SAR tomography for the retrieval of forest biomass and height: Cross-validation at two tropical forest sites in French Guiana

Dinh Ho Tong Minh a,b,⁎, Thuy Le Toan b, Fabio Rocca c, Stefano Tebaldini c, Ludovic Villard b, Maxime Réjou-Méchain d,e, Oliver L. Phillips f, Ted R. Feldpausch b, Pascale Dubois-Fernandez h, Klaus Scipal i, Jérôme Chave d

a Institut national de Recherche en Sciences et Technologies pour l'Environnement et l'Agriculture (IRSTEA), UMR TETIS, Montpellier, France
b Centre d'Etudes Spatiales de la Biosphère (CESBIO), UMR CNRS 5126, University of Paul Sabatier, Toulouse, France
c Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milano, Italy
d Laboratoire Evolution et Diversité Biologique, UMR CNRS 5174, University of Paul Sabatier, Toulouse, France
e Laboratoire d'Écologie Funcionnelle et Gestion des Systèmes Forestiers, MNHN, Université de Paris, CNRS, Jussieu, France
f School of Geography, University of Leeds, University Road, Leeds LS2 9JT, UK
g Geography, College of Life and Environmental Sciences, University of Exeter, UK
h Office National d'Etudes et de Recherches Aerospatiales (ONERA), Toulouse, France
i European Space Research and Technology Centre (ESTRACK), Noordwijk, the Netherlands

A R T I C L E   I N F O
Article history:
Received 19 March 2014
Received in revised form 17 December 2015
Accepted 30 December 2015
Available online xxxx

Keywords:
Aboveground biomass
BIOMASS mission
French Guiana
Paracou
Nouragues
TropiSAR
P-band SAR tomography
Tomography phase
Vertical forest structure

A B S T R A C T
Developing and improving methods to monitor forest carbon in space and time is a timely challenge, especially for tropical forests. The next European Space Agency Earth Explorer Core Mission BIOMASS will collect synthetic aperture radar (SAR) data globally from employing a multiple baseline orbit during the initial phase of its lifetime. These data will be used for tomographic SAR (TomoSAR) processing, with a vertical resolution of about 20 m, a resolution sufficient to decompose the backscatter signal into two to three layers for most closed-canopy tropical forests. A recent study, conducted in the Paracou site, French Guiana, has already shown that TomoSAR significantly improves the retrieval of forest aboveground biomass (AGB) in a high biomass forest, with an error of only 10% at 1.5-ha resolution. However, the degree to which this TomoSAR approach can be transferred from one site to another has not been assessed. We test this approach at the Nouragues site in central French Guiana (ca 100 km away from Paracou), and develop a method to retrieve the top-of-canopy height from TomoSAR. We found a high correlation between the backscatter signal and AGB in the upper canopy layer (i.e., 20–40 m), while lower layers only showed poor correlations. The relationship between AGB and TomoSAR data was found to be highly similar for forests at Nouragues and Paracou. Cross validation using training plots from Nouragues and validation plots from Paracou, and vice versa, gave an error of 16–18% of AGB using 1-ha plots. Finally, using a high-resolution LiDAR canopy model as a reference, we showed that TomoSAR has the potential to retrieve the top-of-canopy height with an error to within 2.5 m. Our analyses show that the TomoSAR-AGB retrieval method is accurate even in hilly and high-biomass forest areas and suggest that our approach may be generalizable to other study sites, having a canopy taller than 30 m. These results have strong implications for the tomographic phase of the BIOMASS spaceborne mission.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Forests play a key role in the global carbon cycle, and hence in the global climate (Wright, 2005; Pan et al., 2011). However, this role remains poorly characterized quantitatively, as compared to other ecosystems due to the practical difficulties in measuring forest biomass stocks over broad scales. Over the past few years, considerable progress has been made in mapping forest ecosystem biomass stocks using a range of remote sensing technologies (Saatchi et al., 2011b; Baccini et al., 2012; Mitchard et al., 2009; Mermoz et al., 2015). However, these studies have limitations associated with limited sensor sensitivity to biomass, inappropriate sampling intensity, and limited validation of the methodology. These maps are least accurate in high carbon stock forests, predominantly found in the tropics, where existing large-scale remotely-sensed biomass maps conflict substantially and with field-based estimates of spatial biomass patterns (e.g., Mitchard et al., 2005).
Tropical forests are highly complex, varied, and often threatened. In this context there is a critical need to develop new technologies that can help survey and monitor tropical forests.

Delivering accurate global maps of forest aboveground biomass (AGB) and height is the primary objective of BIOMASS, the next European Space Agency (ESA) Earth Explorer Core Mission (Le Toan et al., 2011). The BIOMASS satellite is planned for a 2020 launch date. To achieve the goal of wall-to-wall mapping of forest AGB, the BIOMASS mission features, for the first time from space, a fully polarimetric, P-band (435 MHz, ~69 cm wavelength, and 6 MHz bandwidth) Synthetic Aperture Radar (SAR). The low frequency ensures that the transmitted wave can penetrate the vegetation down to the ground even in dense multi-layer tropical forests (Smith-Jonforsen, Ulander, & Luo, 2005; Ho Tong Minh, Le Toan, et al., 2014). The satellite will operate in two different observation phases. The tomographic phase will last for one year and will result in one global forest AGB and total canopy height map at 200-m resolution. It will be followed by an interferometric phase, which will last for four years and will provide updated global forest AGB maps every six months (Ho Tong Minh, Tebaldini, Rocca, Le Toan, Villard, et al., 2015).

The algorithm for forest AGB retrieval based on P-band SAR has been developed during the BIOMASS Mission Assessment Phase (Phase A), based on airborne data collected over boreal and tropical forests (Sandberg, Ulander, Fransson, Holmgren, & Le Toan, 2010; Ho Tong Minh, Le Toan, et al., 2014; Villard & Le Toan, 2015). It makes full use of information on Polarimetric SAR (PolSAR) backscatter intensity and the Polarimetric Interferometric (PollnSAR) phase information. PolSAR algorithms combine statistical and physical models to derive AGB based on intensity measurements in all polarizations (Le Toan, Beaudoin, Riom, & Guyoni, 1992; Sandberg et al., 2011). These algorithms usually perform better for low biomass values (typically less than 200 t/ha in dry matter units), whereas at high AGB, signal intensity exhibits a saturation effect that affects biomass retrieval. PollnSAR technique combines two PolSAR measurements from slightly different orbits to obtain an estimate of forest height; this canopy height is subsequently converted into AGB using field-derived allometric equations (Saatchi et al., 2011a; Le Toan et al., 2011). By combining AGB estimates from these two complementary techniques, AGB maps may be produced with less than 20% root mean square error (RMSE), at a resolution of 4-ha (Le Toan et al., 2011). To achieve this performance, however, AGB estimation algorithms need to be accurately tuned, so as to take into account noise factors that affect radar measurements, primarily terrain topography and ground moisture status (Ho Tong Minh, Le Toan, et al., 2014; Van Zyl, 1993).

The analysis and evaluation of data collected during the tomography phase is essential to achieving the goals of the BIOMASS mission. The satellite’s orbit is designed to gather multiple acquisitions over the same sites from slightly different orbital positions, so as to image forest vertical structure through SAR tomography (henceforth referred to as TomoSAR) (Reigber & Moreira, 2000; Ho Tong Minh, Tebaldini, Rocca, Le Toan, Villard, et al., 2015). Hence, for the first time, BIOMASS will provide quantitative information on forest structure through P-band TomoSAR from space.

The potential of P-band TomoSAR to characterize forest structure was previously assessed in a number of studies relating forest vertical structure to forest biomass (Tebaldini & Rocca, 2012; Mariotti d’Alessandro, Tebaldini, & Rocca, 2013; Ho Tong Minh, Le Toan, et al., 2014). The TropiSAR campaign carried out in 2009 in French Guiana offered the first opportunity to test TomoSAR for tropical forest areas (Dubois-Fernandez et al., 2012). TropiSAR data have been acquired for TomoSAR processing at two forest sites, the Paracou forest and the Nouragues forest, about 100 km apart. In a previous study we conducted at the Paracou site, the signal at P-band coming from upper vegetation layers was found to be strongly correlated with forest AGB, for values ranging from 250 t/ha to 450 t/ha (Ho Tong Minh, Le Toan, et al., 2014). This finding was used to construct a simple AGB model having a RMSE of only 10% at a resolution of 1.5 ha. These results suggest that TomoSAR methods hold promise for accurately mapping forest biomass in tropical areas.

The robustness of the TomoSAR algorithm, however, needs further evaluation to different sites. Here we provide the first such assessment by performing a cross-comparison between two French Guiana tropical forest sites, namely Paracou and Nouragues. In addition we report on the performance of forest top height retrieved from the TomoSAR data at both sites. Specifically, we address the following questions: (1) Can the TomoSAR algorithm be parameterized for a landscape on hilly terrain?; (2) Is the relationship between TomoSAR and AGB transferable across tropical forest sites?; (3) Is the forest top height retrieval algorithm transferable? Finally we discuss the implications of these findings for the tomographic phase of the BIOMASS spaceborne mission.

2. Methods

2.1. Field data

The present study was conducted at two sites in French Guiana. The first site, the Nouragues Ecological Research Station, is located 120 km south of Cayenne, French Guiana (4 05′ N, 52 40′ W). This site is a protected natural reserve characterized by a lowland moist tropical rainforest. The climate is humid with a mean annual rainfall of 2861 mm/year (average 1992–2012), a short dry season in March and a longer 2-month dry season from late August to early November. The site is topographically heterogeneous, with a succession of hills ranging between 26 and 280 m above sea level (asl) and a granitic outcrop (Inselberg) reaching 430 m asl (the mean ground slope is greater than 5 at a 100-m resolution). The study area encompasses three main types of geological substrates, a weathered granitic parent material with sandy soils of variable depths, a laterite crust issued from metavolcanic rock of the Paramaca formation with clayey soils and a metavolcanic parent material. There has been no obvious forest disturbance by human activities in the past 200 years. One hectare of forest includes up to 200 tree species with a diameter at breast height (DBH) > 10 cm. Top-of-canopy height reaches up to 55 m with an average value around 35 m. At Nouragues, ground-based AGB was inferred from two large and long term permanent plots, namely Grand Plateau (1000 × 100 m²) and Petit Plateau (400 × 300 m²), both established in 1992–1994 and regularly surveyed to the present. The two plots were subdivided in 100 × 100 m² subplots, resulting in 22 study plots of 1-ha. We used tree census data conducted at the end of 2008. Five additional plots were also considered in the analyses, three of 1-ha (100 × 100 m²) in terra-firme forest (Pararé-ridge established in 2010; Lor in 2010; Ringler in 2012) and two 0.25-ha plots (50 × 50 m²) in permanently flooded forests (Bas_fond 1 and Bas_fond 2 both in 2012).

The second study area is located at the Paracou station, near Sinnamary, French Guiana (5 18′ N, 52 55′ W). The climate is also humid with a mean annual rainfall of 2980 mm/year (30 years period) and a 2-month dry season occurring from late August to early November. The Paracou site is fairly flat and has a homogeneous topography (5–50 m asl), but with deep drainage gullies flowing into the Sinnamary River. The most common soils at Paracou are shallow ferrallitic soils which are limited in depth by a more or less transformed loamy saprolithes (Gourlet-Fleury, Guehl, & Laroussinie, 2004). Following forest censuses, the number of tree species is estimated to be approximately 140–160 species/ha (trees with DBH ≥ 10 cm). Top-of-canopy height reaches up to 45 m with the average value around 30 m. In Paracou, in-situ forest measurements were available from 16 permanent plots established since 1984. There are 15 plots of 250 × 250 m² (6.25 ha) and one plot of 500 × 500 m² (25 ha). From 1986 to 1988, nine of these 15 6.25-ha plots underwent three different mild to severe logging treatments to study forest regeneration after logging (Gourlet-Fleury et al., 2004). Logging treatments had a significant impact on current...
AGB stocks (Blanc et al., 2009). As at the Nouragues site, we subdivided these large plots in 100 × 100 m². This resulted in 85 field plot units for the Paracou site. To match the BIOMASS resolution, we also subdivided all large plots in 200 × 200 m² subplots, resulting in 19 4-ha plots.

At both sites, the two forests are moist closed-canopy tropical forests. Nouragues forest has a slightly higher top canopy and above-ground biomass stock and is on a more hilly terrain. However, the floristic composition is largely similar (dominant tree families are Fabaceae, Sapotaceae, Burseraceae, Lecythidaceae, Chrysobalanaceae, and Moraceae), and is typical of most forests at the north-eastern end of the long pan-Amazon floristic gradient (e.g., (ter Steege et al., 2006)).

In each permanent sampling plot, living trees ≥ 10 cm DBH were mapped, diameter measured to the nearest 0.5 cm at 1.3 m above the ground, and botanically identified when possible. For trees with buttresses, stilt roots or irregularities, stem diameter was measured 30 cm above the highest irregularity. The point of measurement was marked with permanent paint on the stem. Trees ≤ 10 cm DBH and lianas were disregarded in the census, but these contribute a small fraction of the total AGB.

A subset of tree heights was measured at Nouragues (2462 trees) and Paracou (1157 trees). These were used to construct plot-specific height-diameter allometries in each plot using a model of the form:

$$\ln(H) = a + b \times \ln(DBH) + c \times \ln(DBH)^2$$  \hspace{1cm} (1)

where $H$ is the total tree height (Rejou-Mechain et al., 2015). In Paracou, a single height diameter model was used for all 62.5-ha plots but a specific model was used for the 25-ha plot as this is known to have more slender trees (Vincent, Sabatier, & Ruitersha, 2014).

Above-ground biomass of each tree ($AGB_t$) was estimated using the equation in (Chave et al., 2005):

$$AGB_t = 0.0509 \times \rho \times DBH^2 \times \overline{H}$$  \hspace{1cm} (2)

where $\overline{H}$ is the tree height estimated using the height-diameter Eq. (1) and $\rho$ is the oven-dry wood specific gravity in g/cm³. A more recent allometric equation was published in Chave et al. (2014) but it gave essentially identical AGB values (within 2%). Wood specific gravity $\rho$, was inferred from the species identification of the trees using a global wood density database (Chave et al., 2009). We assigned a $\rho$ value to each tree corresponding to the mean $\rho$ for species found in the database. Only $p$ measurements made in tropical South America (4182 trees) were considered in order to limit the bias due to regional variation of wood density (Muller-Landau, 2004; Chave et al., 2006). When no reliable species identification or no wood density information at the species level was available, the mean wood density at higher taxonomic level (i.e. genus, family) or at the plot level was attributed to the tree. In each plot, AGB was summed across trees and normalized by plot area to obtain AGB density in t/ha, in dry biomass units (note that AGB in dry biomass units may be converted into carbon units using a 0.48 ratio).

2.2. LiDAR data

Airborne LiDAR campaigns were also conducted in the study sites to serve as a reference repository of canopy height estimates. In the Nouragues site, an airborne LiDAR survey was conducted in 2012, covering an area of 2400 ha. A canopy height model was generated from the cloud data at 1-m resolution using the FUSION software ((McGaughey, 2012); Details on canopy model construction can be found in Rejou-Mechain et al. (2015)). At the Paracou study site, an airborne LiDAR survey was conducted in 2008, covering an area of 1200 ha. The canopy model was generated by the ALTOA society using the TerraScan software ((Terrasolid, 2008); Details on the LiDAR data can be found in Vincent et al. (2012)).

2.3. SAR data-sets

The TropiSAR study was conducted in the summer of 2009, and SAR airborne campaigns covered both Nouragues and Paracou sites flying multiple baselines, so as to allow tomographic processing. The SAR system used in the TropiSAR campaign was the ONERA airborne system SETHI (Dubois-Fernandez et al., 2012). The P-band SAR had a bandwidth of 335–460 MHz (125 MHz) and the resolution was 1 m in slant range and 1.245 m in azimuth direction (Dubois-Fernandez et al., 2012). Datasets of the TropiSAR campaign are available as an ESA archive through the EOPI portal (http://eopiesaint).

At Nouragues, tomographic data-sets consisted of five fully polarimetric Single Look Complex (SLC) images at P-band acquired on 14 August 2009. The baselines have been spaced vertically with a spacing of 15 m. The flight trajectory was lower than the reference line (3962 m) with a vertical shift of 15 m, 30 m, 45 m and 60 m, respectively. At Paracou, tomographic data-sets consisted of 6 fully polarimetric SLC images at P-band (and L-band) acquired on 24 August 2009. As for Nouragues, the baselines had a spacing of 15 m with a reference line of 3962 m, but an additional vertical shift at 75 m. In both data-sets, with the vertical shift of 15 m, the height of ambiguity was 110 m in near range and 210 m in far range, enabling unambiguous imaging of the forest volume.

Since the tomographic flight lines were in a vertical plane rather than in a horizontal plane, the phase to height factor and the height of ambiguity had a small variation across the scene swath (Dubois-Fernandez et al., 2012). The resulting vertical resolution is 20 m, whereas forest height ranges from 20 m to over 40 m. These features make it possible to map the 3-D distribution of the reflectivity by a coherent focusing, see Section 2.4.

The Nouragues and Paracou SAR images are shown in Fig. 1. In the Nouragues image, almost the whole scene is forested except the Arataye river in the south and the top of the Inselberg in the northwest. In the Paracou image, the Sinnamary river and the bare terrain areas can be observed. In both images, the texture of the river and the bare terrain areas are uniform as compared to the forested areas.

2.4. TomoSAR processing

The rationale of TomoSAR is to employ multiple flight tracks, nearly parallel to each other, as shown in the left panel of Fig. 2. The ensemble of all flight lines allows formation of a 2-D synthetic aperture, with the possibility to focus the signal in the whole 3-D space. In other words, by exploiting TomoSAR, multi-baseline SLC data can be converted into a new multi-layer SLC data stack where each layer represents scattering contributions associated with a certain height, as shown in the right panel of Fig. 2.

Let us consider a multi-baseline data-set of SLC SAR images acquired by flying the sensor along N parallel tracks, and let $y_n(r, x)$ denote the pixel at slant range, azimuth location $(r, x)$ in the $n-th$ image. Assuming that each image within the data stack has been resampled on a common master grid, and that phase terms due to platform motion and terrain topography have been compensated, the following model holds (Bamler & Hartl, 1998; Reigber & Moreira, 2000; Tabetaldi, 2010):

$$y_n(r, x) = \int S(\xi, r, x) \exp \left( \frac{4\pi}{\lambda} b_n \xi \right) d\xi$$  \hspace{1cm} (3)

where: $b_n$ is the normal baseline relative to the $n-th$ image with respect to a common master image; $\lambda$ is the carrier wavelength; $\xi$ is the cross range coordinate, defined by the direction orthogonal to the Radar Line-of-Sight (LOS) and the azimuth coordinate; $S(\xi, r, x)$ is the average scene complex reflectivity within the slant range, azimuth, cross range resolution cell, as shown in Fig. 3. Eq. (3) states that SAR multi-baseline data and the cross range distribution of the
scene reflectivity constitute a Fourier pair. Accordingly, the latter can be retrieved by taking the Fourier Transform of the data along the baseline direction.

\[
S(\xi, r, x) = \sum_{n=1}^{N} y_n(r, x) \exp\left(-j\frac{4\pi \lambda b_n \xi}{\lambda r}\right)
\]

As a result, TomoSAR processing allows us to retrieve the cross range distribution of the scene complex reflectivity at each range and azimuth location, hence providing fully 3-D imaging capabilities. The final conversion from cross range to height is then obtained through straightforward geometrical arguments. The resulting vertical resolution is approximately (Reigber & Moreira, 2000):

\[
\Delta z \approx \frac{\lambda r \sin \theta}{2 b_{\text{max}}}
\]

where \( \theta \) is the radar look angle and \( b_{\text{max}} \) the overall normal baseline span. Eq. (5) defines the so called Rayleigh limit. This way of processing does not optimize vertical resolution but ensures good radiometric accuracy in the vertical direction. An alternative approach would be to resort to sophisticated spectral estimation techniques such as MUSIC, CAPON, RELAX, or Compressive sensing algorithms (Zhu & Bamler, 2010; Gini, Lombardini, & Montanari, 2002; Lombardini & Reigber, 2003). Such algorithms, however, are optimized for the problem of detecting and localizing point targets, whereas they result in poor radiometric accuracy in the case of distributed targets.

To apply the simple approach depicted above, it is usually necessary to take a number of factors into account (Ho Tong Minh, Le Toan et al., 2014). First, the baseline distribution is not uniform due to atmospheric turbulences affecting the airborne flight trajectory. Second, the phases of the SLC data are affected by slow varying phase disturbances caused by uncompensated platform motion. Both factors affect tomographic focusing, leading to a blurring of the processed data, and hence need to be

[Fig. 1. P-band SAR image (8 km × 6 km) in Pauli false color (R:|HH-VV|, G: 2|HV|, B: |HH + VV|, where H and V refer to horizontal and vertical linear polarizations, respectively). The North is on the top. (a) Nouragues, the near range is on the left. (b) Paracou, the near range is on the right. The in situ AGB measurements are outlined with a label identifying the plot name. The white dash rectangles are relative to the area where LiDAR forest height data is available. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 2. Left panel: a schematic view of the tomography acquisition. Right panel: multi-layer images, each of which represents scattering contributions associated with a certain height.]
corrected. Third, terrain topography has to be considered, as it plays a key role for studying the relation between TomoSAR and in-situ measurements.

After these pre-processing steps, tomographic imaging is performed simply by taking the Fourier Transform (with respect to the normal baseline) of the multi-baseline SLC data set at every slant range, azimuth location. The result of this operation is a multi-layer SLC stack, where each layer is referred to a fixed height above the terrain. We will hereinafter refer to each image within the multi-layer data stack simply by the associated height (i.e.: 15 m layer, 30 m layer…), or as ground layer for the image focused at 0 m. A detailed step by step description of the processing is given in Ho Tong Minh, Le Toan, et al. (2014). Fig. 4a and b show the HV backscatter for layers at ground layer 0 m, 15 m, and 30 m over the Nouragues and Paracou sites, respectively. To provide a comparison we also show the backscatter relative to one image from the original multi-baseline data-stack (i.e. non-tomographic).

We then evaluated the relationship between backscatter for different layer heights and in-situ AGB using the slope of a least-square linear regression and the Pearson coefficient $r_P$. It is well-known that the cross-polarization HV have a better correlation with AGB than the co-polarization HH or VV (see for instance (Ho Tong Minh, Le Toan, et al., 2014)). Hence to focus the discussion we only report on the HV results in this paper.

We define a simple AGB model assuming a classical log law:

$$A GB = a \times \log_{10}(P_t) + b,$$

(6)

where $A GB$ is the estimated forest AGB, $P_t$ is the HV backscatter of a given tomographic layer, and $a$, $b$ are two parameters to be calibrated using training data. These parameters were estimated by using 10 training samples selected randomly out of 112 plots (i.e. calibration dataset). To assess model performance, the retrieved AGB values were then compared with the in-situ AGB of the remaining samples (i.e. validation dataset) to estimate the RMSE of the model.

Finally, to simulate BIOMASS equivalent data we reprocessed the high-resolution airborne data (125 MHz of bandwidth) to generate a new data stack with 6 MHz bandwidth and an azimuth resolution of 12 m. The overall baseline span was fixed to the critical value of BIOMASS (4610 m), 6 passes were used, resulting in the height of ambiguity 110 m and the vertical resolution 20 m (Ho Tong Minh, Tebaldini, Rocca, Le Toan, Villard, et al., 2015). Based on this reprocessed data-set we examined the relationship of TomoSAR products to biomass. The reader is referred to Ho Tong Minh, Tebaldini, Rocca, Le Toan, Villard, et al. (2015) for the description of the BIOMASS simulator, for which BIOMASS tomographic data were emulated at the Paracou site.

2.5. Forest top height retrieval

In tropical rainforests, where canopy structure is more complex than any other forest type, estimating forest top height in the field is a challenging task because it is often hard to clearly identify the top leaf or branch of a tree in the canopy. Due to its ability to accurately characterize the vertical structure of tropical forests, TomoSAR can be used to estimate forest top height. Forest vertical structure can be observed by taking a tomographic profile, i.e. a slice of the multi-layer data stack (Fig. 5).

By retrieving the 3-D backscatter distribution from the multi-layer SLC, it is possible to show the vertical backscatter distribution function. Each vertical distribution is characterized by an effective scattering center, where most of the backscatter is concentrated, the so called phase center.
center $H_c$. This can be written in formula,

$$H_c(r, x) = \arg \max\{P(z, r, x)\}, \tag{7}$$

where $P(z, r, x)$ is the vertical backscatter at slant range, azimuth location $(r, x)$ in vertical direction $z$. Fig. 6a shows an example of HV vertical backscatter distribution with respect to the phase center at 10 m, 20 m and 30 m, from the 3-D backscatter distribution in Nouragues site. Fig. 6b shows a schematic view of the vertical backscatter distribution, in which it can be assumed that the shape of the distribution can be divided into three zones. The first corresponds to the zone where most of the backscatter is concentrated, i.e. the phase center zone. The second is the power loss zone, where the backscatter undergoes a loss along the vertical direction from the phase center location. Further away, the backscatter is dominated by noise, unlikely to be associated with any physically relevant components. Therefore, by identifying the power loss from the phase center location in the upper envelope of the profile, forest top height $H$ can be retrieved (Tebaldini & Rocca, 2012; Ho Tong Minh, Tebaldini, Rocca, Le Toan, Villard, et al., 2015). This can be written in formula,

$$H(r, x) = \arg \min\{|P(z', r, x) - P(H_c, r, x) - K|\}, \tag{8}$$

where $P(H_c, r, x)$ is the backscatter at phase center $H_c$, $K$ is the power loss value, $z'$ is the height values ranging from $H_c$ to the upper envelope of the profile, e.g. 60 m.

Since the forest top height retrieval is dependent on the choice of the power loss value $K$, we used top-of-canopy height LiDAR models to select an optimal power loss value.

3. Results

The three tomographic layers (0, 15 and 30 m) were found to be different in their information content, with the upper vegetation layer (30 m) having the highest correlation between the backscatter and AGB (Fig. 7). For this layer, the Pearson correlation was 0.75 and the slope indicates an increase of $\sim 1.8$ dB per 100 t/ha for a range of AGB of 200–600 t/ha. For the lower layers, the linear correlations were weak, and even negative for the ground layer. Our results thus show that the best TomoSAR estimator to retrieve AGB was based on the HV backscatter at 30 m. Results of the calibration and validation with field data are reported in Fig. 8 and showed a model RMSE of 15%.

Second, to test the robustness and transferability of the relationship between AGB and TomoSAR data, we used 27 plots from Nouragues for training and 85 samples from Paracou for validation, and vice versa. The RMSE values from these cross-validation models were only slightly higher than to those obtained by using both training and validation samples from the same study site (Fig. 9).

Third, we retrieved top heights from the tomographic profile (Fig. 5). Using the top-of-canopy height LiDAR model we evaluated the forest top height location corresponding to a power loss value, with respect to the phase center, ranging from 0 to $-10$ dB (Fig. 10). In both the Nouragues and Paracou sites, the bias associated with the TomoSAR top-height retrieval decreased regularly with the power loss but the RMSE was significantly lower at a power loss of 2 dB reaching only 2.5 m and 2 m in Nouragues and Paracou, respectively. Using a power loss value of $-2$ dB at the Nouragues and Paracou site, we then extrapolated the TomoSAR top-of-canopy height retrieval estimates over the whole area covered by the LiDAR campaigns for comparison purpose (Fig. 11). Results show that the relative differences between the top-of-canopy height LiDAR and TomoSAR estimates were 15% for Nouragues and 10% for Paracou (Fig. 11 right panel).

In the Paracou forest the results from the emulated 6 MHz-bandwidth system were found to be similar with those obtained from the airborne dataset in spite of the significant resolution loss. At the
resolution of 4-ha, the RMSE was 11% (Pearson correlation of 0.79). As shown in Ho Tong Minh, Tebaldini, Rocca, Le Toan, Villard, et al. (2015), it was possible to retrieve forest top height, in which the RMSE was 2.5 m, whereas the relative difference was 10%.

4. Discussion

In this work we show that TomoSAR approaches can be used to characterize the vertical structure of tropical forests accurately, even over terrain with strong topography. The present analysis confirms the performance of the TomoSAR approach for aboveground biomass mapping in the tropics. AGB average relative errors were 15% at a 1-ha resolution, for both Nouragues and Paracou. Further, we demonstrate the stability of the TomoSAR retrieval method for different forest areas. Finally, we showed that canopy height retrieval may be performed efficiently even in tropical forests on hilly terrain. Forest top height RMSE was estimated to be 2.5 m and 2 m for Nouragues and Paracou, respectively. Together these results considerably reinforce the proposal that BIOMASS, during its tomographic phase, will be able to provide highly accurate wall-to-wall AGB mapping even in high carbon stock forests worldwide.

First, we showed that the same analysis conducted originally at a coastal tropical forest site of French Guiana, Paracou, could be replicated at another site (Nouragues), some 100 km away, and with independent ground data. This was expected to be challenging because the Nouragues area has a considerably more undulating terrain than Paracou, and this terrain is more typical of the Guiana Shield. Our study confirms that P-band SAR tomographic data can retrieve AGB even on this terrain. This is reassuring given that many of the remaining mature tropical forests today are on steep slopes, inappropriate for cultivation (see Table S4 in Réjou-Méchain et al. (2014)).

In this paper, we also investigated whether our TomoSAR approach can be generalizable to other sites than the study site originally studied (Paracou), an important issue for the BIOMASS mission. The relationship between AGB and TomoSAR data at Nouragues was found to be highly similar to the one observed in Paracou. In particular, we found that the best correlations hold in the upper layer (e.g., 30 m), whereas the ground and middle layers were poorly correlated to AGB. AGB retrieval using training plots from Nouragues and validation plots from Paracou, and vice versa, resulted in a RMSE of 16–18% using 1-ha plots, for AGB ranging from 200 to 600 t/ha. This is a key result of this paper as it shows that the TomoSAR based biomass retrieval method is generalizable to other study sites at least to those forests with similar physiognomy, i.e. with canopy height ranging from 20 to 40 m. Hence, we provide support to the possibility to transfer training samples from

Fig. 7. Sensitivity of HV backscatter at different layers produced by TomoSAR to above-ground biomass. The top left panel is the HV backscatter associated with the original SAR image. \( r_p \) is the Pearson correlation coefficient. Slope is referred to the angular coefficient of the resulting linear fit.

Fig. 8. Comparison between in-situ AGB and AGB derived from inversion of the P-band HV 30 m layer, for both Paracou and Nouragues. The RMSE in retrieved AGB is 15.3% using 1-ha plots.
one site to another, even if further studies should be conducted in other forests to assess the generality of our approach.

As previously discussed in Ho Tong Minh, Le Toan, et al. (2014), the physical interpretation of these results is as follows. The correlation between the backscatter and AGB was very weak for the ground layer. Scatterers are indeed likely to be dispersed in the ground layer because dominant scattering mechanisms are mostly influenced by local topographical or soil moisture variation. The relationship even tends to be negative, most probably because the signal extinction at the ground level is likely to be higher in the presence of tall trees, and hence high AGB. In the 15-m layer, the correlation between backscatters and AGB was also weak. One possible explanation is that almost all trees from the stand may be represented in a rather similar way across sites in the 15-m layer. In recent studies Stegen et al. (2011) and Slik et al. (2013), showed that only the largest trees (> 70 cm of diameter) drive the difference in AGB among sites and that smaller trees conveys no information on cross-sites differences in AGB. This may explain why the backscatter exhibited a strong significant correlation with AGB in upper layers (20 m layer and higher), where the influence of large trees on backscatters prevails. Further, TomoSAR processing removes the ground contributions in the upper layers, minimizing the perturbing effects (e.g. local topography and/or soil moisture) associated with ground backscatter and thus improving the relationship between AGB and backscatters.

We point out that the quality of our retrieval depended strongly on the availability of tomographic acquisitions. To place this result in perspective, we also used non-tomographic data (i.e. PolSAR) to infer AGB. The non-tomographic data exhibit a much lower sensitivity ($r_P = 0.37$) than the tomographic data of the 30 m layer ($r_P = 0.75$, see top left panel of Fig. 7). The non-tomographic backscatter signals are more dispersed because they integrate noise signals from the ground, that need to be corrected with elaborate techniques (e.g. Villard & Le Toan, 2015), and signals from the middle layer that convey little information on AGB.

By evaluating the vertical forest structure from tomographic profiles, forest top height can be retrieved. Using the LiDAR model as a reference, for Nouragues and Paracou, the same power loss value of $-2$ dB with respect to the phase center was used to retrieve forest height with no bias and minimum errors. The RMSE was estimated to be 2.5 m and 2 m, whereas the relative difference is 15% and 10%, for Nouragues and Paracou, respectively. This shows that the Nouragues hilly terrain is not a major limitation for the implementation of a canopy height retrieval algorithm with TomoSAR.

We note that the same power loss value cannot be straightforwardly transferred to the case of other campaigns. As shown in Tebaldini and Rocca (2012) in the frame of the BioSAR 2008 campaign, the power loss should be varied in space due to a strong variation of the vertical resolution across the scene swath.

The results obtained above have to be carefully assessed in the context of a spaceborne satellite mission. In the case of the BIOMASS mission the limited pulse bandwidth of 6 MHz needs to be taken into account (ITU-2004, 2004). This low bandwidth has a significant effect on the resolution and quality of the TomoSAR products. At the proposed
The middle panels present the results from tomography. The right panels report the relative difference, defined by $|H_{\text{tomography}} - H_{\text{LiDAR}}| / H_{\text{LiDAR}}$.

Incidence angle of 23–32 of BIOMASS, the bandwidth reduction translates into a resolution loss not only in the horizontal direction but also in the vertical direction. Despite these effects, our simulation of BIOMASS-like data suggests that the performance loss of the TomoSAR derived products is not significant. Thus, our TomoSAR approach will be directly applicable to the BIOMASS mission.

In addition to resolution effects also other effects need to be taken into account when extrapolating the results of this study to the spaceborne case. These include ionosphere disturbances and temporal decorrelation effects. However, the impact of ionosphere, i.e. Faraday Rotation, was found not to be critical to TomoSAR (Tebaldini & Iannini, 2012). BIOMASS will acquire fully polarimetric data, therefore allowing estimation of Faraday Rotation to within an accuracy that will ensure a negligible impact on TomoSAR results. The impact of temporal decorrelation is under analysis in the frame of the TropiSAR campaign activities (Ho Tong Minh et al., 2013; Ho Tong Minh, Tebaldini et al., 2014). Temporal depends heavily on the repeat interval, which in the tomographic phase of the BIOMASS mission has been minimized to 3–4 days. The first attempt is provided in Ho Tong Minh, Tebaldini, Rocca, Le Toan (2015), in which the resulting tomograms and forest heights were observed to change acceptably as long as the revisit time is 4 days or less.

To conclude, our results reinforce the science basis for the BIOMASS spaceborne mission. TomoSAR appears to be a promising technique to be used by BIOMASS for the retrieval of tropical forest biomass and height, and for the development of a training/validation strategy during the BIOMASS interferometric phase.

Acknowledgments

We thank the TropiSAR team for providing the TropiSAR datasets of excellent quality. We are grateful to the people and institutes that have contributed to the field data, including L Blanc and B Hérault for Paracou, P Gaucher, P Châtelet, E Courtois, S Fauset, A Montague, H Richard and B Tymen for Nouragues, and G Vincent for tree height data in Paracou. Both Paracou and Nouragues are part of the Guayaf network. We gratefully acknowledge financial support from ESA (TropiSAR and TropiSAR project), CNES (TOSCA program), from ‘Investissement d’Avenir’ grants managed by Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01; TULIP: ANR-10-LABX-0041; ANAEE-Services: ANR-11-INBS-0001) and from NERC (‘AMAZONICA’ consortium) and the Gordon and Betty Moore Foundation for contributing funding for field measurements at Nouragues through the RAINFOR project (www.rainfor.org).

References


Fig. 11. Comparison between LiDAR and tomography retrieval of forest top height in both sites. (a) Nouragues. (b) Paracou. The left panels show LiDAR height $H_{\text{LiDAR}}$, available, see in Fig. 1. The middle panels present the results from tomography $H_{\text{tomography}}$. The right panels report the relative difference, defined by $|H_{\text{tomography}} - H_{\text{LiDAR}}| / H_{\text{LiDAR}}$. 

www.rainfor.org.