This is a repository copy of *Forest biomass retrieval approaches from earth observation in different biomes*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/142129/

Version: Published Version

**Article:**

# Forest biomass retrieval approaches from earth observation in different biomes

Pedro Rodríguez-Veiga[a,b,*], Shaun Quegan[b,c], Joao Carreiras[b,c], Henrik J. Persson[d],
Johan E.S. Fransson[d], Agata Hoscilo[e], Dariusz Ziółkowski[e], Krzysztof Stereńczak[f],
Sandra Lohberger[g], Matthias Stängel[g], Anna Berninger[g], Florian Siegert[g,h], Valerio Avitabile[i],
Martin Herold[i], Stéphane Mermoz[j], Alexandre Bouvet[j], Thuy Le Toan[j], Nuno Carvalhais[k,l],
Maurizio Santoro[m], Oliver Cartus[m], Yrjö Rauste[n], Renaud Mathieu[o,p], Gregory P. Asner[q],
Christian Thiel[r], Carsten Pathe[s], Chris Schmullius[s], Frank Martin Seifert[t], Kevin Tansey[a],
Heiko Balzter[a,b]

[a] University of Leicester, Centre for Landscape and Climate Research, United Kingdom
[b] National Centre for Earth Observation, United Kingdom
[c] University of Sheffield, United Kingdom
[d] Sveriges Lantbruksuniversitet, Sweden
[e] Institute of Geodesy and Cartography, Poland
[f] Forest Research Institute, Poland
[g] Remote Sensing Solutions, Germany
[h] Ludwig-Maximilians-University Munich, Germany
[i] Wageningen University & Research, the Netherlands
[j] CESBIO, Université de Toulouse, CNES, CNRS, IRD, UPS, France
[k] Max Planck Institute for Biogeochemistry, Germany
[l] Universidade NOVA de Lisboa, CENSE, Departamento de Ciências e Engenharia do Ambiente, Faculdade de Ciências e Tecnologia, Portugal
[m] GAMMA Remote Sensing, Switzerland
[n] VTT Technical Research Centre of Finland Ltd, Finland
[o] CSIR, South Africa
[p] University of Pretoria, Department of Geography, Geomatics and Meteorology, South Africa
[q] Carnegie Institution for Science, USA
[r] German Aerospace Agency, Germany
[s] Friedrich-Schiller-University Jena, Germany
[t] European Space Agency ESRIN, Italy

## ARTICLE INFO

## ABSTRACT

The amount and spatial distribution of forest aboveground biomass (AGB) were estimated using a range of regionally developed methods using Earth Observation data for Poland, Sweden and regions in Indonesia (Kalimantan), Mexico (Central Mexico and Yucatan peninsula), and South Africa (Eastern provinces) for the year 2010. These regions are representative of numerous forest biomes and biomass levels globally, from South African woodlands and savannas to the humid tropical forest of Kalimantan. AGB retrieval in each region relied on different sources of reference data, including forest inventory plot data and airborne LiDAR observations, and used a range of retrieval algorithms. This is the widest inter-comparison of regional-to-national AGB maps to date in terms of area, forest types, input datasets, and retrieval methods. The accuracy assessment of all regional maps using independent field data or LiDAR AGB maps resulted in an overall root mean square error (RMSE) ranging from 10 t ha$^{-1}$ to 55 t ha$^{-1}$ (37% to 67% relative RMSE), and an overall bias ranging from $-1$ t ha$^{-1}$ to $+5$ t ha$^{-1}$ at pixel level. The regional maps showed better agreement with field data than previously developed and widely used pan-tropical or northern hemisphere datasets. The comparison of accuracy assessments showed commonalities in error structures despite the variety of methods, input data, and forest biomes. All regional retrievals resulted in overestimation (up to 63 t ha$^{-1}$) in the lower AGB classes, and underestimation (up to 85 t ha$^{-1}$) in the higher AGB classes. Parametric model-based algorithms present advantages due to their low

demand on in situ data compared to non-parametric algorithms, but there is a need for datasets and retrieval methods that can overcome the biases at both ends of the AGB range. The outcomes of this study should be considered when developing algorithms to estimate forest biomass at continental to global scale level.

## 1. Introduction

### 1.1. Background

Forests cover around one third of the Earth's land surface, are an essential socio-cultural element of modern society, support biodiversity and influence the climate system via coupled carbon-water-energy cycles (Bonan, 2008). Quantifying forest aboveground woody biomass (AGB), i.e. the amount of woody matter within a forest, has profound social and economic importance, since it is a source of materials and energy for direct human use, and its structure and temporal dynamics exert substantial influence on the functioning of terrestrial ecosystems, with direct impacts on biodiversity, as well as on the carbon and energy cycles and consequently the whole Earth system (e.g. Bonan, 2008; Le Quéré et al., 2018; Pan et al., 2011). As such, AGB can be used to evaluate the dynamics of global vegetation and Earth system models (e.g. Thurner et al., 2017; Carvalhais et al., 2014), was recognised by the Global Climate Observing System (GCOS) as an Essential Climate Variable (ECV) (Bojinski et al., 2014), and plays an important role in several essential biodiversity variables (EBV) (Pereira et al., 2013). However, quantification of AGB still presents a scientific challenge with significant implications for our current knowledge about the Earth system (Pan et al., 2011; Le Quéré et al., 2018).

Knowledge of the spatial distribution of forest AGB is typically derived from ground measurements collected by national forest inventories. From these, regional- to national-scale summary data are generated for the FAO's quinquennial Global Forest Resource Assessment (FRA) reports (FAO, 2005, 2010; FAO, 2015), aiming at giving a global portrait of biomass stocks and their changes in time. Vast areas covered by forests mean that ground-based forest inventories need a large amount of resources to provide accurate information on the extent, spatial distribution and dynamics of forest AGB. However, forest inventory data in developing countries can be fairly inaccurate (Saatchi and Moghaddam, 2000) and often many years out of date (Shvidenko and Schepaschenko, 2014). A review of the country FRA reports (FAO, 2010) showed that 45 countries (i.e. around 20%) indicated high quality for the reference data used (mostly located in Europe and North America), while 171 did not report on quality (most African and Asian countries). In addition, forest inventory data are not always comparable and biomass estimates may be biased due to differing national forest definitions and differences in methods used to produce the estimates, such as the choice of the minimum tree diameter sampled (Searle and Chen, 2017) and plot size (Réjou-Méchain et al., 2014). The only practical approach for consistent global or regional woody biomass estimation therefore lies in systematic use of Earth Observation (EO) data, either in parameterised model-based approaches or in combination with high-quality reference data. Satellite data have long been used for forest cover mapping, clear-cut or burnt area monitoring and detection of disturbances (Hansen et al., 2013; Healey et al., 2005; Fraser and Li, 2002; Rignot et al., 1997). However, without biomass information this is insufficient to quantify the role of forests in the global carbon and energy cycles and other biogeochemical cycles. In addition, financial mechanisms aiming to reduce emissions and enhance carbon stocks, such as the Reducing Emissions from Deforestation and Forest Degradation (REDD+) initiative and carbon trading schemes, require credible and consistent measurement, reporting and verification (MRV) systems that are spatially explicit with a wall-to-wall extension and provide a full carbon account of forest ecosystems (Steffen et al., 1998).

### 1.2. Current status of biomass estimation from space

Studies aiming at wall-to-wall estimation of AGB at regional and global scale have used passive optical, active or passive microwave, and LiDAR data obtained from Earth Observation space platforms either stand-alone or in synergy (e.g. Saatchi et al., 2011; Baccini et al., 2012; Thurner et al., 2014; Gallaun et al., 2010; Hu et al., 2016, Liu et al., 2015). Multispectral optical imagery contains information on the photosynthetic parts of vegetation rich in chlorophyll, while microwave active sensors, such as Synthetic Aperture Radar (SAR), contain information on the dielectric (essentially moisture content) and structural properties of objects, soil surface and plants. The main advantage of microwave radar sensors is that, unlike optical imagery, radar images are unaffected by cloud cover, allowing usable image acquisitions even in the cloudiest places on Earth. Spaceborne LiDAR sensors, on the other hand, give a sampled retrieval pattern along the orbit and to measure elapsed time between emitted and received light pulses which can be used to estimate forest canopy height at each footprint location. However, these datasets present different degrees of saturation to AGB, where saturation refers to the AGB level at which the sensitivity of the signal (i.e. backscatter, reflectance) becomes too small to be measurable, or where the signal fails to penetrate the canopy (Fagan and DeFries, 2009). This is particularly relevant for dense tropical forest, which is a key biome where accurate biomass information is needed.

The search for consistent approaches over forested areas in the tropics prompted the use of satellite data calibrated against in situ biomass, with special emphasis on forest height estimates derived from the first spaceborne LiDAR, the Geoscience Laser Altimeter System (GLAS) on board the Ice, Cloud and land Elevation Satellite (ICESat) (Lefsky, 2010). This led to the development of two pan-tropical biomass maps (Saatchi et al., 2011; Baccini et al., 2012) at reporting grid size of 1 km and 500 m respectively. The former reported a relative RMSE at pixel level of approximately 30%, while the latter reported similar figures in terms of RMSE (38–50 t ha$^{-1}$). These maps exhibited significant regional differences, although these decreased when biomass estimates were aggregated to country or biome scale (Mitchard et al., 2013, 2014; Rodriguez-Veiga et al., 2016).

Avitabile et al. (2016) fused Saatchi et al. (2011), and Baccini et al. (2012) datasets into a 1 km pan-tropical map using a bias-removal approach by incorporating additional field observations and locally-calibrated high-resolution biomass maps. The bias in the overall mean AGB was reduced to + 5 t ha$^{-1}$, compared with the biases in the input maps of + 21 t ha$^{-1}$ and + 28 t ha$^{-1}$ respectively.

Using very long time series of C-band radar data from Envisat ASAR, Santoro et al. (2015a) produced a Growing Stock Volume (GSV) map for the northern hemisphere at 1 km spatial resolution. The relative RMSE of the retrievals at provincial level was between 12% and 45% (average 29%) when compared to National Forest Inventory data from the major forested countries. This map provided the basis for a carbon stock map of the boreal and temperate forests (Thurner et al., 2014).

A first composite global dataset of forest AGB was developed within the European Commission-funded GEO-CARBON project. The product merged, at a pixel size of 0.01°, the Saatchi et al. (2011), and Baccini et al. (2012) pan-tropical datasets with the boreal and temperate dataset (Santoro et al., 2015a; Thurner et al., 2014) and used the IPCC Tier 1 biomass values for the few remaining areas not covered by these datasets (Avitabile et al., 2014, 2016). This exercise, despite being hindered by limitations in the input EO data used by individual biomass maps, approximations in the retrieval approaches and the fact that the individual maps were based on data acquired at different times between

2000 and 2010, is still the most consistent global AGB map to date. Hu et al. (2016) also published a global AGB map at 1 km resolution derived using GLAS metrics interpolations, MODIS NDVI and Land Cover products and the SRTM DEM, together with climate data. However, the dataset used to calibrate the map consisted of 3348 forest inventory plots of different sizes (including very small plots of 0.05 ha). The calibration dataset was also geographically biased as the plots were mostly located in continental China (> 55% of plots) and Brazil (23% of plots), while almost no plots were used from Europe, North America, Australia, and Africa. These issues might explain the large differences observed in this map when compared to previous global and pan-tropical maps (Rodríguez-Veiga et al., 2017; Hu et al., 2016).

Liu et al. (2015) used vegetation optical depth (VOD) retrieved from several passive microwave satellite sensors to map time series of AGB for all vegetation types globally over the period 1993–2012 at 27.5 km resolution. Unfortunately, this approach was calibrated using the Saatchi et al. (2011) map, which added the uncertainties from this product to the final map, and make it difficult to validate due to the coarse resolution of the product.

At continental scale, MODIS data and forest inventory plots have been used to map AGB over Europe (Gallaun et al., 2010) at 500 m resolution, and Africa (Baccini et al., 2008) at 1 km resolution. The woodlands and savannas of Africa were also mapped at 25 m spatial resolution using ALOS PALSAR data (Bouvet et al., 2018).

At national and regional scales, several examples have been published, such as for Mexico (Rodriguez-Veiga et al., 2016; Cartus et al., 2014), Canada (Beaudoin et al., 2014), Cameroon (Mermoz et al., 2014), China (Yin et al., 2015; Piao et al., 2005; Liu et al., 2012), the Amazon basin (Saatchi et al., 2007), Russia (Houghton et al., 2007), USA (Kellndorfer et al., 2011) and Colombia (Anaya et al., 2009), with spatial resolutions ranging from 30 m to 1 km and in most cases using a combination of optical and SAR imagery. Regional approaches use field AGB measurements to calibrate the algorithms, often complemented with airborne LiDAR datasets (Asner et al., 2014, 2013; Perrin et al., 2016; Lu et al., 2012). These regionally-calibrated products can use a wider variety of datasets, as well as regional expertise, to provide the best possible estimates of biomass. In contrast, global, pantropical or continental products suffer from limitations in the amount and representativeness of data available for calibration and validation. Pan-tropical maps from Saatchi et al. (2011), and Baccini et al. (2012) circumvented the paucity of ground data for calibrating their non-parametric machine learning approaches at large scale by using AGB estimated from LiDAR footprints from the space-borne ICESAT-GLAS instrument. However, they still used a fair amount of ground-based values of AGB to calibrate the relationship between AGB and LiDAR footprint metrics. An algorithm that avoids the use of in situ data for model training is the BIOMASAR algorithm (Santoro et al., 2015a, 2011; Cartus et al., 2012); the algorithm is, however, constrained with information on maximum biomass which are derived from inventory data, regional and national statistics, as well as remote sensing-based biomass estimates (Santoro et al., 2015a, 2011; Cartus et al., 2012). Inaccurate data sources ultimately translate into local estimation biases (Santoro et al., 2011).

The long list of AGB datasets presented above highlights that biomass mapping methods are largely driven by data availability and are scale-dependent. National and regional products can be generated by different parametric and non-parametric approaches. Non-parametric methods, such as machine learning techniques, usually out-perform parametric approaches (Evans and Cushman, 2009) and are preferred at national and regional level if enough ground data are available. At global or continental level the lack of representative in situ measurements is the motivation for using physically-based approaches that require few ground data (if any).

This paper describes a diverse set of regional approaches to AGB mapping in different biomes carried out during the European Space Agency (ESA) Data User Element GlobBiomass project (GlobBiomass,

2015; Schmullius et al., 2015; Balzter et al., 2016; Schmullius, 2017). This study aimed to produce spatially consistent and accurate maps of AGB, using all available EO data and regional knowledge with the objective of supporting the development of global biomass retrieval algorithms and the assessment of thereof resulting estimates. These maps can be used for direct estimation of carbon emission factors or emissions contributing to greenhouse gas inventories. Further aims were: i) to better understand the strengths and weaknesses of existing methods to map AGB using available EO datasets, ii) to establish how differences in forest structure and reference data affect methods to invert EO data to AGB. Five regional AGB maps derived using reference data and EO imagery and various retrieval methods were generated for the year 2010 at 25–100 m spatial resolution. The regions were selected to encompass a wide range of biomes and forest types. Each region was at least 300,000 km² in size; it was either nationwide for Poland (temperate forest) and Sweden (boreal forest), or covered a substantial part of Indonesia (Kalimantan, tropical forest), Mexico (the Yucatan peninsula & Central Mexico, tropical forest-woodland transition), and South Africa (Eastern forest belt, subtropical dry forest). All maps were evaluated quantitatively against an independent dataset, and qualitatively by local experts. They were also compared with existing continental scale AGB maps (Saatchi et al., 2011; Baccini et al., 2012; Thurner et al., 2014) where these overlapped the study areas.

## 2. Study regions

The study regions cover the most common range of woody AGB from low (< 50 t ha⁻¹) to high (> 300 t ha⁻¹) and are representative of major climates and forested biomes, including boreal, temperate, dry tropical savanna and wet tropics (Fig. 1).

Sweden is mostly situated in the boreal region, while Poland lies in the temperate forest zone. Sweden and Poland occupy approximately 447,000 km² and 313,000 km², of which 60% and 30% are forests, respectively. Coniferous forests predominate, though broadleaved forests occupy a significant area in Poland.

The study areas in Central Mexico, the Yucatan peninsula and Kalimantan represent a wide variety of tropical and subtropical forest
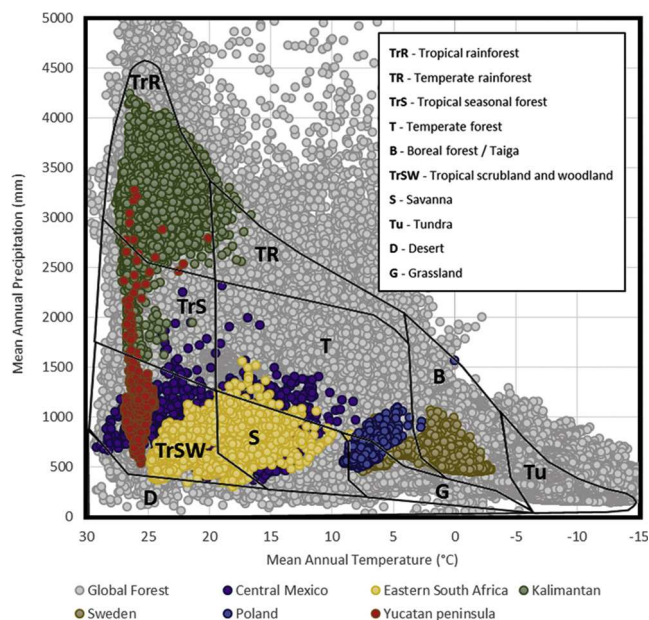


**Fig. 1.** Mean annual precipitation vs. mean annual temperature (Fick and Hijmans (2017) for global forests and study sites sampled at 0.5 ˚ grid scale. The climate space is divided into global terrestrial biomes (Whittaker, 1962, 1970). The global distribution of forests is according to the Global Land Cover (GLC2000) map (Bartholomé and Belward, 2005).

ecosystems. The Yucatan peninsula study area with approximately 160,000 km$^2$ comprises a mix of tropical moist, tropical dry forests and mangroves, whilst an area of 353,000 km$^2$ in Central Mexico covers subtropical coniferous forest, tropical dry forest, tropical moist forest, xeric shrublands, and includes some of the forests with the highest biomass density in Mexico (the Oyamel forest). The Indonesian part of Borneo (Kalimantan) covers 73% of Borneo's land mass (approximately 540,000 km2). The ecosystems of Kalimantan include different forest types: mangrove forests, peat swamp and freshwater swamp forests, the most extensive extent of heath forests in Southeast Asia, lowland dipterocarp forests, ironwood forests, forests on limestone and ultrabasic soils, hill dipterocarp forests and various montane formations (MacKinnon et al., 1996).

The South African study area of approximately 334,000 km$^2$ is situated along a 1300 km North-South transect running to the East of the country next to Zimbabwe, Mozambique, Swaziland, and the Indian Ocean, and is dominated by forested landscape. This area contains various forest types: savanna (68% of the area), commercial plantations, and scattered remnants of indigenous dense forests (Mucina and Rutherford, 2006).

## 3. Data

Remote sensing imagery from different airborne and satellite sensors (optical, LiDAR and SAR) were utilized in this study (Table 1). Except for Sweden, the main dataset used was the freely available ALOS PALSAR 2009 and 2010 mosaics of gamma nought ($\gamma° = \sigma°/cos\theta$, where $\sigma°$ is the radar backscattering coefficient and $cos\theta$ is the local incidence angle) produced by JAXA at 25 m pixel spacing in HH and HV polarizations (Shimada et al., 2014). The ALOS PALSAR mosaics are processed according to a standard protocol (Shimada et al., 2014) which involves calibration, multi-looking (16 looks), projection, ortho-rectification and slope correction using the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) data. A destriping process (Shimada and Isoguchi, 2002) was also applied to try and equalise intensity differences between neighbouring strips normally attributed to seasonal and daily differences in soil moisture conditions. As part of our methods, if significant strip effects still remained locally, substitution by another year's mosaic or histogram matching of the problematic strip with neighbouring strips was performed. A multi-temporal multichannel filter (Quegan and Yu, 2001) using a $7 \times 7$ window was also applied to all the annual mosaics. At this point, the remaining speckle effect was considered negligible.

Landsat 5 and 7 ETM + Surface Reflectance (SR) imagery computed by the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) method (http://ledaps.nascom.nasa.gov/) (Masek et al., 2006) were used to generate annual ( ± 1 yr) median value composites from good quality pixels for all spectral bands in Mexico and Poland. Landsat Percent Tree Cover (PTC) products (Hansen et al., 2013; Sexton et al., 2013) for the year 2010 were acquired from USGS Land Cover Institute (https://landcover.usgs.gov/), and the Global Land Cover Facility (http://glcf.umd.edu/) for Mexico and Eastern South Africa, respectively. Additionally, freely available 30 m spatial resolution elevation data from the Shuttle Radar Topography Mission (void-filled SRTM Plus NASA V3) was obtained for Mexico, Eastern South Africa and Kalimantan from the USGS Earth Explorer repository (http://earthexplorer.usgs.gov/). SPOT-4 High Resolution Visible Infrared (HRVIR) and SPOT-5 High Resolution Geometric (HRG) data were acquired between 2008 and 2010 (approximately 80% from 2010) for Sweden; all images were geometrically precision-corrected to the Swedish National Grid, and the pixel size for all bands was resampled to 25 m using cubic convolution.

The accuracies of the resulting maps were evaluated using either AGB forest inventory plots or airborne LiDAR-derived AGB maps, collected and produced according to different protocols (Table 1), and with characteristics specific to each region. Airborne LiDAR-derived

**Table 1**
Datasets and algorithms used for each 2010 regional map.

| Regional Map | Method | Method type | | Datasets | Plot size / LiDAR point density | ALOS PALSAR | Landsat 5/7 | Landsat PTC | SPOT 4 & 5 | SRTM DEM |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Reference data | | | | | | |
| Indonesia Kalimantan | Two step LiDAR + SAR multiple linear regression | P | DD | 247 field plots & 1245 LiDAR samples calibration / 501 LiDAR samples validation | 0.08 ha - 0.38 ha / 2 - 5 points m$^{-2}$ | x | | | | x |
| Eastern South Africa | WCM + Bayesian inversion | P | MB | 37 field plots calibration / 3188 LiDAR samples validation | 1 ha / 0.5 - 4 points m$^{-2}$ | x | | x | | x |
| Sweden | kNN | NP | DD | 18,401 field plots calibration / 3989 field plots validation | 0.015 ha - 0.030 ha | x | | | x | |
| Central Mexico & Yucatan peninsula | MaxEnt | NP | DD | 5140 field plots calibration / 700 field plots validation | 1 ha (4 x 0.04 ha sampled) | x | x | x | | x |
| Poland | Random Forest | NP | DD | 285 field plots calibration / 84 field plots validation | 0.05 ha | x | x | | | x |

P (parametric), NP (non-parametric), DD (data-driven), MD (model-based).

AGB maps were used in Kalimantan for calibration and validation, but only for validation in Eastern South Africa (Naidoo et al., 2015). A subset of LiDAR-derived AGB maps or the field data (15–30%) was excluded in each region and used as an independent validation dataset. The subset was extracted by stratifying the reference dataset into different AGB classes in order to have a similar distribution of AGB in both the calibration and validation datasets.

# 4. Methods

## 4.1. Biomass estimation methods

Both parametric and non-parametric methods were used to predict AGB. These can be further grouped into data-driven and model-based methods. The method for each region was selected based on data availability and the expertise of each regional research group. Teams working in areas with forest inventory and other in situ data of sufficient number and quality for calibration purposes used non-parametric machine learning algorithms, while areas with insufficient in situ data used parametric models, such as model-based regression, and when available, complemented the ground observations with airborne LiDAR biomass predictions (Table 1).

The probabilistic outputs from the non-parametric MaxEnt algorithm (Phillips et al., 2006, 2004) were used for Mexico (Rodriguez-Veiga et al., 2016). Machine learning algorithms Random Forest (Breiman, 2001; Cartus et al., 2014) and a k-Nearest Neighbours (kNN) (Tomppo et al., 2008; Reese et al., 2003) were applied in Poland and Sweden respectively, both requiring large amounts of field plots for calibration. The parametric method used in Kalimantan used a two-step calibration approach where field plots are first used to calibrate airborne LiDAR measurements covering a larger area, and these were then used to calibrate a multivariate linear regression model with backscatter intensity and texture parameters from the SAR imagery as predictors (Englhart et al., 2011, 2012). Bayesian inversion of a semi-empirical model (the water cloud model - WCM) was used to relate PALSAR backscatter to AGB in South Africa (Bouvet et al., 2018). This method relies on a small number of at least 1 ha in situ AGB plots, ancillary data and simulations from the Multi-static Interferometric and Polarimetric Electromagnetic model for Remote Sensing (MIPERS) (Villard and Borderies, 2007; Villard, 2009) for parameterization.

## 4.2. Accuracy assessment methods

A standardised accuracy assessment was carried out for all regional AGB maps by making use of the independent reference data. The assessment was based on stratifying the reference AGB into contiguous ranges of values and quantifying the estimation bias, the standard deviation of the error and the Root Mean Square Error (RMSE) within each range. The selected ranges varied with test site, depending on the maximum value of biomass for the site and the need to have a sufficient number of reference data within each range.

In more detail, we have a set of reference AGB values, $B_{ref}^{(i)}$ (from in situ or LiDAR data), and their estimates, $B_{est}^{(i)}$, where the reference values are restricted to a given range, $B_1 < B_{ref}^{(i)} < B_2$. For this range we define the bias, $b$, as the average value of the error $B_{est}^{(i)} - B_{ref}^{(i)}$, and we also calculate the standard deviation (SD) of the errors, $\sigma$. The RMSE in the given range is then given by $\sqrt{\sigma^2 + b^2}$, and the relative RMSE as $Rel.\ RMSE = RMSE / B_{ref} \times 100$. Also of interest is the Coefficient of Variation (CV) of the error, given by $\sigma / b$. When the CV exceeds 1, the RMSE is dominated by random error, but when it is less than 1 the dominant error source is bias in the estimator. In particular, if CV = 10, then bias makes up 10% of the RMSE, while if CV = 0.48 it contributes 90%.

# 5. Results

The constructed AGB maps for the year 2010 were generated with a pixel size of 25 m for Mexico-Yucatan Peninsula, Central Mexico,
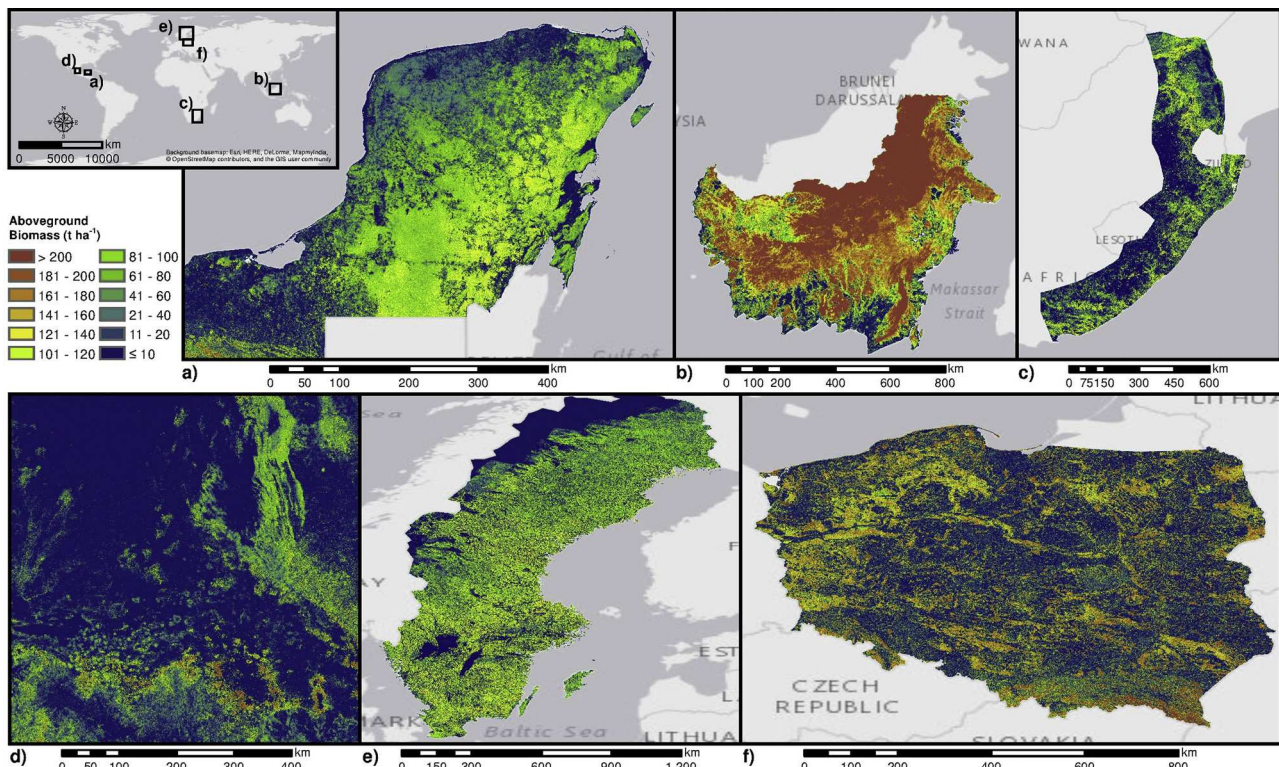


**Fig. 2.** Above-ground biomass (AGB) maps for a) the Mexico-Yucatan Peninsula, b) Indonesia-Kalimantan, c) Eastern South Africa, d) Central Mexico, e) Sweden, and f) Poland. Warmer colours indicate higher AGB.
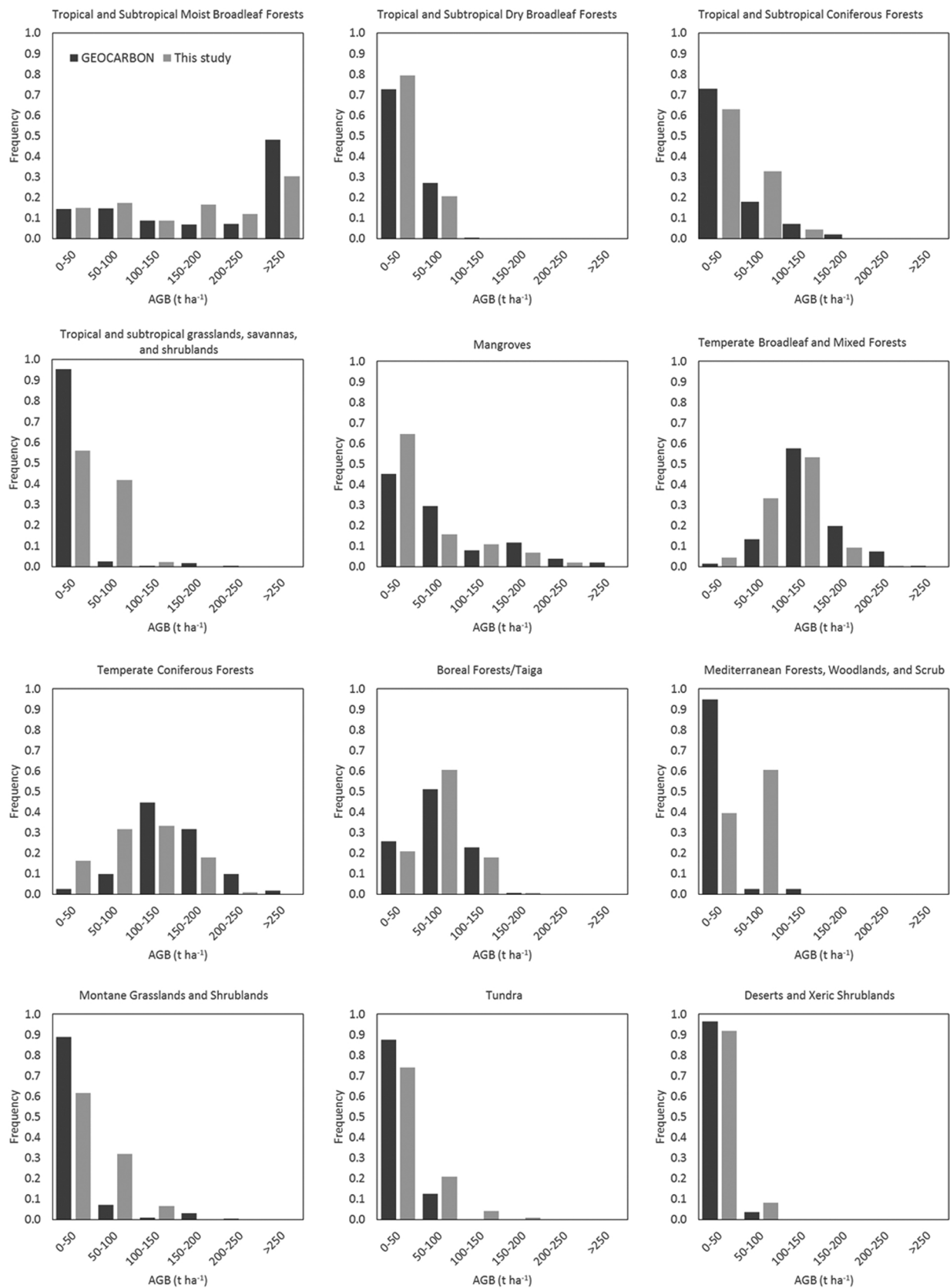
**Fig. 3.** AGB histograms per forest biome (Olson et al., 2001) derived from the combined regional maps in this study (light grey) aggregated to 1 km spatial resolution compared to the AGB histogram from the GEO-CARBON global map (dark grey) within the study areas. Flooded grasslands and savannas biome is not included in this analysis due to the small amount of data available from the study regions, while the temperate grasslands, savannas, and shrublands biome was not encountered in the study regions.

Eastern South Africa, Sweden, and Poland, and 100 m for Indonesia-Kalimantan (Fig. 2).

The regional maps cover the whole range of expected woody AGB densities from low to high biomass. The histogram of the combined regional maps was comparable to the histogram of AGB extracted from the global AGB map of the GEO-CARBON project in the different forest biomes (Olson et al., 2001) covered by this study (Fig. 3). The most substantial differences are on the tropical and subtropical grasslands, savannas and shrublands, and in the montane grasslands and shrublands biomes, where the GEO-CARBON map showed a strongly skewed histogram towards low AGB ($< 50$ t ha$^{-1}$) and very low frequency of higher AGB classes, while this study showed a more distributed declining trend from low to high AGB classes in those biomes. Additionally, the GEO-CARBON map shows the same skewed histogram towards low AGB in the Mediterranean forests, woodlands, and scrub biome, while this study's histogram showed to be slightly skewed towards higher AGB (50-100 t ha$^{-1}$). In the tropical and subtropical moist broadleaf forests biome histograms are similarly skewed towards high AGB, but GEO-CARBON's histogram showed higher frequency at the highest AGB range ($> 250$ t ha$^{-1}$).

The accuracies of the regional AGB maps were assessed using the independent validation datasets (Table 2, Figs. 4 and Fig. 5). LiDAR AGB datasets were aggregated to the corresponding satellite map resolution for validation, while plot datasets used the average value of the pixels within the plot boundaries.

The accuracy analysis reveals several commonalities but also some important differences between study regions. All regions over-estimate AGB for the lower AGB ranges and, with the exception of Eastern S. Africa which covers a very limited biomass range, under-estimate in the upper ranges, especially in the highest AGB class. The bias in the lowest AGB class is substantial in absolute terms for all regions except Central Mexico and Eastern S. Africa, but even in these regions it has values that are 47% and 40% of the mid-range values (i.e., 15 and 10 t ha$^{-1}$), respectively. For most regions the bias decreases in absolute value before increasing again. This is expected for regression-based approaches, which ensure that the regression curve passes through the point defined by mean of the reference and estimated data, but occurs for all methods.

There are striking differences in the balance between bias and random error in the RMSE, as is clear from Table 2 and Fig. 4. In Kalimantan, bias and random error are of similar magnitude except in the middle AGB ranges, where random error dominates. For Central Mexico and Eastern S. Africa, random error is dominant except for the highest AGB class in Central Mexico, where it is comparable to bias. Note that in these two regions the bias is roughly constant across all ranges (except for the highest range in Central Mexico) so it decreases sharply relative to the mid-range values. In the Yucatan Peninsula and Sweden, bias and random error are comparable in the lower biomass ranges, the middle ranges are dominated by random error, while bias is the largest component of error in the highest AGB ranges. For Poland, bias is by the

**Table 2**

Accuracy assessment of the regional AGB maps stratified by reference AGB range: Sample size (N), Root Mean Square Error (RMSE), Relative RMSE (Rel. RMSE), Standard Deviation of the error (SD), Bias, and Coefficient of Variation (CV) of the error (when CV > 1 the random error dominates, when CV < 1 the bias does).

| Study region | Reference AGB range (t ha$^{-1}$) | N | RMSE (t ha$^{-1}$) | Rel. RMSE (%) | Bias (t ha$^{-1}$) | SD (t ha$^{-1}$) | CV |
|---|---|---|---|---|---|---|---|
| **Indonesia Kalimantan** | 0-50 | 141 | 54 | 415 | 37 | 39 | 1.1 |
| | 50-100 | 21 | 71 | 91 | 52 | 50 | 1.0 |
| | 100-150 | 38 | 66 | 50 | 39 | 54 | 1.4 |
| | 150-200 | 117 | 42 | 24 | 16 | 39 | 2.4 |
| | > 200 | 184 | 60 | 24 | −40 | 44 | 1.1 |
| | **Overall** | **501** | **55** | **37** | **4** | **55** | **13.8** |
| **Central Mexico** | 0-30 | 83 | 21 | 159 | 7 | 19 | 2.7 |
| | 30-60 | 48 | 30 | 69 | 8 | 29 | 3.6 |
| | 60-90 | 25 | 34 | 47 | −3 | 34 | 11.3 |
| | 90-120 | 12 | 42 | 40 | 2 | 42 | 21.0 |
| | 120-150 | 6 | 24 | 19 | 4 | 24 | 6.0 |
| | > 150 | 2 | 64 | 32 | −41 | 49 | 1.2 |
| | **Overall** | **176** | **28** | **67** | **5** | **28** | **5.6** |
| **Mexico Yucatan peninsula** | 0-30 | 130 | 33 | 237 | 21 | 26 | 1.2 |
| | 30-60 | 109 | 33 | 75 | 15 | 30 | 2.0 |
| | 60-90 | 111 | 23 | 31 | 6 | 22 | 3.7 |
| | 90-120 | 85 | 25 | 24 | −14 | 21 | 1.5 |
| | 120-150 | 54 | 46 | 34 | −42 | 17 | 0.4 |
| | > 150 | 35 | 67 | 40 | −65 | 16 | 0.2 |
| | **Overall** | **524** | **35** | **50** | **−1** | **35** | **35.0** |
| **Eastern South Africa** | 0-20 | 2,216 | 8 | 100 | 4 | 8 | 2.0 |
| | 20-40 | 734 | 14 | 48 | 7 | 11 | 1.6 |
| | 40-60 | 233 | 15 | 33 | 7 | 11 | 1.6 |
| | > 80 | 5 | 7 | 11 | 3 | 5 | 1.7 |
| | **Overall** | 3188 | **10** | **63** | **5** | **9** | **1.8** |
| **Sweden** | 0-30 | 901 | 38 | 271 | 30 | 23 | 0.8 |
| | 30-60 | 871 | 35 | 76 | 24 | 25 | 1.0 |
| | 60-90 | 850 | 29 | 39 | 12 | 26 | 2.2 |
| | 90-120 | 606 | 29 | 28 | −6 | 29 | 4.8 |
| | 120-150 | 361 | 43 | 32 | −29 | 32 | 1.1 |
| | 150-180 | 245 | 61 | 37 | −50 | 35 | 0.7 |
| | > 180 | 155 | 82 | 42 | −74 | 36 | 0.5 |
| | **Overall** | 3989 | **32** | **39** | **−13** | **29** | **2.2** |
| **Poland** | 0-50 | 13 | 67 | 258 | 63 | 23 | 0.4 |
| | 50-100 | 19 | 54 | 67 | 49 | 24 | 0.5 |
| | 100-150 | 17 | 26 | 20 | 8 | 25 | 3.1 |
| | 150-200 | 16 | 32 | 18 | −21 | 25 | 1.2 |
| | 200-250 | 11 | 73 | 30 | −70 | 20 | 0.3 |
| | > 250 | 8 | 86 | 32 | −85 | 16 | 0.2 |
| | **Overall** | **84** | **54** | **39** | **3** | **54** | **18.0** |
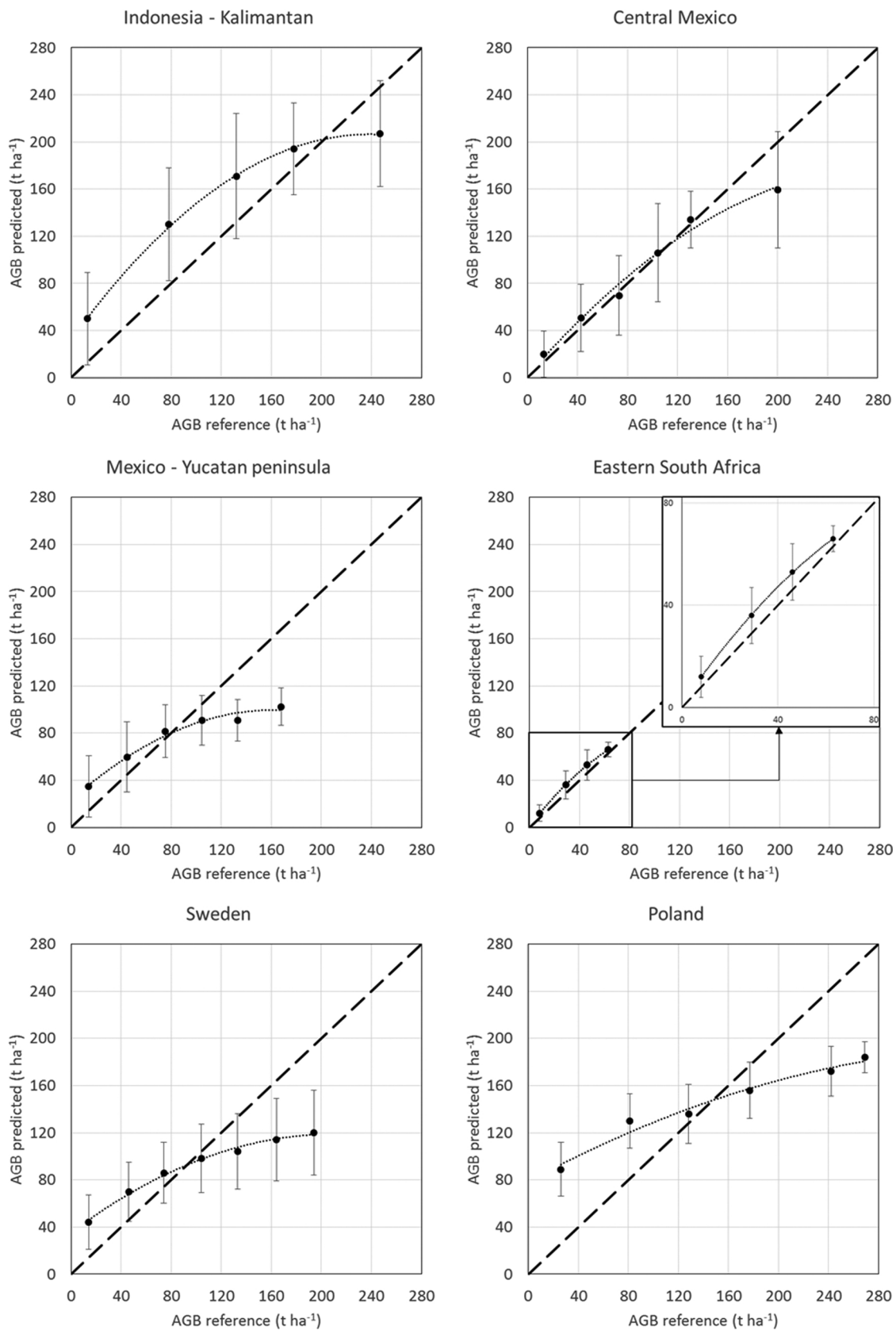
**Fig. 4.** Scatterplots of average predicted AGB versus average reference AGB per reference AGB range. Error bars indicate the standard deviation (random error) of the predicted AGB per reference AGB range. The dotted line indicates a fitting curve to the calculated points (second order polynomial) and the dashed line corresponds to the y = x line. If the error bars do not overlap the y = x line then bias is the dominant error in that AGB range.
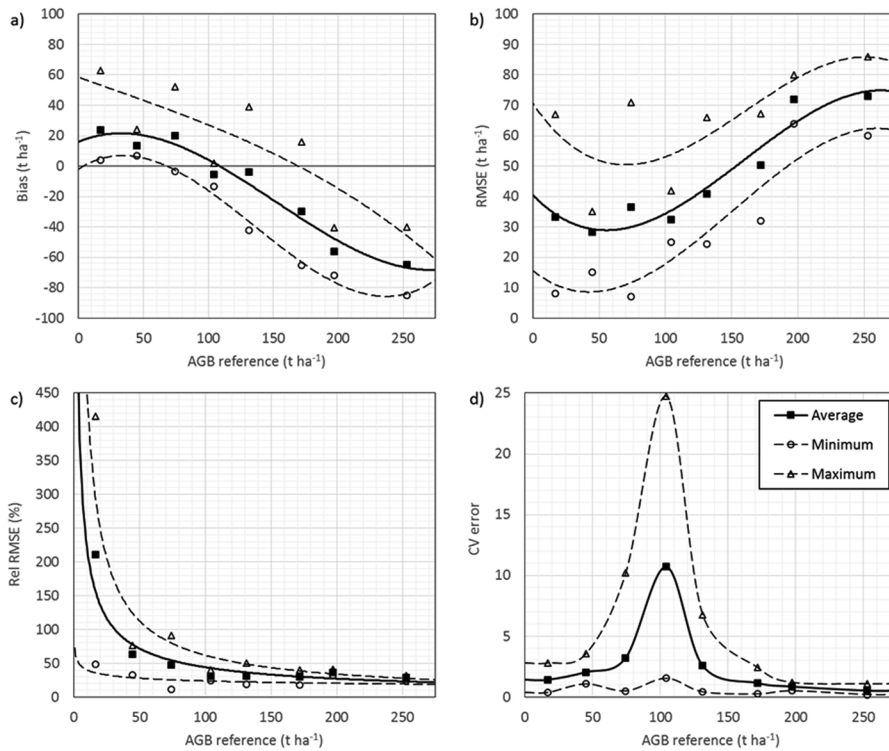
**Fig. 5.** Distribution of a) bias, b) relative root mean square error, c) relative RMSE, and d) coefficient of variation of the error across the reference AGB range for all regional maps.
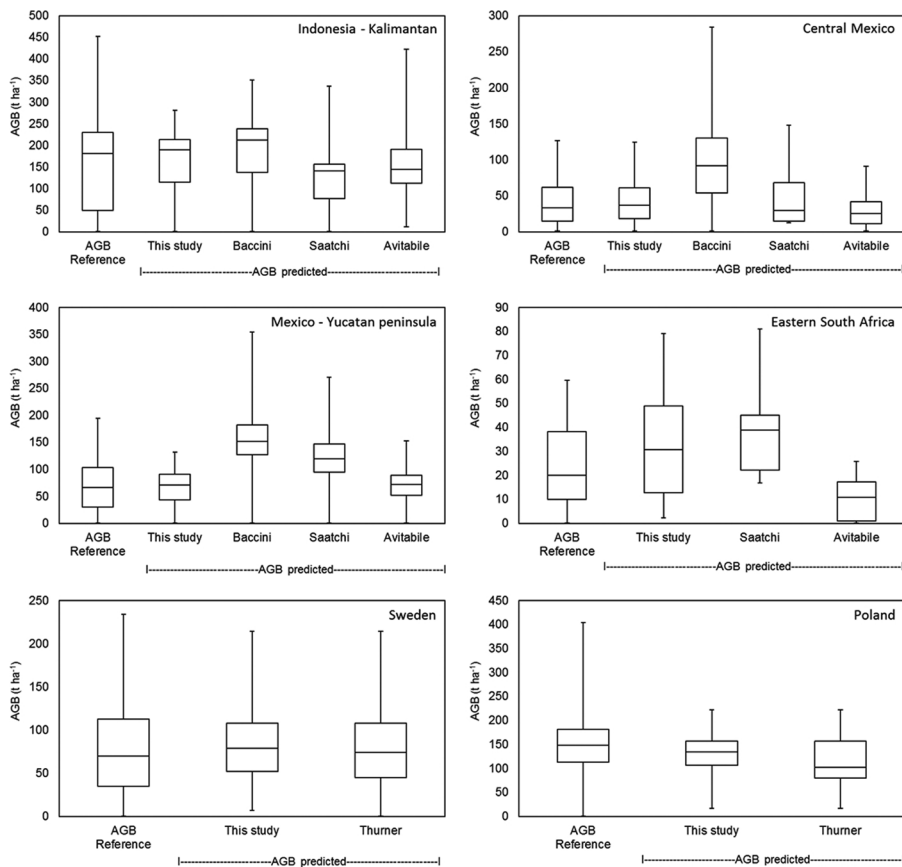


**Fig. 6.** Boxplots by region of reference and predicted AGB in this study, Baccini et al. (2012); Saatchi et al. (2011); Avitabile et al. (2016) and Thurner et al. (2014). The central mark of each box represents the median of the distribution, while the lower and upper limits of the box represent the 25th and 75th percentiles. The whiskers cover the range of extreme values. Outliers are not included. All the maps were aggregated to 1 km spatial resolution.

far the dominant error source apart from the two middle ranges of AGB.

It is also noticeable that the SD of the error does not vary greatly across the different AGB ranges for each region, though is markedly different between regions (see Table 2 and Fig. 4). Hence, roughly speaking, the random error is not strongly dependent on the true AGB, and its value relative to the true AGB decreases as AGB increases (Fig. 5). It can also be seen that the contribution of the random error to the RMSE increases as AGB increases up to approximately 100 t ha$^{-1}$, then reduces sharply with increasing AGB (Fig. 5).

The values given for the overall bias in each region are close to zero (Table 2). This implies that all the methods cause the fitting curve (or its equivalent) to go through the point defined by the averages of the reference and estimated data, in common with normal regression methods which force this to happen. This explains why the overall RMSE and SD are nearly the same, since SD$^2$ = (RMSE)$^2$ – (overall bias)$^2$ and the overall bias is constrained to be nearly zero (Table 2).

Previously published pan-tropical (Saatchi et al., 2011; Baccini et al., 2012; Avitabile et al., 2016) and northern hemisphere (Thurner et al., 2014) mapping studies show AGB distributions and spatial patterns that are different from those for the regional maps from this study and the independent validation data (Figs. 6 and 7). In particular, in Mexico the distributions from Baccini et al. (2012) are shifted towards much higher values than is found in the present study and the validation data.

Similar estimates can be found in areas with high AGB levels such as Kalimantan, but they deviate from the reference AGB in areas of lower AGB such as Eastern South Africa, Yucatan and Central Mexico. In Central Mexico, the Saatchi et al. (2011) AGB distribution is similar to those in our study and the validation data, but for the Yucatan peninsula and Eastern South Africa it is shifted towards higher values while for Kalimantan is shifted towards lower values. In Avitabile et al. (2016) and this study, the AGB distribution is similar to that obtained from the validation data in Yucatan and Central Mexico but in Eastern South Africa the Avitabile et al. (2016) data are shifted toward lower values while this study is shifted towards higher values. In Kalimantan, the distributions of the validation data and all the maps, except for Avitabile et al. (2016), are highly skewed towards high values. Only the Avitabile et al. (2016) AGB map provides estimates similar to the true

AGB in the Mexican sites. In Sweden, the AGB distributions estimated in this study and in Thurner et al. (2014) largely agree with the validation data. In Poland, the AGB distribution from Thurner et al. (2014) looks highly skewed and shifted towards lower values than the estimates from this study and in the validation subset.

## 6. Discussion

### 6.1. Evaluation of the maps

The analysis of the regional maps affirms important properties of AGB estimation methods, some of which have been previously reported in the literature (e.g. Cartus et al., 2014; Carreiras et al., 2012; Englhart et al., 2012; Baccini et al., 2008; Sandberg et al., 2011; Rauste, 2005; Avitabile et al., 2016; Avitabile and Camia, 2018). The accuracy assessment shows underestimation in the upper AGB ranges in which the major error component of the RMSE originates from bias (Table 2). The exception is in Eastern South Africa, where the reference values of AGB are below 80 t ha$^{-1}$. Additionally, although less apparent in Eastern South Africa and Central Mexico, AGB values below 100 t ha$^{-1}$ are overestimated, regardless of the choices of data and method used to retrieve AGB (Table 2 & Fig. 4). Fig. 4 is derived from an independent dataset not used in model fitting, but similar behaviour is seen for data used in model fitting. This means that, although direct linear regression is not being used, all model fits used to predict AGB have an intercept at zero AGB that is too high, and/or the gradient of the model is too large for lower AGB values.

A major problem with the observed biases is that they depend on the true value of the AGB. If not, the data could easily be calibrated to remove them. Even though the analysis quantifies how the bias depends on the true AGB in each region, this does not lead to any way to correct the estimated values. Although some of the methods incorporate bias reduction measures, e.g. MaxEnt (Xu et al., 2016; Saatchi et al., 2011), and post-processing bias reduction techniques are also available, there is a risk of undesirable effects such as inflation of the overall mean square error due to an increase of the variance (Kosmidis, 2014; Xu et al., 2016). Addressing this problem requires new algorithms that intrinsically remove the bias (if this is possible), new data that do not
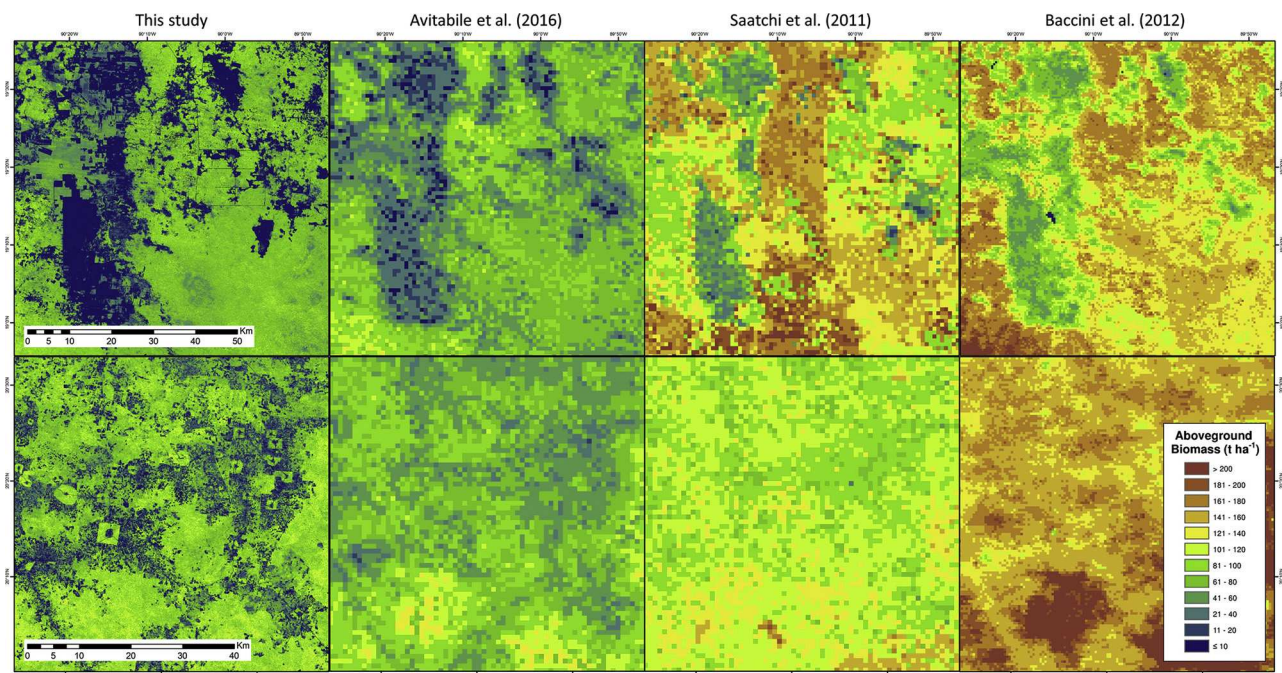


**Fig. 7.** Comparison of four AGB maps over two sites in the Yucatan peninsula (Mexico), showing the regional AGB map presented in this study, and the maps by Avitabile et al. (2016); Saatchi et al. (2011), and Baccini et al. (2012).

suffer saturation of the signal for higher biomass (assuming this is the primary cause of the observed underestimation in this range of biomass), or accepting that such bias will occur because AGB is only indirectly related to the remote sensing observables considered in these studies.

However, as has been shown, for some regions and methods bias is not the dominant effect (i.e. Kalimantan and Central Mexico); instead, the most important type of error is random scatter of the data points in the model inversions, i.e., the model inversions are noisy. Some of this scatter may be reducible if its source can be identified, e.g., there may be errors due to inaccuracies in the ground data, geolocation errors (so that the reference data and inversions are spatially mismatched), radar speckle may have been insufficiently reduced (though our methodology ensures this should be of negligible importance in our analysis), or the remote sensing signal may be weakly correlated to AGB due to the limited sensitivity to biomass of the sensor. Moreover, if the dominant error term comes from scatter, it can be reduced by spatial averaging (whereas bias cannot) at the expense of reduced spatial resolution and aggregation to coarser spatial units.

Pan-tropical (Saatchi et al., 2011; Baccini et al., 2012; Avitabile et al., 2016) and northern hemisphere maps (Thurner et al., 2014) were compared to in situ data and the estimates from this study. Some similarities were found between our study and Avitabile et al. (2016) in the Mexico sites and Thurner et al. (2014) in Sweden. Even though Avitabile's map is a fusion of Saatchi's and Baccini's data, the fusion method used additional ground measurements and higher resolution regional maps to correct for the bias of the original estimates. This is the case in Mexico, where the Avitabile et al. (2016) product was calibrated with another regional AGB product (i.e. Cartus et al., 2014). However, in Eastern South Africa, Avitabile's map does not have calibration data and systematically underestimates AGB with maximum predicted AGB values just above $20\,t\,ha^{-1}$, close to a factor of three less than the higher AGB values reported in the validation dataset ($60\,t\,ha^{-1}$). In general, the distributions of the AGB estimates from this study are closer to those of the independent validation data than those of the published AGB global maps.

### 6.2. Strengths and weaknesses of proposed retrieval methods and available datasets

Several methods were used to predict AGB and showed specific strengths and weaknesses (The requirement for only a limited number of ground data points is an advantage of semi-empirical methods, such as the Water Cloud Model (WCM) + Bayesian inversion used in Eastern South Africa where only one parameter of the regression needs to be derived from in situ biomass data. However, the formulation of the WCM for this region does not produce estimates for AGB above ˜100 t

ha-1, and it is tuned for regions such as savannas with biomass below the saturation level at L-band (Mermoz et al., 2015). A method for global mapping, BIOMASAR (Santoro et al., 2011), is also based on the WCM and does not need in situ data to fit the model parameters. However, the BIOMASAR algorithm is understood to estimate growing stock volume (GSV), so a further step is needed to estimate AGB, which requires spatial information on specific wood density and biomass expansion factors, or similar proxies.

Table 3). A key factor in the performance of regional methods was the amount and quality of available in situ data. In regions with abundant in situ data from a forest inventory, non-parametric data-driven methods, such as k-NN and MaxEnt were chosen, but in regions where data were scarce a model-based parametric approach was selected. Large numbers of forest plots were available in Sweden (22,548 plots) and Mexico (5140 plots). Poland (285 plots) and Kalimantan (247 forest plots) had fewer, but in Kalimantan airborne LiDAR biomass maps were developed to increase the size of the training and validation dataset. In the data-scarce region of the South Africa savannas only 37 1-ha AGB plots were available, but they were complemented with airborne LiDAR-based biomass maps derived from locally developed LiDAR models calibrated against field data (Naidoo et al., 2015) for validation. LiDAR airborne data was used for calibration and validation in Kalimantan, while only as validation in Eastern South Africa. This study assumed the AGB predicted by the LiDAR airborne to be error-free. However, the use of LiDAR data might introduce substantial errors in the AGB prediction originated from the ground-to-LiDAR model used (Saarela et al., 2016; Holm et al., 2017), which are not accounted for in this study.

The requirement for only a limited number of ground data points is an advantage of semi-empirical methods, such as the Water Cloud Model (WCM) + Bayesian inversion used in Eastern South Africa where only one parameter of the regression needs to be derived from in situ biomass data. However, the formulation of the WCM for this region does not produce estimates for AGB above ˜100 t ha$^{-1}$, and it is tuned for regions such as savannas with biomass below the saturation level at L-band (Mermoz et al., 2015). A method for global mapping, BIOMASAR (Santoro et al., 2011), is also based on the WCM and does not need in situ data to fit the model parameters. However, the BIOMASAR algorithm is understood to estimate growing stock volume (GSV), so a further step is needed to estimate AGB, which requires spatial information on specific wood density and biomass expansion factors, or similar proxies.

The most used dataset in this study was the L-band SAR ALOS PALSAR annual mosaics (Shimada et al., 2014; Shimada and Ohtaki, 2010), which were used in all regional methods except for Sweden. The saturation level of L-band SAR was found in previous studies to vary between 40 t ha$^{-1}$ and 150 t ha$^{-1}$ (Balzter et al., 2002a, b, Tansey et al.,

**Table 3**

Strengths and weaknesses of proposed methods/datasets when considering global implementation.

| METHOD | STRENGTHS | WEAKNESSES |
|---|---|---|
| Two step LiDAR + SAR multiple linear regression | • Low demand on *in situ* data<br>• Can combine different EO datasets | • Requires airborne LiDAR (costly and not always available)<br>• Overestimation at low AGB<br>• Underestimation at high AGB |
| WCM + Bayesian inversion | • Low demand on *in situ* data | • Requires large *in situ* plots (1 ha) (not always available)<br>• Assumes saturation of L-band SAR above 100 t ha$^{-1}$ (i.e., only applicable on low AGB areas) |
| Random Forests | • Can combine different EO datasets | • Medium / high demand on *in situ* data<br>• Overestimation at low AGB<br>• Underestimation at high AGB |
| MaxEnt | • Can combine different EO datasets | • High demand on *in situ* data<br>• Overestimation at low AGB<br>• Underestimation at high AGB |
| kNN | • Can combine different EO datasets | • SPOT data not freely available<br>• High demand on *in situ* data<br>• Overestimation at low AGB<br>• Underestimation at high AGB |

2004, Lucas et al., 2010; Hame et al., 2013) and generally HV gave higher saturation levels than HH polarization (Le Toan et al., 1992; Saatchi et al., 2007; Mitchard et al., 2009; Hamdan et al., 2011; Mitchard et al., 2012; Saatchi et al., 2011; Englhart et al., 2011; Hamdan et al., 2014). The saturation level at L-band depends on the geometry of the radar measurements, and therefore on forest type and environmental effects (Yu and Saatchi, 2016), as it can be observed in the different relationships found in the Yucatan peninsula and Central Mexico (Fig. 8). Use of the annual mosaics was also a limitation, as better results have been obtained with the use of multi-temporal SAR datasets due to the decrease in the retrieval error in AGB ranges to which the sensitivity of the SAR signal is weak (Antropov et al., 2013; Santoro et al., 2015b, 2006; Thiel and Schmullius, 2016; Antropov et al., 2017; Cartus et al., 2012). However, this was necessary due to the cost of acquiring the multiple ALOS PALSAR images used to generate the mosaics over large areas and for a given year.

For the model-based approaches, information provided by optical sensors, such as Landsat Percent Tree Cover (e.g. in Eastern South Africa), was needed for parameterisation (Bouvet et al., 2018). Such datasets were also used in the non-parametric machine learning methods, where they contributed towards improving model performance. Three out of the five regional methods used optical data as predictor variables. DEM data from SRTM were also used in Mexico and Poland as a predictor variable and in South Africa and Kalimantan for correcting or masking terrain effects, respectively. The use of topographic information by machine learning approaches for forests located in mountainous areas contributed to the estimation of AGB in Central Mexico.

The evaluation of the maps showed that a crucial limitation of the retrievals is that underestimation occurs at high AGB ranges, and overestimation at low AGB ranges. Remote sensing of AGB (using either reflectance or radar backscatter) is subject to decreasing sensitivity to AGB as biomass increases. Hence changes in AGB above a saturation level result in changes in the remotely sensed variable that are small compared to the variability in the signal. In these circumstances it is readily understood how linear regression would lead to these effects. The regression line always passes through the point defined by the mean of the reference data and the mean of the estimates, and the fitting effectively rotates the line about this point in order to reduce the sum of squared differences between the linear estimates and the reference data. For a concave curve, such as is produced by saturation, it is then inevitable that over-estimation will occur for low biomass and under-estimation for high biomass: getting a good fit for low biomass tends to steepen the line, while for high biomass it reduces the slope, and the regression line trades one against the other. In the case just discussed the model does not properly capture the relationship between the signal and the reference AGB, either due to insufficient calibration

data in the upper AGB range, or by fitting an inappropriate model to the AGB observations. However, if instead a fitting curve is used that correctly represents the saturation, bias is still to be expected for higher values of biomass. This occurs because, by definition, the backscatter values in the saturation zone are the result of random variation around the saturation level, with at most a weak dependence on biomass. Hence, although it is possible to use an estimator that assigns values of biomass above the saturation level, these represent the scatter in the data and the estimator will be biased towards the saturation level.

However, for correctly fitted data the overestimation at low AGB ranges is more difficult to explain. For SAR datasets, it could be connected to the high variability of the signal under soil moisture changes, as well as soil roughness (Mattia et al., 2009), but these do not apply to an optical-based method or to methods using SAR and optical data in synergy. Alternatively, it could be linked to the underestimation at high AGB ranges, as models provide the best fit by minimising an overall error term or cost function. The overestimation in the lower reference AGB ranges may stem from the model compensating for its inability to predict high AGB values accurately. In Kalimantan, the estimations were the highest, above 300 t ha$^{-1}$, and underestimation can only be observed above 200 t ha$^{-1}$. That underestimation occurs at such a high AGB level might be due to the large number of reference data points at high AGB levels compared to all other regions, which resulted in a fitted model with the largest errors and biases in the mid AGB ranges, between 50 and 150 t ha$^{-1}$, for which fewer calibration data were available (Table 2 and Fig. 4).

Avitabile et al. (2016) also found overestimations in the low AGB range and underestimations in the high AGB range when validating pan-tropical datasets (Saatchi et al., 2011; Baccini et al., 2012) against reference data. Several studies using the Random Forests regression algorithm have found similar behaviour at both ends of the AGB range, and report it as an effect of the averaging of tree-based algorithms (Baccini, 2004; Baccini et al., 2008; Avitabile et al., 2011; Urbazaev et al., 2018). Blackard et al. (2008) reported the same using a tree-based method (i.e. recursive partitioning regression), but suggested that saturation of optical data could explain the underestimation for high AGB densities, and scaling issues between plot and pixel could explain the underestimations at low AGB. This effect, characteristic of tree-based algorithms, could also explain the results in Poland, but cannot explain them in the other regions. Additionally other studies, such as Tsui et al. (2013) which used Kriging, Chopping et al. (2011) which used a geometric–optical canopy reflectance model, Del Frate and Solimini (2004) which used Neural Networks, or Sun et al. (2011), and Chi et al. (2015) which used multiple linear regression methods showed the same effects. Kattenborn et al. (2015) also reported the same effect for four semi- or non-parametric regression models (Random Forest, Generalized Additive Model, Generalized Boosted Regression Models
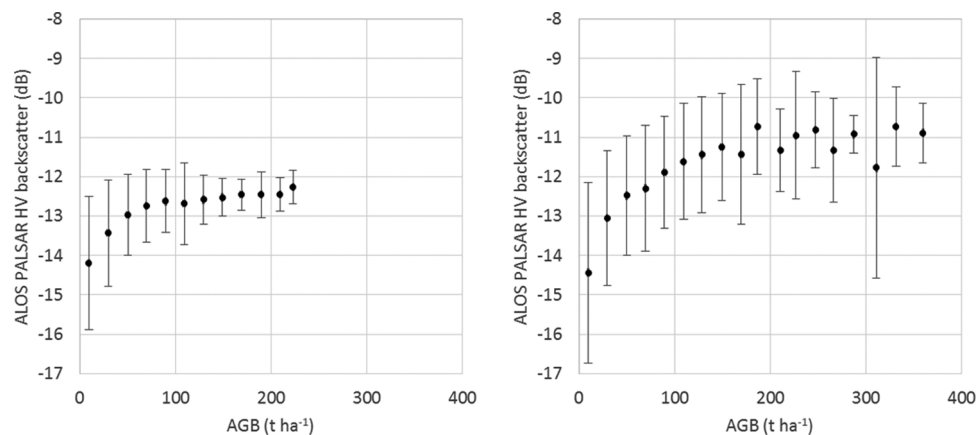


**Fig. 8.** Scatterplots of average ALOS PALSAR HV backscatter versus average reference AGB by reference AGB range. Error bars indicate the standard deviation of the ALOS PALSAR HV backscatter per reference AGB range for the Yucatan peninsula (left) and Central Mexico (right).

and Boosted Generalized Additive Model), and suggested insufficient calibration data at the low and high AGB ends as the main cause. However, the consistency across all these studies suggests that there is a fundamental problem in retrieving biomass from the available data, which may only be solved by the use of SAR data with higher sensitivity to small scattering elements such as C-band (e.g. Sentinel-1) on the lower AGB range, and data with greater sensitivity to large scattering objects (i.e. high biomass) such as the P-band (Villard and Le Toan, 2015) to be exploited by the ESA BIOMASS mission.

## 7. Conclusions

The regional forest AGB mapping methods presented here reflect both the variety of training data available in different regions and the diverse range of algorithmic choices of each regional team. However, the retrieved AGB values agree better with independent in situ data than those published recently (Saatchi et al., 2011; Baccini et al., 2012; Avitabile et al., 2016; Thurner et al., 2014). As the different methods were not compared on the same site, we cannot comment on relative performance. However, we can form conclusions based on the commonalities observed from the comparison of standardized accuracy assessments.

The key EO dataset used in most methods was the L-band ALOS PALSAR mosaics, which provides the highest sensitivity to AGB of the currently available spaceborne sensors. However, it is clear that all current spaceborne sensors (SAR and optical) are inadequate for accurately estimating AGB beyond 100-150 t ha$^{-1}$. The case studies presented here highlight challenges of using sub-optimal datasets for this task. Any estimation beyond this range was dominated by negative bias or presented large errors for any of the given study sites. The assessment indicates, however, that one could push this limit in certain conditions, as seen in Kalimantan or Central Mexico. This could be linked to the use of large amounts of in situ data in the case of Kalimantan. In Central Mexico, a forest structure which leads to a higher L-band saturation level, or the contribution of additional datasets (i.e. the DEM) could be the cause.

There is also a general problem with overestimation at low AGB densities which cannot be entirely explained by the datasets used, but rather as an intrinsic problem of the proposed algorithms to correctly capture the relationship between EO data and AGB in the low AGB range. This means that we might have to consider alternative regression schemes, or accept the biases at both ends of the biomass range, provided that the modelling framework captures the relationship between observations and biomass.

This aspect shall deserve substantial attention in future studies as currently existing models for large-scale biomass estimation rely on simplifying assumptions that may not fully encompass the complex interaction of the remote sensing signal with vegetation

The amount and type of reference data is also very relevant in terms of achieving the most reasonable AGB prediction model. Eastern South Africa used large plots (i.e. 1 ha), more suitable for calibration of EO methods (Réjou-Méchain et al., 2014), which might explain the good results for the low AGB ranges in this study area. Mexico, Sweden, Poland and Kalimantan relied on datasets of numerous small plots for calibration, supplemented with LiDAR in the case of Kalimantan. In the future, similar research should be based on homogeneous field-based datasets to avoid possible discrepancies resulting from the training data.

Better quality and more abundant large plots for calibration of the algorithms (Réjou-Méchain et al., 2014), the use of SAR time series (from Sentinel-1, ALOS 2 PALSAR 2, or future NovaSAR and NISAR) (Santoro et al., 2011; Antropov et al., 2017), the increasing availability of airborne or spaceborne LiDAR sensors like GEDI (Dubayah et al., 2014; Goetz et al., 2015) and MOLI (Kimura et al., 2017), satellites specifically designed for biomass estimation such as NISAR (Rosen et al., 2016, 2015), TanDEM-L (Moreira et al., 2015), and the P-band

BIOMASS SAR mission (Le Toan et al., 2011), and algorithms capable of reducing the errors in the low and high AGB ranges, are a promising way forward to improve global biomass estimates and reduce biases and errors in the map products in all biomass ranges.

## Author contributions

## Acknowledgements

## References

Anaya, J.A., Chuvieco, E., Palacios-Orueta, A., 2009. Aboveground biomass assessment in Colombia: a remote sensing approach. For. Ecol. Manage. 257, 1237–1246.

Antropov, O., Rauste, Y., Ahola, H., Hame, T., 2013. Stand-level stem volume of boreal forests from spaceborne SAR imagery at L-band. Ieee J. Sel. Top. Appl. Earth Obs. Remote. Sens. 6, 35–44.

Antropov, O., Rauste, Y., Häme, T., Praks, J., 2017. Polarimetric ALOS PALSAR time series in mapping biomass of boreal forests. Remote Sens. (Basel) 9, 999.

Asner, G.P., Mascaro, J., Anderson, C., Knapp, D.E., Martin, R.E., Kennedy-Bowdoin, T., Van Breugel, M., Davies, S., Hall, J.S., Muller-Landau, H.C., 2013. High-fidelity national carbon mapping for resource management and REDD+. Carbon Balance Manag. 8, 7.

Asner, G.P., Knapp, D.E., Martin, R.E., Tupayachi, R., Anderson, C.B., Mascaro, J., Sinca, F., Chadwick, K.D., Higgins, M., Farfan, W., 2014. Targeted carbon conservation at national scales with high-resolution monitoring. Proc. Natl. Acad. Sci. 111, E5016–E5022.

Avitabile, V., Camia, A., 2018. An assessment of forest biomass maps in Europe using harmonized national statistics and inventory plots. For. Ecol. Manage. 409, 489–498.

Avitabile, V., Herold, M., Henry, M., Schmullius, C., 2011. Mapping biomass with remote sensing: a comparison of methods for the case study of Uganda. Carbon Balance Manag. 6, 7.

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., 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.

Avitabile, V., Herold, M., Lewis, S., Phillips, O., Aguilar-Amuchastegui, N., Asner, G., Brienen, R., Devries, B., Gatti, R.G., Feldpausch, T., 2014. Comparative analysis and fusion for improved global biomass mapping. Book of Abstracts of the International Conference Global Vegetation Monitoring and Modeling (GV2M). pp. 251–252.

Baccini, A., 2004. Forest biomass estimation over regional scales using multisource data. Geophys. Res. Lett. 31.

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.

Baccini, A., Goetz, S.J., Walker, W.S., Laporte, N.T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P.S.A., Dubayah, R., Friedl, M.A., Samanta, S., Houghton, R.A., 2012. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nat. Clim. Change 2, 182–185.

Balzter, H., Baker, J.R., Hallikainen, M., Tomppo, E., 2002a. Retrieval of timber volume and snow water equivalent over a Finnish boreal forest from airborne polarimetric synthetic aperture radar. Int. J. Remote Sens. 23, 3185–3208.

Balzter, H., Talmon, E., Wagner, W., Gaveau, D., Plummer, S., Yu, J.J., Quegan, S., Davidson, M., Toan, T.L., Gluck, M., 2002b. Accuracy assessment of a large-scale forest cover map of central Siberia from synthetic aperture radar. Can. J. Remote. Sens. 28, 719–737.

Balzter, H., Rodriguez-Veiga, P., Tansey, K., Quegan, S., Carreiras, J., Persson, H.J., Fransson, J.E.S., Hoscilo, A., Ziolkowski, D., Dabrowska-Zielinska, K., Sterenczak, K., Englhart, S., Stängel, M., Siegert, F., Avitabile, V., Mermoz, S., Bouvet, A., Le Toan, T., Carvalhais, N., Soja, M., Eriksson, L., Santoro, M., Pathe, C., Schmullius, C., 2016. GLOBBIOMASS regional case studies - preparing the ground for global forest biomass mapping. ESA Living Planet Symposium.  Prague.

Bartholomé, E., Belward, A., 2005. Glc2000: a new approach to global land cover mapping from Earth observation data. Int. J. Remote Sens. 26, 1959–1977.

Beaudoin, A., Bernier, P., Guindon, L., Villemaire, P., Guo, X., Stinson, G., Bergeron, T., Magnussen, S., Hall, R., 2014. Mapping attributes of Canada's forests at moderate resolution through k NN and MODIS imagery. Can. J. For. Res. 44, 521–532.

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., Salajanu, D., Weyermann, 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.

Bojinski, S., Verstraete, M., Peterson, T.C., Richter, C., Simmons, A., Zemp, M., 2014. The concept of essential climate variables in support of climate research, applications, and policy. Bull. Am. Meteorol. Soc. 95, 1431–1443.

Bonan, G.B., 2008. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. Science 320, 1444–1449.

Bouvet, A., Mermoz, S., Le Toan, T., Villard, L., Mathieu, R., Naidoo, L., Asner, G.P., 2018. An above-ground biomass map of African savannahs and woodlands at 25m resolution derived from ALOS PALSAR. Remote Sens. Environ. 206, 156–173.

Breiman, L., 2001. Statistical Modeling: the two cultures (with comments and a rejoinder by the author). Stat. Sci. 16, 199–231.

Carreiras, J.M.B., Vasconcelos, M.J., Lucas, R.M., 2012. Understanding the relationship between aboveground biomass and ALOS PALSAR data in the forests of Guinea-Bissau (West Africa). Remote Sens. Environ. 121, 426–442.

Cartus, O., Santoro, M., Kellndorfer, J., 2012. Mapping forest aboveground biomass in the Northeastern United States with ALOS PALSAR dual-polarization L-band. Remote Sens. Environ. 124, 466–478.

Cartus, O., Kellndorfer, J., Walker, W., Franco, C., Bishop, J., Santos, L., Fuentes, J., 2014. A national detailed map of forest aboveground carbon stocks in Mexico. Remote Sens. 6, 5559–5588.

Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M.U.M., Saatchi, S., Santoro, M., Thurner, M., Weber, U., Ahrens, B., Beer, C., Cescatti, A., Randerson, J.T., Reichstein, M., 2014. Global covariation of carbon turnover times with climate in terrestrial ecosystems. Nature 514, 213–217.

Chi, H., Sun, G., Huang, J., Guo, Z., Ni, W., Fu, A., 2015. National forest aboveground biomass mapping from icesat/GLAS data and MODIS imagery in China. Remote Sens. 7, 5534–5564.

Chopping, M., Schaaf, C.B., Zhao, F., Wang, Z., Nolin, A.W., Moisen, G.G., Martonchik, J.V., Bull, M., 2011. Forest structure and aboveground biomass in the southwestern United States from MODIS and MISR. Remote Sens. Environ. 115, 2943–2953.

Del Frate, F., Solimini, D., 2004. On neural network algorithms for retrieving forest biomass from SAR Data. Geosci. Remote Sens. IEEE Trans. On 42, 24–34.

Dubayah, R., Goetz, S., Blair, J., Fatoyinbo, T., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., 2014. The global ecosystem dynamics investigation. AGU Fall Meet. Abs. 07.

Englhart, S., Keuck, V., Siegert, F., 2011. Aboveground biomass retrieval in tropical forests — the potential of combined X- and L-band SAR data use. Remote Sens. Environ. 115, 1260–1271.

Englhart, S., Keuck, V., Siegert, F., 2012. Modeling aboveground biomass in tropical forests using multi-frequency SAR Data- a comparison of methods. Selected topics in applied earth observations and remote sensing. IEEE J. 5, 298–306.

Evans, J.S., Cushman, S.A., 2009. Gradient modeling of conifer species using random forests. Landsc. Ecol. 24, 673–683.

Fagan, M., Defries, R., 2009. Measurement and monitoring of the world's forests: a review and summary of remote sensing technical capability, 2009-2015. Resour. Future.

Fao, 2005. Global forest resources assessment 2005. FAO Forestry Paper. Food and Agriculture Organization of the United Nations.

Fao, 2010. Global forest resources assessment 2010. FAO Forestry Paper. Food and Agriculture Organization of the United Nations.

Fao, 2015. Global Forest Resources Assessment 2015. Food and Agriculture Organization of the United Nations, Rome.

Fick, S.E., Hijmans, R.J., 2017. Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol.

Fraser, R.H., Li, Z., 2002. Estimating fire-related parameters in boreal forest using SPOT VEGETATION. Remote Sens. Environ. 82, 95–110.

Gallaun, H., Zanchi, G., Nabuurs, G.-J., Hengeveld, G., Schardt, M., Verkerk, P.J., 2010. Eu-wide maps of growing stock and above-ground biomass in forests based on remote sensing and field measurements. For. Ecol. Manage. 260, 252–261.

GLOBBIOMASS, 2015. Globbiomass Project Website [Online].   Available:. Friedrich-Schiller-University Jena, Institute of Geography, Department Of Earth Observation, Jena, Germany.  Http://Globbiomass.Org/2016.

Goetz, S.J., Hansen, M., Houghton, R.A., Walker, W., Laporte, N., Busch, J., 2015. Measurement and monitoring needs, capabilities and potential for addressing reduced emissions from deforestation and forest degradation under REDD + . Environ. Res. Lett. 10, 123001.

Hamdan, O., Aziz, H.K., Rahman, K.A., 2011. Remotely sensed L-band SAR data for tropical forest biomass estimation. J. Trop. For. Sci. 318–327.

Hamdan, O., Hasmadi, I.M., Aziz, H.K., 2014. Combination of SPOT-5 and ALOS PALSAR images in estimating aboveground biomass of lowland dipterocarp forest. IOP Conference Series: Earth And Environmental Science.  IOP Publishing, 012016.

Hame, T., Rauste, Y., Antropov, O., Ahola, H.A., Kilpi, J., 2013. Improved Mapping of Tropical Forests With Optical and SAR Imagery, Part II: Above Ground Biomass Estimation. Ieee J. Sel. Top. Appl. Earth Obs. Remote. Sens. 6, 92–101.

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.

Healey, S.P., Cohen, W.B., Yang, Z.Q., Krankina, O.N., 2005. Comparison of tasseled cap-based Landsat data structures for use in forest disturbance detection. Remote Sens. Environ. 97, 301–310.

Holm, S., Nelson, R., Ståhl, G., 2017. Hybrid three-phase estimators for large-area forest inventory using ground plots, airborne lidar, and space lidar. Remote Sens. Environ. 197, 85–97.

Houghton, R., Butman, D., Bunn, A.G., Krankina, O., Schlesinger, P., Stone, T., 2007. Mapping Russian forest biomass with data from satellites and forest inventories. Environ. Res. Lett. 2, 045032.

Hu, T., Su, Y., Xue, B., Liu, J., Zhao, X., Fang, J., Guo, Q., 2016. Mapping global forest aboveground biomass with spaceborne LiDAR, optical imagery, and forest inventory data. Remote Sens. 8, 565.

Kattenborn, T., Maack, J., Faßnacht, F., Enßle, F., Ermert, J., Koch, B., 2015. Mapping forest biomass from space–Fusion of hyperspectral EO1-hyperion data and Tandem-X and WorldView-2 canopy height models. Int. J. Appl. Earth Obs. Geoinf. 35, 359–367.

Kellndorfer, J., Walker, W., Lapoint, L., Bishop, J., Cormier, T., Fiske, G., Kirsch, K., Westfall, J., 2011. NACP Aboveground Biomass and Carbon Baseline Data (NBCD 2000), USA, 2000. Data Set.   [Online.] Available at. ORNL DAAC, Oak Ridge, Tennessee.  daac.ornl.gov..

Kimura, T., Imai, T., Sakaizawa, D., Murooka, J., Mitsuhashi, R., 2017. The overview and Status of vegetation lidar Mission, moli. 2017. Ieee International Geoscience and Remote Sensing Symposium (Igarss). pp. 4228–4230.

Kosmidis, I., 2014. Bias in parametric estimation: reduction and useful side-effects. Wiley Interdiscip. Rev. Comput. Stat. 6, 185–196.

Le Quéré, C., Andrew, R.M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A.C., Korsbakken, J.I., Peters, G.P., Canadell, J.G., Jackson, R.B., 2018. Global carbon budget 2017. Earth Syst. Sci. Data 10, 405.

Le Toan, T., Beaudoin, A., Riom, J., Guyon, D., 1992. Relating forest biomass to SAR data. Geosci. Remote Sens., IEEE Trans. On 30, 403–411.

Le Toan, T., Quegan, S., Davidson, M.W.J., Balzter, H., Paillou, P., Papathanassiou, K., Plummer, S., Rocca, F., Saatchi, S., Shugart, H., Ulander, L., 2011. The BIOMASS mission: mapping global forest biomass to better understand the terrestrial carbon cycle. Remote Sens. Environ. 115, 2850–2860.

Lefsky, M.A., 2010. A Global Forest canopy height map from the moderate resolution imaging spectroradiometer and the geoscience laser altimeter system. Geophys. Res. Lett. 37, L15401.

Liu, S., Zhou, T., Wei, L., Shu, Y., 2012. The spatial distribution of forest carbon sinks and sources in China. Chin. Sci. Bull. 57, 1699–1707.

Liu, Y.Y., Van Dijk, A.I.J.M., De Jeu, R.A.M., Canadell, J.G., Mccabe, M.F., Evans, J.P., Wang, G., 2015. Recent reversal in loss of global terrestrial biomass. Nat. Clim. Change 5 (5), 470–474.

Lu, D., Chen, Q., Wang, G., Moran, E., Batistella, M., Zhang, M., Vaglio Laurin, G., Saah, D., 2012. Aboveground forest biomass estimation with Landsat and Lidar data and uncertainty analysis of the estimates. Int. J. For. Res. 2012, 16.

Lucas, R., Armston, J., Fairfax, R., Fensham, R., Accad, A., Carreiras, J., Kelley, J., Bunting, P., Clewley, D., Bray, S., Metcalfe, D., Dwyer, J., Bowen, M., Eyre, T., Laidlaw, M., Shimada, M., 2010. An evaluation of the ALOS PALSAR L-Band backscatter - above ground biomass relationship Queensland, Australia: impacts of surface moisture condition and vegetation structure. Selected topics in applied earth observations and remote sensing. IEEE J 3, 576–593.

MacKinnon, Kathy, Gusti, Hatta, Mangalik, Arthur, Halim, Hakimah, 1996. The Ecology of Kalimantan, vol. 3 Oxford University Press.

Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Gao, F., Kutler, J., Lim, T.-K., 2006. A Landsat surface reflectance dataset for North America, 1990-2000. Geosci. Remote Sens. Lett. IEEE 3, 68–72.

Mattia, F., Satalino, G., Pauwels, V., Alexander, L., 2009. Soil moisture retrieval through a merging of multi-temporal L-band SAR data and hydrologic modelling. Hydrol. Earth Syst. Sci. 13, 343–356.

Mermoz, S., Le Toan, T., Villard, L., Réjou-Méchain, M., Seifert-Granzin, J., 2014. Biomass assessment in the Cameroon savanna using ALOS PALSAR data. Remote Sens. Environ. 155, 109–119.

Mermoz, S., Réjou-Méchain, M., Villard, L., Le Toan, T., Rossi, V., Gourlet-Fleury, S., 2015. Decrease of L-band SAR backscatter with biomass of dense forests. Remote Sens. Environ. 159, 307–317.

Mitchard, E., Saatchi, S., Baccini, A., Asner, G., Goetz, S., Harris, N., Brown, S., 2013. Uncertainty in the spatial distribution of tropical forest biomass: a comparison of pan-tropical Maps. Carbon Balance Manag. 8, 10.

Mitchard, E.T.A., Saatchi, S.S., Woodhouse, I.H., Nangendo, G., Ribeiro, N.S., Williams, M., Ryan, C.M., Lewis, S.L., Feldpausch, T.R., Meir, P., 2009. Using satellite radar backscatter to predict above-ground woody biomass: a consistent relationship across four different African landscapes. Geophys. Res. Lett. 36, L23401.

Mitchard, E.T.A., Meir, P., Ryan, C.M., Woollen, E.S., Williams, M., Goodman, L.E., Mucavele, J.A., Watts, P., Woodhouse, I.H., Saatchi, S.S., 2012. A novel application of satellite radar data: measuring carbon sequestration and detecting degradation in a Community Forestry Project in Mozambique. Plant Ecol. Divers. 6, 159–170.

Mitchard, E.T.A., Feldpausch, T.R., Brienen, R.J.W., Lopez-Gonzalez, G., Monteagudo, A., Baker, T.R., Lewis, S.L., Lloyd, J., Quesada, C.A., Gloor, M., Ter Steege, H., Meir, P., Alvarez, E., Araujo-Murakami, A., Aragão, L.E.O.C., Arroyo, L., Aymard, G., Banki, O., Bonal, D., Brown, S., Brown, F.I., Cerón, C.E., Chama Moscoso, V., Chave, J., Comiskey, J.A., Cornejo, F., Corrales Medina, M., Da Costa, L., Costa, F.R.C., Di Fiore, A., Domingues, T.F., Erwin, T.L., Frederickson, T., Higuchi, N., Honorio Coronado, E.N., Killeen, T.J., Laurance, W.F., Levis, C., Magnusson, W.E., Marimon, B.S., Marimon Junior, B.H., Mendoza Polo, I., Mishra, P., Nascimento, M.T., Neill, D., Núñez Vargas, M.P., Palacios, W.A., Parada, A., Pardo Molina, G., Peña-Claros, M., Pitman, N., Peres, C.A., Poorter, L., Prieto, A., Ramirez-Angulo, H., Restrepo Correa, Z., Roopsind, A., Roucoux, K.H., Rudas, A., Salomão, R.P., Schietti, J., Silveira, M., De Souza, P.F., Steininger, M.K., Stropp, J., Terborgh, J., Thomas, R., Toledo, M., Torres-Lezama, A., Van Andel, T.R., Van Der Heijden, G.M.F., Vieira, I.C.G., Vieira, S., Vilanova-Torre, E., Vos, V.A., Wang, O., Zartman, C.E., Malhi, Y., Phillips, O.L., 2014. Markedly divergent estimates of Amazon forest carbon density from ground plots and Satellites. Glob. Ecol. Biogeogr. 23, 935–946.

Moreira, A., Krieger, G., Hajnsek, I., Papathanassiou, K., Younis, M., Lopez-Dekker, P., Huber, S., Villano, M., Pardini, M., Eineder, M., 2015. Tandem-L: A highly innovative bistatic SAR mission for global observation of dynamic processes on the Earth's surface. Ieee Geosci. Remote. Sens. Mag. 3, 8–23.

Mucina, L., Rutherford, M.C., 2006. The vegetation of South Africa. South African National Biodiversity Institute, Lesotho and Swaziland.

Naidoo, L., Mathieu, R., Main, R., Kleynhans, W., Wessels, K., Asner, G., Leblon, B., 2015. Savannah woody structure modelling and mapping using multi-frequency (X-, C-and L-band) Synthetic Aperture Radar data. Isprs J. Photogramm. Remote. Sens. 105, 234–250.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., 2001. Terrestrial ecoregions of the world: a new map of life on earth. BioScience 51, 933–938.

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., 2011. A large and persistent carbon sink in the World's Forests. Science 333, 988–993.

Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S., Cardoso, A., 2013. Essential biodiversity variables. Science 339, 277–278.

Perrin, I.F., Nicolas, B., Jean-Stéphane, B., Nicolas, B., Valéry, G., Bruno, H., Mahmoud El, H., Frédéric, F., José, 2016. Regional scale rain-forest height mapping using regression-kriging of space borne and airborne LiDAR data: application on French Guiana. Remote Sens. 8, 240.

Phillips, S.J., Dud, M., Schapire, R.E., 2004. A maximum entropy approach to species distribution modeling. Proceedings of the Twenty-First International Conference on Machine Learning. 2004. 1015412: ACM, 83.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Modell. 190, 231–259.

Piao, S., Fang, J., Zhu, B., Tan, K., 2005. Forest biomass carbon stocks in China over the past 2 decades: estimation based on integrated inventory and satellite data. J. Geophys. Res. Biogeosci. 110.

Quegan, S., Yu, J.J., 2001. Filtering of multichannel SAR images. Geosci. Remote Sens. IEEE Trans. 39, 2373–2379.

Rauste, Y., 2005. Multi-temporal Jers SAR data in boreal forest biomass mapping. Remote Sens. Environ. 97, 263–275.

Reese, H., Nilsson, M., Pahlén, T.G., Hagner, O., Joyce, S., Tingelöf, U., Egberth, M., Olsson, H., 2003. Countrywide estimates of forest variables using satellite data and field data from the national forest inventory. Ambio A J. Hum. Environ. 32, 542–548.

Réjou-Méchain, M., Muller-Landau, H.C., Detto, M., Thomas, S.C., Toan, T., Saatchi, S.S., 2014. Local spatial structure of forest biomass and its consequences for remote sensing of carbon stocks. Biogeosciences 11.

Rignot, E., Salas, W.A., Skole, D.L., 1997. Mapping deforestation and secondary growth in Rondonia, Brazil, using imaging radar and thematic mapper data. Remote Sens. Environ. 59, 167–179.

Rodriguez-Veiga, P., Saatchi, S., Tansey, K., Balzter, H., 2016. Magnitude, spatial distribution and uncertainty of forest biomass stocks in Mexico. Remote Sens. Environ. 183, 265–281.

Rodríguez-Veiga, P., Wheeler, J., Louis, V., Tansey, K., Balzter, H., 2017. Quantifying forest biomass carbon stocks from space. Curr. For. Rep. 1–18.

Rosen, P.A., Hensley, S., Shaffer, S., Veilleux, L., Chakraborty, M., Misra, T., Bhan, R., Sagi, V.R., Satish, R., 2015. The nasa-isro sar mission-an international space partnership for science and societal benefit. Radar Conference (RadarCon). IEEE, 2015.

Rosen, P., Hensley, S., Shaffer, S., Edelstein, W., Kim, Y., Kumar, R., Misra, T., Bhan, R., Satish, R., Sagi, R., 2016. An update on the nasa-isro dual-frequency Dbf Sar (Nisar) mission. Geoscience and Remote Sensing Symposium (IGARSS) 2106–2108.

Saarela, S., Holm, S., Grafström, A., Schnell, S., Næsset, E., Gregoire, T.G., Nelson, R.F., Ståhl, G., 2016. Hierarchical model-based inference for forest inventory utilizing three sources of information. Ann. For. Sci. 73, 895–910.

Saatchi, S.S., Moghaddam, M., 2000. Estimation of crown and stem water content and biomass of boreal forest using polarimetric Sar imagery. Ieee Trans. Geosci. Remote. Sens. 38, 697–709.

Saatchi, S.S., Houghton, R., Dos Santos Alvala, R., Soares, J., Yu, Y., 2007. Distribution of aboveground live biomass in the Amazon Basin. Glob. Chang. Biol. 13, 816–837.

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. Proc. Natl. Acad. Sci. 108, 9899–9904.

Sandberg, G., Ulander, L.M.H., Fransson, J.E.S., Holmgren, J., Le Toan, T., 2011. L- and P-Band backscatter intensity for biomass retrieval in hemiboreal forest. Remote Sens. Environ. 115, 2874–2886.

Santoro, M., Eriksson, L., Askne, J., Schmullius, C., 2006. Assessment of stand-wise stem volume retrieval in boreal forest from Jers-1 L-band Sar backscatter. Int. J. Remote Sens. 27, 3425–3454.

Santoro, M., Beer, C., Cartus, O., Schmullius, C., Shvidenko, A., Mccallum, I., Wegmüller, U., Wiesmann, A., 2011. Retrieval of growing stock volume in boreal forest using hyper-temporal series of Envisat ASAR ScanSAR backscatter measurements. Remote Sens. Environ. 115, 490–507.

Santoro, M., Beaudoin, A., Beer, C., Cartus, O., Fransson, J.E., Hall, R.J., Pathe, C., Schmullius, C., Schepaschenko, D., Shvidenko, A., 2015a. Forest growing stock volume of the northern hemisphere: spatially explicit estimates for 2010 derived From Envisat Asar. Remote Sens. Environ. 168, 316–334.

Santoro, M., Eriksson, L.E., Fransson, J.E., 2015b. Reviewing Alos Palsar backscatter observations for stem volume retrieval in Swedish Forest. Remote Sens. 7, 4290–4317.

Schmullius, C., 2017. Globbiomass – estimates of biomass on global and regional scales. 37th International Symposium On Remote Sensing Of Environment (Isrse-37).

Schmullius, C., Seifert, F.M., Thiel, C., Pathe, C., Matejka, E., Quegan, S., Carreiras, J., Wegmüller, U., Santoro, M., Balzter, H., Tansey, K., Rodriguez-Veiga, P., Le Toan, T., Mermoz, S., Bouvet, A., Dabrowska-Zielinska, K., Hoscilo, A., Ziółkowski, D., Siegert, F., Englhart, S., Fransson, J., Herold, M., Avitabile, V., Häme, T., Rauste, Y., Shvidenko, A., Fritz, S., Schepaschenko, D., Sterenczak, K., Eriksson, L., Ulander, L., Reichstein, M., Carvalhais, N., 2015. Due globbiomass - estimates of biomass on a global scale. Isrse (May), 11–15 2015 Berlin.

Searle, E.B., Chen, H.Y.H., 2017. Tree size thresholds produce biased estimates of forest biomass dynamics. For. Ecol. Manage. 400, 468–474.

Sexton, J.O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., Kim, D.-H., Collins, K.M., Channan, S., Dimiceli, C., 2013. Global, 30-M resolution continuous fields of tree cover: Landsat-based rescaling of Modis vegetation continuous fields with Lidar-based estimates of error. Int. J. Digit. Earth 6, 427–448.

Shimada, M., Isoguchi, O., 2002. Jers-1 Sar mosaics of Southeast Asia using calibrated path images. Int. J. Remote Sens. 23, 1507–1526.

Shimada, M., Ohtaki, T., 2010. Generating large-scale high-quality Sar Mosaic datasets: application to Palsar data for global monitoring. Selected topics in applied earth observations and remote sensing. Ieee J. 3, 637–656.

Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., Lucas, R., 2014. New global forest/non-forest maps from Alos Palsar Data (2007–2010). Remote Sens. Environ. 155, 13–31.

Shvidenko, A.Z., Schepaschenko, D.G., 2014. Carbon budget of Russian Forests. Siberian J. For. Sci. 69–92.

Steffen, W., Noble, I., Canadell, J., 1998. The Terrestrial Carbon Cycle: Implication For The Kyoto Protocol. Science 280, 1393–1394.

Sun, G., Ranson, K.J., Guo, Z., Zhang, Z., Montesano, P., Kimes, D., 2011. Forest biomass mapping from Lidar and radar synergies. Remote Sens. Environ. 115, 2906–2916.

Tansey, K.J., Luckman, A.J., Skinner, L., Balzter, H., Strozzi, T., Wagner, W., 2004. Classification of forest volume resources using Ers Tandem Coherence and Jers Backscatter Data. Int. J. Remote Sens. 25, 751–768.

Thiel, C., Schmullius, C., 2016. The potential of Alos Palsar backscatter and Insar Coherence for Forest growing stock volume estimation in Central Siberia. Remote Sens. Environ. 173, 258–273.

Thurner, M., Beer, C., Santoro, M., Carvalhais, N., Wutzler, T., Schepaschenko, D., Shvidenko, A., Kompter, E., Ahrens, B., Levick, S.R., Schmullius, C., 2014. Carbon stock and density of northern boreal and temperate forests. Glob. Ecol. Biogeogr. 23, 297–310.

Thurner, M., Beer, C., Ciais, P., Friend, A.D., Ito, A., Kleidon, A., Lomas, M.R., Quegan, S., Rademacher, T.T., Schaphoff, S., Tum, M., Wiltshire, A., Carvalhais, N., 2017. Evaluation of climate-related carbon turnover processes in global vegetation models for boreal and temperate forests. Glob. Chang. Biol. 23, 3076–3091.

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.

Tsui, O.W., Coops, N.C., Wulder, M.A., Marshall, P.L., 2013. Integrating Airborne Lidar and Space-Borne Radar via Multivariate Kriging to estimate above-ground biomass. Remote Sens. Environ. 139, 340–352.

Urbazaev, M., Thiel, C., Cremer, F., Dubayah, R., Migliavacca, M., Reichstein, M., Schmullius, C., 2018. Estimation of forest aboveground biomass and uncertainties by integration of field measurements, Airborne Lidar, And Sar and optical satellite data in Mexico. Carbon Balance Manag. 13, 5.

Villard, L., 2009. Forward & Inverse Modeling For Synthetic Aperture Radar Observables In Bistatic Configuration. Applications In Forest Remote Sensing. Institut Supérieur De L'aéronautique Et De L'espace.

Villard, L., Borderies, P., 2007. backscattering border effects for forests At C-band. Piers Online 3, 731–735.

Villard, L., Le Toan, T., 2015. Relating P-band Sar intensity to biomass for tropical dense forests in hilly terrain: Γ0 Or T0? Ieee J. Sel. Top. Appl. Earth Obs. Remote. Sens. 8, 214–223.

Whittaker, R.H., 1962. Classification of natural communities. Bot. Rev. 28, 1–239.

Whittaker, R.H., 1970. Communities and Ecosystems. Communities and Ecosystems.

Xu, L., Saatchi, S.S., Yang, Y., Yu, Y., White, L., 2016. performance of non-parametric algorithms for spatial mapping of tropical forest structure. Carbon Balance Manag. 11, 18.

Yin, G., Zhang, Y., Sun, Y., Wang, T., Zeng, Z., Piao, S., 2015. Modis based estimation of forest aboveground biomass in China. PLoS One 10, E0130143.

Yu, Y., Saatchi, S., 2016. Sensitivity of L-band sar backscatter to aboveground biomass of global forests. Remote Sens. 8, 522.