Improving Forest Height-to-Biomass Allometry with Structure Information: A TanDEM-X Study

Changhyun Choi, Matteo Pardini, Member, IEEE, Michael Heym, and Konstantinos P. Papathanassiou, Fellow, IEEE

Abstract—Allometric relations that link forest above ground biomass to top forest (i.e. canopy) height are of particular significance in the context of lidar and interferometric Synthetic Aperture Radar (SAR) remote sensing, as both techniques allow accurate height measurements at ecologically relevant spatial scales. Besides the often unknown allometry itself, its spatial variation in heterogenous forest environments restricts the performance when using a single fixed height-to-biomass allometric relation.

This paper addresses how forest structure information derived from interferometric TanDEM-X data can be used to locally adapt the height-to-biomass allometry in heterogeneous forests, and to improve biomass estimation performance. The analysis is carried out using TanDEM-X interferometric measurements in three tropical forest test sites in Gabon. A structure index expressing forest density is derived from the TanDEM-X data. Then, a continuous relationship between the structure index and the allometric level that defines the forest height-to-biomass allometry is reconstructed from the available lidar data, and used to vary the height-to-biomass relationship. Finally, the potential of the derived structure index to support an allometric relationship common to all sites is evaluated. The experimental results show the appropriateness of TanDEM-X data for characterizing structure and in this way improving the biomass estimation performance.

Index Terms — Above-ground biomass, forest height, forest structure, synthetic aperture radar (SAR), interferometry, TanDEM-X.

I. INTRODUCTION

BIOMASS has a direct relationship to carbon content and is a measure of forest and ecosystem productivity. Estimation of biomass is very inaccurate at local, regional and supraregional scales. Ground measurements of biomass in natural forests often exhibit errors much greater than 20%. Particularly large are the deviations in tropical and natural forests due to their spatial heterogeneity. Dynamic changes of biomass and their spatial distribution are a direct measure of the exchange of carbon between the terrestrial ecosystem and the atmosphere [1]. At the same time, they characterize the variation of forest growth and productivity induced by water or climate stress [2].

The Above Ground Biomass (AGB) $B_T$ of a single tree can be expressed as the product of the tree volume $V$ with its (species specific) wood density $\rho$ [3], [4]. Different adequate standards of tree volume can be chosen depending on the individual application. In forestry, the focus on the dominant stem contribution suggests the use of stem volume that leads to

$$B_T = V \cdot \rho = F \cdot \left(\frac{\pi D^2}{4}\right) \cdot H \cdot \rho$$

where $D$ is the stem diameter at breast height (also known as DBH), $H$ is the tree height, and $F$ is a factor that accounts for the shape of the stem [3], [4]. Accounting for the total tree volume including its leaves and branches is more common for ecological applications and carbon stock estimates.

However, the use of (1) in the context of remote sensing is rather limited to high resolution airborne implementations because the spatial resolution of conventional and especially spaceborne configurations does not allow to measure single tree parameters. With a typical spatial resolution on the order of few to ten meters, such configurations can provide forest parameter estimates at some tenths of meters, representative of patches of trees or forest stands rather than of single trees. However, the transition of (1) from a single tree to a stand biomass relation by replacing the individual tree parameters with the stand means (e.g. mean forest height, mean diameter or alternatively basal area, and mean wood density) is not straightforward and strongly depends on the stand density and heterogeneity in terms of species and age composition [5], [6].

The potential of either spaceborne lidar or interferometric Synthetic Aperture Radar (SAR) configurations to measure forest height at spatial scales of 1 hectare (ha) or even below motivated the use of the so-called forest height-to-biomass allometry at stand level. The AGB $B$ is expressed in terms of (top) height by means of an exponential relationship [4]-[8]:

$$B = a_0 \cdot H^{\beta_0}.$$  (2)

$H$ is the maximum height within the stand area equivalent to the top canopy height, and $(a_0, \beta_0)$, are the so-called allometric factor and reference (allometric) exponent, respectively. The reference exponent $\beta_0$ defines the underlying allometric relation accounting for species composition and growing conditions of a certain stand. The allometric factor accounts

C. Choi is with the German Aerospace Center (DLR), Microwaves and Radar Institute, 82234 Wessling, Germany, and also with the Institute of Environmental Engineering, Swiss Federal Institute of Technology (ETH) Zurich, 8093 Zurich, Switzerland.

M. Pardini, and K. Papathanassiou are with the German Aerospace Center (DLR), Microwaves and Radar Institute, 82234 Wessling, Germany.

M. Heym is with the Bavarian State Institute of Forestry (Bayerische Landesanstalt für Wald und Forstwirtschaft), 85354 Freising, Germany.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
then for (anthropogenic and natural) density variations across stands of the same composition. When the reference exponent $\beta_0$ is fixed, $\alpha_0$ is also known as the allometric level as it scales a set of height-to-biomass relations with the same reference exponent accounting for different stand densities (e.g. different basal areas at a given age) or stand ages. A successful implementation requires the two allometric parameters to be either a-priori known or estimated from reference (for example inventory) measurements. Indeed, allometric relationships in form of (2) have been successfully used to derive biomass estimates from height estimates [6]-[11], however for rather homogeneous stand and forest conditions.

However, the spatial variability in stand density and/or structure reduces the performance of a single height-to-biomass relation with fixed $\alpha_0$ and $\beta_0$. Accordingly, in spatially heterogeneous (in terms of density and structure) forests a fixed height-to-biomass allometry leads to a poor performance or even breaks down [4]-[6], [9]. To visualize this insufficiency three stands with the same $H$ are depicted in Fig. 1. The tree density decreases from left to right with the stand biomass to decrease accordingly. In this particular case, if only height is used and $\alpha_0$, $\beta_0$ are fixed, the same biomass is estimated across the three stands. Using allometric parameters $(\alpha_0, \beta_0)$ fitting the denser stand will overestimate the biomass in the sparser stand, and vice versa.

Summarizing, two main factors limit the use of the height-to-biomass allometry in a wide context: (i) the large uncertainty in the knowledge of the allometric parameters $(\alpha_0,\beta_0)$ for the individual forest conditions (arising from the insufficiency or complete lack of appropriate reference measurements), and (ii) the inability to adapt $(\alpha_0,\beta_0)$ to the spatial variability of forests.

However, today both limitations appear less restrictive making a revisit of the height-to-biomass allometry attractive. Terrestrial lidar scanning (TLS) techniques have the potential to make plot inventory measurements more accurate and faster [12]. At the same time, spaceborne waveform lidar configurations sample forest height in a more or less dense grid, and provide a set of waveform metrics that allow to estimate AGB (using empirically derived models) [13], [14]. TLS and/or spaceborne waveform lidars are able to provide enough forest height and biomass measurements to define a general height-to-biomass allometry at regional or even finer scales. Hence, they resolve, at least to a large extent, the first limitation. On the other hand, the spatial variability of allometry can be accounted with remote sensing configurations able to estimate not only forest height, but also forest structure information with a sub- hectare spatial resolution. Indeed, SAR Interferometry [15]-[21] and especially SAR Tomography have proven to be able to characterize physical forest structure, the latter relying on the reconstruction of the 3D radar reflectivity [22]-[25].

The open question is how far the knowledge of forest structure can be used to adapt (and improve) the general height-to-biomass allometry to local scales. This paper addresses this question in the context of TanDEM-X (i.e. bistatic interferometric acquisitions at X-band) and waveform lidar measurements. Height and biomass estimates from the waveform lidar measurements are used to establish a general height-to-biomass allometry. Then, the horizontal structure index $HS$ as defined in [24] is reconstructed from TanDEM-X measurements and used to account for the spatial variability (by means of the allometric level) of the height-to-biomass allometry within the test site:

$$B = \alpha(HS) \cdot H^{\beta_0}.$$ (3)

There are two arguments for using $HS$ to adapt $\alpha$: first, its relation to the well-established Stand Density Index [26] (and thus to basal area) as discussed and demonstrated in [24], and second, the ability to obtain $HS$ estimates from TanDEM-X data.

Accordingly, in the following sections the estimation of $H$ and $HS$ from TanDEM-X data, the performance of conventional height-to-biomass allometry and the use of $HS$ to improve the height-to-biomass allometry performance for different tropical forest types and conditions are discussed.

II. TEST SITES AND DATA SETS

A. Test sites

The AfriSAR campaign was carried out over tropical forest sites in Gabon in 2015 and 2016 [27], [28]. The objective of the campaign was to acquire air- and space-borne polarimetric SAR interferometric and tomographic data sets complemented by
airborne waveform lidar [28] and field measurements [29] for the development and validation of forest height, structure and biomass estimation algorithms. Three of the AfriSAR sites are considered in this study: Lopé, Mabounie, and Mondah. Their locations are shown in Fig. 2.

The Lopé site is located within the Lopé National Park near the geographic center of Gabon. The site consists of a variety of structure types ranging from open savannas to undisturbed tall (sometimes exceeding 50 m) dense forest stands. Colonizing forest (sparse forest stands mixed up with savanna) or monodominant Okoume (dense, mono-layered, tall and dense forest stands) are two particular cases [29]-[31]. Biomass ranges between 10 t/ha in savanna areas and 600 t/ha in dense forest areas. The terrain is hilly with many local slopes steeper than 20°.

The Mabounie site is part of the “Maboumine” mining project started in 2005. The site is covered by mature stands with canopy height between 40 to 60 m and biomass levels up to 400 t/ha, but is signed by local degradation caused by roads, buildings, and other infrastructure. The site includes also partially flooded areas containing swamp mixed forest [28].

Finally, Mondah is a partially flooded area containing mangrove and mahogany woodlands in northwest Gabon. The western part is a primary forest taller than 50 m, with dense and homogenous stands while the eastern part is secondary forest [28]. The topography is fairly flat within the site.

**B. Lidar acquisitions**

Waveform lidar data were acquired by NASA’s Land and Vegetation and Ice Sensor (LVIS) in February and March 2016 [28], [32], [33]. During AfriSAR, LVIS was operated at a nominal flight altitude of 24000 ft (equivalent to 7315 m), and acquired data with partially overlapping footprints on ground with an average diameter around 22 m [28].

For each test site, the digital terrain model (DTM) and relative height (RH) metrics were derived from the waveforms [33]. Each RH metric expresses the height above ground at which a certain percentage of the total waveform energy is received [14]. For instance, the RH100 represents to the height above the ground in which the total energy is received. RH metrics have been used to estimate biomass and to describe the canopy vertical structure [13], [14].

In the following the RH100 is considered as the top canopy height within a footprint, and is used as a reference for the validation of the heights estimated from TanDEM-X coherences. Consistently with the formulation in Section I, the LVIS top height $H_{LVIS}$ has been calculated by taking the maximum RH100 for every 1 ha resolution cell on ground.

Similarly, the waveform-derived AGB estimates $B_{LVIS}$ and their uncertainty at 1 ha resolution, estimated by means of an allometric relationship linking height and RH metrics to the LVIS AGB and parameterized by using the available field inventory plots, have been used as reference [28], [34]. An overall root mean square error around 70 t/ha across all
inventory plots was documented [28], [34]. LVIS RH and biomass data sets are ideal for the experiments in this paper as they cover large areas allowing to appreciate structure gradients, and provide a statistically (very) large number of samples for comparison at the desired 1 ha resolution.

Fig. 3 shows the maps of $H_{LVIS}$ and $B_{LVIS}$ for the three selected sites. The largest continuous LVIS coverage is over Lopé. $H_{LVIS}$ is larger than 50 m in the northwestern part and lower in the southeastern part, but overall without relevant spatial gradients. At the same time, $B_{LVIS}$ varies significantly stronger suggesting a likewise significant structure dependency. In Mabouinée, many footprints are lost because of cloud coverage. In Mondah, the short stands ($H_{LVIS} < 30$ m) in the central part of the scene have a low biomass ($B_{LVIS} < 100$ t/ha). In contrast, in the western part $B_{LVIS}$ can reach up to 700 t/ha for the taller primary stands ($H_{LVIS}$ around 50 m).

### C. TanDEM-X acquisitions

Single-polarimetric stripmap bistatic TanDEM-X data were acquired close in time to the LVIS flights. The relevant acquisition parameters are summarized in Table I. For every data set the interferometric (InSAR) complex coherence has been calculated as:

$$
\gamma(\kappa_z) = \frac{<S_1 \cdot S_2^*>}{\sqrt{<S_1 \cdot S_1^*> <S_2 \cdot S_2^*>}}
$$

(4)

where $S_1$ and $S_2$ are single-look complex amplitudes of the two images, and $(\cdot)^*$ and $<>$ denote the complex conjugation and the spatial averaging operators. For forest height estimation, the spatial averaging has been performed using cells measuring 25 m × 25 m (corresponding to ~150 independent looks) in ground range – azimuth, respectively. For the estimation of the horizontal structure index (see Section IV.B) a finer resolution of 10 m × 10 m (corresponding to 25 independent looks) has been chosen. Finally, biomass, horizontal structure index and top height have been calculated at the scale of 1 ha (100 m × 100 m), consistent to the reference lidar data.

In (4), $\kappa_z$ represents the vertical wavenumber, which expresses the sensitivity (i.e., the derivative) of the InSAR phase difference with respect to (vertical) height [35], [36]. For the bistatic mode:

$$
\kappa_z = \frac{2\pi}{\text{HoA}} \approx \frac{2\pi}{\lambda} \frac{B}{R \sin \theta_i}
$$

(5)

where HoA is the InSAR Height of Ambiguity, $B$ is the perpendicular InSAR baseline, $\lambda$ is the radar wavelength, and $R$ is the slant range distance. $\theta_i$ is the slope-corrected incidence angle, which corresponds to the difference between the incidence angle and the local terrain slope in range direction. The terrain slopes have been calculated using the TanDEM-X digital elevation model (DEM). The vertical wavenumber characterizes also the performance of parameter inversion from InSAR coherence measurements. In the case of forest height estimation, one single $\kappa_z$ allows unbiased and accurate inversion only for a limited range of forest heights [36]. In Lopé a mean $\kappa_z$ of 0.1 rad/m leads to an optimum performance range of 15 m to 40 m appropriate to cover most of the forest height in the scene, however with some (significant) loss of performance for the tallest stands with height close to the HoA (around 62 m), far above the optimum performance range. For Mabouinée and Mondah, the $\kappa_z$ values are lower, increasing further the HoA and shifting the optimum performance range to larger heights.

For bistatic acquisitions, the absence of temporal decorrelation leaves signal-to-noise ratio (SNR) decorrelation $\gamma_{SNR}$, range (spectral) decorrelation $\gamma_{rg}(\kappa_z)$, quantization decorrelation $\gamma_q$ and volume decorrelation $\gamma_{vol}(\kappa_z)$ as the remaining decorrelation contributions [35]:

$$
\gamma(\kappa_z) = \gamma_{SNR} \cdot \gamma_{rg}(\kappa_z) \cdot \gamma_q \cdot \gamma_{vol}(\kappa_z)
$$

(6)

Height is estimated from $\gamma_{vol}(\kappa_z)$ obtained from $\gamma(\kappa_z)$ after compensating all the other contributions. $\gamma_{SNR}$ and $\gamma_{rg}(\kappa_z)$ have been calculated and compensated using the procedure described in [35]. A fixed value of 0.97 was assumed for $\gamma_q$ according to the analysis in [37].

### III. FOREST BIOMASS ESTIMATION FROM TANDEM-X HEIGHT

#### A. Forest height inversion

After coherence calibration (see Section II.C), the vertical reflectivity function (i.e. the vertical distribution of scatterers) underlying $\gamma_{vol}(\kappa_z)$ is usually represented by a two-layer model accounting for the ground and volume (back-) scattering contributions. However, at X-band and for dense(r) forest conditions, the ground contribution is often neglected so that [35], [36], [38]:

$$
\gamma_{vol}(\kappa_z) = \exp(i \phi_0) \cdot \gamma_v(\kappa_z).
$$

(7)

The phase term $\phi_0 = \kappa_z z_0$ is the InSAR phase corresponding to the ground height $z_0$ and,

$$
\gamma_v(\kappa_z) = \frac{\int_{z_0}^{h_f} f_v(z) \exp(i \kappa_z z) \, dz}{\int_{z_0}^{h_f} f_v(z) \, dz}
$$

(8)
where $f_V(z)$ is the volume-only vertical reflectivity function and $h_V$ is the top volume (i.e. forest) height with respect to the ground topography. For $f_V(z)$ an exponential distribution of scatterers is widely used [35, 36, 38]:

$$f_V(z) = \exp\left[\frac{2\sigma}{\cos(\theta)}\right]$$  \hspace{1cm} (9)

where $\sigma$ is a coefficient defining the shape of the reflectivity function, interpreted as a mean extinction value.

The inversion of (7) with (8) and (9) using a single baseline is an underdetermined problem as the number of unknowns ($\phi_0, h_V, \sigma$) exceeds the number of measurements ($\gamma_{vol}(\kappa)$). In this sense the use of an external DTM for calculating $\phi_0$ has two advantages. First, it allows a balanced inversion. Second, it allows an unbiased estimation of $h_V$ even if there is no penetration until the ground [35]. On the other hand, any offset (e.g. arising from residual TanDEM-X orbit errors) between $\phi_0$ and $\gamma_{vol}(\kappa)$ on surfaces must be compensated before the inversion [35].

The desired $h_V$ is then obtained as the solution of [35]:

$$\min_{h_V, \sigma} \| \gamma_{vol}(\kappa) - \exp(i\phi_0) \cdot \gamma_V(\kappa; h_V, \sigma) \|.$$  \hspace{1cm} (10)

Finally, similarly to the calculation of $H_{LVIS}$, the top height $H_{TDX}$ is obtained by taking the maximum $h_V$ for every 1 ha resolution cell on ground.

The inversion (10) was implemented and applied over the three sites using the LVIS DTMs to calculate $\phi_0$. The inversion was carried out only for $|\gamma_{vol}(\kappa)| > 0.25$ with $\kappa < 0.12$ rad/m ($HoA > 55$ m) corresponding to 85% of the forested areas in Lopé, 90% in Mabouine and 100% in Mondah. The obtained maps of $H_{TDX}$ and the 2D histograms validating $H_{TDX}$ against $H_{LVIS}$ are shown in Fig. 4. The overall RMSE amounts to 3.5 m, 3.5 m and 4.1 m, and the bias amounts to -1.9 m, -0.1 m and -2.9 m in Lopé, Mabouine and Mondah, respectively. For all the sites, $H_{TDX}$ covers the same height range of $H_{LVIS}$, and the two height maps are well correlated. The lower coherence level in taller stands especially in Lopé and Mabouine increase the standard deviation of $H_{TDX}$. $H_{TDX}$ is clearly underestimated (i.e. has a negative bias) for $H_{LVIS} < 30$ m in Lopé and even more in Mondah probably as a result of a more open canopy areas (e.g. colonizing forest) [7, 35, 36]. In Lopé, heights are slightly underestimated also for $H_{LVIS} > 50$ m with a bias around -3 m. This is the effect of the relatively high $\kappa_z$ that limits the sensitivity of the inversion for taller stands.

B. AGB estimation using height

Here the estimation of AGB by means of the height-to-biomass allometry is discussed. Using the LVIS height and biomass estimates the allometric level $\alpha_0$ and the reference exponent $\beta_0$ of a height-to-biomass allometric relation with the form of (2)
have been estimated by means of an ordinary least squares regression:

$$\min_{\alpha, \beta} \left\| b_{LVIS} - \alpha_h h_{LVIS}^{\beta_h} \right\|^2,$$

(11)

where $b_{LVIS}$ is a vector containing the biomass values $B_{LVIS}$ and $h_{LVIS}$ contains the associated $H_{LVIS}$ values. For each site, the regression has been performed individually accounting for the fact that certain height ranges may be insufficiently represented leading to $\alpha_h = 0.473, 1.25, 3.4 \times 10^{-5}$ and $\beta_h = 1.72, 1.44, 4.73$ for Lopé, Mabounie and Mondah, respectively. While the allometric exponents in Lopé and Mabounie are similar, both of them differ significantly from the one in Mondah. This probably reflects the large structural heterogeneity and presence of secondary and disturbed (open) forest stands in Mondah in contrast to the other two test sites.

After the definition of the height-to-biomass allometry, forest height estimates, either $H_{TDX}$ or $H_{LVIS}$ can be transformed to AGB estimates. The obtained AGB maps at 1 ha resolution are shown in Fig. 5. For all three sites the estimated AGB range is smaller than the LVIS AGB range: high AGB levels are consistently underestimated (by 100 t/ha or even more) especially in Lopé and Mondah. In contrast, low AGB levels tend to be overestimated, especially in Lopé and Mabounie. This behavior is common whether TanDEM-X or LVIS heights are used.

Fig. 5. (a) AGB maps (1 ha resolution) estimated from TanDEM-X top heights ($H_{TDX}$), and 2D validation histograms of the biomass estimates against the reference biomass using (b) LVIS top heights ($H_{LVIS}$) and (c) TanDEM-X top heights ($H_{TDX}$) for Lopé, Mabounie, and Mondah. The overall estimation RMSE is reported.
The insufficiency of a single allometric relation to describe accurately the height to biomass relationship even within a single site becomes apparent in the 2D histograms of $B_{\text{LVIS}}$ against $H_{\text{LVIS}}$ plotted for each site in Fig. 6. The middle one of the three dotted lines indicates the height-to-biomass allometry as obtained from (11). In Lopé the AGB ranges from 350 up to 500 t/ha at a height of 50 m. A single height-to-biomass relationship cannot describe this spread. Two additional forest height-to-biomass allometries are plotted defined by the same reference allometric exponent and an allometric level increased (or decreased) by 30% with respect to the original one. It becomes clear that a smaller $a_0$ is able to compensate the overestimation of lower biomass levels (e.g. for sparse(r) stands) seen in Fig. 5. A larger $a_0$ fits better the higher AGB levels (e.g. for dense(r) stands) compensating their underestimation seen in Fig. 5. In conclusion, a single value of $a_0$ is not sufficient for an accurate biomass estimation. To account for this the use of a variable allometric level is attempted next.

### IV. FOREST BIOMASS ESTIMATION FROM TANDEM-X

#### A. Derivation of HS from TanDEM-X

The basic idea behind the structure estimation framework proposed in [23], [24] is to quantify the variability in height of the top canopy “surface” from tomographically reconstructed 3D reflectivity. In the case of TanDEM-X, where usually only one single-pass interferometric acquisition is available (with an appropriate vertical wavenumber), the reconstruction of a vertical reflectivity profile is not possible; at least not in a conventional tomographic way. However, the histogram of the interferometric phases (or alternatively of the converted phase center height) over a large enough area provides in many cases an approximation of the vertical reflectivity profiles [39]-[41]. Such profiles will be referred in the following as canopy height profiles (CHP). Note that because of the different resolutions, the CHP can deviate significantly from the vertical reflectivity underlying the interferometric coherence. Nevertheless, the high attenuation at X-band combined with the high spatial resolution of the TanDEM-X interferograms support the correlation of phase center height variation to (top) canopy height variation allowing the use of the derived CHP to extract relevant horizontal structure information [41].

Interferometric coherences have been estimated with about 25 looks on a 10 m × 10 m ground range-azimuth cell (see Section II). The LVIS DTM is converted to phase (by multiplying with the local vertical wavenumber) and subtracted from the interferometric phase in order to compensate any terrain-induced phase center variations. The obtained phase is then converted to phase center height (by dividing by the local vertical wavenumber). In this way, only the height variations induced by the canopy variability are relevant. Finally, vertical scattering profiles have been obtained within a 25 m × 25 m cell. An example of such CHPs along a 1 km transect across the Lopé site is shown in Fig. 7 and compared with the corresponding LVIS waveforms along the same transect. The peaks of the CHP are distributed closer to the canopy top, while the waveform peaks are distributed much wider between canopy top and the ground. Despite this difference, it is apparent that just based on the top canopy variability, the CHPs allow to distinguish a sparser area (marked as “Area 1”) from a denser one (marked as “Area 2”).

With reference to Fig. 8, the calculation of the horizontal structure index $HS$ proposed in [24] is based on “counting” the number of profile (i.e. reflectivity) peaks (associated to scattering contributions) located within a predefined top layer in a structure resolution cell. In practice, considering only the top layer allows to separate the top canopy height variations from the (vertical) variations induced by volume scattering contributions [24]. As each CHP typically contributes with no more than a couple of peaks, the larger the number of peaks close to the canopy top, the more homogeneous is the forest height, and the higher is the local forest density. Referring the height of the highest peak in the structure resolution cell with $H_p$, the top height layer is defined as the fraction $\Delta \cdot H_p$ below $H_p$. The horizontal structure index is then calculated as [24]:

$$HS = 1 - \frac{N_p}{N_{p,max}}$$  \hspace{1cm} (12)$$

where $N_p$ is the number of CHP peaks in the top layer within...
the structure resolution cell, and \( N_{P_{\text{max}}} \) is a reference maximum often assumed as the maximum \( N_P \) within the site / scene. \( HS \) equals 0 in dense structure resolution cells, and 1 in sparse structure resolution cells corresponding to the cases illustrated in Fig. 8. The obtained \( HS \) value depends on the top layer extent: wider top layers contain more CHP peaks than thinner ones. As a consequence, the choice of an appropriate \( \Delta \) is crucial for the ability of the index to reflect the physical structure (density) and its local variation.

Horizontal structure indices have been derived from the TanDEM-X CHPs and the LVIS waveforms (using the same waveform maxima counting procedure and same top layer as with the CHPs) denoted as \( HS_{\text{TDX}} \) and \( HS_{\text{LVIS}} \) respectively. The CHPs have been additionally “multi-looked” in height by means of a 10 m Gaussian-shaped moving window in order to suppress peaks caused by the (interferometric phase) noise. A (sliding) 1 ha structure resolution cell has been used to aggregate a statistically relevant number of CHP’s (or waveforms). Each structure resolution cell then contains 16 (4 × 4) profiles corresponding to non-overlapping areas on ground. A top layer extent \( \Delta = 0.35 \) has been chosen. This choice is rather empirical, although motivated by the analysis in [24].

However, both indices appear widely robust against the choice of the top layer extent: \( HS_{\text{TDX}} \) and \( HS_{\text{LVIS}} \) does not change significantly for \( \Delta \) changing within 0.3 and 0.45, in accordance with the results reported in [41]. The derived \( HS_{\text{TDX}} \) and \( HS_{\text{LVIS}} \) maps for the three test sites are shown in Fig. 9 and reveal very similar structure patterns. In Lopé both indices distinguish the denser southeastern part from the sparser colonizing forest stands at the border to the savannah [30]. In Mabounie the mature forest stands are characterized as dense by both indices. Finally, Mondah’s disturbed open forest stands in the central part of the scene (indicated as sparse) are clearly distinguished from the older tall stands in the western part (indicated as dense) by both indices.

Fig. 10 shows the 2D scatterplots of \( B_{\text{LVIS}} \) against \( H_{\text{LVIS}} \) for each site. The color of each point corresponds to its \( H_{\text{LVIS}} \) value. Despite the dispersion, the dependency of the height-to-biomass allometry on the horizontal structure index \( HS_{\text{LVIS}} \) becomes clearly visible: for a given height level \( B_{\text{LVIS}} \) increases with decreasing \( HS_{\text{LVIS}} \) (i.e., with increasing forest density) or decreases with increasing \( HS_{\text{LVIS}} \) (i.e., with decreasing forest density). Interestingly, tall stands with top heights of about 40 m appear to have very low biomass levels (< 100 t/ha). Recalling the maps in Fig. 9, these stands are composed by few isolated tall trees and shorter vegetation, like for example the colonizing forest stands in Lopé or the sparse forest stands in the center of the Mondah site. This agrees with Fig. 6, and supports the idea of using the horizontal structure index to adapt \( \alpha_0 \) on local stand conditions.

**B. Biomass estimation from TanDEM-X forest height and horizontal structure**

In this Section, the improvement of the height-to-biomass allometry by exploring the dependency of the allometric level on the horizontal structure index is addressed.

To define the relation of the allometric level \( \alpha_0 \) with the horizontal structure index \( HS \) the LVIS estimates \( B_{\text{LVIS}}, H_{\text{LVIS}} \) and \( HS_{\text{LVIS}} \) are used. In the range from 0 to 1 \( HS_{\text{LVIS}} \) has been segmented into \( N_{\alpha} \) non-overlapping and uniformly distributed intervals centered around the structure values \( \{HS_{\text{LVIS}, i}\}_{i=1}^{N_{\alpha}} \). For each of these intervals the allometric level \( \{\alpha_0(HS_{\text{LVIS}, i})\}_{i=1}^{N_{\alpha}} \) and a common reference allometric exponent \( \beta_0 \) are jointly estimated by means of a least squares optimization, similarly to (11):

\[
\min_{\alpha_0(HS_{\text{LVIS}, i}), \beta_0} \left\| \begin{bmatrix} b_{\text{LVIS}, 1} \\ b_{\text{LVIS}, 2} \\ \vdots \\ b_{\text{LVIS}, N_{\alpha}} \end{bmatrix} - \begin{bmatrix} \alpha_0(HS_{\text{LVIS}, 1}) \beta_0^{b_{\text{LVIS}, 1}} \\ \alpha_0(HS_{\text{LVIS}, 2}) \beta_0^{b_{\text{LVIS}, 2}} \\ \vdots \\ \alpha_0(HS_{\text{LVIS}, N_{\alpha}}) \beta_0^{b_{\text{LVIS}, N_{\alpha}}} \end{bmatrix} \right\|^2
\]

where \( b_{\text{LVIS}, i} \) and \( h_{\text{LVIS}, i} \) are the vectors containing the (reference) \( B_{\text{LVIS}} \) and \( H_{\text{LVIS}} \) values, for the generic \( i \)-th \( HS_{\text{LVIS}} \) interval centered at \( HS_{\text{LVIS}, i} \). The obtained allometric exponents for the three test sites Lopé, Mabounie and Mondah, are \( \beta_0 = 1.8, 1.3 \) and 2.5 respectively. Note that the obtained reference allometric exponent \( \beta_0 \) are very similar (at least in Lopé and...
Mabounie), but not identical to the ones obtained from (11) as the two optimisation problems are different.

The general behavior of $\alpha_0(H_{SLVIS})$ is obtained from (2) by

$$ \alpha_0(H_{SLVIS}) = \frac{B_{LVIS}}{H_{SLVIS}} $$

and is visualized by means of a 2D histogram in Fig. 11(a) for the Lopé site. The $\alpha_0(H_{SLVIS})$ obtained from (14) is indicated by the white dotted line. For all three sites a decreasing trend of $\alpha_0(H_{SLVIS})$ for increasing $H_{SLVIS}$ is obtained. This agrees with Fig. 6 and Fig. 10, and confirms the relevance of the structure index for adjusting the allometric level. The distribution becomes wider for large $H_{SLVIS}$ values (0.75 in Lopé, 0.9 in Mabounie and 0.75 in Mondah). At larger $H_{SLVIS}$, levels $\alpha_0(H_{SLVIS})$ jumps suddenly to very small values, and its distribution becomes asymmetrical. If $H_{STDX}$ is used instead of $H_{SLVIS}$, the distribution of $\alpha_0(H_{STDX})$ [Fig. 11(b)] becomes slightly wider, but its behavior does not change significantly.

AGB could be estimated by using $H_{TDX}$ and $\alpha_0(H_{TDX})$ in (3). The validation of the obtained AGB against the LVIS AGB are shown in Fig. 12. It becomes clear that the use of $\alpha_0(H_{TDX})$ successfully compensates for the density and/or structure induced variation of the single allometry with fixed allometric level. Indeed, the estimation performance increases drastically: the RMSE decreases down to 15 - 25% while the correlation coefficient increases to up to 0.7 - 0.9 for the different sites. The strong overestimation especially for low

Fig. 9. Horizontal structure maps at 1 ha resolution (a) from LVIS waveforms ($H_{SLVIS}$) and (b) from TanDEM-X CHPs ($H_{STDX}$) for Lopé, Mabounie, and Mondah. Sparse forest stands appears in red; dense stands appear in blue.

Fig. 10. Scatterplots between the reference LVIS top height ($H_{LVIS}$) and AGB ($B_{LVIS}$) for the three test sites. The color indicates the value of the LVIS horizontal structure index ($H_{SLVIS}$).
biomass values in Fig. 5 is widely compensated. The remaining residual overestimation, especially in Lopé, is a result of the ambiguity to relate $\alpha_0$ to a single value of $H_{SX}$ at large $H$ values (see Fig. 11). The optimization of (13) is equivalent to a minimization of the mean square error, including both the estimation bias and the standard deviation. Choosing $\alpha_0(H)$ as the one minimizing the bias between the estimates and LVIS AGB compensates almost complete the bias in the low biomass regions at the cost of an overall increase of the RMSE by 10%. However, AGB can still be underestimated in the high biomass region as a result of the underestimation of forest height.

C. Common height and structure-to-biomass allometry across test sites

The generalization of the height-to-biomass allometry across the three sites requires the estimation of a common (constant) allometric exponent $\beta_{OC}$ and a common $\alpha_{OC}(H_{TX})$ relation for all sites. For this, the procedure outlined in Sections IV.B has been applied to 50000 $B_{LVIS}$, $H_{LVIS}$ and $H_{SVIS}$ samples from all three sites. The samples are randomly selected across biomass levels in order to equally represent each site. In this way, an allometric exponent of $\beta_{OC} = 1.7$ and the $\alpha_{OC}(H_{TX})$
relationship shown in Fig. 13(a) are obtained. The behavior of $\alpha_{0c}(H_{TX})$ is consistent to the single-site relations obtained in the previous Section. The AGB estimates obtained from the common allometry and its comparison with the reference AGB is shown in Fig. 13(b). The overall RMSE is about 79 t/ha while the correlation is larger than 0.8. An underestimation of about 50 t/ha appears at around 300 t/ha. The overall performance is however convincing, keeping in mind the very different characteristics of the three sites.

### D. Effect of reduced penetration on structure calculation

TanDEM-X forest height estimation is limited by the insufficient penetration capability at X-band due to the high attenuation rates especially in tall/dense/wet forest conditions. This limitation no longer exists if the ground topography is known (e.g. an external DTM is available). In this sense, the lack of a DTM becomes critical for forest height estimation in tall/dense/wet forest conditions. Different height estimation algorithms can be followed [42]. Any additional height estimation inaccuracy resulting from this is directly reflected into an additional inaccuracy of $B_{TX}$. By differentiating (3), it is readily found that the sensitivity of a biomass estimation error is directly proportional to $a_0(H_{TX})$. From Fig. 11 or Fig. 13 it is apparent that a height error induces a larger biomass error for smaller $H_{TX}$ (i.e. denser forest stands) for a fixed allometric exponent.

With respect to the horizontal structure index, the high attenuation rates become an advantage increasing the sensitivity to the top-canopy height variations. Therefore, it is expected that the unavailability of an external DTM is less critical for its estimation [15], [41]. The external DTM allows to separate the top-canopy height variations from the topographic induced height variations. However, as the two contributions are expected to occur at different spatial scales this may allows to separate them from each other. As reported in an earlier study a low resolution TanDEM-X digital elevation model (DEM) can be used to compensate topographic induced height variations with sufficient accuracy [43]. A spatial resolution in the order of 100 m appears optimum [15], [43]. After DEM phase compensation both the CHP and the $H_{TX}$ are estimated as described.

In order to test the validity of these expectations $H_{TX}$ calculated with and without DTM are compared in Fig. 14. In order to increase the significance of this comparison the same top layer width has been used. The good agreement obtained in all sites confirms that the reduced penetration is the key feature for the retrieval of structure information. The availability of a DTM plays only a secondary role. However, it can still lead to a loss of performance especially in areas characterized by small-scale topographic variations as it has been seen in Lopé and Mabounie.

It is worth remarking that in the used structure estimation framework an error in the top layer width becomes an (relative) error in $H_{TX}$. $H_{TX}$ increases for an underestimated width as the top layer includes more CHP peaks, and vice versa. In this sense, compared to the knowledge of the ground topography, the misidentification of the top layer can cause a larger performance degradation. However, as discussed above, the performance remains comparable even for significant variations (50%) of the selected top layer.

### V. Conclusions

The potential of using forest structure information to adapt a general height-to-biomass allometry to local stand conditions for improving biomass estimation performance is investigated. It is addressed in terms of forest height and structure indices.
derived from remote sensing measurements particularly in the TanDEM-X context.

The discussed concepts have been demonstrated and validated with TanDEM-X data acquired in 2015-2016 over three different test sites in Gabon in the frame of the AfriSAR campaign. Reference forest height and biomass data as well as (horizontal) forest structure indices have been derived from airborne (LVIS) full waveform data acquired almost at the same time as the TanDEM-X data. TanDEM-X forest height, structure index and biomass estimates have been derived at a spatial scale of 1 ha and compared against the lidar reference data. The general height-to-biomass allometry used to transform forest height estimates to biomass has been established using the lidar height and biomass estimates. The lidar (LVIS) DTM has been used to compensate for the interferometric phase component induced by the terrain topography supporting in this way both the forest height inversion and the estimation of the horizontal structure index from TanDEM-X data.

Forest top height has been estimated with an RMSE well within 20%. In contrast, the biomass values obtained from a single height-to-biomass allometry is affected by large biases independently of the height estimation error.

A continuous relationship between the allometric level and the horizontal structure index was derived by means of a least squares minimization of the biomass RMSE. The achieved results indicate that the horizontal structure index is able to adapt a more general height-to-biomass relation to local forest (density) conditions by changing the allometric level. The estimation biases appearing in the conventional single height-to-biomass allometry are widely compensated improving the overall biomass RMSE up to 30%. In addition, the ability to establish a common height-to-biomass allometry supported by the horizontal structure index for all three sites with a reasonable performance is an important result. The identification of the top canopy layer in which the height variability is evaluated remains a critical choice in the estimation of the horizontal structure index.

Regarding the role of the lidar data, they are indispensable for the definition of the forest height-to-biomass allometry. However, their role in the estimation of forest height and the horizontal structure index is less critical as both can be performed even in the absence of any lidar measurements. More affected from the lack of lidar measurements is the estimation of forest height in tall / dense / wet forest conditions due to the high attenuation rates at X-band [42], rather than the estimation of the horizontal structure index itself.

There are three important features that make TanDEM-X especially appropriate for the quantitative characterization of the horizontal forest structure: i) the high attenuation rates at X-band and the associated limited penetration into the forest volume that maximize the sensitivity of the TanDEM-X (complex) coherence to the spatial variability of the top canopy layer; ii) the high interferometric accuracy of TanDEM-X (driven by the single-pass implementation) that allows to capture the structure induced (complex) coherence variation; and iii) the high spatial resolution of the TanDEM-X (complex) coherence measurements and its continuous measurement nature allows the estimation of structure induced variations at spatial scales relevant for the characterization of the horizontal forest structure.

ACKNOWLEDGEMENT

The lidar AfriSAR data sets were provided by the Land, Vegetation and Ice Sensor (LVIS) team in Code 61A at NASA Goddard Space Flight Center with support from the University of Maryland, College Park, Maryland, USA. This work was performed in the frame of a joint research project between DLR and NASA’s Carbon Monitoring System (CMS) program in support of NASA Grant #80NSSC20K023 to the University of Maryland.

REFERENCES

[1] R. A. Houghton, F. Hall, and S. J. Goetz, “Importance of biomass in the global carbon cycle,” J. Geophys. Res.-Biogeog., vol. 114, p. G00E03, 2009.
[2] M. Reichstein and N. Carvalhais, “Aspects of forest biomass in the earth system: Its role and major unknowns, Surveys in Geophysics, vol. 40, no. 4, pp. 693–707, Jul. 2019.
[3] J. Chave, C. Andalo, S. Brown, M. A. Cairns, J. Q. Chambers, D. Eamus, H. Fölster, F. Fromard, N. Higuchi, T. Kira, J.-P. Lescure, B. W. Nelson, H. Ogawa, H. Puig, B. Rífera, and T. Yamakura, “Tree allometry and improved estimation of carbon stocks and balance in tropical forests,” Oecologia, vol. 145, no. 1, pp. 87–99, 2005.
[4] T. Mette, K. Papanastassiou, I. Hajnsek, H. Pretzsch and P. Biber, “Applying a common allometric equation to convert forest height from Pol-InSAR data to forest biomass,” IGARSS 2004 - 2004 IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK, 2004, pp. 269–272.
[5] G. P. Asner, J. Mascaro, H. C. Muller-Landau, G. Viellellent, R. Vaudry, M. Rasamoelina, J. S. Hall, and M. van Beugel, “A universal airborne LiDAR approach for tropical forest carbon mapping,” Oecologia, vol. 168, no. 4, pp. 1147–1160, Apr. 2012.
[6] T. Mette, K. Papanastassiou and I. Hajnsek, “Biomass estimation from polarimetric SAR interferometry over heterogeneous forest terrain,” IGARSS 2004 - 2004 IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK, 2004, pp. 511–514.
[7] A. Torano Caicoya, F. Kugler, I. Hajnsek, and K. Papanastassiou, “Large-Scale Biomass Classification in Boreal Forests with TanDEM-X Data,” IEEE Trans. Geosci. Remote Sens., vol. 54, no. 10, pp. 5935–5951, Oct. 2016.
[8] N. Knapp, R. Fischer, and A. Huth, “Linking lidar and forest modeling to assess biomass estimation across scales and disturbance states,” Remote Sens. Environ., vol. 205, Feb. 2020, 111597.
[9] N. Knapp, R. Fischer, V. Cazcarra-Bes and A. Huth, “Structure metrics to generalize biomass estimation from lidar across forest types from different continents,” Remote Sens. Environ., vol. 237, pp. 199–209, Feb. 2020.
[10] M. Schlund, S. Erasmi and K. Scipal, “Comparison of Aboveground Biomass Estimation From InSAR and LiDAR Canopy Height Models in Tropical Forests,” IEEE Geoscience and Remote Sensing Letters, vol. 17, no. 3, pp. 367–371, Mar. 2020.
[11] M. J. Soja, H. J. Persson, and L. M. H. Ulander, “Estimation of Forest Biomass from Two-level Model Inversion of Single-pass InSAR data,” IEEE Trans. Geosci. Remote Sens., vol. 53, no. 9, pp. 5083–5099, 2015.
[12] A. E. L. Stovall, H. H. Shugart, “Improved biomass calibration and validation with terrestrial lidar: implications for future LiDAR and SAR missions,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 11, no. 10, pp. 3527–3537, Oct. 2018.
[13] J. B. Drake, R. O. Dubayah, D. B. Clark, R. G. Knox, J. B. Blair, M. A. Hofton, R. L. Chazdon, J. F. Weishampel, and S. Prince, “Estimation of tropical forest structure characteristics using large-footprint lidar,” Remote Sens. Environ., vol. 79, no. 2-3, pp. 305–319, Feb. 2002.
[14] R. O. Dubayah, S. L. Sheldon, D. B. Clark, M. A. Hofton, J. B. Blair, G. C. Hurt, and R. L. Chazdon, “Estimation of tropical forest height and
