Remote Sensing of Environment 253 (2021) 112153

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Mapping above-ground biomass in tropical forests with ground-cancelled P-band SAR and limited reference data

Maciej J. Soja [a,b,*], Shaun Quegan [c], Mauro M. d'Alessandro [d], Francesco Banda [e], Klaus Scipal [f], Stefano Tebaldini [d], Lars M.H. Ulander [g]

[a] MJ Soja Consulting, Hobart, Tasmania, Australia
[b] University of Tasmania, Hobart, Tasmania, Australia
[c] University of Sheffield, Sheffield, UK
[d] Politecnico di Milano, Milan, Italy
[e] Aresys, Milan, Italy
[f] European Space Agency, ESTEC, Noordwijk, The Netherlands
[g] Chalmers University of Technology, Gothenburg, Sweden

ARTICLE INFO

Keywords:
Above-ground biomass (AGB)
Estimation
Tropical forests
P-band synthetic aperture radar (SAR)
ESA BIOMASS
7th Earth Explorer
Ground cancellation

ABSTRACT

This paper introduces the CASINO (CAnopy backscatter estimation, Subsampling, and Inhibited Nonlinear Optimisation) algorithm for above-ground biomass (AGB) estimation in tropical forests using P-band (435 MHz) synthetic aperture radar (SAR) data. The algorithm has been implemented in a prototype processor for European Space Agency's (ESA's) 7th Earth Explorer Mission BIOMASS, scheduled for launch in 2023. CASINO employs an interferometric ground cancellation technique to estimate canopy backscatter (CB) intensity. A power law model (PLM) is then used to model the dependence of CB on AGB for a large number of systematically distributed SAR data samples and a small number of calibration areas with a known AGB. The PLM parameters and AGB for the samples are estimated simultaneously within pre-defined intervals using nonlinear minimisation of a cost function. The performance of CASINO is assessed over six tropical forest sites on two continents: two in French Guiana, South America and four in Gabon, Africa, using SAR data acquired during airborne ESA campaigns and processed to simulate BIOMASS acquisitions. Multiple tests with only two randomly selected calibration areas with AGB $>$ 100 t/ha are conducted to assess AGB estimation performance given limited reference data. At 2.25 ha scale and using a single flight heading, the root-mean-square difference (RMSD) is $\leq$ 27% for at least 50% of all tests in each test site and using as reference AGB maps derived from airborne laser scanning data. An improvement is observed when two flight headings are used in combination. The most consistent AGB estimation (lowest RMSD variation across different calibration sets) is observed for test sites with a large AGB interval and average AGB around 200–250 t/ha. The most challenging conditions are in areas with AGB $<$ 200 t/ha and large topographic variations. A comparison with 142 1 ha plots distributed across all six test sites and with AGB estimated from *in situ* measurements gives an RMSD of 20% (66 t/ha).

1. Introduction

Our current understanding of the global carbon cycle and its role in climate warming is strongly limited by the large uncertainty in both the CO_2 emissions from the land surface due to land use change (mainly tropical deforestation) and forest degradation, and the uptake of CO_2 from the atmosphere by forest growth. The most recent overview (Friedlingstein et al., 2019) indicates that land use change emissions have the largest relative uncertainty (exceeding 50%) of all the terms in the global carbon cycle, while the land uptake term has the largest absolute uncertainty. Intimately involved in both terms is forest biomass (approximately 50% of which is carbon), which is lost during land use change and gained by forest growth. This dual role of biomass in the climate system is why above-ground biomass (AGB) is recognised by the Global Climate Observing System (GCOS) as an Essential Climate Variable (GCOS, 2016). (Below-ground biomass is also important but is difficult to measure so is normally estimated using allometric relations or models, *e.g.*, Thurner et al., 2014.) The primary objective of the

* Corresponding author at: MJ Soja Consulting, Hobart, Tasmania, Australia.
E-mail addresses: maciej@mjsoja.com, maciej.soja@utas.edu.au (M.J. Soja).

https://doi.org/10.1016/j.rse.2020.112153
Received 20 January 2020; Received in revised form 15 October 2020; Accepted 20 October 2020
Available online 5 November 2020

European Space Agency (ESA) 7th Earth Explorer Mission BIOMASS (ESA, 2008, 2012; Quegan et al., 2019) is therefore to determine the worldwide distribution of forest AGB in order to reduce these major uncertainties in calculations of terrestrial carbon stocks and fluxes (Herold et al., 2019).

BIOMASS is scheduled for launch in 2023 and will provide near-global (subject to Space Object Tracking Radar restrictions; Carreiras et al., 2017) maps of forest AGB, height and disturbance. Achieving this involves several innovative technologies. It will carry the first space-borne synthetic aperture radar (SAR) sensor operating at P-band (435 MHz), chosen because this frequency penetrates deep into forest canopies and provides access to the large branches and trunks making up most of the biomass (Hajnsek et al., 2009a; Sandberg et al., 2011; Dubois-Fernandez et al., 2012; Soja et al., 2013; Quegan et al., 2019). Moreover, these large forest elements are temporally more stable, which will allow BIOMASS to employ coherent techniques under a suitable repeat-pass orbit pattern (Albinet et al., 2012; Minh et al., 2014a; Monteith and Ulander, 2018; Tebaldini et al., 2019). As a result, the mission will provide the first systematic spaceborne measurements of forest height with polarimetric-interferometric SAR (PolInSAR) (Quegan et al., 2019) and the first 3D measurements of forest backscatter using tomographic SAR (TomoSAR) (Tebaldini et al., 2019). The P-band frequency will also allow the first mapping of sub-forest topography in densely vegetated areas (d'Alessandro and Tebaldini, 2019; Gatti et al., 2011).

While forest height can be measured at P-band using well-established techniques without the need for ground reference data (Cloude and Papathanassiou, 1998; Papathanassiou and Cloude, 2001), even in relatively sparse boreal forests (Garestier et al., 2008), AGB is a complex quantity determined not only by the geometric structure of the forest (tree architecture and distribution) and its dielectric properties, but also by wood density (Baker et al., 2004; Chave et al., 2008; Phillips et al., 2019). Therefore, although effective AGB estimation from P-band SAR data has been demonstrated on many datasets (Le Toan et al., 2011; Quegan et al., 2019; Sandberg et al., 2011; Soja et al., 2013; Villard and Le Toan, 2015, and references therein), the development of a globally applicable algorithm or set of algorithms still presents a serious challenge.

For BIOMASS, this challenge has several factors:

- Deep penetration into the forest comes at the expense of significant contamination of the signal from the forest canopy by scattering from the ground, either as direct backscatter or through double-bounce interactions. This gives rise to variations in the return caused by topography and soil moisture, not AGB (Minh et al., 2014b; Monteith and Ulander, 2018; Soja et al., 2013; Ulander et al., 2018; Villard and Le Toan, 2015).
- Airborne campaigns are costly and time-consuming, so the available experimental data are spatially and temporally sparse (Dubois-Fernandez et al., 2011; Hajnsek et al., 2008, 2009a, 2009b, 2017; Ulander et al., 2011). Hence, we have limited empirical knowledge about the effects of moisture, seasonal variations, and different forest types and biomes on P-band SAR polarimetric backscatter.
- *In situ* inventories are essential for AGB estimation from space (Chave et al., 2019), but they are costly and time-consuming. Reference AGB data will only be available for a limited number of strategically located sites.
- BIOMASS is planned to have an unusual orbit configuration, which is organised into two main phases, an initial 14-month TomoSAR phase and the subsequent PolInSAR phase. The latter will provide data derived from triplets of observations separated by 3 days, giving global coverage every 228 days (~7.5 months) until the end of the 5-year mission. Hence, successive global cycles will exhibit seasonal shifts.
- The combination of ascending and descending orbits yields observations with different geometries and observation patterns that vary with latitude and longitude.
- There are difficulties in transferring inferences from airborne data to the BIOMASS case because of the steeper incidence angles (between 23° and 34°), the coarser spatial resolution (about 60 m and 8.3 m in ground range and azimuth, respectively, for single-look BIOMASS data), and the much larger coverage (about 150 km × 150 km in ground range and azimuth).

These factors have all played a part in the conceptual development of AGB estimation algorithms for P-band SAR.

The main aim of this paper is to introduce the CASINO (CAnopy backscatter estimation, Subsampling, and Inhibited Nonlinear Optimisation) algorithm for AGB estimation from P-band SAR backscatter, and to assess its performance across six tropical forest test sites in two countries: Gabon, Africa and French Guiana, South America. CASINO begins by minimizing topographic and soil moisture effects using a ground cancellation technique (d'Alessandro et al., 2020; Soja et al., 2018), which isolates the scattering from the forest canopy. The dependence of canopy backscatter (CB) on AGB is then described using a power law model (PLM), which is supported by numerous empirical studies and asymptotic derivations from theoretical models (Le Toan et al., 2011; Quegan et al., 2019; Schlund et al., 2018; Soja et al., 2013, 2018). The PLM is fitted to CB data using constrained nonlinear minimisation of a cost function (Soja et al., 2017), calculated for a large number of data samples. AGB estimation bias is minimised using local calibration data and constraints on AGB and PLM parameter intervals and their spatial, temporal and polarisation variability (Soja et al., 2018). This paper shows that the combination of these principles leads to AGB estimation performance across all six test sites that is similar to other P-band studies, but using only two reference AGB measurements for calibration. As a result, the CASINO algorithm has been selected for initial implementation in the BIOMASS ground processor (Banda et al., 2020).

The paper is structured as follows. Section 2 introduces the cornerstones of the CASINO algorithm, while Section 3 describes the sites and the experimental data used to assess its performance. In Sections 4 and 5, the results are presented and discussed. Finally, Section 6 provides some implications for future use of CASINO with BIOMASS data.

2. The CASINO algorithm

The CASINO algorithm consists of three steps:

1) estimation of canopy backscatter (CB) using a ground cancellation technique;
2) estimation of power law model (PLM) parameters using samples of CB data and constrained nonlinear minimisation of a cost function;
3) estimation of AGB for the entire scene from CB data using the estimated PLM parameters.

2.1. Canopy backscatter estimation

Canopy backscatter is estimated using a ground cancellation technique (d'Alessandro et al., 2020), which decreases ground contributions to the signal arising from direct ground backscatter and double-bounce interactions between ground and vegetation. These contributions are strongly influenced by ground topography, moisture and roughness, which are independent of AGB and difficult to model (Minh et al., 2014b; Monteith and Ulander, 2018; Soja et al., 2013; Ulander et al., 2018; Villard and Le Toan, 2015). Several studies have shown that removal of ground-level backscatter yields higher correlation with AGB and better potential for AGB estimation (Blomberg et al., 2018; d'Alessandro et al., 2020; Minh et al., 2014b; Soja et al., 2018; Tebaldini et al., 2019).

Given two co-registered single-look complex (SLC) images, one master and one slave, the ground-cancelled SLC image (s_{GC}) is formed by subtracting the ground-steered SLC master image (s_M) from the

corresponding SLC slave image (s_S), where ground-steering refers to removal of the interferometric phase due to ground topography (d'Alessandro et al., 2020):

$$s_{GC} = s_S - s_M \approx s_M\left(e^{ik_z z} - 1\right) = 2is_M e^{\frac{ik_z z}{2}} \sin\left(\frac{k_z z}{2}\right) \quad (1)$$

where k_z is the phase-to-height conversion factor (Papathanassiou and Cloude, 2001), and where it is assumed that $s_S \approx s_M e^{ik_z z}$, following the principles of SAR interferometry (Bamler and Hartl, 1998).

The ground-cancelled intensity is then given by:

$$|s_{GC}|^2 = 4|s_M|^2 \sin^2\left(\frac{k_z z}{2}\right) = 4|s_M|^2 \sin^2\left(\frac{\pi z}{z_{2\pi}}\right) \quad (2)$$

where $z_{2\pi} = \frac{2\pi}{k_z}$ is the height-of-ambiguity (HOA), which is the height shift corresponding to a 2π shift of the interferometric phase. The squared sine modulation in (2) strongly attenuates scattering from objects close to the ground and enhances targets located around HOA/2.

The shape of the modulation function depends on k_z, which depends on the baseline and acquisition geometry. To maximise the contribution of the canopy, the baseline should be selected so that the first maximum falls at the height within the canopy where the correlation and sensitivity to AGB is highest. The selection of such an optimal baseline is challenging for both airborne and spaceborne interferometric SAR systems. In the airborne case, unstable flight trajectories and large incidence angle variations cause significant baseline fluctuations in the range and azimuth directions. In the spaceborne case, the incidence angle variability is smaller due to the higher altitude, and the flight trajectory is more stable because there is no perturbing atmosphere. However, the baselines are more difficult to control and are subject to complex latitudinal and longitudinal variations.

To mitigate the effect of sub-optimal baselines, a technique called equalisation is employed (d'Alessandro et al., 2020). This uses all available SLC images within an interferometric or tomographic stack (*i.e.*, a set of co-registered images of the same scene) to interpolate a synthetic SLC image with the required baseline relative to the master image. It also reduces baseline fluctuations within a scene caused by incidence angle variability and flight trajectory fluctuations.

Finally, the ground-cancelled and equalised backscatter intensity $|s_{GCE}|^2$ is normalised to σ^0 to provide an estimate of CB:

$$\sigma^0_{cnp} = c\langle|s_{GCE}|^2\rangle \cos\psi \quad (3)$$

where c is a calibration constant, $\langle\cdot\rangle$ denotes spatial averaging over a window, and ψ is the angle between the ground surface normal and the image plane normal, as depicted in Fig. 1 (Ulander, 1996).

CB estimation is performed for each polarisation and stack of SLC images, and results in one CB image per polarisation and stack, independent of the number of images within the individual stacks.

2.2. Estimation of power law model parameters

The dependence of CB on AGB is described using a power law model (PLM), which is fitted to the data using an approach based on sampling the SAR data and constrained nonlinear minimisation of a cost function.

2.2.1. The power law model (PLM)

The PLM describes the dependence of σ^0_{cnp} (3) on AGB (W, in t/ha) and the local incidence angle (ϑ, the angle between the range direction and the ground surface normal, see Fig. 1). It requires three parameters for each polarisation PQ: a scaling factor (A_{PQ}) and two exponents, one for AGB (α_{PQ}) and one for the incidence angle normalisation (n_{PQ}) (Banda et al., 2020; Quegan et al., 2019; Soja et al., 2018):

$$\left(\sigma^0_{cnp}\right)_{PQ} = A_{PQ} W^{\alpha_{PQ}} \cos^{n_{PQ}}\vartheta \quad (4)$$

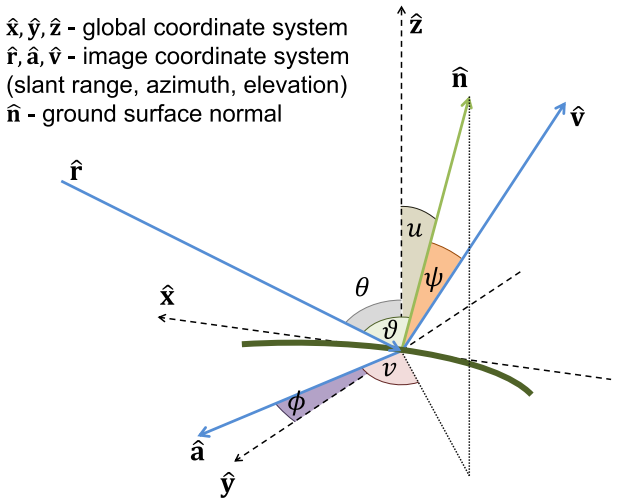

Fig. 1. Geometry of the data acquisition for a left-looking SAR system: ϕ is the flight heading angle, θ is the global incidence angle, ϑ is the local incidence angle, and u and v are the ground slope and slope aspect angles. The angle between the image surface normal and the ground surface normal is referred to as ψ and is used for accurate calibration to σ^0 (Ulander, 1996).

The PLM is well-supported by experimental results from nearly three decades of research (Le Toan et al., 2011; Quegan et al., 2019; Sandberg et al., 2011; Schlund et al., 2018; Soja et al., 2013, and references therein) and is an asymptotic form of more advanced theoretical models available in the literature. As an example, the model proposed in Truong-Loï et al. (2015) predicts direct canopy backscatter as the product of a power law function and an attenuation function:

$$\left(\sigma^0_{cnp}\right)_{PQ} = \underbrace{A_{PQ} W^{\alpha_{PQ}}}_{\text{power law}} \underbrace{\left(1 - \exp\left(-\frac{B_{PQ} W^{\beta_{PQ}}}{\cos\theta}\right)\right)}_{\text{attenuation}} \cos\theta \quad (5)$$

where B_{PQ} and β_{PQ} are polarisation-dependent model parameters and θ is the global incidence angle (the angle between the range vector and the vertical direction, see Fig. 1). This combines elements of two models commonly used to describe the dependence of backscatter on AGB:

- A power law model, which is most commonly used for AGB estimation from backscatter at lower frequencies (Schlund et al., 2018). In the special case of backscattering from a single thin dielectric cylinder, backscatter intensity is proportional to the square of its volume; this approximation has been used in the past to model low-frequency SAR backscatter from boreal forests (Brolly and Woodhouse, 2012; Smith-Jonforsen et al., 2005).
- An attenuation model, rooted in the water cloud model (Attema and Ulaby, 1978). The water cloud model is the basis for both the interferometric water cloud model, primarily used for AGB and stem volume estimation from interferometric SAR data (Askne et al., 1997, 2013; Soja et al., 2017), and the random volume over ground model commonly used for forest height estimation from PolInSAR data (Cloude and Papathanassiou, 1998; Hajnsek et al., 2009a; Papathanassiou and Cloude, 2001).

While the results in Truong-Loï et al. (2015) are promising, the approach used for parameter estimation requires substantial amounts of reference and simulation data, which makes it unsuitable for large-scale applications. One way to estimate the parameters is by fitting (5) to data acquired during airborne campaigns. Although this approach has given some positive results (Soja et al., 2018), further studies have concluded that it is numerically unstable, probably because of the interactions between the power law and attenuation parts. In fact, for reasonable

values of the parameters (non-negative and within limits supported by experimental studies and theoretical arguments), both parts of (5) are increasing functions of AGB, and separation of these effects is difficult.

Two simpler asymptotic cases of (5) can be distinguished (Quegan et al., 2019). For high attenuation, *i.e.*, when $B_{PQ}W^{\beta_{PQ}} \gg 1$, the exponential vanishes, and the model transforms into a power law function (Soja et al., 2018):

$$\left(\sigma^0_{cnp}\right)_{PQ} \approx A_{PQ} W^{\alpha_{PQ}} \cos\theta \tag{6}$$

Conversely, for low attenuation, *i.e.*, when $0 < B_{PQ}W^{\beta_{PQ}} \ll 1$, a first-order Taylor expansion of the exponential term also yields a power law function, but without the cosine term:

$$\left(\sigma^0_{cnp}\right)_{PQ} \approx A'_{PQ} W^{\alpha'_{PQ}} \tag{7}$$

where $A'_{PQ} = A_{PQ}B_{PQ}$ and $\alpha_{PQ}{}' = \alpha_{PQ} + \beta_{PQ}$.

The PLM in (4) can be seen as a generalised case of (6) and (7), with the exponent n_{PQ} emulating the transition from the high-attenuation case (with $n_{PQ} = 1$) to the low-attenuation case (with $n_{PQ} = 0$). While not intended as a rigorous derivation, this is the motivation for selecting (4) as the basis of the algorithm. Note that in (4), the local incidence angle (ϑ) is used, while (5) and (6) use the global incidence angle (θ). This is a commonly used first-order correction for topographic effects (Small, 2011; Ulander, 1996; Villard and Le Toan, 2015).

Fitting to data uses the log-transformed version of (4), which converts it to the simple form:

$$s_{PQ} = l_{PQ} + \alpha_{PQ} w + n_{PQ} c \tag{8}$$

where $s_{PQ} = 10\ \mathrm{lg}\ [k_{PQ}(\sigma^0_{cnp})_{PQ}]$ (with $k_{PQ} = 1$ for the HH and VV channels and 2 for the HV channel), $c = 10\ \mathrm{lg}\ (\cos\vartheta)$, $l_{PQ} = 10\ \mathrm{lg}\ A_{PQ}$, $w = 10\ \mathrm{lg}\ W$, and lg denotes base-10 logarithm. The polarisation weight k_{PQ} is used to correctly account for the total backscattered power.

Note that (8) is nonlinear when l_{PQ}, α_{PQ}, n_{PQ} and w are all unknown parameters and the observed quantities are s_{PQ} (one for each polarisation) and c. Consequently, iterative methods are needed to simultaneously estimate w and the three PLM parameters. Only in the special case when w is known (*i.e.*, when reference AGB data are available) does the model become linear, whereupon the model parameters can be estimated using linear regression.

2.2.2. The estimation approach

The model (8) is fitted to data using constrained nonlinear minimisation of a cost function J (*e.g.*, the Python SciPy implementation of the BFGS algorithm; Nocedal and Wright, 2006), where the estimated values are constrained to lie in the ranges shown in Table 1. These intervals were selected based on physical interpretation of the PLM parameters and the AGB levels expected for the studied forests; the minimum values for W and α_{PQ} were found empirically to give better numerical stability.

The cost function is expressed as the sum of squared differences between the CB coefficient modelled with (8) and the logarithm of the CB coefficient estimated from experimental data with (3). It is calculated for a large number of SAR data samples, which provide a good representation of the backscatter and incidence angle variability across the entire scene. Two types of sampling areas are used:

Table 1
Constraints imposed in the minimisation on AGB and PLM parameters for polarisation PQ.

Parameter	Unit	Lower limit (p_{min})	Upper limit (p_{max})
AGB (W)	t/ha	1	700
l_{PQ}	dB	−60	0
α_{PQ}	dB/dB	0.01	2.0
n_{PQ}	1	0.0	3.0

- calibration (CAL) areas, consisting of a few areas with AGB available from a reliable external source (*e.g.*, *in situ* measurements),
- estimation (EST) areas, consisting of a large number of systematically distributed areas with unknown AGB, but well-represented in terms of incidence angle and backscatter variability.

Consequently, J can be expressed as the sum of two cost functions, one for the calibration areas (J_{CAL}) and one for the estimation areas (J_{EST}):

$$J(\boldsymbol{l}, \boldsymbol{\alpha}, n, w) = J_{CAL}(\boldsymbol{l}, \boldsymbol{\alpha}, n) + J_{EST}(\boldsymbol{l}, \boldsymbol{\alpha}, n, w) \tag{9}$$

where $\boldsymbol{l} = [\ldots, l_{PQ}, \ldots]$, $\boldsymbol{\alpha} = [\ldots, \alpha_{PQ}, \ldots]$ and $\boldsymbol{n} = [\ldots, n_{PQ}, \ldots]$ are vectors containing the PLM parameters for all polarisations PQ, and $\boldsymbol{w} = [w_1, \ldots, w_i, \ldots, w_{N_{EST}}]$ is a vector containing the AGB in decibels (*i.e.*, w) for all N_{EST} estimation areas.

The cost functions for the calibration and estimation areas are:

$$J_{CAL}(\boldsymbol{l}, \boldsymbol{\alpha}, n) = \frac{1}{N_{CAL}} \sum_{i=1}^{N_{CAL}} \sum_{j=1}^{M} \sum_{PQ} \left(l_{PQ} + \alpha_{PQ} w_i^{CAL} + n_{PQ} c_{ij}^{CAL} - s_{PQij}^{CAL} \right)^2 \tag{10}$$

$$J_{EST}(\boldsymbol{l}, \boldsymbol{\alpha}, n, w) = \frac{1}{N_{EST}} \sum_{i=1}^{N_{EST}} \sum_{j=1}^{M} \sum_{PQ} \left(l_{PQ} + \alpha_{PQ} w_i + n_{PQ} c_{ij}^{EST} - s_{PQij}^{EST} \right)^2 \tag{11}$$

where M is the number of independent stacks used to obtain the CB images, N_{CAL} is the number of calibration areas, and superscripts CAL and EST are used to indicate the measured values for sampling area i and stack j (this includes w for calibration areas). Note that (10) and (11) assume that all calibration and estimation areas are covered by all stacks and all polarisations, and that the AGB and PLM parameters are the same for all stacks. If these assumptions become invalid, additional parameters and indexing need to be introduced in (10) and (11), but that is outside the scope of this paper.

To obtain physically meaningful results, all parameters are constrained within pre-defined intervals by parameterising each parameter p (where p can be any of l_{PQ}, α_{PQ}, n_{PQ}, or w_i) with a new, unconstrained parameter x_p. The desired (constrained) parameter value is calculated from x_p by:

$$p(x_p) = p_{max} + (p_{max} - p_{min}) \sin^2 x_p \tag{12}$$

where p_{min} and p_{max} are the smallest and largest permissible values of p, as per Table 1. The cost function J is minimised with respect to x_p rather than p. The function (12) was selected because it is simple, easily differentiable, and has provided reliable results in this study. Other functions may be used, but selection of the optimal function is the subject of further study.

2.2.3. Sampling considerations

The proposed approach simultaneously estimates PLM parameters and AGB for the estimation areas. Although this involves a large number of unknown parameters, the even larger number of measured quantities and the additional constraints in the form of calibration data and predefined parameter intervals make this possible. In fact, assuming three polarisations, only one stack of SLC data ($M = 1$), and two calibration areas ($N_{CAL} = 2$), the total number of observations is $3(2 + N_{EST})$ (one backscatter coefficient per polarisation for each calibration and estimation area), while the number of unknown parameters is $N_{EST} + 9$ (one AGB value for each estimation area and one parameter value per polarisation for each of the three parameters). Hence, even with only two estimation areas ($N_{EST} = 2$), the number of unknown parameters is smaller than the number of observables (11 and 12, respectively). However, practical applications require many more estimation areas to account for similarities between multiple observations and data imperfections. We ensure $N_{EST} > 150$, which was found to be a good trade-off

between computational efficiency and numerical stability for the resolution and spatial coverage of the SAR data used in this paper (see Section 3).

The two terms in the cost function J (9) serve different purposes. By including many estimation areas in the minimisation, the forward model can be fitted to a wide range of different combinations of CB, AGB and local incidence angle across the scene, and thus provide a better representation of the spatial and polarisation variability in the data. Moreover, the constraints applied during fitting restrict the predicted AGB values to a pre-defined, physically motivated interval, *e.g.*, based on *a priori* knowledge of the studied forest.

The cost function for the calibration areas is necessary to provide unbiased AGB estimation, as can be seen if only J_{EST} in (11) is minimised. Then any value of w related to the logarithm of the true AGB, w_0, by:

$$w = q_0 + q_1 w_0 + q_2 c, \tag{13}$$

where q_0, q_1 and q_2 are arbitrary, would be consistent with the known values of s_{PQ} and c for an appropriate set of values of l_{PQ}, α_{PQ} and n_{PQ}. This is easily verified by substituting (13) into (8):

$$\begin{aligned} s_{\mathrm{PQ}} = l_{\mathrm{PQ}} + \alpha_{\mathrm{PQ}} w + n_{\mathrm{PQ}} c = (l_{\mathrm{PQ}} + \alpha_{\mathrm{PQ}} q_0) + \alpha_{\mathrm{PQ}} q_1 w_0 + \\ (\alpha_{\mathrm{PQ}} q_2 + n_{\mathrm{PQ}}) c \equiv l'_{\mathrm{PQ}} + \alpha'_{\mathrm{PQ}} w_0 + n'_{\mathrm{PQ}} c \end{aligned} \tag{14}$$

Unless dealt with, this ambiguity will cause biases in the AGB estimation.

Including the calibration areas, for which the relation between w, w_0 and c is fixed, in the total cost function J can resolve this ambiguity. The approach is flexible with respect to calibration data, as there is no implicit requirement on the reference AGB data and the calibration is performed using all available information. It is also flexible with respect to available SAR data: the same formulation can be used with single and multiple SAR data stacks, both fully and partially overlapping (*e.g.*, acquisitions from adjacent orbits or acquisitions from the ascending and descending directions), and fully polarimetric data are not strictly required (although having multiple polarisations is recommended). Finally, AGB and PLM parameter constraints can be easily enforced in the proposed formulation.

The four main challenges of the approach are a complex algorithm, a high computational load, numerical instabilities related to the nonlinear fitting, and uncertainties related to the availability and quality of calibration data. The first three challenges can be addressed with adequate computational implementation, so they will not be addressed further in this paper. The fourth issue is more critical, as it relates to the available reference AGB measurements, which are costly and time-consuming to acquire. As it is paramount that the algorithm performs well with the available reference data, this is the main focus of what follows.

A different way to resolve the ambiguity expressed in (13) is to estimate the three parameters q_0, q_1 and q_2 using a set of calibration data with known values for w, w_0 and c. Once the three parameters are estimated, (13) can be used to post-calibrate the obtained AGB estimates for all estimation areas. However, with this approach it is difficult to enforce constraints on the AGB and PLM parameter estimates. Moreover, to estimate the three parameters in (13), a minimum of three independent calibration measurements is needed. As shown later, the proposed approach can provide reliable AGB estimates using only two calibration measurements.

2.3. AGB estimation

Once the PLM parameters l_{PQ}, α_{PQ} and n_{PQ} are known for all polarisations, the AGB can be estimated for any area with available measurements of $s_{\mathrm{PQ}j}$ and c_j using:

$$\widehat{w} = \sum_{\mathrm{PQ}} \sum_{j=1}^{M} \lambda_{\mathrm{PQ}} \widehat{w}_{\mathrm{PQ}j} \tag{15}$$

where

$$\widehat{w}_{\mathrm{PQ}j} = \frac{s_{\mathrm{PQ}j} - l_{\mathrm{PQ}} - n_{\mathrm{PQ}} c_j}{\alpha_{\mathrm{PQ}}} \tag{16}$$

is the estimated AGB in decibels for polarisation PQ and stack j, obtained by solving (8) for w, and

$$\lambda_{\mathrm{PQ}} = \frac{\alpha_{\mathrm{PQ}}^2}{M \sum_{\mathrm{PQ}} \alpha_{\mathrm{PQ}}^2} \tag{17}$$

is the weight for polarisation PQ. Note that the estimate of the AGB (in dB) in (15) is a weighted sum of the individual estimates obtained for each polarisation and stack, where the weight is determined by parameter α, *i.e.*, the sensitivity of the particular polarisation to the logarithm of AGB. Effectively, the polarisations and stacks with better sensitivity to AGB make a greater contribution to the final estimate.

The final AGB estimates are obtained by transforming (15) back to linear units (t/ha) and correcting for logarithmic bias (Finney, 1941) using the method proposed by Snowdon (1991):

$$\widehat{W} = \rho \cdot 10^{0.1 \widehat{w}} \tag{18}$$

where the multiplicative correction factor ρ is estimated from CAL data using:

$$\rho = \frac{\sum_{i=1}^{N_{\mathrm{CAL}}} W_i^{\mathrm{CAL}}}{\sum_{i=1}^{N_{\mathrm{CAL}}} \widehat{W}_i^{\mathrm{CAL}}}$$

where $\widehat{W}_i^{\mathrm{CAL}}$ is the AGB estimate (in t/ha) for CAL area i, obtained by transforming the results from (15) back to linear units, and where W_i^{CAL} is the corresponding reference AGB estimate.

3. Data

3.1. Test sites

This study uses data from two test sites in French Guiana (Paracou and Nouragues), South America, and four located in Gabon, in equatorial Africa (Lopé, Mabounié, Mondah, and Rabi), see Fig. 2.

3.1.1. French Guiana

Nouragues (4.06°N, 52.68°W) is situated about 105 km south-southwest of Cayenne, French Guiana. For the parts of the test site relevant to this paper, altitude above sea level varies between 26 and 280 m, with rolling hills causing significant topographic variations. The site contains about 145 tree species per hectare and is covered by both high-canopy, old-growth forest and other forest types, including periodically flooded forest. The forest is protected and no major disturbances have occurred in the last 100 years.

Paracou (5.27°N, 52.93°W) is situated roughly 75 km west-northwest of Cayenne. Altitude above sea level varies between 5 and 45 m and the topographic variations are somewhat gentler than in Nouragues. There are about 140 tree species per hectare and the site is dominated by moist evergreen rainforest.

3.1.2. Gabon

Lopé (0.20° S, 11.59° E) is located approximately 250 km east-southeast of Libreville. Topographic variations are moderate, with altitudes above sea level varying between 200 and 600 m. The site consists of an undisturbed, primary forest-savanna mosaic that is part of a protected national park with an average species density of 35 ha^{-1}.

Mabounié (0.76°S, 10.56°E) is located roughly 180 km south-east of Libreville. Altitudes above sea level vary between 25 and 230 m, with significant topographic undulations. The site has a species density of

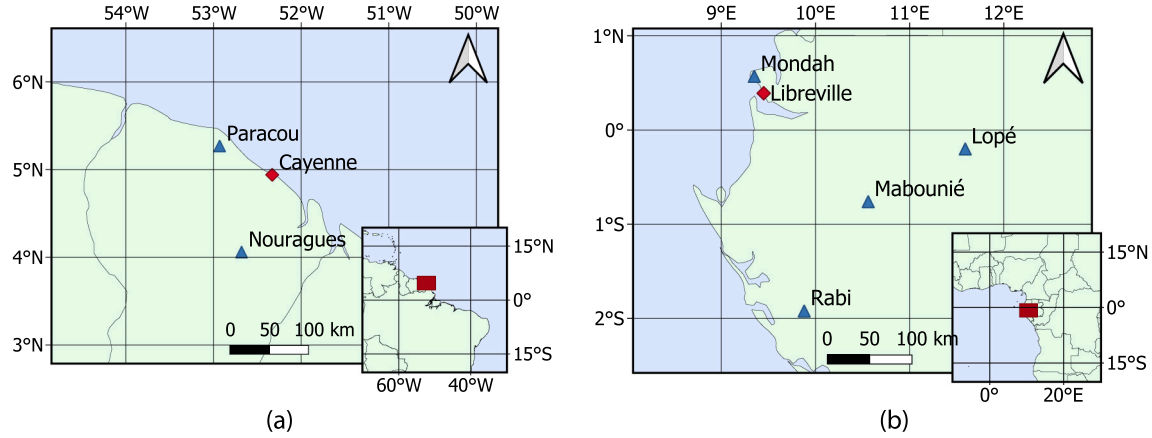

Fig. 2. The locations of the six studied test sites (blue triangles) in (a) French Guiana and (b) Gabon and the locations of the respective capitals Cayenne and Libreville (red diamonds).

about 55 ha^{-1} and the landscape is mostly forested, with areas of substantial forest degradation.

Mondah (0.57°N, 9.35°E) is located roughly 25 km north-west of Libreville. Most of the test site has altitude less than 50 m above sea level. The site has a similar species density to Lopé and the landscape consists of varying vegetation types, including mixed and *Aucoumea*-dominated forests.

Rabi (1.92° S, 9.88° E) is located roughly 260 km south of Libreville. Altitudes above sea level vary between 30 and 80 m. Vegetation is dominated by lowland tropical rainforest with a species density of about 85 ha^{-1}.

3.2. Insitu data

Field inventories were carried out for all six test sites, but on different occasions, by different research groups, following different protocols, and acquiring slightly different measurements. All field data were later jointly analysed using the same approach (Labrière et al., 2018). For consistency, only trees with diameter at breast height (dbh) greater than 10 cm were selected for the analysis. AGB estimates were obtained for each tree using the measured dbh, species-specific wood density, and an AGB estimation model available within the R BIOMASS package (Réjou-Méchain et al., 2017). AGB values for all trees within plots were then aggregated to plot-level estimates. For this study, only plots with size 1 ha were included. Additional information about the plot-level data can be found in Table 2 and in Hajnsek et al. (2017) and Labrière et al. (2018).

3.3. ALS-based AGB maps

For all six test sites, airborne laser scanning (ALS) surveys were conducted within a few years of the field inventories and/or SAR acquisitions. Somewhat different methodologies were used for the different datasets, and point densities varied between 2.4 m^{-2} for Lopé and 30.5 m^{-2} for Mondah. The returns were first classified as ground or vegetation, and then a digital terrain model (DTM) and a canopy height model (CHM) were created on a 1 m × 1 m grid. From the CHM, a mean canopy height (MCH) map was created with a grid posting of 50 m × 50 m. Using the plot-level AGB estimates described in Section 3.2 and the associated MCH values, one empirical MCH-AGB relationship was established for the six test sites described in Section 3.1 (Labrière et al., 2018). For each test site, this empirical model was then used to create a 50 m × 50 m AGB map for the entire area covered by ALS data. The estimated root-mean-square error of the MCH-AGB relationship for the 1 ha *in situ* plots was 47.5 t/ha and the estimated bias was 2.1%. Additional information about the ALS data and AGB maps can be found in Table 2 and in Hajnsek et al. (2017) and Labrière et al. (2018).

3.4. SAR data

Airborne P-band SAR data were collected at Nouragues and Paracou in August 2009 (Table 2) by the SETHI system from the French Aerospace Lab (ONERA) as part of the ESA TropiSAR campaign (Dubois-Fernandez et al., 2011). The SETHI system was left-looking and used a centre frequency of 440 MHz, a bandwidth of 70 MHz, and it covered incidence angles between 24° and 62°. Polarimetric SAR data were acquired with various temporal and/or spatial baselines over four test sites, of which Nouragues and Paracou were selected for this study due to the availability of multi-baseline data acquired on the same day. These provided one data stack for each test site. Due to calibration uncertainties observed in near range in the TropiSAR data (Banda et al., 2020), only data with incidence angles above 33° in Nouragues and 30° in Paracou were used.

The Lopé, Mabounié, Mondah and Rabi data were collected in February 2016 by the F-SAR system from the German Aerospace Center (DLR) as part of the AfriSAR campaign (Hajnsek et al., 2017). The F-SAR system was right-looking and used a centre frequency of 435 MHz, a bandwidth of 50 MHz, and it covered a swath with incidence angles between 25° and 55°. Over 80 polarimetric SAR datasets with various baselines and multiple headings were acquired over the four test sites. We considered only multi-baseline data acquired on the same day, resulting in four stacks with different flight headings for Lopé and one each for the other three test sites.

A total of nine SAR data stacks were therefore available, and these were organised into seven datasets (Table 2). Six single-heading datasets were created, one for each test site. In Lopé, the single-heading dataset consisted of four stacks with different flight headings, while in the remaining five test sites, each dataset consisted of one stack. Additionally, one dual-heading dataset was created in Lopé, consisting of all pairs of the four stacks, which resulted in a total of six stack pairs with unique flight heading combinations. For the single-heading datasets, AGB estimation was performed individually for each stack, while for the dual heading dataset in Lopé, AGB estimation was performed individually for each stack pair.

In order to simulate BIOMASS data, each image was filtered to 6 MHz and multi-looked to a resolution of 50 m in both ground range and azimuth directions, which is close to that of the 6-look BIOMASS data (Quegan et al., 2019). A DTM was then estimated from P-band SAR tomography using the approach described in d'Alessandro and Tebaldini, 2019. Ground cancellation with equalisation was applied following the algorithm described in d'Alessandro et al., 2020, where the estimated DTM was used for ground steering. A HOA of 100 m was used during equalisation for all nine stacks, which was found to be a good trade-off between ground rejection performance and susceptibility to

Table 2
Summary of the seven datasets used in this study, which were constructed from the nine SAR data stacks, individually or in pairs. Two datasets are available in Lopé: a single-heading dataset (Lo) consisting of four independent stacks made with different flight headings, and a dual-heading dataset (Lo-2), consisting of six different pairs of these stacks. The datasets for the other five test sites each consists of one single-heading stack. The table gives key information about the SAR and ALS data, the ALS-based AGB maps, and the sampling and calibration set-up of the CASINO algorithm. "Acq. date" is the date on which the stack of SAR data was acquired. "No of stacks" indicates the number of independent TomoSAR or PolInSAR stacks used (1 for single-heading datasets, 2 for dual-heading). "Stack size" is the number of co-registered, polarimetric SLC images within each stack. "Heading(s)" is the ϕ angle (or angles) depicted in Fig. 1 and measured relative to true north. "Inv. year(s)" is the *in situ* inventory year (or years). "No of plots" is the number of 1 ha *in situ* plots covered by the relevant stack or stacks (in parenthesis: the total number of plots available within the test site). "Plot stats" are the statistics for *in situ* plots covered by the relevant stack, given in the following format: mean (min, median, max). "Acq. year" is the year of ALS data acquisition. "Area" is the total area covered by ALS data and at least one SAR acquisition. "Dens." is the point density of the original ALS dataset (in parenthesis: ground return density). "Map stats" are the statistics for the area of the AGB map covered by either of the SAR acquisitions over the test site, given in the following format: mean (5th percentile, median, 95th percentile). "Size" is the dimension of the square sampling areas, while "Grid" is the separation of the sampling areas in both directions. N_{EST} and N_{CAL} are the number of estimation and calibration areas, respectively. "No of tests" is the total number of unique calibration datasets.

			SAR data						*In situ* data (1 ha plots)			ALS data & AGB map				Subsampling and calibration setup				
Dataset ID	Stack ID(s)	Test site	System	Acq. date	No of stacks	Stack size (s)	Inc. angles [°]	Heading (s) [°]	Inv. year(s)	No of plots	Plot stats [t/ha]	Acq. year	Area [km^2]	Dens. [m^{-2}]	Map stats [t/ha]	Size [m]	Grid [m]	N_{EST}	N_{CAL}	No of tests
No	1	Nouragues	ONERA SETHI	2009-08-24	1	5	33–49	179	2010, 2012	16 (33)	376 (255, 370, 577)	2012	13.3	19.9 (0.3)	360 (168, 381, 465)	150	250	236	2	500
Pa	2	Paracou				6	30–50	8	2009	79 (89)	359 (239, 368, 465)	2009	9.1	5.7 (0.1)	345 (294, 345, 398)		200	279		
Lo	3	Lopé	DLR F-SAR	2016-02-15		3	19–56	124	2016	4 (9)	179 (0, 172, 370)	2015	16.0	2.4 (0.1)	277 (6, 339, 436)		200	230		
	4							230		5 (9)	205 (0, 312, 370)						250	209		
	5							275									250	212		
	6							320									200	300		
Lo-2	3 & 4				2	3 & 3		124 & 230		4 (9)	179 (0, 172, 370)						200	214		
	3 & 5							124 & 275									200	223		
	3 & 6							124 & 320									200	174		
	4 & 5							230 & 275									250	204		
	4 & 6							230 & 320		5 (9)	205 (0, 312, 370)						200	282		
	5 & 6							275 & 320									200	299		
Ma	7	Mabounié		2016-02-11	1	4	21–57	180	2012	7 (12)	309 (175, 312, 465)	2007	77.1	4.3 (0.1)	307 (211, 311, 383)		550	222		
Mo	8	Mondah		2016-02-04		11	22–57	90	2016	10 (15)	94 (2, 74, 264)	2011	40.8	30.5 (2.3)	208 (31, 181, 434)		400	213		
Ra	9	Rabi		2016-02-07		4	23–51	310	2010–2012	25 (25)	314 (222, 310, 533)	2015	9.1	2.5 (0.05)	277 (184, 280, 358)		200	161		

ambiguities for the tropical forests studied in this paper. The data were then calibrated to σ^0 using the ψ angle calculated from the estimated DTM, and interpolated onto the same coordinate systems as the ALS-based AGB maps described in Section 3.3. All other angles (local incidence angle ϑ, slope angle u, and slope aspect angle v, see Fig. 1) were also calculated from the estimated DTM.

3.5. AGB estimates

The PLM parameters and AGB were estimated from single-heading data for all six test sites and the dual-heading data available only for Lopé.

3.5.1. Single-heading datasets (all test sites)

For each of the nine stacks, a square grid of square sampling areas was created within the region covered by both the SAR data and the AGB map. The size and spacing of the sampling areas were chosen to give between 150 and 300 areas per stack (found to be a good compromise between computational efficiency and AGB estimation performance) and at least 50 m separation between them (to decrease the risk of correlation between adjacent samples). This led to the sampling areas having size 150 m × 150 m for all stacks, and the spacing of their centres varying between 200 m × 200 m for the smaller stacks (Rabi, Paracou and two headings in Lopé) and 550 m × 550 m for the largest stack (Mabounié), see Table 2. For each of the sampling areas, CB and local incidence angle data were extracted from the images by averaging the corresponding nine pixels in the 50 m × 50 m images of σ^0_{cnp} (in linear units) and ϑ.

For each stack, 500 independent tests were carried out with different sets of calibration and estimation areas. In each test, two sampling areas were chosen at random and treated as calibration areas with AGB known from the ALS-based AGB maps. The AGB in the other sampling areas was then estimated. A different pair of calibration areas was used for each test and all calibration areas were required to have AGB > 100 t/ha. The number of estimation areas, N_{EST}, varied across the stacks (Table 2).

In each test, the PLM was fitted using the approach described in Section 2.2 to yield the nine PLM parameters (three parameters for each of the three polarisations) and N_{EST} estimates of AGB (in decibels), which were converted to linear units (t/ha) and corrected for logarithmic bias using (18). AGB maps with 50 m resolution were created from the CB and local incidence angle images using the approach described in Section 2.3.

The single-heading results were grouped into six datasets: one with 500 tests for each of the five test sites with one available flight heading (Nouragues, Paracou, Mabounié, Mondah and Rabi), and one for Lopé with 2000 tests, consisting of 500 tests for each of the four flight headings.

3.5.2. Dual-heading dataset (Lopé only)

For the dual-heading case, six pairs of stacks were available at Lopé. The estimation procedure was as in Section 3.5.1, except that the sampling areas were required to be covered by SAR data from both stacks, and the estimated AGB map was only created for the area of overlap between the two stacks. This resulted in fewer sampling areas and/or smaller separation between them than in the single-heading case. The results were grouped into a single dataset of 3000 tests, consisting of 500 tests for each of the six pairs.

4. Results

This section evaluates the performance of the CASINO approach using the data described in Section 3. Section 4.1 addresses the influence of calibration data on AGB estimation, while Section 4.2 investigates the impact of different AGB levels, local incidence angles, ground slope angles, and test sites. In Section 4.3, the AGB mapping potential is studied using selected tests with good AGB estimation results, while Section 4.4 assesses AGB estimation performance against *in situ* AGB estimates at 1 ha plots. The relation between AGB estimation performance and test site properties is studied in Section 4.5. Finally, the stability of the estimated PLM parameters is studied across test sites and polarisations in Section 4.6.

The following statistical measures, defined in Table 3, are used for quantitative analysis:

- bias (t/ha);
- absolute root-mean-square difference (RMSD, in t/ha);
- relative RMSD, measured in percent (%) of the average reference AGB;
- coefficient of determination (R^2).

4.1. Overall performance across calibration sets

For each dataset, each test yields values of AGB for each of the N_{EST} 150 m × 150 m (2.25 ha) estimation areas. Forming the difference between these estimates and the reference AGB in each estimation area allows the bias and RMSD for that test to be calculated. A distribution of bias, RMSD and relative RMSD is then formed from all tests for each of the seven datasets. These distributions are shown in Fig. 3 as boxplots, which indicate the median and the 5th, 25th, 75th, and 95th percentile values for each metric and each dataset.

For all datasets, the median bias is small (Fig. 3(a)). In the single-heading case, bias dispersion is smallest for Mondah and Lopé, where for 90% of the tests it lies in the range − 46 to 56 t/ha and − 56 to 63 t/ha, respectively. In Rabi and Mabounié, the spread is larger, and larger still for the French Guiana sites. In Nouragues, 90% of tests have bias between −143 t/ha and 85 t/ha, with similar values for Paracou. The dual-heading case for Lopé results in 90% of the tests having bias between −48 and 55 t/ha, which is better than for the single-heading case.

The smallest single-heading RMSDs are observed in Lopé and Mondah (Fig. 3(b)): for 75% of the tests, RMSD ≤ 68 t/ha for Mondah and RMSD ≤ 82 t/ha for Lopé. As before, the largest differences are seen in Nouragues and Paracou, with RMSD ≤ 120 t/ha and 116 t/ha, respectively, for 75% of the tests. For the dual-heading Lopé case, RMSD ≤ 71 t/ha is observed for 75% of the tests, which is significantly better than for the single-heading case in Lopé.

The relative RMSD is similar for all datasets (Fig. 3(c)). In the single-heading case, the median relative RMSD lies between 20% and 27% for all test sites. In Rabi, 75% of all tests have relative RMSD ≤ 29%, while the corresponding values are 32% for Mondah, 33% for Paracou and Lopé, 34% for Nouragues, and 35% for Mabounié. The dispersion is highest in Paracou and Mabounié, with 5% of the tests showing RMSD ≥ 61% and 56%, respectively. In the dual-heading Lopé case, the median relative RMSD is 22%, and 75% of the tests have an RMSD ≤ 28%, which is significantly better than the single-heading case.

Fig. 4 shows the distributions of bias and RMSD for the CASINO algorithm and two simplified cases in which the PLM parameters were estimated from CAL data only: one with constraints on PLM parameters

Table 3
Definitions of the performance metrics: W are the AGB estimates, W_0 are the reference AGB values, $\langle\cdot\rangle$ denotes averaging over the given set of measurements, and RMSD is root-mean-square difference.

Quantity	Unit	Formula
Bias	t/ha	$\langle W - W_0 \rangle$
Absolute RMSD	t/ha	$\sqrt{\langle (W - W_0)^2 \rangle}$
Relative RMSD	%	$\frac{100}{\langle W_0 \rangle} \sqrt{\langle (W - W_0)^2 \rangle}$
Coefficient of determination (R^2)	%	$100 \left(1 - \frac{\langle (W - W_0)^2 \rangle}{\langle (W_0 - \langle W_0 \rangle)^2 \rangle} \right)$

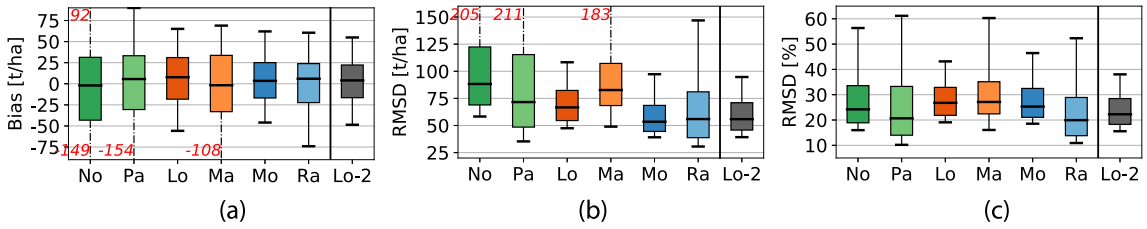

Fig. 3. Distributions of (a) bias in t/ha, (b) absolute RMSD in t/ha, and (c) relative RMSD in percent, across the seven datasets. The boxplots show the statistics for each of the three performance metrics (Table 3). The centrelines indicate the medians, boxes indicate the 25th and 75th percentiles, and whiskers indicate the 5th and 95th percentiles, calculated from 2000 (Lo), 3000 (Lo-2) or 500 (No, Pa, Ma, Mo, Ra) individual values for each metric. Whisker values outside the scale are indicated with red digits, while the black vertical lines separate the single-heading results from the dual-heading results.

and one without. To avoid unrealistic AGB estimates in both simplified cases, the obtained AGB values were truncated to the interval shown in Table 1. Due to the low number of CAL areas, the simplified cases could only be used with the dual-heading Lopé case (Lo-2, see Table 2). In both simplified cases, the results are worse than for the CASINO algorithm. For the simplified case with constraints on PLM parameters, the median bias is 8 t/ha and the bias lies between −147 and 66 t/ha for 90% of the calibration datasets. For 75% of the tests, the RMSD is below 89 t/ha (or 36% of the average AGB). When no constraints are applied, the performance is much worse: the median bias is significantly larger (48 t/ha), for 90% of the calibration datasets the bias is between −44 and 261 t/ha, and for 75% of the tests, the RMSD is below 210 t/ha (84%).

4.2. Performance under different conditions

Fig. 5 shows the dependence of bias and absolute and relative RMSD on AGB, local incidence angle, ground slope angle, and ground slope direction relative to the SAR antenna. Each variable is divided into three groups, as indicated in Fig. 5, and the performance metrics are calculated for each group and dataset, using all tests and available 150 m × 150 m estimation areas. Histograms of occurrence frequency of the variables across the groups are also shown. Values are only displayed in Fig. 5(a-l) when the occurrence frequency is at least 10%.

For all datasets except Nouragues and Mabounié, the bias is between ±25 t/ha and there is no strong dependence on any of the reference parameters (Fig. 5(a-d)). In Nouragues, AGB is significantly underestimated for high incidence angles ($\vartheta > 45°$), large slopes ($u > 8°$), and slopes away from the antenna. These three effects are connected: large, away-facing slopes will lead to an increased local incidence angle. Overestimation is observed for small slopes ($u < 4°$), which occur in around 10% of all samples. There is similar underestimation for away-facing slopes in Mabounié, but no clear trends with either incidence angle or ground slope. Note that Nouragues and Mabounié have the most marked topographic variation amongst all the test sites (Fig. 5(o)).

The RMSD tends to be higher in Nouragues, Paracou and Mabounié, which have the highest average AGB levels (Fig. 5(e-h) and Table 2), and it typically increases with increasing AGB, increasing ground slope, and decreasing incidence angles. In Lopé, the RMSD is about 10 t/ha lower for dual-heading than for single-heading data.

The relative RMSD has similar levels for all six test sites (Fig. 5(i-l)), and there is a clear dependence on AGB: for AGB < 200 t/ha, the relative RMSD is between 50% and 80% for Lopé and Mondah (the only test sites with a significant contribution to this group), while for AGB > 200 t/ha, the relative RMSD is between 18% and 33% for all six test sites. The use of dual-heading data at Lopé improves the relative RMSD from 25% and 21% for the medium- and high-AGB groups to 21% and 18%, respectively. For Mondah, relative RMSD increases from up to 25% for $\vartheta < 45°$ to 51% for $\vartheta > 45°$.

4.3. Examples of best-case performance

Figs. 6–8 show some of the best AGB mapping results for each of the seven datasets. This evaluation was done to investigate the potential of CASINO for mapping of spatial trends in AGB with only two CALs, as compared with the ALS-based AGB maps. For each dataset, a single test was manually selected for which the absolute bias ≤ 5 t/ha and relative RMSD ≤ 20%, and the spatial features visually matched each other as well as possible. The results are shown as maps and scatterplots, together with performance metrics, for both the 2.25 ha estimation areas and for the 0.25 ha pixels. Canopy backscatter for HV polarisation (s_{HV}) is also shown.

The visual features in the estimated and reference AGB maps for Nouragues are comparable (Fig. 6(a)), but there is large pixel-to-pixel variability and a tendency to overestimate AGB for AGB < 300 t/ha. Nevertheless, AGB estimation was successful for the two calibration areas in Nouragues, as indicated by the scatterplot in Fig. 6(a). The coefficient of determination for the 0.25 ha pixels is low (27%) and the RMSD is 84 t/ha (23%). For Paracou (Fig. 6(b)), the RMSD is only 63 t/ha (18%), but few features are correctly represented. However, the reference map covers only a narrow range of AGB, as indicated in

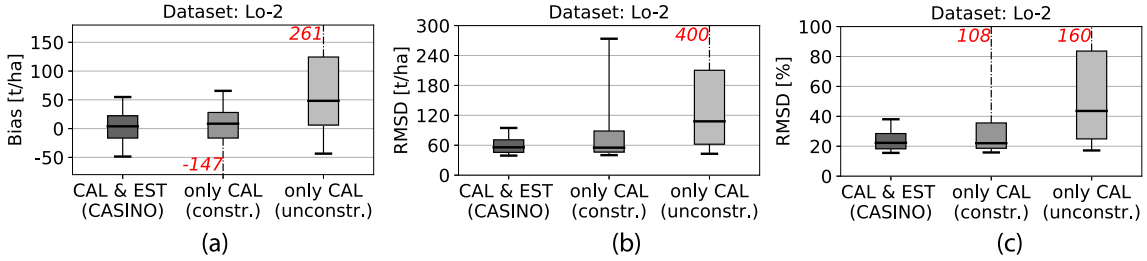

Fig. 4. Comparison between the CASINO algorithm (using both CAL and EST data to estimate AGB) and two simplified cases in which the PLM parameters are estimated directly from CAL data, with and without constraints on the PLM parameters. The figures show the distributions of (a) bias in t/ha, (b) absolute RMSD in t/ha, and (c) relative RMSD in percent, as defined in Table 3. The centrelines indicate the medians, boxes indicate the 25th and 75th percentiles, and whiskers indicate the 5th and 95th percentiles, calculated from 3000 individual values for each metric. Whisker values outside the scale are indicated with red digits. For the two simplified cases using only CAL data, AGB estimates were truncated to the interval given in Table 1. The results for CASINO are the same as those for the Lo-2 dataset in Fig. 3.

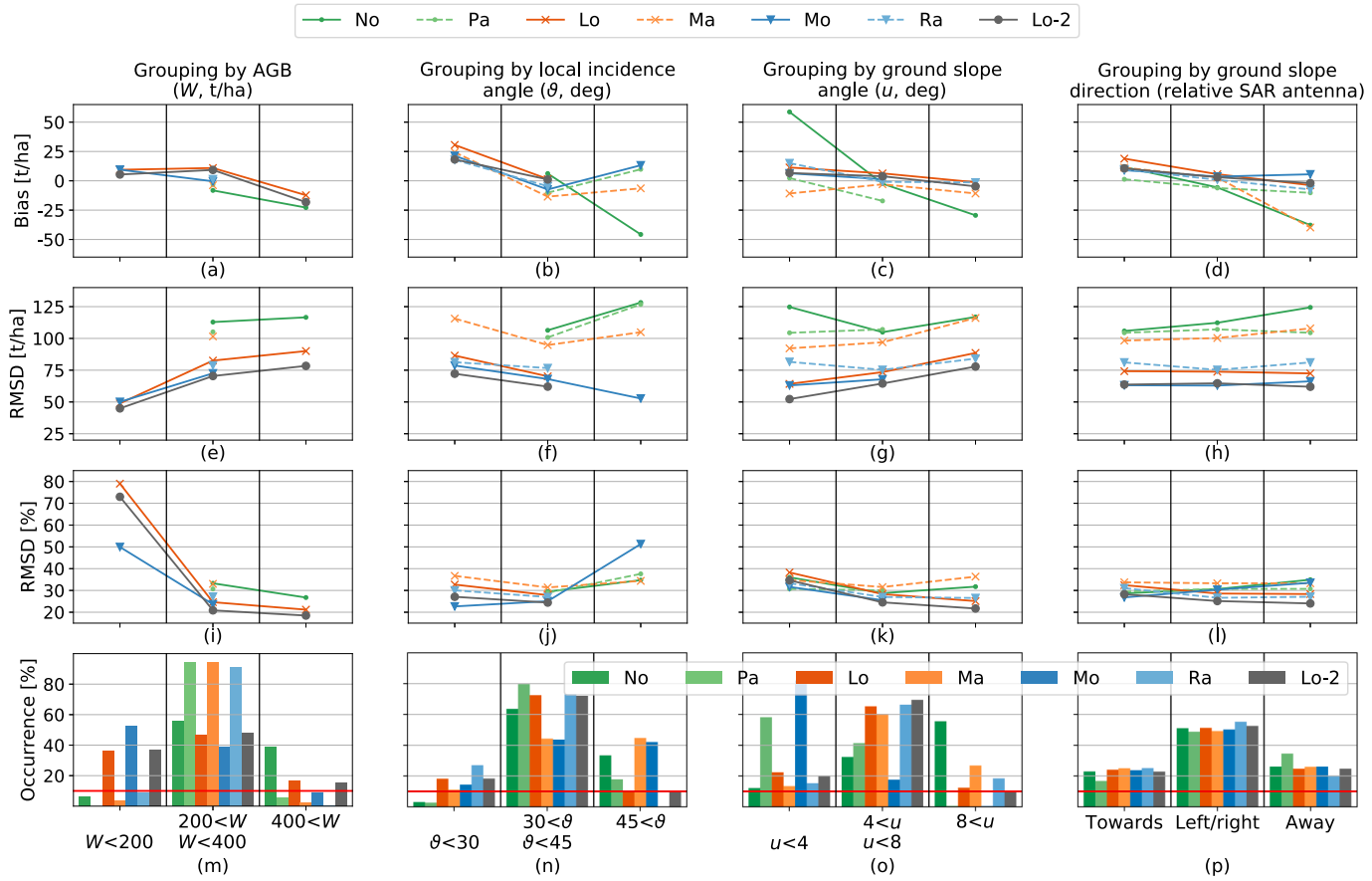

Fig. 5. The dependence on AGB, local incidence angle, ground slope angle, and ground slope direction (2.25 ha scale) for: (a-d) bias in t/ha; (e-h) absolute RMSD in t/ha; (i-l) relative RMSD in percent. (m-p) Distribution of samples across the groups. In (a-l), only results from groups with at least 10% of all samples are included; the 10% occurrence threshold is marked with red lines in (m-p).

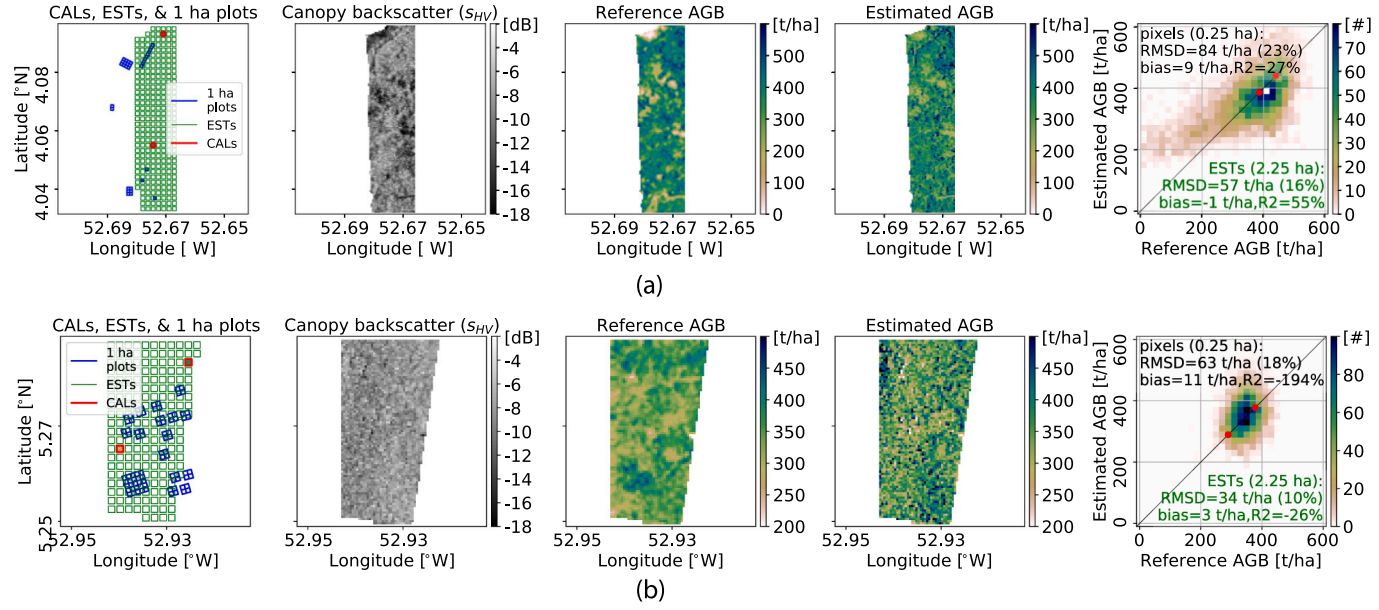

Fig. 6. Best-case AGB estimation results in (a) Nouragues and (b) Paracou, in French Guiana. The leftmost plots show the 150 m × 150 m estimation (EST) areas, two 150 m × 150 m calibration (CAL) areas, and the 1 ha *in situ* plots. The next three plots show the canopy backscatter for HV polarisation (s_{HV}), the reference AGB map from ALS data, and the best-case AGB map estimated from SAR with CASINO. The rightmost plots show 2D histograms created using all pixels from the reference and estimated AGB maps, with the colour scale indicating the number of pixels; these plots include performance metrics for the 2.25 ha EST areas (bottom right corner) and the 0.25 ha pixels (top left corner). The red dots in the scatterplots show estimation results for the two CAL areas. The SAR data were acquired with the left-looking SETHI system.

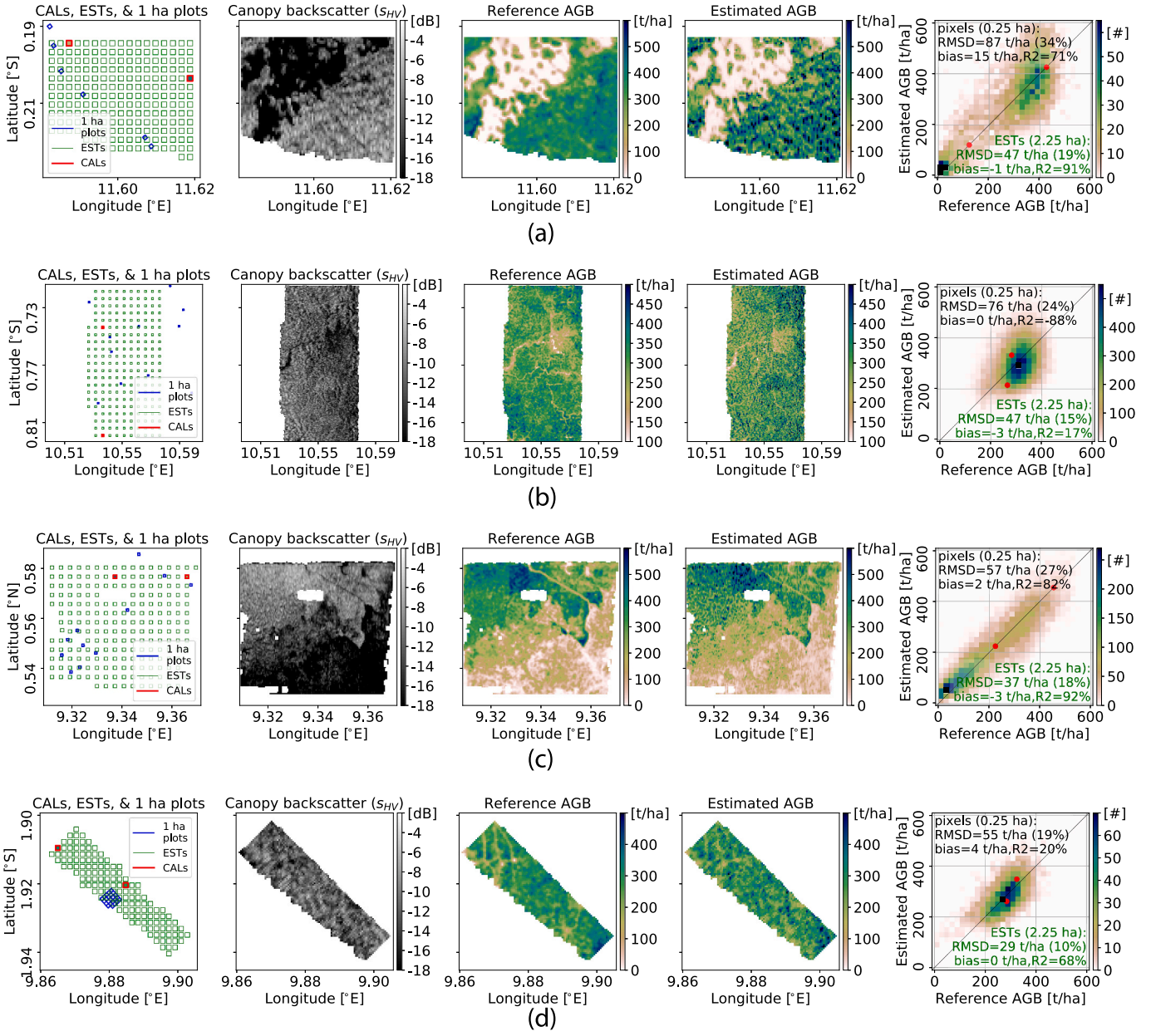

Fig. 7. Best-case AGB estimation results in (a) Lopé (heading: 275 deg), (b) Mabounié, (c) Mondah, and (d) Rabi, in Gabon. The leftmost plots show the 150 m × 150 m estimation (EST) areas, two 150 m × 150 m calibration (CAL) areas, and the 1 ha *in situ* plots. The next three plots show the canopy backscatter for HV polarisation (s_{HV}), the reference AGB map from ALS data, and the best-case AGB map estimated from SAR with CASINO. The rightmost plots show 2D histograms created using all pixels from the reference and estimated AGB maps, with the colour scale indicating the number of pixels; these plots include performance metrics for the 2.25 ha EST areas (bottom right corner) and the 0.25 ha pixels (top left corner). The red dots in the scatterplots show estimation results for the two CAL areas. The SAR data were acquired with the right-looking F-SAR system.

Table 2.

For the Gabon single-heading datasets (Fig. 7), the estimated AGB maps match most of the large-scale features within the reference scenes, but there is some residual local incidence angle dependence in Mabounié (Fig. 7(b)) and unexplained fine-scale noise-like variability in the estimated AGB maps for Lopé, Mabounié and Mondah. The RMSDs for the 0.25 ha pixels are 87 t/ha (34%) for Lopé, 76 t/ha (24%) for Mabounié, 57 t/ha (27%) for Mondah, and 55 t/ha (19%) for Rabi. For the two test sites with the largest AGB intervals, Mondah and Lopé, the coefficients of determination for pixel-level AGB estimates are 82% and 71%, respectively.

The dual-heading Lopé data (Fig. 8) yield better results than the single-heading data (Fig. 7(a)), with an RMSD for the 0.25 ha pixels of 60 t/ha (24%) and a coefficient of determination of 86%. The fine-scale variability is less pronounced and there is less variability for AGB > 300 t/ha.

4.4. *Assessment against in situ data*

Fig. 9 compares AGB estimates for the 1 ha plots shown in Figs. 6–7 and derived from: (i) *in situ* measurements (Section 3.2); (ii) ALS-based AGB maps (Section 3.3); and (iii) single-heading SAR-based AGB maps for the best-case tests (Section 4.4). The AGB estimates derived from the 50 m × 50 m maps were formed by first block-averaging to 100 m × 100 m and then linearly interpolating to the plot centres.

When compared with ALS-based AGB estimates, the best-case single-

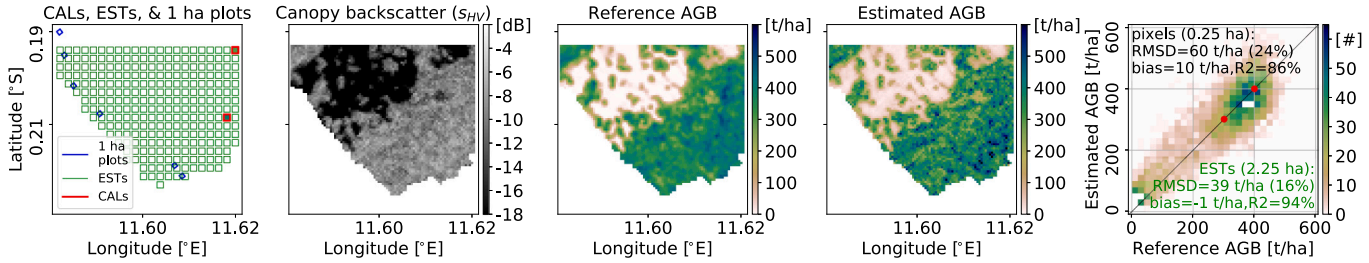

Fig. 8. Best-case AGB estimation results in Lopé, Gabon using dual-heading data (heading pair: 230° and 320°). The leftmost plot shows the 150 m × 150 m estimation (EST) areas, two 150 m × 150 m calibration (CAL) areas, and the 1 ha *in situ* plots. The next three plots show the canopy backscatter for HV polarisation (s_{HV}), the reference AGB map from ALS data, and the best-case AGB map estimated from SAR with CASINO. The rightmost plot shows 2D histograms created using all pixels from the reference and estimated AGB maps, with the colour scale indicating the number of pixels; these plots include performance metrics for the 2.25 ha EST areas (bottom right corner) and the 0.25 ha pixels (top left corner). The red dots in the scatterplots show estimation results for the two CAL areas. The SAR data were acquired with the right-looking F-SAR system.

heading SAR estimates for all six test sites give an RMSD of 46 t/ha (14%) and a coefficient of determination of 70% (Fig. 9(a)). When SAR-based estimates are compared with *in situ* AGB, the RMSD increases to 66 t/ha (20%) and the coefficient of determination decreases to 58% (Fig. 9(b)). Note that the comparison between ALS and *in situ* AGB is only slightly better (Fig. 9(c)), with an RMSD of 59 t/ha (18%) and a coefficient of determination of 67%. Both SAR and ALS overestimate AGB for nearly all plots when *in situ* AGB < 200 t/ha.

4.5. The impact of test site properties on AGB estimation performance

Sections 4.1–4.3 show that the best AGB estimation results were obtained at Lopé and Mondah, while the performance was worst at Paracou. Moreover, performance was affected by conditions such as AGB level and topography. Here the AGB estimation performance is assessed with respect to properties of each test site. For compactness, we only discuss relative RMSD for estimation areas where the reference AGB > 200 t/ha.

Fig. 10 shows scatterplots, for all six test sites, of median relative RMSD and its range between the 5th and 95th percentiles against: (i) average slope, calculated as the mean ground slope angle u; (ii) average AGB, calculated as the mean AGB; and (iii) AGB range, calculated as the difference between the 5th and 95th percentiles of AGB. Scatterplots in Fig. 10 show those cases where systematic behaviour was observed. The test site properties were calculated at 2.25 ha scale for all estimation areas, including those with AGB < 200 t/ha, and the dual-heading results for Lopé are included.

The median relative RMSD increases with increasing slope and increasing average AGB (Fig. 10(a-b)), while its range increases with increasing average AGB but decreases with increasing AGB range (Fig. 10(c-d)). The use of two flight headings at Lopé significantly reduces the median relative RMSD, but gives only a small improvement in the relative RMSD range (although the single heading Lopé estimates already have the smallest range amongst all sites).

4.6. Parameter variability across sites

Fig. 11 shows the distributions of the estimated PLM parameters for different datasets and polarisations, together with the permissible parameter intervals (Table 1). In the single-heading case, 2000 PLM parameter sets were derived for Lopé and 500 sets for the remaining five test sites. In the dual-heading Lopé case, 3000 sets were derived.

For Nouragues, Lopé, Mabounié and Rabi, the median values for A_{PQ} lie between −45 dB and − 30 dB, while for Paracou and Mondah, they are occasionally higher (Fig. 11(a-c)). For all test sites except Paracou, A_{PQ} is a few dB lower for HV than for HH and VV. No clear differences are observed between the single- and dual-heading cases in Lopé.

For all seven datasets, the median sensitivity, α_{PQ}, of CB to AGB is typically between 0.5 and 1.5 dB/dB (Fig. 11(d-f)). The lowest sensitivity and largest dispersion occur for Paracou. There is little dependence on polarisation except for Paracou, where VV has lower sensitivity to AGB than both HH and HV. The differences between the single- and dual-heading cases in Lopé are insignificant.

For all single-heading datasets except Paracou, the highest values for n_{PQ} are obtained for HH, mainly between 2 and 3, and the maximum permissible value of 3 is often reached (Fig. 11(g-i)), while $n_{PQ} \approx 2$ for HV and VV. For Paracou, the lowest median value (0.6) is observed for HH, a higher value (1.2) is observed for HV, and the highest value (2.7) is observed for VV. For Lopé, n_{PQ} is clearly different between the single- and dual-heading cases, with respective values of 2.7 and 1.5 for HH, 1.8

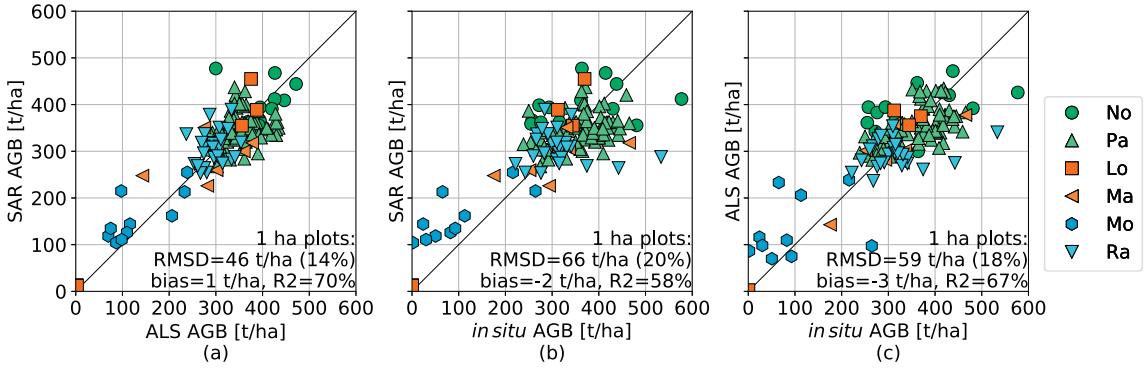

Fig. 9. Scatterplots showing pairwise comparisons of best-case single heading SAR-, ALS- and *in situ*-based AGB estimates for the 1 ha plots located within the six test sites, together with the associated performance metrics (bottom right corner). The SAR and ALS maps and *in situ* plots are shown in Figs. 6–7. AGB estimates were extracted from SAR and ALS maps by first averaging them to 100 m × 100 m and then linearly interpolating the AGB data to plot centres. Subplots show: (a) SAR AGB *vs.* ALS AGB, (b) SAR AGB *vs. in situ* AGB, and (c) ALS AGB *vs. in situ* AGB.

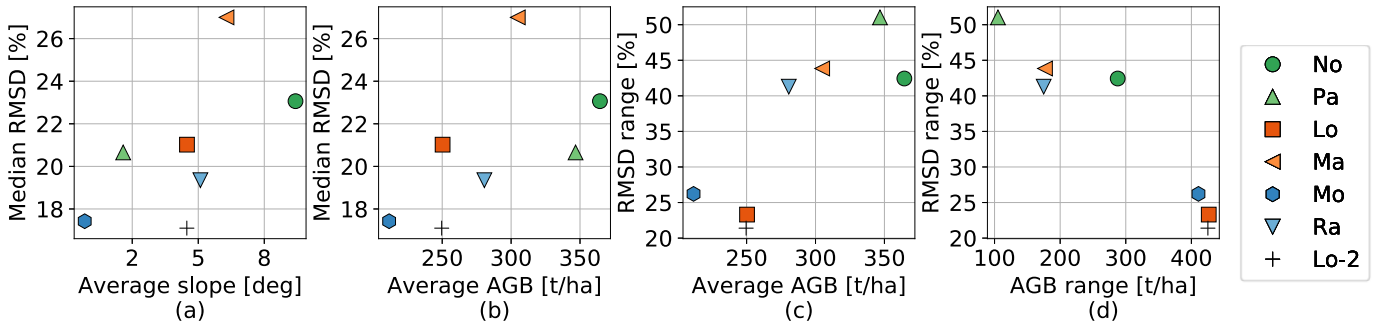

Fig. 10. Scatterplots of median RMSD against (a) average ground slope angle and (b) average AGB and relative RMSD range against (c) average AGB and (d) AGB range for all 6 test sites. The AGB statistics were calculated for estimation areas with AGB > 200 t/ha and the average test site properties were estimated at 2.25 ha scale.

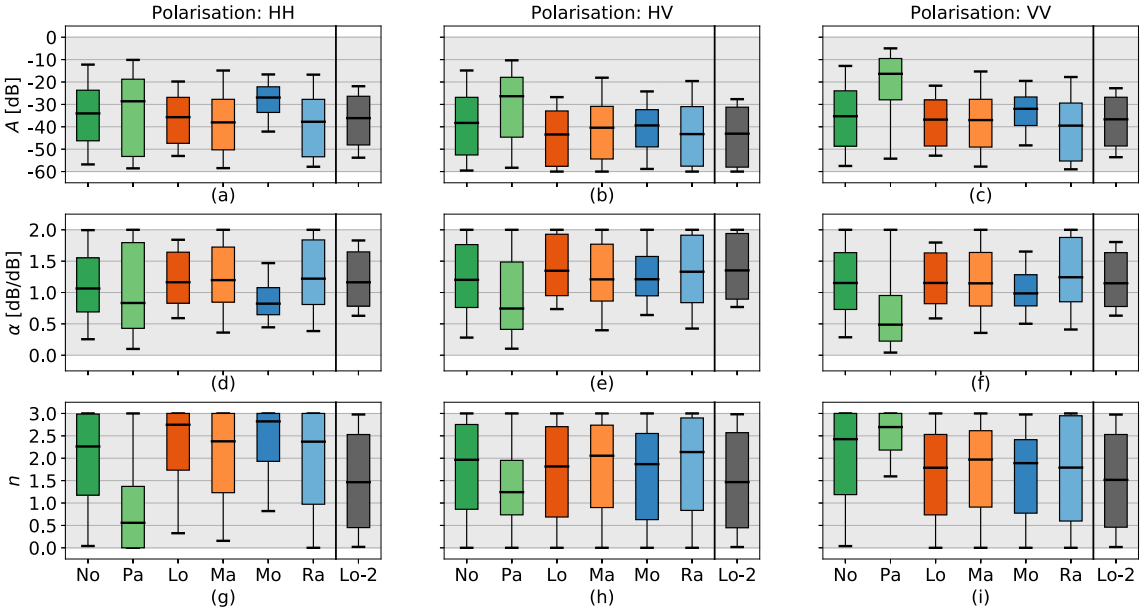

Fig. 11. Distributions of the estimated PLM parameters across datasets and polarisations. The number of PLM parameter sets varied from 500 (No, Pa, Ma, Mo, Ra) to 2000 (Lo) and 3000 (Lo-2). The centrelines indicate the medians, boxes indicate the 25th and 75th percentiles, and whiskers indicate the 5th and 95th percentiles. The grey background indicates the permissible parameter intervals (Table 1), while the black vertical lines separate the single-heading and dual-heading results.

and 1.5 for HV, and 1.8 and 1.5 for VV.

5. Discussion

Section 4 indicates that using only two calibration areas with reference AGB > 100 t/ha, the CASINO algorithm can provide AGB estimation results with low bias and a relative RMSD at 2.25 ha resolution better than 22% in 25% of the tests, better than 27% in 50% of the tests, and better than 35% in 75% of the tests. These results are better than when the power law model is fitted directly to the calibration data without subsampling or constraints on AGB (Fig. 4). Other studies using P-band SAR in tropical forests can achieve similar AGB estimation performance only by using much larger reference datasets for model training (Villard and Le Toan, 2015) and, in the case of Minh et al., 2014b and Tebaldini et al. (2019), SAR tomography.

5.1. Calibration and validation methodology

Throughout this paper, we used the term "root-mean-square difference (RMSD)" rather than "root-mean-square error (RMSE)" to highlight the fact that the reference data are themselves estimates and prone to measurement errors. Under the assumption that the errors in the two types of estimate are uncorrelated, the expected value of the RMSD^2 is given by

$$\langle \mathrm{RMSD}^2 \rangle = (b_e - b_r)^2 + V_e + V_r \quad (19)$$

where b_e and b_r are the biases in the AGB estimates and the reference data respectively, and V_e and V_r are the corresponding error variances. Eq. (19) indicates that the accuracy of the estimates is likely to be better than the values suggested by the RMSD, except in cases when bias is the dominant source of error.

For calibration, this study used AGB maps derived from ALS data, while validation used both the AGB maps and *in situ* measurements. Calibration with ALS-based AGB maps allowed us to assess the performance of the CASINO algorithm against a variety of different calibration datasets and to investigate whether it reproduced spatial features seen in the reference AGB maps. However, the ALS estimates of AGB are prone to significant biases, *e.g.*, in topographic areas (Labrière et al., 2018; Leitold et al., 2015), which may be transferred to the SAR-based maps when the ALS data are used for calibration. This manifested itself as similar biases observed when the ALS- and SAR-based AGB estimates were compared with 1 ha plot measurements (Fig. 9). Meanwhile, direct comparison between ALS- and SAR-based estimates showed closer agreement. For this reason and in line with Chave et al., 2019, we

recommend that the global AGB estimation scheme for BIOMASS relies on calibration with AGB estimates from *in situ* inventories, which are less prone to systematic errors (although random errors with standard deviation of 10% or more are possible in the tropics, see Chave et al., 2005). Note that this study disregarded temporal offsets between the SAR, ALS and/or *in situ* measurements, which could be as high as nine years (for the ALS and SAR data in Mabounié). Therefore, temporal changes due to logging, forest degradation and growth could have contributed to the observed discrepancies.

5.2. Variability across test sites

Overall, the test sites in French Guiana gave poorer results than those in Gabon, but it is unclear if this reflects different properties of South American and African forests or the particularities of each test site, most notably topography and AGB distribution within the scene. The least ideal conditions were in French Guiana: Nouragues has the most significant topographic variation, while Paracou has the narrowest range of AGB. The two test sites with the smallest topographic variation and largest AGB intervals were both in Gabon (Lopé and Mondah).

The best and most consistent performance across different calibration datasets was found for test sites with a large AGB interval within the scene and an average AGB around 200–250 t/ha. Large AGB intervals give a wide range of canopy backscatter, making it easier to fit the PLM to the sampled data. In addition, the sensitivity to AGB tends to be larger away from the highest levels of AGB. The large variability in estimation performance with different calibration datasets in Paracou is mainly due to the narrow AGB interval and high average AGB. In contrast, there is good performance at Lopé and Mondah, where the AGB intervals are large ($>$ 400 t/ha) and the average AGB levels are in the interval 200–250 t/ha.

AGB estimation performance was worse in areas of significant topography and/or AGB $<$ 200 t/ha. This effect was most prominent in Nouragues, where a clear overestimation trend was observed at low AGB levels (Fig. 6(a)). Topographic variation is well-known to hinder low-frequency SAR mapping of forests. Even after ground cancellation, some topographic effects remain, especially in sparse forests where attenuation by the forest canopy is small. The selection of the height-of-ambiguity (HOA) during equalised ground cancellation also plays a role (d'Alessandro et al., 2020). For consistency, a constant HOA of 100 m was used throughout the paper, emphasising scattering from objects located about 50 m above ground. However, in forest whose height is significantly less than 50 m (which is typical when AGB $<$ 200 t/ha), the relative contribution of the canopy is then reduced, which is likely to worsen the AGB estimation. This suggests that forest height, which can be estimated from the data used in ground cancellation, should be used to select the HOA best matched to the forest being observed. Additionally, the PLM only contains a simple, first order correction for topographic effects, which is not effective for the test sites with most marked topography (Nouragues and Mabounié) (Small, 2011; Villard and Le Toan, 2015). In the case of Nouragues (Fig. 6(a)), the PLM was not able to correctly model AGB variability in the ground-cancelled data across the entire AGB interval, so it adapted to the AGB levels of the calibration data. Further work with ground cancellation and topographic corrections in the PLM is therefore needed to improve algorithm performance in hilly terrain.

This study also showed relatively large intervals of estimated PLM parameter values (Fig. 11). However, it is difficult to compare these values due to the varying forest properties (moisture, wood density and structure), AGB distributions, topographic conditions, and SAR systems used in the AfriSAR and TropiSAR campaigns. Physical interpretation of the parameters is also challenging, as the PLM is a simplified case of a more complex scattering model (Truong-Loï et al., 2015), generalised to cover both the low- and high-attenuation cases. As shown in Banda et al., 2020, the same PLM parameters can in some cases provide good AGB estimation results at different test sites, while in other cases the estimated parameters are clearly site-specific. In the future, algorithmic stability of CASINO could be improved through stricter constraints on the PLM parameters, *e.g.*, based on their physical interpretation and/or empirical outcome.

5.3. Algorithm implementation

Even with a single flight heading, the AGB estimation performance at Lopé was better than for most other datasets, but it improved further when two headings were combined. This was mainly because of better estimation of the parameter n_{PQ}, which is related to topographic correction, and which was the only parameter with significantly different statistics when two headings were used. In contrast, using two headings did not decrease the variability in the AGB estimates induced by different calibration datasets. This is important for BIOMASS, which will acquire both ascending and descending data, but their separation in time may cause variation in the PLM parameters that will need to be accounted for in the global algorithm.

We also investigated whether there are any clear criteria that calibration areas must satisfy. For compactness, this paper focussed on AGB estimation with only two calibration areas. A larger number of calibration areas will improve estimation performance and bring it closer to the median values shown in Fig. 3. Also, with more calibration areas, multiple tests with subsets of calibration data can be used to provide insight into AGB estimate variability. Based on empirical tests, an important requirement on calibration data was that AGB $>$ 100 t/ha, largely because forests of less 100 t/ha are typically sparse and/or short. As discussed earlier, ground cancellation with 100 m HOA will not be effective in this case and the measured backscatter will contain a significant contribution from the ground. No other investigated criteria led to significant improvements in AGB estimation performance.

This paper highlights the importance of including many estimation samples with unknown AGB in the cost function. At least two reasons emerge: it allows us to apply constraints on both the PLM parameters and the predicted AGB, and it allows the algorithm to separate incidence angle trends from AGB trends, without requiring extensive calibration data. The latter occurs because with many estimation samples, it is likely that there are multiple measurements with the same AGB but different incidence angles. If there are at least two estimation samples with the same AGB, then q_2 in (13) can be estimated without knowing w_0, *i.e.*, the reference AGB. Therefore, the CASINO algorithm can estimate three model parameters per polarisation using only two reference AGB measurements and single-heading data. Although in the dual-heading case the PLM parameters could be estimated directly from the calibration data (Fig. 4), the CASINO algorithm still provided better results because constraints could be applied to the estimated AGB.

An important issue is to identify pathways to improve the performance of the algorithm in the lead-up to the launch of BIOMASS, and at least four possible avenues should be explored:

- *Exploit more fully the physical content of the scattering model:* A tacit assumption of the approach used in this paper is that the model given by (5) is a complete description of the scattering from the canopy, so all the physical effects are captured by the parameters of the model. However, as explained in Truong-Loï et al. (2015), this is a simplification of more complex models in which the links to the scattering physics (dielectric and geometry) are made explicit (*e.g.*, Saatchi and Moghaddam, 2000). It may be productive to evaluate whether use of a more complete model can provide better AGB estimation performance and/or improve understanding and mitigation of topographic and moisture effects (Soja et al., 2013).
- *Impose other constraints on the problem:* Santoro and Fransson, 2019, describe an approach in which a variety of ancillary datasets are used to adapt the parameters of the water cloud model (Attema and Ulaby, 1978) to local or regional conditions; a similar approach may be possible with the model used here. Other BIOMASS products and

remote sensing data may be used to constrain model parameters and/or for calibration, including forest height and tomographic profiles from BIOMASS and spaceborne lidar data from the Global Ecosystem Dynamics Investigation (GEDI) mission (Dubayah et al., 2020).

- *Expand the datasets available for algorithm calibration:* The range of available *in situ* and ALS data is continually expanding, not least because of the calibration and validation requirements of space missions devoted to forest structure and biomass (Chave et al., 2019; Herold et al., 2019). It is essential to keep abreast of these developments and work them into a global approach to algorithm calibration for BIOMASS.
- *Study and improve optimisation approaches:* We have not investigated whether the observed performance of the algorithm is affected by the particular solution scheme being used or is an inherent feature of the problem.

6. Conclusion

The analysis in this paper indicates that the CASINO algorithm for AGB estimation from P-band SAR can provide reliable estimates of AGB with as few as two calibration measurements. The best results were obtained for areas with large AGB variability and small topographic variation.

These observations provide the basis for the CASINO algorithm currently implemented in the prototype ground processor of the BIOMASS satellite (Banda et al., 2020). In fact, the BIOMASS data will have properties that are better suited to CASINO than the airborne data used here, including lower variability in incidence angle, more stable baselines and larger coverage, resulting in less spatial variability caused by the SAR system and more variability induced by the forest (*e.g.*, larger AGB variability). However, BIOMASS data will also present new challenges: the incidence angles will be steeper than for airborne geometries, giving higher levels of ground scattering and more pronounced topographic effects, while the larger coverage will result in more spatial variability of moisture and forest type. Ground cancellation performance will also differ for BIOMASS, due to different geometries and acquisition patterns. Further research should focus on improving AGB estimation performance in hilly areas and forests with AGB < 200 t/ha. Additionally, follow-up studies should address AGB estimation in temperate and boreal forests, as well as the effects of canopy moisture variations.

Although further work is needed before P-band SAR can provide the global wall-to-wall AGB maps needed by various stakeholders (Herold et al., 2019), CASINO is a viable initial approach, with good potential for future development and which minimises the reliance on large quantities of reference data.

Declaration of Competing Interest

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

Acknowledgements

This work was funded by the European Space Agency under contract 4000119231/167/NL/CT: “Biomass implementation study”. The entire project team is acknowledged for their contributions to this work. The authors would also like to thank Dr. Nicolas Labrière for his comments about the field data and the three anonymous reviewers for their excellent feedback and suggestions.

References

Albinet, C., Borderies, P., Koleck, T., Rocca, F., Tebaldini, S., Villard, L., Le Toan, T., Hamadi, A., Ho Tong Minh, D., 2012. TropiSCAT: a ground based Polarimetric Scatterometer experiment in tropical forests. IEEE J. Sel. Top. Appl. Earth Observat. Remote Sensing 5, 1060–1066. https://doi.org/10.1109/JSTARS.2012.2201917.

Askne, J.I.H., Dammert, P.B.G., Ulander, L.M.H., Smith, G., 1997. C-band repeat-pass interferometric SAR observations of the forest. IEEE Transactions on Geoscience and Remote Sensing 35 (1), 25–35. https://doi.org/10.1109/36.551931.

Askne, J.I.H., Fransson, J.E.S., Santoro, M., Soja, M.J., Ulander, L.M.H., 2013. Model-based biomass estimation of a hemi-boreal forest from multitemporal TanDEM-X acquisitions. Remote Sens. 5, 5574–5597. https://doi.org/10.3390/rs5115574.

Attema, E.P.W., Ulaby, F.T., 1978. Vegetation modeled as a water cloud. Radio Sci. 13, 357–364. https://doi.org/10.1029/RS013i002p00357.

Baker, T.R., Phillips, O.L., Malhi, Y., Almeida, S., Arroyo, L., Di Fiore, A., Erwin, T., Killeen, T., Laurance, S.G., Laurance, W.F., Lewis, S.L., Lloyd, J., Vásquez Martínez, R., 2004. Variation in wood density determines spatial patterns in Amazonian forest biomass. Global Change Biol. 10 (5) https://doi.org/10.1111/j.1365-2486.2004.00751.x.

Bamler, R., Hartl, P., 1998. Synthetic aperture radar interferometry. Inverse Problems 14 (4), R1–R54.

Banda, F., Giudici, D., Le Toan, T.L., Mariotti d’Alessandro, M.M., Papathanassiou, K., Quegan, S., Riembauer, G., Scipal, K., Soja, M., Tebaldini, S., Ulander, L., Villard, L., 2020. The BIOMASS level 2 prototype processor: design and experimental results of above-ground biomass estimation. Remote Sens. 12, 985. https://doi.org/10.3390/rs12060985.

Blomberg, E., Ferro-Famil, L., Soja, M.J., Ulander, L.M.H., Tebaldini, S., 2018. Forest biomass retrieval from L-band SAR using tomographic ground backscatter removal. IEEE Geosci. Remote Sens. Lett. 15 (7), 1030–1034. https://doi.org/10.1109/LGRS.2018.2819884.

Brolly, M., Woodhouse, I.H., 2012. A matchstick model of microwave backscatter from a forest. Ecol. Model. 237–238, 74–87. https://doi.org/10.1016/j.ecolmodel.2012.04.014.

Carreiras, J.M.B., Quegan, S., Le Toan, T., Ho Tong Minh, D., Saatchi, S.S., Carvalhais, N., Reichstein, M., Scipal, K., 2017. Coverage of high biomass forests by the ESA BIOMASS mission under defense restrictions. Remote Sens. Environ. 196, 154–162. https://doi.org/10.1016/j.rse.2017.05.003.

Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.-P., Nelson, B.W., Ogawa, H., Puig, H., Rièra, B., Yamakur, T., 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forest. Oecologia 145, 87–99. https://doi.org/10.1007/s00442-005-0100-x.

Chave, J., Olivier, J., Bongers, F., Châtelet, P., Forget, P.-M., van der Meer, P., Norden, N., Riéra, B., Charles-Dominique, P., 2008. Above-ground biomass and productivity in a rain forest of eastern South America. J. Trop. Ecol. 24, 355–366. https://doi.org/10.1017/S0266467408005075.

Chave, J., Davies, S.J., Phillips, O.L., Lewis, S.L., Sist, P., Schepaschenko, D., Armston, J., Baker, T.R., Coomes, D., Disney, M., Duncanson, L., Hérault, B., Labrière, N., Meyer, V., Réjou-Méchain, M., Scipal, K., Saatchi, S., 2019. Ground data are essential for biomass remote sensing missions. Surv. Geophys. 40, 863–880. https://doi.org/10.1007/s10712-019-09528-w.

Cloude, S.R., Papathanassiou, K.P., 1998. Polarimetric SAR interferometry. IEEE Trans. Geosci. Remote Sens. 36 (5), 1551–1565. https://doi.org/10.1109/36.718859.

d’Alessandro, M.M., Tebaldini, S., 2019. Digital terrain model retrieval in tropical forests through P-band SAR tomography. IEEE Trans. Geosci. Remote Sensing 57 (9), 6774–6781. https://doi.org/10.1109/TGRS.2019.2908517.

d’Alessandro, M.M., Tebaldini, S., Quegan, S., Soja, M.J., Ulander, L.M.H., Scipal, K., 2020. Interferometric ground cancellation for above ground biomass estimation. IEEE Trans. Geosci. Remote Sensing 58 (9), 6410–6419. https://doi.org/10.1109/TGRS.2020.2976854.

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The global ecosystem dynamics investigation: high-resolution laser ranging of the Earth’s forests and topography. Science of Remote Sensing 1, 100002. https://doi.org/10.1016/j.srs.2020.100002.

Dubois-Fernandez, P., Le Toan, T., Chave, J., Blanc, L., Daniel, S., Oriot, H., Arnaubec, A., Rejou-Mechain, M., Villard, L., Lasne, Y., Koleck, T., 2011. TropiSAR 2009: technical assistance for the development of airborne SAR and geophysical measurements during the TropiSAR 2009 experiment (final report). European Space Agency, Noordwijk, The Netherlands. ESA contract no. 22446/09. CNES contract no. 9292903/08/09.

Dubois-Fernandez, P.C., Toan, T.L., Daniel, S., Oriot, H., Chave, J., Blanc, L., Villard, L., Davidson, M., Petit, M., 2012. The TropiSAR airborne campaign in French Guiana: objectives, description, and observed temporal behavior of the backscatter signal. IEEE Trans. Geosci. Remote Sens. 50 (8), 3228–3241. https://doi.org/10.1109/TGRS.2011.2180728.

European Space Agency, 2008. Candidate Earth Explorer core missions – Reports for assessment: BIOMASS. ESA SP-1313/2 (6 volume series). European Space Agency. Noordwijk, The Netherlands.

European Space Agency, 2012. Report for mission selection: BIOMASS. ESA SP-1324/1 (3 volume series). European Space Agency. Noordwijk, The Netherlands.

Finney, D.J., 1941. On the distribution of a variate whose logarithm is normally distributed. Suppl. J. R. Stat. Soc. 7 (2), 155–161. https://doi.org/10.2307/2983663.

Friedlingstein, P., Jones, M.W., O’Sullivan, M., Andrew, R.M., Hauck, J., Peters, G.P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D.C.E., Canadell, J.G., Ciais, P., Jackson, R.B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L.P., Currie, K.I., Feely, R. A., Gehlen, M., Gilfillan, D., Gkritzalis, T., Goll, D.S., Gruber, N., Gutekunst, S.,

Harris, I., Haverd, V., Houghton, R.A., Hurtt, G., Ilyina, T., Jain, A.K., Joetzjer, E., Kaplan, J.O., Kato, E., Klein Goldewijk, K., Korsbakken, J.I., Landschützer, P., Lauvset, S.K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Marland, G., McGuire, P.C., Melton, J.R., Metzl, N., Munro, D.R., Nabel, J.E.M.S., Nakaoka, S.-I., Neill, C., Omar, A.M., Ono, T., Peregon, A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P.P., Tian, H., Tilbrook, B., Tubiello, F.N., van der Werf, G.R., Wiltshire, A.J., Zaehle, S., 2019. Global carbon budget 2019. Earth Syst. Sci. Data 11 (4), 1783–1838. https://doi.org/10.5194/essd-11-1783-2019.

Garestier, F., Dubois-Fernandez, P.C., Champion, I., 2008. Forest height inversion using high-resolution P-band Pol-InSAR data. IEEE Trans. Geosci. Remote Sens. 46 (11), 3544–3559. https://doi.org/10.1109/TGRS.2008.922032.

Gatti, G., Tebaldini, S., Mariotti d'Alessandro, M., Rocca, F., 2011. ALGAE: a fast algebraic estimation of Interferogram phase offsets in space-varying geometries. IEEE Trans. Geosci. Remote Sens. 49 (6), 2343–2353. https://doi.org/10.1109/TGRS.2010.2091278.

GCOS, 2016. The Global Observing System for Climate: Implementation Needs. https://doi.org/10.13140/RG.2.2.23178.26566.

Hajnsek, I., Scheiber, R., Keller, M., Horn, R., Lee, S.-K., Ulander, L.M.H., Gustavsson, A., Sandberg, G., Le Toan, T., Tebaldini, S., Guarnieri, A.M., Rocca, F., 2009b. BioSAR 2008: Technical assistance for the development of airborne SAR and geophysical measurements during the BioSAR 2008 experiment (final report). European Space Agency, Noordwijk, The Netherlands. ESA contract no. 22052/08/NL/CT.

Hajnsek, I., Scheiber, R., Ulander, L.M.H., Gustavsson, A., Sandberg, G., Tebaldini, S., Guarnieri, A.M., Rocca, F., Lombardini, F., Pardini, M., 2008. BioSAR 2007: technical assistance for the development of airborne SAR and geophysical measurements during the BioSAR 2007 experiment (final report). European Space Agency, Noordwijk, The Netherlands. ESA contract no. 20755/07/NL/CB.

Hajnsek, I., Kugler, F., Lee, S.-K., Papathanassiou, K.P., 2009a. Tropical-forest-parameter estimation by means of Pol-InSAR: the INDREX-II campaign. IEEE Trans. Geosci. Remote Sens. 47 (2), 481–493. https://doi.org/10.1109/TGRS.2008.2009437.

Hajnsek, I., Pardini, M., Jäger, M., Horn, R., Kim, J.-S., Jörg, H., Papathanassiou, K.P., Dubois-Fernandez, P., Depuis, X., Wasik, V., 2017. Technical assistance for the development of airborne SAR and geophysical measurements during the AfriSAR experiment (final report). European Space Agency, Noordwijk, The Netherlands. ESA contract no. 4000114293/15/NL/CT.

Herold, M., Carter, S., Avitabile, V., Espejo, A.B., Jonckheere, I., Lucas, R., McRoberts, R. E., Næsset, E., Nightingale, J., Petersen, R., Reiche, J., Romijn, E., Rosenqvist, A., Rozendaal, D.M.A., Seifert, F.M., Sanz, M.J., De Sy, V., 2019. The role and need for space-based forest biomass-related measurements in environmental management and policy. Surv. Geophys. 40, 757–778. https://doi.org/10.1007/s10712-019-09510-6.

Labrière, N., Tao, S., Chave, J., Scipal, K., Le Toan, T., Abernethy, K., Alonso, A., Barbier, N., Bissiengou, P., Casal, T., Davies, S.J., Ferraz, A., Herault, B., Jaouen, G., Jeffery, K.J., Kenfack, D., Korte, L., Lewis, S.L., Malhi, Y., Memiaghe, H.R., Poulsen, J.R., Rejou-Mechain, M., Villard, L., Vincent, G., White, L.J.T., Saatchi, S., 2018. In situ reference datasets from the TropiSAR and AfriSAR campaigns in support of upcoming Spaceborne biomass missions. IEEE J. Sel. Top. Appl Earth Observations Remote Sensing 11 (10), 3617–3627. https://doi.org/10.1109/JSTARS.2018.2851606.

Le Toan, T., Quegan, S., Davidson, M.W.J., Balzter, H., Paillou, P., Papathanassiou, K., Plummer, S., Rocca, F., Saatchi, S., Shugart, H., Ulander, L.M.H., 2011. The BIOMASS mission: mapping global forest biomass to better understand the terrestrial carbon cycle. Remote Sens. Environ. 115 (11), 2850–2860. https://doi.org/10.1016/j.rse.2011.03.020.

Leitold, V., Keller, M., Morton, D.C., Cook, B.D., Shimabukuro, Y.E., 2015. Airborne lidar-based estimates of tropical forest structure in complex terrain: opportunities and trade-offs for REDD+. Carbon Balance Manag. 10 (3), 1–12. https://doi.org/10.1186/s13021-015-0013-x.

Minh, D.H.T., Tebaldini, S., Rocca, F., Toan, T.L., Borderies, P., Koleck, T., Albinet, C., Hamadi, A., Villard, L., 2014a. Vertical structure of P-band temporal decorrelation at the Paracou forest: results from TropiScat. IEEE Geosci. Remote Sens. Lett. 11 (8), 1438–1442. https://doi.org/10.1109/LGRS.2013.2295165.

Minh, D.H.T., Toan, T.L., Rocca, F., Tebaldini, S., d'Alessandro, M.M., Villard, L., 2014b. Relating P-band synthetic aperture radar tomography to tropical forest biomass. IEEE Trans. Geosci. Remote Sens. 52 (2), 967–979. https://doi.org/10.1109/TGRS.2013.2246170.

Monteith, A.R., Ulander, L.M.H., 2018. Temporal survey of P- and L-band polarimetric backscatter in boreal forests. IEEE J. Sel. Top. Appl Earth Observations Remote Sensing 11 (10), 3564–3577. https://doi.org/10.1109/JSTARS.2018.2814825.

Nocedal, J., Wright, S., 2006. Quasi-Newton methods. In: Numerical optimization. Springer, New York, NY.

Papathanassiou, K.P., Cloude, S.R., 2001. Single-baseline polarimetric SAR interferometry. IEEE Trans. Geosci. Remote Sens. 39 (11), 2352–2363. https://doi.org/10.1109/36.964971.

Phillips, O.L., Sullivan, M.J.P., Baker, T.R., Monteagudo Mendoza, A., Vargas, P.N., Vásquez, R., 2019. Species matter: wood density influences tropical Forest biomass at multiple scales. Surv. Geophys. 40, 913–935. https://doi.org/10.1007/s10712-019-09540-0.

Quegan, S., Le Toan, T., Chave, J., Dall, J., Exbrayat, J.-F., Minh, D.H.T., Lomas, M., d'Alessandro, M.M., Paillou, P., Papathanassiou, K., Rocca, F., Saatchi, S.S., Scipal, K., Shugart, H., Smallman, T.L., Soja, M.J., Tebaldini, S., Ulander, L.M.H., Villard, L., Williams, M., 2019. The European Space Agency BIOMASS mission: measuring forest above-ground biomass from space. Remote Sens. Environ. 227, 44–60. https://doi.org/10.1016/j.rse.2019.03.032.

Réjou-Méchain, M., Tanguy, A., Piponiot, C., Chave, J., Hérault, B., 2017. BIOMASS: an R package for estimating above-ground biomass and its uncertainty in tropical forests. Methods Ecol. Evol. 8, 1163–1167. https://doi.org/10.1111/2041-210X.12753.

Saatchi, S.S., Moghaddam, M., 2000. Estimation of crown and stem water content and biomass of boreal forest using polarimetric SAR imagery. IEEE Trans. Geosci. Remote Sens. 38 (2), 697–709. https://doi.org/10.1109/36.841999.

Sandberg, G., Ulander, L.M.H., Fransson, J.E.S., Holmgren, J., Le Toan, T., 2011. L- and P-band backscatter intensity for biomass retrieval in hemiboreal forest. Remote Sens. Environ. 115, 2874–2886. https://doi.org/10.1016/j.rse.2010.03.018.

Santoro, M., Fransson, J.E.S., 2019. Integrating SAR backscatter, ICESAT GLAS metrics and allometric functions towards an improved estimation of forest biomass. IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 6320–6323. https://doi.org/10.1109/IGARSS.2019.8900407.

Schlund, M., Scipal, K., Quegan, S., 2018. Assessment of a power law relationship between P-band SAR backscatter and aboveground biomass and its implications for BIOMASS mission performance. IEEE J. Sel. Top. Appl Earth Observations Remote Sensing 11 (10), 3538–3547. https://doi.org/10.1109/JSTARS.2018.2866868.

Small, D., 2011. Flattening gamma: radiometric terrain correction for SAR imagery. IEEE Trans. Geosci. Remote Sens. 49 (8), 3081–3093. https://doi.org/10.1109/TGRS.2011.2120616.

Smith-Jonforsen, G., Ulander, L.M.H., Luo, X., 2005. Low VHF-band backscatter from coniferous forests on sloping terrain. IEEE Trans. Geosci. Remote Sens. 43 (10), 2246–2260. https://doi.org/10.1109/TGRS.2005.855134.

Snowdon, P., 1991. A ratio estimator for bias correction in logarithmic regressions. Can. J. For. Res. 21 (5), 720–724. https://doi.org/10.1139/x91-101.

Soja, M.J., d'Alessandro, M.M., Quegan, S., Tebaldini, S., Ulander, L.M.H., 2018. Model-based estimation of tropical forest biomass from notch-filtered P-band SAR backscatter. IEEE Geoscience and Remote Sensing Symposium (IGARSS) 8617–8620. https://doi.org/10.1109/IGARSS.2018.8517614.

Soja, M.J., Sandberg, G., Ulander, L.M.H., 2013. Regression-based retrieval of boreal Forest biomass in sloping terrain using P-band SAR backscatter intensity data. IEEE Trans. Geosci. Remote Sens. 51 (5), 2646–2665. https://doi.org/10.1109/TGRS.2012.2219538.

Soja, M.J., Askne, J.I.H., Ulander, L.M.H., 2017. Estimation of boreal forest properties from TanDEM-X data using inversion of the Interferometric water cloud model. IEEE Geosci. Remote Sens. Lett. 14 (7), 997–1001. https://doi.org/10.1109/LGRS.2017.2691355.

Tebaldini, S., Ho Tong Minh, D., Mariotti d'Alessandro, M., Villard, L., Le Toan, T., Chave, J., 2019. The status of technologies to measure forest biomass and structural properties: State of the art in SAR tomography of tropical forests. Surv. Geophys. 40, 779–801. https://doi.org/10.1007/s10712-019-09539-7.

Thurner, M., Beer, C., Santoro, M., Carvalhais, N., Wutzler, T., Schepaschenko, D., Shvidenko, A., Kompter, E., Ahrens, B., Levick, S.R., Schmullius, C., 2014. Carbon stock and density of northern boreal and temperate forests: carbon stock and density of boreal and temperate forests. Glob. Ecol. Biogeogr. 23, 297–310. https://doi.org/10.1111/geb.12125.

Truong-Loï, M.-L., Saatchi, S., Jaruwatanadilok, S., 2015. Soil moisture estimation under tropical forests using UHF radar polarimetry. IEEE Trans. Geosci. Remote Sens. 53 (4), 1718–1727. https://doi.org/10.1109/TGRS.2014.2346656.

Ulander, L.M.H., 1996. Radiometric slope correction of synthetic aperture radar images. IEEE Trans. Geosci. Remote Sens. 34, 1115–1122. https://doi.org/10.1109/36.536527.

Ulander, L.M.H., Gustavsson, A., Flood, B., Murdin, D., Dubois-Fernandez, P., Depuis, X., Sandberg, G., Soja, M.J., Eriksson, L.E.B., Fransson, J.E.S., Holmgren, J., Wallerman, J., 2011. BioSAR 2010: Technical assistance for the development of airborne SAR and geophysical measurements during the BioSAR 2010 experiment (final report). European Space Agency, Noordwijk, The Netherlands. ESA contract no. 4000102285/10/NL/JA/ef.

Ulander, L.M.H., Monteith, A.R., Soja, M.J., Eriksson, L.E.B., 2018. Multiport vector network analyzer radar for tomographic forest scattering measurements. IEEE Geosci. Remote Sens. Lett. 15 (12), 1897–1901. https://doi.org/10.1109/LGRS.2018.2865673.

Villard, L., Le Toan, T., 2015. Relating P-band SAR intensity to biomass for tropical dense forests in hilly terrain: γ0 or t0? IEEE J. Sel. Top. Appl Earth Observations Remote Sensing 8 (1), 214–223. https://doi.org/10.1109/JSTARS.2014.2359231.