Topographic Correction for Biomass Retrieval from P-band SAR Data in Boreal Forests

Maciej J. Soja[†], Gustaf Sandberg[†] and Lars M. H. Ulander[†‡]

[†]Chalmers University of Technology, Sweden
[‡]Swedish Defence Research Agency (FOI), Sweden

Abstract—The influence of the ground slope on radar backscatter has been proven to be greater for lower radar frequencies due to deeper canopy penetration. In this study, multiple heading, P-band SAR data of boreal forest in Sweden was used to find a model for topographic correction for improved biomass retrieval. Eleven models were tested and the best model was selected. The selected model was then used for biomass retrieval. Even by means of the most simplified approach, forest biomass could be established with a root-mean-square error of approximately 50 t/ha for HV and 66 t/ha for HH.

I. INTRODUCTION

With the continued threat of global warming, the need for forest biomass mapping on global scale is growing stronger. Several studies (eg. [1]–[4]) have shown that P-band backscatter has a great potential for forest biomass estimation. However, most of these studies are done on areas with little or no topographic variability. Thus, there is a need to investigate the impact of ground slope on P-band backscatter data. To address this issue, the BioSAR 2008 campaign, initiated and funded by European Space Agency (ESA), was conducted in northern Sweden during October 2008. In this experiment, a test site with strong topographic variations was mapped using P- and L-band synthetic aperture radar (SAR). To enable detailed analysis of the topographic influence, several flight directions (headings) were used. Although no significant studies of P-band backscatter from forest growing on undulating terrain have been performed using multiple headings, previous studies of VHF-data have proven the usefulness of this approach [5]. An overview of the BioSAR 2008 campaign follows. A detailed description of the BioSAR 2008 experiment can be found in [6]. Moreover, some references to the BioSAR 2007 campaign will be made in this text. The description of the campaign can be found in [4], [7].

A. Test Site

The studied test site is located in northern Sweden (64°14'N, 19°46'E). It is a part of Vindeln Experimental Forest and belongs to the unit for field-based forest research at the Faculty of Forest Sciences, Swedish University of Agricultural Sciences (SLU). The test site has been extensively used for forestry and remote sensing research. The forest is classified as boreal and is dominated by mixed coniferous forest. The bedrock consists almost entirely of gneiss. The dominating soil is moraine with variations in thickness.

B. In-situ Data

A total of 31 forest stands, ranging in size from 2.4 to 26.3 ha, were inventoried during the summer of 2008. In the selected stands, circular sample plots with a radius of 10 m were laid out in a systematic way with a grid spacing selected to obtain approximately 10 plots for each stand. In addition, 110 sample plots were positioned outside the 31 stands. On the sample plots, all trees with a dbh (diameter at breast height) greater than 4 cm were callipered and the species registered. Approximately 1.5 trees per stand were randomly selected for measurement of tree height and age. Site variables such as vegetation and soil type were also recorded for each plot. The *in-situ* measurements were used to estimate several stand level measurements, including biomass and tree number density. Detail of the field inventory and estimation of forest parameters can be found in [6]. To complement the *in-situ* measurements, all but five of the 31 inventoried forest stands were visited during the campaign period and current conditions for the stands, as for example air temperature and percipitation, were noted. Measurements of soil moisture were also made for 10 of the stands. Finally, weather data from a research station located within the test site was also used.

C. Laser scanning data

A laser scanning of the test site was performed on between the 5th and 6th of August 2008 with the TopEye system S/N 425 mounted on a helicopter, at a flight altitude of approximately 500 m above ground level for main strips and 250 m above ground level for cross strips. Approximately 70 km^2 were covered using an average point density of approximately 5 points per square meter in the main strips and 15 points per square meter in the cross strips. The laser data was processed to obtain a Digital Elevation Map (DEM) and a map containing difference between minimum and maximum (vertical) z-value in a resolution cell. The second map corresponds to the vegetation height in the green areas. Using *in-situ* data from all available sample plots, a model relating laser scanning data to biomass was obtained. Biomass estimation was then made for 10 m by 10 m cells in a raster covering almost the entire test site. Additional details on the laser scanning can be found in [6].
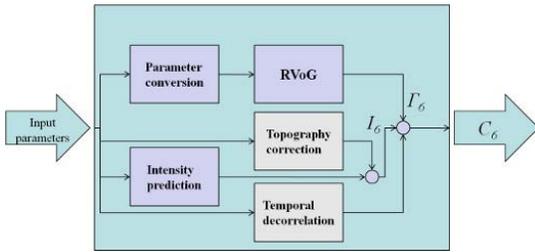
Fig. 1. A block diagram of the forward model. The chosen input parameters (including forest biomass, ground topography, and radar system setup) are first converted to input parameters which can then be plugged into the Random Volume over Ground model (RVoG). The model then gives the complex coherence matrix Γ_6, which may be corrected for temporal decorrelation. The intensity matrix I_6 is predicted using regression of BioSAR. Eventually, the coherence and intensity parts are combined giving the required extended correlation matrix C_6.

D. SAR data

P- and L-band SAR data was acquired using the Experimental SAR (ESAR) platform between the 14th and 15th of October 2008. For an area of 10 km (azimuth) by 3 km (ground range), large area polarimetric and interferometric SAR (PolInSAR) data was acquired for both frequency bands. Two PolInSAR data sets with 6 flight passes each were acquired for each frequency band. Moreover, polarimetric SAR data for 5 (P-band) and 4 (L-band) additional flight passes were acquired for a 3 km by 3 km area located within the larger area. The radiometric calibration of the images was checked using trihedral corner reflectors. Measured deviations from expected values were limited by about 1 dB, which was within specification values. Moreover, amplitude and phase imbalance between HH and VV polarizations were within the specifications for the ESAR system. Additional details on the SAR data can be found in [6].

II. METHOD

Within the framework of the ESA project "Development of Algorithms for Forest Biomass Retrieval", a request on a polarimetric and interferometric forward model is in demand. The model should predict the magnitude and phase of each element in the extended covariance matrix:

$$c_{w_M, w_S} = E\left[s_{w_M} s_{w_S}^*\right], \qquad (1)$$

where w_M and w_S refer to the polarization states of the master and slave image, respectively. With one single baseline two images are available, and using three independent polarization modes (HH, HV, and VV) the size of the covariance matrix becomes 6x6 elements, and thus the matrix is referred to as C_6.

Each element of C_6 can be re-written as a product of complex coherence $\widetilde{\gamma}$ and backscatter intensity i:

$$c_{w_M, w_S} = \widetilde{\gamma}_{w_M, w_S} \cdot i_{w_M, w_S}. \qquad (2)$$

Both elements will now be presented and discussed. For a block diagram of the forward model, see Figure 1.
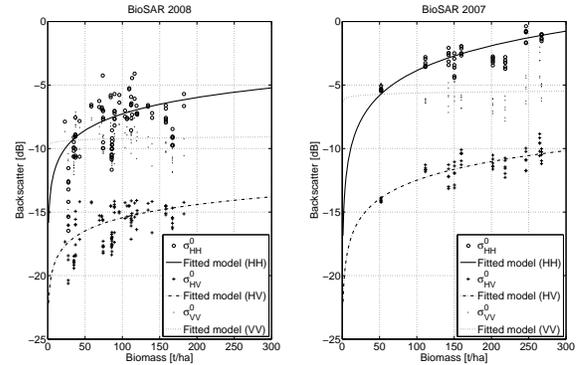

Fig. 2. The figures above show that backscatter is more susceptible to topography in BioSAR 2008 campaign (shown in the figure to the left) than in BioSAR 2007 campaing (shown in the figure to the right). Also, it can be seen that HV polarization is less prone to topography variations than HH and VV.

A. Coherence Prediction

1) Random Volume over Ground: In the chosen version of the forward model, the complex coherence is to be modelled using Random Volume over Ground (RVoG) model [8], [9]. This model predicts the complex coherence due to a volume of randomly located and oriented point scatterers positioned above a coherently scattering ground. A general expression for the model is:

$$\widetilde{\gamma}_\omega(h_V, f_z, \phi_0, m_\omega) = \exp(i\phi_0)\frac{\widetilde{\gamma}_V(h_V, f_z) + m_\omega}{1 + m_\omega}, \quad (3)$$

$$\widetilde{\gamma}_V(h_V, f_z) = \frac{\int_0^{h_V} f_z \exp(-ik_z z)dz}{\int_0^{h_V} f_z dz}, \quad (4)$$

where ω describes a certain polarization setup and $\widetilde{\gamma}_V$ represents the volume contribution. The model receives a set of input parameters (volume height h_V, scatterer distribution f_z, ground phase ϕ_0, and ground-to-volume ratio m_ω) which not neccessarily are connected to the geophysical properties of forest (eg. biomass and tree height). In order to provide RVoG with the neccessary parameters, some methods for conversion between the parameters need to be developped. Relations between the RVoG input parameters and forest parameters may be found by inversion of SAR data (eg. BioSAR 2007-data, see [7]). In the block diagram in Figure 1, both the RVoG-model and the parameter converter are marked.

2) Temporal Decorrelation: It is a well-known fact that coherence of a forest decreases with time (temporal decorrelation due to wind, percipitation, growing, etc.). This fact needs to be included in the model. Since the temporal decorrelation is quite difficult to study experimentally (due to the complexity of the governing processes and the costs of airborne SAR campaigns), an assumption about the decorrelation form is made:

$$\widetilde{\gamma} = \widetilde{\gamma}_0 e^{-t/\tau}, \qquad (5)$$

where τ is a temporal decorrelation constant and $\widetilde{\gamma}_0$ is the complex coherence without considering temporal decorrela-

TABLE I
TESTED MODELS

Model ID	Nonzero coefficients	f_{ij}
1	a_0, a_1	-
2	a_0, a_1, a_3	$\log(\cos\theta_i)$
3	a_0, a_1, a_3	$\log(1 - \tan\theta_0 \cdot \tan u \cdot \sin v)$
4	a_0, a_1, a_3	$\log(1 - \tan u \cdot \sin v)$
5	a_0, a_1, a_3	$\log(1 - \sin v)$
6	a_0, a_1, a_2, a_3	$\log(\cos\theta_i)$
7	a_0, a_1, a_2, a_3	$\log(1 - \tan\theta_0 \cdot \tan u \cdot \sin v)$
8	a_0, a_1, a_2, a_4	$\log(\cos\theta_i)$
9	a_0, a_1, a_2, a_4	$\log(1 - \tan\theta_0 \cdot \tan u \cdot \sin v)$
10	All	$\log(\cos\theta_i)$
11	All	$\log(1 - \tan\theta_0 \cdot \tan u \cdot \sin v)$

Fig. 4. Root-mean-square errors (RMSE) between the true and the modelled backscatter values for the eleven different models presented in Table I. Models number 8 and 9 give the lowest RMSE values. Model 9 gives lowest RMSE in both HH and VV while model 8 is slightly better in HV.
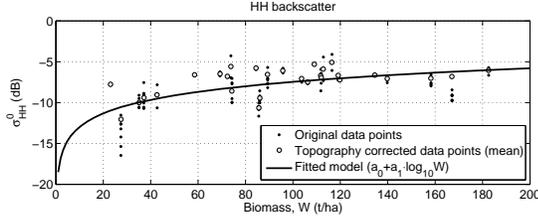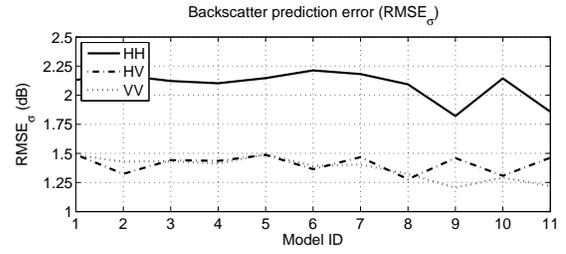
Fig. 3. Backscatter from 31 stands in Krycklan (BioSAR 2008) varies more for low-biomass stands. Lower tree density implies greater ground contribution and thus higher susceptibility to topography. Also, it can be noted that mean backscatter values for each stand predicted by the best found model (number 9) give realistic results that can be used for improved biomass estimation.
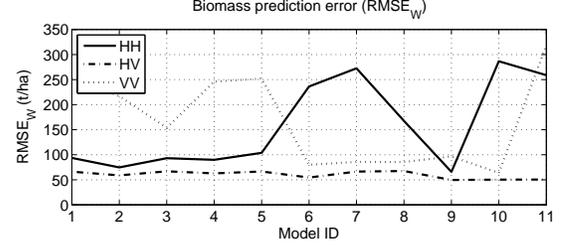
Fig. 5. Here, RMSE values for biomass retrieval after topography correction for all eleven models are presented. Even here, the best results are achieved using model 9.

tion.

B. Intensity Prediction

It can be easily verified that each element i_{w_M, w_S} of the intensity matrix I_6 can be re-written as a combination of backscatter intensities for master and/or slave in different polarizations (denoted as σ^0_{PQ}).

The intensity dependence on biomass and geometry will be predicted using regression of BioSAR data. The biomass dependence can be modelled using:

$$[\sigma^0_{PQ}]_{dB} = a_1 + a_2 \log_{10}(W), \quad (6)$$

where $[\sigma^0_{PQ}]_{dB}$ is the backscatter amplitude of the PQ-channel in dB, W is the forest biomass, and a_i are constants found using least-squares approach. In Figure 2, one can see that fitting of model (6) gives much better results if the topography is negligible (as it does in BioSAR 2007 campaign). Also, one can see that the influence of topography is generally lowest for HV-channel (due to the least ground contribution in this channel).

In order to improve the model performance, some topography dependent elements should be added to equation 6. The procedure of selection of the best model will now be presented.

C. Topography Correction

A good understanding of topography influence on radar backscatter is important for two reasons: 1) better forward model performance, 2) better biomass inversion performance. The selection of a suitable forward model will be made first. After that, the model will be used for biomass inversion.

Some theory suggests that backscatter is proportional to $\cos^n\theta_i$, where n lies between 1 and 2. In this text, we assume that it is true and employ the following model:

$$[\sigma^0_{ij}]_{dB} = a_0 + a_1 \cdot \log_{10} W + a_2 \cdot (\log_{10} W)^2 + \\ + a_3 \cdot f_{ij}(u, v, \theta_i) + a_4 \cdot f^2_{ij}(u, v, \theta_i), \quad (7)$$

where constants a_n and function f_{ij} have their values according to Table I. Angle θ_i is the local incident angle (θ_0 is the global incident angle). Angles u and v are defined in [10].

By fitting all eleven models from Table I to the BioSAR 2008 data, one can compute the difference between the measured backscatter values and predicted model values in terms of root-mean-square deviation. In the graph in Figure 4, one can see the RMSE values for all eleven models. The best fit for two of three polarizations is given by model 9, and the final form of the model is thus:

$$[\sigma^0_{ij}]_{dB} = a_0 + a_1 \cdot \log_{10} W + \\ + a_3 \cdot \log(1 - \tan\theta_0 \cdot \tan u \cdot \tan v) + \\ + a_4 \cdot [\log(1 - \tan\theta_0 \cdot \tan u \cdot \tan v)]^2. \quad (8)$$

In Figure 3, one can see the original data points together with the values predicted using model described above.

D. Biomass Retrieval

Biomass can be retrieved from radar backscatter by inversion of the forward model presented in equation (8). Assuming that the "true" relation between biomass and backscatter is the one presented in equation (6), the second part can be seen as topography influence only. By subtracting the topographic influence part from the SAR data, one can obtain modified,
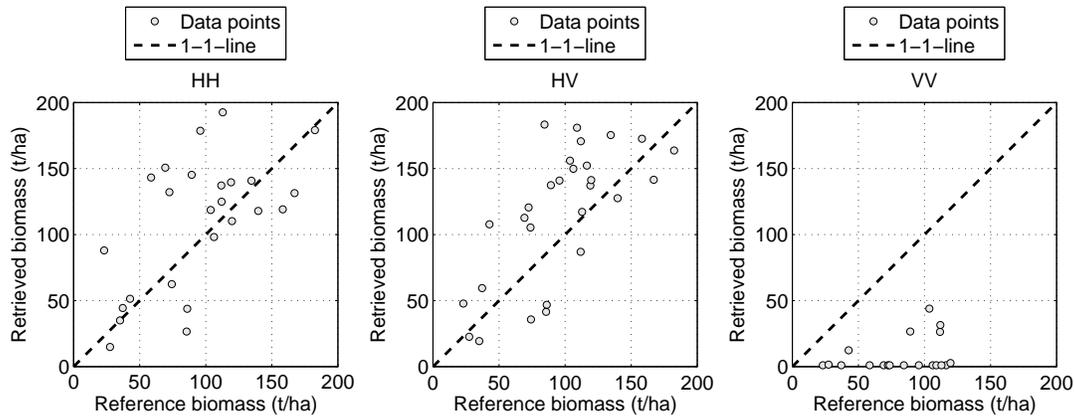
Fig. 6. Results from biomass inversion after topography correction. HH and HV give, as expected, rather good estimates of biomass. VV channel fails to estimate biomass correctly.

topography corrected backscatter values (shown in Figure 3). A line representing the model in equation (6) can be fitted to the data (also shown in Figure 3). The inverted version of that fit can then be used for biomass retrieval.

III. RESULTS

Retrieval of biomass using the method described in section II-D gives good results for HH and HV polarizations (see Figure 6). For VV, the inversion process fails (VV has not been shown to contain meaningful information for biomass retrieval). A root-mean-square error of around 50 t/ha can be observed for HV and 66 t/ha for HH. Different RMSE values depending on the type of correction function used are shown in Figure 5. HV shows in general the best performance. Also, backscatter from low-biomass stands is more susceptible to underlying ground topography compared to high biomass stands. This is due to the higher canopy density and lower ground penetration.

IV. CONCLUSION

The results described above are based on a simplified statistical analysis. Methods that include mixed effect models, splitting of data, and more advanced error analysis will hopefully give better results and are under development. Also, some other correction functions may give better results. Keeping in mind that the presented approach is a great simplification, an RMSE of around 50 t/ha for single-channel biomass estimation is still a very promising result.

REFERENCES

[1] A. Beaudoin, T. L. Le Toan, S. Goze, E. Nezry, A. Lopes, E. Mougin, C. C. Hsu, H. C. Han, J. A. Kong, and R. T. Shin, "Retrieval of Forest Biomass from SAR Data," *International Journal of Remote Sensing*, vol. 15, pp. 2777–2796, 1994.

[2] T. Le Toan, A. Beaudoin, J. Riom, and D. Guyon, "Relating Forest Biomass to SAR Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, pp. 403–411, 1992.

[3] S. Saatchi, K. Halligan, D. Despain, and R. Crabtree, "Estimation of Forest Fuel Load from Radar Remote Sensing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 1726–1740, 2007.

[4] G. Sandberg, L. M. H. Ulander, J. E. S. Fransson, J. Holmgren, and T. Le Toan, "L-band versus P-band SAR for Biomass Retrieval in Hemiboreal Forest," *Remote Sensing of the Environment*, 2009, accepted for publication.

[5] K. Folkesson, G. Smith-Jonforsen, and L. M. H. Ulander, "Model-based Compensation of Topographic Effects for Improved Stem-volume Tretrieval from CARABAS-II VHF-band SAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 1045–1055, 2009.

[6] I. Hajnsek, R. S. M. Keller, R. Horn, S. Lee, L. M. H. Ulander, A. Gustavsson, G. Sandberg, T. Le Toan, S. Tebaldini, A. M. Guarnieri, and F. Rocca, "BioSAR 2008 Technical Assistance for the Development of Airborne SAR and Geophysical Measurements during the BioSAR 2008 Experiment: Draft Final Report - BioSAR Campaign," European Space Agency, Tech. Rep., 2009, ESA contract no. 22052/08/NL/CT.

[7] I. Hajnsek, R. Scheiber, L. Ulander, A. Gustavsson, G. Sandberg, S. Trebaldini, A. M. Guarnieri, F. Rocca, F. Bombardini, and M. Pardini, "BioSAR 2007 Technical Assistance for the Development of Airborne SAR and Geophysical Measurements during the BioSAR 2007 Experiment: Final Report without Synthesis," European Space Agency, Tech. Rep., 2008, ESA contract no. 20755/07/NL/CB.

[8] K. Papathanassiou and S. Cloude, "Single-baseline Polarimetric SAR Interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, pp. 2352–2363, 2001.

[9] T. Mette, "Forest Biomass Estimation from Polarimetric SAR Interferometry," Ph.D. dissertation, Deuches Zentrum für Luft- und Raumfahrt, Institut für Hochfrequenztechnik und Radarsysteme, Oberpfaffenhofen, 2007.

[10] L. M. H. Ulander, "Radiometric Slope Correction of Synthetic Aperture Radar Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 5, pp. 1115–1122, 1996.