Journal of Forestry Research (2014) 25(1): 43–52
DOI 10.1007/s11676-014-0427-4

ORIGINAL PAPER

Estimating *Pinus palustris* tree diameter and stem volume from tree height, crown area and stand-level parameters

C.A. Gonzalez-Benecke • Salvador A. Gezan • Lisa J. Samuelson
Wendell P. Cropper Jr. • Daniel J. Leduc • Timothy A. Martin

Received: 2013-2-29 Accepted: 2013-04-22
© Northeast Forestry University and Springer-Verlag Berlin Heidelberg 2014

Abstract: Accurate and efficient estimation of forest growth and live biomass is a critical element in assessing potential responses to forest management and environmental change. The objective of this study was to develop models to predict longleaf pine tree diameter at breast height (dbh) and merchantable stem volume (V) using data obtained from field measurements. We used longleaf pine tree data from 3,376 planted trees on 127 permanent plots located in the U.S. Gulf Coastal Plain region to fit equations to predict dbh and V as functions of tree height (H) and crown area (CA). Prediction of dbh as a function of H improved when CA was added as an additional independent variable. Similarly, predictions of V based on H improved when CA was included. Incorporation of additional stand variables such as age, site index, dominant height, and stand density were also evaluated but resulted in only small improvements in model performance. For model testing we used data from planted and naturally-regenerated trees located inside and outside the geographic area used for model fitting. Our results suggest that the models are a robust alternative for dbh and V estimations when H and CA are known on planted stands with potential for naturally-regenerated stands, across a wide range of ages. We discuss the importance of these models for use with metrics derived from remote sensing data.

Keywords: Longleaf pine, diameter-height relationships, crown area, individual-tree stem volume, growth and yield modeling

Project funding: This research was supported by the U.S. Department of Defense, through the Strategic Environmental Research and Development Program (SERDP).

The online version is available at http://link.springer.com

C.A. Gonzalez-Benecke (✉) • Salvador A. Gezan
Wendell P. Cropper Jr. • Timothy A. Martin
School of Forest Resources and Conservation, P.O. Box 110410, University of Florida, Gainesville, FL 32611, U. S. A.
Tel.: 352-8460851; Fax: 352-8461277; E-mail: cgonzabe@ufl.edu

Lisa J. Samuelson
School of Forestry and Wildlife Sciences, 3301SFWS Building, Auburn University, Auburn, AL 36849, U. S. A

Daniel J. Leduc
USDA Forest Service, Southern Research Station, Alexandria Forestry Center, 2500 Shreveport Hwy, Pineville, LA 71360, U. S. A

Corresponding editor: Chai Ruihai

Introduction

Forests are a significant proportion of the terrestrial biosphere and are important not only economically, but also for the ecosystem services associated with forested landscapes (Thompson et al. 2011; Goldstein et al. 2012). One important ecosystem service is carbon sequestration (McKinley et al. 2011). Forest biomes are a significant sink for sequestration of atmospheric CO2 with a potentially important role in climate change mitigation (Luyssaert et al. 2008; Fahey et al. 2010). A standard method for assessing forest productivity and carbon sequestration in pine dominated forests has been the use of growth and yield models (e.g. Gonzalez-Benecke et al. 2010, 2011). These models are typically applicable to closed canopy forests, but may be more difficult to apply to savannas and woodlands, particularly those with un-even aged tree populations.

Individual tree biomass, stem volume and carbon content can be estimated with field measurements of stem diameter and height coupled with the application of allometric equations. The equations used for field-based tree biomass estimation typically include height and diameter as independent prediction variables (Satoo and Madgwick 1982). However, direct field measurement of a large number of individual trees is an expensive and time-consuming process. LiDAR (light detection and ranging) is a technique well suited to characterizing trees and other measurements through remote sensing (Lefsky et al. 2002; Roberts et al. 2005; Andersen et al. 2006; Popescu 2007; Dean et al. 2009; Lee et al. 2009, 2010). In addition, satellite interferometry (Ulander et al. 1995) may also be used to characterize tree height profiles.

Given a good algorithm for individual tree detection (Lee et al.

2010; Kaartinen et al. 2012; Li et al. 2012), remote sensing techniques, such as LiDAR, can be used to provide accurate tree height and crown area estimates (Nelson et al. 1988; Nakai et al. 2010), and individual tree diameter in some cases (Popescu 2007; Dalponte et al. 2011). Here, LiDAR estimation of tree height can efficiently feed equations used to calculate biomass of a large number of trees.

Longleaf pine savannas in the southern Coastal Plain are fire-dependent ecosystems characterized by large canopy gaps (Jose et al. 2006). Remote sensing-based methods for individual tree volume estimates are particularly useful for savannas or woodlands that are far from canopy closure. Estimates of savanna biomass have included direct harvesting (Menaut and Cesar 1979) or measurements of individual tree diameters and heights applied to allometric equations (Chen et al. 2003; Sawadogo et al. 2010).

The objective of this study was to develop models to predict longleaf pine tree diameter at breast height (dbh, cm) and merchantable volume (V, m^3). Our models are applicable to direct estimates of stem volume or forest carbon sequestration or as a component of individual-tree-based models of longleaf pine savannas (Drake and Weishampel 2001; Loudermilk et al. 2011) where tree height is the principal state variable that can be derived from remote sensing or obtained from field measurements.

Materials and methods

Data description

We used a dataset consisting of 127 permanent plots measured and maintained by the U.S. Forest Service's Laboratory at Pineville, LA (Goelz and Leduc 2001). Data were collected from regularly remeasured permanent plots in a combination of seven studies exploring the effects of spacing and thinning on longleaf plantations distributed through the Western Gulf Coastal Plain, U.S., from Santa Rosa County in Florida to Jasper County in Texas, representing a large portion of the current range of longleaf pine in the southern Coastal Plain (Goelz and Leduc 2001; Leduc and Goelz 2009).

Total tree height (H, m), dbh, living crown width (CW, m) and V were measured on 3,420 trees. CW was measured in two opposite directions, and living crown area (CA, m^2) was determined assuming tree crown shape as an ellipse. Tree stem volume outside bark (VOB, m^3) was determined by measuring diameter outside bark and its height at 5.08 cm diameter taper steps along the bole from the stump to the 5.08 cm top. Tree stem volume inside bark (VIB, m^3) was determined after estimating diameter inside bark at each diameter taper steps using an equation reported by Gonzalez-Benecke et al. (2013). Stand-level variables including basal area (BA, $m^2 \cdot ha^{-1}$), number of trees per hectare (N, ha^{-1}) and dominant height (Hdom, m) were determined for each plot to be used as additional independent variables in prediction equations. Site index (SI, m), defined as the Hdom at a reference age of 50 years, was calculated using an equation reported by Gonzalez-Benecke et al. (2012). In order to eliminate broken and malformed individuals, trees with form factor F = H/dbh ($m \cdot cm^{-1}$) less than 0.54 $m \cdot cm^{-1}$ and trees with F greater than 13.5 $m \cdot cm^{-1}$ were excluded. Hence, a total of 44 trees were discarded from further analysis.

From the whole dataset, 20 plots (16% of total) were randomly selected and removed to be used for model validation and the rest (i.e. 107 plots) were used for model fitting. The model validation and fitting datasets contained 425 and 2,951 trees, respectively. In order to test the robustness of the models, an independent second source of validation data, that included stands planted outside the geographic range of the data was used. This dataset contains 469 trees measured in four stands on the U.S. Department of Defense's army base at Fort Benning, GA. Selected stands had ages of 12, 21, 64 and 87 years. The 12 year-old and 21 year-old stands corresponded to plantations and the 64 year-old and 87 year-old stands were naturally regenerated. In each stand, four 0.04 ha inventory plots were measured and H and dbh recorded for all 469 trees (the total number of trees measured per each stand was 149, 292, 15 and 13, respectively) and CW was recorded in 120 trees (all trees were measured in the 64 and 87 year-old stand and trees in younger stands were subsampled and included 44 and 48 trees for ages 12 and 21 year-old stands). In a subset of 11 trees (5 from the 21 year-old stand, three from the 64 year-old stand and three from the 87-year-old stand), stem volume over bark was directly measured by destructive harvesting the trees and measuring bole diameter over bark at 2 m steps from stump to a minimum diameter of 5 cm. Stand-level variables BA, N, Hdom and SI were also calculated for each plot. Details of tree and stand characteristics of the three datasets are shown in Table 1. General relationships between dbh and H, CA, dbh and VOB are shown in Fig. 1.

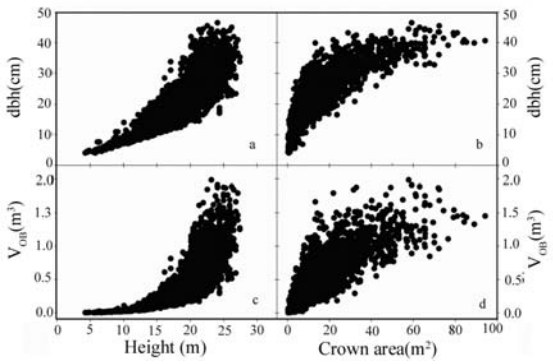

Fig. 1. Relationships between total tree height (H) and crown area (CA) with stem diameter outside bark at 1.37 m height (dbh) (a, b) and stem volume outside bark (V_{OB}) (c, d) for the model development dataset.

Model description

Data from permanent plots with repeated measures, such as the one used in this study, often suffer from within-plot and temporal correlation (Gregoire et al. 1995). In order to address this potential problem of autocorrelation, linear mixed models were fitted to the data by including in the model a random effect of plot (to model spatial correlation) and by specifying an autoregressive error structure (to model temporal correlation). The models fitted

to estimate dbh and V corresponded to a modified version of the model proposed by Nakai et al. (2010), as shown below:

$$\ln(\text{dbh}) = a_1 + a_2 \cdot \ln(\text{H} - 1.37) + a_3 \cdot \ln(\text{CA}) + p_{j1} + \varepsilon_1 \quad (1)$$

$$\ln(\text{V}) = b_1 + b_2 \cdot \ln(\text{H}) + b_3 \cdot \ln(\text{CA}) + p_{j2} + \varepsilon_2 \quad (2)$$

where a_1 to a_3 and b_1 to b_3 are curve fit parameters. The random effects p_{j1} and p_{j2} are effects associated to the j^{th} plot, with $p_{ji} \sim N(0, \sigma_{pi}^2)$, where σ_{pi}^2 corresponds to the variance between plot effects for the dbh and the V models. The error terms ε_1 and ε_2 were modeled assuming and autoregressive error structure of order one, with $\varepsilon_i \sim N(0, \sigma_i^2 \cdot \rho_i^d)$, where σ_i^2 is the residual variance, ρ_i is the year-to-year temporal correlation and d is the lag (in years) between consecutive measurements.

In addition to the explanatory variables H and CA, stand-level variables were included as covariates into the model in order to improve the general equation. The general model used was:

$$\ln(\text{dbh}) = a_1 + a_2 \cdot \ln(\text{H} - 1.37) + a_3 \cdot \ln(\text{CA}) + a_4 \cdot \ln(\text{N}) + a_5 \cdot \ln(\text{H}_{\text{dom}}) + a_6 \cdot \ln(\text{SI}) + p_{j3} + \varepsilon_3 \quad (3)$$

$$\ln(\text{V}) = b_1 + b_2 \cdot \ln(\text{H}) + b_3 \cdot \ln(\text{CA}) + b_4 \cdot \ln(\text{N}) + b_5 \cdot \ln(\text{H}_{\text{dom}}) + b_6 \cdot \ln(\text{SI}) + p_{j4} + \varepsilon_4 \quad (4)$$

where, a_1 to a_6 and b_1 to b_6 are curve fit parameters, and as before, p_{j3} and p_{j4} are the random plot effect terms associated to plot measurements, and ε_3 and ε_4 were modeled using an autoregressive error structure of order one as indicated earlier.

Stand age was not included as a stand-level covariate as it was highly correlated with H_{dom}. Also, BA was not included as its calculation depends on dbh values, which are assumed to be unknown. For the same reason other stand parameters that depend directly on BA, as do stand density index and quadratic mean diameter, were not considered in the general model. For comparison, we also fitted the models presented in Eq. 1 and 2 without CA as a predicting variable, resulting in local models for dbh and V that only rely on H as the predictive variable.

Table 1. Summary of individual-tree and stand-level characteristics for measured longleaf pine

Variable	Model development dataset (n = 2,951)				Model validation dataset inside geographical range (n = 425)				Model validation dataset outside geographical range (Fort Benning) (n = 469)			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
Age	31.9	8.2	20.0	45.0	31.6	7.9	20.0	45.0	28.7	25.5	12	87
dbh	21.3	8.3	4.1	46.5	21.1	7.7	4.3	40.4	16.1	12.1	2.5	57.4
H	18.5	4.2	4.3	27.4	18.4	4.1	4.6	25.9	12.4	7.0	2.4	32.4
CW	3.9	1.8	0.6	11.0	3.8	1.6	0.6	9.3	3.6	268	0.8	15.0
CA	14.4	13.6	0.3	94.3	13.3	11.1	0.3	67.9	15.0	28.2	0.4	175.0
V_{OB}	0.45	0.37	0.00	2.01	0.43	0.32	0.00	1.33	0.82	0.97	0.03	2.95
N	923	495	49	2,056	899	506	204	2,145	819	795	50	2,150
BA	24.7	6.9	6.7	47.6	25.1	6.2	11.9	39.6	14.2	6.3	4.5	24.2
H_{dom}	20.7	3.0	14.9	26.9	20.5	2.9	15.5	26.4	18.3	8.5	8.7	32.4
SI	27.0	2.2	21.3	33.6	26.8	2.4	21.5	33.3	22.2	3.3	13.5	27.4

Age: stand age (yr.); dbh: average diameter at 1.37 m height (cm); H: average tree height (m); CW: average tree crown width (m); CA: average tree crown area (m^2); V_{OB}: average stem volume outside bark (m^3); N: trees per hectare (ha^{-1}); BA: stand basal area (m^2 ha^{-1}); H_{dom}: average height of dominant and codominant trees (m); SI: site index at base age of 50 yrs. (m); SD: standard deviation; n: number of trees. **Note**: At Fort Benning, n = 120 for CA and n = 11 for V_{OB}.

Model validation

The predictive ability of the equations previously described, including the reduced local model without CA, was compared against the data from the plots selected from the validation database. The models were also evaluated against the data from the four stands in Fort Benning, GA. In all cases, four measures of accuracy were used to evaluate the goodness-of-fit between observed and predicted (simulated) values for each variable from the model validation dataset. These were: (1) mean absolute error (MAE); (2) root mean square error (RMSE); (3) mean bias error (Bias); and (4) coefficient of determination (R^2) (Fox 1981; League and Green 1991; Kobayashi and Salam 2000).

All statistics were obtained using SAS 9.3 (SAS Inc., Cary, NC, USA). The procedure MIXED was used in order to model spatial and temporal correlations, and statistical comparison of fitted models was done using a likelihood ratio test. The variance inflation factor (VIF) was calculated to detect multicollinearity between predicting variables, discarding all variables included in the model with VIF larger than 5 (Neter et al. 1996).

Results

Model Fitting

The covariance structure estimated from the analysis indicated a significant spatial and temporal correlation for all dbh, VOB and

VIB models. The year-to-year residual correlation values ranged from 0.85 to 0.99 where the largest values were obtained, as expected, for dbh (Table 2). The average relative magnitudes of the plot variance components (in relation to the error component) showed moderate levels of the spatial component but important changes across models. The largest relative plot components were found on models that did not depend on CA or any stand-level variable (i.e. models dbh1, VOB1 and VIB1), reflecting the effect of these extra factors on explaining additional between plot variability.

The prediction equations and parameter estimates for planted longleaf pine trees are presented in Table 3. All parameter estimates were significant at p <0.001. Parameter estimates for intercepts in all equations (a1, b1 and c1) include the correction proposed by Snowdon (1991) for logarithm transformation of the response variable.

Table 2. Variance components and correlation estimates from fitted models of dbh and stem volume predicted from tree height and stand characteristics for planted longleaf pine trees (n = 2,951)

Model		Covariance Component	Parameter Estimate	SE	p-value
dbh1	$\ln(dbh) = a_1 + a_2 \cdot \ln(H - 1.37)$	σ_p^2	0.031549	0.004655	< 0.0001
		ρ	0.988316	0.000776	< 0.0001
		σ^2	0.037139	0.001473	< 0.0001
dbh2	$\ln(dbh) = a_1 + a_2 \cdot \ln(H - 1.37) + a_3 \cdot \ln(CA)$	σ_p^2	0.002741	0.000635	< 0.0001
		ρ	0.905937	0.008053	< 0.0001
		σ^2	0.013623	0.000463	< 0.0001
dbh3	$\ln(dbh) = a_1 + a_2 \cdot \ln(H - 1.37) + a_3 \cdot \ln(CA) + a_4 \cdot \ln(N) + a_5 \cdot \ln(H_{dom}) + a_6 \cdot \ln(SI)$	σ_p^2	0.007011	0.001860	< 0.0001
		ρ	0.954079	0.005971	< 0.0001
		σ^2	0.015963	0.000818	< 0.0001
$V_{OB}1$	$\ln(V_{OB}) = b_1 + b_2 \cdot \ln(H)$	σ_p^2	0.094202	0.014834	< 0.0001
		ρ	0.974914	0.001870	< 0.0001
		σ^2	0.125801	0.005271	< 0.0001
$V_{OB}2$	$\ln(V_{OB}) = b_1 + b_2 \cdot \ln(H) + b_3 \cdot \ln(CA)$	σ_p^2	0.006006	0.001363	< 0.0001
		ρ	0.849064	0.011258	< 0.0001
		σ^2	0.050265	0.001521	< 0.0001
$V_{OB}3$	$\ln(V_{OB}) = b_1 + b_2 \cdot \ln(H) + b_3 \cdot \ln(CA) + b_4 \cdot \ln(N) + b_5 \cdot \ln(H_{dom}) + b_6 \cdot \ln(SI)$	σ_p^2	0.002819	0.000975	0.0019
		ρ	0.875927	0.009473	< 0.0001
		σ^2	0.049234	0.001584	< 0.0001
$V_{IB}1$	$\ln(V_{IB}) = b_1 + b_2 \cdot \ln(H)$	σ_p^2	0.108387	0.017007	< 0.0001
		ρ	0.976026	0.001762	< 0.0001
		σ^2	0.141412	0.005899	< 0.0001
$V_{IB}2$	$\ln(V_{IB}) = b_1 + b_2 \cdot \ln(H) + b_3 \cdot \ln(CA)$	σ_p^2	0.006982	0.001592	< 0.0001
		ρ	0.852428	0.011022	< 0.0001
		σ^2	0.055810	0.001698	< 0.0001
$V_{IB}3$	$\ln(V_{IB}) = b_1 + b_2 \cdot \ln(H) + b_3 \cdot \ln(CA) + b_4 \cdot \ln(N) + b_5 \cdot \ln(H_{dom}) + b_6 \cdot \ln(SI)$	σ_p^2	0.003301	0.001141	0.0019
		ρ	0.879683	0.009257	< 0.0001
		σ^2	0.055202	0.001793	< 0.0001

dbh: diameter outside-bark at 1.37 m height (cm); H: total tree height (m); CA: average tree crown area (m^2); V_{OB}: stem volume outside-bark up to 5.08 cm diameter limit (m^3); V_{IB}: stem volume inside-bark up to 5.08 cm diameter limit (m^3); N: trees per hectare (ha^{-1}); H_{dom}: average height of dominant and codominant trees (m); SI: site index at base age of 50 yrs. (m). σ_p^2 is the plot variance component; ρ is the year-to-year residual temporal correlation; and σ^2 is the error variance.

The model that estimates dbh using H as the only dependent variable (local model dbh1) had a coefficient of variation (CV, RMSE as a percentage of observed mean value) of 18.4% and R^2 of 0.77 (Table 3). When CA was included into the model (local model dbh2), the fit of the model was improved considerably reducing CV to 12.1% and increasing R^2 to 0.90. When stand parameters N, H_{dom} and SI were also included into the model (general model dbh3) all variables were significant in the final model (Eq. 3). This final general model showed little improvement when compared with the local model dbh2 (CV of 11.1% and R^2 of 0.92 (Table 3). The parameter SI had a positive effect on dbh (positive value of parameter estimate): as site productivity increased, trees had larger dbh for any given H and CA. Contrary, the parameters N and H_{dom} had a negative effect on dbh (negative value of parameter estimates): in stands with higher stand density or in older stands, trees had smaller dbh for any given fixed value of H, CA and SI. In all cases multicollinearity between explanatory variables was small (VIF < 3.37; data not shown).

The models that estimate V_{OB} and V_{IB} using H as the only independent variable (local models $V_{OB}1$ and $V_{IB}1$, respectively) had an R^2 of 0.78 and 0.77, respectively (Table 3). The fit of

both models improved when CA was included (local models $V_{OB}2$ and $V_{IB}2$), reducing CV by about 38% compared to models $V_{OB}1$ and $V_{IB}1$, and increasing R^2 to 0.91 and 0.92, respectively. Similar to the dbh3 model, all stand-level variables were significant in the final general model (Eq. 4). These final general models $V_{OB}3$ and $V_{IB}3$ showed little improvement when compared with local models $V_{OB}2$ and $V_{IB}2$ (Table 3). The parameters N, H_{dom} and SI had a negative effect on V_{OB} and V_{IB}: as stand density and site productivity increased, likely due to tapering changes, trees had smaller V for any given fixed value of H and CA. In all cases multicollinearity between explanatory variables was small (VIF < 3.54; data not shown).

Table 3. Parameter estimates and fit statistics of equations for predicting dbh and stem volume from height and stand characteristics for planted longleaf pine trees (n = 2,951)

Model		Parameter	Parameter estimate	SE	R^2	RMSE	CV (%)
dbh1	$\ln(\text{dbh}) = a_1 + a_2 \cdot \ln(\text{H} - 1.37)$	a_1 *	0.403079	0.032711	0.771	3.93	18.39
		a_2	0.970048	0.009821			
dbh2	$\ln(\text{dbh}) = a_1 + a_2 \cdot \ln(\text{H} - 1.37) + a_3 \cdot \ln(\text{CA})$	a_1 *	0.342191	0.027786	0.901	2.59	12.11
		a_2	0.804174	0.011660			
		a_3	0.185289	0.003761			
dbh3	$\ln(\text{dbh}) = a_1 + a_2 \cdot \ln(\text{H} - 1.37) + a_3 \cdot \ln(\text{CA}) + a_4 \cdot \ln(\text{N}) + a_5 \cdot \ln(\text{H}_{\text{dom}}) + a_6 \cdot \ln(\text{SI})$	a_1 *	0.288848	0.204670	0.916	2.38	11.11
		a_2	0.930729	0.015051			
		a_3	0.132246	0.004031			
		a_4	-0.065363	0.005838			
		a_5	-0.337178	0.026633			
		a_6	0.393273	0.058204			
$V_{OB}1$	$\ln(V_{OB}) = b_1 + b_2 \cdot \ln(\text{H})$	b_1 *	-9.944543	0.078449	0.784	0.17	38.21
		b_2	3.123691	0.024536			
V_{OB} 2	$\ln(V_{OB}) = b_1 + b_2 \cdot \ln(\text{H}) + b_3 \cdot \ln(\text{CA})$	b_1 *	-9.686492	0.062915	0.909	0.11	24.84
		b_2	2.661325	0.025422			
		b_3	0.372844	0.007370			
V_{OB} 3	$\ln(V_{OB}) = b_1 + b_2 \cdot \ln(\text{H}) + b_3 \cdot \ln(\text{CA}) + b_4 \cdot \ln(\text{N}) + b_5 \cdot \ln(\text{H}_{\text{dom}}) + b_6 \cdot \ln(\text{SI})$	b_1 *	-6.480444	0.308427	0.919	0.10	23.38
		b_2	2.953958	0.031336			
		b_3	0.327533	0.008107			
		b_4	-0.092607	0.011486			
		b_5	-0.820639	0.053347			
		b_6	-0.259927	0.087629			
$V_{IB}1$	$\ln(V_{IB}) = b_1 + b_2 \cdot \ln(\text{H})$	b_1 *	-10.531556	0.082652	0.768	0.14	41.48
		b_2	3.216861	0.025772			
$V_{IB}2$	$\ln(V_{IB}) = b_1 + b_2 \cdot \ln(\text{H}) + b_3 \cdot \ln(\text{CA})$	b_1 *	-10.270262	0.066367	0.902	0.09	27.02
		b_2	2.729643	0.026808			
		b_3	0.395527	0.007778			
V_{IB} 3	$\ln(V_{IB}) = b_1 + b_2 \cdot \ln(\text{H}) + b_3 \cdot \ln(\text{CA}) + b_4 \cdot \ln(\text{N}) + b_5 \cdot \ln(\text{H}_{\text{dom}}) + b_6 \cdot \ln(\text{SI})$	b_1 *	-6.960066	0.328228	0.913	0.08	25.43
		b_2	3.025198	0.033191			
		b_3	0.346722	0.008570			
		b_4	-0.100300	0.012157			
		b_5	-0.833450	0.056431			
		b_6	-0.264805	0.093337			

dbh: diameter outside-bark at 1.37 m height (cm); H: total tree height (m); CA: average tree crown area (m^2); V_{OB}: stem volume outside-bark up to 5.08 cm diameter limit (m^3); V_{IB}: stem volume inside-bark up to 5.08 cm diameter limit (m^3); N: trees per hectare (ha^{-1}); H_{dom}: average height of dominant and codominant trees (m); SI: site index at base age of 50 yrs. (m); SE: standard error; R^2: coefficient of determination; RMSE; root mean square error; CV: coefficient of variation (100·RMSE/mean). For all parameter estimates: $p < 0.001$. **Note**: Parameters estimates for a_1 and b_1 include the correction proposed by Snowdon (1991).

Model validation

The relationship between predicted and observed values of V_{OB} and dbh using the local models, that only depends on H (local models $V_{OB}1$ and dbh1; Fig. 2a and 2b, respectively), showed a tendency to underestimate trees with V_{OB} and dbh larger than about 1 m^3 and 30 cm, respectively. When CA was included in the local models the relationship between observed and predicted values improved considerably (local models $V_{OB}2$ and dbh2; Fig. 2c and 2d, respectively) and there was no important departure in residuals. When stand parameters N, H_{dom} and SI were also

included into the local models $V_{OB}2$ and dbh2, the relationship between observed and predicted values resulted in little improvement (general models $V_{OB}3$ and dbh3; Fig. 2e and 2f, respectively).

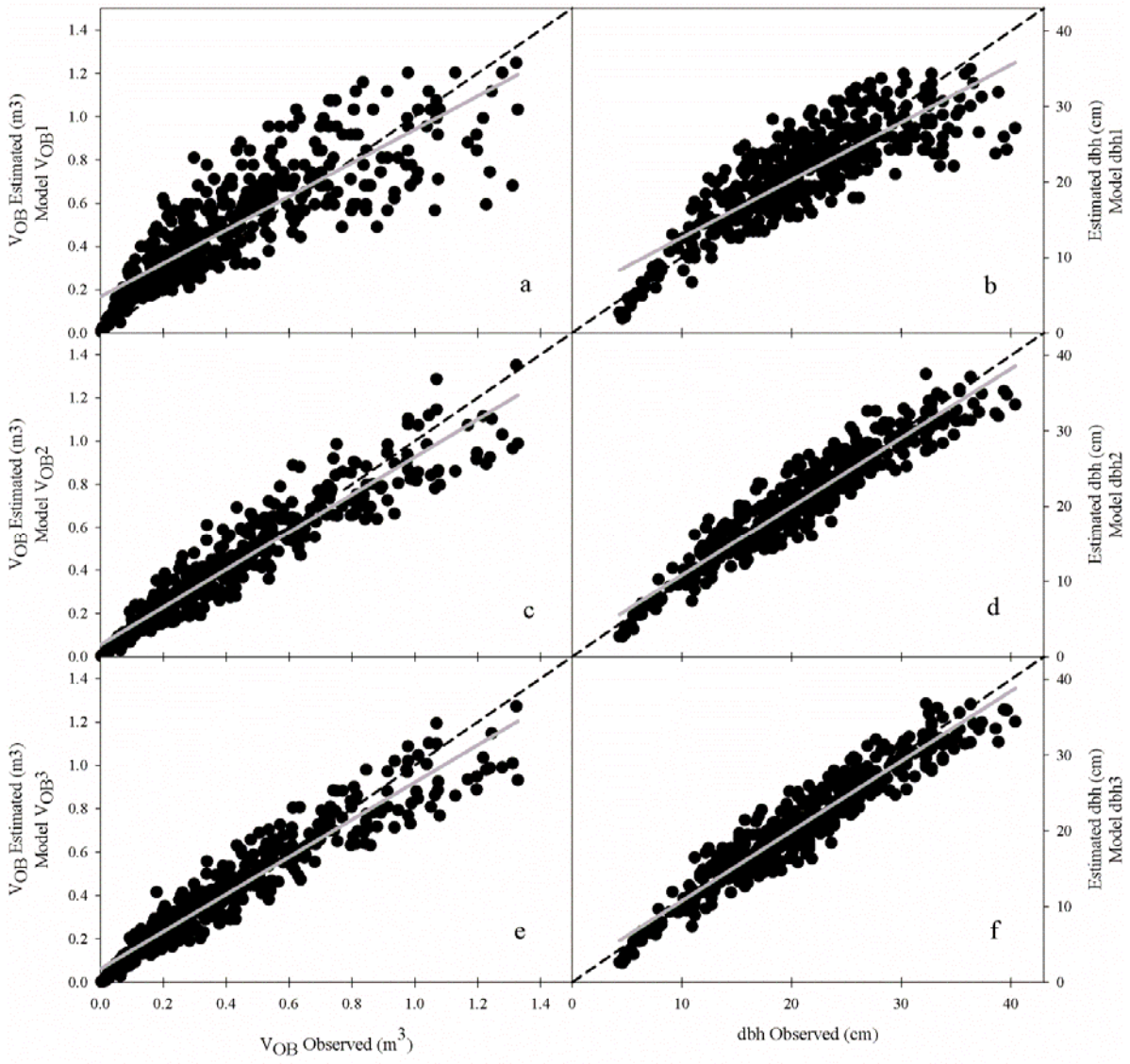

Fig. 2. Examples of validation of stem volume outside bark (V_{OB}) (a, c, e) and stem diameter outside bark at 1.37 m height (dbh) (b, d, f) models using data inside geographical range of fitting plots. Observed v/s predicted (simulated) values for V_{OB} and dbh using local model $V_{OB}1$ (a) and dbh1 (b) (only use H as independent variable), local model $V_{OB}2$ (c) and dbh2 (d) (use H and CA as independent variables), and general model $V_{OB}3$ (e) and dbh3 (f) (use H, CA, N, H_{dom} and SI as independent variables). Grey lines represents linear regression fit.

All model performance tests showed that agreement between dbh and V observed and estimated values improved when CA was included in the local model (Table 4). For example, using the validation dataset, MAE and RMSE for dbh estimations were reduced from 15 and 20% (local model dbh1) to about 9 and 11% (local model dbh2), respectively; and R^2 increased from 0.71 to 0.90. In the case of VOB estimations, MAE and RMSE were reduced from 33 and 43% (local model VOB1) to 16 and 24% (local model VOB2), respectively; and R^2 increased from 0.70 to 0.90. Bias was highly reduced when CA was included into the local model to estimate VOB (from 15% overestimations to 0.8% underestimations). Performance tests showed that dbh, VOB and VIB estimations that used H and CA as explanatory variables with little improved when stand parameters were added in the general models (Table 4).

Model validation using external data

When models to estimate dbh and V_{OB} were evaluated using trees measured outside the geographical range of the model development dataset (i.e. Fort Benning, GA), across stand ages, there was no difference between observed and predicted values for any of the predicting models reported ($p > 0.21$, paired t-test). For dbh estimations, the model that only depended on H (local model dbh1; Table 4; Fig. 3a) shows a tendency to underestimate dbh, with larger absolute errors as tree dbh increased. Similar to

previous results, when CA was included the relationship between observed and estimated values improved considerably (local model dbh2; Table 4; Fig. 3c). As before, there was little improvement when stand parameters N, H_{dom} and SI were included in the local model dbh2 (general model dbh3; Table 4; Fig. 3e). Fig. 3b, 3d and 3f show plot of residuals (as a proportion of observed values) versus observed dbh; here, CA reduced residuals dispersion but stand-level parameters resulted in an increase in Bias for younger stands. For dbh estimations, when CA was included, Bias was reduced from 4.2 to 1.8% overestimations. For V_{OB} estimations, the Bias reduction was from 15.3 to 5.7% underestimations when CA was included (Table 4). As only 11 trees were measured for V_{OB} at Fort Benning, the model validation was only carried out for means differences, and as stated previously, there was no difference between mean observed and predicted values for any of the models to predict V_{OB}.

Table 4. Summary of model validation statistics for dbh, V_{OB} and V_{IB}, estimations using validation datasets inside geographical range (n = 425) and outside geographical range (Fort Benning; n = 120 for dbh; n = 11 for V_{OB}) of model development plots.

Validation Dataset	Model	Independent Variables	O	P	MAE	RMSE	Bias	R^2
Inside geographical range	dbh1	H	21.15	21.06	3.16 (15.0)	4.16 (19.7)	-0.08 (-0.4)	0.708
	dbh2	H, CA	21.15	20.94	1.90 (9.0)	2.44 (11.5)	-0.20 (-1.0)	0.900
	dbh3	H, CA, N, H_{dom}, SI	21.15	21.01	1.83 (8.6)	2.31 (10.9)	-0.14 (-0.7)	0.910
	$V_{OB}1$	H	0.428	0.493	0.139 (32.6)	0.185 (43.3)	0.065 (15.3)	0.702
	$V_{OB}2$	H, CA	0.428	0.424	0.070 (16.4)	0.104 (24.3)	-0.003 (-0.8)	0.896
	$V_{OB}3$	H, CA, N, H_{dom}, SI	0.428	0.427	0.066 (15.5)	0.097 (22.6)	-0.001 (-0.2)	0.913
	$V_{IB}1$	H	0.312	0.363	0.109 (34.9)	0.145 (46.7)	0.051 (16.5)	0.684
	$V_{IB}2$	H, CA	0.312	0.310	0.054 (17.4)	0.082 (26.3)	-0.002 (-0.7)	0.890
	$V_{IB}3$	H, CA, N, H_{dom}, SI	0.312	0.311	0.052 (16.5)	0.076 (24.5)	0.000 (-0.1)	0.908
Outside geographical range (Fort Benning)	dbh1	H	16.80	17.51	3.27 (19.5)	4.51 (26.9)	0.71 (4.2)	0.920
	dbh2	H, CA	16.80	17.10	2.05 (12.2)	3.36 (20.0)	0.30 (1.8)	0.929
	dbh3	H, CA, N, H_{dom}, SI	16.80	18.10	3.17 (18.8)	4.00 (23.8)	1.30 (7.8)	0.916
	$V_{OB}1$	H	0.824	0.698	0.193 (23.4)	0.327 (39.7)	-0.126 (-15.3)	0.947
	$V_{OB}2$	H, CA	0.824	0.777	0.251 (30.5)	0.381 (46.3)	-0.047 (-5.7)	0.846
	$V_{OB}3$	H, CA, N, H_{dom}, SI	0.824	0.780	0.219 (26.6)	0.328 (39.8)	-0.044 (-5.3)	0.877

dbh: diameter outside-bark at 1.37 m height (cm); H: total tree height (m); CA: average tree crown area (m^2); V_{OB}: stem volume outside-bark up to 5.08 cm diameter limit (m^3); V_{IB}: stem volume inside-bark up to 5.08 cm diameter limit (m^3); N: trees per hectare (ha^{-1}); H_{dom}: average height of dominant and codominat trees (m); SI: stand site index (m); O: mean observed value; P: mean predicted value; *n*: number of observations; MAE: mean absolute error; RMSE: root of mean square error; Bias: absolute bias estimator; R^2: coefficient of determination. Values in parenthesis are percentage relative to observed mean.

Note: MAE, RMSE and Bias are presented in the same units as dependent variable.

Discussion

The set of prediction equations for longleaf pine trees reported in this study represents a useful tool for the study and management of the species. We anticipate that the main use of the equations reported here will be for cases when stem diameter is unknown, and tree height and crown width measurement s are available, as is the case when data are obtained though remote sensing, such as LiDAR or satellite interferometry. General and local models are presented for dbh, V_{OB} and V_{IB} estimations. The user should decide which model to use depending on data availability (stand and tree level) and the desired accuracy.

The inclusion of CA into the models greatly improved the accuracy of the predictions. The relationships between dbh, H, V and CA, which are shown in Fig. 1, help to explain that response: as longleaf pines age, they continue growing in diameter and volume and also developing larger crowns, but they tend to plateau in height. When H and CA are known, small improvements were found when additional stand parameters were included, implying that CA implicitly incorporates the effects of stand stocking and productivity on the allometric relationships of longleaf pine trees.

The additional elements incorporated to model spatial and temporal correlations resulted in improved fittings, where, for all variables and models, a significant spatial components were noted, and, as expected, its relative importance decreased as other explanatory variables where included. In addition, important year-to-year correlations were detected reflecting the non-independent nature of this data.

The observed relationship between dbh and H on longleaf pine trees indicated that dbh continues growing when H reached its asymptote (see Fig. 1a) and dbh estimations from known large H values had low accuracy (see Tables 1 and 2). Similar observations have been reported for other pine species, such as *Pinus sylvestris* (Kalliovirta and Tokola 2005) and *Pinus radiata* (Bi et al. 2012). The final model selected for dbh estimations showed similar fitting to non-linear models proposed by Bi et al. (2012), or a linear model (after root-square transformation) proposed by Kalliovirta and Tokola (2005). The inclusion of CA into the models to estimate dbh and V greatly improved the accuracy of

the predictions. Similar results were reported for *Pinus densiflora* and *Cryptomeria japonica* by Nakai et al. (2010) and for *P. sylvestris, Picea abies* and *Betula pendula* by Kalliovirta and Tokola (2005).

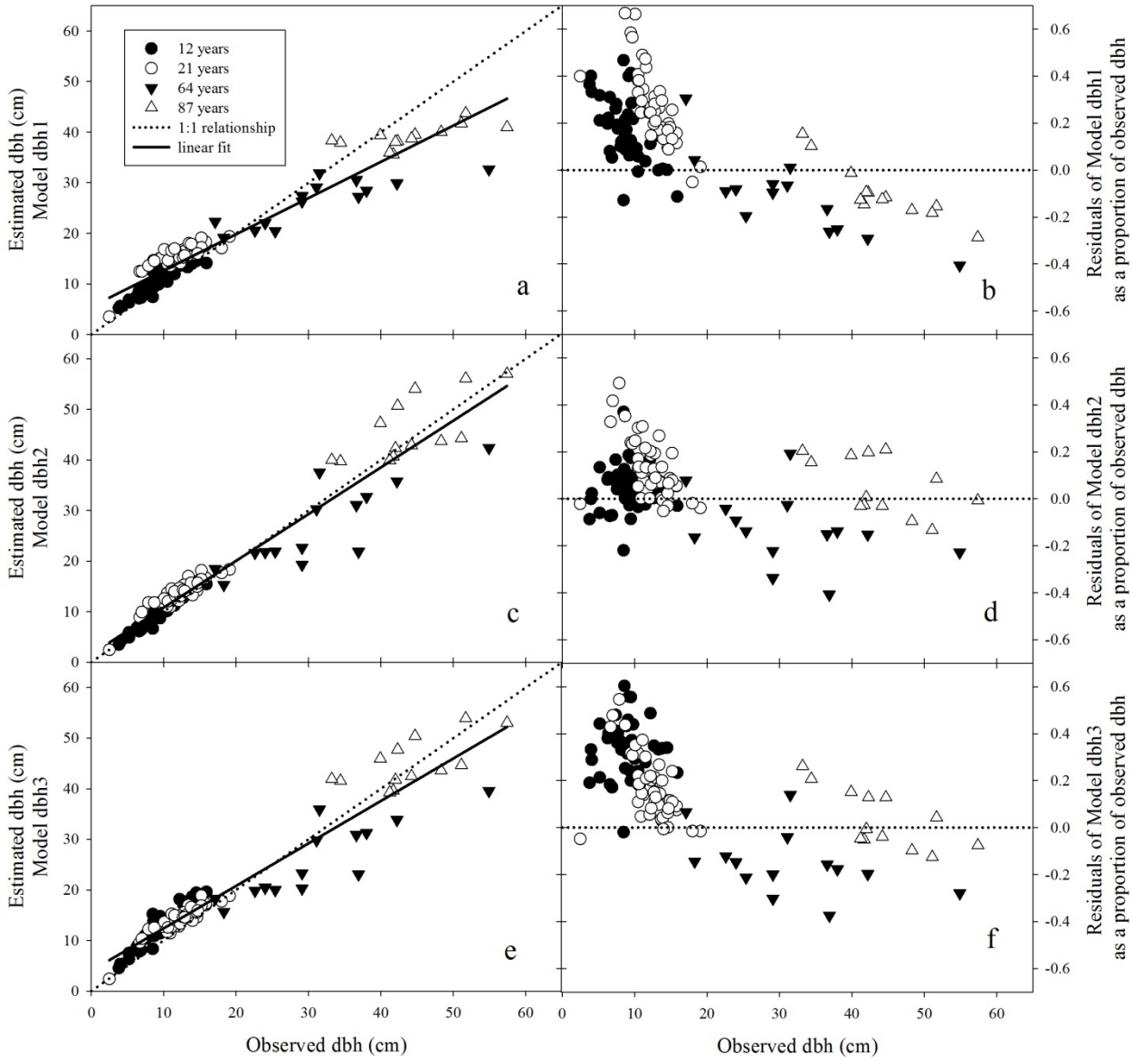

Fig. 3. Validation of stem diameter outside bark at 1.37 m height (dbh) models using data outside geographical range of fitting plots (Fort Benning, GA), for stands of different ages. Observed versus predicted (simulated) values (a, c and e) and residuals (predicted-observed) versus observed values (b, d and f) of dbh using model dbh1 (a, b) (only use H as independent variable), local model dbh2 (c, d) (use H and CA as independent variables), and general model dbh3 (e, f) (use H, CA, N, H_{dom} and SI as independent variables).

The stand parameters included in the models reported by Kalliovirta and Tokola (2005) for *P. sylvestris* showed only small improvements in model performance. In our general models, stand variables related to density (N) and site quality (H_{dom} and SI) were significant into the final model selected, but only marginally improved model fitting.

When the equations to estimate dbh and V_{OB} were tested in a dataset obtained in stands located outside the geographical zone of the data used for model fitting (i.e. Fort Benning, GA), the results support the robustness of the models that included CA, even in naturally-regenerated stands. When stand parameters were included into the general models, there were minimal further improvements in dbh prediction for any stand age tested. Due to costs constrains, our validation of V_{OB} with data from Fort Benning was carried out only on 11 trees with ages ranging from 21 to 87 years.

Our results suggest that the fitted models are a robust tool for dbh and V_{OB} estimations when H and CA measurements are available on planted stands (and perhaps naturally-regenerated as well) with non-overlapping crowns across a wide range of ages. One alternative for dbh and V estimation from H measurements consists on determining dbh by solving the inverse of an equa-

tion fitted for H estimations from known dbh, and determining V by solving an individual tree-level equation fitted for V estimations from dbh^2 or $dbh^2 \cdot H$, where dbh was previously solved from H. This option is not recommended as completely different response variables is the one minimized originally (i.e. H instead of dbh). For comparison purposes we estimated dbh and V_{OB} using local and general equations reported by Gonzalez-Benecke et al. (2013). These estimates had similar Bias but, as expected, larger MAE and RMSE (data not shown). Even though in the general model stand-level parameters improved the fitting, the estimates still showed a considerable Bias for trees with dbh larger than about 40 cm (data not shown). Therefore, we do not recommend the use of this alternative estimation methodology. The same previous analysis applies to allometric equations to estimate biomass from dbh^2 or $dbh^2 \cdot H$.

Forestry and forest ecology studies often produce estimates of the standing biomass or of carbon stocks and how these variables change over time. These estimates are relevant to assessments of economic value as well as non-market issues such as carbon sequestration. Climate change mitigation through reduced loss of stored carbon (Putz et al. 2008) and provision of other ecosystem services (Tallis and Polasky 2009) are important applications that require accurate stand estimations. Traditional estimates using allometric equations require time consuming measurements of large numbers of individual trees. Our models provide a straightforward method of estimating tree diameter, volume and later biomass without necessarily resorting to ground-based measurements. Work is underway to develop tree-level biomass equations that use H and CA as independent variables.

An ideal example of potential model applicability is in situations where LiDAR data have been collected for large areas of longleaf pine forests with no overlapping crowns (Loudermilk et al. 2011). In this case, a typical application of the model would be for estimations of stand-level volume and/or biomass. LiDAR techniques are now well developed for tree-level H and CA indirect measurements (Nilsson 1996; Popescu et al. 2003; Popescu and Wynne 2004; Popescu 2007; Li et al. 2012). After applying equations 1 and 2, the dbh and V can be determined for each tree. To determine biomass, allometric equations, such as those reported by Baldwin and Saucier (1983), can be applied, as dbh and H, the independent variables of the biomass equations, would be known. After applying the appropriate scaling factor, stand level basal area, volume and stand- level biomass can be easily determined.

In conclusion, the models reported in this study performed well for both independent validation datasets within and outside of the geographical range of the fitting dataset. The equations presented here provide a valuable tool for supporting present and future longleaf pine research and management decisions, and for facilitating the use of remote sensing data for quantifying longleaf pine stand structure and function. Work is underway to assess stand-level estimates using field-based and LiDAR data collected on the same stands.

Acknowledgements

This research was supported by the U.S. Department of Defense, through the Strategic Environmental Research and Development Program (SERDP). The authors acknowledge the U.S. Forest Service Southern Research Station for their assistance and for providing the long term datasets. Special thanks go to Mr. Tom Stokes, Ms. Ann Huyler, Mr. Justin Rathal and Mr. Jake Blackstock for their help with field data collection.

References

Andersen H-E, Reutebuch SE, McGaughey RJ. 2006. A rigorous assessment of tree height measurements obtained using airborne lidar and conventional field methods. *Canadian Journal of Remote Sensing*, **32**: 355−366.

Baldwin VC, Saucier JR. 1983. Aboveground weight and volume of unthinned, planted longleaf pine on West Gulf forest sites. USDA Forest Service, Southern Forest Experiment Station, New Orleans, LA. Research Paper SO-191, p.25.

Bi H, Fox JC, Li Y, Lei Y, Pang Y. 2012. Evaluation of nonlinear equations for predicting diameter from tree height. *Canadian Journal of Forest Research*, **42**:1−18.

Chen X, Hutley LB, Eamus D. 2003. Carbon balance of a tropical savanna of northern Australia. *Oecologia*, **137**: 405−416.

Dalponte M, Bruzzone L, Gianelle D. 2011. System for the estimation of single-tree stem diameter and volume using multireturn LIDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, **49**: 2481−2490.

Dean TJ, Cao QV, Roberts SD, Evans DL. 2009. Measuring heights to crown base and crown median with LiDAR in a mature, even-aged loblolly pine stand. *Forest Ecology and Management*, **257**: 126−133.

Drake JB, Weishampel JF. 2001. Simulating vertical and horizontal multifractal patterns of a longleaf pine savanna. *Ecological Modelling*, **145**: 129−142.

Fahey TJ, Woodbury PB, Battles JJ, Goodale CL, Hamburg SP, Ollinger SV, Woodall CW. 2010. Forest carbon storage: ecology, management, and policy. *Frontiers in Ecology and the Environment*, **8**:245−252.

Fox DG. 1981. Judging air quality model performance. *Bulletin of the American Meteorological Society*, **62**: 599−609.

Goelz JC, Leduc DJ. 2001. Long-term studies on development of longleaf pine plantations. *In*: J. S. Kush (ed), *Proceedings of the Third Longleaf Alliance Regional Conference, Forests for our Future*, October 16-18, 2000, Alexandria, LA. The Longleaf Alliance and Auburn University, Auburn, AL, pp.116−118.

Goldstein JH, Caldarone G, Duarte TK, Ennaanay D, Hannahs N, Mendoza G, Polasky S, Wolny S, Daily GC. 2012. Integrating ecosystem-service tradeoffs into land-use decisions. *Proceedings of the National Academy of Sciences of the United States of America*, **109**: 7565−7570.

Gonzalez-Benecke CA, Martin TA, Cropper Jr. WP, Bracho R. 2010. Forest management effects on *in situ* and *ex situ* slash pine forest carbon balance. *Forest Ecology and Management*, **260**: 795−805.

Gonzalez-Benecke CA, Martin TA, Jokela EJ, de la Torre R. 2011. A flexible hybrid model of life cycle carbon balance for loblolly pine (*Pinus taeda* L.) management systems. *Forests*, **2**: 749−776.

Gonzalez-Benecke CA, Gezan SA, Leduc DJ, Martin TA, Cropper Jr. WP, Samuelson LJ. 2012. Modeling survival, yield, volume partitioning and their response to thinning for longleaf pine plantations. *Forests*, **3**(4): 1104−1132.

Gonzalez-Benecke CA, Gezan SA, Martin TA, Cropper Jr. WP, Samuelson LJ, Leduc DJ. 2013. Individual tree diameter, height and volume functions for

longleaf pine. *Forest Science*, dx.doi.org/10.5849/forsci.12-144.

Gregoire TJ, Schabenberger O, Barrett JP. 1995. Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. *Canadian Journal of Forest Research,* **25**:137−156.

Jose S, Jokela E, Miller D. 2006. The longleaf pine ecosystem: An Overview. *In*: S. Jose, E. Jokela and D. Miller (eds), *The longleaf pine ecosystem. Ecology, silviculture, and restoration*. New York: Springer, pp 3−8.

Kaboyashi K, Salam MU. 2000. Comparing simulated and measured values using mean squared deviation and its components. *Agronomy Journal*, **92**: 345−352.

Kalliovirta J, Tokola T. 2005. Functions for estimating stem diameter and tree age using tree height, crown width and existing stand database information. *Silvae Fenica*, **39**: 227−248.

Kaartinen H, Hyyppä J, Yu X, Vastaranta M, Hyyppä H, Kukko A, Holopainen M, Heipke C, Hirschmugl M, Morsdorf F, Næsset E, Pitkänen J, Popescu S, Solberg S, Wolf BM, Wu JC. 2012. An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing*, **4**: 950−974.

Leduc D, Goelz J. 2009. A height-diameter curve for longleaf pine plantations in the Gulf Coastal Plain. *Southern Journal of Applied Forestry*, **33**: 164−170.

Lee H, Slatton KC, Roth BE, Cropper Jr. WP. 2009. Prediction of forest canopy light interception using three-dimensional airborne LiDAR data. *International Journal of Remote Sensing*, **30**: 189−207.

Lee H, Slatton KC, Roth BE, Cropper Jr. WP. 2010. Adaptive clustering of airborne LiDAR data to segment individual tree crowns in managed pine forests. *International Journal of Remote Sensing,* **31**: 117−139.

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

Li W, Guo Q, Jakubowski MK, Kelly M. 2012. A new method for segmenting individual trees from the Lidar point cloud. *Photogrametric Engineering and Remote Sensing*, **78**: 75−84.

Loague K, Green RE. 1991. Statistical and graphical methods for evaluating solute transport models: Overview and application. *Journal of Contaminant Hydrology*, **7**: 51−73.

Loudermilk EL, Cropper Jr. WP, Mitchell RJ, Lee H. 2011. Longleaf pine and hardwood dynamics in a fire-maintained ecosystem: A simulation approach. *Ecological Modelling*, **222**: 2733−2750.

Luyssaert S, Schulze ED, Boerner A, Knohl A, Hessenmoeller D, Law BE, Ciais P, Grace J. 2008. Old-growth forests as global carbon sinks. *Nature,* **455**: 213−215.

McKinley DC, Ryan MG, Birdsey RA, Giardina CP, Harmon ME, Heath LS, Houghton RA, Jackson RB, Morrison JF, Murray BC, Pataki DE, Skog KE. 2011. A synthesis of current knowledge on forests and carbon storage in the United States. *Ecological Applications*, **21**: 1902−1924.

Menaut JC, Cesar J. 1979. Structure and primary productivity of Lamto savannas, Ivory Coast. *Ecology,* **60**: 1197−1210.

Nakai Y, Hosoi F, Omasa K. 2010. Estimation of coniferous standing tree volume using airborne LiDAR and passive optical remote sensing. *Journal of Agricultural Meteorology*, **66**: 111−116.

Nelson R, Swill R, Krabill W. 1988. Using airborne lasers to estimate forest canopy and stand characteristics. *Journal of Forestry*, **86**: 31−38.

Neter J, Kutner MH, Wasserman W, Nachtsheim CJ. 1996. *Applied linear statistical models*. McGraw-Hill/Irwin, 4th edition. Irwin Series in Statistics, 770 pp.

Nilsson M. 1996. Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment*, **56**: 1−7.

Popescu SC, Wynne RH, Nelson RF. 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing*, **29**: 564−577.

Popescu SC, Wynne RH. 2004. Seeing the trees in the forest: using LIDAR and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photogrammetric Engineering and Remote Sensing*, **70**: 589−604.

Popescu SC. 2007. Estimating biomass of individual pine trees using airborne LiDAR. *Biomass and Bioenergy*, **31**: 646−655

Putz FE, Zuidema PA, Pinard MA, Boot RGA, Sayer JA, Sheil D, Sist P, Elias, Vanclay JK. 2008. Improved tropical forest management for carbon retention. *PLOS Biology*, **6**: 1368−1369

Roberts SD, Dean TJ, Evans DL, McCombs JW, Harrington RL, Glass PA. 2005. Estimating individual tree leaf area in loblolly pine plantations using LiDAR-derived measurements of height and crown dimensions. *Forest Ecology and Management*, **213**: 54−70.

Sawadogo L, Savadogo P, Tiveau D, Dayamba SD, Zida D, Nouvellet Y, Oden PC, Guinko S. 2010. Allometric prediction of above-ground biomass of eleven woody tree species in Sudanian savanna-woodland of West Africa. *Journal of Forestry Research*, **21**: 475−481.

Satoo T, Madwick HAI. 1982. *Forest Biomass*. The Hague / Boston / London: Martinus Ni1hoff / **Dr** W. Junk Publishers, p.152.

Snowdon P. 1991. A ratio estimator for bias correction in logarithmic regressions. *Canadian Journal of Forest Research*, **21**: 720−724.

Tallis H, Polasky S. 2009. Mapping and valuing ecosystem services as an approach for conservation and natural-resource management. *Annals of the New York Academy of Sciences*, **1162**: 265−283.

Thompson ID, Okabe K, Tylianakis JM, Kumar P, Brockerhoff EG, Schellhorn NA, Parrotta JA, Nasi R. 2011. Forest biodiversity and the delivery of ecosystem goods and services: translating science into policy. *Bioscience,* **61**: 972−981.

Ulander LMH, Patrik BG, Hagberg JO. 1995. Measuring tree height using ERS-1 SAR Interferometry. *In: Proceeding of Geoscience and Remote Sensing Symposium* IGARSS '95, Florence, Italy, pp. 2189−2191.