Synergistic use of very high-frequency radar and discrete-return lidar for estimating biomass in temperate hardwood and mixed forests

Banskota, Randolph Wynne, Johnson, Emessiene

Annals of Forest Science (2011) 68:347–356
DOI 10.1007/s13595-011-0023-0

ORIGINAL PAPER

Synergistic use of very high-frequency radar and discrete-return lidar for estimating biomass in temperate hardwood and mixed forests

Asim Banskota · Randolph H. Wynne · Patrick Johnson · Bomono Emessiene

Received: 8 April 2010 / Accepted: 14 September 2010 / Published online: 16 March 2011

Abstract
• ***Introduction*** Accurate estimation of aboveground biomass is essential to better understand the carbon cycle in forest ecosystems.
• ***Methods*** The objective of this study was to determine whether biomass in temperate hardwood forests is better estimated using very high-frequency radar data (from BioSAR) alone or in combination with small-footprint discrete-return lidar data (both profiling and scanning). The study area was in the Appomattox-Buckingham State Forest, Virginia, USA (78°41′W, 37°25′N). Aboveground biomass for 28 stands was estimated using 131 basal area factor 10 point samples. The resulting stand biomass estimates were used as the dependent variable in a multiple linear regression. Descriptors of the lidar distributions (both profiling and scanning) and averaged normalized radar cross-sections in each of these stands were used as independent variables.
• ***Results*** Regression results revealed the following: (1) neither BioSAR nor scanning lidar data alone are good predictors of stand biomass (R^2=0.57, root mean squared error (RMSE)= 31.0 tonnes/ha and R^2=0.64, RMSE=28.5 tonnes/ha, respectively); (2) BioSAR data combined with small-footprint discrete lidar data (either profiling or scanning) are the best predictors of stand biomass (R^2=0.80, RMSE=21.3 tonnes/ha and R^2=0.76, RMSE=24.2 tonnes/ha, respectively); and (3) when used with BioSAR data for stand biomass estimation, less costly profiling lidar data convey the same information as more costly scanning lidar data.
• ***Conclusion*** Useful synergy can be realized by considering lidar and radar measurements jointly in estimating aboveground biomass in hardwood and mixed forests.

Keywords BioSAR · Scanning lidar · Profiling lidar · Aboveground biomass · Best subsets regression · Carbon

Handling Editor: Matthias Dobbertin

A. Banskota (✉)
Geospatial and Environmental Analysis Program,
Virginia Polytechnic Institute and State University,
313 Cheatham Hall,
Blacksburg, VA 24061, USA
e-mail: asimb@vt.edu

R. H. Wynne
Department of Forest Resources and Environmental Conservation,
Virginia Polytechnic Institute and State University,
Blacksburg, VA, USA
e-mail: wynne@vt.edu

P. Johnson · B. Emessiene
Zimmerman Associates, Inc.,
Fairfax, VA, USA

P. Johnson
e-mail: pjohnson@terresense.com

B. Emessiene
e-mail: bemessiene@zai-inc.com

1 Introduction

Biomass forms an essential part of the active carbon pools in the global carbon cycle. Prediction of forest biomass is therefore important for understanding major carbon cycle uncertainties and monitoring carbon sequestration (Hyde et al. 2007; Lucas et al. 2000; Skole and Tucker 1993). The traditional field methods used for measuring aboveground biomass are accurate but very expensive and labor intensive (Lefsky et al. 1999; Lim et al. 2003). In recent years, various remote sensing techniques have been explored in estimating aboveground biomass, starting with optical passive remote sensing using the visible and infrared part of the solar electromagnetic spectrum (Fraser and Li 2002; Sader et al.

1989; Steininger 2000). In forested environments, these sensors measure reflected energy that is largely a function of canopy architecture (leaf area, leaf angle, and clumping), leaf pigments, soil background, and underground vegetation (Goel 1989). A direct physical relationship between aboveground biomass and optical remote sensing data does not exist; hence, the latter only provide indirect estimates of aboveground biomass (Chopping et al. 2008; Hyde et al. 2007). Different approaches have been utilized with varying degrees of success to estimate biomass from optical remote sensing data, for example, establishing the relationships between biomass and vegetation indices (Peddle et al. 2001; Sader et al. 1989), spectral bands (Boyd et al. 1999; Foody et al. 2003; Steininger 2000), image texture (Lu 2005), and combinations of texture and spectral information (Lu and Batistella 2005). Vegetation reflectance and/or reflectance derivatives (for example, vegetation indices, texture, etc.) saturate in moderate to high biomass forests (Anaya et al. 2009; Lefsky et al. 2002). For example, Steininger (2000) showed that the canopy reflectance saturated when aboveground biomass approached 150 tonnes/ha or vegetation age reached 15 years in the tropical secondary successional forests in Manaus, Brazil.

A number of studies have investigated the role of synthetic aperture radar (SAR) data to estimate aboveground biomass (Austin et al. 2003; Harrell et al. 1997; Luckman et al. 1998; Santos et al. 2002). SAR data are sensitive to geometric properties of the forest; thus, they are more directly related to measurements of aboveground biomass than optical remote sensing data. Studies utilizing SAR data have found that the sensitivity of the radar backscatter coefficient to biomass disappears in forests with biomass exceeding 100 tonnes/ha (Dobson et al. 1992; Fransson et al. 2000; Imhoff et al. 2000; Imhoff 1995). The usual 1- to 10-GHz frequencies are often intercepted by the crown layer of the canopy in dense forests, and twigs, needles, and smaller branches in the top layers of the canopy are the major scatterers at these wavelengths (Fransson et al. 2000). Using lower frequencies, e.g., the P-band (around 440 MHz), the sensitivity to biomass increases, but saturation still occurs at 100–200 tonnes/ha biomass. Hence, even with the P-band, 81% of global forests are above the saturation limit and not within the range of biomass assessment (Imhoff 1995; Nelson et al. 2007). Therefore, even lower SAR frequencies are necessary to push the saturation point upward and facilitate biomass retrieval for medium to high biomass forest stands.

The pressing need for the use of lower-frequency SAR systems to address saturation-related issues accelerated the development of BioSAR. BioSAR is a very high-frequency (VHF) SAR that was specifically designed by NASA and Zimmerman Associates, Inc., Vienna, Virginia, USA to retrieve forest biomass estimates without saturating. It operates in the frequency range of 80–120 MHz. Results have shown good correlation between backscatter measured with BioSAR and forest biomass up to values of 250 tonnes/ha without saturation (Imhoff et al. 2000). For VHF SAR, the dominant scattering mechanism is ground-trunk dihedral scattering, since backscattering from branches, needles, and the forest floor and corresponding attenuation due to the forest canopy are very low. Hence, the strength of backscattering can be directly related to the trunk volume, with improved biomass discrimination and no saturation effects (Jonforsen et al. 2005).

The use of airborne lidar remote sensing in estimating aboveground biomass has also become more prevalent. A lidar sensor can directly measure components of vegetation canopy structure, such as canopy height, and provides indirect information about the vertical distribution of canopy structure such as leaves and branches (Drake et al. 2003; Lim et al. 2003). Because the components of the vertical structure of the canopy are closely linked with aboveground biomass, several studies have shown good correlation between biomass and lidar metrics. Studies that used lidar for aboveground biomass estimation utilized both "discrete return" systems and "full waveform" systems. Bortolot and Wynne (2005) and Popescu et al. (2003) utilized an individual tree-based approach, and van Aardt et al. (2008) utilized a stand-based approach with discrete lidar data to estimate biomass (<350 tonnes/ha) in a hardwood and mixed forest in Virginia, USA. Zhao et al. (2009) utilized a scale invariant approach to estimate biomass (<300 tonnes/ha) in a plantation and deciduous forest in eastern Texas, USA and obtained root mean squared errors (RMSEs) between measured and model-estimated biomass ranging from 14.3 to 36.7 tonnes/ha. Boudreau et al. (2008) used waveform lidar in the different vegetation zones in the Canadian province of Québec (biomass up to 215 tonnes/ha) and estimated aboveground biomass with RMSEs ranging from 14.2 to 38.7 tonnes/ha. Lefsky et al. (1999) and Means et al. (1999) estimated aboveground biomass with waveform lidar in a deciduous forest of eastern Maryland, USA (biomass up to 450 tonnes/ha, R^2 between measured and estimated biomass up to 0.81) and with large footprint scanning airborne lidar in the Western Cascades of Oregon, USA (biomass ranging from 500 to 1,500 tonnes/ha, RMSE=132 tonnes/ha), respectively. Nelson et al. (2007) reported that no lidar research to date has noted any saturation effect for biomass estimation. Lidar makes use of the close link between canopy height and aboveground biomass. The absence of saturation in the relationship between lidar-derived variables and biomass is likely because higher biomass forests usually have taller canopies. However, there are some shortcomings of both discrete-return and waveform lidar data. For waveform data, these include the often large footprint and typically low horizontal sampling densities (van Aardt et al. 2006). For discrete-return data, the primary

shortcomings are the poor results in mixed and deciduous forests (Boudreau et al. 2008; Popescu et al. 2003) and (for scanning lidar) the high cost.

The main objective of this study was to investigate the usefulness of BioSAR, used individually or with lidar, to estimate biomass in hardwood and mixed forests. The current study is a follow-on to work reported by Nelson et al. (2007). They investigated the use of the VHF radar system BioSAR in conjunction with a simple profiling lidar, the portable airborne laser system (PALS), in estimating aboveground biomass in an industrial, intensively managed pine forest. They found no practical, exploitable lidar–radar synergy in their studies. However, the authors recommended continuing studies to compare lidar and BioSAR data sets to look for exploitable synergies in temperate and hardwood mixed forests. Their assumption was that the utility of radar–lidar combination might be manifested in hardwood and mixed forests, where biomass estimation has been shown to be more problematic than in coniferous forests. Deliquescent tree forms in hardwood forests do not exhibit apical dominance as in coniferous forests, thereby making the height–volume or height–biomass relationship noisier (Boudreau et al. 2008). Hence, adding SAR variables (whose value is directly related to bole or large branch volume) to the height-related information from lidar variables might provide better estimates of biomass in hardwood and mixed forests. The current study augments Nelson et al. (2007) by looking at hardwood and mixed forest stands using small-footprint discrete-return scanning lidar (hereinafter referred to as scanning lidar) data in addition to BioSAR and PALS data.

2 Methods

2.1 Study area and ground data

The study area is in the Appomattox–Buckingham State Forest, Virginia, USA (78°40′30″ W, 37°25′9″ N). This Virginia Piedmont region consists of various coniferous, upland hardwood, and mixed forest stands. The tree species were white oak (*Quercus alba*), chestnut oak (*Quercus prinus*), northern red oak (*Quercus rubra*), southern red oak (*Quercus falcata*), yellow poplar (*Liriodendron tulipifera*), red maple (*Acer rubrum*), and three species of pines: Virginia pine (*Pinus virginiana*), loblolly pine (*Pinus taeda*), and shortleaf pine (*Pinus echinata*). The stand age varied, being approximately 25 years for the pine plantations, 45–65 years for the pine-hardwood mixed stands, and up to 110–125 years for the upland hardwood stands. A mean elevation of 185 m, with a minimum of 159 m and a maximum of 238 m, and rather gentle slopes characterize the topography of the study area. The location of the study area and a canopy height model from the scanning lidar data used are shown in Fig. 1.

Field data consisted of 131 basal area plots (basal area factor 10) out of 256 previously mapped plots on a 16 columns by 16 rows, 201.17-m grid (van Aardt et al. 2006). Not all locations could be measured this time because of recent or active harvests or lack of permission to sample some plots on private property. Field data were collected during November and December of 2007. Differentially corrected plot location, plot basal area, diameter at breast height (dbh), height, and species were measured for all plots and tallied trees. Biomass equations for per-tree calculations found in Popescu et al. (2004) were used in this study. This latter study was situated within the same geographical boundaries, with the same species being studied. Biomass was calculated on an individual tree basis for each plot and expanded to per-hectare values. Descriptive statistics for all basal area plots are given in Table 1.

2.2 Data collection and preprocessing

2.2.1 BioSAR data

The BioSAR sensor was carried aboard a Twin Otter twin-engine aircraft to collect data over the study area in October 2007. BioSAR acquires volumetric returns on a 600-m radius circle directly beneath the aircraft in the 80–120 MHz range. As reported in Nelson et al. (2007), the size of the area illuminated by the radio pulse is determined by the beam width and flight altitude, nominally 300 m AGL. The frequency of the emitted pulse was designed to avoid contamination from local FM radio station signals.

BioSAR returns from a 600-m target area were range-gated and Doppler-sorted to produce 300 m across-track × 30 m along-track pixels. Doppler sorting was used to collect radar signal in 5° incidence angle increments fore and aft of nadir. Radar response values were averaged together to provide a single response value for 18 incidence angle bins (5°–45°) for each Doppler resolution cell. The angular volumetric BioSAR responses were radar cross-sections for a given incidence angle, defined as the ratio of the received power to the transmitted power, adjusted for receiver and antenna gains. Each radar cross-section was then divided by the effective target area to produce the normalized radar cross-section (NRCS).

2.2.2 Profiling lidar

Profiling lidar data were acquired using a portable airborne laser system (PALS) coincident with the BioSAR along 13 flight lines. PALS was specifically designed as a sampling tool to estimate forest volume, biomass, and carbon across larger regions. The laser used in the acquisition was a Riegl LD90-

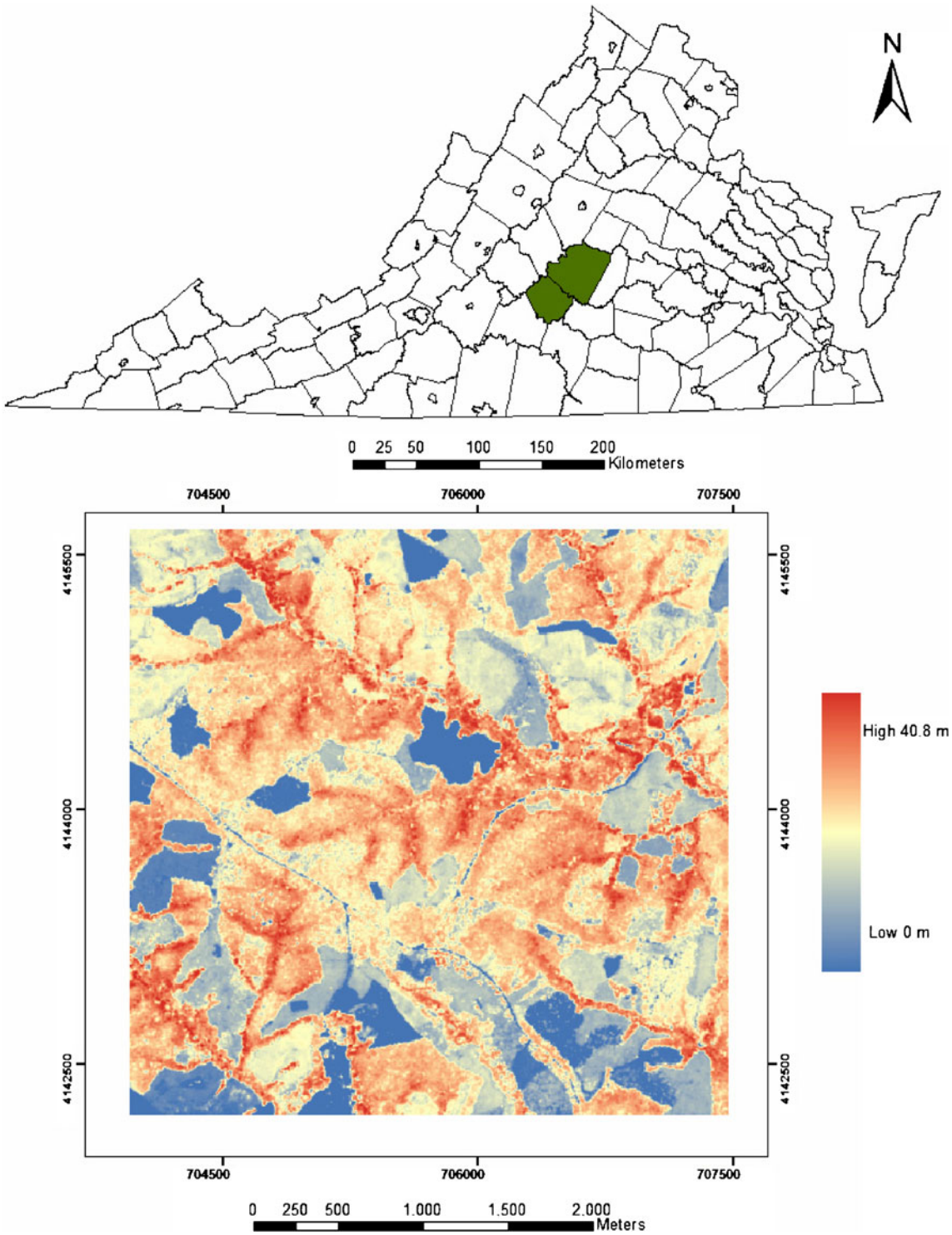

Fig. 1 Map of the Commonwealth of Virginia, USA indicating the study area location; the canopy height model from the scanning lidar data set used in this study is shown below (height in meters). Projected coordinate system of map: NAD 1983 UTM Zone 17N

3800 VHS FLP, a near-infrared lidar (λ=0.905 μm), programmed to supply sequential, alternating first and last returns. The PALS spot size was 60 cm at target, and the along-track post spacing was 16.7 cm. Since first and last returns are sequentially acquired, the top of the canopy is described by pulses spaced 33.4 cm apart. The interleaved last returns were used to identify ground beneath the forest canopy by identifying local minima and connecting these minima with a spline function. Once the ground line was defined, canopy height was calculated for each first return pulse.

Table 1 General descriptive statistics for deciduous and coniferous plots

Type	Parameter	unit	Min	Max	Mean	σ
Coniferous	Height	m	8.44	31.56	18.12	3.92
	DBH	cm	12.70	53.10	24.10	7.87
	Biomass	tonnes/ha	11.60	202.00	96.90	47.80
Deciduous	Height	m	2.10	38.12	22.10	5.47
	DBH	cm	10.41	130.56	34.80	14.99
	Biomass	tonnes/ha	16.33	267.20	140.50	57.30

"Deciduous" or "coniferous" types were defined with majority contribution, from either deciduous or coniferous species, respectively. The last four columns show the minimum, maximum, mean, and standard deviation of the measured (height and DBH) and calculated (biomass) parameters

DBH *diameter at breast height*

2.2.3 Scanning lidar

Scanning lidar data were acquired in August 2008 (during "leaf on" conditions) using an Optech ALTM 3100 sensor. The laser was operated with a pulse rate of 100 kHz, with a scan angle less than 15°, point densities of five pulses or better per square meter, a vertical accuracy of 15 cm or better over bare ground, and 0.5 m or better horizontal accuracy. The range captured included up to four measurements with first, second, third, and last returns. Bare ground returns were supplied by the data provider and were based on a proprietary algorithm. Ground returns were interpolated to a 1-m spatial resolution grid using regular kriging, which was found to be the most accurate interpolation technique using similar data for the same study area (Popescu et al. 2003; van Aardt et al. 2006). Canopy height was computed as the difference between all returns and the corresponding ground return values.

2.2.4 Stand-based metrics calculation

A stand-based strategy was adapted to directly compare model-estimated biomass with that of ground-observed biomass. We utilized a recently updated Appomattox–Buckingham State Forest (ABSF) stand map of the area for that purpose. ABSF stands were buffered inward by 50 m in order to exclude fuzzy stand boundary areas and to address potential misregistration between the field and remotely sensed data. Only those BioSAR pixels falling cleanly inside the buffered stands and lying within 100 m of plot centers were extracted. BioSAR measurements within the stands were averaged, resulting in 28 stand-level observations for the 18 variables (Table 2). Lidar pulses were collected across the same 28 stands to get stand-level PALS distributional metrics as listed in Table 3. Stand-level metrics were calculated for all hits and canopy hits (pulse with a height≥3 m). Similar distributional metrics were also calculated for scanning lidar for heights derived from first, second, and third returns.

2.3 Statistical analysis

Best subsets linear regression (MINITAB 15) was used to create models of stand biomass as a function of the remote sensing variables (Tables 2 and 3). We separated the variables into 11 sensor-specific sets, as follows: one set of BioSAR variables, two sets of PALS variables (metrics from all hits and canopy hits), and eight sets of scanning lidar variables (metrics from "canopy hits" and "all hits" for first, second, third, and all lidar returns). These sensor-specific sets were considered separately and in combination with each other resulting in 21 analysis sets of independent variables. For each analysis set, correlation analysis was used to remove all but one (that most correlated with biomass) within each group of highly correlated ($r>0.9$) variables. Each (now paired) analysis set was used to predict biomass measured in situ. Only models of six or fewer predictor variables were considered in each case. Only the best performing (highest adjusted R^2) analysis sets for each sensor or sensor combination were retained, reducing the final number of analysis sets from 21 to five (BioSAR only, PALS all hits, BioSAR plus PALS all hits, scanning lidar all return canopy hits, and BioSAR plus scanning lidar all return canopy hits). Selected models were subject to the following constraints: (1) all independent variables were significant at $\alpha=0.05$, (2) Mallows C_p was less than $p+1$, where p is the number of predictor variables used, and (3) there could not be multi-collinearity, i.e., all variance inflation factors (VIFs) were less than 10. Significance testing and calculation of VIFs were done in MATLAB (v. 7.1). Consideration was also given to model parsimony, i.e., a model with fewer variables was preferred to one with many variables. Once the best models were

Table 2 BioSAR variables used in this study

Angular bins (in degrees)	Variable name
0–5	a00, b00
5–10	a05, b05
10–15	a10, b10
15–20	a15, b15
20–25	a20, b20
25–30	a25, b25
30–35	a30, b30
35–40	a35, b35
40–45	a40, b40

Variables "*a*" and "*b*" refer to the mean forward and backward NRCS response, respectively

Table 3 Metrics derived from PALS

PALS metric	Description
h_a, h_c	Mean height, all hits and canopy hits
std_{al}, std_c	Standard deviation, all hits and canopy hits
sk_{al}, sk_c	Skewness, all hits and canopy hits
k_{al}, k_c	Kurtosis, all hits and canopy hits
h_{qal}, h_{qc}	Quadratic mean canopy height, all hits and canopy hits
p_{10}, p_{20}, $\cdots p_{100}$	Percentile heights, all hits
p_{c10}, p_{c20}, $\cdots p_{c100}$	Percentile heights, canopy hits

Similar set of metrics were developed for height derived from first, second, third, and all returns of scanning lidar. A canopy hit is a pulse with a height≥3 m

chosen, leave-one-out cross validation was used to validate them. In each iteration of the cross validation, one of the 28 observations was set aside, a predictive equation was calculated using the remaining 27 observations, and the single observation set aside was used to validate the model. This procedure was repeated 28 times. The cross-validation coefficient of determination (CV-R^2) and prediction sum of squares (PRESS) were calculated to assess the prediction capability of the best models.

3 Results

The final results of best subset regression analyses are presented in Table 4. The first column shows the type of independent variables used in the model. The second, third, fourth, and fifth columns show the model coefficient of determination (R^2), adjusted R^2, model Mallows C_p, and root mean square error (RMSE in tonnes per hectare). The sixth and seventh columns show the cross-validation R^2 (CV-R^2) and PRESS, respectively. The last column shows the best subset regression equations between aboveground biomass (in tonnes per hectare) and subsets of remote sensing variables selected based upon different models' statistics (adjusted R^2, model mallows C_p, VIF, and statistical significance).

The models derived from the scanning lidar metrics provided the best fit (R^2=0.64, RMSE=28.5 tonnes/ha) when the regression models were built using measurements from a single sensor. However, the models fitted using composite variables from two sensors, BioSAR and PALS (R^2=0.80, RMSE=21.3 tonnes/ha) and BioSAR and scanning lidar (R^2= 0.76, RMSE=24.2 tonnes/ha), provided better estimates of biomass compared to the models with single sensor variables. The predicted versus observed biomass plots for the best two models are provided in Figs. 2 and 3.

No model had more than four variables. Across all models, the selected variables were as follows: a15, mean NRCS response at forward 15°to 20° angle bins (decibels); b20, mean NRCS response at backward 20° to 25° angle bins (decibels); b40, mean NRCS response at backward 40° to 45° angle bins (decibels); p_{60}, 60th percentile of PALS height, all hits (meters); p_{70}, 70th percentile of PALS height, all hits (meters); k_{al}, kurtosis of PALS height, all hits; k_{cI}, kurtosis of scanning lidar height, all return canopy hits; and p_{c30I}, 30th percentile of scanning lidar height, all return canopy hits (meters). While space precludes illustration of the relationship of each of these variables to field-measured biomass, we include two to help the reader understand why the simple linear models are sufficient to predict biomass. Observed biomass versus b20 is shown in Fig. 4. Observed biomass versus p_{c30I} is shown in Fig. 5.

Table 4 Regression results between stand-based biomass and lidar (PALS or scanning lidar) and/or BioSAR variables

Variables	R^2	Adjusted R^2	C_p	RMSE	CV-R^2	PRESS	Best model
BioSAR only	0.57	0.54	−2.6	31.0	0.49	28,955	$Y = 151.2 - 4.5(a15) + 4.4(b20)$
PALS only	0.55	0.51	−2	31.6	0.47	29,596	$Y = -1.7 + 7.5(k_{al}) + 6.1(p_{60})$
BioSAR+PALS	0.80	0.78	4	21.3	0.70	17,358	$Y = -163.3 + 9.0(k_{al}) + 9.7(p_{70}) - 8.4(b40_c)$
Scanning lidar only	0.64	0.61	2.8	28.5	0.55	25,491	$Y = 178 - 14.3(p_{c30I}) - 84.8(k_{cI})$
BioSAR+scanning lidar	0.76	0.72	−1.4	24.2	0.67	18,816	$Y = 144.9 + 10.3(p_{c30I}) - 49.8(k_{cI}) - 2.8(a15) + 2.4(b20)$

The first column shows the type of independent variables used in the model. The second, third, fourth, and fifth columns show the model coefficient of determination (R^2), adjusted R^2, model Mallows C_p, and RMSE (in tonnes/ha), respectively. The sixth and seventh columns show the CV-R^2 and PRESS, respectively

Y, biomass (tonnes/ha); a15, mean NRCS response at forward 15° to 20° angle bins (decibels); b20, mean NRCS response at backward 20° to 25° angle bins (decibels); b40, mean NRCS response at backward 40° to 45° angle bins (decibels); p_{60}, 60th percentile of PALS height, all hits (meters); p_{70}, 70th percentile of PALS height, all hits (meters); k_{al}, kurtosis of PALS height, all hits; k_{cI}, kurtosis of scanning lidar height, all return canopy hits; p_{c30I}, 30th percentile of scanning lidar height, all return canopy hits (meters)

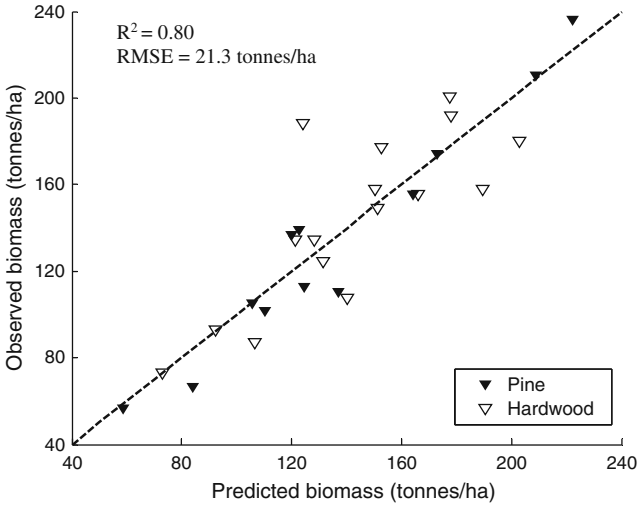

Fig. 2 Predicted versus observed biomass plot for best model with BioSAR variables (b40, mean NRCS response at backward 40°–45° angle bins (decibels)) and PALS variable (p_{70}, k_{al}, 70th percentile of PALS height, all hits (meters), and kurtosis of PALS height, all hits). *Solid* and *hollow triangles* stand for the pine and hardwood stands, respectively

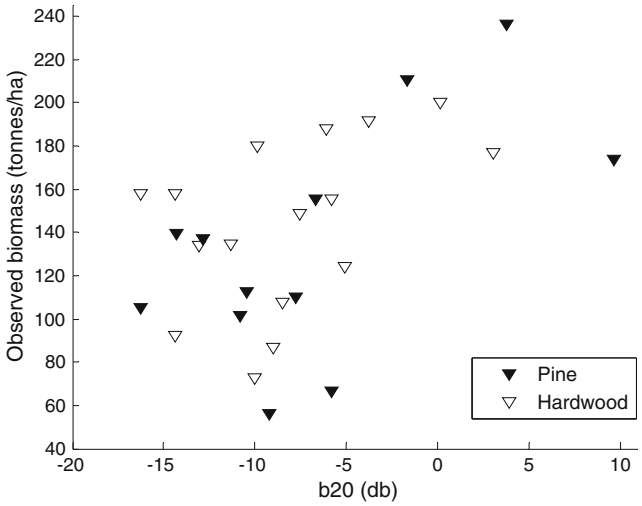

Fig. 4 Observed biomass versus b20, the mean NRCS response at backward 20° to 25° angle bins. *Solid* and *hollow triangles* stand for the pine and hardwood stands, respectively

4 Discussion

The results indicate that significant synergy may be realized by using lidar and radar data in tandem for estimating biomass in hardwood and mixed forests. The model statistics in this study were better than those for an object- and stand-level study by van Aardt et al. (2006) and a plot-level lidar study by Popescu et al. (2004) in the same study area. The best model in this study had an R^2 of 0.80 and RMSE of 21.3 tonnes/ha, versus 0.66 and 20.3 tonnes/ha (object level) and 0.46 and 41.2 tonnes/ha (stand level) found by van Aardt et al. (2006). Similarly, Popescu et al. (2004) obtained an R^2 of 0.82 and RMSE of 29.0 tonnes/ha (for coniferous species) and an R^2 of 0.33 and RMSE of 44.4 tonnes/ha (for deciduous species). These results are different from those reported by Hyde et al. (2007) and Nelson et al. (2007). In those studies, lidar, whether scanning or profiling, estimated biomass more accurately and precisely than any of the radar data considered in the two studies. Further consideration of lidar and radar jointly did not, for all practical purposes, improve forest biomass

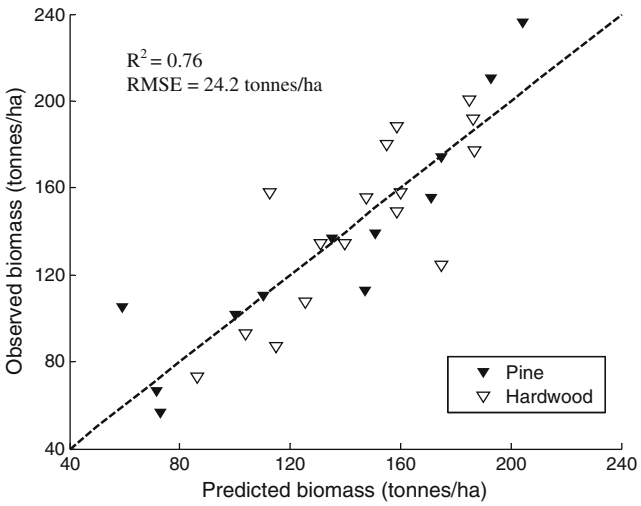

Fig. 3 Predicted versus observed biomass plot for the best model with BioSAR variables (a15, b20, mean NRCS response at forward 15° to 20°, backward 20° to 25° angle bins (decibels)) and scanning lidar variable (k_{cI}, p_{c30I}, kurtosis of scanning lidar height, all return canopy hits, and 30th percentile of scanning lidar height, all return canopy hits (meters)). *Solid* and *hollow triangles* stand for the pine and hardwood stands, respectively

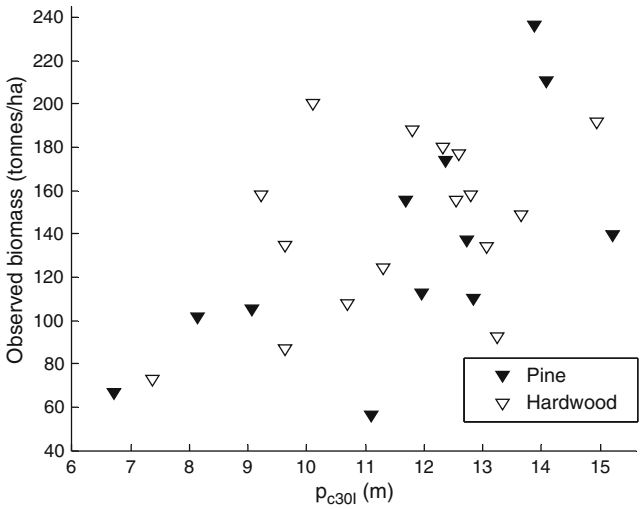

Fig. 5 Observed biomass versus p_{c30I}, the 30th percentile of scanning lidar height, all return canopy hits. *Solid* and *hollow triangles* stand for the pine and hardwood stands, respectively

estimation. However, those two studies were undertaken in industrial, heavily managed, even-aged forests of loblolly pine and ponderosa pine, respectively. Lidar has been shown to accurately estimate canopy height in even-aged, monospecific forests with an apically dominant growth form (Nelson et al. 2007). These forests exhibit greater correlation between individual tree height and biomass (Popescu et al. 2003). Hence lidar-based biomass estimates are highly accurate and precise in studies conducted in these type of forests (cross-validation R^2=93.3 and RMSE=33.9 tonnes/ha in Nelson et al. 2007; cross-validation R^2=82.6 and RMSE=26.05 tonnes/ha in Hyde et al. 2007), and little is gained by further adding radar-derived variables. On the other hand, canopy height estimation in hardwood and mixed forests using small footprint lidar data is less accurate than in coniferous forests, mainly because of (1) increased understory and (2) failure to sample the tops of the relatively broad trees (Clark et al. 2004; Lefsky et al. 2002; Means et al. 1999). Combined with the greater uncertainty in canopy height estimation is the decreased correlation between tree height and biomass in hardwood forests compared to coniferous forests (Popescu et al. 2003). The lower correlation between tree height and biomass can be attributed to hardwoods' deliquescent form, as hardwood trees tend to distribute their biomass in their lateral branches, weakening the height–biomass relationships (Nelson et al. 2007). As such, we posit that the relatively poor relationship between biomass and lidar-derived height variables in this study (used alone) is explained by the combined effect of greater uncertainty in canopy height estimates and the weaker biomass–height relationship in hardwood and mixed forests.

Studies using CARABAS, a radar sensor similar to BioSAR, have shown that the VHF backscatter mechanisms are mainly dominated by ground-trunk double-bounce scattering mechanisms for simple forest structures such as gently sloped coniferous forests (Fransson et al. 2000; Hallberg et al. 2005). The longer wavelength radar has good foliage penetration; hence, the foliage-floor and trunk-foliage double-bounce scattering can be considered negligible. The backscattering amplitude due to trunk-ground scattering has been shown to be proportional to stem volume (Israelsson et al. 1997; Jonforsen et al. 2005). In hardwood forests, not only the main trunk but also the primary branches are expected to dominate the backscattering through branch–ground interactions and direct backscatter from the branches themselves (Israelsson et al. 1997). Understanding the exact mechanisms of this scattering and the sensitivity of backscatter to canopy structure (size, distribution, and density of the stem branches) is outside the scope of this study, but it is expected that BioSAR returns measured at multiple incidence angles respond to the volume (and thus the biomass) both in the main trunk and the major branches in hardwood trees of deliquescent form as in this study. This is borne out by Fig. 4, in which there is more variability in the mean NRCS response at lower (e.g., 80–120 tonnes/ha) biomasses (and thus, given similar stocking, smaller diameter boles) than at higher (e.g., 160–200 tonnes/ha) biomasses (with larger diameter boles more on a par with the wavelength of the radar). The lidar data provide additional information on canopy height that is directly related to observed biomass (Fig. 5). As can be seen, at lower biomasses, there is less height variability at a given percentile—nicely complementing the BioSAR data with the opposite situation (lower biomass, more variability). The combined information on canopy height (from the lidar) and tree volume (from the VHF backscatter) provided better estimates of aboveground biomass than either used alone.

The predicted versus observed biomass for the best two BioSAR–PALS and BioSAR–scanning lidar models are provided in Figs. 2 and 3, respectively. These figures, together with the model statistics, show that PALS and BioSAR variables provided slightly better synergy over scanning lidar and BioSAR variables. We can visually infer from the figures that biomass values for pine stands were better predicted with the BioSAR–PALS model, and the predictions for hardwood stands were more or less consistent between the two models. Given the greater coverage and higher pulse density of scanning lidar, it is difficult to put a finger on what caused the PALS data to be the better lidar data set. One of the potential causes might be the time discrepancy between the field data and scanning lidar data collection. Though we do not expect a significant growth in biomass in less than a single year, some trees might have been harvested or lost through natural disturbance during that time, for example. Based on this study, we can argue that simple profiling lidar measurements might suffice when considered jointly with BioSAR volumetric returns, rendering high cost and extensive processing of scanning lidar data unnecessary in estimating stand-based aboveground biomass in hardwood and mixed forests. Cross-validation results reiterate the adequacy of profiling lidar, as the BioSAR–profiling lidar model (CV-R^2=0.70, PRESS=17,358) has better predictive ability than the BioSAR–scanning lidar model (CV-R^2=0.67, PRESS=18,816). Though the time difference of data collection is an issue, this really points in favor of profiling lidar because both BioSAR and PALS sensors are carried on the same aircraft and can be used to collect the data concomitantly.

The most important lidar variables consistently represented in the biomass equations in this study were kurtosis of lidar heights (k_{al} and k_{cl}). Other important lidar variables were the 30th percentile of lidar heights derived from canopy hits (p_{c30l}) for scanning lidar and the 60th and 70th percentiles of heights derived from all hits for profiling lidar. Among

different BioSAR variables, a15, b20, and b40, mean response at forward 15°–20°, backward 20°–25°, and backward 40°–45° angle bins, respectively, were the most useful predictor variables. Our findings are comparable to those of Imhoff et al. (2000), who also identified radar returns at a 25° look angle as best for predicting forest dry biomass. However, Nelson et al. (2007) found the backward-looking BioSAR NRCS response at 8.5° to be the best single BioSAR predictor of stem green biomass using 1.4° Doppler bins rather than the 5° bins used in this study. The observed differences in which BioSAR variable performed best might be attributed to the differences among the heavily managed pine forests studied by Nelson et al. (2007), the hardwood and mixed forests in this study and the oak/pine forests of Big Thicket Forest Preserve in eastern Texas studied by Imhoff et al. (2000).

5 Conclusion

Hardwood and mixed forests comprise a large portion of the earth's land surface; they are important carbon pools with substantial fluxes. Our results suggest that useful synergy can be realized by considering lidar and radar measurements jointly in estimating aboveground biomass in hardwood and mixed forests.

Acknowledgments This material is based in part upon work supported by the U.S. National Science Foundation under grant number IIP-0711992. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. National Science Foundation. We are grateful to Kathryn C. Hollandsworth (Department of Forest Resources and Environmental Conservation, Virginia Polytechnic Institute and State University, USA) for her editorial assistance.

References

Anaya JA, Chuvieco E, Palacios-Orueta A (2009) Aboveground biomass assessment in Colombia: a remote sensing approach. For Ecol Manag 257:1237–1246

Austin JM, Mackey BG, Van Niel KP (2003) Estimating forest biomass using satellite radar: an exploratory study in a temperate Australian Eucalyptus forest. For Ecol Manag 176:575–583

Bortolot ZJ, Wynne RH (2005) Estimating forest biomass using small footprint LiDAR data: an individual tree-based approach that incorporates training data. ISPRS J Photogramm Remote Sens 59:342–360

Boudreau J, Nelson RF, Margolis HA, Beaudoin A, Guindon L, Kimes DS (2008) Regional aboveground forest biomass using airborne and spaceborne LiDAR in Québec. Remote Sens Environ 112:3876–3890

Boyd DS, Foody GM, Currans PJ (1999) The relationship between the biomass of Cameroonian tropical forests and radiation reflected in middle infrared wavelengths (3.0–5.0 mm). Int J Remote Sens 20:1017–1023

Clark ML, Clark DB, Roberts DA (2004) Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. Remote Sens Environ 91:68–89

Chopping M, Moisen GG, Su L, Laliberte A, Rango A, Martonchik JV, Peters DPC (2008) Large area mapping of southwestern forest crown cover, canopy height, and biomass using the NASA Multiangle Scanning Spectro-Radiometer. Remote Sens Environ 112:2051–2063

Dobson MC, Ulaby FT, LeToan T, Beaudoin A, Kasischke ES, Christensen N (1992) Dependence of radar backscatter on coniferous forest biomass. IEEE Trans Geosci Remote Sens 30:412–415

Drake JB, Knox RG, Dubayah RO, Clark DB, Condit R, Blair JB, Hofton M (2003) Above-ground biomass estimation in closed canopy Neotropical forests using lidar remote sensing: factors affecting the generality of relationships. Glob Ecol Biogeogr 12:147–159

Foody GM, Boyd DS, Cutler ME (2003) Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. Remote Sens Environ 85:463–474

Fransson JES, Walter F, Ulander LMH (2000) Estimation of forest parameters using CARABAS-II VHF SAR data. IEEE Trans Geosci Remote Sens 38:720–727

Fraser RH, Li Z (2002) Estimating fire-related parameters in boreal forest using SPOT VEGETATION. Remote Sens Environ 82:95–110

Goel NS (1989) Inversion of canopy reflectance models for estimation of biophysical parameters from reflectance data. In: Asrar G (ed) Theory and application of remote sensing. Wiley, New York, pp 205–252

Hallberg B, Smith-Jonforsen G, Ulander LMH (2005) Measurements on individual trees using multiple VHF SAR images. IEEE Trans Geosci Remote Sens 43:2261–2269

Harrell PA, Kasischke ES, Bourgeau-Chavez LL, Haney EM, Christensen NL (1997) Evaluation of approaches to estimating aboveground biomass in southern pine forests using SIR-C data. Remote Sens Environ 59:223–233

Hyde P, Nelson RF, Kimes D, Levine E (2007) Exploring LiDAR-RaDAR Synergy - Predicting aboveground biomass in a southwestern ponderosa pine forest using LiDAR, SAR, and InSAR. Remote Sens Environ 106:28–38

Imhoff ML (1995) Radar backscatter and biomass saturation: ramifications for global biomass inventory. IEEE Trans Geosci Remote Sens 33:511–518

Imhoff ML, Johnson P, Holford W, Hyer J, May L, Lawrence W, Harcombe P (2000) BioSAR™: an inexpensive airborne VHF multiband SAR system for vegetation biomass measurement. IEEE Trans Geosci Remote Sens 38:1458–1463

Israelsson H, Ulander RMH, Martin T, Askne J, Fransson J, Frölin P, Gustavsson A, Hellsten H (1997) Retrieval of forest stem volume using VHF SAR. IEEE Trans Geosci Remote Sens 35:36–40

Jonforsen GS, Ulander LMH, Luo X (2005) Low VHF-band backscatter from coniferous forests on sloping terrain. IEEE Trans Geosci Remote Sens 43:2246–2260

Lefsky MA, Harding D, Cohen WB, Parker G, Shugart HH (1999) Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. Remote Sens Environ 67:83–98

Lefsky MA, Cohen WB, Parker GG, Harding DJ (2002) Lidar remote sensing for ecosystem studies. Bioscience 52:19–30

Lim K, Treitz P, Wulder M, St-Onge B, Flood M (2003) LiDAR remote sensing of forest structure. Prog Phys Geogr 27:88–106

Lu D (2005) Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon Basin. Int J Remote Sens 26:2509–2525

Lu D, Batistella M (2005) Exploring TM image texture and its relationships with biomass estimation in Rondônia, Brazilian Amazon. Acta Amazon 35:261–268

Lucas RM, Moghaddam M, Cronin N (2000) Microwave scattering from mixed-species forests, Queensland, Australia. IEEE Trans Geosci Remote Sens 42:2142–2159

Luckman A, Baker J, Honzák M, Lucas R (1998) Tropical forest biomass density estimation using JERS-1 SAR: seasonal variation, confidence limits, and application to image mosaics. Remote Sens Environ 63:126–139

Means JE, Acker SA, Harding DJ, Blair JB, Lefsky MA, Cohen WB, Harmon ME, McKee WA (1999) Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the western cascades of Oregon. Remote Sens Environ 67:298–308

Nelson RF, Hyde P, Johnson P, Emessiene B, Imhoff ML, Campbell R, Edwards W (2007) Investigating RaDAR-LiDAR synergy in a North Carolina pine forest. Remote Sens Environ 110:98–108

Peddle DR, Brunke SP, Hall FG (2001) A comparison of spectral mixture analysis and ten vegetation indices for estimating boreal forest biophysical information from airborne data. Can J Remote Sens 27:627–635

Popescu SC, Wynne RH, Nelson RF (2003) Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. Can J For Res 29:564–577

Popescu SC, Wynne RH, Scrivani JA (2004) Fusion of small-footprint lidar and multispectral data to estimate plot-level volume and biomass in deciduous and pine forests in Virginia, U.S.A. For Sci 50:551–565

Sader SA, Waide RB, Lawrence WT, Joyce AT (1989) Tropical forest biomass and successional age class relationships to a vegetation index derived from Landsat TM data. Remote Sens Environ 28:143–156

Santos JR, Lacruz MSP, Araujo LS, Keil M (2002) Savanna and tropical rainforest biomass estimation and spatialization using JERS-1 data. Int J Remote Sens 23:1217–1229

Skole D, Tucker C (1993) Tropical deforestation and habitat fragmentation in the Amazon: satellite data from 1978 to 1988. Science 260:1905–1909

Steininger MK (2000) Satellite estimation of tropical secondary forest above-ground biomass: data from Brazil and Bolivia. Int J Remote Sens 21:1139–1157

van Aardt JAN, Wynne RH, Oderwald RG (2006) Forest volume and biomass estimation using small-footprint lidar-distributional parameters on a per-segment basis. For Sci 52:636–649

van Aardt JAN, Wynne RH, Scrivani JA (2008) Lidar-based mapping of forest volume and biomass by taxonomic group using structurally homogenous segments. Photogramm Eng Remote Sens 74:1033–1044

Zhao K, Popescu S, Nelson R (2009) Lidar remote sensing of forest biomass: a scale-invariant estimation approach using airborne lasers. Remote Sens Environ 113:182–196