MODEL-BASED ESTIMATION OF TROPICAL FOREST BIOMASS FROM NOTCH-FILTERED P-BAND SAR BACKSCATTER

Maciej J. Soja[1,2], Mauro M. d'Alessandro[3], Shaun Quegan[4], Stefano Tebaldini[3], and Lars M. H. Ulander[5]

[1] Horizon Geoscience Consulting, Belrose, NSW, Australia
[2] University of Tasmania, Hobart, TAS, Australia
[3] Politecnico di Milano, Milan, Italy
[4] University of Sheffield, Sheffield, UK
[5] Chalmers University of Technology, Göteborg, Sweden

Abstract *– This paper presents a new algorithm for forest biomass estimation from P-band synthetic-aperture radar (SAR) backscatter data, notch-filtered at ground-level. A semi-empirical model is fitted to spatial and polarization trends in the backscatter data and no reference biomass data are needed for training. An evaluation on airborne P-band SAR data from a tropical test site in Gabon results in a root-mean-square error lower than 20% and a correlation better than 90%.*

1. INTRODUCTION

European Space Agency's (ESA's) 7th Earth Explorer Mission and the first P-band (center frequency: 435 MHz, bandwidth: 6 MHz) synthetic aperture radar (SAR) in space, BIOMASS [1], is scheduled for launch in 2021. The primary goal of the mission is to provide a platform for global measurements of forest biomass, height, and disturbance, which are important for improved carbon cycle modelling and climate change prediction.

Previous studies have shown that P-band SAR provides good sensitivity to biomass across different biomes [1]. However, all proposed biomass estimation algorithms rely on the availability of local training data, which greatly reduces their applicability at global scale given the sparsity of suitable *in situ* data, especially in the tropics. Other challenges with global estimation of biomass from P-band SAR data lie in defining the form of the relationship between biomass and SAR observables and the significant sensitivity of the observables to scatterer size and shape distributions, soil and vegetation moisture, and ground topography. Thus, an accurate biomass retrieval model, suitable for global mapping and robust against these effects, is thus yet to be proposed for BIOMASS.

Due to the temporal stability of the main P-band scatterers (larger branches, trunks, and ground), repeat-pass interferometric and tomographic, polarimetric SAR data will be acquired. During its main operational phase, BIOMASS will make several near-global acquisitions in a repeat-pass interferometric mode, producing datasets of three images separated by three days. During an initial tomographic phase, BIOMASS will acquire one full coverage of 7-pass tomographic SAR data.

Since many challenges are related to ground-level scattering (which includes the ground-trunk and ground-canopy interactions), a simple yet robust solution is to decrease the ground-level contribution in the SAR data. This can be done using an interferometric notch filter at ground level [2,3]. The filtered SAR images should then be dominated by contributions from the forest canopy, which is expected to be simpler to model. A similar approach, but using SAR tomography, has been used in the past [4].

In this paper, we propose a new biomass estimation algorithm using a semi-empirical canopy backscatter model based on [5], together with an optimization approach based on least-squares minimization, but not requiring any reference biomass data. We show, using airborne P-band SAR data acquired over a tropical forest, that the proposed approach provides excellent results, and we briefly discuss its global application aspects.

2. INVERSION ALGORITHM

The basic inversion algorithm relies on minimizing the sum of squared differences between the measured backscatter, notch-filtered at ground level, and canopy backscatter predicted with a forward model f, with respect to the unknown forward model parameters:

$$(1)\ \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{\mathrm{PQ}} V_{\mathrm{PQ}}\left(f\left(W_i, \mathbf{p}_{\mathrm{PQ}j}\right) - \sigma^0_{\mathrm{PQ}ij}\right)^2$$

where N is the number of regions of interest (ROIs), M is the number of multi-temporal

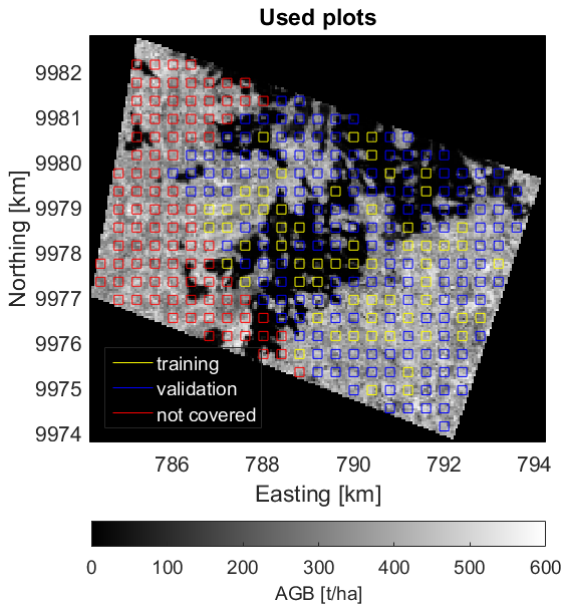

Figure 1: Training and validation plots overlaid on lidar biomass map.

acquisitions, W_i is biomass for ROI i (assumed constant between acquisitions), $\mathbf{p}_{\mathrm{PQ}j}$ is a vector of model parameters for polarization PQ and acquisition j, and V_{PQ} is a weighting factor for polarization PQ (here: 1 for HH and VV, 4 for HV).

Note that we apply several constraints in (1), which reduce the dimensionality of the forward model parameters: W_i can vary in space (across ROIs) but it cannot vary between polarizations and is assumed constant in time, while $\mathbf{p}_{\mathrm{PQ}j}$ can vary with polarization and time, but not in space. These constraints ensure that spatial trends in backscatter are explained by changing biomass, whereas polarization trends are explained by changing model parameters. Spatial variations in moisture are neglected in the current approach. Furthermore, if structural and dielectric changes between acquisitions can be neglected and the same well-calibrated system is used for all acquisitions, the dimensionality can be reduced even further by removing the time dependence in $\mathbf{p}_{\mathrm{PQ}j}$.

Given $M = 3$ (a typical BIOMASS scenario), 4 model parameters varying in time and polarization, and 3 unique polarizations (HH, HV and VV), the number of observables is $9N$, while the number of unknown parameters is $N + 36$. This system of equations is only solvable if $9N > N + 36$, i.e. $N > 4.5$, so at least 5 ROIs should be used, although a larger number of ROIs, representative in terms of spatial variability, is recommended to obtain a reliable solution.

During implementation, standard methods for non-linear optimization can be used, but bounded optimization is recommended to constrain parameter estimates within reasonable limits. To obtain confidence intervals for the estimated model parameters, an unconstrained fit can be conducted afterwards.

3. FORWARD MODEL

The following forward model is used for canopy backscatter:

$$(2)\ f(W,\mathbf{p}) = AW^{\alpha}\cos\theta\left(1-\exp\left(-\frac{BW}{\cos\theta}\right)\right)+N$$

where W is biomass, θ is the local incidence angle, and $\mathbf{p} = [A, \alpha, B, N]$ is a vector of model parameters, where A is a scaling constant, α is an exponent describing the dependence of backscatter on the biomass of individual scatterers, B is a constant describing the attenuation of the volume of scatterers, and N represents the residual contributions (noise). This model is based on the volume model in [5]. All model parameters are assumed non-negative. All polarizations are modelled with (2), but with different values of $\mathbf{p}$.

Note that (2) consists of a power law function and an exponential function, both of which are increasing for non-negative parameter values. This may render model inversion difficult due to parameter interactions.

In low-biomass areas, attenuation is expected to be small and we expect $B_{PQ}W \ll 1$. Model (2) can then be simplified to a simple power-law model for backscatter normalized to σ^0:

$$(3)\ f_{\mathrm{la}}(W,\mathbf{p}) = ABW^{\alpha+1} + N.$$

By fitting the low-attenuation model (3) to filtered backscatter data for low-biomass ROIs, we can obtain an estimate of α, which can subsequently be used as a constant when the full canopy model (2) is used with all data. This procedure is expected to make the inversion of the full model (2) more stable.

An inherent scaling issue is observed when both biomass and model parameters are estimated at the same time. It is straightforward to show that if W in (2) is a scaled version of the true biomass W_0 ($W = cW_0$, where c is a scaling constant), then the same model predictions can be obtained if parameters A and B are scaled accordingly ($A = \frac{A_0}{c^{\alpha}}$ and $B = \frac{B_0}{c}$, where A_0 and

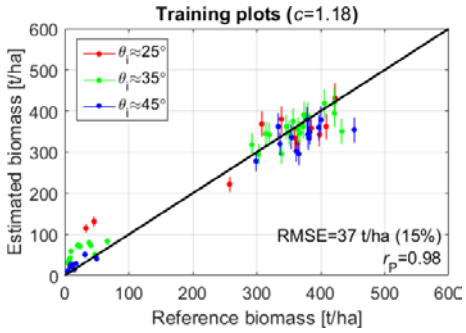

Figure 2: Estimated biomass and 95% confidence intervals for the 61 training plots, for which model parameters were also estimated.

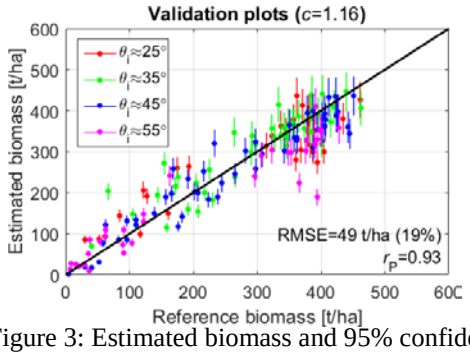

Figure 3: Estimated biomass and 95% confidence intervals for the 170 validation plots, for which model parameters were provided by earlier estimation steps.

B_0 are the true model parameters). This means that if W, A, B, and α are estimated simultaneously from the same data, biases in the estimated W, A and B are expected. This problem can be diminished to some extent by a sensible choice of initial values and boundary conditions. However, to ensure unbiased biomass estimation, *a posteriori* bias correction should be applied using, e.g., an unbiased, low-resolution estimate of the average biomass for the ROIs.

4. NOTCH FILTERING

SAR is an imaging process that maps the three-dimensional world onto a two-dimensional image. SAR imaging cannot distinguish between targets placed at the same distance from the sensor and at the same along-track coordinate; in other words, it is blind to the vertical direction. When sensing a semi-transparent medium, radar echoes associated with several targets placed at different elevations are mixed in the same resolution cell. The integral of the vertical reflectivity profile of the scene $s(z)$ determines the complex value of the SLC image as stated by [2]:

$$(4)\ y_{\mathrm{M}} = \int s(z)\mathrm{d}z.$$

The very same reflectivity profile, when observed from a slightly different position, is projected onto a complex sinusoid whose spatial frequency is proportional to the normal baseline:

$$(5)\ y_{\mathrm{S}} = \int s(z)\exp(jk_z z)\,\mathrm{d}z$$

The joint exploitation of these two interferometric images provides sensitivity to the vertical direction. The difference of y_{M} and y_{S} emphasizes certain elevations and attenuates others. The resulting filtered image takes the form:

$$(6)\ y_N = \int s(z)(1-\exp(jk_z z))\mathrm{d}z$$

$$(7)\ y_N = 2j\int s(z)\sin\left(\frac{k_z z}{2}\right)\exp\left(\frac{jk_z z}{2}\right)\mathrm{d}z$$

Equation (7) states that the vertical reflectivity profile of the scene $s(z)$ is modulated by a sine function before SAR projection. Elevations surrounding the zero of the sine wave are attenuated during the notch image formation process. The availability of the local topography allows the image pair to be modulated so that the zero of the sine is at ground level for each range azimuth pixel [3]. In this case, the resulting backscattered power can be completely ascribed to the canopy, and the contributions due to direct ground scattering and ground-trunk and ground-canopy interactions vanish.

5. SAR DATA

For this paper, we use three polarimetric SAR acquisitions from the AfriSAR campaign [6], acquired with the German F-SAR system on 2016/02/10 over La Lopé, a tropical forest site in Gabon. Our study uses 231 200 m × 200 m plots on a systematic, 400 m × 400 m grid, of which 61 are used for model parameter estimation (training plots, requirements: slope $< 10°$, incidence angle between $20°$ and $50°$, not on a forest/non-forest boundary) and the remaining 170 are used for biomass estimation with known model parameters (validation plots). Reference biomass is estimated for each of the 231 plots from the 50 m × 50 m biomass map derived from ALS and *in situ* data and provided within the AfriSAR campaign [6]. The training and validation plots used in this study are shown in Figure 1.

The low-attenuation model (3) was first used with those of the 61 training plots with AGB < 100 t/ha. The full model (2) was then used with all 61 plots, using α values from the previous step. Finally, biomass was estimated for the 170

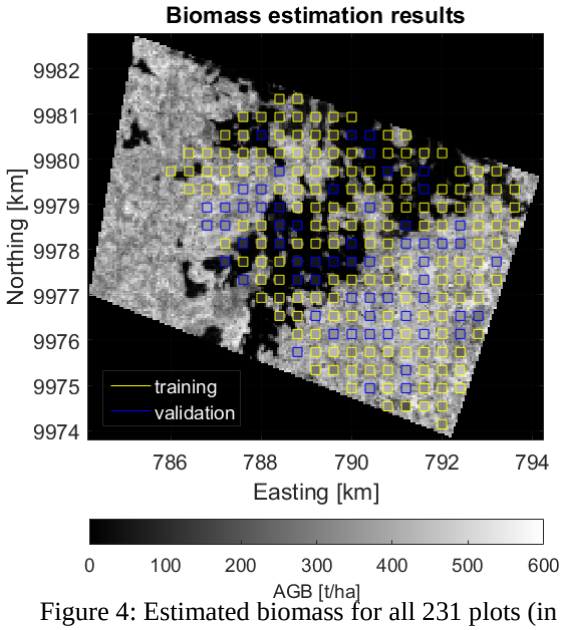

Figure 4: Estimated biomass for all 231 plots (in foreground) and reference biomass (in background).

validation plots with all four model parameters estimated in the previous step. Bias correction was applied by forcing the average of the estimated biomass values to equal the average of the reference biomass values.

6. RESULTS

Figures 2 and 3 show scatterplots comparing the estimated and reference biomass values for the 61 training plots and 170 validation plots, respectively. In both cases, the root-mean-square error (RMSE) for biomass estimation is below the 20% required for BIOMASS [1] and the correlation is better than 90%. The performance is good for all different incidence angle groups. The observed outliers in Figure 3 are thought to be caused by extreme topography.

Figure 4 shows the estimated values overlaid on the reference AGB map and illustrates how a potential global estimation can be performed. Biomass and model parameters are first estimated for a set of training plots selected to represent a wide range of backscatter values, but excluding complicated scenarios. Subsequently, biomass is estimated for the remaining plots or for all pixels within the image using the model parameters estimated earlier using training plots.

7. CONCLUSIONS

The proposed algorithm performs well for the challenging test site in La Lopé, Gabon, with biomass RMSE<20% and correlation above 90%. No reference biomass values are required for model training, although an unbiased, low-resolution estimate is needed for bias correction and a low-biomass mask is needed for initial estimation of α. Such information is, however, expected to be available from other sources (optical data, other SAR techniques and sensors, *in situ* data). Moreover, data acquired during the tomographic phase and overlaps between neighboring acquisitions can further aid parameter estimation. Plots used for parameter estimation should be large and homogeneous within the boundaries (to reduce speckle effects) and should represent a wide range of backscatter variation. Other SAR techniques can be used for plot selection.

Further work includes evaluating this approach on temperate and boreal data, where the ground contribution is expected to be more significant and biomass estimation from the filtered backscatter is expected to be more challenging.

ACKNOWLEDGEMENTS

Most of this work was conducted within the BIOMASS Level 2 Algorithm Implementation Study, funded by ESA. The authors would like to thank all team members for their constructive comments and support.

REFERENCES

[1] ESA (2012), "Report for Mission Selection: BIOMASS", ESA SP-1324/1 (3 volume series), European Space Agency, Noordwijk, The Netherlands.

[2] S. Tebaldini *et al.*, "Phase Calibration of Airborne Tomographic SAR Data via Phase Center Double Localization", *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1775-1792, March 2016.

[3] M. M. d'Alessandro *et al.*, "Phenomenology of Ground Scattering in a Tropical Forest Through Polarimetric Synthetic Aperture Radar Tomography", *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 8, pp. 4430-4437, Aug. 2013.

[4] E. Blomberg *et al.*, "Forest Biomass Retrieval from L-band SAR using Tomographic Ground Backscatter Removal", *IEEE Geosci. Remote Sens. Lett.* (accepted)

[5] M.-L. Truong-Loï *et al.*, "Soil Moisture Retrieval under Tropical Forests using UHF Radar Polarimetry", *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1718-1727, Apr. 2015

[6] I. Hajnsek *et al., AfriSAR:* "Technical Assistance for the Development of Airborne SAR and Geophysical Measurements during the AfriSAR Experiment: Final Report", ESA contract no. 4000114293/15/NL/CT, 2017