Tropical Forest Vertical Structure Characterization: From GEDI to P-Band SAR Tomography

Yen-Nhi Ngo, Yue Huang, Dinh Ho Tong Minh, Member, IEEE, Laurent Ferro-Famil, Member, IEEE, Ibrahim Fayad, and Nicolas Baghdadi

Abstract—Estimating tropical forests’ vertical structure using remote sensing is a challenge. Active sensors such as low-frequency synthetic aperture radar (SAR) operating at the P-band, with a wavelength of ∼69 cm, and light detection and ranging (LiDAR) are able to penetrate thick vegetation layers. While NASA’s Global Ecosystem Dynamics Investigation (GEDI) is collecting spaceborne LiDAR data, the ESA’s next Earth Explorer BIOMASS mission will acquire multiple acquisitions over the same areas to form 3-D images through the SAR tomography (TomoSAR) technique. Our study shows the potential value of GEDI and TomoSAR acquisitions in producing accurate estimates of forests’ vertical structure. By analyzing the airborne P-band TomoSAR, airborne LiDAR, and spaceborne GEDI LiDAR at a tropical forest site in Paracou, French Guiana, South America, we show that both GEDI and P-band TomoSAR can directly measure surface, vegetation heights, and vertical profiles with high resolution and precision. Airborne TomoSAR is of higher quality than GEDI due to better penetration properties and precision. However, GEDI vegetation height root-mean-square error (RMSE) is less than 5 m, for an average forest height value around 30 m at the Paracou site, which is similar to the expected performance of the future spaceborne BIOMASS mission. These results suggest that GEDI measurements, i.e., shots with sensitivity greater than 98%, will provide a good reference of forest structure to calibrate the BIOMASS mission algorithms.

Index Terms—Canopy height model (CHM), digital terrain model (DTM), Global Ecosystem Dynamics Investigation (GEDI), Paracou, tomography, vertical structure profile.

I. INTRODUCTION

TROPICAL forests are a key component of the global carbon cycle. Many remote sensing technologies have been used to study their forest parameters such as canopy heights or biomass at various scales using active sensors such as synthetic aperture radar (SAR) [1] and its 3-D mapping capabilities when operated in the tomographic mode [2], [3] and spaceborne light detection and ranging (LiDAR) [4]. Each technology can be associated with strengths and weaknesses for studying forest areas.

The recent spaceborne LiDAR system is the Global Ecosystem Dynamics Investigation (GEDI) LiDAR on board the International Space Station (ISS), which was launched in December 2018 with on-orbit in April 2019 [5]. The GEDI instrument is specifically optimized for retrieving vegetation vertical structure. GEDI is a full-waveform (FW)-based LiDAR system that presents a new opportunity for the observation of forest structures globally. GEDI comprises three lasers emitting 1064-nm light, at a high sampling rate of 242 Hz (in comparison, the previous ICESat-1 FW LiDAR was at 40 Hz). One of the lasers’ power is split into two, and referred to as coverage lasers, while the remaining two lasers remain at full power, and they are referred to as full power lasers. GEDI measures vertical structures with echoed waveforms digitized to a maximum of 1246 bins and a vertical resolution of 15 cm, corresponding to a maximum of 186.9 m of height ranges. At any given moment, eight tracks of data are produced, separated by ∼600 m across-track, with a footprint diameter of ∼25 m, versus ∼60 m for ICESat-1, with a distance between footprint centers of 60 m along-track [5]. In addition, the GEDI mission provides a wide array of information about canopy structure, biomass, and topography. Many studies have shown the potential of GEDI for estimating the canopy height. Fayad et al. [6], [7] assessed GEDI’s LiDAR data for the estimation of canopy height of Eucalyptus plantations in Brazil. The results showed that over low-slopped terrain the canopy height estimates were obtained with an error lower than 2 m and \( R^2 \) higher than 0.85. Over natural forests, there have been a few studies that relied on GEDI data to propose global canopy height maps [5], [8], [9].

The seventh European Space Agency Earth Explorer Core Mission BIOMASS will collect the P-band (e.g., ∼69-cm wavelength) SAR data from using a multiple baseline orbit [10]. These data can be used for SAR tomography (TomoSAR) processing, resulting in a vertical information in forested areas from space [2], [3]. The TomoSAR signal is sensitive to the contribution from the vegetation layer, and hence biomass and forest height [11]. BIOMASS has a mission requirement to provide a 200-m gridded biomass product with a standard error less than 20%. To meet this requirement, there

Manuscript received 30 June 2022; revised 9 September 2022; accepted 19 September 2022. Date of publication 22 September 2022; date of current version 7 October 2022. This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
is a strong need to have a good reference forest inventory to calibrate and train mission algorithms. In this letter, we investigate the capability of GEDI measurements for supporting the BIOMASS mission algorithms since both of them can provide forest vertical structure parameters. Specifically, we will focus on digital terrain model (DTM), canopy height model (CHM), and vertical profile metrics.

II. Method
A. Study Area—Paracou Tropical Forest

The Paracou study site is located in a lowland tropical rain forest near Sinnamary, French Guiana, South America. The forest in Paracou is characterized by thick and dense vegetation layers with more than 100 species per hectare. This is a fundamental test site for the algorithm development of the BIOMASS mission [12], [13]. The available SAR, LiDAR, and GEDI datasets are shown in Fig. 1.

The tomographic P-band SAR dataset is available from the airborne TropiSAR campaign conducted in the summer of 2009. The datasets of the TropiSAR campaign are available as an ESA archive through the EOPI portal (https://earth.esa.int/eogateway). An airborne LiDAR survey was conducted in 2008 (see [14]), covering an overlap area of $4 \times 3$ km with the SAR data.

B. GEDI Processing

The GEDI collected waveforms (Level 1) and their processing (Level 2) are provided by NASA’s Land Processes Distributed Active Archive Center (LP DAAC). The Level 1 data product, namely, the L1B data product, includes the geolocated waveforms as collected by the GEDI system. The Level 2 data product includes footprint-level elevation and canopy heights metrics (L2A) and the footprint-level canopy cover and vertical profile metrics (L2B). In this study, we extracted from Version 2 of the L1B data product the geolocated waveforms (rxwaveform) [see Fig. 1(b)], the geolocation (longitude, latitude, and elevation) of the received waveforms, and their acquisition time. The L2A data product processes the waveforms using six threshold settings’ groups (henceforth referred to as algorithms [5]). The six algorithms are defined by different signal smoothing widths and thresholds on noise to suit a wide variety of applications. For forestry applications, the choice of the best algorithm depends on the characteristics of the forests studied. The algorithm tuning parameters can impact the values of waveform metrics used for canopy heights’ retrieval. In tropical areas with dense vegetation, the ground peak in the GEDI waveforms is usually low intensity and therefore more difficult to separate from the background noise. However, the algorithm setting group number a5 has a lower waveform signal end threshold compared with the other setting groups, allowing better detection of weak ground returns. In less dense forest ecosystems, however, a low threshold can result in the interpretation of noise below the actual ground as the ground peak, leading to an overestimation of canopy heights. Over our study area, as suggested in [15], we selected the a5 algorithm which is the one that has lower forward and backward thresholds among the six GEDI algorithm setting groups in the L2A dataset. From L2A, we extracted the following variables.

1) The number of detected modes (num-detectedmodes).
2) The acquisition date and time of the shot.
3) The latitude, longitude, and elevation of the lowest mode (ELM) which represents the ground return.
4) The latitude, longitude, and elevation of the highest return (EHR) which represents the canopy top.
5) The relative height metrics (RHn) with $n$ from 0 (lowest detectable return, ground position) to 100% (highest detectable return, canopy top). RHn represents the height between the ground position and the location at $n\%$ of cumulative energy. Fig. 1(b) shows an example of the GEDI waveform and the associated RHn. Based on a
previous study [16], we selected RH100 as the GEDI dominant height estimates. GEDI-derived DTM (GEDI-DTM) and GEDI-derived CHM (GEDI-CHM) values used from the L2B dataset are the “elev-lowestmode” and “RH100” values, respectively.

6) Shot Sensitivity: The sensitivity of an acquired shot is the probability of a given shot over a given canopy cover to reach the ground. Shot sensitivity is also affected by the acquiring laser. Indeed, while the full power lasers are expected to reach the ground over forests with a tree cover up to 98%, the coverage lasers are expected to reach the ground for tree cover up to 95% [9].

Over Paracou site [see Fig. 1(a)], more than 1000 shots were acquired in the period between April 2019 and August 2021. However, not all these shots are usable due to unfavorable atmospheric conditions (e.g., clouds) that affect them. Therefore, a waveform was not investigated further if it met any of the following criteria.

1) Shots without any detected modes (num-detectmodes = 0). These shots are noisy signals.
2) Shots where the absolute difference between the ELM and the corresponding SRTM DEM is higher than 75 m (|ELM-SRM| > 75).
3) Shots where RH100 < 3 m. These shots most likely correspond to bare soil or low vegetation.

After the filtering scheme, 417 shots collocated among the GEDI, TomoSAR, and airborne LiDAR datasets were retained for further analysis [see Fig. 1(a)].

C. TomoSAR Processing

Forest vertical structure indicators are extracted from 3-D tomographic images, using reflectivity-based approaches aiming at estimating the elevation of the ground and the height of the observed trees. As one may observe on the tomographic profile displayed in Fig. 1(c), the P-band tomograms generally exhibit a local peak of reflectivity at the vertical position of the ground, caused by several scattering mechanisms, such as direct scattering from the ground surface, and more importantly double-bounce scattering between the ground and trunks [17] and the ground and the forest canopy volume [18]. Tree top height may be estimated from the tomographic profiles as the upper elevation limit at which significant reflectivity can be measured.

In this study, the polarimetric single-look complex (SLC) images are first processed through the phase calibration...
Fig. 5. Vertical profiles at different layers from both TomoSAR and GEDI.

technique proposed in [19] to compensate the residual phase deviations and motion errors which may affect the accuracy of 3-D reflectivity reconstruction. Polariometric TomoSAR covariance matrices are then estimated in the slant range domain using a $20 \times 20$ m boxcar filter and decomposed into ground and volume components, using the approach exposed in [20], which consists of a refinement of the technique originally introduced in [21], with improved convergence and stability properties. The spatial covariance matrices of both the components are focused in the vertical direction using Capon’s adaptive beamformer. The ground elevation is computed as the coordinate of the maximum of the ground vertical reflectivity profile, whereas the tree top elevation is estimated as the maximal elevation at which the volume component reflectivity is larger than a fixed percentage of its maximum value, as proposed in [22], [23], and [24]. The percentage, or equivalently the amount of dB, is a fixed value over the whole scene which ensures globally unbiased tree height estimates. One may note that this quantity could be fixed using any external unbiased tree top elevation estimates, such as those provided by GEDI. Consistency with the values retrieved with the GEDI data process used in this study is enforced by further applying a median filter over values estimated within circles with a radius of 12.5 m, centered around the sampled GEDI coordinate. The vertical tomographic profiles are also computed using the classical beamformer focusing technique, around these geo-referenced ground coordinates, to compare the ability of the P-band TomoSAR and GEDI to retrieve the structure of tropical forests.

III. RESULTS

The DTM and CHM obtained by TomoSAR and GEDI are shown in Figs. 2 and 3, respectively, as a function of the airborne LiDAR estimates. For the GEDI signals, there are two measurements which are composed of 140 coverage power points and 277 full power points, resulting in 417 GEDI total points. The root-mean-square error (RMSE) and coefficient of determination ($R^2$) are used to report the statistical comparison.

TomoSAR has the best performance for both DTM and CHM, as shown in Figs. 2(a) and 3(a). GEDI full power measurements are characterized by a better agreement with the airborne LiDAR data [see Figs. 2(d) and 3(d)]. One may observe in Figs. 2(c) and 3(c) that in the coverage power mode, the GEDI-derived ground elevation is overestimated, whereas tree height is generally underestimated. This is due to the fact that the coverage lasers have reduced penetration capabilities in comparison to the full power lasers [9].

To showcase the importance of beam sensitivity in GEDI signals, we estimated the performances from only the full power laser points and those obtained all GEDI points. In the first case, we did not find the beam sensitivity signal’s effects on the performance. In the second case, the use of the GEDI sensitivity criterion allows us to obtain a good performance for both the DTM and CHM parameters (see Fig. 4). For beam sensitivity higher than 98% (222 points), the accuracy DTM is improved from 4.47 to 3.5 m (and $R^2$ from 0.84 to 0.89), whereas the CHM error decreases from 5.6 to 4.7 m (and $R^2$ from 0.21 to 3.10). The performance of GEDI beams with sensitivity $>98\%$ is similar to full power lasers.

Finally, the vertical profiles from both TomoSAR and GEDI are reported, after compensation of the ground elevation, in Fig. 5. The tomographic vertical power distributions are represented for different polarimetric channels HH, HV, and VV, where V and H refer to the vertical and horizontal linear polarizations, respectively. The dynamics in the forest vertical structure at different layers (i.e., 20–30 m and 30–40 m) in the Paracou area can be observed from the GEDI measurements and through the tomographic behaviors in each of the polarizations. Similar to GEDI signals, significant contributions are observed at the canopy levels (i.e., 10–40 m). However, for all the polarizations, the ground layer is more visible in the tomographic power profiles, confirming a better penetration capability of the P-band SAR with respect to the GEDI signals.

IV. CONCLUSION AND DISCUSSION

In this letter, the potential value of GEDI and TomoSAR in mapping ground topography, canopy height, and vertical profile has been addressed. The Paracou study site is located on tropical forest areas, a very complex environment due to thick and dense vegetation layer. We show that both GEDI and P-band TomoSAR can be used to estimate ground
elevation, canopy height, and vertical profiles in Paracou with a high accuracy. There is a better performance on the ground elevation parameters with respect to CHM. Thus, these demonstrated results confirm the great potential of the forest vertical structure from GEDI and from the P-band TomoSAR.

For GEDI measurements, the penetration capability is due to their small wavelengths (i.e., 1064 nm) that reflect from individual elements (e.g., top of canopies) or penetrate through small gaps to reflect from the ground. The penetration is also related to the power of GEDI and the elevation of ISS [8], [9]. For the P-band TomoSAR data, thanks to its long wavelength (∼69 cm), the signals can penetrate to the ground and quantify the whole forest vertical structure (see Fig. 5). Therefore, both GEDI and P-band TomoSAR can directly measure surface, vegetation heights, and vertical profile with high resolution and precision.

We show that the P-band TomoSAR outperforms GEDI for both DTM and CHM estimation [see Figs. 2(a) and 3(a)]. This is mainly due its higher sensitivity to the forest vertical structure. However, the TomoSAR result obtained here cannot be directly extrapolated to spaceborne configurations, such as the one of the BIOMASS mission, as the involved systems have very different characteristics, mainly in terms of resolution and temporal decorrelation [2], [12]. Furthermore, tropical forest structure estimates derived from the BIOMASS may show a performance degradation with respect to the airborne configurations, even if alternative algorithms, having higher resolution at the cost of increased complexity [23], may replace the tomographic processing chain used in this study. Thus, GEDI measurements are very likely to play an important role in the training of the BIOMASS algorithms and in the determination of site-specific processing parameters with reliable priors. GEDI CHM could be used to fit these parameters at sparse coordinates, and TomoSAR estimates would provide dense and accurate canopy height maps. The present work shows that attention must be paid for selecting shots with the highest probability to detect the ground return, i.e., shots with sensitivity greater than 98% or GEDI beams with full power lasers.

It is important to point out that time acquisition is different among datasets. However, we do not observe the bias between the full power GEDI and LiDAR [see Fig. 3(d)]. Future works can be dedicated to adding the annual growth increment in the Paracou LiDAR data to better compare with GEDI CHM. For example, the growth data can be based on a model-based estimate of growth age relationships [25].

REFERENCES

[1] D. Ho Tong Minh, E. Nidkumana, G. Vieilledent, D. McKey, and N. Baghdadi, “Potential value of combining ALOS PALSAR and Landsat-derived tree cover data for forest biomass retrieval in madagascar,” Remote Sens. Environ., vol. 213, pp. 206–214, Aug. 2018.

[2] D. H. T. Minh, S. Tebaldini, F. Rocca, T. Le Toan, L. Villard, and P. C. Dubois-Fernandez, “Capabilities of BIOMASS tomography for investigating tropical forests,” IEEE Trans. Geosci. Remote Sens., vol. 53, no. 2, pp. 965–975, Feb. 2015.

[3] H. Aghababaei et al., “Forest SAR tomography: Principles and applications,” IEEE Geosci. Remote Sens. Mag., vol. 8, no. 2, pp. 30–45, Feb. 2020.

[4] L. Duncanson et al., “Aboveground biomass density models for NASA’s global ecosystem dynamics investigation (GEDI) lidar mission,” Remote Sens. Environ., vol. 270, Mar. 2022, Art. no. 112845.

[5] R. Dubayah et al., “The global ecosystem dynamics investigation: High-resolution laser ranging of the Earth’s forests and topography,” Sci. Remote Sens., vol. 1, Jun. 2020, Art. no. 100002.

[6] I. Fayad et al., “Assessment of GEDI’s LiDAR data for the estimation of canopy heights and wood volume of eucalyptus plantations in Brazil,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 14, pp. 7095–7110, 2021.

[7] I. Fayad et al., “A CNN-based approach for the estimation of canopy heights and wood volume from GEDI waveforms,” Remote Sens. Environ., vol. 265, Nov. 2021, Art. no. 112652.

[8] G. Sun, K. J. Ranson, D. S. Kimes, J. B. Blair, and K. Kovacs, “Forest vertical structure from GLAS: An evaluation using LVIS and SRTM data,” Remote Sens. Environ., vol. 112, no. 1, pp. 107–117, Jan. 2008.

[9] S. Hancock et al., “The GEDI simulator: A large-footprint waveform lidar simulator for calibration and validation of spaceborne missions,” Earth Space Sci., vol. 6, no. 2, pp. 294–310, Feb. 2019.

[10] S. Quegan et al., “The European Space Agency BIOMASS mission: Measuring forest above-ground biomass from space,” Remote Sens. Environ., vol. 227, pp. 44–60, Jun. 2019.

[11] D. H. T. Minh et al., “SAR tomography for the retrieval of forest biomass and height: Cross-validation at two tropical forest sites in French Guiana,” Remote Sens. Environ., vol. 175, pp. 138–147, Mar. 2016.

[12] D. H. T. Minh et al., “Vertical structure of P-band temporal decorrelation at the paracou forest: Results from tropiscat,” IEEE Geosci. Remote Sens. Lett., vol. 11, no. 8, pp. 1438–1442, Aug. 2014, Art. no. 6708421.

[13] D. H. T. Minh, T. Le Toan, F. Rocca, S. Tebaldini, M. M. D’Alessandro, and L. Villard, “Relating P-band synthetic aperture radar tomography to tropical forest biomass,” IEEE Trans. Geosci. Remote Sens., vol. 52, no. 2, pp. 967–979, Feb. 2014.

[14] G. Vincent et al., “Accuracy of small footprint airborne LiDAR in its predictions of tropical moist forest stand structure,” Remote Sens. Environ., vol. 125, pp. 23–33, Oct. 2012.

[15] I. Fayad, N. Baghdadi, and K. Lahmim, “An assessment of the GEDI lasers’ capabilities in detecting canopy Tops and their penetration in a densely vegetated, tropical area,” Remote Sens., vol. 14, no. 13, p. 2969, 2022.

[16] M. Adam, M. Urbazaev, C. Dubois, and C. Schmullius, “Accuracy assessment of GEDI terrain elevation and canopy height estimates in European temperate forests: Influence of environmental and acquisition parameters,” Remote Sens., vol. 12, no. 25, p. 3948, Dec. 2020.

[17] M. M. D’Alessandro, S. Tebaldini, and F. Rocca, “Phenomenology of ground scattering in a tropical forest through polarimetric synthetic aperture radar tomography,” IEEE Trans. Geosci. Remote Sens., vol. 51, no. 8, pp. 4430–4437, Aug. 2012.

[18] R. Abdo, L. Ferro-Famil, F. Boutet, and S. Allain-Bailhache, “Analysis of the double-bounce interaction between a random volume and an underlying ground, using a controlled high-resolution PolTomoSAR experiment,” Remote Sens., vol. 13, no. 4, p. 636, Feb. 2021.

[19] S. Tebaldini, F. Rocca, M. M. D’Alessandro, and L. Ferro-Famil, “Phase calibration of airborne tomographic SAR data via phase center double localization,” IEEE Trans. Geosci. Remote Sens., vol. 54, no. 3, pp. 1775–1792, Mar. 2016.

[20] L. Ferro-Famil and S. Tebaldini, “ML tomography based on the MB RVG model: Optimal estimation of a covariance matrix as a sum of two Kronecker products,” in Proc. POLinSAR Workshop, Frascati, Italy, Jan. 2013, p. 147.

[21] S. Tebaldini, “Algebraic synthesis of forest scenarios from multibaseline PolInSAR data,” IEEE Trans. Geosci. Remote Sens., vol. 47, no. 12, pp. 4132–4142, Dec. 2009.

[22] S. Tebaldini and F. Rocca, “Multibaseline polarimetric SAR tomography of a boreal forest at P- and L-bands,” IEEE Trans. Geosci. Remote Sens., vol. 50, no. 1, pp. 232–246, Jan. 2012.

[23] Y. Huang, Q. Zhang, and L. Ferro-Famil, “Forest height estimation using a single-pass airborne L-band polarimetric and interferometric SAR system and tomographic techniques,” Remote Sens., vol. 13, no. 3, p. 487, 2021.

[24] P. Q. Yang, S. Tebaldini, M. M. d Alessandro, and M. Liao, “Tropical forest height retrieval based on P-band multibaseline SAR data,” IEEE Geosci. Remote Sens. Lett., vol. 17, no. 3, pp. 451–455, Mar. 2020.

[25] D. Requena Suarez et al., “Estimating aboveground net biomass change for tropical and sub-tropical forests: Refinement of IPCC default rates using forest plot data,” Global Change Biol., vol. 25, no. 11, pp. 3669–3684, Nov. 2019.