



# Nationwide native forest structure maps for Argentina based on forest inventory data, SAR Sentinel-1 and vegetation metrics from Sentinel-2 imagery

Eduarda M.O. Silveira<sup>a,\*</sup>, Volker C. Radeloff<sup>a</sup>, Sebastián Martinuzzi<sup>a</sup>, Guillermo J. Martinez Pastur<sup>b</sup>, Julieta Bono<sup>c</sup>, Natalia Politi<sup>d</sup>, Leonidas Lizarraga<sup>e</sup>, Luis O. Rivera<sup>d</sup>, Lucia Ciuffoli<sup>c</sup>, Yamina M. Rosas<sup>f</sup>, Ashley M. Olah<sup>a</sup>, Gregorio I. Gavier-Pizarro<sup>g</sup>, Anna M. Pidgeon<sup>a</sup>

<sup>a</sup> SILVIS Lab, Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, 1630 Linden Drive, Madison, WI 53706, USA

<sup>b</sup> Centro Austral de Investigaciones Científicas (CADIC), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Houssay 200 (9410) Ushuaia, Tierra del Fuego, Argentina

<sup>c</sup> Dirección Nacional de Bosques, Ministerio de Ambiente y Desarrollo Sostenible de la Nación, Buenos Aires, Argentina

<sup>d</sup> Instituto de Ecoregiones Andinas (INECOA), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Universidad Nacional de Jujuy (UNJu), Juan Bautista Alberdi 47 (Y4600DTA), Jujuy, Argentina

<sup>e</sup> Dirección Regional Noroeste, Administración de Parques Nacionales, Santa Fe 23 (CP4400), Salta, Argentina

<sup>f</sup> Department of Geosciences and Natural Resource Management, University of Copenhagen, Rolighedsvej 23, 1958 Frederiksberg, Denmark

<sup>g</sup> Instituto Nacional de Tecnología Agropecuaria (INTA), Buenos Aires, Argentina

## ARTICLE INFO

Edited by Marie Weiss

### Keywords:

DBH  
Basal area  
Mean height  
Dominant height  
Volume  
Canopy cover  
EVI  
DHIs  
VV polarization  
VH polarization  
Radar  
Optical

## ABSTRACT

Detailed maps of forest structure attributes are crucial for sustainable forest management, conservation, and forest ecosystem science at the landscape level. Mapping the structure of broad heterogeneous forests is challenging, but the integration of extensive field inventory plots with wall-to-wall metrics derived from synthetic aperture radar (SAR) and optical remote sensing offers a potential solution. Our goal was to map forest structure attributes (diameter at breast height, basal area, mean height, dominant height, wood volume and canopy cover) at 30-m resolution across the diverse 463,000 km<sup>2</sup> of native forests of Argentina based on SAR Sentinel-1, vegetation metrics from Sentinel-2 and geographic coordinates. We modelled the forest structure attributes based on the latest national forest inventory, generated uncertainty maps, quantified the contribution of the predictors, and compared our height predictions with those from GEDI (Global Ecosystem Dynamics Investigation) and GFCH (Global Forest Canopy Height). We analyzed 3788 forest inventory plots (1000 m<sup>2</sup> each) from Argentina's Second Native Forest Inventory (2015–2020) to develop predictive random forest regression models. From Sentinel-1, we included both VV (vertical transmitted and received) and VH (vertical transmitted and horizontal received) polarizations and calculated 1st and 2nd order textures within 3 × 3 pixels to match the size of the inventory plots. For Sentinel-2, we derived EVI (enhanced vegetation index), calculated DHIs (dynamic habitat indices (annual cumulative, minimum and variation) and the EVI median, then generated 1st and 2nd order textures within 3 × 3 pixels of these variables. Our models including metrics from Sentinel-1 and 2, plus latitude and longitude predicted forest structure attributes well with root mean square errors (RMSE) ranging from 23.8% to 70.3%. Mean and dominant height models had notably good performance presenting relatively low RMSE (24.5% and 23.8%, respectively). Metrics from VH polarization and longitude were overall the most important predictors, but optimal predictors differed among the different forest structure attributes. Height predictions ( $r = 0.89$  and  $0.85$ ) outperformed those from GEDI ( $r = 0.81$ ) and the GFCH ( $r = 0.66$ ), suggesting that SAR Sentinel-1, DHIs from Sentinel-2 plus geographic coordinates provide great opportunities to map

\* Corresponding author.

E-mail addresses: [esilveira@wisc.edu](mailto:esilveira@wisc.edu) (E.M.O. Silveira), [radeloff@wisc.edu](mailto:radeloff@wisc.edu) (V.C. Radeloff), [martinuzzi@wisc.edu](mailto:martinuzzi@wisc.edu) (S. Martinuzzi), [gpastur@conicet.gov.ar](mailto:gpastur@conicet.gov.ar) (G.J. Martinez Pastur), [jbono@ambiente.gob.ar](mailto:jbono@ambiente.gob.ar) (J. Bono), [npoliti@conicet.gov.ar](mailto:npoliti@conicet.gov.ar) (N. Politi), [llizarraga@apn.gob.ar](mailto:llizarraga@apn.gob.ar) (L. Lizarraga), [lrivera@conicet.gov.ar](mailto:lrivera@conicet.gov.ar) (L.O. Rivera), [lciffoli@ambiente.gov.ar](mailto:lciffoli@ambiente.gov.ar) (L. Ciuffoli), [ymro@ign.ku.dk](mailto:ymro@ign.ku.dk) (Y.M. Rosas), [ahannah@wisc.edu](mailto:ahannah@wisc.edu) (A.M. Olah), [gavierpizarro.g@inta.gob.ar](mailto:gavierpizarro.g@inta.gob.ar) (G.I. Gavier-Pizarro), [apidgeon@wisc.edu](mailto:apidgeon@wisc.edu) (A.M. Pidgeon).

<https://doi.org/10.1016/j.rse.2022.113391>

Received 12 October 2022; Received in revised form 16 November 2022; Accepted 21 November 2022

Available online 30 November 2022

0034-4257/© 2022 Elsevier Inc. All rights reserved.

multiple forest structure attributes for large areas. Based on our models, we generated spatially-explicit maps of multiple forest structure attributes as well as uncertainty maps at 30-m spatial resolution for all Argentina's native forest areas in support of forest management and conservation planning across the country.

## 1. Introduction

Accurate, large-area yet fine-resolution information on forest resources is needed by global, national, and local agencies to support sustainable management and conservation planning (Coops et al., 2021; Moreno et al., 2016; Silveira et al., 2019a; Zald et al., 2016), to maintain ecosystem services (Matasci et al., 2018; Zald et al., 2016), and for scientific studies (White et al., 2014). Forest inventories are conducted to collect forest structure attributes (Wulder et al., 2020) such as diameter at breast height (DBH), dominant height, and canopy cover, and these metrics are used to predict wood volume, aboveground biomass (AGB), carbon storage, wildlife habitat, and biodiversity value (Beaudoin et al., 2014; Blackard et al., 2008; Coops et al., 2021; Hyde et al., 2006). However, detailed and accurate maps of forest attributes, especially for large areas, are rare because extensive field data and reliable predictors are needed.

One of NASA's GEDI (Global Ecosystem Dynamics Investigation Lidar) mission goals is to map forest attributes for large areas. GEDI measures forest vertical structure using 25-m diameter footprints separated by about 600 m across flight tracks within a  $\sim 4.2$  km swath. From this sample a 1-km resolution grid of mean canopy height is derived (Dubayah et al., 2020, 2021b), as well as aboveground biomass (Dubayah et al., 2021a). While these GEDI products are valuable, they have limits for forest management because the gridded maps are relatively coarse (1-km) (Silva et al., 2021), do not cover high latitudes (between 51.6°N and 51.6°S) and represent a point sample of a limited portion of the land area (Potapov et al., 2021). Also, there is high-frequency noise in the GEDI data itself, and substantial geolocation uncertainty (Roy et al., 2021) as well as reduced accuracy in terrain height estimates for steep slopes and in dense forests, which may lead to inaccurate estimates of canopy height (Liu et al., 2021). Furthermore, GEDI products have only been validated with field data from a few countries (Dubayah et al., 2020), and it is unclear how accurate height predictions are for those parts of the globe that did not contribute field data to the GEDI validations.

To overcome the problem that GEDI data itself is only a sample of the Earth's surface, which means that only coarse-scale grids can be derived from GEDI data by themselves, these data have been combined with Landsat multitemporal surface reflectance data to generate a 30-m resolution global map of forest canopy height for the year 2019, the GFCH (Global Forest Canopy Height) product (Potapov et al., 2021). However, any errors in the GEDI data propagate into the GFCH dataset, and the GFCH provides only canopy height information, not other forest structure attributes. Thus, both GEDI and GFCH provide global data on forest structure, but there remains a need to validate them over large areas, including a range of forest ecosystems and types and topographic conditions.

In general, mapping forest structure for large areas requires combining field-based measurements with data from passive or active remote sensors, plus other available auxiliary information, such as elevation (Saatchi et al., 2011), environmental and terrain-related data (Silveira et al., 2019b) or geographic coordinates (Matasci et al., 2018). As part of that framework, remote sensing provides proxies that can be used to model forest attributes and provide accurate measurements over large areas at much lower cost than traditional field inventories (Bouvier et al., 2015). However, most of the forest structure maps developed at continental or nation level are based on coarse-resolution imagery (Avitabile et al., 2016; Baccini et al., 2008; Silveira et al., 2019c), or only on optical imagery, such as medium-resolution Landsat data (Matasci et al., 2018). For example, in Finland and Sweden, as part of the national

forest inventory, 10,000 field plots are measured annually and are combined with satellite data, and other georeferenced digital data to produce volume of the growing stock at a pixel size of 25-m (Tomppo et al., 2008). In Canada, remotely-sensed derived maps of forest structure are available at 30-m resolution (Matasci et al., 2018) while for the conterminous US, Alaska and Puerto Rico, such maps are available at 250-m (Blackard et al., 2008) and 1 ha spatial resolution (Yu et al., 2022).

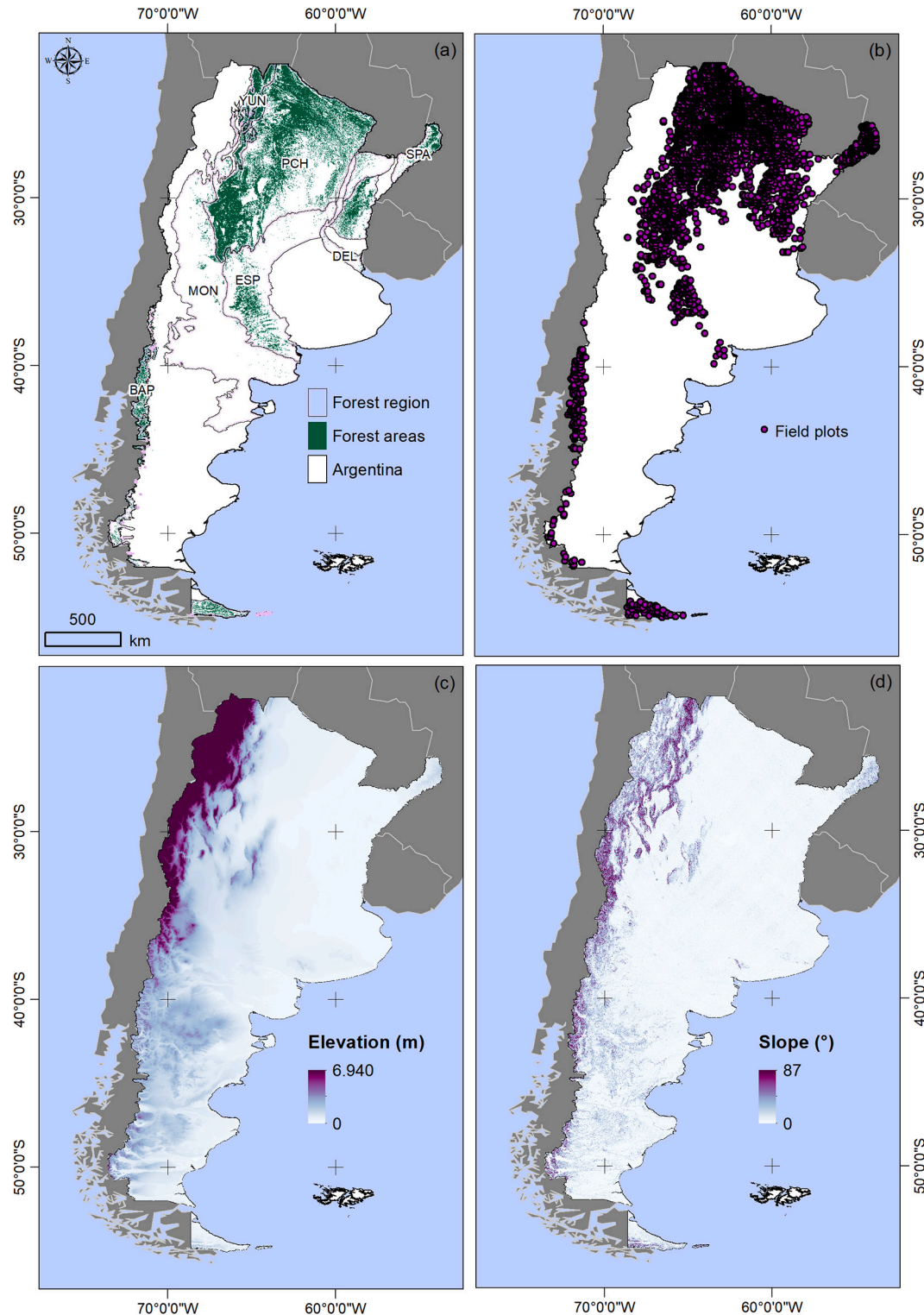
Optical satellite data are suitable for mapping vegetation types (Silveira et al., 2022) and retrieving vegetation structure. Features derived from Sentinel-2 optical data alone can predict forest structure attributes including AGB (Castillo et al., 2017; Majasalmi and Rautiainen, 2016; Puliti et al., 2020), wood volume (Chrysafis et al., 2017; Puliti et al., 2018), DBH, basal area, height (Astola et al., 2019; Lang et al., 2019; Wittke et al., 2019), and canopy cover (Eskandari et al., 2020; Korhonen et al., 2017). However, optical sensors have limitations including saturation problems in high biomass forests, and the influence of weather conditions, atmosphere, moisture, and vegetation phenology (Lu et al., 2016). For example, differences in vegetation phenology due to environmental conditions are an important consideration when selecting image acquisition date for evaluations of the relationships between forest structure and spectral data in subtropical dry forests of Argentina (Gasparri et al., 2010; Gasparri and Baldi, 2013). While optical sensors can capture some aspects of horizontal vegetation structure, integration of lidar and optical data (e.g., Landsat) using empirical modelling procedures greatly improves estimate and maps of forest canopy height and other forest structure attributes (Hudak et al., 2002). In many countries airborne laser scanning (ALS) plays a major role. The combination of national forest inventory, ALS and other remotely sensed data improves models of forest structural variables (Hauglin et al., 2021), however, ALS is not always available.

Radar data can also support the estimation of vertical vegetation structure, such as canopy height. Synthetic aperture radar (SAR) sensors provide information on the backscattered energy from the illuminated target, which means that they penetrate the forest canopy, and are weather-independent (Saatchi, 1997). SAR Sentinel-1C-band satellite data, for example, penetrates through leaves and is scattered by small branches (Ghasemi et al., 2011). SAR Sentinel-1 data are available in C-band HH (horizontal transmit-horizontal) + HV (horizontal transmit-vertical) or VV (vertical transmit-vertical) + VH (vertical transmit-horizontal) polarizations. Time series from SAR Sentinel-1 and derived-metrics (e.g., backscatter, slope, correlation coefficients, and texture metrics) can be used to predict forest structure variables (Bruggisser et al., 2021; Periasamy, 2018), however the C-band does not penetrate deep into the canopy (Karjalainen et al., 2012), which limits the accuracy of forest structure maps in dense vegetation (Nizalapur et al., 2010). The backscattering of radar is affected by surface parameters (e.g., soil moisture and surface roughness) and geometric factors related to the surface attributes of targets (Dobson et al., 1995). For example, slopes facing towards the radar will have small local incidence angles, causing relatively strong backscattering to the sensor, which results in a bright-toned appearance in an image, complicating the estimation of forest attributes.

However, the combination of SAR and optical imagery offers exciting opportunities for accurate forest structure maps for large areas (e.g., Forkuor et al., 2020; Navarro et al., 2019; Pötzschner et al., 2022). In particular, the complementary information provided by SAR Sentinel-1 and Sentinel-2 (Nuthammachot et al., 2020), enhances the accuracy of forest attribute predictions (Forkuor et al., 2020; Morin et al., 2019). When modelling mangrove plantation aboveground biomass in Senegal,

SAR Sentinel-1 (VH backscatter) outperformed Sentinel-2 (spectral bands and vegetation indices), but the combination of the two performed best (Navarro et al., 2019). In tropical savannas and woodlands, Sentinel-2 outperformed SAR Sentinel-1, but again the combination of the two produced the best results (Forkuor et al., 2020). In addition to the pixel-level metrics from either SAR Sentinel-1 polarizations or Sentinel-2 spectral indices, it is advantageous to calculate textures

metrics (Haralick et al., 1973) within moving windows. For example, textures accurately predicted aboveground biomass in China (Chen et al., 2018), and basal area, height, and tree density in a *Pinus pinaster* forest located in France (Morin et al., 2019). However, the empirical evidence that combining SAR Sentinel-1 and 2 can map forest attributes accurately is limited to small study areas so far, and it is not clear if data from these two satellite sensors can also map forest attributes accurately



**Fig. 1.** (a) Location of Argentina's forest regions: Bosque Andino-Patagónico (BAP), Delta e Islas del Río Paraná (DEL), Espinal (ESP), Monte (MON), Parque Chaqueño (PCH), Selva Paranaense (SPA), and Yungas (YUN), (b) forest inventory field plots, (c) elevation and (d) slope.

for large areas and across a wide range of forest types.

Our goal was to evaluate SAR Sentinel-1 and Sentinel-2 and their derivatives for mapping multiple forest attributes across the 463,000 km<sup>2</sup> of Argentina's forests. Specifically, we tested the combination of metrics derived from VV and VH polarizations of SAR Sentinel-1 data and from vegetation metrics derived from Sentinel-2 data to map six forest structure attributes: DBH, basal area, mean height, dominant height, wood volume and canopy cover. Our objectives were to: (1) model forest structure attributes based on forest inventory data, metrics derived from SAR Sentinel-1, Sentinel-2 and geographic coordinates to generate high spatial resolution maps of forest structure attributes across Argentina's forests; (2) map uncertainty for each forest structure attribute; (3) quantify the relative contribution of the predictors to models of forest structure attributes; and (4) compare the accuracy of our height predictions with the accuracies of predictions from GEDI and GFCH using forest inventory plots as validation.

## 2. Methods

### 2.1. Study area

We analyzed the native forest landscape of Argentina (~463,000 km<sup>2</sup>) as delineated by the national government ([Dirección Nacional de Bosques, 2021](#)), in South America, between 20°S and 60°S latitude, and 50°W and 80°W longitude. This map of native forest areas was developed by Argentina's government as part of the National Native Forest Monitoring System and is based on satellite imagery (mainly Landsat and Sentinel-2) and ground-truth data. Argentina's national forest law 26,331 and complementary regulations define native forest as natural woods where trees are dominant species reaching >3-m height at maturity, and where canopy cover is >20% in a minimum 0.5 ha area ([Fig. 1a](#)).

The country has seven regions of native forests ([Peri et al., 2021](#)): (1) Bosque Andino-Patagónico, (2) Delta e Islas del Río Paraná, (3) Espinal, (4) Monte, (5) Parque Chaqueño, (6) Selva Paranaense, and (7) Yungas ([Fig. 1a](#)). Bosque-AndinoPatagónico includes temperate forests with low-diversity tree species assemblages of different types including broadleaved deciduous, evergreen species, and coniferous species. Espinal, Monte and Parque Chaqueño are mostly comprised of subtropical and dry deciduous or evergreen broadleaved forests. Selva Paranaense and Yungas are subtropical rainforests which include different types of evergreen and deciduous trees in a variety of moisture and elevational contexts. Finally, Delta e Islas del Río Paraná forests are located along the main rivers of the region and include subtropical rainforests, deciduous and evergreen broadleaved forests ([Matteucci et al., 2021](#)). Among regions, there is strong topographic contrast between the mountainous western and flat eastern parts of the country ([Fig. 1c-d](#)).

### 2.2. Data

#### 2.2.1. Forest attributes

To obtain the multiple forest attributes, we analyzed a total of 3788 circular field plots (1000-m<sup>2</sup>), which are systematically distributed over the forested area ([Fig. 1b](#)), as part of Argentina's Second Native Forest Inventory ([Dirección Nacional de Bosques, 2021](#)). Given the long-time window of the forest inventory survey (2015–2020), we checked for deforested areas from 2015 to 2020, and excluded 27 deforested plots based on the Global Forest Change map ([Hansen et al., 2013](#)).

Inventory plots were designed based on a grid of 10-km by 10-km, and on this grid the selection was made with high resolution satellite images where there were forests. Then, when the plots were established, exact coordinates were collected with a GPS unit at the center of the plot. A photo of the GPS was taken in order to verify that the coordinates had been typed correctly when loading the data. The GPS error was between 3 and 8 m. During the field survey, technicians measured the

diameter at breast height (DBH, 1.3 m) and the total height of all trees with a minimum DBH of 5 or 10-cm (with different DBH thresholds depending on the region) (Table S1). To check if these different thresholds affected nation-wide analysis, we recalculated the plot-level forest attributes in those regions where the smaller threshold was adopted (i.e., ESP, MON, PCH), based on a) all trees, and b) only trees >10 cm DBH. We found only miniscule differences suggesting that the 5–10 cm DBH trees did not affect plot-level estimates of the six forest attributes, and hence our subsequent model results, noticeably (Table S2). In the field, technicians also measured canopy cover (%) as the projection of tree canopies based on a north-south transect crossing the center of the plot ([Fig. S1](#)).

Then, to obtain our forest structure attributes, for DBH and height, we calculated mean values inside each plot, expressed in cm and m, respectively. We calculated height of dominant trees (expressed in meters), defined as the mean height of the 20% tallest trees. We also calculated basal area (m<sup>2</sup>/ha), and total wood volume (m<sup>3</sup>/ha). Detailed information on the forest inventory is at [SGayDS-Secretaría de Gobierno de Ambiente y Desarrollo Sustentable de la Nación \(2019\)](#). Descriptive statistics of the forest attributes are shown in [Table 1](#).

#### 2.2.2. Remotely sensed predictor variables and auxiliary information

We derived thirty-six (36) remotely-sensed predictors from the Copernicus mission satellites Sentinel-1 and Sentinel-2 at 10-m spatial resolution as well as auxiliary information (two extra variables: latitude and longitude). For SAR Sentinel-1, we used high-resolution ground range detected products (GRD), which we acquired in the Interferometric Wide (IW) swath mode in descending pass direction, available in Google Earth Engine ("COPERNICUS/S1\_GRD"). We used both VV (vertical transmitted and vertical received) and VH (vertical transmitted and horizontal received) backscatter polarizations from 2015 to 2020 to match the forest inventory survey dates. Backscatter is the portion of the outgoing radar signal that the target redirects directly back towards the radar antenna. It is a measure of the reflective strength of a radar target. We did not test HH-HV polarizations because these polarizations are generally used for sea-ice detection and in polar regions ([Wang and Li, 2021](#)).

Thus, based on VV and VH backscatter polarizations, we calculated median values and applied a 3 × 3 moving window to calculate 1st and 2nd order image texture metrics ([Table 2](#)). We chose that window size to match the field plot size of 1000 m<sup>2</sup>. That area is equal to that of a 35.68 by 35.68 m square. Given that each Sentinel pixel size is 10 by 10 m, we summarized the satellite data in a moving window of 3 × 3 pixels, or 30 by 30 m, to approximate the area of the field plots (900 versus 1000 m<sup>2</sup>) ([Fig. S2](#)). All analyses, including both the model fitting, and the model predictions, were done at the 30-m resolution. The 1st order texture metrics are statistical summaries of pixel gray-levels within the processing extent, while 2nd order texture metrics are derived from the gray-level co-occurrence matrix (GLCM) ([Haralick et al., 1973](#)). We

**Table 1**

Descriptive statistics of the forest structure attributes from the 3788 forest inventory plots (1000 m<sup>2</sup>). SD = standard deviation; CV = coefficient of variation; DBH = diameter at breast height; BA = basal area; H = mean height (H); DH = dominant height; WV = wood volume (WV); CC = canopy cover.

Statistic	DBH (cm)	BA (m <sup>2</sup> / ha)	H (m)	DH (m)	WV (m <sup>3</sup> / ha)	CC (%)
Minimum	5.0	0.1	2.3	2.3	0.2	0.0
1st quartile	14.5	5.2	5.4	7.9	21.8	36.6
Mean	19.5	12.0	7.4	11.2	73.4	61.4
Median	17.2	8.9	6.4	10.1	44.1	64.8
3rd quartile	21.6	14.4	8.0	12.8	82.7	90.1
Maximum	92.2	106.9	34.4	44.0	1379.8	100.0
SD	8.5	11.6	3.4	5.1	104.9	31.1
CV (%)	43.4	96.5	46.2	45.2	143.0	50.7



**Table 2**

SAR Sentinel-1 and Sentinel-2 derived features and their derivation from first and second order texture measures. VV = vertical transmitted and received; VH = vertical transmitted and horizontal received; SD = standard deviation; HOM = homogeneity; CORR = correlation; SH = shade. EVI = enhanced vegetation index.

Satellite	Polarization	Metric	Feature
Sentinel-1	VV	1st - Mean	S1_VV_MEAN
		1st - Standard deviation	S1_VV_SD
		2nd - Homogeneity	S1_VV_HOM
		2nd - Uniformity	S1_VV_UNI
		2nd - Correlation	S1_VV_CORR
		2nd - Shade	S1_VV_SH
		1st - Mean	S1_VH_MEAN
	VH	1st - Standard deviation	S1_VH_SD
		2nd - Homogeneity	S1_VH_HOM
		2nd - Uniformity	S1_VH_UNI
		2nd - Correlation	S1_VH_CORR
		2nd - Shade	S1_VH_SH
Sentinel-2	EVI_median	1st - Mean	S2_EVI_median_MEAN
		1st - Standard deviation	S2_EVI_median_SD
		2nd - Homogeneity	S2_EVI_median_HOM
		2nd - Uniformity	S2_EVI_median_UNI
		2nd - Correlation	S2_EVI_median_CORR
		2nd - Shade	S2_EVI_median_SH
	EVI_DHI_Cumulative	1st - Mean	S2_EVI_DHI_Cum_MEAN
		1st - Standard deviation	S2_EVI_DHI_Cum_SD
		2nd - Homogeneity	S2_EVI_DHI_Cum_HOM
		2nd - Uniformity	S2_EVI_DHI_Cum_UNI
		2nd - Correlation	S2_EVI_DHI_Cum_CORR
		2nd - Shade	S2_EVI_DHI_Cum_SH
	EVI_DHI_Minimum	1st - Mean	S2_EVI_DHI_Min_MEAN
		1st - Standard deviation	S2_EVI_DHI_Min_SD
		2nd - Homogeneity	S2_EVI_DHI_Min_HOM
		2nd - Uniformity	S2_EVI_DHI_Min_UNI
		2nd - Correlation	S2_EVI_DHI_Min_CORR
		2nd - Shade	S2_EVI_DHI_Min_SH
	EVI_DHI_Variation	1st - Mean	S2_EVI_DHI_Var_MEAN
		1st - Standard deviation	S2_EVI_DHI_Var_SD
		2nd - Homogeneity	S2_EVI_DHI_Var_HOM
		2nd - Uniformity	S2_EVI_DHI_Var_UNI
		2nd - Correlation	S2_EVI_DHI_Var_CORR
		2nd - Shade	S2_EVI_DHI_Var_SH

selected six texture metrics that are not strongly correlated (less than + – 0.7) (mean, standard deviation, homogeneity, uniformity, correlation, shade). For 2nd order textures calculations, we rescaled values to 8-bit (GLCMs with dimensions of 256 rows  $\times$  256 columns) to avoid zeros in the matrix that occurred in 16-bit imagery (65,536 rows  $\times$  65,536 columns).

For Sentinel-2, we acquired surface reflectance data, available in Google Earth Engine (“COPERNICUS/S2\_SR”) from 2015 to 2020. The Sentinel-2 processing includes an atmospheric correction applied to Top-Of-Atmosphere Level-1C orthoimage products computed by running sen2cor (Müller-Wilm, 2016). We used the Sentinel-2 scene classification band to mask the pixels that were labelled as cloud, shadow, and snow. We calculated the enhanced vegetation index (EVI) and generated a composite image by calculating EVI median values for 2015–2020. In addition to EVI median values, we also calculated annual composite DHIs (dynamic habitat indices) based on EVI times series from 2015 to 2020. The DHIs summarize productivity of vegetation in

three ways: (i) annual cumulative productivity (cumulative DHI), a measure of available energy and integrates productive capacity over a year; (ii) minimum productivity (minimum DHI), a measure of resource limitation; and (iii) annual variation or seasonality (variation DHI), that captures seasonality (Berry et al., 2007; Coops et al., 2009; Duro et al., 2007). To calculate DHIs, first, we selected the highest EVI value in each month, to reduce artifacts. From these, we selected the median monthly values for the six years, to reduce the effects of extreme years. Based on the median monthly values, we calculated the cumulative DHI as the sum of the monthly values, the minimum DHI, and the variation DHI as the coefficient of variation. We then calculated the 1st and 2nd order texture metrics within a moving window of  $3 \times 3$  pixels (Table 2). For 2nd order textures calculations, we rescaled values to 8-bit.

As auxiliary information, we included latitude and longitude coordinates in our model to capture trends associated with the regional temperature and precipitation gradients across the country (Matasci et al., 2018). We did not use temperature or precipitation as predictors because the climate dataset across Argentina is at 1-km spatial resolution (Fick and Hijmans, 2017), and thus too coarse for our predictions at 30-m spatial resolution.

### 2.3. Analysis

#### 2.3.1. Random forest models and uncertainty maps

Machine learning techniques such as k-nearest neighbor (k-NN; Dudani, 1976) and random forest (RF; Breiman, 2001) have been successfully used to model forest structure attributes (Latifi and Koch, 2012; Silveira et al., 2019c; Tomppo et al., 2008). The principle of k-NN is to predict the target variable based on the closest observations in the training dataset, while the RF algorithm is based on decision trees. In models of forest structure attributes, RF performed better than k-NN in some studies (Fassnacht et al., 2014; Latifi and Koch, 2012; Tompalski et al., 2019). In addition, k-NN technique is computationally intensive when there is a large number of plots (Tomppo et al., 2008). In addition, RF provides the ‘Variable Importance’ to evaluate the importance of predictors, is less sensitive to noise in the training data, can accommodate many predictor variables without overfitting, and tends to result in more accurate predictions than other regression frameworks (Baccini et al., 2008; Shataee et al., 2012). Because of these advantages, we selected RF for our models.

We used the randomForest package (Liaw and Wiener, 2002) available in R (R Core Team, 2021). For calibration and training, we randomly selected 70% of the field plots balanced across Argentina’s forest regions (according to the data histogram, capturing all class intervals for training). Based on the training data, we selected optimal values of the hyperparameters (ntree and mtry) for each random forest model and employed the random search method to tune the hyperparameters (Bergstra and Bengio, 2012). The remaining 30% were our validation data. For validation, we generated scatterplots and computed the coefficient of determination ( $R^2$ , in %), the root-mean-square error (RMSE) and mean error (ME).

Understanding pixel-level uncertainty is critical to understanding the utility of the predictive maps. However, because Random forest is a non-parametric approach, there is no direct quantification of prediction uncertainty at the pixel level. Some studies that assessed forest structure attributes, and their related uncertainties, did so based on model-assisted (e.g., Gregoire et al., 2011; Saarela et al., 2015), or model-based inference (Magnussen, 2015), or via Monte Carlo simulation (Coulston et al., 2016). However, these methods require a fairly large probability sample of field data and tend to underperform in sparse areas of the distribution (Coulston et al., 2016). One alternative is to provide uncertainties in terms of RMSE (Csillik et al., 2019, 2020; Dos Reis et al., 2020; Saarela et al., 2015). The RMSE is a key performance metric when modelling forest attributes (Coops et al., 2021). Thus, we generated uncertainty maps of each forest structure attribute to assess the correspondence between what was predicted and what was observed in each

pixel. For this, we calculated both the absolute RMSE (Eq. (1)) and relative RMSE (i.e., the RMSE as a proportion of the mean; Eq. (2)).

To calculate RMSE, first we grouped the predicted values of each forest structure variable into 10 bins (based on natural breaks) and then computed the RMSE (both absolute and relative) for each bin. Second, we fitted polynomial regression models using the 10 RMSE bin values and the predicted mean values of each variable per bin. Based on these regression models we then mapped the RMSE of each forest attribute. Based on the relative uncertainty maps, we also evaluated results by forest types, elevation and slope gradients. For forest types we used a phenocluster map of Argentina at 30-m spatial resolution where forests are characterized based on land surface phenology and climate patterns (Silveira et al., 2022). Elevation and slope were derived from the SRTM (shuttle radar topography mission).

$$\text{RMSE absolute} = \sqrt{\frac{\sum (\text{predicted} - \text{observed})^2}{N}} \quad (1)$$

$$\text{RMSE relative} = \frac{\text{RMSE absolute}}{\sum \text{observed}} * 100 \quad (2)$$

To quantify the relative contribution of the predictors to models of each forest structure attribute, we employed the RRelief algorithm (Kononenko et al., 1996). RRelief is a classic filter-based method suitable for regression tasks with continuous output values and calculates a feature score which ranks each feature for selection, improving the performance of the algorithms (Chandrashekar and Sahin, 2014; Dos Reis et al., 2020). RRelief estimates the quality of attributes in regression problems and gives higher scores to higher quality attributes (Kononenko et al., 1996; Urbanowicz et al., 2018). Based on the ranking attributes, we used the function *cutoff.biggest.diff* to identify the subset of predictors which were top contributors to model performance (R Core Team, 2021). In addition to our initial models based on variables from Sentinel-1 plus Sentinel-2 plus geographic coordinates, we also tested models based on subsets of predictors derived from (1) Sentinel-1, (2) Sentinel-2, (3) Sentinel-1 plus Sentinel-2.

### 2.3.2. Forest structure attribute maps

To map the forest structure attributes DBH, basal area, mean height, dominant height, wood volume and canopy cover, we applied the random forest regression models to each pixel containing the selected features' continuous values. Because the original remotely-sensed predictors are 10-m spatial resolution, we resampled the selected features' continuous values to 30-m spatial resolution (900-m<sup>2</sup> pixel) to approximately match resolutions. All maps are in geographic coordinate system and datum World Geodetic System (WGS) 1984. One of the challenges when generating independent models for each forest attribute is that the different predictions for a given pixel may not match, e.g., there may be cases with a high predicted basal area, but low predicted volume. To check for this, we randomly distributed 2000 points over the study area, and calculated Pearson correlation between the different variables both in the sampling plots, and in our 2000 randomly points. We found strong positive correlations (e.g., height and dominant height), moderate positive correlations (e.g., basal area and height) and low positive correlations (e.g., diameter at breast height and canopy cover) for both plot-level estimates and predicted locations. For the correlations between basal area, mean height, dominant height, wood volume and canopy cover, we found no differences in the correlations for plots and predicted locations suggesting that nonsensical predictors were not an issue (Fig. S3). Only the correlations between DBH and the other forest attributes were lower for the predicted locations, suggesting that DBH predictions may not be quite as consistent with the other variables.

### 2.3.3. Height predictions versus GEDI and GFCH

We compared our predictions of mean height with predictions from GEDI and GFCH. The objective of this analysis was to evaluate which

height estimate is most accurate in Argentina. To do so, first we downloaded GEDI 3 Gridded Land Surface Metrics, in which mean canopy height is provided, calculated as the mean of values within each 1-km cell, from 2019 and 2020. Canopy height is provided as the mean height (in meters) above the ground of the received waveform signal that was the first reflection off the top of the canopy (RH100) (Dubayah et al., 2020). Accordingly, we expected that the GEDI canopy height would correlate most strongly with dominant height. Because the original GEDI dataset is an Equal-Area Scalable Earth (EASE)-Grid projection in datum WGS 1984 we converted it to geographic coordinates to match our dataset. Second, we downloaded the GFCH (Potapov et al., 2021), a 30-m spatial resolution global forest canopy height map for 2019, derived from GEDI extrapolation based on Landsat multitemporal metrics (Potapov et al., 2019, 2020). To obtain the GFCH estimates, for each footprint, a set of relative height metrics (RH75, RH90, RH95, and RH100) are extracted and the mean RH metric is calculated for each footprint (Potapov et al., 2021). We expected that the GFCH estimates would correlated most strongly with mean height.

To compare our predictions with GEDI, we resampled our mean height map to 1-km spatial resolution. Then, we selected the validation plots (i.e., the remaining 30% plots that were not used to train our random forest models) that also intersected with GEDI gridded metric (total 240 plots). Based on these plots, we computed Pearson correlation (*r*) and the root-mean square error (RMSE) between (1) mean height of 240 validation plots versus our mean height predictions, and (2) mean height of 240 validation plots versus GEDI predictions. To compare our predictions with the GFCH, we used our mean height map at 30-m spatial resolution. We also selected the validation plots that also intersected with the GFCH map (total 550 plots). Based on these plots, we also computed Pearson correlation (*r*) and the root-mean square error (RMSE) between (1) mean height of 550 validation plots versus our mean height predictions, and (2) mean height of 550 validation plots versus GFCH predictions. Finally, we analyzed the influence of elevation and slope on the differences between GEDI and GFCH versus our validation plots. To do so, we computed Pearson correlation between GEDI/GFCH residual errors with elevation and slope (obtained from SRTM) and generated boxplots by forest region.

In a second evaluation, we repeated the previous analysis by comparing our measure of dominant height (instead of mean height) to GEDI and GFCH. However, because mean and dominant height are highly correlated (*r* = 0.92), and the results were very similar, we present this extra analysis based on dominant height in our supplementary material (Fig. S4).

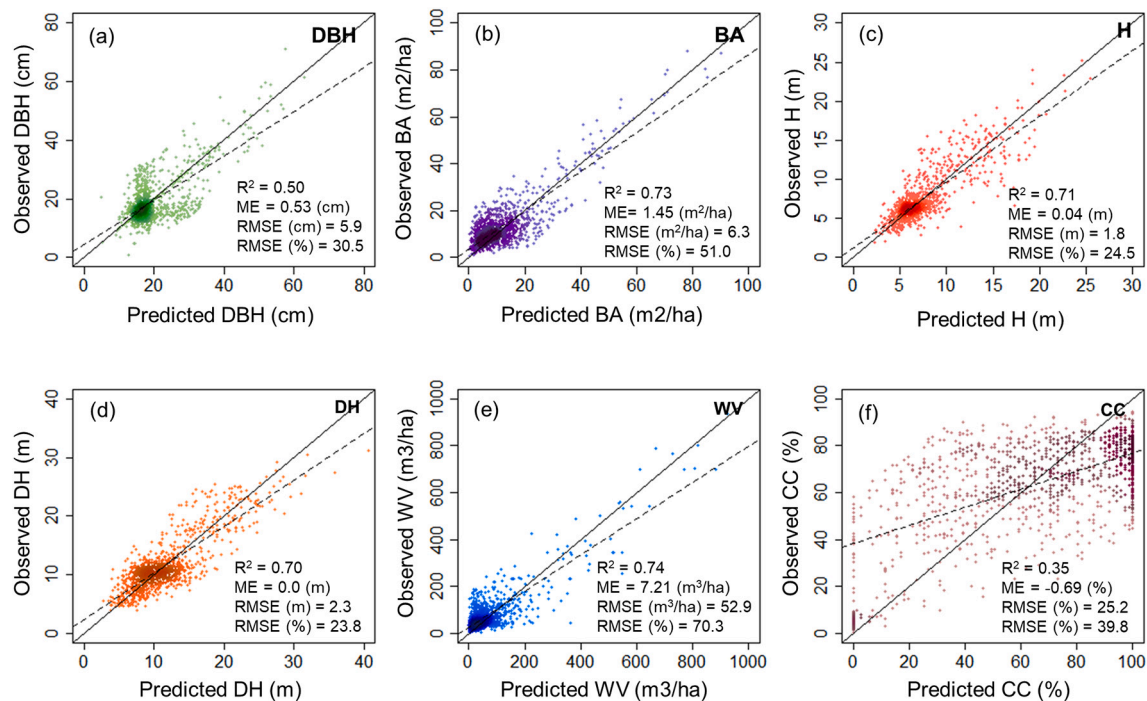
## 3. Results

### 3.1. Random forest models

The models of forest structure attributes captured varying levels of attribute variability. Mean and dominant height models had the best performance when considering the *R*<sup>2</sup> (0.71 and 0.70, respectively) and RMSE (24.5% and 23.8%, respectively) together. The model *R*<sup>2</sup> metrics ranged from 0.35 (canopy cover) to 0.74 (wood volume) and relative RMSE values ranged from 23.8% (dominant height) to 70.3% (wood volume). RMSE was lower than or equal to 50% for DBH, mean height, dominant height, and canopy cover. However, for basal area and wood volume RMSE reached 51.0% and 70.3%, respectively. Scatterplots showed that the relationships with predictions for basal area, mean height, dominant height, and wood volume were close to the 1:1 line (*R*<sup>2</sup> = 0.73, 0.71, 0.70 and 0.74, respectively), while DBH and canopy cover deviated more from the 1:1 line (*R*<sup>2</sup> = 0.50 and 0.35, respectively) (Fig. 2).

### 3.2. Forest structure attributes spatial distribution

Our forest structure maps at 30-m spatial resolution across



**Fig. 2.** Scatterplots for the regression model of forest structure attributes: (a) diameter at breast height (DBH), (b) basal area, (c) mean height, (d) dominant height (DH), (e) wood volume and (f) canopy cover. 1:1 line in solid black and fitted line in dashes.  $R^2$  = coefficient of determination; ME = mean error; RMSE = root-mean-square error.

Argentina's forests demonstrate the potential of combining SAR Sentinel-1, vegetation metrics from Sentinel-2 and geographic coordinates to capture detailed forest structure patterns at landscape level (Fig. 3). Overall, all forest structure attributes values were higher in Bosque Andino-Patagónico, Yungas and in Selva Paranaense and lower in Parque Chaqueño, Espinal and Monte. Across Argentina, according to our predictions, DBH ranged from 5 to 42-cm, while basal area and wood volume ranged from 1 to 106 m<sup>2</sup>/ha and 1–1240 m<sup>3</sup>/ha, respectively. Mean height and dominant height ranged from 2 to 31-m and 3–32-m, respectively and canopy cover ranged from 0 to 98% (Fig. 4).

The main descriptive statistics (min, mean, max and standard deviation) of the observed (field plots) and predicted values (our maps) matched relatively closely across the country (Table 3). According to our predictions, mean values across Argentina's forests of DBH, basal area and wood volume were 20.1 cm, 13.8 m<sup>2</sup>/ha and 82.14 m<sup>3</sup>/ha, respectively. Predicted mean height and dominant height were 7.5-m and 11.5-m, respectively, while canopy cover was 62% on average.

### 3.3. Uncertainty maps

To generate uncertainty maps, we used the polynomial models for both the relative RMSE (%) and the absolute RMSE. Those models generally had a good fit, with the  $R^2$  ranging from 0.74 to 0.95 for the relative RMSE (Fig. 5a) and from 0.63 to 0.94 (for the absolute RMSE (Fig. 5b)). Overall, the highest relative errors were mostly in Delta e Islas del Río Paraná, Espinal, Monte and in Parque Chaqueño, forest regions (Fig. 6; Fig. S5), where trees are small, while the highest absolute errors were concentrated in Bosque Andino-Patagónico, Selva Paranaense and Yungas, and forest regions (Fig. 7; Fig. S6), where trees are dense and tall. The exception was canopy cover, which also had higher absolute RMSE in all forest regions.

In evaluating relative RMSE, the uncertainties in DBH were <50% and the highest errors were in predicted values of DBH lower than 15 cm (mostly in Espinal and Monte forest regions). For basal area, relative uncertainties were higher than 50% where values were higher than 40 m<sup>2</sup>/ha, mainly in Parque Chaqueño, Espinal and Monte forest regions.

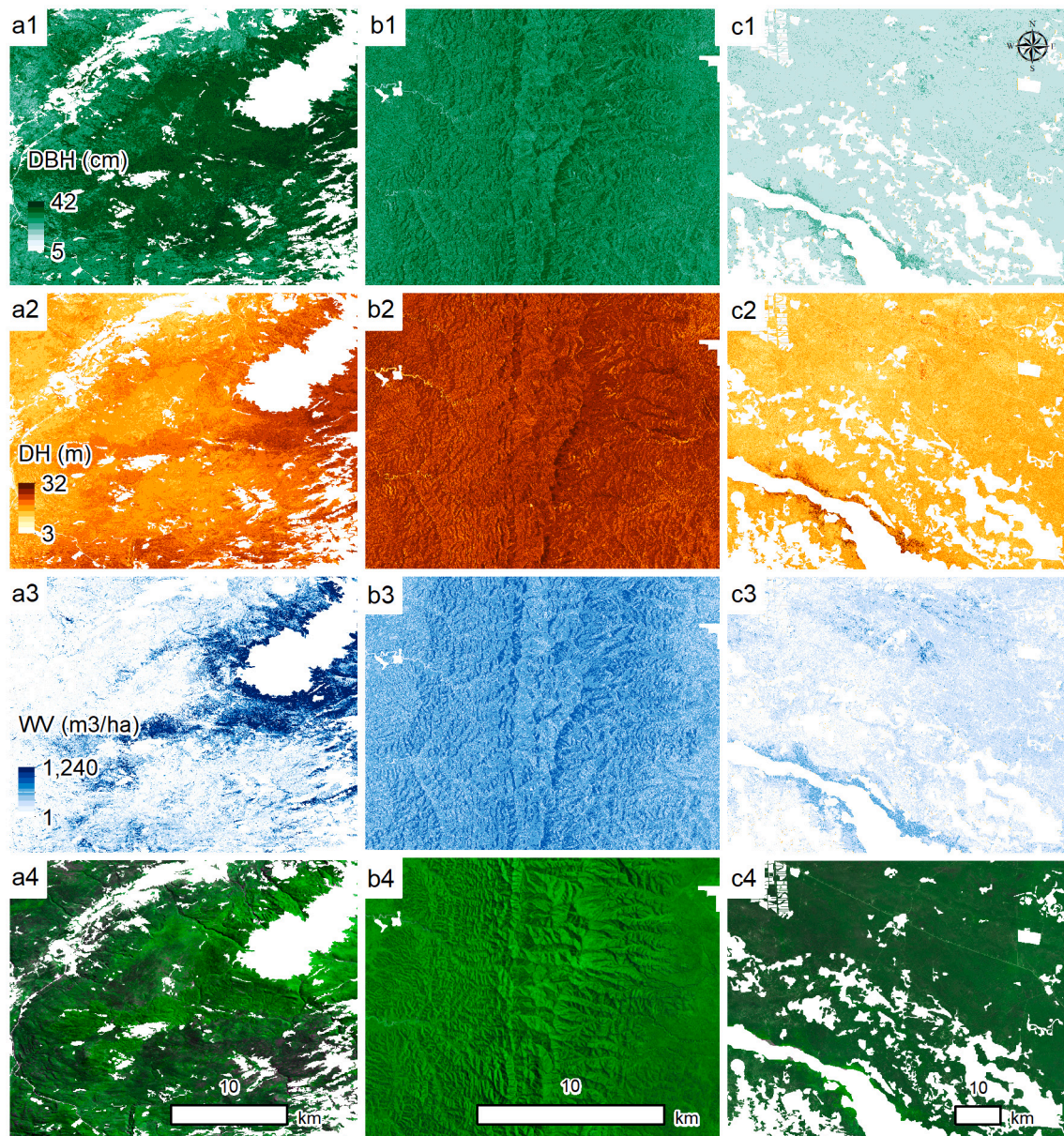
For mean and dominant height, uncertainty was <40% and the highest errors we found in predicted values of height lower than 5 m (mostly in Monte region). The relative uncertainties in wood volume were higher than 50% for lower predicted values of wood volume (< 400 m<sup>3</sup>/ha), especially in Parque Chaqueño, Espinal and Monte. For canopy cover, uncertainties were higher than 50% for predictions lower than 60%, in Espinal and Monte (Fig. 5a; Fig. 6).

Regarding absolute RMSE, the uncertainties in DBH were >6-cm for predicted values of DBH higher than 30-cm, in Bosque Andino-Patagónico, Yungas and Selva Paranaense. For basal area, absolute uncertainties were higher than 6 m<sup>2</sup>/ha where basal areas were higher than 40 m<sup>2</sup>/ha, such as in Yungas, Bosque Andino-Patagónico and in Selva Paranaense forest regions. For height and dominant height, we found uncertainties of >3-m where trees were taller than 15-m (e.g., Yungas, Bosque Andino-Patagónico and Selva Paranaense). Finally, the absolute uncertainties for wood volume were higher in Bosque Andino-Patagónico, where values were higher than 300 m<sup>3</sup>/ha (Fig. 5b; Fig. 7).

When analyzing relative RMSE by forest type, results showed that six phenoclusters mainly represented by riparian forests (phenoclusters 14, 17, 18, 20, 21 and 23) within Delta e Islas del Río Paraná and Espinal forest regions presented the highest errors for all predicted forest attributes. For DBH, basal area and wood volume, lower relative RMSE are in Bosque Andino-Patagónico where broadleaved forests are dominant. For height and dominant height, lower relative RMSE are within the broadleaved forests of Bosque Andino-Patagónico, but also in degraded forests in Selva Paranaense. In addition, canopy cover had higher relative RMSE in degraded Calden forests in the Espinal forest region (Fig. S7–S12).

When analyzing relative RMSE by elevation, forest structure attributes had low errors in flat terrain (0–100-m), and highest errors in higher altitudes (500–1000-m) (Fig. S13). Regarding slopes gradients, intermediate classes of 3–5° and 10–15° had the lowest uncertainty (Fig. S14).





**Fig. 3.** Example of detailed pattern of diameter at breast height (DBH), dominant height (DH) and wood volume (WV) within (a) Bosque Andino-Patagónico (BAP), (b) Yungas (YUN) and (c) Parque Chaqueño (PCH). Rows: (1) DBH; (2) DH; (3) WV and (4) Sentinel-2 bands 4, 3, and 2 in RGB.

### 3.4. Predictor variable importance

The optimal subset of predictor variables included eight (8) of our thirty-eight (38) initial predictors. Of the eight predictors, two were the geographic coordinates latitude and longitude, and the remaining six were mean derived texture metrics from Sentinel-1 and Sentinel-2 (Fig. 8). The remaining 1st order texture metric (standard deviation) and the 2nd order textures homogeneity, uniformity, correlation and shade derived from both SAR Sentinel-1 and Sentinel-2 were not selected in models of any forest attribute. Latitude and longitude improved the r-square relative to models based on Sentinel-1 and 2 only by 2.1 to 5.7 percentage point (Table S3).

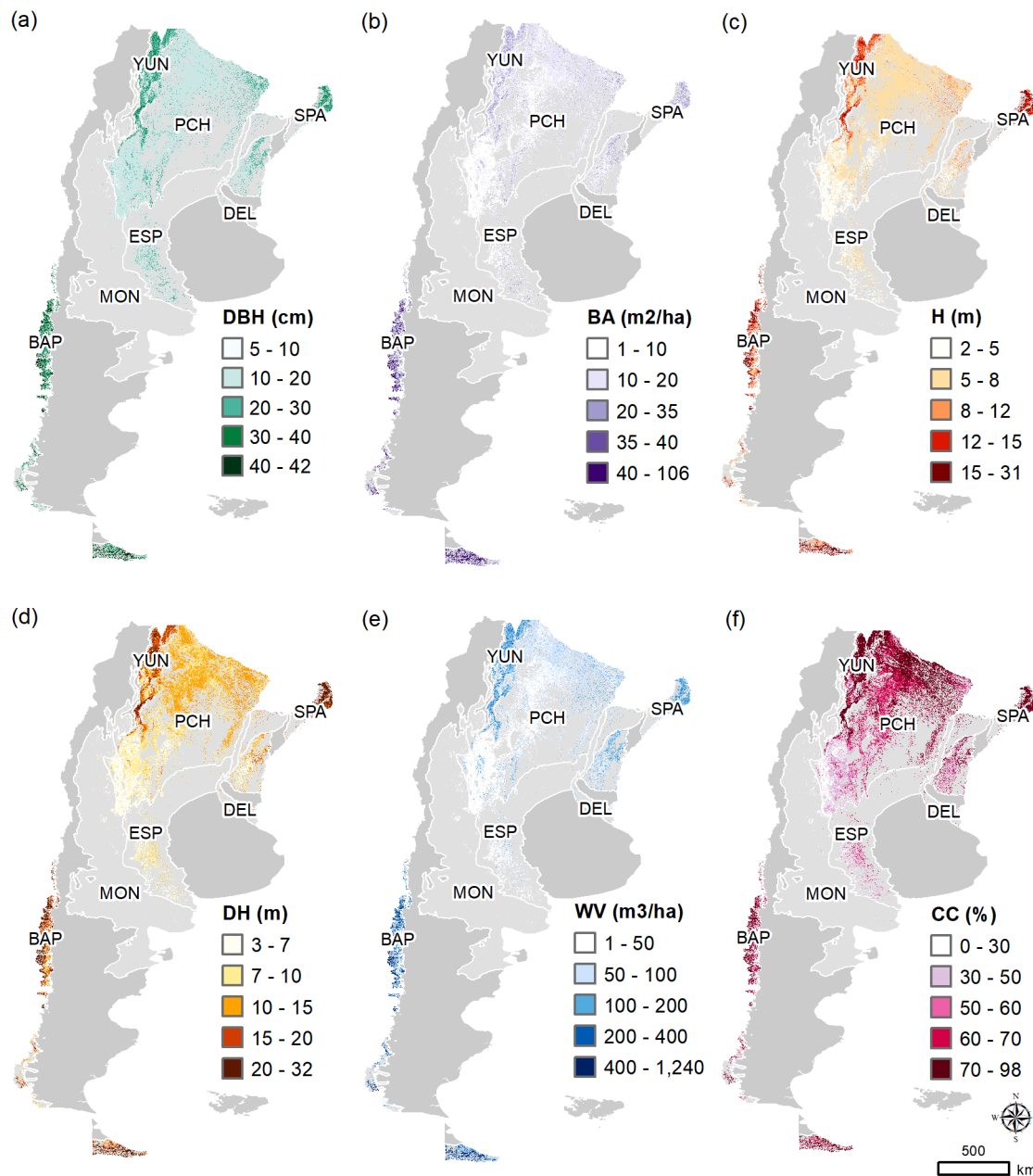
We ranked the eight predictors according to their relative importance in each model of forest structure attributes. The most important variable for modelling the forest structure attributes was the mean texture metric of VH polarization from SAR Sentinel-1 (S1\_VH\_MEAN), which was selected in all six forest structure attribute models. Longitude was the 2nd most important variable, also included in all six models. For

each forest attribute, at least one geographic coordinate, one polarization variable from SAR Sentinel-1 and one vegetation metric from Sentinel-2 was selected, but optimal predictors differed in their strength of association with each forest structure attribute. For DBH and basal area the VH polarization from SAR Sentinel-1 (S1\_VH\_MEAN) was the most important (relative importance of 38% and 26%, respectively). For mean height and dominant height, mean of minimum DHI from Sentinel-2 (S2\_EVI\_DHI\_Min\_MEAN) was the most important variable (relative importance of 20% and 34%, respectively). For wood volume, median EVI from Sentinel-2 (S2\_EVI\_median\_MEAN) was the most important predictor (relative importance of 27%) and for canopy cover the mean of VV polarization of Sentinel-2 (S1\_VV\_MEAN) had 31% of the relative importance of predictors in the model (Fig. 8).

### 3.5. Height predictions versus GEDI and GFCH

Our height predictions based on SAR Sentinel-1 combined with vegetation metrics from Sentinel-2 and geographic coordinates clearly





**Fig. 4.** Estimates of (a) diameter at breast height (DBH), (b) basal area (BA), (c) mean height (H), (d) dominant height (DH), (e) wood volume (WV), and (f) canopy cover (CC) across Argentina's forest regions.

outperformed GEDI and the GFCH predictions when compared with field-plot validation measures of mean height. Our predicted mean height values had a Pearson correlation of 0.89 with the field data versus 0.81 for GEDI. In addition, GEDI had absolute and relative RMSEs of 4.0 m and 53.5% while our mean height predictions had RMSEs of 1.5-m and 19.7% (Fig. 9a, b). When comparing our maps with the GFCH, our predicted mean height values had a Pearson correlation of 0.85, versus 0.66 for GFCH. The RMSE also revealed that our mean height predictions presented lower RMSE (1.7-m and 22.9% for relative and absolute, respectively) than the GFCH (3.3-m and 52.3%) (Fig. 9c, d).

Topographic conditions affected GEDI and GFCH estimates. We found higher errors on steeper slopes and at higher elevations (Table S4). In addition, we found higher errors of both GEDI and GFCH estimates within BAP and YUN forest regions, which are characterized by mountain forests (Fig. S15).

#### 4. Discussion

We modelled and mapped six forest structure attributes in Argentina using SAR Sentinel 1- and Sentinel-2 data, and  $R^2$  values ranged from 0.48 to 0.74 with RMSE ranging from 23.82% to 70.33% when validated against native forest inventory plots. Texture metrics of SAR Sentinel-1 polarizations combined with metrics of DHIs from Sentinel-2 were complementary in our models. The optimal subset of predictors differed for DBH, basal area, mean height, dominant height, wood volume and canopy cover, but the most important variable was the mean texture metric of VH polarization from SAR Sentinel-1 (S1\_VH\_MEAN). Our height predictions outperformed GEDI and the global forest canopy height product (GFCH), highlighting the potential of our predictors to map forest structure attributes, especially mean height, and dominant height, at broad scale more accurately than was possible previously

**Table 3**

Summary statistics for observed and predicted forest attributes across Argentina. Mean DBH = diameter at breast height; Min = minimum; Max = maximum. DBH = diameter at breast height; BA = basal area; H = mean height (H); DH = dominant height; WV = wood volume (WV); CC = canopy cover. Forest regions: Bosque Andino-Patagónico (BAP), Delta e Islas del Río Paraná (DEL), Espinal (ESP), Monte (MON), Parque Chaqueño (PCH), Selva Paranaense (SPA), and Yungas (YUN).

Forest region	Forest attribute	Observed			Predicted		
		Min	Mean	Max	Min	Mean	Max
BAP	DBH (cm)	10.4	37.6	92.2	15.3	36.8	42.74
	BA (m <sup>2</sup> /ha)	0.3	39.8	106.9	3.6	42.7	106.0
	H (m)	3.2	12.9	34.4	3.9	13.5	31.0
	DH (m)	3.5	17.7	44.0	5.7	18.4	32.0
	WV (m <sup>3</sup> /ha)	0.4	305.5	1379.8	4.3	287.6	1240
	CC (%)	0.0	66.7	100.0	0.0	65.4	86.9
DEL	DBH (cm)	11.6	19.3	31.3	14.2	19.3	22.7
	BA (m <sup>2</sup> /ha)	2.9	11.0	21.1	7.4	13.4	20.6
	H (m)	4.2	8.5	17.3	4.7	7.6	13.8
	DH (m)	4.9	11.3	21.0	5.0	10.9	19.0
	WV (m <sup>3</sup> /ha)	23.3	75.7	248.9	39.1	68.2	88.1
	CC (%)	14.1	50.4	97.2	0.0	56.3	75.9
ESP	DBH (cm)	7.6	18.7	51.0	9.0	16.9	35.7
	BA (m <sup>2</sup> /ha)	0.3	10.2	39.8	2.2	10.2	35.8
	H (m)	3.1	5.3	11.0	3.1	5.0	9.2
	DH (m)	3.5	7.4	16.7	4.0	7.5	18.1
	WV (m <sup>3</sup> /ha)	0.3	58.7	205.5	14.4	55.5	155.5
	CC (%)	0.0	51.7	100.0	0.0	46.6	92.7
MON	DBH (cm)	5.0	10.4	21.9	5.2	13.0	18.4
	BA (m <sup>2</sup> /ha)	0.5	5.2	16.7	2.2	4.6	14.8
	H (m)	2.3	3.7	6.0	2.9	3.7	5.2
	DH (m)	3.8	5.7	14.6	4.6	5.5	6.7
	WV (m <sup>3</sup> /ha)	0.6	6.5	29.1	6.3	9.6	25.9
	CC (%)	0.0	22.7	80.3	0.0	24.7	70.6
PCH	DBH (cm)	5.7	16.8	47.2	5.8	18.3	36.0
	BA (m <sup>2</sup> /ha)	0.1	9.2	63.8	0.9	11.0	54.9
	H (m)	2.3	6.3	17.5	2.3	6.4	14.0
	DH (m)	2.3	9.9	23.9	3.9	9.9	19.7
	WV (m <sup>3</sup> /ha)	0.2	47.9	1017.5	8.6	58.0	358.5
	CC (%)	0.0	60.1	100.0	0.0	61.5	94.4
SPA	DBH (cm)	12.6	27.9	57.6	14.4	23.7	49.5
	BA (m <sup>2</sup> /ha)	0.4	17.3	64.7	3.7	19.6	34.6
	H (m)	4.0	13.7	22.5	9.7	14.4	18.2
	DH (m)	4.5	20.2	33.9	16.5	22.1	25.6
	WV (m <sup>3</sup> /ha)	1.1	115.1	709.7	34.8	130.8	271.6
	CC (%)	0.0	66.2	100.0	0.0	76.3	93.3
YUN	DBH (cm)	10.1	28.7	59.0	14.6	24.2	48.4
	BA (m <sup>2</sup> /ha)	0.3	16.7	73.7	0.6	18.2	57.0
	H (m)	3.5	12.0	27.3	5.4	11.6	16.5
	DH (m)	3.5	17.8	36.3	7.0	17.6	30.5
	WV (m <sup>3</sup> /ha)	0.3	136.7	1016.0	23.5	141.3	554.7
	CC (%)	0.0	82.2	100.0	0.0	78.9	98.0

based on GEDI and the GFCH.

#### 4.1. Forest structure modelling

The combination of features from SAR Sentinel-1, Sentinel-2, geographic coordinates, and our high-quality field data, contributed to the high predictive power of our models. The optimal subset of predictor variables included metrics from both SAR Sentinel-1 and Sentinel-2, plus latitude and longitude, highlighting the gain of combining radar, optical data and geographic coordinates.

Other studies that predicted forest attributes based on radar plus optical data also found complementarity of different features in top models (Chen et al., 2018; Forkuor et al., 2020; Laurin et al., 2018; Navarro et al., 2019; Pötzschner et al., 2022). While radar data can characterize structure information, optical data captures canopy density and foliage information (Forkuor et al., 2020). For example, for maps of aboveground biomass in tropical savannas and woodlands of Africa, the combination of SAR Sentinel-1 and Sentinel-2 data resulted in models with RMSE of 45.4% compared to 78.6% when only using SAR Sentinel-

1 data, and 60.6% for Sentinel-2 data alone (Forkuor et al., 2020). Similarly, in the dry Chaco of Argentina, aboveground biomass models that combined radar and optical predictors had an RMSE of 15.1 ton/ha compared to 27.1 ton/ha for optical predictors only, and 25.8 ton/ha for radar predictors alone (Pötzschner et al., 2022).

We found that our models of mean and dominant height performed best, whereas models of wood volume, basal area and DBH had lower predictive power (Fig. 2). Those findings match the results obtained for boreal forests in Southern Finland where mean height predictions based on Sentinel-2 data, had the lowest RMSE (30.4% compared to 24.5% in our results) and wood volume predictions the highest (59.3% compared to 70.3% in our results) (Astola et al., 2019). Similarly, models of forest growing stock volume based on Sentinel-2 data in a heterogeneous Mediterranean forest in northeastern Greece had an RMSE of 63.1 m<sup>3</sup>/ha (Chrysafis et al., 2017), compared to 70.3 m<sup>3</sup>/ha in our result. In Sentinel-2 models of vegetation height in Gabon (tropical Africa) and in Switzerland, low absolute RMSE of 5.6 and 3.4-m, respectively were achieved (Lang et al., 2019) compared to 1.8-m in our models. Options to improve modelling of wood volume, basal area and DBH could be to include additional predictor variables, such as soils and elevation that may vary more by region (Silveira et al., 2019b).

Among our predictors, the most important ones were the mean texture metric of VH polarization from SAR Sentinel-1 (S1\_VH\_MEAN) and longitude. In general, predictions of forest structure attributes based on radar data with HV and VH polarizations (cross-polarizations) yield better result than HH or VV polarizations (co-polarizations) (Rauste et al., 1994; Sinha et al., 2015). The reason is that co-polarized data are sensitive to the varying surface conditions and not as sensitive to radar scattering. Similar to our results, in Northeast China SAR Sentinel-1 VH polarizations had more robust and accurate results than VV polarization when mapping mean height and biomass (Liu et al., 2019). The VH polarization was also best for estimating tree height, basal area and total biomass in swamp forests in the coastal plains of southern Mississippi, USA (Wu and Sader, 1987). Similarly, in temperate forest of New England, USA, the VH polarization had higher potential than the VV polarization to estimate AGB (Huang et al., 2018). Longitude also plays an important role in our models, probably because it is correlated with temperature and precipitation gradients across the country (Matasci et al., 2018).

#### 4.2. Height comparison with GEDI and the GFCH

Our mean height predictions clearly outperformed those from GEDI and the GFCH. Our predictions based on SAR Sentinel-1, Sentinel-2, and geographic coordinates had considerably higher correlations ( $r = 0.89$  and  $0.85$  compared to  $0.81$ -GEDI and  $0.66$ -GFCH) and lower RMSE (19.7% and 22.9% compared to 53.5%-GEDI and 52.3%-GFCH) when related to independent validation plots.

While some of the lower accuracy of the GEDI/GFCH data may be due to forest growth and disturbance in the period between GEDI/GFCH data (2019–2020) and forest inventory surveys (2015–2020), another reason may be geolocation errors of GEDI (Guerra-Hernández and Pascual, 2021; Pascual et al., 2021). In addition, the signals returned from vegetated surfaces are also influenced by topography and forest structure (Guerra-Hernández and Pascual, 2021). For example, tall, dense canopies and steep slopes are challenging for GEDI and other Lidar missions such as ICESat-2 (Duncanson et al., 2020). Indeed, GEDI canopy height predictions are affected by slope, vegetation height, and beam sensitivity in forests, such as the temperate forests in central Germany (Adam et al., 2020). In our study area, GEDI and GFCH errors were higher on steep slopes than on lower slopes (Pearson correlation of  $0.44$  and  $0.37$ , respectively; Table S4), corroborating previous studies.

Thus, although GEDI and the GFCH provide great datasets for forest structure applications, topography, forest structure and the geolocation uncertainty must be taken into account when considering use of GEDI products (Fayad et al., 2021; Roy et al., 2021). For example, in spatially

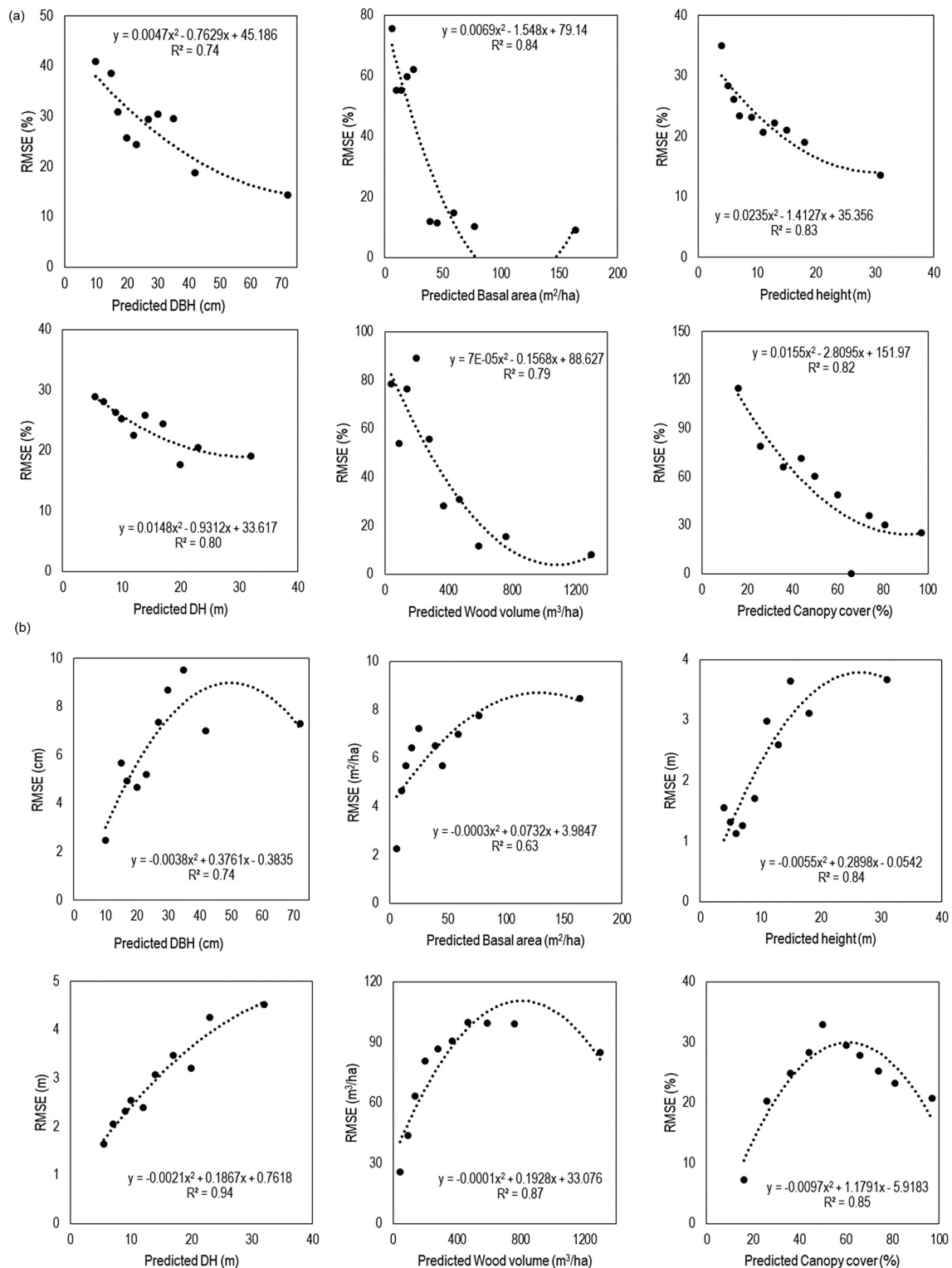


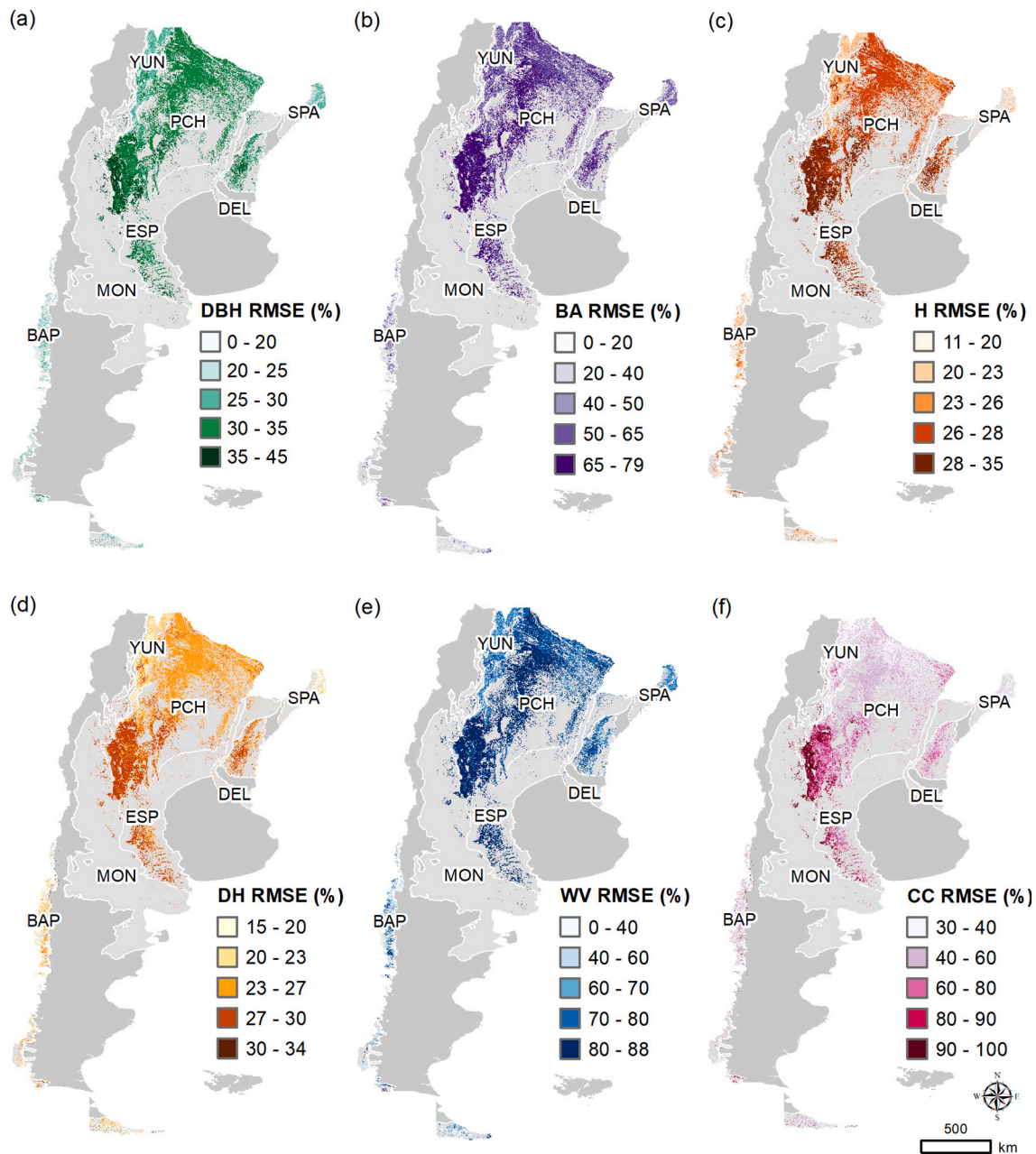
Fig. 5. Relative (a) and absolute (b) RMSE uncertainty of the predicted forest structure attributes with the polynomial function fitted.

heterogeneous canopies (e.g., secondary forests), unreliable GEDI canopy height retrievals are likely to be associated with geolocation uncertainty because of their spatially fragmented and heterogeneous three-dimensional structure. A recent preliminary assessment of the geolocation of GEDI indicates a  $1\sigma$  horizontal geolocation error of 23.8-m instead the geolocation requirement of 10-m (Beck et al., 2021). This means that reported and actual GEDI footprint location can be different by more than a GEDI 25-m footprint diameter suggesting caution in the

use of GEDI data acquired over spatially heterogeneous canopies, such as small forest stands, over small and/or fragmented features in the canopy, and over forest edges (Roy et al., 2021).

The future of mapping forest structure attributes is bright thanks to several upcoming satellite missions. For example, in 2023, the NASA-Indian Space Research Organization (ISRO) Synthetic Aperture Radar (NISAR) partnership mission plans to launch a SAR satellite sensitive to forest structure, which will provide an annual global land biomass map





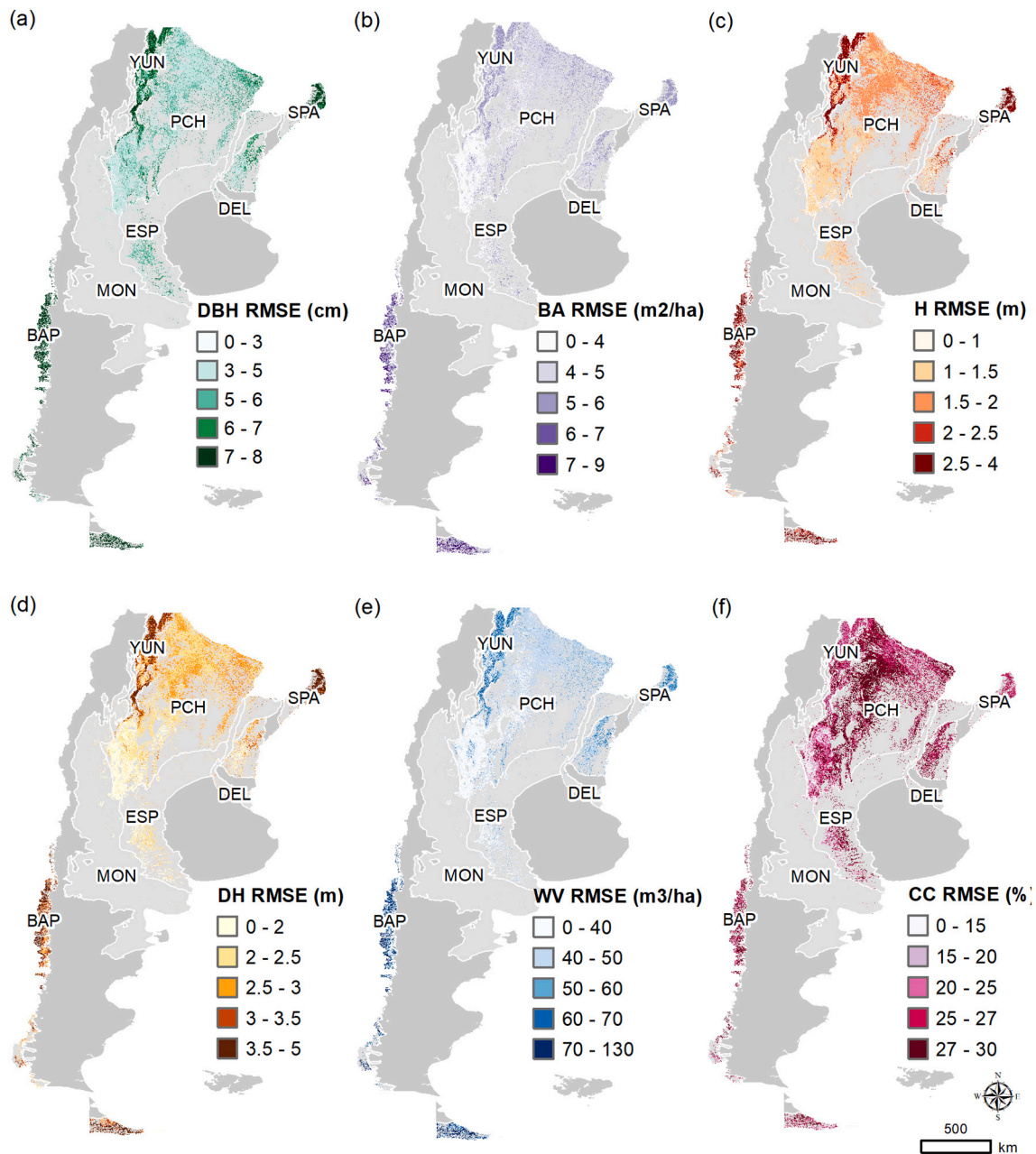
**Fig. 6.** Uncertainty (RMSE, in%) maps for (a) diameter at breast height (DBH), (b) basal area (BA), (c) mean height (H), (d) dominant height (DH), (e) wood volume (WV) and (f) canopy cover (CC).

at 1 ha spatial resolution with a RMSE of 20 Mg/ha for at least 80% of forests with AGB < 100 Mg/ha (given L-band limitations at higher biomass regimes) (Duncanson et al., 2020). Furthermore, the European Space Agency (ESA) Earth Explorer missions will include the BIOMASS mission. BIOMASS will implement the first P-band radar with multi-baseline interferometric and fully-polarimetric capabilities in space. This configuration will provide unprecedented sensitivity for characterizing forest aboveground biomass, and penetration capability down to the underlying terrain. Planned products include global 200-m resolution maps of both aboveground biomass and forest height, and a 50-m resolution map of forest disturbance (Quegan et al., 2019). Thus, in the near future, gains will be made by combining NISAR, BIOMASS, GEDI and other related data to predict forest structure attributes.

#### 4.3. Forest management and conservation implications

Our method of combining remote sensing with field measurements to estimate large-scale forest attributes retains the advantages of the accuracy of field measurements with the advantages of contiguous spatial coverage of remote sensing (Du et al., 2014; Wulder et al., 2008). Our approach results in relevant information for both forest management inventories (FMI) and national forest inventories (NFI). For example, FMI need accurate information of biometrics at stand level, which our 30-m predictions can provide, while NFI requires consistent broad-scale information, which our nation-wide maps offer.

Forest inventory data has historically been collected to assess the productivity of forests prior to harvesting, but now they are increasingly also used for monitoring wildlife habitat and forest biodiversity (Chirici et al., 2012; Corona et al., 2011; Lehtomäki et al., 2015). National-scale maps of forest structure attributes are also required to support science,

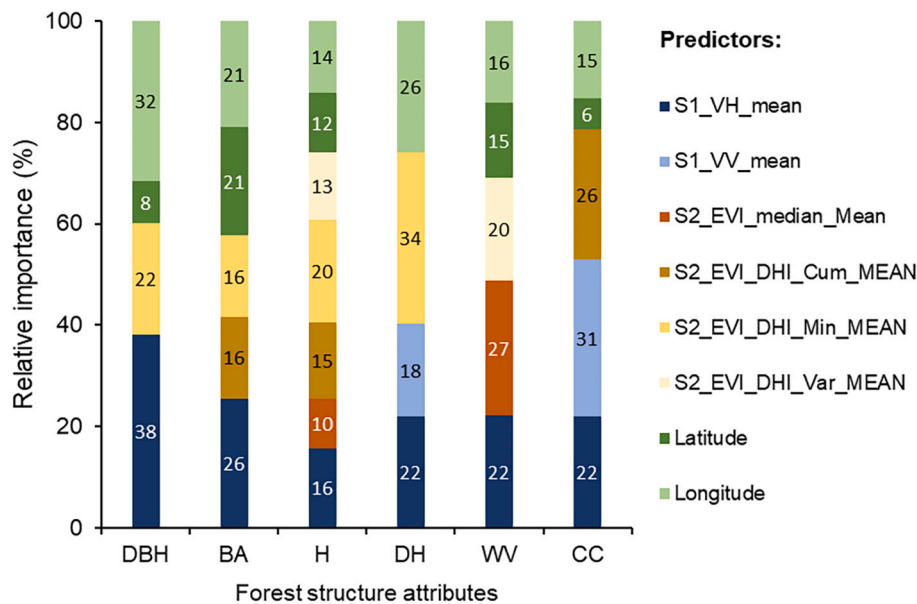


**Fig. 7.** Uncertainty (RMSE absolute) maps for (a) diameter at breast height (DBH), (b) basal area (BA), (c) mean height (H), (d) dominant height (DH), (e) wood volume (WV), and (f) canopy cover (CC).

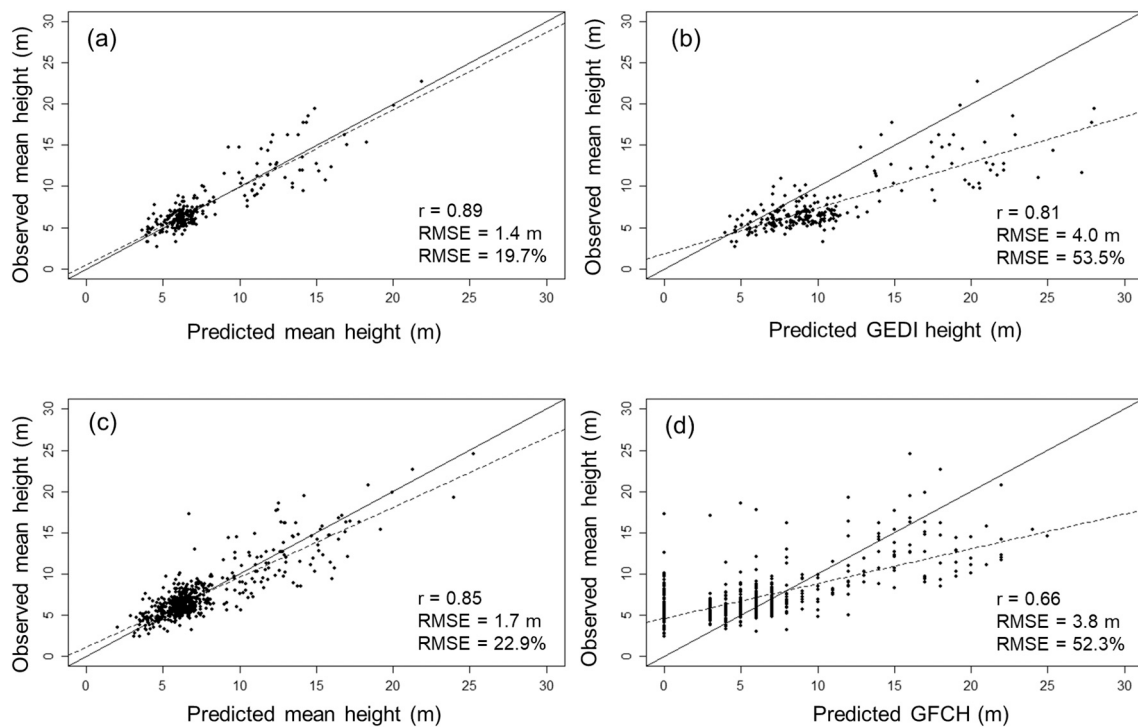
policy development and reporting obligations (White et al., 2014), which is another reason many countries have national forest inventories (NFIs). The products that we developed here (30-m resolution maps of DBH, basal area, mean height, dominant height, wood volume and canopy cover), amplify the value of the Second National Native Forest Inventory conducted by Argentine National Forest Service, and can support many applications including: (i) zoning of native forests in multiple uses according to different management and conservation criteria, (ii) evaluation of habitat quality for species associated with particular structures of the native forest; (iii) monitoring over time to achieve sustainable forest management; (iv) identification of priority areas for conservation or restoration; and (v) species distribution modelling.

The products can be useful for specific applications in each forest region, especially in Bosque Andino-Patagónico, Selva Paranaense, and Yungas forest regions (characterized by dense and tall forests), where

our approach had lower relative errors (Fig. 6; Fig. S5). For example, in Bosque Andino-Patagónico, the maps of forest structure can help to define the forested area in the province (e.g., tree-line forests), timber forests (e.g., up to 15-m dominant height and  $> 40 \text{ m}^2/\text{ha}$  of basal area), monitoring (e.g., forest recovery after harvest, fire and beaver impacts), silvopastoral planning (e.g., understory AGB based on dominant height and crown cover), and for identifying forest stands valuable for conservation purposes. Another specific application is in the Yungas region, where endemic threatened parrots depend on trees with cavities, and higher density of nests are found in forest stands with high DBH, basal area and tree height (Rivera et al., 2022). Thus, the information obtained in this study can be useful to model high quality habitat for species and for developing conservation and management strategies accordingly.



**Fig. 8.** Relative importance of the predictor variables in random forest models of diameter at breast height (DBH), basal area (BA), mean height (H), dominant height (DH), wood volume (WV), and canopy cover (CC).



**Fig. 9.** Scatterplots for the regression model of mean height from the second forest inventory survey versus predictions obtained from (a) 240 plots of Sentinel-1 plus Sentinel-2 derived features at 1-km resolution, (b) 240 plots of GEDI at 1-km resolution and (c) 550 plots of Sentinel-1 plus Sentinel-2 derived features at 30-m resolution, (d) and 550 plots of the GFCH at 30-m resolution.

## 5. Conclusion

In summary, we demonstrated the utility of combining extensive field inventory measures with features from SAR Sentinel-1, vegetation metrics from Sentinel-2 and geographic coordinates to map multiple forest structure attributes across the 463,000 km<sup>2</sup> of Argentina's forests. We developed an approach to obtain accurate maps at 30-m spatial resolution across large areas with highly heterogeneous and diverse forests and derived uncertainty maps of each forest structure attribute.

The best predictors of forest structure attributes were a combination of Sentinel-1 and 2, plus latitude and longitude. We predicted forest height at broad scale more accurately than available GEDI and GFCH datasets. Our models yielded reliable predictions of forest structure attributes, especially mean and dominant height, providing a strong basis for forest management and conservation planning across Argentina. The forest structure attribute maps are freely available at: [http://silvis.forest.wisc.edu/webmaps/forest\\_structure\\_maps\\_for\\_argentina/](http://silvis.forest.wisc.edu/webmaps/forest_structure_maps_for_argentina/) and at <https://doi.org/10.5061/dryad.vx0k6djwg>.



## Credit author statement

E.M.O-S was responsible for processing the remote sensing analysis and writing the paper. V.C.R and A.M.P were responsible for creating the research design, writing, and editing. S.M, G.J.M.P, J.B, N.P, L.N, L.O.R., L.C., Y.M.R., A.M.O., and G.I.P. were responsible for writing and editing. All authors discussed the results and contributed to the final manuscript.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

I have shared the link to my data in the manuscript

## Acknowledgements

We gratefully acknowledge support for this work by the National Aeronautics and Space Administration (NASA) Biodiversity and Ecological Forecasting Program, project 80NSSC19K0183 and the National Forest Service of the Ministry of the Environment and Sustainable Development of Argentina for the database of the Second National Native Forest Inventory and support the analyses. Three anonymous reviewers provided valuable comments that greatly improved our manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.rse.2022.113391>.

## References

- Adam, M., Urbazaev, M., Dubois, C., Schmullius, C., 2020. Accuracy assessment of GEDI terrain elevation and canopy height estimates in European temperate forests: influence of environmental and acquisition parameters. *Remote Sens.* 12, 1–28.
- Astola, H., Häme, T., Sirro, L., Molinier, M., Kilpi, J., 2019a. Comparison of Sentinel-2 and landsat 8 imagery for forest variable prediction in boreal region. *Remote Sens. Environ.* 223, 257–273.
- Avitabile, V., Herold, M., Heuvelink, G.B.M., Lewis, S.L., Phillips, O.L., Asner, G.P., Armston, J., Ashton, P.S., Banin, L., Bayol, N., Berry, N.J., Boeckx, P., de Jong, B.H. J., Devries, B., Girardin, C.A.J., Kearsley, E., Lindsell, J.A., Lopez-Gonzalez, G., Lucas, R., Malhi, Y., Morel, A., Mitchard, E.T.A., Nagy, L., Qie, L., Quinones, M.J., Ryan, C.M., Ferry, S.J.W., Sunderland, T., Laurin, G.V., Gatti, R.C., Valentini, R., Verbeeck, H., Wijaya, A., Willcock, S., 2016. An integrated pan-tropical biomass map using multiple reference datasets. *Glob. Chang. Biol.* 22, 1406–1420.
- Baccini, A., Laporte, N., Goetz, S.J., Sun, M., Dong, H., 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. *Environ. Res. Lett.* 3, 045011.
- Beck, J., Wirt, B., Luthcke, S., Hofton, M., Armston, J., 2021. Global Ecosystem dynamic investigation (GEDI) level 02 user guide. Document version 2.0, April 2021. US Geological Survey, Earth Resources Observation and Science Center (Sioux Falls, South Dakota, USA).
- Beaudoin, A., Bernier, P.Y., Guindon, L., Villemaire, P., Guo, X.J., Stinson, G., Bergeron, T., Magnussen, S., Hall, R.J., 2014. Mapping attributes of Canada's forests at moderate resolution through kNN and MODIS imagery. *Can. J. For. Res.* 44, 521–532.
- Bergstra, J., Bengio, Y., 2012. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* 13, 281–305.
- Berry, S., Mackey, B., Brown, T., 2007. Potential applications of remotely sensed vegetation greenness to habitat analysis and the conservation of dispersive fauna. *Pac. Conserv. Biol.* 13, 120–127.
- Blackard, J.A., Finco, M.V., Helmer, E.H., Holden, G.R., Hoppus, M.L., Jacobs, D.M., Lister, A.J., Moisen, G.G., Nelson, M.D., Riemann, R., Ruefenacht, B., Salajano, D., Weyerermann, D.L., Winterberger, K.C., Brandeis, T.J., Czaplewski, R.L., McRoberts, R.E., Patterson, P.L., Tymcio, R.P., 2008. Mapping U.S. Forest biomass using nationwide forest inventory data and moderate resolution information. *Remote Sens. Environ.* 112, 1658–1677.
- Bouvier, M., Durrieu, S., Fournier, R.A., Renaud, J.P., 2015. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote Sens. Environ.* 156, 322–334. <https://doi.org/10.1016/j.rse.2014.10.004>.
- Breiman, L., 2001. In: *Random Forests Machine Learning*, 45, pp. 5–32.
- Bruggisser, M., Dorigo, W., Dostálová, A., Hollaus, M., Navacchi, C., Schläpfer, S., Pfeifer, N., 2021. Potential of sentinel-1 c-band time series to derive structural parameters of temperate deciduous forests. *Remote Sens.* 13, 1–30.
- Castillo, J.A.A., Apan, A.A., Maraseni, T.N., Salmo, S.G., 2017. Estimation and mapping of above-ground biomass of mangrove forests and their replacement land uses in the Philippines using sentinel imagery. *ISPRS J. Photogramm. Remote Sens.* 134, 70–85.
- Chandrashekar, G., Sahin, F., 2014. A survey on feature selection methods. *Comput. Electr. Eng.* 40, 16–28.
- Chen, L., Ren, C., Zhang, B., Wang, Z., Xi, Y., 2018. Estimation of forest above-ground biomass by geographically weighted regression and machine learning with sentinel imagery. *Forests* 9, 1–20.
- Chirici, G., McRoberts, R.E., Winter, S., Bertini, R., Bröandli, U.B., Asensio, I.A., Bastrup-Birk, A., Rondeux, J., Barsoum, N., Marchetti, M., 2012. National forest inventory contributions to forest biodiversity monitoring. *For. Sci.* 58, 257–268.
- Chrysafis, I., Mallinis, G., Siachalou, S., Patias, P., 2017. Assessing the relationships between growing stock volume and Sentinel-2 imagery in a Mediterranean forest ecosystem. *Remote Sens. Lett.* 8, 508–517.
- Coops, N.C., Tompalski, P., Goodbody, T.R.H., Queinnee, M., Luther, J.E., Bolton, D.K., White, J.C., Wulder, M.A., van Lier, O.R., Hermosilla, T., 2021. Modelling lidar-derived estimates of forest attributes over space and time: a review of approaches and future trends. *Remote Sens. Environ.* 260, 112477.
- Coops, N.C., Wulder, M.A., Iwanicka, D., 2009. Demonstration of a satellite-based index to monitor habitat at continental-scales. *Ecol. Indic.* 9, 948–958.
- Corona, P., Chirici, G., McRoberts, R.E., Winter, S., Barbati, A., 2011. Contribution of large-scale forest inventories to biodiversity assessment and monitoring. *For. Ecol. Manag.* 262, 2061–2069.
- Coulston, J.W., Blinn, C.E., Thomas, V.A., Wynne, R.H., 2016. Approximating prediction uncertainty for random Forest regression models. *Photogramm. Eng. Remote Sens.* 82 (3), 189–197.
- Csillik, O., Kumar, P., Mascaro, J., O'Shea, T., Asner, G.P., 2019. Monitoring tropical forest carbon stocks and emissions using planet satellite data. *Sci. Rep.* 9, 1–12.
- Csillik, O., Kumar, P., Asner, G.P., 2020. Challenges in estimating tropical forest canopy height from planet dove imagery. *Remote Sens.* 12, 1–18.
- Dirección Nacional de Bosques, 2021. Datos del Segundo Inventario Nacional de Bosques Nativos de la República Argentina. Ministerio de Ambiente y Desarrollo Sostenible de la Nación.
- Dobson, M.C., Ulaby, F.T., Pierce, L.E., Sharik, T.L., Bergen, K.M., Kellndorfer, J., Kendra, J.R., Li, E., Lin, Y.C., Nashashibi, A., Sarabandi, K., Siqueira, P., 1995. Estimation of forest biophysical characteristics in northern Michigan with SIR-C/X-SAR. *IEEE Trans. Geosci. Remote Sens.* 33, 877–895.
- Dos Reis, A.A., Werner, J.P.S., Silva, B.C., Figueiredo, G.K.D.A., Antunes, J.F.G., Esquerdo, J.C.D.M., Coutinho, A.C., Lamparelli, R.A.C., Rocha, J.V., Magalhães, P.S. G., 2020. Monitoring pasture aboveground biomass and canopy height in an integrated crop-livestock system using textural information from PlanetScope imagery. *Remote Sens.* 12.
- Du, L., Zhou, T., Zou, Z., Zhao, X., Huang, K., Wu, H., 2014. Mapping forest biomass using remote sensing and national forest inventory in China. *Forests* 5, 1267–1283.
- Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The global ecosystem dynamics investigation: high-resolution laser ranging of the Earth's forests and topography. *Sci. Remote Sens.* 1, 100002.
- Dubayah, R.O., Armston, J., Kellner, J.R., Duncanson, L., Healey, S.P., Patterson, P.L., Hancock, S., Tang, H., Hofton, M.A., Blair, J.B., Luthcke, S.B., 2021. GEDI L4A Footprint Level Aboveground Biomass Density, Version 1. Oak Ridge, Tennessee, USA.
- Dubayah, R.O., Luthcke, S.B., Sabaka, T.J., Nicholas, J.B., Preaux, S., Hofton, M.A., 2021. GEDI L3 Gridded Land Surface Metrics, Version 1. ORNL DAAC, Oak Ridge, Tennessee, USA.
- Dudani, S.A., 1976. The distance-weighted k-nearest-neighbor rule. *IEEE Trans. Syst. Man Cybernet.* 4, 325–327.
- Duncanson, L., Neuenschwander, A., Hancock, S., Thomas, N., Fatoyinbo, T., Simard, M., Silva, C.A., Armston, J., Luthcke, S.B., Hofton, M., Kellner, J.R., Dubayah, R., 2020. Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California. *Remote Sens. Environ.* 242, 111779.
- Duro, D.C., Coops, N.C., Wulder, M.A., Han, T., 2007. Development of a large area biodiversity monitoring system driven by remote sensing. *Prog. Phys. Geogr.* 31, 235–260.
- Eskandari, S., Jaafari, M.R., Oliva, P., Ghorbanzadeh, O., Blaschke, T., 2020. Mapping land cover and tree canopy cover in zagros forests of Iran: application of Sentinel-2, google earth, and field data. *Remote Sens.* 12, 1912.
- Fassnacht, F.E., Hartig, F., Latifi, H., Berger, C., Hernández, J., Corvalán, P., 2014. Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sens. Environ.* 154, 102–114.
- Fayad, I., Baghdadi, N.N., Alvares, C.A., Stape, J.L., Bailly, J.S., Scolforo, H.F., Zribi, M., Maire, G.L., 2021. Assessment of GEDI's LiDAR data for the estimation of Canopy Heights and wood volume of eucalyptus plantations in Brazil. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 14, 7095–7110.
- Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315.
- Forkuor, G., Benewinde Zougrana, J.B., Dimobe, K., Ouattara, B., Vadrevu, K.P., Tondoh, J.E., 2020. Above-ground biomass mapping in west african dryland forest using Sentinel-1 and 2 datasets - a case study. *Remote Sens. Environ.* 236, 111496.

- Gasparri, N.I., Baldi, G., 2013. Regional patterns and controls of biomass in semiarid woodlands: lessons from the northern Argentina dry Chaco. *Reg. Environ. Chang.* 13, 1131–1144.
- Gasparri, N.I., Parmuchi, M.G., Bono, J., Karszenbaum, H., Montenegro, C.L., 2010. Assessing multi-temporal landsat 7 ETM+ images for estimating above-ground biomass in subtropical dry forests of Argentina. *J. Arid Environ.* 74, 1262–1270.
- Ghasemi, N., Sahebi, M.R., Mohammadzadeh, A., 2011. A review on biomass estimation methods using synthetic aperture radar data. *Int. J. Geomat. Geosci.* 1, 776–788.
- Gregoire, T.G., Ståhl, G., Næsset, E., Gobakken, T., Nelson, R., Holm, S., 2011. Model-assisted estimation of biomass in a LiDAR sample survey in Hedmark County, Norway. *Can. J. For. Res.* 41 (1), 83–95.
- Guerra-Hernández, J., Pascual, A., 2021. Using GEDI lidar data and airborne laser scanning to assess height growth dynamics in fast-growing species: a showcase in Spain. *For. Ecosyst.* 8, 14.
- Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342, 850–853.
- Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. *IEEE Trans. Syst. Man Cybernet.* 3, 610–621.
- Hauglin, M., Rahlf, J., Schumacher, J., Astrup, R., Breidenbach, J., 2021. Large scale mapping of forest attributes using heterogeneous sets of airborne laser scanning and National Forest Inventory data. *For. Ecosyst.* 8, 65.
- Huang, X., Ziniti, B., Torbick, N., Ducey, M.J., 2018. Assessment of forest above ground biomass estimation using multi-temporal C-band Sentinel-1 and polarimetric L-band PALSAR-2 data. *Remote Sens.* 10, 1424.
- Hudak, A.T., Lefsky, M.A., Cohen, W.B., Berterretche, M., 2002. Integration of lidar and landsat ETM+ data for estimating and mapping forest canopy height. *Remote Sens. Environ.* 82, 397–416.
- Hyde, P., Dubayah, R., Walker, W., Blair, J.B., Hofton, M., Hunsaker, C., 2006. Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sens. Environ.* 102, 63–73.
- Karjalainen, M., Kankare, V., Vastaranta, M., Holopainen, M., Hyypä, J., 2012. Prediction of plot-level forest variables using TerraSAR-X stereo SAR data. *Remote Sens. Environ.* 117, 338–347.
- Kononenko, I., Robnik-Šikonja, M., Pompe, U., 1996. Relief for estimation and discretization of attributes in classification, regression, and ILP problems. *Artif. Intell. Methodol. Syst. Appl.* 18, 1–15.
- Korhonen, L., Hadi, Packalen, P., Rautiainen, M., 2017. Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index. *Remote Sens. Environ.* 195, 259–274.
- Lang, N., Schindler, K., Wegner, J.D., 2019. Country-wide high-resolution vegetation height mapping with Sentinel-2. *Remote Sens. Environ.* 233, 111347.
- Laurin, G.V., Balling, J., Corona, P., Mattioli, W., Papale, D., Puletti, N., Rizzo, M., Truckenbrodt, J., Urban, M., 2018. Above-ground biomass prediction by Sentinel-1 multitemporal data in Central Italy with integration of ALOS2 and Sentinel-2 data. *J. Appl. Remote. Sens.* 12 (1), 016008.
- Latifi, H., Koch, B., 2012. Evaluation of most similar neighbour and random forest methods for imputing forest inventory variables using data from target and auxiliary stands. *Int. J. Remote Sens.* 33, 6668–6694.
- Lehtomäki, J., Tuominen, S., Toivonen, T., Leinonen, A., 2015. What data to use for forest conservation planning? A comparison of coarse open and detailed proprietary forest inventory data in Finland. *PLoS ONE* 10, 1–25.
- Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. *R News* 2, 18–22.
- Liu, A., Cheng, X., Chen, Z., 2021. Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. *Remote Sens. Environ.* 264, 112571.
- Liu, Y., Gong, W., Xing, Y., Hu, X., Gong, J., 2019. Estimation of the forest stand mean height and aboveground biomass in Northeast China using SAR sentinel-1B, multispectral sentinel-2A, and DEM imagery. *ISPRS J. Photogramm. Remote Sens.* 151, 277–289.
- Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., Moran, E., 2016. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *Int. J. Digit. Earth* 9, 63–105.
- Magnussen, S., 2015. Arguments for a model-dependent inference? *Int. J. Res.* 88 (3), 317–325.
- Majasalmi, T., Rautiainen, M., 2016. The potential of Sentinel-2 data for estimating biophysical variables in a boreal forest: a simulation study. *Remote Sens. Lett.* 7, 427–436.
- Matasci, G., Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Zald, H.S.J., 2018. Large-area mapping of Canadian boreal forest cover, height, biomass and other structural attributes using landsat composites and lidar plots. *Remote Sens. Environ.* 209, 90–106.
- Matteucci, S.D., Martinez-Pastur, G., Lencinas, M.V., Rovere, A.E., Amoroso, M.M., Barberis, I., Vesprini, J.L., Galetto, L., Torres, C., Villagra, P.E., Malizia, L.R., Blundo, C., Politi, N., Peri, P.L., 2021. Breve descripción de las regiones forestales de la Argentina. Capítulo 2 In: Peri, P.L., Pastur, G. Martínez, Schlichter, T. (Eds.), *En: Uso sustentable del bosque: Aportes desde la silvicultura argentina*. Ministerio de Ambiente y Desarrollo Sostenible de la Nación Argentina, Buenos Aires, Argentina, pp. 48–77 (ISBN978-987-46815-4-6).
- Moreno, A., Neumann, M., Hasenauer, H., 2016. Optimal resolution for linking remotely sensed and forest inventory data in Europe. *Remote Sens. Environ.* 183, 109–119.
- Morin, D., Planells, M., Guyon, D., Villard, L., Mermoz, S., Bouvet, A., Thevenon, H., Dejoux, J.F., Le Toan, T., Dedieu, G., 2019. Estimation and mapping of forest structure parameters from open access satellite images: development of a generic method with a study case on coniferous plantation. *Remote Sens.* 11, 1275.
- Müller-Wilm, U., 2016. In: Sentinel-2 MSI – Level-2A Prototype Processor Installation and User Manual. European Space Agency, (Special Publication) ESA SP 49, pp. 1–51.
- Navarro, J.A., Algeet, N., Fernández-Landa, A., Esteban, J., Rodríguez-Noriega, P., Guillén-Climent, M.L., 2019. Integration of UAV, Sentinel-1, and Sentinel-2 data for mangrove plantation aboveground biomass monitoring in Senegal. *Remote Sens.* 11, 1–23.
- Nizalapur, V., Jha, C.S., Madugundu, R., 2010. Estimation of above ground biomass in indian tropical forested area using multi frequency DLR ESAR data. *Int. J. Geomat. Geosci.* 1, 167–178.
- Nuthammachot, N., Askar, A., Stratoulas, D., Wicaksono, P., 2020. Combined use of Sentinel-1 and Sentinel-2 data for improving above-ground biomass estimation. *Geocarto Int.* 37 (2), 366–376.
- Pascual, A., Guerra-Hernández, J., Cosenza, D.N., Sandoval-Altalerrea, V., 2021. Using enhanced data co-registration to update spanish National Forest Inventories (NFI) and to reduce training data under LiDAR-assisted inference. *Int. J. Remote Sens.* 42, 106–127.
- Peri, P.L., Nahuelhual, L., Martinez-Pastur, G., 2021. Ecosystem services in patagonia: a multi-criteria approach for an integrated assessment. In: Springer Nature: Natural and Social Sciences of Patagonia. Springer, Cham, Switzerland.
- Periasamy, S., 2018. Significance of dual polarimetric synthetic aperture radar in biomass retrieval: an attempt on Sentinel-1. *Remote Sens. Environ.* 217, 537–549.
- Potapov, P., Hansen, M.C., Kommareddy, I., Kommareddy, A., Turubanova, S., Pickens, A., Adusei, B., Tyukavina, A., Ying, Q., 2020. Landsat analysis ready data for global land cover and land cover change mapping. *Remote Sens.* 12, 30426.
- Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M.C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C.E., Armston, J., Dubayah, R., Blair, J. B., Hofton, M., 2021. Mapping global forest canopy height through integration of GEDI and landsat data. *Remote Sens. Environ.* 253, 112165.
- Potapov, P., Tyukavina, A., Turubanova, S., Talero, Y., Hernandez-Serna, A., Hansen, M. C., Saah, D., Tenneson, K., Poortinga, A., Aekakkararungroj, A., Chishtie, F., Towashiraporn, P., Bhandari, B., Aung, K.S., Nguyen, Q.H., 2019. Annual continuous fields of woody vegetation structure in the lower Mekong region from 2000–2017 landsat time-series. *Remote Sens. Environ.* 232, 111278.
- Pötzschner, F., Baumann, M., Gasparri, N.I., Conti, G., Loto, D., Piquer-Rodríguez, M., Kuemmerle, T., 2022. Ecoregion-wide, multi-sensor biomass mapping highlights a major underestimation of dry forests carbon stocks. *Remote Sens. Environ.* 269, 112849.
- Puliti, S., Saarela, S., Gobakken, T., Ståhl, G., Næsset, E., 2018. Combining UAV and Sentinel-2 auxiliary data for forest growing stock volume estimation through hierarchical model-based inference. *Remote Sens. Environ.* 204, 485–497.
- Puliti, S., Hauglin, M., Breidenbach, J., Montesano, P., Neigh, C.S.R., Rahlf, J., Solberg, S., Klingenberg, T.F., Astrup, R., 2020. Modelling above-ground biomass stock over Norway using national forest inventory data with ArcticDEM and Sentinel-2 data. *Remote Sens. Environ.* 36, 111501.
- Quegan, S., Le Toan, T., Chave, J., Dall, J., Exbrayat, J.F., Minh, D.H.T., Lomas, M., D'Alessandro, M.M., Paillou, P., Papathanassiou, K., Rocca, F., Saatchi, S., Scipal, K., Shugart, H., Smallman, T.L., Soja, M.J., Tebaldini, S., Ulander, L., Villard, L., Williams, M., 2019. The European Space Agency BIOMASS mission: measuring forest above-ground biomass from space. *Remote Sens. Environ.* 227, 44–60.
- R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria <https://cran.r-project.org/web/packages/FSelector/FSelector.pdf>.
- Rauste, Y., Häme, T., Pulliainen, J., Heiska, K., Hallikainen, M., 1994. Radar-based forest biomass estimation. *Int. J. Remote Sens.* 15, 2797–2808.
- Rivera, L., Politi, N., Bucher, E.H., Pidgeon, A., 2022. Effect of forest logging on food availability, suitable nesting habitat, nest density and spatial pattern of a neotropical parrot. *For. Ecol. Manag.* 507, 120005.
- Roy, D.P., Kashongwe, H.B., Armston, J., 2021. The impact of geolocation uncertainty on GEDI tropical forest canopy height estimation and change monitoring. *Sci. Remote Sens.* 4, 100024.
- Saarela, S., Grafström, A., Ståhl, G., Kangas, A., Holopainen, M., Tuominen, S., Nordkvist, K., Hyypä, J., 2015. Model-assisted estimation of growing stock volume using different combinations of LiDAR and landsat data as auxiliary information. *Remote Sens. Environ.* 158, 431–440.
- Saatchi, S.S., 1997. Coherent effects in microwave backscattering models for forest canopies. *IEEE Trans. Geosci. Remote Sens.* 35, 1032–1044.
- Saatchi, S.S., Harris, N.L., Brown, S., Lefsky, M., Mitchard, E.T.A., Salas, W., Zutta, B.R., Buermann, W., Lewis, S.L., Hagen, S., Petrova, S., White, L., Silman, M., Morel, A., 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. In: Proceedings of the National Academy of Sciences of the United States of America, 108, pp. 9899–9904.
- SGAyDS-Secretaría de Gobierno de Ambiente y Desarrollo Sustentable de la Nación, 2019. Segundo Inventario Nacional de Bosques Nativos: manual de campo. Buenos Aires: Secretaría de Gobierno de Ambiente y Desarrollo Sustentable de la Nación, Buenos Aires, Argentina.
- Shataee, S., Kalbi, S., Fallah, A., Pelz, D., 2012. Forest attribute imputation using machine-learning methods and ASTER data: comparison of k-NN, SVR and random forest regression algorithms. *Int. J. Remote Sens.* 33, 6254–6280. <https://doi.org/10.1080/01431161.2012.682661>.
- Silva, C.A., Duncanson, L., Hancock, S., Neuenschwander, A., Thomas, N., Hofton, M., Fatoyinbo, L., Simard, M., Marshak, C.Z., Armston, J., Lutchke, S., Dubayah, R., 2021. Fusing simulated GEDI, ICESat-2 and NISAR data for regional aboveground biomass mapping. *Remote Sens. Environ.* 253, 112234.

- Silveira, E.M.O., Terra, M.C.N.S., ter Steege, H., Maeda, E.E., Acerbi Júnior, F.W., Scolforo, J.R.S., 2019a. Carbon-diversity hotspots and their owners in Brazilian southeastern savanna, Atlantic Forest and semi-arid woodland domains. *For. Ecol. Manag.* 452, 117575.
- Silveira, E.M.O., Cunha, L.I.F., Galvão, L.S., Withey, K.D., Acerbi Júnior, F.W., Scolforo, J.R.S., 2019b. Modelling aboveground biomass in forest remnants of the Brazilian Atlantic Forest using remote sensing, environmental and terrain-related data. *Geocarto Int.* 36 (3), 281–298.
- Silveira, E.M.O., Espírito Santo, F.D., Wulder, M.A., Acerbi Júnior, F.W., Carvalho, M.C., Mello, C.R., Mello, J.M., Shimabukuro, Y.E., Terra, M.C.N.S., Carvalho, L.M.T., Scolforo, J.R.S., 2019c. Pre-stratified modelling plus residuals kriging reduces the uncertainty of aboveground biomass estimation and spatial distribution in heterogeneous savannas and forest environments. *For. Ecol. Manag.* 445, 96–109.
- Silveira, E.M.O., Radeloff, V.C., Martínez Pastur, G.J., Martinuzzi, S., Politi, N., Lizarraga, L., Rivera, L.O., Gavier-Pizarro, G.I., Yin, H., Rosas, Y.M., Calamari, N.C., Navarro, M.F., Sica, Y., Olah, A.M., Bono, J., Pidgeon, A.M., 2022. Forest phenoclusters for Argentina based on vegetation phenology and climate. *Ecol. Appl.* 32 (3), 1–21.
- Sinha, S., Jeganathan, C., Sharma, L.K., Nathawat, M.S., 2015. A review of radar remote sensing for biomass estimation. *Int. J. Environ. Sci. Technol.* 12, 1779–1792.
- Tompalski, P., White, J.C., Coops, N.C., Wulder, M.A., 2019. Demonstrating the transferability of forest inventory attribute models derived using airborne laser scanning data. *Remote Sens. Environ.* 227, 110–124.
- Tomppo, E., Olsson, H., Ståhl, G., Nilsson, M., Hagner, O., Katila, M., 2008. Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sens. Environ.* 112, 1982–1999.
- Urbanowicz, R.J., Meeker, M., La Cava, W., Olson, R.S., Moore, J.H., 2018. Relief-based feature selection: introduction and review. *J. Biomed. Inform.* 85, 189–203.
- Wang, Y.R., Li, X.M., 2021. Arctic Sea ice cover data from spaceborne synthetic aperture radar by deep learning. *Earth Syst. Sci. Data* 13, 2723–2742.
- White, J.C., Wulder, M.A., Hobart, G.W., Luther, J.E., Hermosilla, T., Griffiths, P., Coops, N.C., Hall, R.J., Hostert, P., Dyk, A., Guindon, L., 2014. Pixel-based image compositing for large-area dense time series applications and science. *Can. J. Remote. Sens.* 40, 192–212.
- Wittke, S., Yu, X., Karjalainen, M., Hyyppä, J., Puttonen, E., 2019. Comparison of two-dimensional multitemporal Sentinel-2 data with three-dimensional remote sensing data sources for forest inventory parameter estimation over a boreal forest. *Int. J. Appl. Earth Obs. Geoinf.* 76, 167–178.
- Wu, S.T., Sader, S.A., 1987. Multipolarization SAR data for surface feature delineation and Forest vegetation characterization. *IEEE Trans. Geosci. Remote Sens.* GE-25, 67–76.
- Wulder, M.A., White, J.C., Fournier, R.A., Luther, J.E., Magnussen, S., 2008. Spatially explicit large area biomass estimation: three approaches using forest inventory and remotely sensed imagery in a GIS. *Sensors* 8, 529–560.
- Wulder, M.A., Hermosilla, T., White, J.C., Coops, N.C., 2020. Biomass status and dynamics over Canada's forests: disentangling disturbed area from associated aboveground biomass consequences. *Environ. Res. Lett.* 15, 094093.
- Zald, H.S.J., Wulder, M.A., White, J.C., Hilker, T., Hermosilla, T., Hobart, G.W., Coops, N.C., 2016. Integrating landsat pixel composites and change metrics with lidar plots to predictively map forest structure and aboveground biomass in Saskatchewan, Canada. *Remote Sens. Environ.* 176, 188–201.
- Yu, Y., Saatchi, S., Domke, G.M., Walters, B., Woodall, C., Ganguly, S., Li, S., Kalia, S., Park, T., Nemani, R., Hagen, S.C., Melendy, L., 2022. Making the US national forest inventory spatially contiguous and temporally consistent. *Environ. Res. Lett.* 17, 065002.