GDAL’) y acceso a operaciones de proyección / transformación desde la biblioteca ‘PROJ’.<sup>4</sup>

**sp.** Este paquete contiene funciones para datos espaciales, tales como trazar mapas con datos, métodos para recuperación de coordenadas, para subconjuntos, impresión, resumen, etc.

**raster.** Este paquete presenta funciones para la lectura, la escritura, la manipulación, el análisis y el modelado de datos espaciales en forma de matriz. El paquete implementa funciones básicas y de alto nivel para el procesamiento de archivos raster.

**sf.** Este paquete tiene funciones simples para soporte de otros paquetes, entre ellos están una forma estandarizada de codificar datos de vectores espaciales. Junto con el paquete ‘GDAL’, se usa para leer y escribir datos, y con ‘GEOS’, para operaciones geométricas, así como con ‘PROJ’, para conversiones de proyección y transformaciones de datos.

**glcm.** Este paquete calcula diferentes medidas de textura de las imágenes usando matrices de coocurrencia (Haralick *et al.*, 1973).

**randomForest.** Este paquete implementa el algoritmo *Random Forest* para clasificación y regresión.

**ModelMap.** Este paquete tiene funciones para construir modelos predictivos de respuestas continuas o discretas, utilizando *random forest*. Valida estos modelos con un conjunto de datos independientes o con validación cruzada, crea gráficos y tablas del modelo, realiza un diagnóstico de validación y aplica estos modelos a archivos de imágenes predictoras para crear superficies continuas o mapas (Freeman *et al.*, 2009).

**prettymapr.** Este paquete automatiza el proceso para la creación de establecer márgenes, trazar el mapa, la barra de escala, la flecha norte y restablecer los parámetros gráficos una vez finalizado.

Los paquetes usados en este manual pueden ser instalados usando la función `install.packages()`, como se indica al inicio de cada capítulo, o bien, usando el código para la instalación siguiente:

```
# Instalación de los paquetes que se requieren para
# mapear la biomasa aérea de bosques en este manual
```

```
install.packages("openxlsx")
install.packages("BIOMASS")
install.packages("knitr")
install.packages("httr")
install.packages("dplyr")
if (!require(devtools)) install.packages('devtools')
devtools::install_github("giswqs/whiteboxR", force = TRUE)
whitebox::wbt_init()
install.packages("rgdal") #versión v1.4-4
install.packages("spatial.tools") #versión 1.6.0
install.packages("sp")
install.packages("glcm")
install.packages("rgeos")
install.packages("sf")
install.packages("randomForest")
install.packages("ModelMap")
install.packages("prettymapr")
```

En el caso del paquete **raster** es necesario instalar la versión 2.9-23, o bien, usando el archivo **"raster\_2.9-23.tar"**.

```
# Instalar el paquete raster con el archivo "raster_2.9-23.tar"
```

```
install.packages(choose.files(), repos = NULL, type = "source")
```



### 3. Cálculo de la biomasa aérea en campo

EN ESTE CAPÍTULO SE DESCRIBE EL procedimiento para calcular la biomasa aérea del bosque para cada conglomerado del Inventario Nacional Forestal y de Suelos; sin embargo, los programas desarrollados pueden ser utilizados para cualquier inventario con parcelas de campo con cualquier diseño. Para el cálculo de la biomasa aérea de cada árbol se utilizan ecuaciones alométricas, las cuales toman en cuenta la densidad de la madera por especie. Por lo tanto, inicialmente se revisa la taxonomía de las especies y se corrigen los nombres de especies en caso de errores. Con la lista de especies depurada es posible extraer valores de la densidad de la madera de una base de datos global, e incorporar datos de densidades de madera locales. Por último, se obtiene un resumen de la biomasa aérea del bosque para cada conglomerado en unidades de toneladas por hectárea.

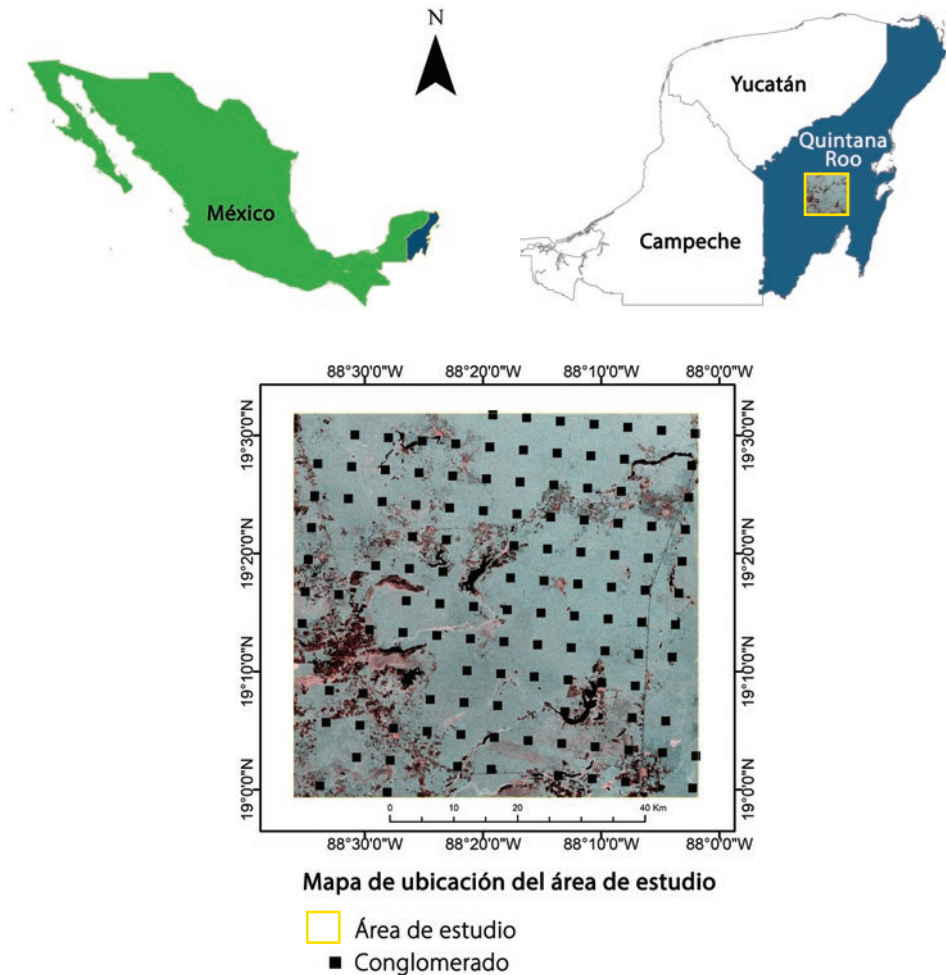
#### 3.1 Área de estudio y unidades de observación

El área de estudio donde se tomaron los datos de campo para mapear la biomasa aérea del bosque, comprende un paisaje de 3600 km<sup>2</sup> y se ubica en el municipio de Felipe Carrillo Puerto, en el estado de Quintana Roo. El área está dominada por selvas medianas subperennifolias, además de contar con áreas de selvas bajas inundables. Aproximadamente entre 25 y 30 % de las especies pierden sus hojas en la temporada de sequía. El clima del lugar es cálido subhúmedo, con un período de sequía que abarca los meses de febrero a mayo; tiene una temperatura media anual de 26 °C y una precipitación anual de entre 1100 y 1400 mm; la topográfica es plana (Flores y Espejel, 1994). El uso dominante de la tierra después del forestal son los pastizales para la cría de ganado, aunque también se practica la agricultura tradicional de roza, tumba y quema, así como la extracción de madera. Todos estos aprovechamientos han producido un mosaico de selvas en diferentes etapas de sucesión, junto con áreas abiertas para actividades agropecuarias (Hernández-Stefanoni *et al.*, 2018). El dosel de este bosque tiene de 15 a 25 m de altura aproximadamente y las comunidades vegetales son estructuralmente complejas con dos o tres estratos verticales. Dentro de las especies más abundantes, se encuentran: *Manilkara zapota* (L.) P.

Royen, *Piscidia piscipula* (L.) Sarg., *Bursera simaruba* (L.) Sarg., *Guettarda combsii* Urb. y *Lonchocarpus rugosus* Benth., entre otras (Hernández-Stefanoni *et al.*, 2018).

Se obtuvieron todos los conglomerados medidos entre 2009 y 2014 por el Inventario Nacional Forestal y de Suelos (INFyS) que cayeron dentro de los 3 600 km<sup>2</sup> que comprende el área de estudio (fig. 2). El diseño de las unidades de observación o conglomerados consiste en cuatro parcelas circulares de 400 m<sup>2</sup>, dispuestas en forma de Y invertida y cubren una superficie de una hectárea (Conafor, 2013).

**Figura 2.** Ubicación del área de estudio con las unidades de muestreo



### 3.2 Creación de la base de datos

Para facilitar los procesos que se ejecutan con R es recomendable definir un espacio de trabajo específico, esto es, la ruta de una carpeta para leer y escribir archivos por default. Para definir este espacio, se utiliza la función `setwd()`, y para saber cuál es el espacio de trabajo se utiliza la función `getwd()`. En este manual, la ruta de trabajo para el cálculo de la biomasa se definió como “`c:/Mapeo_biomasa/FCP`”.

```
# Establecer directorio de trabajo
```

```
setwd("c:/Mapeo_biomasa/FCP")
getwd()
# [1] "c:/Mapeo_biomasa/FCP"
```

Una vez definido el espacio de trabajo, el archivo con los datos de los conglomerados que caen dentro del área de estudio es `CONGLOMERADOS_FCP_INFYS.xlsx`, el cual se encuentra en la carpeta `/Mapeo_biomasa/FCP`. Este archivo Excel contiene la información para cada uno de los árboles medidos durante el muestreo en campo (diámetro, altura, identificación a nivel de especie, familia, etc.), así como información de la unidad de muestreo (identificación, localización, etc.), el estado de la república donde se levantó el inventario, el tipo de vegetación, entre otros. Para cargar este archivo se utiliza el paquete `openxlsx` y la función `read.xlsx(file.choose())`. Al ejecutar función `file.choose()` aparece una ventana en donde se debe especificar la carpeta donde están guardados los datos y el nombre del archivo a leer.

```
# Abrir los paquetes, instalarlos si es necesario con
# install.packages("openxlsx")
```

```
library("openxlsx")
library(dplyr)
# Attaching package: 'dplyr'
# The following objects are masked from 'package:stats':
#   filter, lag
# The following objects are masked from 'package:base':
#   intersect, setdiff, setequal, union
data_all <- read.xlsx(file.choose())
```

Durante este proceso, el archivo leído es asignado a un objeto en formato tabular (*dataframe*) nombrado `data_all`. En la parte superior derecha de la figura 3 se

Figura 3. Lectura del archivo con la base de datos

The screenshot shows the RStudio interface. The main window displays a data table with the following columns: Conglomerado, cgl\_sit\_arb, Año, Estado, Municipio, and Format. The data rows are as follows:

Conglomerado	cgl_sit_arb	Año	Estado	Municipio	Format	
1	58680	58680_4_24	2009	Quintana Roo	Felipe Carrillo Puerto	Selva
2	59191	59191_1_8	2010	Quintana Roo	Felipe Carrillo Puerto	Selva
3	59191	59191_2_13	2010	Quintana Roo	Felipe Carrillo Puerto	Selva
4	59191	59191_2_26	2010	Quintana Roo	Felipe Carrillo Puerto	Selva
5	58685	58685_2_6	2009	Quintana Roo	Felipe Carrillo Puerto	Selva
6	58686	58686_3_2	2010	Quintana Roo	Felipe Carrillo Puerto	Selva
7	58686	58686_3_29	2010	Quintana Roo	Felipe Carrillo Puerto	Selva
8	58686	58686_4_2	2010	Quintana Roo	Felipe Carrillo Puerto	Selva
9	58686	58686_4_8	2010	Quintana Roo	Felipe Carrillo Puerto	Selva

Below the table, the console window shows the following R code and output:

```

C:/Mapaeo_biomasa/FCP/
filter, lag
The following objects are masked from 'package:base':
  intersect, setdiff, setequal, union
> data_all <- read.xlsx(file.choose())
> view(data_all)
>
Showing 1 to 10 of 19,008 entries, 51 total columns
    
```



observa que `data_all` tiene 19 008 obs. (registros) y 51 variables. Al dar doble clic en `data_all` aparecerá en la ventana superior derecha el archivo leído con las variables y los datos de cada registro como se observa en la parte izquierda de la figura 3.

La función `names(data_all)` muestra todas las variables que contiene este *dataframe*, es decir, las columnas de la base de datos de campo.

```
names(data_all)
# [1] "Conglomerado"          "cgl_sit_arb"          "Anio"
# [4] "Estado"                "Municipio"           "Formato"
# [7] "Cve_veg_SV"           "Tipo_veg_SV"         "Veg_prim_levantada"
# [10] "Veg_sec_levantada"    "Sitio"                "Registro"
# [13] "Arbol"                 "NumeroTallos"        "FormaFuste"
# [16] "TipoTocon"            "Familia_APG"         "NombreCientifico_APG"
# [19] "Condicion"            "NomComun"            "X"
# [22] "Y"                     "Altitud"              "Vigor"
# [25] "VigEtapa"             "PosicionCopa"         "porc_CopaViva"
# [28] "ExposicionLuz"        "DensidadCopa"         "TransparenciaCopa"
# [31] "MuerteRegresiva"     "DiametroBasal"       "Azimut"
# [34] "distancia"            "Edad"                 "GrosorCorteza"
# [37] "Long10Anillos"       "NumAnillos25"        "Troza1"
# [40] "Troza2"               "Troza3"               "Troza4"
# [43] "Troza5"               "Troza6"               "Troza7"
# [46] "Troza8"               "AlturaFusteLimpio"   "AlturaComercial"
# [49] "Forma_Biologica_1"   "Altura_total"        "Diametro_normal"
```

Tanto el archivo de Excel con la base de datos (`CONGLOMERADOS_FCP_INFYS.xlsx`), como el *dataframe* (`data_all`) contienen una lista de árboles con alturas y diámetros, así como su identificación a nivel de especie. Para el mapeo de la biomasa aérea solamente se utilizaron aquellos conglomerados que contenían las cuatro parcelas de 400 m<sup>2</sup>. Para ello se creó una lista única de los diferentes conglomerados y se imprimó con el nombre de la lista `"list"`. En esta lista existen 125 conglomerados, de los cuales aquí se muestran los tres primeros y los tres últimos.

```
# Crear lista de conglomerados
```

```
list<- data.frame(unique(data_all$Conglomerado))
list

#   unique.data_all.Conglomerado.
# 1                               58680
# 2                               59191
# 3                               58685
# .
# .
# .
#
# 12359703
# 12458681
# 12557168
```

Posteriormente, se identificaron en la lista los conglomerados que cumplían con la condición de tener las cuatro parcelas de 400 m<sup>2</sup>. Aquí se muestra el primero que cumple con la condición, junto con los últimos cinco. Adicionalmente, se identificaron los registros de la base de datos que cumplían con la condición de tener las cuatro parcelas, registrándose dicha condición en la variable **FLsit**. Ahora **data\_all** contiene esa nueva variable, por lo que **data\_all** tiene 19 008 observaciones y 52 variables.

```
# Identificar los conglomerados que contengan cuatro parcelas
```

```
for (i in 1:length(list[,1])){
  list[i,2]<-length(unique((filter(data_all,
    Conglomerado == list[i,1]))$sitio))=4
}
list

#   unique.data_all.Conglomerado.   V2
# 1                               58680   TRUE
# .
# .
# .
```

```
#
# 121 57171      TRUE
# 122 59186      FALSE
# 123 59703      FALSE
# 124 58681      FALSE
# 125 57168      FALSE
```

```
# Identificar en la base de datos los registros que no cumplen la condición
```

```
data_all$FLsit <- NA_character_
for(row in seq_len(nrow(list))){
  data_all$FLsit[substr(
    data_all$Conglomerado,0,nchar(
      list$unique.data_all.Conglomerado.[row] ) ) == list$unique.data_all.Conglomerado.[row] ] <- list$V2[row]
}
```

En seguida se crearon dos *dataframes*, el primero (**data.list**) contiene todos los registros u observaciones que cumplen con la condición, con 18 758 observaciones y 52 variables, es decir, 250 registros no cumplieron con la condición de tener registros de árboles en las cuatro parcelas. El segundo *dataframe* elimina todas las variables que no son útiles para el cálculo de la biomasa. Este *dataframe* llamado “**data**” contiene 18 758 observaciones y 12 variables.

```
# Crear el dataframe con los registros que cumplen la condición de cuatro sitios
```

```
data.list<- data_all[which(data_all$FLsit == TRUE),]
```

```
# Crear un nuevo dataframe con las columnas útiles en el cálculo de la biomasa
```

```
data<- data.list[ , c(1,3,8,9,11,13,18,21,22,49,50,51)]
```

Para poder revisar la taxonomía de las especies en la base de datos y utilizar esta información para extraer valores de densidad de la madera, es necesario dividir la variable “**NombreCientifico\_APG**” en dos variables “**GENERO**” y “**ESPECIE**”. Al hacerlo “**data**” contiene 18 758 observaciones y 14 variables.

```
# Dividir nombre científico en género y especie
# y agregarlo al dataframe en nuevas columnas
```

```
for (i in 1:(length(data$Conglomerado))){
  cad <- unlist( strsplit(data$NombreCientifico_APG[i], " "))
  data$GENERO[i] <- cad[1]
  data$ESPECIE[i] <- cad[2]
}
```

Una vez creada la base de datos se puede explorar los datos utilizando las funciones `names()` y `head()` como se muestra a continuación.

# Explorar los datos

`names(data)`

```
# [1] "Conglomerado"      "Anio"              "Tipo_veg_SV"
# [4] "Veg_prim_levantada" "Sitio"             "Arbol"
# [7] "NombreCientifico_APG" "X"                 "Y"
# [10] "Forma_Biologica_1" "Altura_total"     "Diametro_normal"
# [13] "GENERO"           "ESPECIE"
```

`head(data)`

#	Conglomerado	Anio	Tipo_veg_SV	Veg_prim_levantada
#1	58680	2009	SELVA BAJA ESPINOSA SUBPERENNIFOLIA	SMQ
#2	59191	2010	SELVA MEDIANA SUBPERENNIFOLIA	SMQ
#3	59191	2010	SELVA MEDIANA SUBPERENNIFOLIA	SMQ
#4	59191	2010	SELVA MEDIANA SUBPERENNIFOLIA	SMQ
#5	58685	2009	SELVA MEDIANA SUBPERENNIFOLIA	SMQ
#6	58686	2010	SELVA MEDIANA SUBPERENNIFOLIA	SMQ

#	Sitio	Arbol	NombreCientifico_APG	X	Y	Forma_Biologica_1
#1	4	24	Bucida buceras	-88.58853	19.23450	Arbol
#2	1	8	Coccoloba barbadensis	-88.36000	19.12333	Arbol
#3	2	13	Protium copal	-88.36000	19.12367	Arbol
#4	2	26	Metopium brownei	-88.36000	19.12367	Arbol
#5	2	6	Drypetes lateriflora	-88.35083	19.21367	Arbol
#6	3	2	Sabal yapa	-88.30339	19.20894	Palma

#	Altura_total	Diametro_normal	GENERO	ESPECIE
#1	9.0	16.7	Bucida	buceras
#2	3.0	15.0	Coccoloba	barbadensis
#3	9.8	13.5	Protium	copal
#4	10.3	41.0	Metopium	brownei
#5	4.3	9.6	Drypetes	lateriflora
#6	6.9	12.1	Sabal	yapa

### 3.3 Correcciones taxonómicas

Durante esta parte del procesamiento se utiliza la función `correctTaxo()` del paquete **BIOMASS**. Dicha función tiene dos parámetros: el género y la especie. Esta función corrige los errores tipográficos para un nombre científico particular de alguna especie, en caso de existir dicho error. Para poder llevar a cabo esta actividad se hace uso de la base de datos Taxonomic Name Resolution Service (TNRS), que es una herramienta en línea automatizada para estandarizar los nombres de las plantas.

Esta función devuelve un *dataframe* con el nombre correcto del género y la especie, además de la indicación si este tiene modificaciones con relación al nombre científico proporcionado. En este caso en particular no se encontraron nombres con errores, pero hubo 106 registros en donde no se encontró un taxón válido o registrado en la base de datos.

```
# Abrir los siguientes paquetes
```

```
library(BIOMASS)
library(knitr)
library("httr")
```

```
# Encontrar la sintaxis de cada especie usar la Fórmula: correctTaxo(genus, species)
```

```
Taxo<-correctTaxo(genus=data$GENERO, species=data$ESPECIE)
```

```
# Cache last modification time : 2020-03-04 10:53:03
```

```
# Source cache:18758
```

```
# Corrections FALSE:18652, TaxaNotFound:106
```

Posteriormente, los nombres modificados se guardan como factores y se despliegan los niveles de cada factor. Por ejemplo, en los nombres modificados, solo existen dos niveles: FALSE y TaxaNotFound, TRUE no está como factor en la base de datos. Además, se explora el *dataframe* "Taxo", que proporciona la lista de géneros especies y si estos fueron o no corregidos.

```
# Los nombres modificados se asignan como factores
```

```
Taxo$nameModified <- as.factor(Taxo$nameModified)
Taxo$speciesCorrected <- as.factor(Taxo$speciesCorrected)
```

```
# Se asignan los niveles de cada factor
```

```
levels(Taxo$nameModified)

# [1] "FALSE"                "TaxaNotFound"

levels(Taxo$speciesCorrected)

# [1] "acapulcensis"        "aculeata"
# [3] "acuminata"           "acuminatus"
# [5] "aegyptium"           "aesculifolia"
# [7] "albicans"            "alicastrum"
```

```
# Explorar el dataframe con los nombres corregidos
```

```
head(Taxo)

#   genusCorrected   speciesCorrected   nameModified
# 1      Bucida         buceras          FALSE
# 2   Coccoloba     barbadensis          FALSE
# 3     Protium         copal            FALSE
# 4   Metopium       brownei            FALSE
# 5   Drypetes     lateriflora          FALSE
# 6     Sabal         yapa              FALSE
```

En la última parte de este procesamiento, se asignan los nombres corregidos de género y especie a la base de datos obtenida en la sección 3.2 en el *dataframe* “**data**”. Adicionalmente, utilizando la función **getTaxonomy()**, que tiene como parámetros el género y la indicación de la familia, se obtuvo la familia a la que pertenece cada género y se agregó al *dataframe* “**data**”. Esto es importante porque podría utilizarse la familia como una categoría para aquellos casos donde no existe densidad de la madera a nivel de género; se puede utilizar el promedio de la familia para asignar un valor de densidad de la madera. Con estos procedimientos, ahora “**data**” tienen 17 variables en lugar de 14.

```
# Agregar los nombres de especies y géneros correctos al dataframe
```

```
data$GENEROCor <- Taxo$genusCorrected
data$ESPECIECor <- Taxo$speciesCorrected
```

```
# Encontrar la familia a partir de los géneros y ordenar si es necesario.
Para ello se usa la siguiente fórmula: getTaxonomy(genus, findOrder = FALSE)
```

```
Fam<-getTaxonomy(data$GENEROCor, findOrder =F)
```

```
# Agregar la familia al dataframe
```

```
data$Familia <- Fam$family
```

```
names(data)
```

```
# [1] "Conglomerado"           "Anio"
# [3] "Tipo_veg_SV"           "Veg_prim_levantada"
# [5] "Sitio"                  "Arbol"
# [7] "NombreCientifico_APG"  "X"
# [9] "Y"                       "Forma_Biologica_1"
# [11] "Altura_total"          "Diametro_normal"
# [13] "GENERO"                 "ESPECIE"
# [15] "GENEROCor"             "ESPECIECor"
# [17] "Familia"
```

### 3.4 Extraer valores de densidad de la madera de bases de datos locales y globales

Para extraer valores de densidad de la madera, que se utilizan en el cálculo de la biomasa de cada individuo, se usa la función `getWoodDensity(genus, species, stand = NULL, family = NULL, region = "World", addWoodDensityData = NULL)` del paquete **BIOMASS**. Esta función tiene cinco parámetros: el género, la especie, la unidad de observación, la familia, una indicación de la región para extraer los datos globales y un *dataframe* con los datos de densidad de la madera locales. Los parámetros indicados con **NULL** son opcionales. Por ejemplo, si se tienen datos locales de densidad de la madera se utilizaría el parámetro `addWoodDensityData = dataframe` con los datos; en caso contrario, no se coloca este parámetro en la función.

Se dispone de una lista con valores de densidad de la madera locales que se ha recolectado en la Unidad de Recursos Naturales del Centro de Investigación Científica de Yucatán (CICY) para la península de Yucatán. Estos datos están disponibles en el archivo `Yucatan_DM.csv` dentro del directorio `c:/Mapeo_biomasa/FCP`. Se leyó este archivo y se asignó al *dataframe* "DMlocal"; además, se realizó una exploración de los datos.

```
DMlocal <- read.table(choose.files(caption = "Selecciona el archivo con datos de
DM"), header = TRUE, sep = ",")
head(DMlocal)
```

#	genus	species	wd
# 1	Acacia	angustissima	0.85
# 2	Acacia	berlandieri	0.60
# 3	Acacia	cochliacantha	0.94
# 4	Acacia	cornigera	0.70
# 5	Acacia	dolichostachya	0.75
# 6	Acacia	gaumeri	0.90

A continuación, se utilizó la función `getWoodDensity()` para extraer los valores de densidad de la madera para cada especie. Se tienen 1011 valores de densidad de la madera, de los cuales, 959 se obtuvieron de la base de datos local. En caso de que una especie no tenga un valor de densidad de la madera, se asigna el promedio a nivel de especies, género o conglomerado; asimismo, se puede utilizar también el promedio a nivel de familia. En este caso, se consideró obtener un promedio de la densidad de la madera con las condiciones ambientales del sitio y, por ello, se optó por el promedio del conglomerado. Los datos de densidad de la madera son alma-



cenados en un *dataframe*. De todos los árboles medidos en la base de datos, se proporciona un resumen de cuántos tuvieron valores de densidad de la madera a nivel de especie, género y conglomerado. En 13 406 registros se obtuvo densidad de la madera para especies, 4 276 para géneros y 1 076 para conglomerado. Finalmente, la medida de la densidad de la madera, así como el nivel y su desviación estándar son asignados a la base de datos obtenida en la sección 3.3. en el *dataframe* “**data**”, que ahora tiene 20 variables.

# Fórmula cuando hay datos locales, además de los globales

```
dataDM<-getWoodDensity(genus=data$GENEROCor,
  species=data$ESPECIECor,
  stand=data$Conglomerado,
  region="Mexico",
  addWoodDensityData=DMlocal)

# The reference dataset contains 1011 wood density values
# Your taxonomic table contains 221 taxa
```

# Ver el número de individuos con valores de densidad de la madera  
estimados a nivel de #Especie, géneros y parcela.  
# A nivel de especie

```
sum(dataDM$levelWD=="species")
# [1] 13406
```

# A nivel de género

```
sum(dataDM$levelWD=="genus")
# [1] 4276
```

# A nivel de parcela

```
sum(!dataDM$levelWD%in%c("genus","species"))
# [1] 1076
```

```
# Agregar la Densidad de la madera al dataframe
```

```
data$DM <- dataDM$meanWD
data$DM_Tipo <- dataDM$levelWD
data$DM_SD <- dataDM$sdWD

names(data)

# [1] "Conglomerado"          "Anio"                  "Tipo_veg_SV"
# [4] "Veg_prim_levantada"    "Sitio"                 "Arbol"
# [7] "NombreCientifico_APG" "X"                     "Y"
# [10] "Forma_Biologica_1"    "Altura_total"         "Diametro_normal"
# [13] "GENERO"               "ESPECIE"              "GENEROCor"
# [16] "ESPECIECor"          "Familia"              "DM"
# [19] "DM_Tipo"              "DM_SD"
```

### 3.5 Cálculo de la biomasa a nivel de árbol individual

Para el cálculo de la biomasa de cada individuo se puede utilizar la función **computeAGB()** del paquete **BIOMASS**. Esta función tiene tres parámetros: el diámetro, la densidad de la madera y la altura, y utiliza la ecuación alométrica desarrollada por Chave *et al.* (2014), la cual estima la biomasa en kilogramos. Una expresión para calcular la biomasa en R sería: **AGBtree<-computeAGB(D=data\$Diametro\_normal, WD=data\$DM, H =data\$Altura\_total)**. Posteriormente, se convierte esta biomasa a toneladas por hectárea. Para ello, primero se divide la biomasa de cada individuo entre 1 000, convirtiendo la biomasa de kilogramos a toneladas; después se multiplica esa cantidad por 6.25, que es el resultado de dividir la superficie de la hectárea entre la superficie de las cuatro parcelas de 400 m<sup>2</sup> (10000/1600 = 6.25), por lo que una expresión para la conversión de la biomasa en R sería **data\$AGB2 <- (AGBtree/1000) \* 6.25**.

Una segunda opción es utilizar ecuaciones alométricas locales. Aquí se utilizaron diferentes ecuaciones locales desarrolladas para diferentes tipos de vegetación, forma de vida y tamaño de los individuos (cuadro 1). En este manual se desarrolló una función en R para el cálculo de la biomasa de los individuos con estas ecuaciones. La función se denomina **AGBcalc()** y tiene ocho parámetros: el identificador del conglomerado, el diámetro, la altura, la forma de vida, la vegetación potencial registrada en la serie V de INEGI, la vegetación primaria definida por Conafor, la densidad de la madera y la superficie de la unidad de observación (1 600 m<sup>2</sup>). Todos estos parámetros se encuentran en el *dataframe* "data" construido en este capítulo. Dentro de la función se hace la conversión a toneladas por hectárea siguiendo el

mismo razonamiento mostrado arriba, por ello, se solicita la superficie de la unidad de observación en campo.

```
# Definir parámetros para usar la función del cálculo de biomasa
```

```
ID <- data$Conglomerado # Identificador de individuos
DAP <- data$Diametro_normal # DAP = Diámetro normal (cm)
H <- data$Altura_total # H = Altura Total (m)
FV <- data$Forma_Biológica_1 # FV = Forma de vida
VSV <- data$Tipo_veg_SV # VSV = Vegetación potencial registrada en la serie V de
INEGI
VP <- data$Veg_prim_levantada # VP = Vegetación primaria definida por Conafor
Ds <- data$DM # Ds = Densidad de la madera
S <- 1600 # S = superficie en m2 de las 4 parcelas de 400 m2
```

```
# Función para el cálculo de la biomasa aérea por árbol en toneladas por hectárea
```

```
AGBcalc <- function (DAP, H,FV, VSV, VP, Ds, S = NULL){

#vegetación definición ----
if (is.null(S)) {
  stop("Falta definir el valor de superficie del sitio de muestreo")
}

if (is.null(VP)) {
  VP = NA
} else {
  VP = VP
}
arbol <- c("Arbol", "Arboles", "ND")
lianas <- c("Liana", "Lianas", "Bejuco", "Bejucos")
palmas <-c("Palma", "Palmas")

if (!is.null(VSV)){
  ifelse (VSV%in% c("AGRICULTURA DE RIEGO", "AGRICULTURA DE TEMPORAL", "ASENTA-
MIENTOS HUMANOS", "MANGLAR",
                  "PALMAR NATURAL","PASTIZAL CULTIVADO", "PASTIZAL INDUCIDO", "SA-
BANA",
```

```

                "SIN VEGETACION APARENTE", "VEGETACION DE PETEN", "TULAR"),
    vegetacion <- VP ,   vegetacion <- VSV)

} else {
    vegetacion <- VP
}

```

# Factor de conversión de superficie a 1 ha, en este caso s = 1600

```

if (is.null(s)){
    stop("S superficie del sitio de muestra no está definido")
} else {
    FS <- 10000 / S
}

```

# Cálculo de la AGB por diferentes formas de vida con diferentes ecuaciones

```

if (FV%in% palmas){
#Ecuación para palmas
    AGB <- ( (-4.51+(7.7*H))/1000)*FS
}
else if (FV%in% lianas) {
#Ecuación para lianas
    AGB <- ((exp(0.049+2.053*log(DAP)))/1000)*FS
}
else if (FV%in% arbol & (vegetación%in% c("VM"))) {

#Ecuación para manglares obtenida de Chave 2003
    AGB <- ((Ds*exp(-1.349+(1.980*log(DAP))+0.207*(log(DAP))^2)-(0.281*(log(DAP))^
3)))/1000)*FS
}
#Ecuaciones para árboles en selvas caducifolias y subcaducifolias
else if (FV%in% arbol & (
    vegetacion%in% c("SELVA BAJA CADUCIFOLIA", "SELVA BAJA ESPINOSA
CADUCIFOLIA", "SELVA BAJA SUBCADUCIFOLIA",
    "SELVA MEDIANA CADUCIFOLIA", "SELVA MEDIANA SUBCADUCIFOLIA",

```

```

        "SMS", "SBC", "SBS", "SMC", "SBK", "VW", "VS", "NA"))){
  ifelse (DAP < 10,
    # para árboles con dap < 10 se usó la ecuación de Ramírez
    AGB <- ((exp(-4.1392+(0.99*log((DAP^2)*H)))+(1.2268*Ds))/1000)*FS,
    # para árboles con dap > 10 se usó la ecuación de Chave 2005
    AGB <- ((Ds*exp(-0.667+(1.784*log(DAP)))+(0.207*(log(DAP))^2)-(0.0281*(log(D
AP))^3))/1000)*FS )
  }
#Ecuaciones para árboles en selvas subperennifolias
  else if (FV%in% arbol & (
    vegetacion%in% c("SELVA ALTA PERENNIFOLIA", "SELVA ALTA SUBPEREN-
NIFOLIA","SELVA BAJA ESPINOSA SUBPERENNIFOLIA",
    "SELVA BAJA SUBPERENNIFOLIA", "SELVA MEDIANA SUBPERENNIFOLIA",
    "SMQ", "SBQ", "SAQ", "PT", "SBQP", "SMP", "PI", "SBP",
    "SAP", "ADV", "VPI", "VT", "VA", "RAS", "VPN", "SIN VEGETACION")))
{

  ifelse (DAP < 10,
    #para árboles con dap < 10 se usó la ecuación de Guyot
    AGB <- ((exp((1.3636*log(DAP)))+(1.615*log(H))-2.9267))/1000)*FS,
    #para árboles con dap > 10 se usó la ecuación de Cairns modificada por Ur-
quiza-Haas
    AGB <- ((exp(-2.173+(0.868*log((DAP^2)*H)))+(0.0939/2)*(Ds/0.7))/1000)*FS)
  } else {
    AGB <- 0.0
  }
}
}

```

Aquí se calcula la biomasa área para cada registro de la base de datos utilizando la función **AGBcalc()**:

```
# Cálculo de biomasa para cada registro
```

```

for (i in 1:(length(ID))){
  data$AGB[i] <- AGBcalc(DAP = DAP[i],
    H = H[i],
    FV = FV[i],

```

```

VSV = VSV[i],
VP = VP[i],
Ds = Ds[i],
S = S)
}

```

El *dataframe* “**data**” ahora tiene 21 variables incluyendo la biomasa de cada individuo en la variable AGB.

```

names(data)

# [1]      "Conglomerado"      "Anio"      "Tipo_veg_SV"
# [4]      "Veg_prim_levantada"  "Sitio"     "Arbol"
# [7]      "NombreCientifico_APG" "X"        "Y"
# [10]     "Forma_Biologica_1"  "Altura_total" "Diametro_normal"
# [13]      "GENERO"      "ESPECIE"    "GENEROCor"
# [16]     "ESPECIECor"      "Familia"    "DM"
# [19]     "DM_Tipo"      "DM_SD"     "AGB"

```

### 3.6 Cálculo de la biomasa y otros atributos de la vegetación por conglomerado

En esta última parte del procesamiento se obtienen resúmenes de los principales atributos de la vegetación, tales como el diámetro, la altura y el área basal para todo el conglomerado. Adicionalmente, se colocan las coordenadas geográficas de cada conglomerado, con el objetivo de relacionar la información de campo con la de las imágenes de ALOS PALSAR. Finalmente, esta base de datos se guarda en el archivo “**FCP\_bio\_resumen.csv**” para utilizarse en los siguientes procesos.

```
# Suma de la biomasa aérea por conglomerado
```

```
AGBconglomerado <- tapply(data$AGB, data$Conglomerado, FUN=sum)
```

**Cuadro 1.** Descripción de las ecuaciones alométricas usadas para estimar la biomasa

AUTOR	ECUACIÓN ALOMÉTRICA	TIPO DE VEGETACIÓN	FORMA DE VIDA/ TAMAÑO
Ramírez <i>et al.</i> (2019)	$EXP(-4.1392 + 0.99 * LN(DBH2 * LENG) + 1.2268 * DENS)$	Selva baja y mediana subcaducifolia	árbol/diámetro <10 cm
Chave <i>et al.</i> (2005)	$DENSI * EXP(-0.667 + 1.784 * LN(DBH) + 0.207 * LN(DBH)^2 - 0.0281 * LN(DBH)^3)$	Selva baja y mediana subcaducifolia	árbol/diámetro ≥10 cm
Guyot (2011)	$EXP(1.3636 * LN(DBH) + 1.615 * LN(LENG) - 2.9267)$	Selva mediana subperenifolia	árbol/diámetro <10 cm
Cairns (2003) modificada por Urquiza-Haas <i>et al.</i> (2007)	$EXP(-2.12605 + 0.868 * LN(DBH2 * TH) + (0.0939/2)) * (DENS/0.7)$	Selva mediana subperenifolia	árbol/diámetro ≥10 cm
Chave <i>et al.</i> (2003)	$EXP(0.049 + 2.053 * LN(DBH))$	Selva baja, mediana subcaducifolia y mediana subperenifolia	Liana/diámetro ≥2.5 cm
Frangi y Lugo (1985)	$-4.51 + (7.7 * LENG)$	Selva baja, mediana subcaducifolia y mediana subperenifolia	Palma/diámetro ≥10 cm

# Cálculo del área basal AB por tallo en hectárea

```
AB <- pi*(data$Diámetro_normal/2)^2
data$AB <- (AB/10000)*6.25
```

# Obtener la sumatoria del área basal por conglomerado

```
ABconglomerado <- tapply(data$AB, data$Conglomerado, sum)
```

# Promedio del diámetro, altura y densidad de la madera por conglomerado

```
data_promedios <- data%>% group_by(Conglomerado)%>%
  summarise_at(.vars = vars(DM, Diámetro_normal ,Altura_total),
    .funs = c(mean="mean"), na.rm=T)
```

# Obtener las coordenadas de cada conglomerado, con las coordenadas de la parcela uno

```
data1 <- filter(data, Sitio == 1)
data2 <- data1%>% group_by(Conglomerado)%>%
  summarise_at(.vars = vars(Anio,X,Y),
    .funs = c(mean="mean"), na.rm=T)
```

```
# Unir los dataframes
```

```
data_resumen <- as.data.frame(cbind(data2,data_promedios,ABconglomerado,AGBconglomerado))
```

```
#eliminar columnas no deseadas
```

```
data_res<- data_resumen[ , c(1,2,3,4,7,8,9,10)]
```

```
# Escribir un archivo CSV
```

```
write.csv(data_res, "FCP_bio_resumen.csv")
```

En la figura 4 se observa la base de datos en donde se calcularon los resúmenes de diferentes atributos de la vegetación por conglomerado, la suma de las áreas basales y la biomasa de todos los individuos dentro de un conglomerado, así como los promedios del diámetro y la altura para cada conglomerado. Esta base de datos contiene, además, la ubicación geográfica de cada conglomerado.



Figura 4. Base de datos con atributos de la vegetación por conglomerado

	A	B	C	D	E	F	G	H	I	J	K	L
1												
2		56919	2013	-88.319166	19.529166	13.83227092	11.46693227	29.66482733	186.5877655			
3		56920	2014	-88.271749	19.52536044	13.56842105	10.68038278	23.22318304	141.5529599			
4		56921	2010	-88.224194	19.52027711	14.339801	10.8520896	24.5132437	142.1027443			
5		56922	2011	-88.176388	19.51608333	14.38625	10.15791667	30.26303566	173.282656			
6		56923	2012	-88.128888	19.51166667	14.05444015	9.597297297	29.79516416	160.3606799			
7		56924	2013	-88.081388	19.50722222	14.4593361	10.29211618	32.47811374	181.9957955			
8		56925	2009	-88.033944	19.50277778	15.11849711	11.41734104	25.52563175	162.7338782			
9		57169	2012	-88.513889	19.50111111	15.72925532	10.3356383	29.85184045	153.9424181			
10		57170	2013	-88.466389	19.49697189	16.03652695	10.11676647	26.58681266	144.4003382			
11		57171	2009	-88.418888	19.49249967	9.02195122	6.148780488	1.69483524	5.297607271			
12		57172	2010	-88.371416	19.488333	15.27451923	10.17115385	28.91894452	166.5655602			
13		57173	2011	-88.323638	19.4838856	14.34090909	11.74747475	24.53091516	161.8107001			
14		57174	2012	-88.27611	19.47972156	14.64463519	10.7583691	30.69376274	174.2055019			
15		57175	2013	-88.22861	19.47527711	15.10882353	10.76176471	27.71621031	158.4081166			
16		57176	2014	-88.181249	19.471416	12.68458333	9.809166667	21.49969058	123.8919908			
17		57177	2010	-88.133611	19.46669378	14.71435644	9.911881188	26.55891139	152.2977578			
18		57179	2012	-88.038583	19.45777778	11.89444444	7.766203704	17.36916874	80.29254509			
19		57420	2009	-88.566111	19.46027778	15.34417476	10.14320388	30.47908888	171.7399334			
20		57421	2010	-88.518333	19.45611111	12.06649215	8.531937173	16.63974492	83.55289167			
21		57422	2011	-88.47086	19.45166667	11.10636364	7.031818182	15.16792349	64.22742378			
22		57423	2012	-88.42336	19.44749967	17.26527778	10.75069444	27.4402214	156.5176179			
23		57424	2013	-88.375833	19.443083	14.07577093	10.58370044	26.96957155	153.2379804			
24		57425	2009	-88.328333	19.4388856	16.38216561	10.22292994	25.95673189	140.6093601			
25		57426	2010	-88.280833	19.43444411	15.27891892	10.33135135	28.08780182	170.1677606			
26		57477	2011	00.333333	10.43037711	14.42410355	10.0244095	22.946575	153.0270150			



## 4. Preprocesamiento de las imágenes ALOS PALSAR

LAS IMÁGENES DEL MOSAICO GLOBAL DE ALOS PALSAR, con una resolución espacial de 25 m, fueron utilizadas en este manual para el mapeo de la biomasa aérea, es por eso que en este capítulo se revisará la forma de descargar estas imágenes, así como la creación de mosaicos con las diferentes escenas obtenidas durante la descarga y el recorte de áreas particulares de estudio. Por otro lado, las distorsiones geométricas del sensor y el efecto de la topografía son corregidos aplicando una ortorrectificación y corrección de la pendiente aplicadas directamente por la JAXA. Es decir, las imágenes disponibles de ALOS PALSAR tienen dichas correcciones. Sin embargo, las imágenes obtenidas en la JAXA se proporcionan en números digitales, lo cual hace necesario realizar una conversión de dichos números a coeficientes de retrodispersión (Shimida y Ohtaki, 2010). Finalmente, para reducir el ruido presente en las imágenes de radar o moteado, sin sacrificar la estructura de la imagen, se recomienda la aplicación de filtros. En este capítulo, se aplicará el filtro propuesto por Lee (1980), el cual ejecuta una función espacial sobre una ventana de 3 x 3 píxeles, en donde el valor del píxel central se reemplaza por un valor calculado utilizando los píxeles vecinos.

### 4.1 ¿Cómo descargar las imágenes de ALOS PALSAR?

Las imágenes ALOS PALSAR-2 pueden ser descargadas de la página de la Agencia de Exploración Aeroespacial de Japón (JAXA). Para poder realizar la descarga, primero es necesario realizar un registro en donde se solicita un nombre de usuario, clave y correo electrónico. Esto puede hacerse mediante el siguiente enlace: [https://www.eorc.jaxa.jp/ALOS/en/palsar\\_fnf/registration.htm](https://www.eorc.jaxa.jp/ALOS/en/palsar_fnf/registration.htm). Cuando algún usuario se ha registrado, se tiene acceso a todas las imágenes históricas de ALOS PALSAR. Para realizar la descarga de las imágenes se debe introducir el usuario y la clave de acceso en el siguiente enlace: [https://www.eorc.jaxa.jp/ALOS/en/palsar\\_fnf/data/index.htm](https://www.eorc.jaxa.jp/ALOS/en/palsar_fnf/data/index.htm).

Debido a que los datos de campo se tomaron entre 2009 y 2014, las imágenes a descargar con la fecha disponible más cercana son las de 2015. Elija este año en la página web. En otra página (fig. 5) se encuentra un mapa con la disponibilidad de imágenes ALOS PALSAR en el mundo. Dar un clic donde está la península de Yucatán, como se muestra en la figura 4. Realizar los acercamientos necesarios para encontrar las escenas identificadas como N19W089 y N20W089, que son las que cubren el área donde se levantaron los datos del Inventario Nacional Forestal para este estudio. Cuando se selecciona la escena, aparece una vista previa de la capa de bosque y no bosque, así como las capas con las polarizaciones **HH** y **HV**; proceder a la descarga de los archivos (fig. 6).

Como se menciona en el capítulo 3, la ruta de trabajo para el cálculo de la biomasa en este manual se definió como “**c:/Mapeo\_biomasa/FCP**”. Por lo tanto, la descarga de los archivos de ALOS PALSAR, se hará en este directorio. Para cada escena hay dos archivos comprimidos (**\*.tar**), uno que contiene un mapa clasificado con las áreas cubiertas con bosque y no bosque, y el otro donde se encuentran las dos polarizaciones **HH** y **HV**.

**Figura 5.** Mapa mundial en donde se encuentran disponibles las imágenes de ALOS PALSAR. Con la flecha roja podrá seleccionar un área para realizar acercamientos

Please click on an area of the map to download the data.

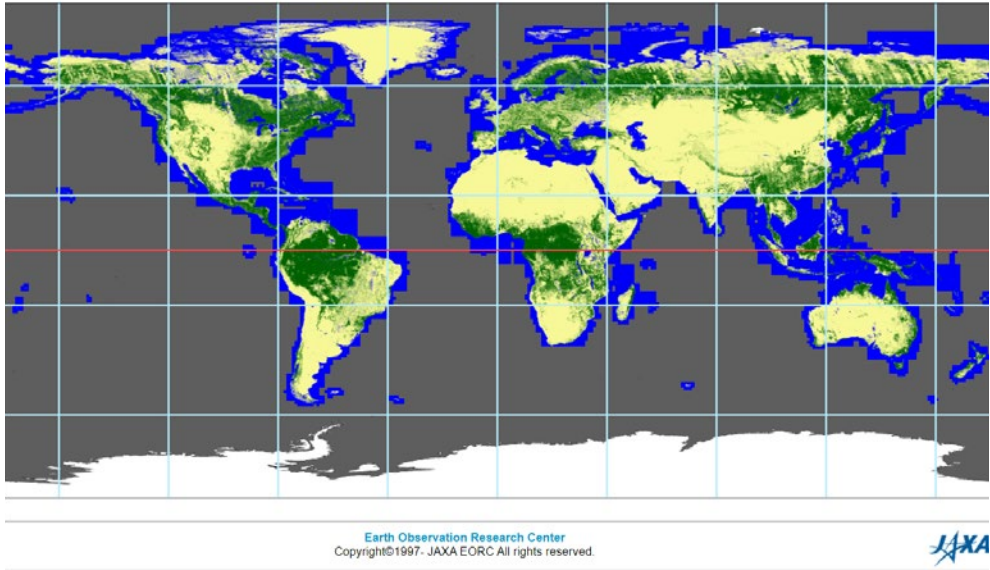
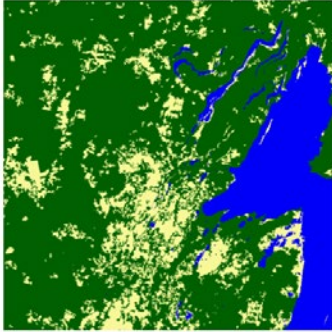

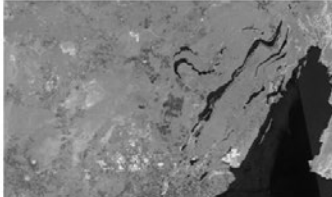


Figura 6. Descarga de una escena en ALOS PALSAR



N19W089\_2015

Year	Type	Image	Download
	FNF		
2015	HH		

## 4.2 Preparación de datos y descompresión de archivos

La primera actividad en el preprocesamiento consiste en definir el espacio de trabajo, que en este manual se definió con la siguiente ruta “c:/Mapeo\_biomasa/FCP”. Adicionalmente, se guardó el directorio de trabajo en la variable “**folder**” para usarlo posteriormente.

```
# Definir directorio de trabajo
```

```
setwd(choose.dir(default = "", caption = "Selecciona el directorio de trabajo"))
getwd()
# [1] "C:/Mapeo_biomasa/FCP"
```

```
folder <- getwd()
```

Se abren los paquetes utilizados en este *script*.

```
# Abrir los siguientes paquetes
```

```
library(rgdal)
library(spatial.tools)
library(sp)
library(raster)
library(whitebox)
wbt_version()
```

El área de estudio donde se tomaron los datos de campo es un polígono de 3600 km<sup>2</sup>, y está guardado en un archivo de tipo vector. Se utilizaron las funciones **shapefile()** y **choosefiles()** para dar lectura al polígono. Al ejecutar la segunda función, se debe buscar el directorio y el archivo de tipo vector que contiene los límites del área de estudio, es decir, se debe ingresar el archivo *shapefile* (**FCP\_square\_poly.shp**), que se encuentra en el directorio “**c:/Mapeo\_biomasa/FCP**”. Después de leer el polígono se pueden ver las coordenadas y la proyección geográfica que este tiene (WGS84).

```
# Cargar el polígono con el área de estudio (archivo shapefile)
```

```
area <- shapefile(choose.files(caption = "selecciona el shapefile con el polígono de
estudio"))
plot(area)
summary(area)
# Object of class SpatialPolygonsDataframe
# Coordinates:
#           min                max
# x      -88.60036          -88.02964
# y       18.98903           19.53101
# Is projected: FALSE
# proj4string :
```

```
# [+proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0]
# Data attributes:
# .....
```

En la última actividad de este apartado, se forma una lista con los archivos de ALOS PALSAR descargados de la JAXA, los cuales se descomprimen usando la función `untar()` y se almacenan en una carpeta llamada “files\_AlosPalsar” ubicada en el directorio de trabajo “C:/Mapeo\_biomasa/FCP/files\_AlosPalsar”. En los resultados se puede observar que la lista se compone de los cuatro archivos comprimidos.

# Crear lista de archivos comprimidos con extensión **tar.gz**

```
list.gz<-list.files(path=".", pattern ="tar.gz")
list.gz
# [1] "N19W089_15_FNF_F02DAR.tar.gz" "N19W089_15_MOS_F02DAR.tar.gz"
# [3] "N20W089_15_FNF_F02DAR.tar.gz" "N20W089_15_MOS_F02DAR.tar.gz"
```

# Crear un folder para los archivos a descomprimir

```
Name.Dir.Decompress <- "Files_AlosPalsar" # Nombre del directorio
{ifelse(!
  dir.exists(file.path(folder,Name.Dir.Decompress)),
  dir.create(file.path(folder,Name.Dir.Decompress)),
  FALSE)
dir.files <- file.path(folder,Name.Dir.Decompress)}
```

# Descomprimir todos los archivos con la función **untar()**

```
for (i in 1:length(list.gz)){
  untar(list.gz[i], exdir = dir.files)
}
```

Por último, se crea una lista con los archivos de ALOS PALSAR que contienen las polarizaciones **HH** y **HV**, en este caso particular son cuatro archivos, dos para cada escena. Esta lista se utiliza en las etapas posteriores de procesamiento. En los resultados se muestran las dimensiones de los archivos, la resolución espacial y la

extensión de los archivos, así como las coordenadas geográficas, el nombre del archivo y su ubicación.

```
# Crear lista de archivos descomprimidos únicamente los que contengan las
# polarizaciones HH y HV. Ambos pueden definirse por los últimos caracteres
# del nombre del archivo
```

```
list.rast <- list.raster.files(path = dir.files,
                              pattern = "HH_F02DAR$|HV_F02DAR$",
                              return_rasters = TRUE,
                              return_bbox = FALSE)

list.rast

# $raster_files
# $raster_files[[1]]
# class      : RasterBrick
# dimensions : 4500, 4500, 20250000, 1 (nrow, ncol, ncell, nlayers)
# resolution : 0.0002222222, 0.0002222222 (x, y)
# extent     : -89, -88, 18, 19 (xmin, xmax, ymin, ymax)
# crs        : +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0
# source     : C:/Mapeo_biomasa/FCP/Files_AlosPalsar/N19W089_15_sl_HH_F02DAR
# names      : N19W089_15_sl_HH_F02DAR
# min values :
# max values :
#           0
#           65535

# $raster_files[[2]]
# class      : RasterBrick
# dimensions : 4500, 4500, 20250000, 1 (nrow, ncol, ncell, nlayers)
# resolution : 0.0002222222, 0.0002222222 (x, y)
# extent     : -89, -88, 18, 19 (xmin, xmax, ymin, ymax)
# crs        : +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0
# source     : C:/Mapeo_biomasa/FCP/Files_AlosPalsar/N19W089_15_sl_HV_F02DAR
# names      : N19W089_15_sl_HV_F02DAR
# min values :
# max values :
#           0
#           65535

# $raster_files[[3]]
# class      : RasterBrick
```



```

# dimensions : 4500, 4500, 20250000, 1 (nrow, ncol, ncell, nlayers)
# resolution : 0.0002222222, 0.0002222222 (x, y)
# extent      : -89, -88, 19, 20 (xmin, xmax, ymin, ymax)
# crs         : +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0
# source      : C:/Mapeo_biomasa/FCP/Files_AlosPalsar/N20W089_15_sl_HH_F02DAR
# names       : N20W089_15_sl_HH_F02DAR
# min values  :
# max values  : 65535

# $raster_files[[4]]
# class       : RasterBrick
# dimensions  : 4500, 4500, 20250000, 1 (nrow, ncol, ncell, nlayers)
# resolution  : 0.0002222222, 0.0002222222 (x, y)
# extent      : -89, -88, 19, 20 (xmin, xmax, ymin, ymax)
# crs         : +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0
# source      : C:/Mapeo_biomasa/FCP/Files_AlosPalsar/N20W089_15_sl_HV_F02DAR
# names       : N20W089_15_sl_HV_F02DAR
# min values  :
# max values  : 65535

```

### 4.3 Creación de un mosaico con las escenas de ALOS PALSAR y recorte del área de estudio

Para hacer un recorte del área de estudio con el mosaico de datos, primero se debe tener la seguridad de que los dos grupos de datos tengan la misma proyección geográfica. Al polígono que limita el área de estudio se le asigna la proyección geográfica de una de las imágenes ALOS PALSAR, utilizando la función **spTransform()**.

```

# Primero se debe homogenizar la proyección del área de estudio
# con la de los rásters de ALOS PALSAR
# Visualizar el polígono

```

```

areaNew <- spTransform(area, CRS(projection(list.rast$raster_files[[1]])))
plot(areaNew)
summary(areaNew)
# Object of class SpatialPolygonsDataframe
# Coordinates:

```

```
#           min           max
# x         -88.60036     -88.02964
# y          18.98903      19.53101
# Is projected: FALSE
# proj4string :
# [+proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0]
# Data attributes:
# .....
```

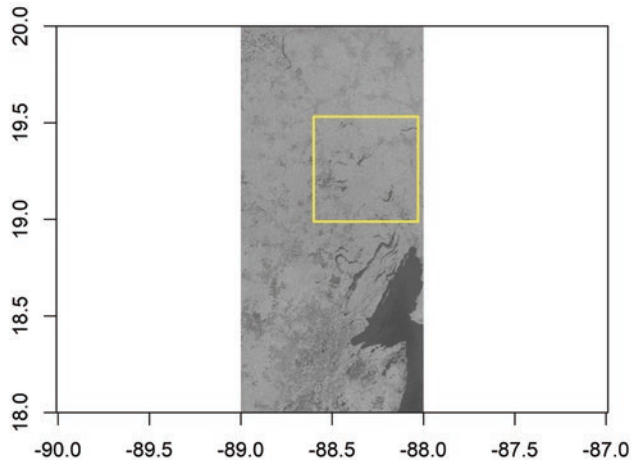
Los mosaicos se crean para los diferentes tipos de datos, las dos capas de las polarizaciones **HH**, otro para las dos capas de **HV** y el último para las dos capas de datos FNF (datos de bosque y no bosque). Para ello, primero se obtienen los nombres de cada tipo de datos, y después se crean listas separadas para cada uno de estos tipos de dato (**HH**, **HV** y **FNF**). Los mosaicos de cada lista se obtienen usando la función **do.call()**. Esta función tienen dos parámetros: el primero corresponde al nombre de la función a ejecutar, en este caso es **merge()**; el segundo parámetro es la lista de archivos a unir en un mosaico. Como ejemplo se grafica el mosaico con las polarizaciones **HV** junto con el polígono de estudio en color amarillo.

```
# Crear un mosaico con las escenas
# Listar los nombres de los archivos por tipo de imagen
```

```
FileHH <- list.files(path = dir.files, pattern = "HH_F02DAR$", full.names = TRUE)
FileHH
FileHV <- list.files(path = dir.files, pattern = "HV_F02DAR$", full.names = TRUE)
FileHV
FileFNF <- list.files(path = dir.files, pattern = "C_F02DAR$", full.names = TRUE)
FileFNF
```

```
# Definir la lista de nombres de archivos en formato raster
```

```
HH.list <- list()
for(i in 1:length(FileHH)){
  HH.list[[i]] <- raster(FileHH[i])
}
HV.list <- list()
for(i in 1:length(FileHV)){
```



```

HV.list[[i]] <- raster(FileHV[i])
}
FNF.list <- list()
for(i in 1:length(FileFNF)){
  FNF.list[[i]] <- raster(FileFNF[i])
}

```

# Unir los raster a un nuevo raster layer

```

HH.merge <- do.call(merge, HH.list)
HV.merge <- do.call(merge, HV.list)
FNF.merge <- do.call(merge, FNF.list)

plot(HV.merge, col = gray.colors(30, start = 0.3, end = 1), legend = FALSE)
plot(areaNew, add = TRUE, lwd = 2, border = "yellow")

```

Para realizar el recorte del área de estudio en los mosaicos de las polarizaciones **HH** y **HV** y en el mapa de bosque no bosque (**FNF**), se utilizan dos funciones. La primera es la función **extent()**, la cual tiene como parámetro un objeto de tipo vector o *raster*, en este caso fue el archivo con el polígono de estudio en formato vector, en el cual se había realizado una proyección geográfica. Esta función devuelve un objeto espacial en la extensión del polígono. La segunda función es **crop()**, que realiza el

recorte del área de estudio, y tiene dos parámetros: uno es el objeto en donde se hará el recorte, los cuales son los mosaicos creados anteriormente en un formato *raster*, el segundo parámetro es la extensión del polígono de estudio. Finalmente, se grafican los *rasters* de las capas **HH**, **HV** y **FNF** (en este caso se presenta solamente el de **HV**).

# Definir el área de extracción

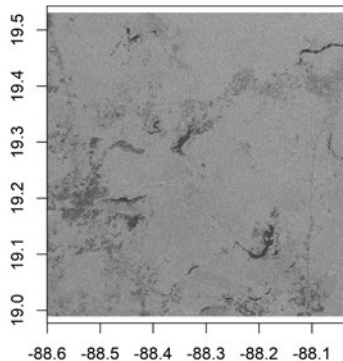
```
ext.shp <- extent(areaNue)
ext.shp

# class      : Extent
# xmin       : -88.60036
# xmax       : -88.02964
# ymin       : 18.98903
# ymax       : 19.53101
```

# Extraer el área de estudio en los mosaicos creados anteriormente

```
HH_crop <- crop(HH.merge, ext.shp)
HV_crop <- crop(HV.merge, ext.shp)
FNF_crop <- crop(FNF.merge, ext.shp)

plot(HH_crop, col = gray.colors(30, start = 0.3, end = 1), legend = FALSE)
plot(HV_crop, col = gray.colors(30, start = 0.3, end = 1), legend = FALSE)
plot(FNF_crop,
     col= c("darkgreen", "khaki1", "navyblue"),
     legend = FALSE)
```



#### 4.4 Convertir valores de números digitales a retrodispersión en las polarizaciones HH y HV

Para realizar la conversión de números digitales a coeficientes de retrodispersión se utilizó la siguiente ecuación (Shimida y Ohtaki, 2010):

$$\gamma^{\circ}(dB) = 10\text{Log}_{10}(DN^2) - 83.0(1)$$

donde  $DN$  es el número digital expresado como entero corto sin signo y  $\gamma^{\circ}$  es el coeficiente de retrodispersión expresado en decibeles.

En esta parte del procesamiento, primero se define el directorio en donde se guardarán las imágenes con valores de retrodispersión; el nombre del directorio es **Bcks** y está ubicado en “**C:/Mapeo\_biomasa/FCP/Bcks/**”.

# Definir un directorio de salida de archivos

```
{Name.Dir.Bcks <- "Bcks" # Nombre del directorio
  ifelse(!
    dir.exists(file.path(folder,Name.Dir.Bcks)),
    dir.create(file.path(folder,Name.Dir.Bcks)),
    FALSE)
  dir.Bcks <- file.path(folder,Name.Dir.Bcks)}
}
```

Se define la función que convierte valores digitales a retrodispersión usando la ecuación (1).

# Definir la función de cálculo

```
fun<- function (x) {
  (10 * log10(x ^ 2) - 83)
}
```

Para generar los archivos *raster* de las polarizaciones **HH** y **HV** con los valores de retrodispersión, se utiliza la función **calc(x, fun, filename='', overwrite=FALSE, format='', ...)**, la cual tienen varios parámetros. El parámetro **x** es el archivo *raster* original; el segundo parámetro corresponde al nombre del archivo, en este caso, está compuesto por el directorio + nombre del archivo, **HH** o **HV** según sea el caso; el

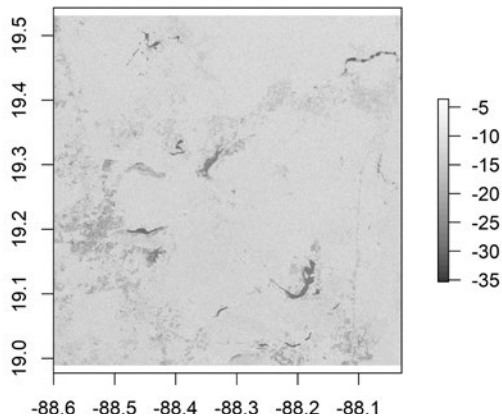
tercer parámetro es para indicar que se desea sobrescribir el archivo. Por último, se define el formato del archivo como **GeoTiff**.

# **Calc()** es una función de cálculo para *raster* a nuevos *raster* mediante una función

```
HH.bck <- calc(HH_crop, fun,
  filename = paste(dir.Bcks, "/", "HH", sep = ""),
  overwrite = TRUE,
  format="GTiff")# Guardar archivos en formato Gtiff
HV.bck <- calc(HV_crop, fun,
  filename = paste(dir.Bcks, "/", "HV", sep = ""),
  overwrite = TRUE,
  format="GTiff")
```

Una vez obtenidas las imágenes convertidas a valores de retrodispersión se grafican como se indica en el siguiente *script* (se muestra el resultado de la imagen **HV**).

```
plot(HH.bck,
  col = gray.colors(30, start = 0.3, end = 1),
  legend = TRUE)
plot(HV.bck,
  col = gray.colors(30, start = 0.3, end = 1),
  legend = TRUE)
```



## 4.5 Aplicación del filtro de Lee

Durante esta parte del procesamiento, primero se define el directorio en donde se guardarán las imágenes que tienen valores de retrodispersión y que serán corregidas por el filtro de Lee para eliminar ruido. El nombre del directorio es **LeeFl**, el cual estará ubicado en “**C:/Mapeo\_biomasa/FCP/LeeFl/**”

```
# Crear directorio para guardar los archivos
```

```
{Name.Dir.Lee <- "LeeFl" # Nombre del directorio

  ifelse(!
    dir.exists(file.path(folder,Name.Dir.Lee)),
    dir.create(file.path(folder, Name.Dir.Lee)),
    FALSE)
  dir.LeeFl <- file.path(folder,Name.Dir.Lee)}
}
```

Después se crea una lista usando la función **list.files()**, con los nombres de archivos de salida del procesamiento anterior ubicados en “**C:/Mapeo\_biomasa/FCP/Bcks/**”. Al desplegar la lista se observan los archivos **HH** y **HV** en formato **tiff**.

```
BckList <- list.files(path = dir.Bcks, pattern = ".tif$")
BckList

# [1] "HH.tif" "HV.tif"
```

Se utiliza la función **wbt\_lee\_sigma\_filter()** para aplicar el filtro de Lee. Dicha función tiene cuatro parámetros. El primero es la ubicación y la lista de nombres de archivos de entrada, el segundo tiene la ubicación y la lista de nombres de los archivos de salida, en el tercero y el cuarto está el número de renglones y columnas que definen el tamaño de la ventana donde se aplicará el filtro de Lee.

```
for (i in 1:length(BckList)){
  wbt_lee_sigma_filter(file.path(dir.Bcks,BckList[i]),
    file.path( dir.LeeFl,BckList[i]),
```

```

        filterx =3,
        filtery =3)
}

```

#### 4.6 Convertir las coordenadas de las imágenes preprocesadas, de geográficas a proyectadas

Se define el directorio en donde se guardarán las imágenes proyectadas. El nombre del directorio es **IMG**, el cual estará ubicado en “**C:/Mapeo\_biomasa/FCP/IMG/**”. Además, se crea una lista usando la función **list.raster.files()**, con los archivos preprocesados (imágenes con conversión a retrodispersión y aplicación del filtro de Lee) ubicados en “**C:/Mapeo\_biomasa/FCP/LeeF1/**”. Esta función, a diferencia de la función **list.files()**, guarda la lista de archivos en lugar de los nombres de los mismos. Al desplegar la lista se pueden observar las características de los archivos preprocesados, como el formato (*raster*), las dimensiones, la extensión, el tipo de proyección (WGS84), la ubicación del archivo y su nombre. Se tienen dos archivos con nombre **HH** y **HV**.

```
# Crear el directorio de salida
```

```

{Name.Dir.IMG <- "IMG" # Nombre del directorio
}ifelse(!
  dir.exists(file.path(folder,Name.Dir.IMG)),
  dir.create(file.path(folder,Name.Dir.IMG)), FALSE
)
dir.IMG <- file.path(folder,Name.Dir.IMG)}
}

FileList <- list.raster.files(path=dir.LeeF1, pattern = ".tif$",
  return_rasters = TRUE,
  return_bbox = FALSE)

FileList
# $raster_files
# $raster_files[[1]]
# class      : RasterBrick
# dimensions : 2439, 2569, 6265791, 1 (nrow, ncol, ncell, nlayers)

```



```
#resolution : 0.0002222222, 0.0002222222 (x, y)
# extent    : -88.60044, -88.02956, 18.98911, 19.53111 (xmin, xmax, ymin, ymax)
# crs       : +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0
# source    : C:/Mapeo_biomasa/FCP/LeeFl/HH.tif
# names     : HH

# $raster_files[[2]]
# class     : RasterBrick
# dimensions : 2439, 2569, 6265791, 1 (nrow, ncol, ncell, nlayers)
# resolution : 0.0002222222, 0.0002222222 (x, y)
# extent    : -88.60044, -88.02956, 18.98911, 19.53111 (xmin, xmax, ymin, ymax)
# crs       : +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0
# source    : C:/Mapeo_biomasa/FCP/LeeFl/HV.tif
# names     : HV
```

En esta fase se utiliza la proyección que usa el INEGI para la república mexicana que es la Cónica Conforme de Lambert (CCL) y el Datum ITRF2008 (INEGI, 2001). Para llevar a cabo esta proyección se utiliza la función **projectRaster(x, crs, filename="", overwrite=, format=, ...)**, esta función contiene varios parámetros: **x** es el archivo *raster* que será proyectado; **crs** es el objeto que describe el sistema de coordenadas; **filename** tiene la dirección y nombre del archivo; **overwrite** para sobrescribir archivos existentes y **format** para asignar el formato de salida del archivo.

# Definir la nueva proyección en formato PROJ4

```
# México ITRF2008/LCC EPSG: 6372
newProj <- "+proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +towgs84=0,0,0,0,0,0 +units=m +no_defs"
#crs(newProj)
```

# Proyección de los archivos en lista **FileList** (HH y HV)

```
proj <- list()
for (i in 1:length(FileList$raster_files)){
  proj[[i]] <- projectRaster(FileList$raster_files[[i]],
    crs = newProj,
```

```

        filename = file.path(dir.IMG,names(FileList$raster_files[[i]])),
        overwrite = TRUE,
        format = "HFA"
    )
}

```

Para terminar con este módulo, las imágenes proyectadas se colocan en una pila utilizando la función `stack()`, es decir, en un objeto multibanda. Además, se utiliza la función `writeRaster(x, filename, overwrite, format, , ...)` para guardar en el disco las imágenes. Esta función contiene varios parámetros: el primero es el objeto *raster* que va a grabar; el segundo corresponde al nombre del archivo y su ubicación (se guardará en el directorio **IMG**, con el nombre **FCP**); el siguiente parámetro especifica que el archivo se sobrescribe con el mismo nombre; finalmente, el último parámetro que utilizamos especifica el formato, que en este caso es **GeoTiff**.

# Guardar la capa HH y HV en un raster multibanda

```

stack <- stack(proj)
writeRaster(stack,
            filename = file.path(dir.IMG,"FCP"),
            overwrite = TRUE,
            format = "GTiff" )

```

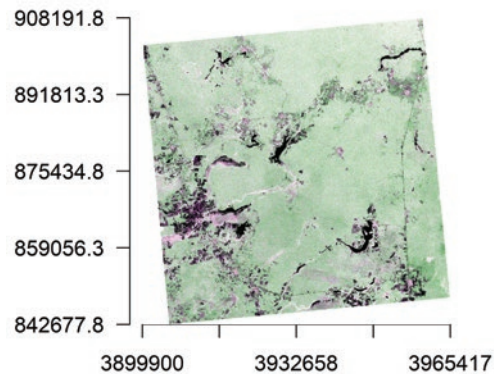
Para visualizar el archivo *raster* multibanda como un compuesto RGB, se utiliza la función `plotRGB()`. Esta función tiene varios parámetros y se usaron los siguientes: el primero es el objeto *raster* multibanda; en el segundo parámetro se especifica que en rojo se coloca la polarización **HH**; después, en verde se coloca la polarización **HV**; en el siguiente parámetro se especifica que para el color azul se coloca de nueva cuenta la polarización **HH**; con otro parámetro se colocan los ejes; finalmente, en el último parámetro se especifica que se realiza un estiramiento lineal de la imagen.

# Visualización imagen SAR como un compuesto RGB

```

plotRGB(stack, r=1, g=2, b=1,axes=TRUE, stretch ="lin" )

```



Para finalizar este módulo se lleva a cabo la proyección del *raster* **FNF**, que contiene el mapa de bosque no bosque, con la función **projectRaster()** similar a como se utilizó anteriormente.

# Proyección de la imagen FNF recortada

```
projectRaster(FNF_crop,
  crs = newProj,
  filename = paste(dir.IMG,"FNF.tif", sep= "/"),
  overwrite = TRUE,
  format = "GTiff") # Guardar en f
```



## 5. Procesamiento de las imágenes ALOS PALSAR

LAS IMÁGENES DE ALOS PALSAR QUE han sido preprocesadas están listas para extraer valores de retrodispersión de las dos polarizaciones **HH** y **HV**, que pueden ser asociados con valores de la biomasa aérea en campo para ajustar modelos de predicción de la biomasa aérea. Sin embargo, como se mencionó en la introducción, existen algunas limitaciones en el uso de los valores de retrodispersión para estimar la distribución espacial de la biomasa. Una de las limitaciones más importantes es que los valores de retrodispersión pueden no cambiar o cambiar muy poco en bosques densos y/o con estructura de la vegetación compleja, efecto conocido como saturación. Algunos cálculos mencionan que el nivel de saturación es cercano a  $150 \text{ mg ha}^{-1}$  (Mermoz *et al.*, 2015; Morel *et al.*, 2011). A pesar de ello, en la selva mediana subperennifolia de la península de Yucatán, se han reportado valores de biomasa superiores a  $300 \text{ Mg ha}^{-1}$  (Hernández-Stefanoni *et al.*, 2014), lo cual hace necesario incorporar análisis adicionales que intenten superar esta limitación. Aquí se calculará el índice diferencial de retrodispersión normalizado (NDBI, por sus siglas en inglés), además de diferentes medidas de textura de segundo orden.

### 5.1 Configuración del directorio de trabajo y lectura de datos

Para iniciar el procesamiento de las imágenes, se cargan los paquetes utilizados en este módulo y se define el espacio de trabajo, que es el directorio donde se encuentran las imágenes multibanda preprocesadas y con el sistema de coordenadas CCL y Datum ITRF2008, es decir, la siguiente ruta: “**c:/Mapeo\_biomasa/FCP/IMG**”.

```
# Cargar paquetes utilizados en este módulo
```

```
library(glcm)
library(rgeos)
library(rgdal)
library(raster)
```

```
library(sp)
library(spatial.tools)
```

```
# Definir directorio de trabajo
```

```
setwd(choose.dir(default = "",
                 caption = "Seleccionar folder con las imágenes en formato IMG"))
getwd()
# [1] "C:/Mapeo_biomasa/FCP/IMG"

folder <- getwd()
```

En este módulo también se carga el polígono con el área de estudio, debido a que se utilizará esta área en la última parte del *script*. Para ello se debe ingresar el archivo *shapefile* (**FCP\_square\_poly.shp**), que se encuentra en el directorio "**c:/Mapeo\_biomasa/FCP**". Este polígono tiene coordenadas geográficas con base en Datum WGS84 y se proyectará a CCL con Datum ITRF2008 usando la función **spTransform()**, que tiene como parámetros un objeto vector y la proyección deseada. Después de ejecutar el código, se observan las características del polígono, ahora su proyección es CCL con el elipsoide GRS80.

```
# Cargar el shapefile del área de estudio
```

```
area <- shapefile(choose.files(caption = "selecciona el shapefile del área de
estudio"))
plot(area)
summary(area)
# Object of class SpatialPolygonsDataframe
# Coordinates:
#
#           min           max
# x      -88.60036      -88.02964
# y       18.98903       19.53101
# Is projected: FALSE
# proj4string :
# [+proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0]
```

```
# Data attributes:
# .....
```

```
# Transformación de coordenadas del polígono de interés
# similar a la proyección de las imágenes
```

```
newProj <- "+proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +towgs84=0,0,0,0,0,0,0 +units=m +no_defs"
```

```
newarea<- spTransform(area,newProj)
```

```
newarea
```

```
# class      : SpatialPolygonsDataframe
# features   : 1
# extent     : 3900026, 3965290, 842796.3, 908057.9 (xmin, xmax, ymin, ymax)
# crs        : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102
+x_0=2500000 +y_0=0 +ellps=GRS80
# +towgs84=0,0,0,0,0,0,0 +units=m +no_defs
# variables  : 5
# names      :      x,      y, BUFF_DIST, ORIG_FID, ORIG_FID_1
# value      : -88.315, 19.26,      30,      0,      0
```

Los últimos datos en leer son los archivos de las imágenes preprocesadas con las polarizaciones **HH** y **HV** (archivos con la terminación **\*.img**), los cuales se almacenan en una lista y para ello se utiliza la función **list.raster.files()**. Como se observa en el resultado de la lectura en cada una de las listas, existe solamente un archivo de cada polaridad, primero el **HH** y luego el **HV**; sin embargo, el *script* podría leer más de dos archivos por lista.

```
# Leer los archivos de cada polarización
```

```
FileListHH <- list.raster.files(path = folder, pattern = "HH(.*).img$",
                               return_bbox = FALSE,
                               return_rasters = TRUE)
FileListHH$raster_files
# [[1]]
```

```

# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs        : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+ y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/HH.img
# names      :      HH
# min values : -27.88649
# max values :  3.94952

FileListHV <- list.raster.files(path = folder, pattern = "HV(.*).img$",
                               return_bbox = FALSE,
                               return_rasters = TRUE)
FileListHV$raster_files
[[1]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs        : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+ y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/HV.img
# names      :      HV
# min values : -31.29588
# max values : -8.525797

```

## 5.2 Cálculo de NDBI

El índice diferencial de retrodispersión normalizado (**NDBI**) se calcula con los coeficientes de retrodispersión **HH** y **HV** por medio de la siguiente ecuación:

$$NDBI = \frac{HH - HV}{HH * HV} \quad (2)$$

Este índice ayuda a distinguir áreas con distinta estructura del bosque debido a la contribución diferenciada de la dispersión dada por las dos polarizaciones. Para



calcular una imagen con este índice se utiliza la función `overlay(x, y, ..., fun, filename="", overwrite=TRUE, format=)`. Esta función del paquete *raster*, crea un nuevo objeto *raster*, usando dos o más objetos *raster* (también se podría usar un solo objeto), pero basado en una función que realice algún cálculo entre el o los objetos. La función `overlay()` puede tener varios parámetros, aquí se utilizaron los siguientes: una lista de objetos *raster* para **HH** (**x**); otra lista de objetos *raster* para **HV** (**y**); la función de cálculo considerando los dos objetos *raster* (**fun**); el nombre del archivo (**filename**); la especificación de permitir escribir el archivo en disco y el formato del mismo.

```
for (i in 1:length(FileListHH$raster_files)){
  overlay(FileListHH$raster_files[[i]],
    FileListHV$raster_files[[i]],
    # Definir la funcion
    fun=function (x,y) {return((x - y)/(x + y))},
    filename= paste(folder,"/", "NDBI", sep = ""),
    overwrite =TRUE,
    format="HFA")
}
```

### 5.3 Cálculo de las medidas de textura en HH, HV y NDBI

Las medidas de textura se calcularon para imágenes ALOS PALSAR de media resolución (pixel de 25 m). Estas medidas permiten cuantificar la variación horizontal de los pixeles que se pueden relacionar con un gradiente de las diferentes condiciones en el dosel del bosque, tales como los fragmentos de bosque con diferentes edades de sucesión y que complementan la penetración vertical que tienen los sensores de radar (Fisher *et al.*, 2019). El análisis de textura se refiere a un procedimiento para caracterizar las variaciones espaciales dentro de las imágenes y puede calcularse con estadísticas de segundo orden, basadas en la matriz de coocurrencia de niveles de grises (GLCM) (Haralick *et al.*, 1973). Se calcularon ocho medidas de textura de segundo orden para ambas polarizaciones (**HH** y **HV**) y para el índice (**NDBI**). Las medidas de textura de segundo orden se calcularon a 0, 45, 90 y 135 grados y se promediaron para obtener un único valor de textura sin dirección, el tamaño de ventana es de 3 x 3 pixeles que permite capturar la variabilidad dentro de una hectárea. La lista de métricas está descrita en Haralick *et al.* (1973).

El procesamiento para calcular las métricas de textura de las imágenes inicia con la creación de una lista de los nombres de los archivos a procesar. Como puede

verse, la lista contiene los archivos con las polarizaciones **HH** y **HV**, así como la del índice **NDBI**.

```
# Generar la lista de nombres de archivos a procesar
```

```
FileListIMG <- list.files(path = folder, pattern = ".img$")
FileListIMG
# [1] "HH.img" "HV.img" "NDBI.img"
```

Se crea una lista de vectores en donde se almacenarán los resultados de las texturas. Ahora los vectores están vacíos y se tiene un vector para cada capa.

```
# Crear lista de vectores para almacenar los resultados de las texturas
```

```
List <- vector(mode = "list",
              length = length(FileListIMG))
List
# [[1]]
# NULL
# [[2]]
# NULL
# [[3]]
# NULL
```

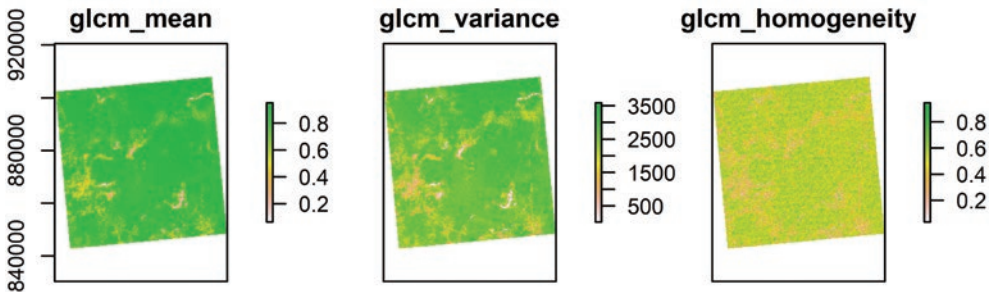
Se utiliza la función `gldm(x, n_grey = 32, window = c(3, 3), shift = c(1, 1), statistics = c("mean", "variance", "homogeneity", "contrast", "dissimilarity", "entropy", "second_moment", "correlation"), min_x=NULL, max_x=NULL, na_opt="any", na_val=NA, scale_factor=1, asinteger=FALSE)`, la cual contiene varios parámetros: en el primero se especifica el objeto *raster* en el que se calcularán las texturas (**x**), que en este caso es una lista de archivos *raster*; en el segundo parámetro se especifica el nivel de grises (**n\_grey**), aquí se elige 64; después se especifica el tamaño de ventana, en este caso fue de 3 x 3 píxeles, que es el área más cercana a una hectárea; en otro parámetro se especifican las direcciones del cálculo de texturas (cuatro texturas), y el último parámetro permite seleccionar las medidas de textura a calcular.

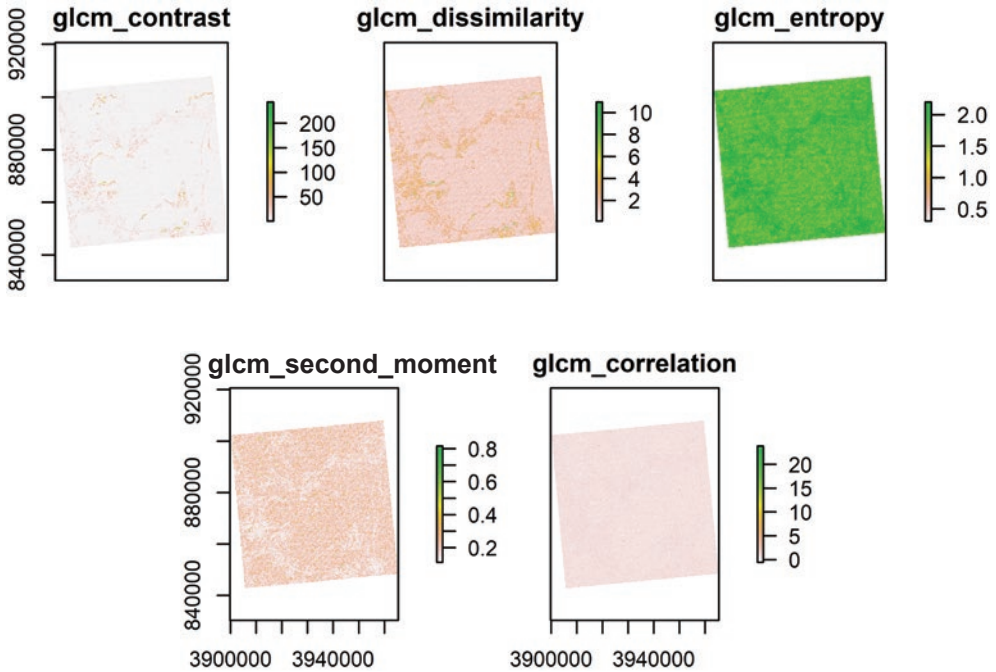
# Cálculo de texturas para todas las **.img** en la lista

```
for (i in 1:length(FileListIMG)){
  List[[i]] <- glcm(raster(FileListIMG[i]),
    n_grey=64,
    window = c(3,3),
    shift=list(c(0,1), c(1,1), c(1,0), c(1,-1)),
    statistics = c("mean", "variance", "homogeneity",
      "contrast", "dissimilarity", "entropy",
      "second_moment", "correlation"),
    min_x=NULL, max_x=NULL,
    na_opt="any", na_val=NA,
    scale_factor=1,
    asinteger=FALSE)
}
```

Una vez obtenidas las imágenes de cada textura para cada una de las tres capas (HH, HV y NDBI) se graficaron para verificar cada métrica (se presenta como ejemplo la capa HV).

```
for (i in 1:length(List)){
  plot(List[[i]])
}
```





Se crea el directorio en donde se guardarán las imágenes que tienen las texturas. El nombre del directorio es **IMG\_texturas**, el cual estará ubicado en **“C:/Mapeo\_biomasa/FCP/IMG/”**.

# Crear el directorio de salida de texturas

```
{Name.Dir.text <- "IMG_texturas"
  ifelse(!
    dir.exists(file.path(folder, Name.Dir.text)),
    dir.create(file.path(folder, Name.Dir.text)), FALSE
  )
  dir.text <- file.path(folder, Name.Dir.text)}
}
```

Al final de este apartado, se guardan en el directorio de salida las ocho imágenes de textura para cada una de las capas **HH**, **HV** y **NDBI**, usando la función **WriteRaster()**,

que tiene como parámetros: un objeto *raster*, es decir, la imagen de textura; el nombre y la ruta del archivo; la indicación de sobrescritura, y el formato de los archivos.

# Guardar las imágenes de textura en formato HFA

```
for (i in 1:length(FileListIMG)){
  writeRaster(List[[i]]$glcm_mean,
    paste(dir.text,"/",
      sub(pattern = "(.*)\\..*$",
        replacement = "\\1", basename(FileListIMG[i])),
      "_mean", sep = ""),
    overwrite =TRUE,
    format="HFA")
  writeRaster(List[[i]]$glcm_variance,
    paste(dir.text,"/",
      sub(pattern = "(.*)\\..*$",
        replacement = "\\1", basename(FileListIMG[i])),
      "_vari", sep = ""),
    overwrite =TRUE,
    format="HFA")
  writeRaster(List[[i]]$glcm_homogeneity,
    paste(dir.text,"/",
      sub(pattern = "(.*)\\..*$",
        replacement = "\\1", basename(FileListIMG[i])),
      "_homo", sep = ""),
    overwrite =TRUE,
    format="HFA")
  writeRaster(List[[i]]$glcm_contrast,
    paste(dir.text,"/",
      sub(pattern = "(.*)\\..*$",
        replacement = "\\1", basename(FileListIMG[i])),
      "_cont", sep = ""),
    overwrite =TRUE,
    format="HFA")
  writeRaster(List[[i]]$glcm_dissimilarity,
    paste(dir.text,"/",
      sub(pattern = "(.*)\\..*$",
```

```

                replacement = "\\1", basename(FileListIMG[i])),
            "_diss", sep = ""),
        overwrite = TRUE,
        format="HFA")
writeRaster(List[[i]]$glcm_entropy,
            paste(dir.text, "/",
                sub(pattern = "(.*)\\..*$",
                    replacement = "\\1", basename(FileListIMG[i])),
                    "_entr", sep = "")),
            overwrite = TRUE,
            format="HFA")
writeRaster(List[[i]]$glcm_second_moment,
            paste(dir.text, "/",
                sub(pattern = "(.*)\\..*$",
                    replacement = "\\1", basename(FileListIMG[i])),
                    "_seco", sep = "")),
            overwrite = TRUE,
            format="HFA")
writeRaster(List[[i]]$glcm_correlation,
            paste(dir.text, "/",
                sub(pattern = "(.*)\\..*$",
                    replacement = "\\1", basename(FileListIMG[i])),
                    "_corr", sep = "")),
            overwrite = TRUE,
            format="HFA")
}

```

## 5.4 Corrección de archivos

En el último paso de este módulo, se corrigen los valores de algunos archivos, debido a que presentan valores nulos y con estos valores no se puede ejecutar el modelo u obtener el mapa final. Inicialmente, se crea una lista de archivos *raster* con las imágenes de textura, utilizando la función `list.raster.files()`. Al imprimir la lista se obtienen las propiedades de los 24 archivos cargados, ocho imágenes de textura para cada una de las capas (**HH**, **HV** y **NDBI**). Como ejemplo se presentan los primeros tres archivos de texturas con el contraste, la disimilaridad y la correlación en la capa **HH**.

## # Lista de raster de salida

```

FileListIMG_t <- list.raster.files(path = dir.text, pattern = ".img$", return_bbox =
FALSE, return_rasters = TRUE)
FileListIMG_t
# $raster_files
# $raster_files[[1]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs        : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_cont.img
# names      : HH_cont
# min values :      0
# max values : 289.8889

# $raster_files[[2]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs        : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_corr.img
# names      : HH_corr
# min values : -0.7375351
# max values :  70.31269

# $raster_files[[3]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)

```

```
# crs      : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source   : C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_diss.img
# names    : HH_diss
# min values :      0
# max values : 13.66667
```

Particularmente, la imagen que tiene valores nulos es la de la correlación, entonces se crea una lista con las imágenes de correlación para cada una de las capas (**HH**, **HV** y **NDBI**) usando la función **list.raster.files()**. Al imprimir la lista, se obtienen las propiedades de los tres archivos cargados en la lista: correlación de la capa **HH**, correlación de la capa **HV** y correlación de la capa **NDBI**.

# Lista de textura **correlation** para la sustitución de valores NA a 0

```
FileListIMG_corr <- list.raster.files(path = dir.text, pattern = "corr.img$", re-
turn_bbox = FALSE, return_rasters = TRUE)
FileListIMG_corr
# $raster_files
# $raster_files[[1]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs       : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_corr.img
# names      : HH_corr
# min values : -0.7375351
# max values : 70.31269

# $raster_files[[2]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
```



```
# resolution : 23.2, 24.4 (x, y)
# extent      : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs         : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source      : C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HV_corr.img
# names       : HV_corr
# min values  : -14.67318
# max values  : 93.79423

# $raster_files[[3]]
# class       : RasterBrick
# dimensions  : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution  : 23.2, 24.4 (x, y)
# extent      : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs         : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source      : C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/NDBI_corr.img
# names       : NDBI_corr
# min values  : -14.63102
# max values  : 106.431
```

En seguida, se realiza la sustitución de los valores nulos por el valor de cero en las imágenes de correlación de la lista con las tres capas **HH**, **HV** y **NDBI**.

# Sustitución de valores NA

```
for (i in 1:length(FileListIMG_corr$raster_files)){
  FileListIMG_corr$raster_files[[i]][is.na(FileListIMG_corr$raster_files[[i]])] <- 0
}
```

Otra corrección importante se refiere a eliminar los valores de los bordes; para ello, se utiliza el polígono del área de estudio obtenido en la sección 5.1. Para llevar a cabo este recorte se utiliza la función **mask(x, mask, ...)** del paquete *raster*. Aquí se utilizaron dos parámetros: el objeto *raster* a recortar (**x**) que, en este caso, es una

lista de archivos y un objeto espacial (**mask**) que, en este caso, es el polígono con el área de estudio.

```
for (i in 1:length(FileListIMG_corr$raster_files)){
  FileListIMG_corr$raster_files[[i]]<-mask(FileListIMG_corr$raster_files[[i]], newarea)
}
```

Para finalizar este módulo, se realiza la escritura de las imágenes de correlación usando la función **writeRaster()** del paquete *raster*, de la misma forma que se utilizó anteriormente al final de la sección 5.3.

# Escritura de los archivos

```
for (i in 1:length(FileListIMG_corr$raster_files)){
  writeRaster(FileListIMG_corr$raster_files[[i]],
             filename = file.path(dir.text, names(FileListIMG_corr$raster_files[[i]])),
             overwrite =TRUE,
             format="HFA")
}
```

## 6. Preparación de bases de datos para la modelación

EN ESTE MÓDULO SE CREARÁ UNA base datos que contenga tanto la información de los datos de biomasa en campo, como la información de las imágenes que comprende los datos de retrodispersión (**HH**, **HV** y **NBBI**) y las medidas de textura para cada una de las capas. Esta base de datos se utilizará para ajustar los modelos que estimen la biomasa relacionando datos de campo y datos de las imágenes ALOS PALSAR. Para ello, se realizarán las siguientes actividades: primero, se crea un archivo espacial de puntos con las coordenadas y los datos del cálculo de la biomasa en campo; después, se extraerán los valores de las medidas de textura usando los puntos; posteriormente, se extraerán los valores promedio de retrodispersión en una hectárea que tiene como centro las coordenadas de los conglomerados en campo; por último, se creará una base de datos con toda la información.

### 6.1 Crear un archivo espacial con la base de datos de campo

Para iniciar con la creación de la base de datos espacial que tiene la información de la biomasa calculada en campo y las coordenadas de cada conglomerado, se define el espacio de trabajo, que es el directorio donde se encuentra la base de datos, es decir, en: `"c:/Mapeo_biomasa/FCP/"`.

```
# Definir el directorio de trabajo
```

```
setwd(choose.dir( default = ".", caption = "Seleccione el directorio de trabajo"))
getwd()
folder <- getwd()
```

Se cargan los paquetes que se usarán en este módulo.

```
library(sp)
library(raster)
```

```
library(spatial.tools)
library(sf)
```

Se lee el archivo con la base de datos de campo y sus coordenadas que se llama **FCP\_bio\_resumen.csv** y que se encuentra en el directorio “**c:/Mapeo\_biomasa/FCP/**” usando la función **read.csv()**. En seguida, con la función **names(data)** se muestran todas las variables que contiene el *dataframe* “**data**”.

```
# Cargar el archivo que contenga las coordenadas
```

```
data <- read.csv(choose.files(caption = "Selecciona el archivo de coordenadas en
formato CSV"),
  header = TRUE,
  sep = ",")
names(data)

# [1]          "X"          "Conglomerado"          "Anio_mean"
# [4]        "X_mean"        "Y_mean"          "Diametro_normal_mean"
# [7]    "Altura_total_mean"    "ABconglomerado"    "AGBconglomerado"
```

Se utiliza la función **st\_as\_sf(x, coor, crs=, ...)** para convertir un objeto *dataframe* a un objeto espacial de tipo vector. La función **st\_as\_sf()** tiene varios parámetros: el primer parámetro contiene al objeto (**x**) que se convertirá a un objeto espacial; el segundo parámetro contiene las coordenadas de cada uno de los puntos, que en el *dataframe* “**data**” son llamadas **X\_mean** y **Y\_mean**; el último parámetro se refiere al sistema de coordenadas de los puntos que se van a ingresar, las coordenadas de los conglomerados están en coordenadas geográficas con base en el Datum WGS84. Por otro lado, la función **summary()**, cuyo parámetro es el objeto *points*, proporciona un resumen estadístico de cada una de las variables de este objeto espacial, además del número de puntos que este objeto tiene, así como el tipo de proyección: lat, long.

```
# Convertir dataframe to spatial feature
# Asignar las coordenadas nativas de los puntos
```

```
points <- st_as_sf(data, coords = c("X_mean", "Y_mean"),
  # definir la proyeccion nativa
```

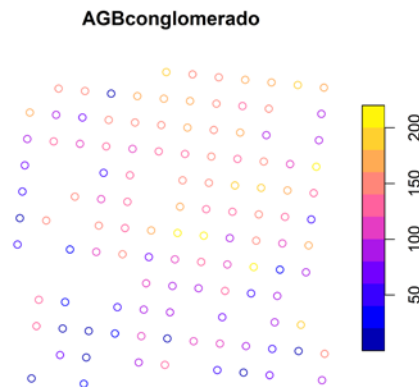
```

crs = "+proj=longlat +ellps=WGS84 +datum=WGS84 +no_defs"
)
summary(points)
#           X           Conglomerado           Anio_mean           Diametro_normal_mean
#   Min.    :56919           Min.    :56919           Min.    :2009           Min.    : 9.022
#   1st Qu.:57672           1st Qu.:57672           1st Qu.:2010           1st Qu.:12.685
#   Median :58187           Median :58187           Median :2011           Median :14.746
#   Mean    :58320           Mean    :58320           Mean    :2011           Mean    :14.483
#   3rd Qu.:59187           3rd Qu.:59187           3rd Qu.:2012           3rd Qu.:16.068
#   Max.    :59947           Max.    :59947           Max.    :2014           Max.    :19.046
#
#   Altura_total_mean   ABconglomerado   AGBconglomerado           geometry
#   Min.    : 2.943           Min.    : 1.182           Min.    : 4.569           POINT           :117
#   1st Qu.: 8.305           1st Qu.:16.640           1st Qu.: 81.076           epsg:4326       : 0
#   Median : 9.809           Median :22.553           Median :121.978           +proj=long...   : 0
#   Mean    : 9.335           Mean    :20.864           Mean    :112.456
#   3rd Qu.:10.474           3rd Qu.:26.587           3rd Qu.:152.073
#   Max.    :13.406           Max.    :35.239           Max.    :218.408

```

La función **plot()** con el parámetro *points* realiza una gráfica del atributo número 7 que es **ABGconglomerado**, es decir, se grafica la distribución espacial de la biomasa en los puntos de muestreo.

```
plot(points[,7])
```



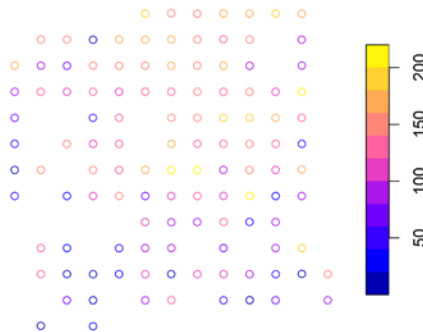
Las imágenes con las métricas de texturas y las polarizaciones tienen una proyección con el sistema de coordenadas CCL Datum ITRF2008, como quedó expuesto en el módulo 4. Para poder utilizar el archivo de puntos y realizar las extracciones de la información de las imágenes, es necesario proyectar el sistema de coordenadas del objeto de puntos con el mismo sistema de coordenadas que el de las imágenes. Se utiliza la función `st_transform()` con el objeto vector de puntos y la nueva proyección como parámetros para proyectar el objeto puntos de coordenadas geográficas a CCL (WSG84 a ITRF2008). Después se grafica la distribución espacial de la biomasa, pero ahora con otro sistema de coordenadas; como puede verse, cambia ligeramente la distribución de los puntos debido al cambio en el sistema de coordenadas.

```
# Nueva proyección de coordenadas # Mexico ITRF2008/LCC EPSG: 6372
```

```
newProj <- "+proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +towgs84=0,0,0,0,0,0,0 +units=m +no_defs"
points <- st_transform(points, newProj)

plot(points[,7])
```

AGBconglomerado



## 6.2 Extracción de los valores de las imágenes de textura con las coordenadas de puntos

Se extraerán los valores de las ocho texturas para cada una de las capas **HH**, **HV** y **NDBI**. Se recordará que las texturas se calcularon con un tamaño de ventana de 3 x 3 píxeles, que fue el tamaño más próximo a la unidad de observación (1 ha). Por otra

parte, la textura se escribe en el punto central de la ventana de 3 x 3 píxeles; por lo tanto, la extracción de los valores de las texturas se realiza tomando como base las coordenadas de cada conglomerado, es decir, se realiza una extracción de valores de las imágenes basada en las coordenadas de los conglomerados (puntos).

Para realizar la extracción de los valores de las texturas, primero se selecciona el directorio donde estas fueron gradadas, es decir, en “**c:/Mapeo\_biomasa/FCP/IMG/IMG\_texturas**”. Posteriormente, se leen los 24 archivos de imágenes que contienen las ocho texturas para cada una de las tres capas (**HH**, **HV** y **NDBI**), y se guardan en una lista de archivos usando la función **list.raster.files()**. Al imprimir la lista se pueden observar las propiedades de cada una de las 24 imágenes de textura, sin embargo, aquí solo se presentan las tres primeras: el contraste de la capa **HH**, la correlación de la capa **HH** y la disimilaridad de **HH**.

# Cargar directorio de imágenes de archivos de texturas

```
dir.IMG_textura <- choose.dir(caption = "Seleccione el directorio de los archivos
de texturas")

# Listas archivos raster
ListIMG_texturas <- list.raster.files(path = dir.IMG_textura, pattern = ".img$",
return_bbox = FALSE,
return_rasters = TRUE)

ListIMG_texturas

# $raster_files
# $raster_files[[1]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs        : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_cont.img
# names      : HH_cont
# min values :      0
# max values : 289.8889
```

```

# $raster_files[[2]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs        : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_corr.img
# names      : HH_corr
# min values : -0.7375351
# max values : 70.31269

# $raster_files[[3]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs        : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_diss.img
# names      : HH_diss
# min values : 0
# max values : 13.66667

```

Para facilitar el procesamiento se crea un archivo multibanda con las 24 imágenes de textura utilizando la función **stack()**, que tiene como parámetro la lista de archivos *raster*. Al imprimir este archivo multibanda, se observan las siguientes características: las dimensiones de las imágenes y el número de bandas que tiene este objeto (en este caso son 24), la resolución espacial, las coordenadas extremas de las imágenes y su proyección geográfica, así como la lista de imágenes y las estadísticas básicas de cada imagen.

```
# Stack de los archivos raster
```

```
texturs <- stack (ListIMG_texturas$raster_files)
```



**texturs**

```

# class      : RasterStack
# dimensions : 2685, 2824, 7582440, 24 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs       : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m
# +no_defs
# names      :      HH_cont,      HH_corr,      HH_diss,      HH_entr,      HH_
homo,      HH_mean,      HH_seco,      HH_vari,      # HV_cont,      HV_corr,
HV_diss,      HV_entr,      HV_homo,      HV_mean,      HV_seco, ...
# min values :  0.0000000, -0.73753512,  0.0000000,  0.0000000,
0.03406450,  0.04578993,  0.11111111,
# 7.08667614,  0.0000000, -14.67317772,  0.0000000,  0.0000000,  0.02115252,
0.02907986,  0.11111111, ...
# max values : 289.8888889,  70.3126907,  13.6666667,  2.1972246,
1.0000000,  0.9437934,  1.0000000,
# 3649.7854563, 461.5833333,  93.7942276,  15.0000000,  2.1972246,
1.0000000,  0.9713542,  1.0000000, ...

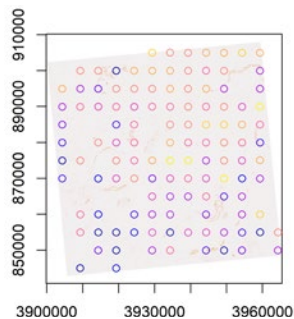
```

Con la finalidad de verificar que exista una coincidencia entre los archivos de las imágenes de textura y los puntos centrales de los conglomerados, se grafican ambas variables, antes de realizar la extracción.

```

plot(ListIMG_texturas$raster_files[[1]], legend= FALSE)
plot(points[,7], add = TRUE)

```



La extracción de información de las imágenes de textura se realiza con la función `extract(x, y, method='simple', df=FALSE, sp=FALSE, ...)` del paquete *raster*, la cual tiene diferentes parámetros. El primer parámetro (**x**) es un objeto de tipo *raster*, en este caso, es un archivo multibanda con las imágenes de textura. El parámetro (**y**) se refiere a un objeto espacial de puntos. En el siguiente parámetro (**method=simple**) se especifica el método de extracción, en este caso, se regresa el valor del pixel en donde cae el punto. El parámetro **df** indica que el resultado será regresado como un *dataframe* espacial. Finalmente, **sp** indica que los valores extraídos serán agregados al objeto espacial de puntos. La función `summary()` se utiliza para imprimir el resumen del *dataframe* espacial "**data.p**" y nos arroja las coordenadas mínimas y máximas, la proyección geográfica, el número de puntos (117 conglomerados), así como las estadísticas básicas de los diferentes atributos: las variables que tenía el objeto espacial "*points*" y las texturas.

# Extraer valores de texturas por punto central

```
data.p <- extract(texturs, points,
                 method = "simple",
                 df = TRUE, sp = TRUE)
summary(data.p)
```

```
# Object of class SpatialPointsDataframe
# Coordinates:
      min          max
# coords.x1 3904323.9 3964331.4
# coords.x2  844936.6  904994.7
# Is projected: TRUE
# proj4string : [+proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102
+x_0=2500000 +y_0=0 +ellps=GRS80
# +towgs84=0,0,0,0,0,0,0 +units=m +no_defs]
# Number of points: 117
# Data attributes:
```

#	X	Conglomerado	Anio_mean	Diametro_normal_mean
#	Min. :56919	Min. :56919	Min. :2009	Min. : 9.022
#	1st Qu.:57672	1st Qu.:57672	1st Qu.:2010	1st Qu.:12.685
#	Median :58187	Median :58187	Median :2011	Median :14.746
#	Mean :58320	Mean :58320	Mean :2011	Mean :14.483

#	3rd Qu.:59187	3rd Qu.:59187	3rd Qu.:2012	3rd Qu.:16.068
#	Max. :59947	Max. :59947	Max. :2014	Max. :19.046
#	Altura_total_mean	ABconglomerado	AGBconglomerado	
#	Min. : 2.943	Min. : 1.182	Min. : 4.569	
#	1st Qu.: 8.305	1st Qu.:16.640	1st Qu.: 81.076	
#	Median : 9.809	Median :22.553	Median :121.978	
#	Mean : 9.335	Mean :20.864	Mean :112.456	
#	3rd Qu.:10.474	3rd Qu.:26.587	3rd Qu.:152.073	
#	Max. :13.406	Max. :35.239	Max. :218.408	
#	HH_cont	HH_corr	HH_diss	
#	Min. : 0.2500	Min. :-0.2418	Min. :0.2500	
#	1st Qu.: 0.6389	1st Qu.: 0.1122	1st Qu.:0.5556	
#	Median : 0.8611	Median : 0.3108	Median :0.6944	
#	Mean : 1.2149	Mean : 0.3367	Mean :0.7509	
#	3rd Qu.: 1.2500	3rd Qu.: 0.5777	3rd Qu.:0.8611	
#	Max. :16.5000	Max. : 0.8608	Max. :2.8889	
#	HH_entr	HH_homo	HH_mean	HH_seco
#	Min. :0.7517	Min. :0.3224	Min. :0.5736	Min. :0.1296
#	1st Qu.:1.3013	1st Qu.:0.6111	1st Qu.:0.6200	1st Qu.:0.1975
#	Median :1.5615	Median :0.6722	Median :0.6328	Median :0.2346
#	Mean :1.5298	Mean :0.6649	Mean :0.6328	Mean :0.2567
#	3rd Qu.:1.7300	3rd Qu.:0.7278	3rd Qu.:0.6465	3rd Qu.:0.3086
#	Max. :2.0817	Max. :0.8750	Max. :0.6749	Max. :0.5741
#	HH_vari	HV_cont	HV_corr	HV_diss
#	Min. :1287	Min. : 0.500	Min. :-0.05146	Min. :0.5000
#	1st Qu.:1525	1st Qu.: 1.167	1st Qu.: 0.33425	1st Qu.:0.8056
#	Median :1594	Median : 1.694	Median : 0.53416	Median :1.0278
#	Mean :1594	Mean : 2.298	Mean : 0.50694	Mean :1.1144
#	3rd Qu.:1683	3rd Qu.: 2.833	3rd Qu.: 0.71825	3rd Qu.:1.3611
#	Max. :1839	Max. :19.028	Max. : 0.93804	Max. :2.9167
#	HV_entr	HV_homo	HV_mean	HV_seco
#	Min. :1.316	Min. :0.3137	Min. :0.7229	Min. :0.1111

#	1st Qu.:1.682	1st Qu.:0.4710	1st Qu.:0.8084	1st Qu.:0.1358
#	Median :1.846	Median :0.5556	Median :0.8296	Median :0.1728
#	Mean :1.846	Mean :0.5486	Mean :0.8301	Mean :0.1759
#	3rd Qu.:2.043	3rd Qu.:0.6361	3rd Qu.:0.8553	3rd Qu.:0.2099
#	Max. :2.197	Max. :0.7500	Max. :0.9060	Max. :0.2963
#	HV_vari	NDBI_cont	NDBI_corr	NDBI_diss
#	Min. :2041	Min. :0.1667	Min. :-0.3857	Min. :0.1667
#	1st Qu.:2579	1st Qu.:0.6667	1st Qu.: 0.2000	1st Qu.:0.5833
#	Median :2735	Median :1.0278	Median : 0.3925	Median :0.7778
#	Mean :2736	Mean :1.2766	Mean : 0.3726	Mean :0.8174
#	3rd Qu.:2920	3rd Qu.:1.5556	3rd Qu.: 0.5667	3rd Qu.:0.9722
#	Max. :3338	Max. :5.6389	Max. : 0.9113	Max. :1.9167
#	NDBI_entr	NDBI_homo	NDBI_mean	NDBI_seco
#	Min. :0.4358	Min. :0.3504	Min. :0.8140	Min. :0.1111
#	1st Qu.:1.4459	1st Qu.:0.5639	1st Qu.:0.8444	1st Qu.:0.1790
#	Median :1.6675	Median :0.6333	Median :0.8529	Median :0.2099
#	Mean :1.6216	Mean :0.6361	Mean :0.8535	Mean :0.2359
#	3rd Qu.:1.8216	3rd Qu.:0.7139	3rd Qu.:0.8639	3rd Qu.:0.2654
#	Max. :2.1972	Max. :0.9167	Max. :0.8852	Max. :0.7346
#	NDBI_vari			
#	Min. :2588			
#	1st Qu.:2814			
#	Median :2884			
#	Mean :2888			
#	3rd Qu.:2967			

### 6.3 Extracción de los valores de las polarizaciones HH, HV y NDBI

Las imágenes de ALOS PALSAR con las polarizaciones **HH** y **HV**, así como la del índice **NDBI** tiene valores independientes por pixel, a diferencia de las texturas que representan una ventana de 3 x 3 pixeles, por lo que la extracción valores de las retrodispersiones **HH**, **HV** y el índice **NDBI** se realizarán creando un *buffer* de una hectárea sobre las coordenadas de cada conglomerado; las extracciones consideran los valores promedio de los pixeles que caen en el polígono circular de una hectárea.

Para realizar la extracción de los valores de estas capas, primero se selecciona el directorio donde están almacenadas dichas imágenes, las cuales fueron previamente preprocesadas (“**c:/Mapeo\_biomasa/FCP/IMG/**”). Leer los tres archivos de imágenes (**HH, HV y NDBI**) por medio de la función **list.raster.files()**, y guardarlos en una lista de archivos *raster*. Al imprimir la lista se pueden observar las propiedades de cada una de las tres imágenes.

```
# Cargar directorio de imágenes de archivos backscatter
```

```
dir.IMG_bcks <- choose.dir(caption = "Seleccione el directorio de los archivos
Backscatter y NDBI")

# Listas archivos raster
ListIMG_bcks <- list.raster.files(path = dir.IMG_bcks, pattern = ".img$",
                                return_bbox = FALSE,
                                return_rasters = TRUE)
ListIMG_bcks

# $raster_files
# $raster_files[[1]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs        : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/HH.img
# names      :      HH
# min values : -27.88649
# max values :   3.94952

# $raster_files[[2]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
```

```

# crs      : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source   : C:/Mapeo_biomasa/FCP/IMG/HV.img
# names    :      HV
# min values : -31.29588
# max values : -8.525797

# $raster_files[[3]]
# class    : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent    : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs      : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m}
# +no_defs
# source   : C:/Mapeo_biomasa/FCP/IMG/NDBI.img
# names    :      NDBI
# min values : -1.78489
# max values : 0.06081292

```

De nueva cuenta, esta lista de archivos se une en un archivo multibanda para facilitar las extracciones; para ello, se utiliza la función **stack()**, que tiene como parámetro la lista de archivos *raster*. Al imprimir este archivo multibanda, se observan las siguientes características: las dimensiones de las imágenes y el número de bandas, tres en este caso; la resolución espacial; las coordenadas extremas de las imágenes; su proyección geográfica, así como la lista de imágenes y las estadísticas básicas de cada una de ellas (**HH**, **HV** y **NDBI**).

# Stack de los archivos *raster*

```

bcks <- stack (ListIMG_bcks$raster_files)
bcks

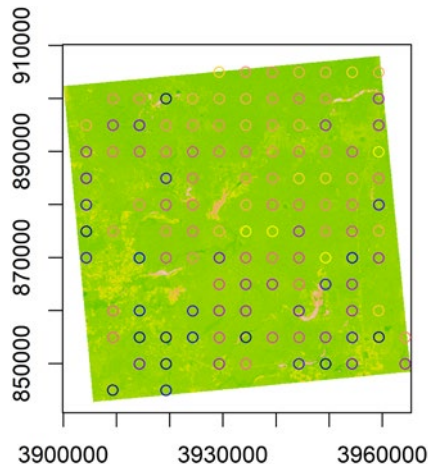
# class      : RasterStack
# dimensions : 2685, 2824, 7582440, 3 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)

```

```
# extent      : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs         : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+y_0=0 +ellps=GRS80 +units=m
# +no_defs
# names       :      HH,          HV,          NDBI
# min values  :    -27.88649,    -31.29588,    -1.78489
# max values  :     3.94951994,   -8.52579662,    0.06081292
```

Para verificar que los archivos de las imágenes de retrodispersión y los puntos centrales de los conglomerados se sobreponen, se obtiene la gráfica con ambas variables, la cual, una vez verificada, se le extrae. Esta vez se utiliza como ejemplo la capa **HH**.

```
plot(ListIMG_bcks$raster_files[[1]], legend= FALSE)
plot(points[,7], add = TRUE)
```



Para la extracción de los valores promedio de retrodispersión en una hectárea, se genera un archivo con polígonos de una hectárea a partir de los puntos centrales de los conglomerados, por lo que el radio para crear dichos polígonos circulares es de 56.41 m. Primero, para conservar los datos extraídos de las texturas y asegurarse de tener un objeto espacial, se utiliza la función **st\_as\_sf()** para convertir el

archivo *dataframe* “**data.p**” a un objeto espacial llamado *points*. Posteriormente, con la función **st\_buffer()**, que tiene como parámetros el objeto espacial *points* y el radio del polígono, se crea el archivo de polígonos “**points\_buffer**”. Aquí se muestra una gráfica de este archivo:

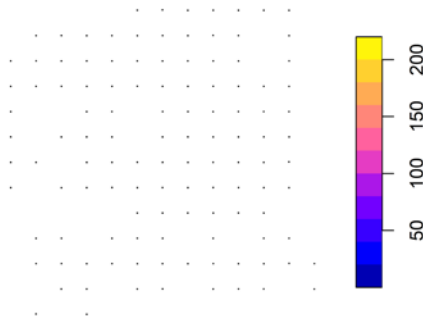
```
# Extraer valores promedio de backstactter en una Ha
```

```
points <- st_as_sf(data.p)
```

```
# Generar polígonos por buffer
```

```
points_buffer <- st_buffer(points, 56.41895835)
plot(points_buffer[,7])
```

**AGBconglomerado**



La extracción de información de las imágenes se realiza con la misma función que se utilizó para extraer las imágenes de las texturas, **extract(x, y, fun=mean, na.rm=TRUE, sp=TRUE, ...)**, donde **x** es un objeto de tipo *raster*, en este caso, un archivo multibanda con las imágenes de ALOS PALSAR y el **NDBI**; el parámetro (**y**) se refiere a un objeto espacial de polígonos; **na.rm** se utiliza para especificar que se utilizará una función en el método de extracción, en este caso, el promedio, y **sp** indica que los valores extraídos serán agregados al objeto espacial de polígonos. Con la función **summary()** se imprimen las características del objeto espacial, las coordenadas mínimas y máximas, la proyección geográfica, el número de polígonos (117),



así como las estadísticas básicas de los diferentes atributos: las variables que tenía el objeto espacial “*points*”, las texturas y ahora, además, las de valores de retrodispersión. Aquí se muestran las primeras (**X**, **Conglomerado**, **Anio\_mean**, **Diametro\_normal\_mean**) y las últimas variables (**HH**, **HV** y **NDBI**).

```
data.b <- extract(bcks, points_buffer,
  fun = mean,
  na.rm= TRUE,
  sp = TRUE)

summary(data.b)
# Object of class SpatialPolygonsDataframe
# Coordinates:
#           min           max
# x      3904267.5      3964387.8
# y      844880.2       905051.2
# Is projected: TRUE
# proj4string : [+proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102
+x_0=2500000 +y_0=0 +ellps=GRS80
# +towgs84=0,0,0,0,0,0,0 +units=m +no_defs]
# Data attributes:
#           X           Conglomerado           Anio_mean           Diametro_normal_mean
#   Min.    :56919      Min.    :56919      Min.    :2009           Min.    : 9.022
#   1st Qu.:57672      1st Qu.:57672      1st Qu.:2010           1st Qu.:12.685
#   Median :58187      Median :58187      Median :2011           Median :14.746
#   Mean    :58320      Mean    :58320      Mean    :2011           Mean    :14.483
#   3rd Qu.:59187      3rd Qu.:59187      3rd Qu.:2012           3rd Qu.:16.068
#   Max.    :59947      Max.    :59947      Max.    :2014           Max.    :19.046

#           NDBI_vari           HH           HV           NDBI
#   Min.    :2588      Min.    :-9.927      Min.    :-14.74      Min.    :-0.2931
#   1st Qu.:2814      1st Qu.: -8.404      1st Qu.: -13.04      1st Qu.: -0.2352
#   Median :2884      Median : -7.974      Median : -12.50      Median : -0.2215
#   Mean    :2888      Mean    : -8.005      Mean    : -12.58      Mean    : -0.2228
#   3rd Qu.:2967      3rd Qu.: -7.589      3rd Qu.: -12.02      3rd Qu.: -0.2075
#   Max.    :3137      Max.    : -6.587      Max.    : -10.92      Max.    : -0.1671
```

## 6.4 Conversión de datos espaciales a *dataframe*

En esta última parte del módulo, se convierte el archivo espacial con los datos extraídos a un *dataframe* y se guarda en disco; para ello, se utiliza la función `as.data.frame()`. Posteriormente, se crea el directorio en donde se guardará el archivo con los datos (“c:/Mapeo\_biomasa/FCP/RF/”) y que serán utilizados para obtener los modelos que relacionen la biomasa medida en campo con información de las imágenes. Finalmente, se utiliza la función `write.csv()` para guardar el archivo de datos con el nombre “datos.csv”.

# Convertir los datos a *dataframe* y guardarlos en CSV

```
data <- as.data.frame(data.b)

{Name.Dir.Lee <- "RF" # Nombre del directorio

  {ifelse(!
    dir.exists(file.path(folder,Name.Dir.Lee)),
    dir.create(file.path(folder, Name.Dir.Lee)),
    FALSE)
    dir.RF <- file.path(folder,Name.Dir.Lee)}
}

write.csv(data, file= paste(dir.RF,"DATOS.csv", sep = "/"))
```

## 7. Modelación y mapeo de la biomasa

EN ESTE MÓDULO SE CONSTRUIRÁ UN modelo para estimar la biomasa aérea del bosque usando Random Forest y como variables predictoras, la retrodispersión y la textura derivada de imágenes ALOS PALSAR. Para ello, se utilizó el paquete **ModelMap** de R (Freeman *et al.*, 2009). Además, la precisión de la estimación de la biomasa forestal se evaluará comparando los valores estimados de la biomasa con un conjunto independiente de datos de campo de las parcelas de inventario nacional forestal. Se utilizará el coeficiente de determinación ( $R^2$ ), así como la raíz del error cuadrático promedio (RMSE, por sus siglas en inglés) y el error cuadrático promedio relativo (%RMSE), para comparar los valores observados y predichos de la biomasa del bosque.

Por otro lado, se crearán mapas con la distribución espacial de la biomasa aérea del bosque y el coeficiente de variación de los valores de biomasa estimados usando el paquete **ModelMap**. Para ello, se utilizará el modelo de Random Forest obtenido, así como las capas de retrodispersión y textura de las imágenes incluidas en el modelo. El modelo Random Forest se realiza considerando la media de todos los árboles de la variable de respuesta, en este caso, la biomasa aérea. Por lo tanto, estas predicciones de árboles individuales también se pueden usar para mapear medidas de incertidumbre, como mapas del coeficiente de variación (dividiendo la desviación estándar por la media). Estos mapas proporcionan una visualización de regiones con mayor incertidumbre.

### 7.1 Construcción del modelo de Random Forest para estimar la biomasa

Para iniciar con la construcción del modelo para la estimación de la biomasa del bosque, se define el espacio de trabajo, que es el directorio donde se encuentra la base de datos de campo con las variables predictoras de retrodispersión y medidas de texturas derivadas de imágenes ALOS PALSAR, el directorio es: “**c:/Mapeo\_biomasa/FCP/RF/**”.

```
# Seleccionar el directorio de trabajo
```

```
setwd(choose.dir(caption = "Seleccione el directorio RF"))
folder <- getwd()
```

Se cargan los paquetes que se usarán en este módulo.

```
# Cargar paquetes que se utilizarán en este módulo
```

```
library(randomForest)
library(ModelMap)
library(raster)
library(sp)
library(spatial.tools)
library(prettymapr)
library(Metrics)
```

Del total de conglomerados del Inventario Nacional Forestal dispuesto para el área de estudio (117), el 70 % de las parcelas (82) servirá para calibrar el modelo, mientras que el 30 % de los conglomerados (35) se utilizará para la validación del modelo. Para llevar a cabo esta actividad, se crearán tres variables que guardarán los nombres de los archivos: el archivo con los datos de campo y los valores de retrodispersión y textura extraídos anteriormente (**DATOS.csv**), el nombre del archivo con los datos de calibración (**DATOS\_TRAIN.csv**), y el nombre del archivo con los datos de validación (**DATOS\_TEST.CSV**).

```
# Definición de tipos de archivos
```

```
qdatafn <- "DATOS.CSV" #archivo de datos
qdata.trainfn <- "DATOS_TRAIN.csv" # archivo de datos para entrenar el modelo
qdata.testfn <- "DATOS_TEST.csv" # archivo de datos para validar el modelo
```

Para dividir el archivo de datos original en aquellos que serán usados durante la calibración y la validación del modelo, se utiliza la función **get.test()** del paquete **ModelMap**. Esta función divide el grupo de datos de manera aleatoria y tiene varios parámetros: el primero especifica la proporción de los datos que permiten la validación (**proportion.test**), en este caso, se usará un 30 % y por ello, el valor de este parámetro es igual a 0.3; el parámetro **qdatafn**, tiene el nombre del archivo con los

datos originales; posteriormente, se define una semilla aleatoria (`seed=49`), y por último, se especifican los nombres de los archivos de datos de calibración y de validación que se delimitaron anteriormente (`qdata.trainfn` y `qdata.testfn`).

```
# Separar el conjunto de datos en dos grupos: datos de entrenamiento y de validación
# La proporción de los datos de validación es de un 30 %
# y se seleccionan de manera aleatoria
```

```
get.test(proportion.test=0.3,
         qdatafn=qdatafn,
         seed=49,
         folder=folder,
         qdata.trainfn=qdata.trainfn,
         qdata.testfn=qdata.testfn)
```

Se continúa con la definición de los parámetros que se utilizarán en el modelo y lo primero es decidir sobre el tipo de modelo. El paquete **ModelMap** tiene tres tipos de modelos distintos, pero la elección ha sido usar el Random Forest, por lo que se coloca **RF**. Adicionalmente, se debe especificar el nombre del archivo donde se guardará la salida del modelo; todos los archivos de salida tendrán el prefijo **AGB** como se especifica a continuación:

```
# Definición de los parámetros del modelo
```

```
model.type <- "RF" # el Modelo es de Random Forest
MODELfn.a <- "AGB"
```

Se define la lista de variables predictoras y cuáles de estas son categóricas. En el estudio de caso de este manual se tienen las capas de las imágenes con los valores de retrodispersión (**HH** y **HV**), el índice **NBDI** calculado con los valores de retrodispersión y las 24 medidas de textura, es decir, ocho medidas de textura calculadas para cada capa (**HH**, **HV** y **NBDI**), en total, 27 variables. Como ninguna de estas es categórica, este parámetro se define como **FALSE**.

```
predList <- c("HH_cont", "HH_corr", "HH_diss",
             "HH_entr", "HH_homo", "HH_mean",
             "HH_seco", "HH_vari", "HV_cont",
             "HV_corr", "HV_diss", "HV_entr",
```

```

"HV_homo", "HV_mean", "HV_seco",
"HV_vari", "NDBI_cont", "NDBI_corr",
"NDBI_diss", "NDBI_entr", "NDBI_homo",
"NDBI_mean", "NDBI_seco", "NDBI_vari",
"HH", "HV", "NDBI" )

```

```
predFactor <- FALSE
```

Se define entonces la variable de respuesta y si esta es continua, binaria o categórica. La variable de respuesta que se calculó con los datos de campo y está en el archivo (**DATOS.csv**) es **"AGBconglomerado"**, que corresponde a la biomasa de cada conglomerado calculada en toneladas por hectárea. Esta variable es continua.

```

response.name.a <- "AGBconglomerado" # variable de respuesta
response.type <- "continuous"

```

Los parámetros que restan por definir son la semilla aleatoria para el modelo y la columna de la base de datos que contiene un identificador único para cada renglón. Este identificador será utilizado para crear el archivo de salida con los valores observados y predichos de la biomasa cuando se ejecute la función para la validación del modelo, que en este caso es el conglomerado.

```

seed.a <- 25
unique.rowname <- "Conglomerado" #identificador único de cada renglón

```

Para construir el modelo se utiliza la función **model.build()** del paquete **ModelMap**. Esta función regresa como resultado un objeto con el modelo. Por otro lado, esta función guarda un archivo de texto con los parámetros que se utilizaron en el modelo. Como el prefijo de nombre de archivo que se utilizó fue **AGB**, entonces, este archivo de texto tiene por nombre **AGB\_model\_biulding\_arguments.txt** y se almacena en el directorio de trabajo establecido.

La función utilizada para construir el modelo tiene los siguientes parámetros que se establecieron arriba: 1) el tipo de modelo (**model.type**); 2) los datos de calibración (**qdata.trainfn**); 3) el directorio donde se guardan los resultados (**folder**);

4) la columna que contiene los identificadores únicos (**unique.rwname**); 5) el nombre del modelo con el que se guardarán los archivos de salida (**MODELfn**); 6) la lista de las variables predictoras (**predList**); 7) el tipo de datos de las variables predictoras (**predFactor**) (como no hay variables categóricas, colocamos **FALSE**); 8) el nombre de la variable de respuesta (**response.name**); 9) el tipo de datos de la variable de respuesta (**response.type**), y 10) la semilla aleatoria.

Después de ejecutar la función, se imprime el objeto **model.obj.AGB**. Se observa que se utilizaron 500 árboles en este modelo y que 18 variables fueron las más importantes. Además, se proporciona el promedio del cuadrado de los residuales (1891.53) que, al aplicarle la raíz cuadrada, se obtiene el RMSE = 43.49 ton ha<sup>-1</sup>. Por último, tenemos el porcentaje de variabilidad explicada por el modelo que, en este caso, es del 31.01 % (o bien R<sup>2</sup> = 0.31).

```

# Construcción del modelo usando Random Forest

model.obj.AGB <- model.build( model.type=model.type,
                             qdata.trainfn=qdata.trainfn,
                             folder=folder,
                             unique.rowname=unique.rowname,
                             MODELfn=MODELfn.a,
                             predList=predList,
                             predFactor=predFactor,
                             response.name=response.name.a,
                             response.type=response.type,
                             seed=seed.a,
                             na.action = "na.omit")

model.obj.AGB

# Call:
# randomForest(x = qdata.x, y = qdata.y, ntree = 500, mtry = 18,          replace =
TRUE, importance = TRUE, proximity =
# FALSE)
#
#           Type of random forest: regression
#           Number of trees: 500
# No. of variables tried at each split: 18
#
#           Mean of squared residuals: 1891.53
#           % Var explained: 31.03

```

## 7.2 Validación del modelo para estimar la biomasa

En la sección anterior se ajustó el modelo para estimar la biomasa del bosque. Aquí se hacen predicciones utilizando ese modelo y un conjunto de datos independiente (30 % de los conglomerados originales) con el objeto de evaluar la precisión de las estimaciones de la biomasa. La función `model.diagnostics()` permite realizar estas actividades. Esta función regresa un *dataframe* que tiene como variables el conglomerado, los valores observados y los valores predichos de biomasa, adicionalmente los guarda en un archivo `*.csv` que, en este caso, es (`AGB_pred.csv`) en el directorio de trabajo. Por otro lado, crea varios gráficos y tablas de resultados. Esta función recurre a algunos parámetros similares a la función `model.build()`. Los parámetros adicionales son: 1) el tipo de predicción (`prediction.type`) que, en este caso, es para los datos de validación (`TEST`), y 2) la especificación del formato en el que se grabarán los archivos de salida (`device.type`). Los archivos de salida corresponden a: 1) gráfica con la importancia de las variables en el modelo grabado con el nombre de archivo (`AGB_pred_importance.pdf`); 2) una gráfica de correlaciones entre las variables predichas (`AGB_pred_corrplot.pdf`); 3) el diagrama de dispersión entre los valores observados y predichos de biomasa (`AGB_pred_scatterplot.pdf`), y 4) los errores en el modelo en función del número de árboles (`AGB_pred_error.pdf`). Enseguida, se ejecuta la función `model.diagnostics()` con sus respectivos parámetros. Con la función `head()` se imprimen el nombre de las variables del *dataframe* “`model.pred.AGB`” y las observaciones de los primeros renglones.

# Validación del **Model** usando un conjunto de datos independiente

```
model.pred.AGB <- model.diagnostics( model.obj=model.obj.AGB,
                                   qdata.testfn=qdata.testfn,
                                   folder=folder,
                                   MODELfn=MODELfn.a,
                                   unique.rowname=unique.rowname,
                                   #parametros de la validación del modelo
                                   prediction.type="TEST",
                                   device.type=c("pdf"),
                                   cex=1.2)
```

```
head(model.pred.AGB)
```

#	Conglomerado	obs	pred
# 56919	56919	186.587765	141.2625



# 56920	56920	141.552960	137.2267
# 56922	56922	173.282656	136.6961
# 56924	56924	181.995796	143.7902
# 57169	57169	153.942418	103.0285
# 57171	57171	5.297607	38.3412

A pesar de que en el archivo **AGB\_pred\_scatterplot.pdf** se presenta el diagrama de dispersión entre los valores observados y predichos de biomasa, se optó por realizar el cálculo de diferentes medias de error como el coeficiente de determinación ( $R^2$ ), así como el error cuadrático promedio (RMSE), y el error cuadrático promedio relativo (%RMSE), además, se obtuvo el diagrama de dispersión entre los valores observados y predichos de la biomasa, así como la regresión lineal entre ambas variables.

```
rmse.val <- rmse(model.pred.AGB$obs, model.pred.AGB$pred)
mean.obs <- mean(model.pred.AGB$obs)
percent.rmse <- (rmse.val / mean.obs) * 100
lm.Val <- lm(obs ~ pred, data = model.pred.AGB)
```

Para observar el coeficiente de determinación del modelo de regresión (0.408), los coeficientes de los parámetros y la demás información del modelo de regresión, se imprime un resumen del objeto **lm.val**, resultado de la regresión lineal, con la función **summary()**.

```
summary(lm.Val)

# Call:
# lm(formula = obs ~ pred, data = model.pred.AGB)

# Residuals:
#      Min       1Q   Median       3Q      Max
# -81.102  -29.392   -0.024   33.827  104.224
# Coefficients:
#      Estimate Std. Error t value Pr(>|t|)
# (Intercept)  -9.2443    26.8225  -0.345   0.733
# pred          1.0742     0.2252   4.770 3.63e-05 ***
```

```
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Residual standard error: 41.85 on 33 degrees of freedom
# Multiple R-squared:  0.4081,    Adjusted R-squared:  0.3901
# F-statistic: 22.75 on 1 and 33 DF,  p-value: 3.631e-05
```

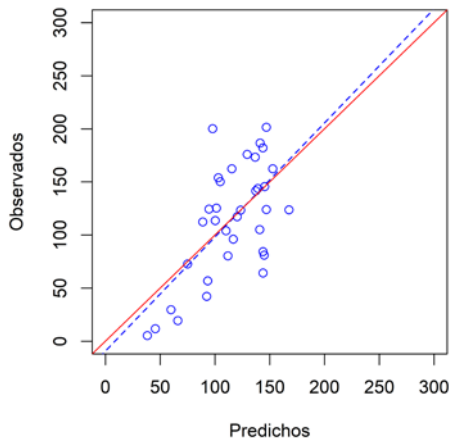
Se imprimen los valores de **RMSE**, el promedio de la biomasa de los valores observados y el error relativo (%**RMSE**).

```
cbind(rmse.val,mean.obs,percent.rmse)

#      rmse.val mean.obs percent.rmse
# [1,] 40.70856 114.1595     35.65937
```

Finalmente, en esta sección se obtiene el diagrama de dispersión entre los valores observados y predichos de la biomasa.

```
plot(x=model.pred.AGB$pred, y=model.pred.AGB$obs, xlab = "Predichos", ylab =
"Observados",col="blue", lwd=1,xlim=c(0,300), ylim=c(0,300))
abline(lm.Val,col="blue", lwd=1.2, lty=2)
abline(0,1,col="red", lty=1, lwd=1)
```



### 7.3 Crear un mapa tipo *raster* con la estimación de la biomasa

Para obtener un mapa con la distribución de la biomasa aérea en el área estudiada, se utiliza el modelo de Random Forest generado en el apartado 7.1, así como las capas de retrodispersión y textura de las imágenes de ALOS PALSAR. Para realizar esta actividad, primero se asigna en el directorio donde se encuentran los datos de retrodispersión “C:/Mapeo\_biomasa/FCP/IMG/” a la variable “`dir.IMG.bcks`” y el directorio donde se encuentran las texturas “C:/Mapeo\_biomasa/FCP/IMG/IMG\_texturas” a la variable “`dir.IMG.txtr`”.

```
# Directorio donde se encuentran las imágenes usadas como predictores en el modelo
```

```
dir.IMG.bcks <- "C:/Mapeo_biomasa/FCP/IMG" #backscatters
dir.IMG.txtr <- "C:/Mapeo_biomasa/FCP/IMG/IMG_texturas" #texturas
```

Se crea una lista de las imágenes con las polarizaciones **HH** y **HV**, así como las del índice **NDBI** (archivos con la terminación `*.img`), utilizando la función `list.raster.files()`, de la misma forma que se ha hecho anteriormente. Al imprimir el contenido de la lista se puede observar que esta tiene tres archivos *raster*, que son las capas con las polarizaciones **HH** y **HV**, así como el índice **NDBI**. En la lista se observan las propiedades de cada una de las imágenes *raster*.

```
# Crear una lista con las imágenes de retrodispersión (Backscatter)
```

```
bcks <- list.raster.files(path = dir.IMG.bcks, pattern = ".img$",
                        return_rasters = TRUE, return_bbox = FALSE)
bcks$raster_files[]

# [[1]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent      : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs         : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+ y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source      : C:/Mapeo_biomasa/FCP/IMG/HH.img
# names       :      HH
# min values  : -27.88649
# max values  :  3.94952
```

```

# [[2]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs        : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+ y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/HV.img
# names      : HV
# min values : -31.29588
# max values : -8.525797

# [[3]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs        : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+ y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/NDBI.img
# names      : NDBI
# min values : -1.78489
# max values : 0.06081292

```

En seguida, se crea un *dataframe* con el nombre de “**rastLUTfn.a**”, que contiene tres columnas: la primera, con el directorio y nombre del archivo de la imagen (**\*.img**); la segunda columna es el nombre de la variable predictora, y por último, el número de capa del *raster* que corresponde a la variable predictora. Al imprimir este *dataframe* se observa la información de las tres imágenes con las tres columnas (**V1**, **V2** y **V3**).

# Crear un *dataframe* de imágenes de *backscatter*

```

rastLUTfn.a <- data.frame()
for (i in 1: length(bcks$raster_files)){

```

```

rastLUTfn.a[i,1]<- paste(dir.IMG.bcks,"/",
                        names(bcks$raster_files[[i]]),".img", sep = "")
rastLUTfn.a[i,2]<- names(bcks$raster_files[[i]])
rastLUTfn.a[i,3]<- 1 #bands number
}
rastLUTfn.a

# 1      C:/Mapeo_biomasa/FCP/IMG/HH.img      HH      1
# 2      C:/Mapeo_biomasa/FCP/IMG/HV.img      HV      1
# 3      C:/Mapeo_biomasa/FCP/IMG/NDBI.img    NDBI    1

```

Se realiza un procedimiento similar para las imágenes de textura, primero se crea la lista de las imágenes de textura con la función **list.raster.files()**. Al imprimir el contenido de la lista se puede observar que esta tiene 24 archivos *raster*, que corresponden a ocho medidas de textura para cada una de las capas utilizadas (**HH**, **HV** y **NDBI**). Aquí solo se muestran las tres primeras medidas de textura de la capa **HH**; de igual manera, se observan las propiedades de cada una de las imágenes *raster*.

```

# Crear una lista con las imágenes de textura

txtr <- list.raster.files(path = dir.IMG.txtr, pattern = ".img$",
                        return_rasters = TRUE, return_bbox = FALSE)
txtr$raster_files

# [[1]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs       : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+ y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_cont.img
# names      : HH_cont
# min values :      0
# max values : 289.8889

```

```

# [[2]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs        : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+ y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_corr.img
# names      : HH_corr
# min values : -0.7375351
# max values : 70.31269

# [[3]]
# class      : RasterBrick
# dimensions : 2685, 2824, 7582440, 1 (nrow, ncol, ncell, nlayers)
# resolution : 23.2, 24.4 (x, y)
# extent     : 3899900, 3965417, 842677.8, 908191.8 (xmin, xmax, ymin, ymax)
# crs        : +proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12 +lon_0=-102 +x_0=2500000
+ y_0=0 +ellps=GRS80 +units=m
# +no_defs
# source     : C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_diss.img
# names      : HH_diss
# min values : 0
# max values : 13.66667

```

Se crea un *dataframe* con el nombre de “**rastLUTfn.b**”, que contiene tres columnas: la primera, con el directorio y el nombre del archivo de la imagen (**\*.img**); la segunda columna es el nombre de la variable predictora, y por último, el número de capas correspondiente a la variable. Se utiliza la función **head()** para imprimir las primeras observaciones del *dataframe* que contiene la información de las 24 imágenes.

# Crear el *dataframe* de las imágenes de textura

```

rastLUTfn.b <- data.frame()
for (i in 1: length(txtr$raster_files)){
  rastLUTfn.b[i,1]<- paste(dir.IMG.txtr,"/",
                          names(txtr$raster_files[[i]]),".img", sep = "")
}

```

```

rastLUTfn.b[i,2]<- names(txtr$raster_files[[i]])
rastLUTfn.b[i,3]<- 1
}
head(rastLUTfn.b)

```

#	V1	V2	V3
# 1	C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_cont.img	HH_cont	1
# 2	C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_corr.img	HH_corr	1
# 3	C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_diss.img	HH_diss	1
# 4	C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_entr.img	HH_entr	1
# 5	C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_homo.img	HH_homo	1
# 6	C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_mean.img	HH_mean	1

Se crea un nuevo *dataframe* con la unión de los dos *dataframe* anteriores para un archivo con la ubicación de cada una de las imágenes que se utilizan como variables explicativas en el modelo de Random Forest. Este archivo es de gran utilidad para obtener el mapa de biomasa. Se utiliza la función **head()** para imprimir las primeras observaciones del *dataframe* que contiene la información de las 27 imágenes. Como puede observarse, las tres primeras corresponden a las imágenes de retrodispersión y las siguientes tres son medidas de textura de la capa **HH**.

# Unir los dos *dataframes*

```

rastLUTfn <- rbind(rastLUTfn.a, rastLUTfn.b)
head(rastLUTfn)

```

#	V1	V2	V3
# 1	C:/Mapeo_biomasa/FCP/IMG/HH.img	HH	1
# 2	C:/Mapeo_biomasa/FCP/IMG/HV.img	HV	1
# 3	C:/Mapeo_biomasa/FCP/IMG/NDBI.img	NDBI	1
# 4	C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_cont.img	H_cont	1
# 5	C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_corr.img	HH_corr	1
# 6	C:/Mapeo_biomasa/FCP/IMG/IMG_texturas/HH_diss.img	HH_diss	1

Se crea el directorio de salida en donde se guardarán los mapas de biomasa e incertidumbre. Dicho directorio es “**c:/Mapeo\_biomasa/FCP/RF/MAPS**”.

```
# Crear el directorio de salida
```

```
{Name.Dir.MAPS <- "MAPS"}
{ifelse(!
  dir.exists(file.path(folder, Name.Dir.MAPS)),
  dir.create(file.path(folder, Name.Dir.MAPS)), FALSE
)}
dir.MAPS <- file.path(folder, Name.Dir.MAPS)}
}
```

Se define el nombre del mapa de salida que, en este caso, es (AGB.img) y se agrega el directorio en el que se guardará, que corresponde al directorio creado recientemente (**dir.MAPS**).

```
# Definir el nombre del mapa de salida
```

```
OUTPUTfn <- "AGB.img"
OUTPUTfn <- paste(dir.MAPS, OUTPUTfn, sep="/")
```

Para obtener el mapa de biomasa y sus respectivos mapas de incertidumbre, se utiliza la función **model.mapmake()** del paquete **ModelMap**. Esta función tiene varios parámetros, aquí se utilizaron: 1) el objeto modelo para la predicción (**model.obj**) que fue obtenido en la sección 7.1, en este caso es **model.obj.AGB**, que usa un modelo de Random Forest; 2) el directorio de salida definido en 5.1; 3) el *dataframe* que contiene la ubicación de las imágenes que se utilizan como variables explicativas (**rastLUTfn**); 4) el nombre del mapa de salida y su ubicación (**OUTPUTfn**); 5) el nombre del prefijo que tendrán todos los archivos de salida definido en la sección 7.1 como **MODELfn.a =AGB**; 6) el último parámetro se especifica si se quiere obtener el mapa de desviaciones estándar y de la media para el cálculo de un mapa de coeficiente de variación.

Esta función genera diferentes archivos de salida: 1) un archivo de texto (**AGB\_projections**) que tiene la proyección geográfica de las 27 capas utilizadas como predictores que, en este caso, tienen el sistema de coordenadas **CCL Datum ITRF2008**; 2) cuatro archivos de imágenes (**\*.img**), el primero, con la distribución espacial de la biomasa aérea (**AGB.img**), mientras que los otros tres tienen la media, la desviación estándar y el coeficiente de variación de la biomasa forestal (**ABG\_mean.img**, **AGB\_stdev.img**, **AGB\_coefv.img**).



```

# Crear una imagen con el mapa predicho de biomasa

model.mapmake( model.obj=model.obj.AGB,
               folder=folder,
               rastLUTfn=rastLUTfn,
               # parámetros del mapa
               OUTPUTfn=OUTPUTfn,
               MODELfn= MODELfn.a,
               map.sd = TRUE
)

# [1] "model.NA.ACTION: NULL"
# [1] "Using default 'na.action' of \"na.omit\""
# [1] "starting production prediction"
# [1] "predFactor: "
# [1] "all predictor layer rasters match"
# .
# .
# .
# [1] "ARGfn: C:/Mapeo_biomasa/FCP/RF/MAPS/AGB_mapmake_arguments.txt"

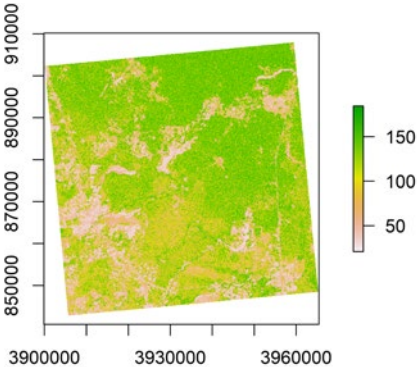
```

Se utiliza la función **raster()** del paquete para crear un objeto de tipo *raster* llamado **mapgrid** que se lee del archivo especificado en **OUTPUTfn**; posteriormente, se grafica.

```

mapgrid <- raster(OUTPUTfn)
plot(mapgrid)

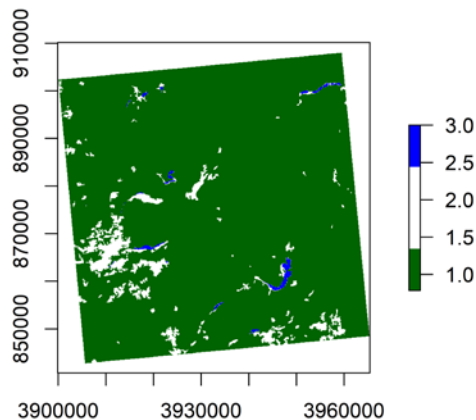
```



En el mapa de biomasa que se acaba de crear, no están excluidas las áreas que no tienen bosque, ni los cuerpos de agua. Para obtener un mapa más realista de la biomasa del bosque se deben eliminar estas áreas. Para ello, se requiere de un mapa que identifique las áreas de bosque, las que están desprovistas de bosque y los cuerpos de agua. Se recordará que cuando se hizo la descarga de las imágenes de ALOS PALSAR, se descargaron estos mapas (fig. 5 en el capítulo 4). Para realizar este proceso, primero se utiliza la función *raster* para crear el objeto *raster* “FNF” y leer el archivo con esta imagen (**FNF.tiff**) ubicada en el directorio “**c:/Mapeo\_biomasa/FCP/IMG/**”. Además, se genera una gráfica con el mismo.

```
# Crear una máscara con áreas de bosque y áreas de no bosque y aplicarla al mapa
FNF <- raster(choose.files(caption = "select raster file FNF", multi = FALSE))

my_col <- c("darkgreen", "white", "blue")
plot (FNF, legend = TRUE, col = my_col)
      legend = c("Forest", "NoForest", "Water bodies"),
      fill = my_col)
```

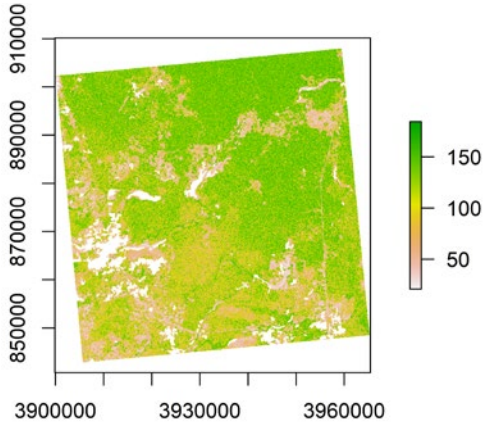


Para excluir las estimaciones de biomasa en las áreas de no bosque y cuerpos de agua, se utilizará la función **mask()** del paquete **raster**. Esta función crea un nuevo objeto *raster* con los mismos valores del objeto *raster*, excepto que, para los píxeles

del objeto *raster* máscara marcados con un valor específico, se actualizarán con un valor nulo. Esto es, la función `mask(x, mask, maskvalue, ..)` se utilizará con tres parámetros: *x* es el objeto *raster* en donde se aplicará la máscara, es decir, el mapa de biomasa; *mask* es otro objeto *raster* clasificado con diferentes categorías, en este caso, con áreas de bosque, no bosque y cuerpos de agua; y *maskvalue* es el valor en el objeto de la máscara, que indica los pixeles que serán actualizados en el objeto *raster* *x* con un valor nulo. En este caso, se pondrán en nulo las áreas de no bosque = 2 y los cuerpos de agua = 3. Por último, se grafica el objeto *raster* en donde se aplicó la máscara para su visualización.

```
mapgrid.mask <- mask(mapgrid,
                    FNF, maskvalue= 2)
mapgrid.mask <- mask(mapgrid.mask,
                    FNF, maskvalue= 3)

plot(mapgrid.mask)
```



Para finalizar con esta sección, se realiza una proyección del *raster* en donde se aplicó la máscara “`mapgrid.mask`” de las coordenadas CCL Datum ITRF2008 a coordenadas geográficas WGS84 usando la función `projectRaster()`. Además, se guarda este objeto *raster* en disco con el nombre `ABG_FCP.img` usando la función `writeRaster()`.

```
# Proyectar las coordenadas del mapa máscara a WGS84
mapgrid.a <- projectRaster(mapgrid.mask,
  crs= "+proj=longlat +ellps=WGS84 +datum=WGS84 +no_defs",
  method = "ngb")
```

```
# Guardar en disco el mapa creado
writeRaster(mapgrid.a,
  paste(dir.MAPS, "AGB_FCP.img", sep= "/"),
  overwrite= TRUE
)
```

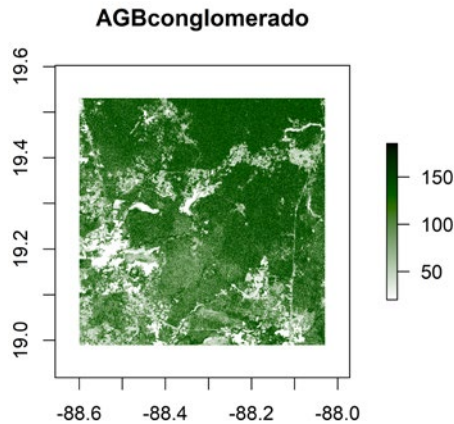
## 7.4 Impresión de mapas

En esta última parte del módulo de modelación y mapeo de la biomasa, se imprimen los mapas de biomasa y el de coeficiente de variación. Se imprime el mapa “**mapgrid.a**” que fue proyectado a coordenadas geográficas WGS84. Para ello, se define la paleta de colores, se identifican los valores máximo y mínimo en el objeto *raster*. Por último, se imprime el mapa usando la función **plot()** con los siguientes parámetros: 1) el objeto *raster* a imprimir que es “**mapgrid.a**”; 2) la paleta de colores utilizados “**col.map**”; 3) el título principal que, en este caso, es el mismo de la variable de respuesta definido en la sección 7.1 “**response.name.a**”, y 4) el rango de valores del objeto *raster*.

```
# Definir la paleta de colores
l <- seq(100,0,length.out=101)
c <- seq(0,100,length.out=101)
col.ramp <- hcl(h = 120, c = c, l = l)

r.range <- c(minValue(mapgrid.a), maxValue(mapgrid.a))

# Imprimir
plot(mapgrid.a,
  col = col.ramp,
  main = response.name.a,
  zlim = r.range)
```



Para finalizar con esta sección, se imprime el mapa del coeficiente de variación para identificar aquellas áreas con mayores errores al momento de hacer las predicciones. Para ello, se lee la imagen que tiene el coeficiente de variación (**AGB\_coefv.img**) ubicada en el directorio “**c:/Mapeo\_biomasa/FCP/RF/MAPS**”. Se aplicará una máscara al igual que en la sección 5.4. Se proyecta a coordenadas geográficas WSG84. Se define la paleta de colores, se imprime y se guarda en disco. Se puede observar que aquellas áreas cercanas a donde no existe bosque son las que tienen un mayor error.

```
# Generar mapa de incertidumbre

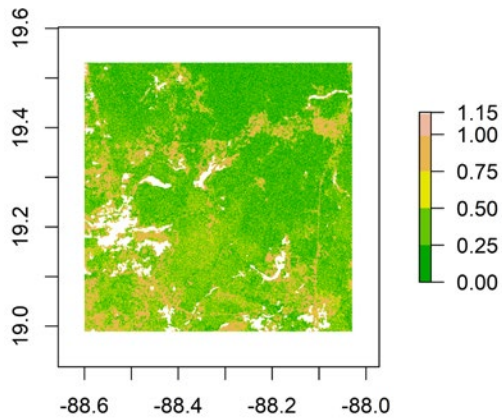
coefv <- raster(choose.files(caption="Seleccionar raster Coeficiente de Variacion"))

coefv.mask <- mask(coefv,
                  FNF, maskvalue= 2)
coefv.mask <- mask(coefv.mask,
                  FNF, maskvalue= 3)

# conversión a WGS84
coefv.a <- projectRaster(coefv.mask,
                        crs= "+proj=longlat +ellps=WGS84 +datum=WGS84 +no_defs",
                        method = "ngb")

ramp <- colorRampPalette(c("green4","yellow"))
plot(coefv.a, breaks = c(0,0.25,0.50,0.75,1, round(maxValue(coefv.a),2)),
```

```
col=terrain.colors(6))  
  
writeRaster(coefv.a,  
            paste(dir.MAPS, "AGB_FCP_coefv.img", sep= "/"),  
            overwrite= TRUE  
)
```



## Consideraciones finales

EN ESTE MANUAL SE PRESENTÓ UNA metodología para mapear la biomasa área de los bosques, basada en la relación entre datos de campo e imágenes de satélite procesadas con el *software* R. Sin embargo, no se puso énfasis en las fuentes de error para mejorar la precisión de las predicciones de la biomasa. Adicionalmente, los mapas disponibles a escala continental y nacional tienen mucha incertidumbre, particularmente en áreas tropicales con altas densidades de vegetación, como es el caso de los bosques tropicales (Rodríguez-Veiga *et al.*, 2019; Avitabile *et al.*, 2016). Los errores en los mapas de biomasa están relacionados con la calidad de los datos de campo que se utilizan para calibrar imágenes de satélite y con la capacidad de las imágenes para mapear la biomasa.

Los datos de campo del Inventario Nacional Forestal y de Suelos tienen algunas limitaciones, una de ellas está relacionada con el diámetro mínimo utilizado. En este inventario solamente se midieron árboles con diámetros > 7.5 cm (Conafor, 2013). La exclusión de los árboles pequeños en el inventario es particularmente importante en los bosques tropicales secos, ya que una gran proporción de árboles son pequeños (Romeo-Duque *et al.*, 2007). Los tallos pequeños (< 10 cm diámetro) representan entre 15 % y 40 % de la biomasa de los bosques maduros de este tipo de selvas en México, y hasta 80 % de la biomasa en bosques jóvenes (Jaramillo *et al.*, 2003; Read *et al.*, 2003). Por lo tanto, los datos de campo de este inventario podrían estar subestimando la biomasa de estos bosques. Otra limitación del Inventario Nacional Forestal y de Suelos es el tiempo que tomó recopilar los datos de las parcelas en campo. El último inventario se realizó entre 2009 y 2014, dentro de este intervalo de tiempo de seis años pudieran ocurrir cambios significativos en la biomasa, principalmente por procesos de crecimiento, reclutamiento y mortalidad de árboles. Para reducir el error debido a estos cambios, Hernández-Stefanoni *et al.* (2020) calibraron las parcelas del Inventario Nacional Forestal con datos de los sitios de monitoreo intensivo de carbono en donde se midieron árboles pequeños. Además, para solucionar las diferencias temporales entre el inventario de las parcelas y la adquisición de imágenes, se utilizaron datos de cambios en biomasa a lo largo del tiempo usando

cronosecuencias de los principales ecosistemas en la península de Yucatán, y con ello, lograron reducir el error de estimación en un 13 %.

En este manual, para mejorar la precisión de los mapas de biomasa y superar problemas como la escasez de imágenes a causa de la nubosidad y los relacionados con la saturación de las imágenes, se usaron tanto los datos de retrodispersión, como métricas de textura de las imágenes de Radar. Sin embargo, en estudios recientes se han utilizado otras fuentes de datos que permiten reducir los errores de estimación. Por ejemplo, utilizado datos climáticos (Hernández-Stefanoni *et al.*, 2020) o bien, datos topográficos derivados de modelos de elevación digital (Cartus *et al.*, 2014). Por último, se han utilizado datos LiDAR para incrementar el tamaño muestra, usando un procedimiento de escalamiento en dos etapas, es decir, de los datos de campo a las áreas cubiertas por LiDAR y utilizando estimaciones de biomasa en las áreas de LiDAR para calibrar la información de imágenes de satélite que cubran toda el área (Wikes *et al.*, 2015; Luther *et al.*, 2019).

Las observaciones anteriores nos indican que el mapeo de la biomasa es un área activa para los investigadores que buscan mejorar la precisión de las estimaciones de la biomasa área.



# Referencias

- Almeida-Filho, R., Y. E. Shimabukuro, A. Rosenqvist, & G. A. Sanchez. 2009. Using dual-polarized ALOS PALSAR data for detecting new fronts of deforestation in the Brazilian Amazônia. *International Journal of Remote Sensing*, 30(14), 3735-3743.
- Avitabile, V., M. Herold, G. B. Heuvelink, S. L. Lewis, O. L. Phillips, G. P. Asner, ... & N. J. Berry. 2016. An integrated pan-tropical biomass map using multiple reference datasets. *Global Change Biology*, 22(4), 1406-1420.
- Baccini, A. G. S. J., S. J. Goetz, W. S. Walker, N. T. Laporte, M. Sun, D. Sulla-Menashe, ... & S. Samanta. 2012. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*, 2(3), 182-185.
- Banin, L., T. R. Feldpausch, O. L. Phillips, T. R. Baker, J. Lloyd, K. Affum-Baffoe, ... & S. Davies. 2012. What controls tropical forest architecture? Testing environmental, structural and floristic drivers. *Global Ecology and Biogeography*, 21(12), 1179-1190.
- Burrough, P. A. 2001. GIS and geostatistics: Essential partners for spatial analysis. *Environmental and Ecological Statistics*, 8(4), 361-377.
- Bastin, J. F., A. Fayolle, Y. Tarelkin, J. Van den Bulcke, T. De Haulleville, F. Mortier, ... & C. De Cannière. 2015. Wood specific gravity variations and biomass of central African tree species: The simple choice of the outer wood. *PloS One*, 10(11), e0142146.
- Cairns, M. A., S. Brown, E. H. Helmer, & G. A. Baumgardner. 1997. Root biomass allocation in the world's upland forests. *Oecologia*, 111(1), 1-11.
- Cairns, M. A., I. Olmsted, J. Granados, & J. Argaez. 2003. Composition and aboveground tree biomass of a dry semi-evergreen forest on Mexico's Yucatan Peninsula. *Forest Ecology and Management*, 186(1-3), 125-132.
- Cartus, O., J. Kellndorfer, W. Walker, C. Franco, J. Bishop, L. Santos, & J. M. M. Fuentes. 2014. A national, detailed map of forest aboveground carbon stocks in Mexico. *Remote Sensing*, 6(6), 5559-5588.
- Comisión Nacional Forestal. 2013. *Inventario Nacional Forestal y de Suelos, Procedimientos de muestreo*. Conafor, México.

- Chave, J., R. Condit, S. Lao, J. P. Caspersen, R. B. Foster, & S. P. Hubbell. 2003. Spatial and temporal variation of biomass in a tropical forest: results from a large census plot in Panama. *Journal of Ecology*, 91(2), 240-252.
- Chave, J., M. Réjou-Méchain, A. Búrquez, E. Chidumayo, M. S. Colgan, W. B. Delitti, ... & M. Henry. 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology*, 20(10), 3177-3190.
- Fahrig, L. 2003. Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 34(1), 487-515.
- Fearnside, P. M., & W. F. Laurance. 2004. Tropical deforestation and greenhouse-gas emissions. *Ecological Applications*, 14(4), 982-986.
- Fischer, R., N. Knapp, F. Bohn, H. H. Shugart, & A. Huth. 2019. The relevance of forest structure for biomass and productivity in temperate forests: New perspectives for remote sensing. *Surveys in Geophysics*, 40(4), 709-734.
- Flores, J. & I. Espejel. 1994. *Tipos de vegetación de la península de Yucatán*. Universidad Autónoma de Yucatán. Mérida. pp. 73-80. (Etnoflora yucatanense, 3).
- Frangi, J. L., & A. E. Lugo. 1985. Ecosystem dynamics of a subtropical floodplain forest. *Ecological Monographs*, 55(3), 351-369.
- Freeman, E., T. Frescino, & G. Moisen. 2009. *ModelMap: An R package for modeling and map production using Random Forest and Stochastic Gradient Boosting*. USDA Forest Service, Rocky Mountain Research Station, 507.
- Foody, G. M., D. S. Boyd, & M. E. Cutler. 2003. Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. *Remote Sensing of Environment*, 85(4), 463-474.
- Galicia, L., V. Saynes, & J. Campo. 2015. Biomasa aérea, biomasa subterránea y necromasa en una cronosecuencia de bosques templados con aprovechamiento forestal. *Botanical Sciences*, 93(3), 473-484.
- García, M., S. Saatchi, S. Ustin, & H. Balzter. 2018. Modelling forest canopy height by integrating airborne LiDAR samples with satellite Radar and multispectral imagery. *International Journal of Applied Earth Observation and Geoinformation*, 66, 159-173.
- Gehring, C., S. Park, & M. Denich. 2004. Liana allometric biomass equations for Amazonian primary and secondary forest. *Forest Ecology and Management*, 195(1-2), 69-83.
- Guyot, J. 2011. *Estimation du stock de carbone dans la végétation des zones humides de la Péninsule du Yucatan*. Mémoire de fin d'études. AgroParis Tech-El Colegio de la Frontera Sur, France.
- Haralick, R. M., K. Shanmugam, & I. H. Dinstein. 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, (6), 610-621.

- Hernández-Stefanoni, J. L., J. M. Dupuy, K. D. Johnson, R. Birdsey, F. Tun-Dzul, A. Peduzzi, ... & D. López-Merlín. 2014. Improving species diversity and biomass estimates of tropical dry forests using airborne LiDAR. *Remote Sensing*, 6(6), 4741-4763.
- Hernández-Stefanoni, J. L., G. Reyes-Palomeque, M. Á. Castillo-Santiago, S. P. George-Chacón, A. H. Huechacona-Ruiz, F. Tun-Dzul, ... & J. M. Dupuy. 2018. Effects of sample plot size and GPS location errors on aboveground biomass estimates from LiDAR in tropical dry forests. *Remote Sensing*, 10(10), 1586.
- Hernández-Stefanoni, J. L., M. Á. Castillo-Santiago, J. F. Mas, C. E. Wheeler, J. Andres-Mauricio, F. Tun-Dzul, ... & J. M. Dupuy. 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.
- Houghton, R. A., B. Byers, & A. A. Nassikas. 2015. A role for tropical forests in stabilizing atmospheric CO<sub>2</sub>. *Nature Climate Change*, 5(12), 1022-1023.
- Huang, H., C. Liu, X. Wang, X. Zhou, & P. Gong. 2019. Integration of multi-resource remotely sensed data and allometric models for forest aboveground biomass estimation in China. *Remote Sensing of Environment*, 221, 225-234.
- Hulshof, C. M., N. G. Swenson, & M. D. Weiser. 2015. Tree height–diameter allometry across the United States. *Ecology and Evolution*, 5(6), 1193-1204.
- Islebe, G., N. Torrescano-Valle, M. Valdez-Hernández, M. Tuz-Novelo, & H. Weissenberger. 2009. Efectos del impacto del huracán Dean en la vegetación del sureste de Quintana Roo, México. *Foresta Veracruzana*, 11(1):1-6
- INEGI. (2011). Geografía, Geodesia. Obtenido de INEGI: [http://www.inegi.org.mx/geo/contenidos/geodesia/doc/el\\_cambio\\_de\\_itr.pdf](http://www.inegi.org.mx/geo/contenidos/geodesia/doc/el_cambio_de_itr.pdf)
- Jaramillo, V. J., J. B. Kauffman, L. Rentería-Rodríguez, D. L. Cummings, & L. J. Ellingson. 2003. Biomass, carbon, and nitrogen pools in Mexican tropical dry forest landscapes. *Ecosystems*, 6(7), 609-629.
- Joshi, N., E. T. Mitchard, M. Brolly, J. Schumacher, A. Fernández-Landa, V. K. Johannsen, ... & R. Fensholt. 2017. Understanding ‘saturation’ of radar signals over forests. *Scientific Reports*, 7(1), 1-11.
- Keith, H., B. G. Mackey, & D. B. Lindenmayer. 2009. Re-evaluation of forest biomass carbon stocks and lessons from the world’s most carbon-dense forests. *Proceedings of the National Academy of Sciences*, 106(28), 11635-11640.
- Lee, J. S. 1980. Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2), 165-168.
- Le Quéré, C., R. Moriarty, R. M. Andrew, J. G. Canadell, S. Sitch, J. I. Korsbakken, ... & R. A. Houghton. 2015. Global carbon budget 2015. *Earth System Science Data*, 7(2), 349-396.

- Lefsky, M. A., W. B. Cohen, & R. A. Spies. 2001. An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Canadian Journal of Forest Research*, 31, 78-87.
- Lu, D., Q. Chen, G. Wang, E. Moran, M. Batistella, M. Zhang, ... & D. Saah. 2012. Aboveground forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates. *International Journal of Forestry Research*, 2012.
- Lu, D., Q. Chen, G. Wang, L. Liu, G. Li, & E. Moran. 2016. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *International Journal of Digital Earth*, 9(1), 63-105.
- Luther, J. E., R. A. Fournier, O. R. van Lier, & M. Bujold. 2019. Extending ALS-based mapping of forest attributes with medium resolution satellite and environmental data. *Remote Sensing*, 11(9), 1092.
- Mas, J. F. 2018. *Análisis Espacial Con R: Usa R como un sistema de información geográfica*. European Scientific Institute Publishing, 105 p.
- Mermoz, S., M. Réjou-Méchain, L. Villard, T. Le Toan, V. Rossi, & S. Gourlet-Fleury. 2015. Decrease of L-band SAR backscatter with biomass of dense forests. *Remote Sensing of Environment*, 159, 307-317.
- Mitchard, E. T., S. S. Saatchi, L. White, K. Abernethy, K. L. Jeffery, S. L. Lewis, ... & P. Meir. 2012. Mapping tropical forest biomass with radar and spaceborne LiDAR in Lopé National Park, Gabon: overcoming problems of high biomass and persistent cloud. *Biogeosciences*, 9(1), 179-191.
- Molto, Q., V. Rossi, & L. Blanc. 2013. Error propagation in biomass estimation in tropical forests. *Methods in Ecology and Evolution*, 4(2), 175-183.
- Pan, Y., R. A. Birdsey, J. Fang, R. Houghton, P. E. Kauppi, W. A. Kurz, ... & P. Ciais. 2011. A large and persistent carbon sink in the world's forests. *Science*, 333(6045), 988-993.
- Ploton, P., N. Barbier, P. Couteron, C. M. Antin, N. Ayyappan, N. Balachandran, ... & V. Droissart. 2017. Toward a general tropical forest biomass prediction model from very high-resolution optical satellite images. *Remote Sensing of Environment*, 200, 140-153.
- Ramírez Ramírez, G., J. M. Dupuy Rada, L. Ramírez y Avilés, & F. J. Solorio Sánchez. 2019. Evaluación de ecuaciones alométricas de biomasa epigea en una selva mediana subcaducifolia de Yucatán. *Madera y Bosques*, 23(2), 163-179.
- Read, L., & D. Lawrence. 2003. Recovery of biomass following shifting cultivation in dry tropical forests of the Yucatan. *Ecological Applications*, 13(1), 85-97.
- Réjou-Méchain, M., A. Tanguy, C. Piponiot, J. Chave, & B. Hérault. 2017. Biomass: An R package for estimating above-ground biomass and its uncertainty in tropical forests. *Methods in Ecology and Evolution*, 8(9), 1163-1167.

- Réjou-Méchain, M., N. Barbier, P. Couteron, P. Ploton, G. Vincent, M. Herold, ... & J. B. Féret. 2019. Upscaling forest biomass from field to satellite measurements: sources of errors and ways to reduce them. *Surveys in Geophysics*, 40(4), 881-911.
- Rodríguez-Veiga, P., S. Quegan, J. Carreiras, H. J. Persson, J. E. Fransson, A. Hoscilo, ... & A. Berninger. 2019. Forest biomass retrieval approaches from earth observation in different biomes. *International Journal of Applied Earth Observation and Geoinformation*, 77, 53-68.
- Rodríguez-Veiga, P., S. Saatchi, K. Tansey, & H. Balzter. 2016. Magnitude, spatial distribution and uncertainty of forest biomass stocks in Mexico. *Remote Sensing of Environment*, 183, 265-281.
- Romero-Duque, L. P., V. J. Jaramillo, & A. Pérez-Jiménez. 2007. Structure and diversity of secondary tropical dry forests in Mexico, differing in their prior land-use history. *Forest Ecology and Management*, 253(1-3), 38-47.
- Saatchi, S. S., N. L. Harris, S. Brown, M. Lefsky, E. T. Mitchard, W. Salas, ... & S. Petrova. 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences*, 108(24), 9899-9904.
- Shimada, M., & T. Ohtaki. 2010. Generating large-scale high-quality SAR mosaic datasets: application to PALSAR data for global monitoring. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3(4), 637-656.
- Sorgonà, A., A. R. Proto, L. M. Abenavoli, & A. Di Iorio. 2018. Spatial distribution of coarse root biomass and carbon in a high-density olive orchard: effects of mechanical harvesting methods. *Trees*, 32(4), 919-931.
- Thapa, R. B., M. Watanabe, T. Motohka, & M. Shimada. 2015. Potential of high-resolution ALOS-PALSAR mosaic texture for aboveground forest carbon tracking in tropical region. *Remote Sensing of Environment*, 160, 122-133.
- Thomson, A. M., K. V. Calvin, L. P. Chini, G. Hurtt, J. A. Edmonds, B. Bond-Lamberty, ... & A. C. Janetos. 2010. Climate mitigation and the future of tropical landscapes. *Proceedings of the National Academy of Sciences*, 107(46), 19633-19638.
- Urbazaev, M., C. Thiel, F. Cremer, R. Dubayah, M. Migliavacca, M. Reichstein, & C. Schmullius. 2018. Estimation of forest aboveground biomass and uncertainties by integration of field measurements, airborne LiDAR, and SAR and optical satellite data in Mexico. *Carbon Balance and Management*, 13(1), 1-20.
- Urquiza-Haas, T., P. M. Dolman, & C. A. Peres. 2007. Regional scale variation in forest structure and biomass in the Yucatan Peninsula, Mexico: Effects of forest disturbance. *Forest Ecology and Management*, 247(1-3), 80-90.
- Wilkes, P., S. D. Jones, L. Suarez, A. Mellor, W. Woodgate, M. Soto-Berelov, ... & A. K. Skidmore. 2015. Mapping forest canopy height across large areas by upscaling ALS estimates with freely available satellite data. *Remote Sensing*, 7(9), 12563-12587.

Zhao, P., D. Lu, G. Wang, L. Liu, D. Li, J. Zhu, & S. Yu. 2016. Forest aboveground biomass estimation in Zhejiang Province using the integration of Landsat TM and ALOS PALSAR data. *International Journal of Applied Earth Observation and Geoinformation*, 53, 1-15.

# Apéndice

El acceso a datos y *scripts* empleados en este manual se encuentran disponibles para descarga en: <https://www.cicy.mx/unidad-de-recursos-naturales/investigador/jose-luis-hernandez-stefanoni/proyecto-mapeo-de-biomasa-aerea-de-bosques>

Se provee un archivo comprimido de la carpeta **Mapeo\_biomasa**, el cual contiene la siguiente información:

1. **Prerrequisitos:** Archivo en formato txt que contiene información para descargar e instalar el *software* R y RStudio, además, contiene líneas de código para la instalación de librerías en R. `.../Mapeo_biomasa/prerrequisitos`
2. **Archivos *script* en formato R** para cada proceso del manual:
  - a. `Calculo_AGB_campo.R`: Cálculo de biomasa aérea en las parcelas de campo. `.../Mapeo_biomasa/1 Calculo_AGB_campo.R`
  - b. `Pre-procesamiento_imagenes.R`: Preprocesamiento de las imágenes ALOS PALSAR. `.../Mapeo_biomasa/2 Pre-procesamiento_imagenes.R`
  - c. `Procesamiento_imagenes.R`: Procesamiento de las imágenes ALOS PALSAR. `2 .../Mapeo_biomasa/3 Procesamiento_imagenes.R`
  - d. `Preparacion_base_datos.R`: Preparación de bases de datos para la modelación. `.../Mapeo_biomasa/3 Procesamiento_imagenes.R`
  - e. `Modelo y mapeo.R`: Modelación y mapeo de la biomasa. `.../Mapeo_biomasa/5 Modelo y mapeo.R`
3. **Raster\_2.9-23.tar.gz** es un archivo de instalación de la librería Raster en la versión 2.9-23 para R. Fuente del archivo: <https://cran.r-project.org/src/contrib/Archive/raster/>. `.../Mapeo_biomasa/raster_2.9-23.tar.gz`
4. **Carpeta FCP**, la cual contiene. `.../Mapeo_biomasa/FCP`
  - a. `CONGLOMERADOS_FCP_INFYS.xlsx` archivo Excel que contiene la información de campo para el cálculo de la biomasa. `.../Mapeo_biomasa/FCP/CONGLOMERADOS_FCP_INFYS.xlsx`
  - b. `Yucatan_DM.csv` es un archivo delimitado por comas que contiene información de densidades de madera para especies arbóreas. `.../Mapeo_biomasa/FCP/Yucatan_DM.csv`
  - c. Polígono de delimitación del área de mapeo en formato shapefile, `FCP_square_poly.shp` y complementos del archivo para poder ser visualizado y proyectado `FCP_square_poly.dbf`, `FCP_square_poly.prj`, `FCP_square_poly.sbn`, `FCP_square_poly.sbx`, `FCP_square_poly.shp.xml` y `FCP_square_poly.shx` `.../Mapeo_biomasa/FCP/ FCP_square_poly.*`

*Mapeo de la biomasa aérea de los bosques mediante datos de sensores remotos y R*

se terminó de imprimir el 30 de enero de 2021 en los talleres de

Editorial Fray Bartolomé de Las Casas,

Pedro Moreno 7, Barrio de Santa Lucía,

C.P. 29250, San Cristóbal de Las Casas, Chiapas.

El cuidado editorial estuvo a cargo de Julio Roldán.

Corrección: Gabriela Herrera.

Diseño de interiores: Sofía Carballo.

Diseño de portada: Alberto Guerra.

La edición consta de 500 ejemplares.



ESTE LIBRO PROPORCIONA UNA GUÍA para llevar a cabo el mapeo de la biomasa aérea del bosque en grandes superficies relacionando datos de campo con la información derivada de imágenes de satélite. La metodología para el mapeo de esta variable utiliza *scripts* desarrollados en el lenguaje R, e incluye la estimación de la biomasa en campo, preprocesamiento de las imágenes, ajuste de modelos de predicción, así como su aplicación para obtener mapas de distribución espacial. Esta obra se dirige a investigadores, estudiantes, técnicos forestales de organizaciones no gubernamentales y del gobierno, y a usuarios de la percepción remota en general que requieran obtener mapas de diferentes atributos de la vegetación en sus sitios de interés.

