FOREST BIOMASS ESTIMATION USING RADAR AND LIDAR SYNERGIES

Siyuan Tian[1], Mihai A. Tanase[1], Rocco Panciera[1], J. Hacker[2], Kim Lowell[1]

[1] Cooperative Research Centre for Spatial Information, The University of Melbourne
[2]Airborne Research Australia, School of the Environment, Flinders University

ABSTRACT

This study investigates the improvement in above ground biomass estimates when using a synergistic model based on lidar derived forest structural information (i.e., canopy cover percentage) and radar backscatter. The results were cross-compared with a radar only model. A two-layered radar backscatter model was also tested. The results showed that lidar-based structural information has the potential to increase the accuracy of biomass estimation by up to 20% depending on polarization and acquisition date. A smaller improvement was observed when using a modeled estimate of the forest canopy cover as would be the case of a future lidar/radar joint space-borne mission. The two-layered vegetation backscatter model did not improve the biomass estimation accuracy with errors being higher when compared to a single-layer vegetation model.

Index terms **-** forest biomass, lidar-radar synergies

1. INTRODUCTION

Global concern about the impacts of increased atmospheric concentrations of greenhouse gases such as CO_2 on climate has focused attention on the dynamics of carbon in terrestrial ecosystems, particularly in forests which account for over 60% of the total carbon stock. Recent research has mainly focused on biomass retrieval from space-borne radar or lidar data. While space-borne lidar sensors only provide samples at isolated locations, radar has the mapping capability required for continuous spatial estimations. Various radar sensors have been used to model forest backscatter response with relatively limited accuracy partly due to the inadequate description of forest structural spatial variability. Previous investigations of radar–lidar synergy [1, 2] found that there was little to be gained by just adding radar metrics to lidar-based regression models. The lidar-only models performed better with radar metrics being essentially redundant [3]. However, no attempt was made to expand radar backscattering models taking into account spatially explicit structural information obtained from lidar sensors. Lidar data could provide *a priori* information on forest structure that can be subsequently used for radar backscatter modeling. The aim of this study was to investigate the utility of embedding lidar-derived vegetation structural information into radar water cloud type backscatter models. In addition, an existing water cloud type radar backscatter model was modified to account for the backscattering components of different vegetation layers related to forest overstory and understory, respectively.

2. STUDY AREA AND DATASETS

The forest area studied (Gillenbah State forest) is located in New South Wales in Australia, near the town of Narrandera. The forest is dominated by white cypress pine (*Calitris glucophylla*) with 10% dispersed grey box trees (*Eucalyptus microcarpa*). The topography is nearly flat. A biometric survey was conducted of 60 circular plots (500 m^2) clustered in 12 sites during September 2011. A cluster site consisted of five circular plots, one in the centre and one at each cardinal direction (Figure 1). The location, circumference and height of all trees with a diameter at breast height (DBH) greater than 5cm was recorded whereas smaller trees were counted and their average height recorded. Information from 10 additional sparsely vegetated plots was also collected. Total above-ground biomass (AGB) and biomass components (leaves, branches, stem wood) were estimated for each tree using species-specific allometric equations [4, 5] and aggregated to plot level. The total AGB for individual plots varied between 1.5 and 180 t ha^{-1}.

The lidar data were acquired by a Riegl Q560 Airborne Laser Scanning (ALS). The ALS Q560 was flown over the forest area in September 2011 at an altitude of about 300m AGL. The system recorded all echo pulses within a small footprint (~15cm). An average first return pulse density of 40 ppm was obtained after combining flight lines acquired from two different directions. The discrete returns were extracted from the raw lidar data and combined with the navigation data to yield geo-referenced point clouds. Accuracies of this procedure are approximately 0.4m horizontally and 0.15m vertically, with higher accuracies within individual scans. The point clouds were then classified into ground and non-ground returns. All non-ground returns were considered vegetation since no man-made features are located within the forest perimeter.

Four dual-polarized ALOS PALSAR datasets (HH and

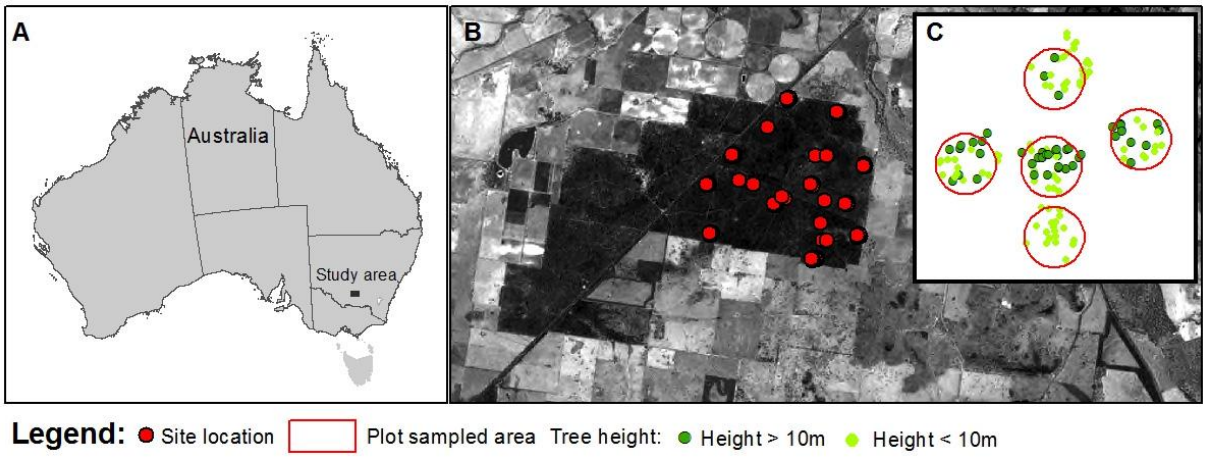

Figure 1. Location of the study area (black square in panel A) and field data sampling locations (red dots in panel B). The inset (C) shows the layout of sampled plots at each site and an example of the tree spatial distribution measured at one sampling location. Panels B has as background a Landsat ETM+ panchromatic image.

HV channels) were used in this study for forest backscatter modelling. The satellite images, acquired the year before ground and lidar data collection (April to August 2010), were multi-looked (2x8) to obtain similar ground pixel spacing in both azimuth and range (25m). The intensity was transformed to radar backscatter coefficient (σ°) after applying the absolute calibration factor [6]. To express the speckle reduction achieved after multi-looking, the number of statistically independent looks was estimated by means of the equivalent number of looks (ENL) using homogeneous areas [7]. Due to the low ENL observed (8) speckle filtering [8] was applied using a 9x9 window size to further reduce data noise. All images were geocoded to the UTM coordinate system using a lookup table that described the transformation between the radar and the map geometry [9].

3. METHODS

The water cloud-type backscatter model proposed in [10] was parameterized and subsequently inverted to obtain biomass estimates from radar data. The model was used in its original form (see eq. 1) that expresses the forest backscatter as a function of the area fill factor (η) and the forest transmissivity (T). The simplified model form (2) proposed in [10] was also used for cross-comparison. The simplified model was initially used for inversion since the area fill factor is not a commonly used parameter for forest inventory [10]. However, for our study the use of the original model form (1) was possible by extracting the area fill factor (i.e., forest percent cover) from the point cloud Lidar data. The total forest backscatter (σ°_{for}) was modeled as the sum of ground scattering (direct and attenuated by the vegetation layer) (σ°_{gr}) and the direct scattering of the vegetation (σ°_{veg}) weighted by the area fill factor and the two-way vegetation transmissivity (T). The two–way transmissivity was initially expressed as a function of stem volume. However, since the stem volume represents a fraction of the forest biomass (taking into account wood density) the total ABG was used instead to estimate the two-way transmissivity (i.e., $T=e^{-\beta AGB}$ where β is an empirically defined coefficient). A two-layered vegetation model (3) was also considered for the biomass estimation. This model is an extension of the model proposed in [10] with the vegetation contributions being split by vegetation layers (σ°_{veg1} and σ°_{veg2}). The area fill factor was derived from the lidar data (first returns over a specified height divided by the total number of first returns) for the entire vegetation strata (η) or by each layer (η_1, η_2) for vegetation heights below (η_1) and above (η_2) 6m (i.e., roughly the regrowth layer height).

$$\sigma^\circ_{for} = (1-\eta)\sigma^\circ_{gr} + \eta[\sigma^\circ_{gr}T + \sigma^\circ_{veg}(1-T)] \quad (1)$$

$$\sigma^\circ_{for} = \sigma^\circ_{gr}e^{-\beta AGB} + \sigma^\circ_{veg}\left(1-e^{-\beta AGB}\right) \quad (2)$$

$$\sigma^\circ_{for} = (1-\eta)\sigma^\circ_{gr} + \eta\sigma^\circ_{gr}T + \left(\eta_1\sigma^\circ_{veg1} + \eta_2\sigma^\circ_{veg2}\right)(1-T) \quad (3)$$

The models were parameterized using the backscatter, the area fill factor and the AGB extracted for the 60 ground assessed plots (500 m^2 each). Ten sparsely vegetated plots were also used for the backscatter modeling. The location of each ground sample plot within the PALSAR scenes was determined using GPS data collected during the field sampling. All models were subsequently inverted (4 to 6) to produce spatially explicit biomass estimates for each SAR acquisition date and SAR polarization. The radar based estimates were compared to the reference biomass measured on the ground to assess the absolute and relative estimation errors. The data were split by biomass intervals (20t ha^{-1} each) and random samples were drawn from each interval to form the training and validation datasets. Ten iterations were carried out to minimize random sampling effects when

$$AGB = -\frac{1}{\beta}\, ln\left(\frac{\sigma^0_{veg}-\sigma^0_{for}}{\sigma^0_{veg}-\sigma^0_{gr}}\right) \quad (4)$$

$$AGB = -\frac{1}{\beta} ln\left(\frac{\sigma^\circ_{for}-(1-\eta)\sigma^\circ_{gr}-\eta\sigma^\circ_{veg}}{\eta(\sigma^\circ_{gr}-\sigma^\circ_{veg})}\right) \quad (5)$$

$$AGB = -\frac{1}{\beta} ln\left(\frac{\sigma^\circ_{for}-(1-\eta)\sigma^\circ_{gr}-\eta_1\sigma^\circ_{veg1}-\eta_2\sigma^\circ_{veg2}}{\eta\sigma^\circ_{gr}-\eta_1\sigma^\circ_{veg1}-\eta_2\sigma^\circ_{veg2}}\right) \quad (6)$$

calculating average error metrics (i.e., Root Mean Squared Error-RMSE, Relative Error-$RE_{\%}$ and correlation coefficient between estimated and measured biomass-*r*).

To apply such enhanced radar backscatter models accurate information on the area fill factor is needed at large scales. Such data could be available from space-borne lidar. However, past and future space-borne lidar missions are not able to provide full spatial coverage with inter-beams areas not being sampled (Figure 2). The area fill factor for inter-beam pixels could be interpolated using the nearest sampled lidar beams. Alternatively, the area fill factor could be modeled as a function of the radar backscatter and the estimates collected over the areas observed by the lidar sensor. The validity of the latter approach was tested for our study area as follows: 1) using the training samples (i.e., assumed as falling inside both the lidar beams and the radar images) a model describing the relationship between the area fill factor and the radar backscatter was parameterized, 2) the area fill factor for the validation samples (i.e. assumed as falling outside the lidar beams but inside the radar images) was computed from the radar backscatter using the parameterized model 3) the radar backscatter model (1) was parameterized using the training samples, and 4) the biomass was estimated for the validation samples by inverting the water cloud model using the modeled area fill factor values (from step 2).

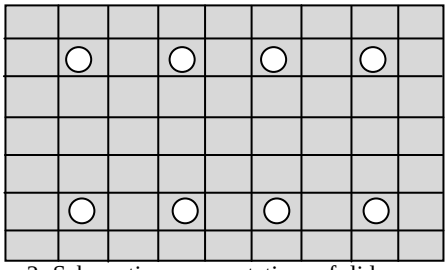

Figure 2 Schematic representation of lidar sample beams (white circles) overlaid on a spatially continuous radar acquisition.

4. RESULTS AND DISCUSSIONS

The RMSE, $RE_{\%}$ and *r* are shown in Table 1 for each ALOS PALSAR acquisition day and polarization. Table 1 also shows the difference in error metrics between models (1) and (2). A negative difference indicates higher accuracy of the original model when compared to the simplified form. For all days, polarizations and error metrics the inclusion of lidar-derived forest structural parameters (i.e., forest cover) significantly improved the estimation accuracy (Table 1). A mean improvement of about 3 t ha^{-1} was observed across days and polarizations which in relative terms corresponds to a 12% decrease of the biomass estimation error. Depending on acquisition date and polarization the relative estimation error improvement was as much as 20%.

The higher improvements observed for the co-polarized channel (HH) is related to the scattering mechanism which is dominated by surface scattering suggesting that the area fill factor provided key information for model parameterization.

Using a more complex two-layered vegetation model (3) did not increase the biomass estimation accuracy. In fact, an opposing effect was observed with the RMSE and $RE_{\%}$ increasing by 13.4 t ha^{-1} and 13% (mean values computed for all polarizations and acquisition dates) when compared to the simplified model form (2). A significant factor when modeling and validating the models was the balanced repartition of the samples by biomass intervals. When using random repartition of the data between training and validation datasets, as opposed to a random stratified repartition, the biomass estimation errors were about 8% higher. This is probably caused by the inefficient parameterization of the model since not all intervals were always covered during the random data repartition due to the small number of samples for some biomass intervals.

When using a radar-based area field factor (Figure 3) an improvement of both the RMSE (3.5 t ha^{-1}) and the $RE_{\%}$ (5%) with respect to the simplified model form (2) was observed while no improvement was observed for the *r* error metric (Table 2). This suggests that even for areas not covered by lidar data the biomass estimation from radar data could be enhanced although the gains in accuracy are somewhat lower than when using an area fill factor directly estimated from lidar data.

Table 1. Error metrics of the water cloud model when using Lidar based forest cover information. The error metrics for the simplified model form are also given. The difference between the two models is displayed in **bold**.

Model	Error metric	2010.04.04		2010.05.20		2010.07.05		2010.08.20		mean
		HH	HV	HH	HV	HH	HV	HH	HV	
(1)	RMSE	28.1	23.7	28.6	23.3	32.0	30.8	27.7	26.0	27.5
(2)	(t ha-1)	34.2	27.4	32.1	26.9	33.8	28.7	33.8	27.1	30.5
Difference		**-6.1**	**-3.7**	**-3.5**	**-3.6**	**-1.7**	**2.2**	**-6.1**	**-1.1**	**-3.0**
(1)	RE%	45.1	52.3	49.7	42.0	47.4	53.6	40.7	52.2	47.9
(2)	(%)	59.8	60.0	69.2	54.1	55.1	61.9	60.0	59.0	59.9
Difference		**-14.7**	**-7.7**	**-19.5**	**-12.1**	**-7.7**	**-8.3**	**-19.2**	**-6.7**	**-12.0**
(1)	*r*	0.4	0.3	0.4	0.3	0.6	0.3	0.2	0.1	0.3
(2)		0.1	0.3	0.2	0.2	0.3	0.5	0.2	0.0	0.2
Difference		**0.3**	**0.0**	**0.1**	**0.1**	**0.3**	**-0.2**	**-0.1**	**0.0**	**0.1**

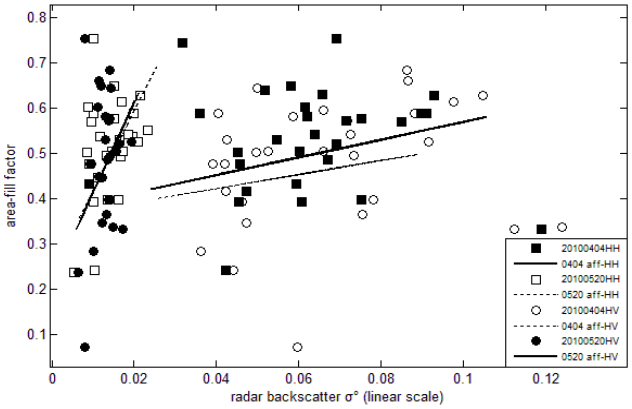

Figure 3 Example of relationships between radar backscatter and the area fill factor for HH and HV polarizations (see Table 2 for correlation coefficients).

Table 2. Error metrics of the water cloud model when using radar modeled area fill factor. The error metrics for the simplified model form are also given for cross-comparison. In bold the difference between the two models.

Model	Error metric	2010.04.04 HH	2010.04.04 HV	2010.05.20 HH	2010.05.20 HV	mean
(1)	RMSE	28.8	24.7	26.9	23.7	26.0
(2)	(t ha-1)	33.2	28.9	31.9	24.4	29.6
Difference		**-4.4**	**-4.2**	**-5.0**	**-0.6**	**-3.6**
(1)	RE%	69.5	45.2	71.6	49.5	58.9
(2)	(%)	73.3	53.2	77.0	51.3	63.7
Difference		**-3.8**	**-8.0**	**-5.4**	**-1.8**	**-4.8**
(1)	r	0.50	0.34	0.46	0.31	0.40
(2)		0.30	0.29	0.31	0.45	0.34
Difference		**0.2**	**0.05**	**0.15**	**-0.14**	**0.06**
r*		0.35	0.52	0.32	0.49	0.42

* Correlation coefficient between the area fill factor and the radar backscatter for the training samples.

5. CONCLUSIONS

This study investigated the potential improvements in biomass estimation accuracy when using ancillary information of forest structure derived from lidar data. Furthermore, a more complex two-layered backscatter model was assessed and the influence of the sample distribution on model parameterization/validation was studied. The results showed that a significant decrease of the biomass estimation error was achieved when using ancillary information on forest structure derived from lidar data as opposed to a model that does not take into account the variability in forest canopy cover especially for co-polarized backscatter data. A two-layer backscatter model did not improve the estimation error over a single-layered model. Applying enhanced backscatter models at large scale would decrease the uncertainty of biomass estimates from radar data when appropriate forest structural information is available.

6. ACKNOWLEDGMENTS

This work was funded by the ARC Super Science Fellowships (grant no. FS100100040) and the Cooperative Research Centre for Spatial Information. The field work was supported by the ARC Discovery Project scheme (grant no. DP0984586).

7. REFERENCES

[1] R. F. Nelson, P. Hyde, P. Johnson, B. Emessiene, M. L. Imhoff, R. Campbell, and W. Edwards, "Investigating RaDAR–LiDAR synergy in a North Carolina pine forest," *Remote Sensing of Environment*, vol. 110, pp. 98-108, 2007.

[2] P. Hyde, R. Nelson, D. Kimes, and E. Levine, "Exploring LiDAR–RaDAR synergy—predicting aboveground biomass in a southwestern ponderosa pine forest using LiDAR, SAR and InSAR," *Remote Sensing of Environment*, vol. 106, pp. 28-38, 2007.

[3] P. Hyde, R. Dubayah, W. Walker\, J. B. Blair, M. Hofton, and C. Hunsaker, "Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy," *Remote Sensing of Environment*, vol. 102, pp. 63-73, 2006.

[4] W. H. Burrows, M. B. Hoffman, J. F. Compton, and P. V. Back, "Allometric relationships and community biomass stocks in White Cypress Pine (Callitris glaucophylla) and associated eucalpyts of the Carnarvon area, central Queensland.," 2001.

[5] S. D. Hamilton, G. Brodie, and C. O'Dwyer, "Allometric relationships for estimating biomass in grey box (Eucalyptus microcarpa)," *Australian Forestry*, vol. 68, pp. 267-273, 2005.

[6] M. Shimada, O. Isoguchi, T. Tadono, and K. Isono, "PALSAR Radiometric and Geometric Calibration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, pp. 3915-3932, 2009.

[7] J. Bruniquel and A. Lopes, "Multi-variate optimal speckle reduction in SAR imagery," *International Journal of Remote Sensing*, vol. 18, pp. 603-627, 1997.

[8] A. Lopes, E. Nezry, and R. Touzi, "Adaptive filters and scene heterogeneity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, pp. 992-1000, 1990.

[9] U. Wegmüller, C. Werner, T. Strozzi, and A. Wiesmann, "Automated and precise image registration procedures," in *Analysis of multi-temporal remote sensing images*, vol. 2, L. Bruzzone and P. Smits, Eds. Singapore: World Scientific 2002, 2002, pp. 37-49.

[10] M. Santoro, L. Eriksson, J. Askne, and C. Schmullius, "Assessment of stand-wise stem volume retrieval in boreal forest from JERS-1 L-band SAR backscatter " *International Journal of Remote Sensing*, vol. 27, pp. 3425-3454, 2006.