Forestry *An International Journal of Forest Research*

Institute of Chartered Foresters

Forestry 2020; **93**, 444–457, doi:10.1093/forestry/cpz073
Advance Access publication 26 December 2019

A novel application of small area estimation in loblolly pine forest inventory

P. Corey Green[1,*], Harold E. Burkhart[1], John W. Coulston[2] and Philip J. Radtke[1]

[1]*Department of Forest Resources and Environmental Conservation, Virginia Polytechnic Institute and State University, 310 W. Campus Drive, Blacksburg, VA 24061, USA*
[2]*United States Forest Service, 1710 Research Center Drive, Blacksburg, VA 24060, USA*

*Corresponding author Phone: +1 (540) 231-7263; E-mail: pcgreen7@vt.edu.

Received 12 September 2019

Loblolly pine (*Pinus taeda* L.) is one of the most widely planted tree species globally. As the reliability of estimating forest characteristics such as volume, biomass and carbon becomes more important, the necessary resources available for assessment are often insufficient to meet desired confidence levels. Small area estimation (SAE) methods were investigated for their potential to improve the precision of volume estimates in loblolly pine plantations aged 9–43. Area-level SAE models that included lidar height percentiles and stand thinning status as auxiliary information were developed to test whether precision gains could be achieved. Models that utilized both forms of auxiliary data provided larger gains in precision compared to using lidar alone. Unit-level SAE models were found to offer additional gains compared with area-level models in some cases; however, area-level models that incorporated both lidar and thinning status performed nearly as well or better. Despite their potential gains in precision, unit-level models are more difficult to apply in practice due to the need for highly accurate, spatially defined sample units and the inability to incorporate certain area-level covariates. The results of this study are of interest to those looking to reduce the uncertainty of stand parameter estimates. With improved estimate precision, managers, stakeholders and policy makers can have more confidence in resource assessments for informed decisions.

Keywords: Small area estimation, Forest inventory, Auxiliary data, Lidar, Loblolly pine.

Introduction

In the southern United States, managed pine plantations occupy ∼14 million hectares of forest land (Zhao *et al.*, 2016). Of the major pine species, loblolly pine (*Pinus taeda* L.) is the most widely planted and has seen large improvements in productivity beginning in the 1950s (Fox *et al.*, 2007). Plantation loblolly pine is important not only as a productive timber species but also for multiple use forest management purposes (Schultz, 1999). The extensive planting and intensive management of loblolly pine have significant implications for wildlife management (Andreu *et al.*, 2008) and the global carbon cycle (Johnsen *et al.*, 2001). In recent years, ownership patterns have seen changes resulting in frequent acquisitions and dispositions of timberlands (Fox *et al.*, 2007). Due to commercial and ecological implications, plantation forestry requires accurate, reliable and expeditious information at the stand-level for informed management decisions. Forest inventory is a primary tool used to obtain estimates of stand parameters. Ground-based forest inventory is typically conducted using either fixed- or variable-radius plots as sample units established with a specific sampling intensity and spatial arrangement with the goal of achieving a certain targeted precision. Common sample measurements include species, diameter at breast height (d.b.h.), total height: (Ht) and stem quality assessments (Burkhart *et al.*, 2019). In practice, it is common for design-based estimates (i.e. estimates derived only from the ground-based sample units) to lack the precision required for management purposes. This is often due to logistical and budgetary constraints which limit the planned sample intensity. In these cases, a class of model-based statistical estimation techniques known as small area estimation (SAE) is an option that can be used to reduce the uncertainty of inventory estimates. SAE models can be broadly classified as ‘area-level’ and ‘unit-level’. Area-level models relate area-based direct estimates to area-level covariates, while unit-level models relate sample unit values to the corresponding sample unit covariates (Rao and Molina, 2015). For many areas where loblolly pine is grown commercially, auxiliary data are available that can be leveraged for use in SAE models.

Light detection and ranging

Light detection and ranging, referred to as lidar, is a form of active remote sensing that includes a scanning laser, an inertial measurement unit, a global positioning system (GPS) which is a form of a global navigation satellite system and a computer containing timing systems and storage (Campbell and Wynne, 2011). Generally, lidar systems are installed on fixed-wing aircraft or helicopters. In aerial lidar analysis, a 'point cloud' contains three-dimensional data that include *x-y-z* coordinates representing both horizontal and vertical structure referenced above the Earth's ellipsoid. In many applications, the point cloud is normalized by subtracting the ground elevation from the heights above the ellipsoid. Starting in the 1980s with work by Nelson *et al.* (1984) and Nelson *et al.* (1988), lidar has been extensively demonstrated to provide useful information in forestry applications. When utilized for forest inventory purposes, two general approaches have seen both research and application: (1) area-based approaches and (2) individual-tree detection approaches. Both methods have been demonstrated to be effective and are recommended for use in plantation forest inventory applications (Maltamo *et al.,* 2014).

Area-based approaches have been successfully applied to estimating mean dominant tree heights in a variety of forest conditions (e.g. Næsset, 1997a, 2004a, b, 2007; Means *et al.,* 2000). In addition to dominant height, the area-based approach has been used to predict total stand volume and biomass through regression approaches with lidar-derived height and canopy cover metrics as predictors (e.g. Næsset, 1997b, 2002, 2004a, b, 2007; Holmgren, 2004). Biomass and volume were estimated by van Aardt *et al.* (2006) through an object-oriented approach. Area-based methods have also been utilized to parameterize models that predict stem density, average diameter and basal area (Means *et al.,* 2000; Næsset, 2002, 2004a, b, 2007; Holmgren, 2004).

All the studies mentioned up to this point have focused on estimating forest characteristics through area-based methods. These techniques do not rely on the detection and delineation of individual tree stems or crowns; rather, they rely on quantiles and distributional features of lidar data as predictors for a given area (Yu *et al.,* 2010). The individual tree detection methods rely on algorithms to locate and measure trees in a point cloud. While they typically require higher pulse densities and greater computational resources, individual tree detection requires less data from field measurements to calibrate (Yu *et al.,* 2010). Successful examples of predicting stand characteristics using the individual tree approach include McCombs *et al.* (2003), Popescu and Wynne (2003), Popescu *et al.* (2003) and Yu *et al.* (2010).

Area-level SAE

The area-level SAE approach was first proposed by Fay III and Herriot (1979) for use in predicting income in low population areas. Using U.S. Census data and other auxiliary information, they were able to improve the estimation precision with their proposed composite estimators. The area-level approach has seen multiple forest research applications. Goerndt *et al.* (2011) used lidar-derived auxiliary information to improve the precision of stand-level estimates of density, quadratic mean diameter, total height and total volume in a variety of cover types in coastal Oregon. Several area-level SAE composite estimators were found to provide comparable gains in precision with the aforementioned stand variables. In addition, Magnussen *et al.* (2017) presented three case studies using lidar and one case study using photogrammetrically derived point clouds as ancillary data in area-level analysis. In a variety of European locations including Spain, Germany, Switzerland and Norway, both forms of auxiliary data provided increased precision of total volume estimates.

Unit-level SAE

The unit-level approach was first introduced for the prediction of crop area in selected Iowa counties. Using Landsat imagery as auxiliary information, the standard errors for the estimates were reduced compared with the direct estimate alone (Battese *et al.,* 1988). In addition, unit-level SAE models have seen application in the forestry literature. Weighted unit-level SAE models were used to estimate the total area occupied by olive trees in Navarra, Spain, using Landsat imagery as auxiliary information (Militino *et al.,* 2006). Using Norwegian National Forest Inventory data, Breidenbach and Astrup (2012) used a photogrammetrically derived point cloud as auxiliary information and found significant improvements in the precision of above-ground forest biomass estimates. Goerndt *et al.* (2013) utilized a variety of auxiliary data including Landsat variables, land cover class, tree cover and elevation to successfully improve the precision of a variety of county-level forest attributes in the Oregon Coast Range.

Unit *vs* area comparisons

Several research studies have compared the effectiveness of the two SAE methods. Mauro *et al.* (2017) investigated area- and unit-level models in predominately coastal coniferous forests located in the Oregon Coast Range. Using auxiliary lidar data, unit-level approaches were found to produce more precise estimates compared with area-level and design-based approaches for all stand variables of interest. Area-level models, however, were found to produce more precise estimates when compared to the direct estimates. In addition to lidar, a comparison between area- and unit-level estimators using photogrammetrically derived point clouds found greater stand-level estimate precision with unit-level models compared with area-level models in most cases (Breidenbach *et al.,* 2018).

Research objectives and questions

To our knowledge, no studies have focused on using SAE techniques specifically in intensively managed loblolly pine plantations. The overall objective of this work was to demonstrate the potential for using SAE techniques to improve the precision of stand-level estimates of total planted volume in operationally managed loblolly pine plantations across a range of common inventory entry points. Specific objectives/questions include:

1. Do lidar-derived auxiliary data improve total planted volume estimate precision with area-level SAE analysis?
2. Do lidar-derived auxiliary data improve total planted volume estimate precision with unit-level SAE analysis?
3. How do area- and unit-level SAE approaches compare in plantation pine forest inventory?
4. What other sources of auxiliary data improve estimates with area- and/or unit-level SAE analysis?

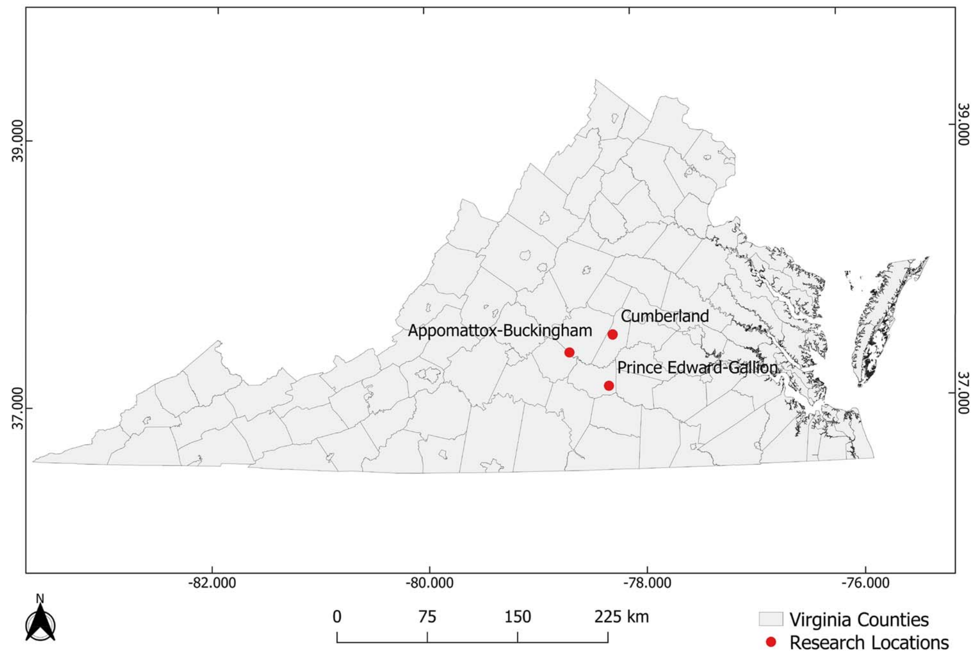

Figure 1 Locations of state forests used in this study.

Data and methods

Study location and ground data

Locations

State forests in the Commonwealth of Virginia are under the administration of the Virginia Department of Forestry (VDOF) and are managed with multiple use objectives including timber management. Three state forests, Appomattox-Buckingham (ABSF), Cumberland (CUSF) and Prince Edward-Gallion (PESF), were selected for this study (Figure 1).

These forests are located in the central portion of the state and are representative of a variety of Piedmont physiographic conditions. ABSF is representative of the upper Piedmont with higher elevations than the other sites evaluated and is on the edge of the natural range of loblolly pine. CUSF exhibited great topographic diversity with many areas that resembled upper Coastal Plain topography. PESF was generally representative of common Piedmont conditions. Managed loblolly pine plantations are prevalent throughout this area, and the VDOF generally follows management strategies common throughout the Piedmont region of the Southeast (e.g. site preparation, planting densities, planting stock and thinning type).

Forty stands were selected within the state forests to cover the range of forest inventory entry points typical of managed pine in the Southeast. Specifically, stages of development that commonly require inventory information for planning decisions include: canopy closure, immediate pre-first thinning, immediate post thinning and pre-final harvest. Both unthinned and thinned stands established from 1976 to 2010 were considered for study (Figure 2). Stands selected varied in initial planting density, genetic origin and silviculture treatments. Apart from three stands nearing rotation age, all thinned stands were thinned within the past 2 years prior to ground measurements in the winter and early spring of 2019.

Sample design and ground data

Stands selected for inventory in the winter and early spring of 2019 were allocated sample units using QGIS Development Team (2019). Approximately 1 sample plot per 1.2 ha were arranged at random with a minimum distance of ~70 to 80 m between plot centers. This was assumed to produce an equal probability sample. Due to time constraints, the full sample intensity was measured for only 22 stands of 40. In stands that did not receive the full sample intensity, measurement plots were randomly chosen while still maintaining spatial coverage. Unthinned stands were inventoried with 0.013 ha fixed-radius plots. In cases where excessive natural regeneration was present, 0.01 ha fixed-radius plots were used instead. Thinned stands were all inventoried with 0.02 ha fixed-radius plots. Plot center locations were established with a Trimble Geo7x GPS capable of submetre accuracy. A minimum of 50 GPS points were collected at each plot center and were differentially post processed based on the nearest continuously referenced base stations. A total of 267 plot locations were measured and 260 locations were collected using the Trimble Geo7x. Missing plot centers were due to a combination of missed field collections and GPS data post-processing limitations.

On each sample unit, all living planted stems were recorded and measured for d.b.h. Only living natural stems d.b.h. $\geq$ 7.62 cm were recorded and measured for d.b.h.. A subset of planted trees was measured for total height. With few exceptions, a minimum of 25 planted tree heights across the diameter distribution were measured in each stand with at least one height measured per plot (except for plots with no trees). For natural loblolly pine, Virginia pine (*Pinus virginiana* Mill.) and shortleaf pine (*Pinus echinata* Mill.) a subset of heights was measured across the diameter distribution. Planted pine heights not measured were predicted with heights measured at the stand level using equation (1). Natural pine heights measured were pooled region-wide and used to fit

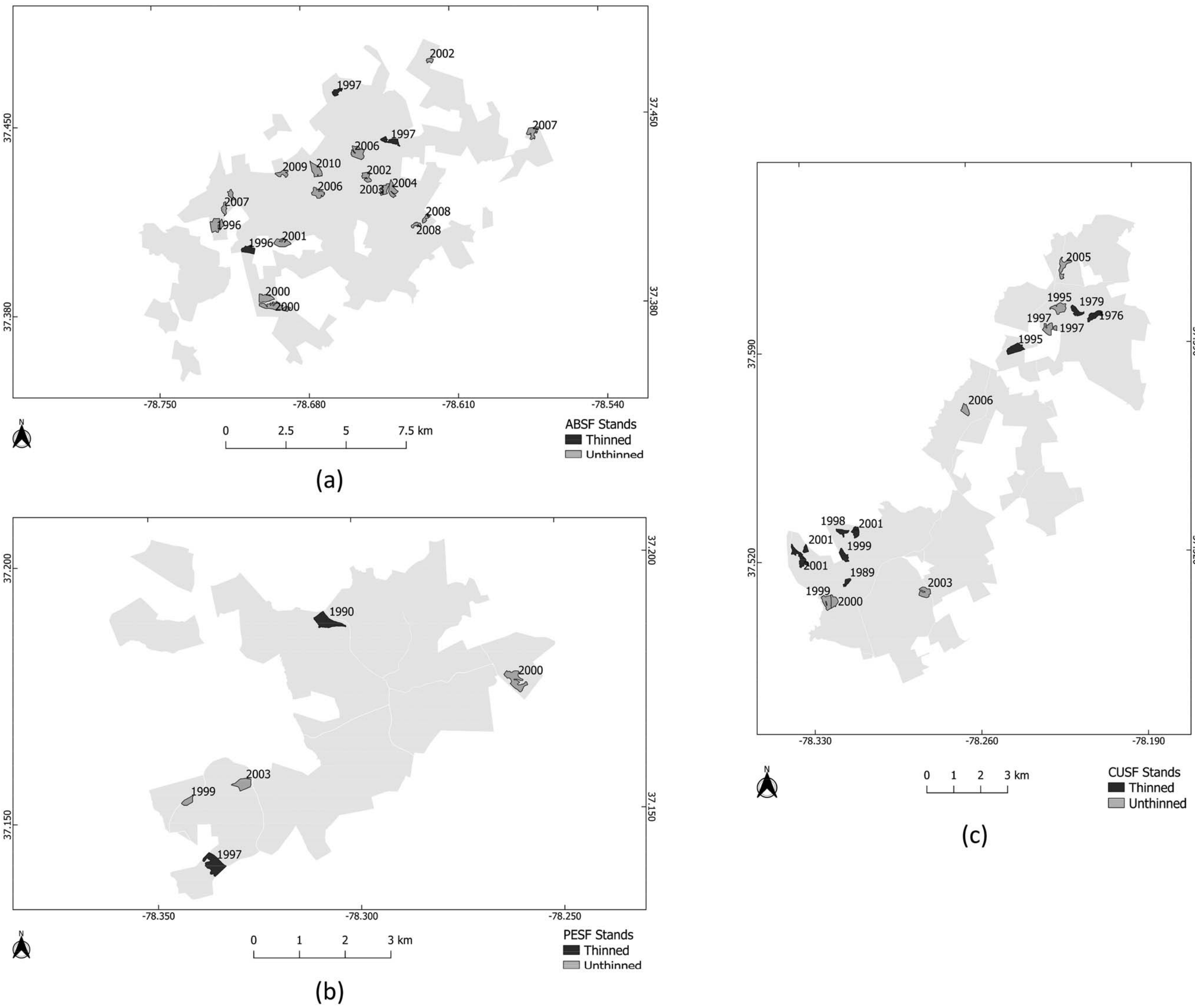

Figure 2 Stands at Appomattox-Buckingham (a), Prince Edward-Gallion (b) and Cumberland (c) state forests inventoried for study. Each stand is labeled with the year it was established (all ground measurements taken in the winter/early spring of 2019). The fill of the polygons indicates thinning status.

the regression model for predicting unmeasured heights, also using the model form in equation (1).

$$\ln(Ht) = b_0 + b_1(\mathrm{DBH}^{-1}) \tag{1}$$

In addition to all coniferous species, hardwood trees (d.b.h. ≥ 7.62 cm) were measured for d.b.h. and a subset of heights was measured. Following definitions of dominant height given by Gyawali and Burkhart (2015), the top 80 per cent of planted pine heights per plot were used in place of mean dominant height. For all trees, volumes were estimated using allometric equations presented in Table 1.

All field data were processed using R (R Core Team, 2018). Additional R packages used for graphics and data processing include the following: ggplot2 (Wickham, 2016), xlsx (Dragulescu and Arendt, 2018) and reshape2 (Wickham, 2007).

Auxiliary information

For the entire study area, 1-m digital elevation models (DEMs) and the associated lidar point clouds used to generate them were obtained from publicly available data maintained by the United States Geological Survey (USGS). ABSF and CUSF were part of the 2015 'Chesapeake Bay' lidar campaign, while PESF was part of the 2014 'Sandy' campaign. Details of these two lidar collections are found in Table 2. The DEMs and the associated lidar point clouds are available from the USGS National Map (USGS, 2017a, b).

Additionally, stand thinning status was used as auxiliary information for both area- and unit-level analysis. Stands and plots that had received at least one thinning treatment (not including pre-commercial thinning) were classified as thinned. No distinction was made between plots and stands that had one thinning treatment and those that had received multiple. Thinning status was obtained from stand attribute information and confirmed during the field inventory.

Table 1 Sources of allometric equations used to estimate total stem volume.

Species/species group	Source
Planted loblolly pine	Tasissa *et al.* (1997) (unthinned coefficients)
Natural loblolly, Virginia, and shortleaf > = 12.7 cm d.b.h.	Tasissa *et al.* (1997) (unthinned coefficients)
Natural loblolly, Virginia, and shortleaf <12.7 cm d.b.h.	Warner and Goebel, 1963
Hardwoods with no measured total height	Clark III *et al.* (1986) (table 10 coefficients)
Hardwoods with measured total height	Clark III *et al.* (1986) (table 14 coefficients)

Table 2 Lidar specifications for the Chesapeake Bay and Sandy projects.

	Chesapeake Bay	Sandy
Collection dates	15 November 2015–30 March 2016	24 March 2014–21 April 2014
Sensor	Riegl 680i	Leica ALS60 or Leica ALS70
Scan angle (degrees)	60	unreported
Point density (pulses m^2)	2.3	unreported
Nominal pulse spacing (m)	0.66	0.7
Flight line overlap	55%	30% (ALS60) or 20% (ALS70)
Pulse rate (kHz)	200	154.3 (ALS60) or 301.6 (ALS70)

Processing auxiliary information

USGS delivers elevation data in tiles that were smaller than the areas of interest for this study. Stand boundaries often overlapped into multiple tiles. To facilitate and simplify analysis, the individual DEM tiles were merged and mosaicked using QGIS to produce a single 1-m DEM for each state forest. The lasmerge tool in the LAStools suite (LAStools, 2018) was used to combine the individual lidar tiles into a single lidar dataset for each state forest. The merged lidar dataset was clipped to each area of interest (stands for area-level analysis and plots for unit-level analysis) using lasclip in the LAStools suite. The DEMs were converted from ascii to DTM format using the ASCII2DTM tool in FUSION needed for further analysis with other FUSION tools (McGaughey, 2018). For all lidar analyses, an area-based approach was utilized due to its relatively low computational demands and ability to accurately estimate stand parameters of interest. Lidar metrics were generated using the FUSION GridMetrics and Cloudmetrics tools. For area-level analysis, each stand was tessellated into a regular grid where the grid size was set to approximately the same size as the sample units used for the specific stand. Unit-level analysis used the Clipdata tool to normalize the lidar point clouds followed by Cloudmetrics to generate plot-level metrics. Lidar metrics were summarized using R. Returns ≥ 30.5 m above ground were excluded as all measured heights were lower than this value. Following this subset, any heights > third quantile +1.5∗(Interquartile range) were removed in the gridded metrics for area-level analysis. The subsets were to account for overlapping, large canopies from remnant trees and adjacent stands that may have caused an overestimation of height percentiles. Filtered grid cell percentiles were averaged to produce stand-level lidar attributes (e.g. a stand level 80th percentile lidar height). A variety of R packages were used for geospatial processing tasks such as reprojecting and subsetting spatial layers throughout the workflow. These packages include the raster package (Hijmans, 2019), the sp package (Pebesma and Bivand, 2005; Bivand *et al.*, 2013) and the rgdal package (Bivand *et al.*, 2019).

Direct estimators

Under the assumption of an equal probability simple random sample in stand i for the parameter of interest θ_i, the direct estimate of the mean is

$$\hat{\theta}_i = \bar{y}_i = n_i^{-1} \sum_{j=1}^{n_i} y_{ij} \quad (2)$$

and the variance of the estimate is

$$\hat{\Psi}_i = \mathrm{Var}\left(\hat{\theta}_i\right) = n_i^{-1} \frac{\sum \left(y_{ij} - \bar{y}_i\right)^2}{n_i - 1} \quad (3)$$

where y_{ij} is sample plot j in stand i and n_i is the sample size for stand i.

Small area estimators

Area-level

Given a properly designed sample in area i, a direct estimator for parameter of interest θ_i is available (equation 4)[1]:

$$\theta_i + e_i \quad (4)$$

where the individual random errors e_i in equation (4) are iid $N(0, \Psi_i)$. In many cases, however, an insufficient sample intensity leads to a direct estimate that is not reliable enough for a given

1 The proceeding descriptions of area-level small area estimators incorporate a combination of notation used in Goerndt *et al.* (2011) and Rao and Molina (2015).

management objective (i.e. an inflated variance Ψ_i leading to an unacceptably wide confidence interval). In cases where auxiliary information is available for area $\boldsymbol{i}$, we assume the parameter of interest θ_i can be linearly related to a set of $\boldsymbol{m}$ auxiliary covariates $\boldsymbol{z}$ through equation (5):

$$\theta_i = \boldsymbol{z}_i^T \boldsymbol{\beta} + b_i v_i \tag{5}$$

where $\boldsymbol{z}_i$ is a vector of area-specific covariates, $\boldsymbol{\beta}$ is the vector of regression coefficients, $\boldsymbol{b}_i$ are positive constants (assumed to equal 1) and area-specific random effects $\boldsymbol{v}_i$that are iid $N\left(0, \sigma_v^2\right)$. Combining equations (4) and (5) leads to the following mixed model:

$$\hat{\theta}_i = \boldsymbol{z}_i^T \boldsymbol{\beta} + b_i v_i + e_i \tag{6}$$

where all terms are as previously described in equations (4) and (5).

Prior to estimation, sampling error variance (Ψ_i) and random error variance (σ_v^2) must be estimated and partitioned. In this study, variances calculated directly from the sample units were used as an estimate for Ψ_i. The estimated sample variance ($\hat{\Psi}_i$) was utilized to estimate $\hat{\sigma_v^2}$ using the restricted maximum likelihood (REML) method as implemented in the R SAE package (Molina and Marhuenda, 2015). The empirical best linear unbiased predictor (EBLUP) was then obtained with the composite estimator in equation (7):

$$\hat{\theta}_i^H = \hat{\gamma}_i \hat{\theta}_i + (1 - \hat{\gamma}_i)\, \boldsymbol{z}_i^T \hat{\beta} \tag{7}$$

where $\hat{\gamma}$ is a weight using both sources of error accounted for previously and is given in equation (8).

$$\hat{\gamma}_i = \hat{\sigma_v^2} / \left(\hat{\sigma_v^2} + \hat{\Psi}_i\right) \tag{8}$$

Final EBLUP and Mean squared error (MSE) estimates were obtained through the 'mseFH' function in the R SAE package. The default REML method was used, as were the default maximum iterations (100) and the default precision (0.0001). The MSE estimation involves three components:

$$g_{i1}\left(\hat{\sigma_v^2}\right) = \hat{\gamma}_i \hat{\Psi}_i \tag{9}$$

$$g_{i2}\left(\hat{\sigma_v^2}\right) = (1 - \hat{\gamma}_i)^2 \boldsymbol{z}_i^T \left(\left(\frac{\hat{\boldsymbol{\gamma}}}{\hat{\sigma_v^2}} \boldsymbol{Z}\right)^T \boldsymbol{Z}\right)^{-1} \boldsymbol{z}_i \tag{10}$$

$$g_{i3}\left(\hat{\sigma_v^2}\right) = (1 - \hat{\gamma}_i)^2 \frac{2}{\sum \left(\frac{\hat{\gamma}}{\hat{\sigma_v^2}}\right)^2} \left(\hat{\sigma_v^2}\, \hat{\Psi}_i\right)^{-1} \tag{11}$$

where $\boldsymbol{Z}$ is an $I\ x\ m$ matrix of $\boldsymbol{z}_i^T$for each domain and $\hat{\boldsymbol{\gamma}}$ is an I-*dimensional* vector of $\hat{\gamma}_i$ for each domain. The combination of the components leads to the final MSE estimate:

$$MSE\left(\hat{\theta_i^H}\right) = g_{i1}\left(\hat{\sigma_v^2}\right) + g_{i2}\left(\hat{\sigma_v^2}\right) + 2g_{i3}\left(\hat{\sigma_v^2}\right) \tag{12}$$

The previous notation was reconstructed from the R SAE package code (Molina and Marhuenda, 2015). For details of REML fitting to obtain EBLUP estimates, readers are directed to Rao and Molina (2015); Datta and Lahiri (2000) provide additional details of the MSE estimation.

Unit-level

When data are available at the unit-level (individual sample plots in this study), unit-level SAE methods can be utilized.[2] The nested error unit-level model specifies the observed attribute *Y*on plot $\boldsymbol{j}$ in area $\boldsymbol{i}$ as:

$$Y_{ij} = \boldsymbol{x}_{ij}^T \boldsymbol{\beta} + u_i + e_{ij} \tag{13}$$

where $\boldsymbol{u}_i$ are area effects and are iid $N\left(0, \sigma_u^2\right)$ and $\boldsymbol{e}_{ij}$ are individual errors iid $N\left(0, \sigma_e^2\right)$.

The EBLUP for a particular small area $\boldsymbol{i}$ is expressed in equation (14):

$$\hat{\bar{Y}}_i^{EBLUP} = f_i \bar{y}_{is} + \left(\bar{\boldsymbol{X}}_i - f_i \bar{\boldsymbol{x}}_{is}\right)^T \hat{\boldsymbol{\beta}} + (1 - f_i)\, \hat{u}_i \tag{14}$$

where

$$f_i = {}^{n_i}\!/_{N_i} \tag{15}$$

where $\boldsymbol{N}_i$ is the total number of units for stand $\boldsymbol{i}$. This was obtained by dividing the area of the stand by the size of the sample unit (plot size).

$$\bar{y}_{is} = n_i^{-1} \sum\nolimits_{j \in s_i} Y_{ij} \tag{16}$$

where $\boldsymbol{j} \in \boldsymbol{s}_i$ indicates units j contained in stand i.

$$\bar{\boldsymbol{x}}_{is} = n_i^{-1} \sum\nolimits_{j \in s_i} \boldsymbol{x}_{ij} \tag{17}$$

$$\hat{u}_i = \hat{\gamma}_i \left(\bar{y}_{is} - \bar{\boldsymbol{x}}_{is}^T \hat{\boldsymbol{\beta}}\right) \tag{18}$$

$$\hat{\gamma}_i = \hat{\sigma}_u^2 \Big/ \left(\hat{\sigma}_u^2 + \frac{\hat{\sigma}_e^2}{n_i}\right) \tag{19}$$

and $\bar{\boldsymbol{X}}_i$ are true population totals for the auxiliary data. The true lidar height percentiles were obtained from grid-level point cloud summaries.

Final unit-level EBLUP estimates and the associated MSE values were obtained with the 'pbmseBHF' function in the R SAE package. This function uses a REML procedure for fitting and a parametric bootstrap approach for estimating MSE values. The default value of 200 bootstrap samples was utilized. The parametric bootstrap method was utilized to relax some of the restrictions encountered with analytical error estimation methods. For details regarding EBLUP fitting using the REML method, readers are directed to Rao and Molina (2015); González-Manteiga *et al.* (2008) outline further details regarding the MSE estimation using the bootstrap procedure.

Results

Stand-level summaries of trees per hectare, basal area per hectare, total volume per hectare and dominant height are presented in Figure 3. Density, basal area and volume are partitioned into planted, natural and total (the sum of planted and natural). Dominant height was not considered for natural

2 *The proceeding descriptions of unit-level small area estimators incorporate a combination of notation used in* Molina and Marhuenda (2015) *and* Rao and Molina (2015).

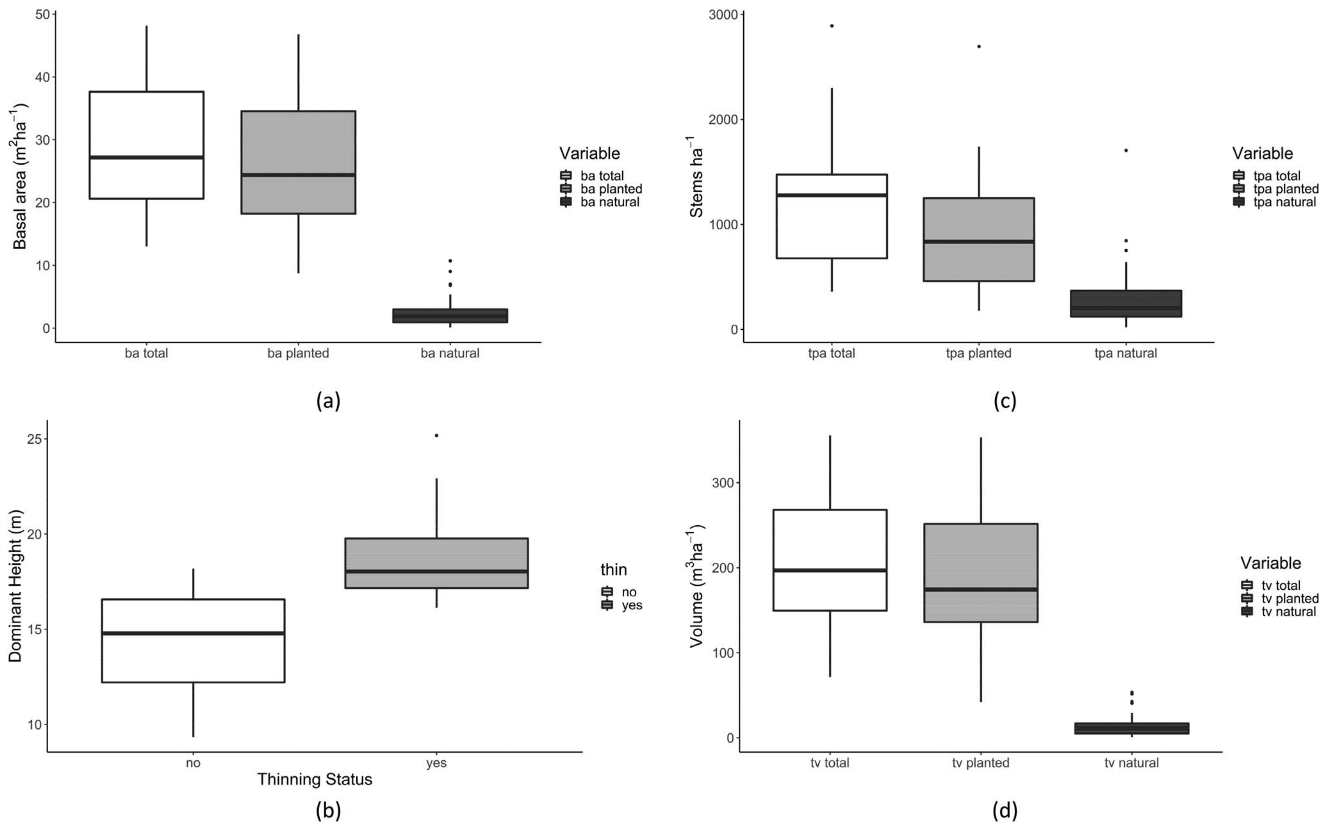

Figure 3 Boxplots (Tukey, 1977) of stand variables for basal area per hectare (a), dominant height (b), live stems per hectare (c) and volume per hectare (d). Note: Dominant height was not considered for natural trees.

trees; thus, dominant height is presented by stand thinning status. A goal of this work was to cover the full range of common entry points in loblolly pine plantation forest inventory. The distributions of stand variables capture a wide range of commonly encountered conditions from canopy closure to final harvest. A common condition encountered was significant natural Virginia pine regeneration which made up a large amount of the natural component as seen in Figure 3. This is common throughout areas on the periphery of the loblolly pine range including the Piedmont of Virginia. A noted limitation in coverage was the small number of stands that received first thinning in the 'typical' window of 12–19 years common in managed pine plantations (Figure 2).

All SAE methods evaluated assume a linear relationship between the variable of interest and the auxiliary information. The relationship between the auxiliary information (lidar-derived 80th percentile and thinning status) and total planted volume is confirmed in Figure 4. The remaining lidar height metrics (90th, 95th and 99th percentiles) had very similar relationships; thus, the figures are not presented.

For the following results and discussion, relative error ratios (RER) were constructed for comparisons between the precision of the model EBLUP and direct estimates. For SAE models, the RER for small area ***i*** is

$$\mathrm{RER}_i\ (\%) = \frac{\sqrt{\mathrm{MSE}\,(\mathrm{EBLUP}_i)}}{\mathrm{EBLUP}_i} * 100 \qquad (20)$$

and the RER for the direct estimate is

$$\mathrm{RER}_i\ (\%) = \frac{\sqrt{\hat{\Psi}_i}}{\hat{\theta}_i} * 100 \qquad (21)$$

where $\hat{\theta}_i$ and $\hat{\Psi}_i$ are as defined in equations (2) and (3), respectively.

A RER is similar to a coefficient of variation in that it standardizes the variation of the estimate to the estimate itself. The RERs were visually compared for each stand using 1–1 scatterplots. In the following figures, any point falling below the 1–1 line indicated a smaller model RER for the estimate type on the *y*-axis.

Area-level SAE

Area-level SAE models that utilized only a lidar height percentile resulted in small gains in precision for some stands (Figure 5a) The incorporation of both a lidar height percentile and thinning status resulted in larger gains in precision (Figure 5b). Modeled total volume estimates generally followed the 1–1 relationship with the direct estimates for both model forms (Figure 5c,d). Additional lidar height percentiles were evaluated (90th, 95th and 99th) resulting in similar goodness of fit statistics (Table 3). The Akaike information criterion (AIC) and Bayesian information criterion (BIC), two forms of penalized likelihood criteria, were

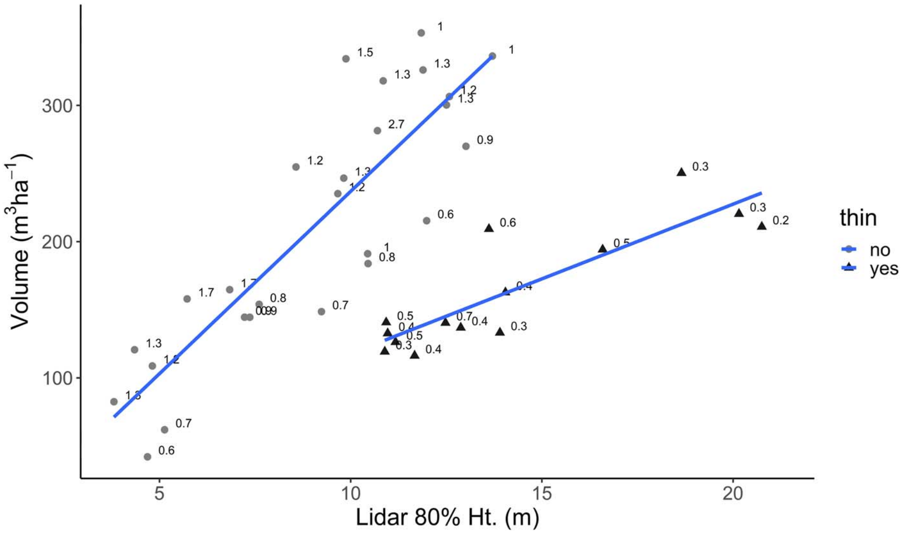

Figure 4 Linear relationship between auxiliary information (80th height percentile from lidar), with dependent variable of interest (total cubic volume per hectare) by thinning status (light gray round points are unthinned and dark gray triangular points are thinned). Points are labeled with direct estimate of surviving planted stem density in thousands of trees per hectare at the time of inventory. Linear fit by thinning status indicates the importance of including thinning status as explanatory variable.

Table 3 Goodness of fit summaries for area-level models considered for estimating total volume.

Model fit	AIC	BIC
TV_Planted ~ P80	675.80	680.87
TV_Planted ~ P90	675.91	680.97
TV_Planted ~ P95	675.86	680.93
TV_Planted ~ P99	675.46	680.53
TV_Planted ~ P80 + thin_status	640.08	646.83
TV_Planted ~ P90 + thin_status	640.91	647.66
TV_Planted ~ P95 + thin_status	641.40	648.16
TV_Planted ~ P99 + thin_status	642.44	649.20

calculated for model performance. Models that used the lidar 80th height percentile exhibited the lowest AIC and BIC values. For consistency, models that use the 80th percentile were used for all figures and further discussion.

Unit-level SAE

Unit-level SAE models were evaluated using lidar height percentiles alone and with thinning status included. The unit-level approach resulted in large increases in precision for some of the stands evaluated; however, a decrease in precision of estimates was observed for some stands, especially those with low variability in the direct estimate (Figure 6a). The unit-level models tended to produce estimates that were larger than the associated direct estimates (Figure 6c,d). In addition to the lidar 80th height percentile, the 90th, 95th, and 99th percentiles were evaluated (Table 4). Models that used the lidar 80th height percentile exhibited the lowest AIC and BIC values. For consistency, models that use the 80th percentile was used for all figures and further discussion.

Table 4 Goodness of fit summaries for unit-level models considered for estimating total volume.

Model fit	AIC	BIC
TV_Planted ~ P80	4219.43	4233.67
TV_Planted ~ P90	4237.72	4251.96
TV_Planted ~ P95	4243.65	4257.90
TV_Planted ~ P99	4254.06	4268.31
TV_Planted ~ P80 + thin_status	4140.88	4158.68
TV_Planted ~ P90 + thin_status	4164.39	4182.20
TV_Planted ~ P95 + thin_status	4172.33	4190.14
TV_Planted ~ P99 + thin_status	4186.52	4204.33

Values not calculated directly in the SAE R function. Values calculated from the definitions of AIC and BIC and the log likelihood value calculated with the SAE function.

Comparison

Area- and unit-level models are visually compared in Figure 7. A comparison was first made between the area- and unit-level models that utilized only the same 80th percentile height as the auxiliary information. As seen in Figure 7a, the unit-level model can potentially result in larger gains in precision compared to the area-level model, particularly at some higher levels of variation. The inclusion of thinning status resulted in similar relationships (Figure 7b). Despite these potential improvements, the mean RERs were very similar or higher for all model forms evaluated (Table 5). Area-level models more often and, on average, resulted in larger decreases in uncertainty compared with the unit-level models evaluated.

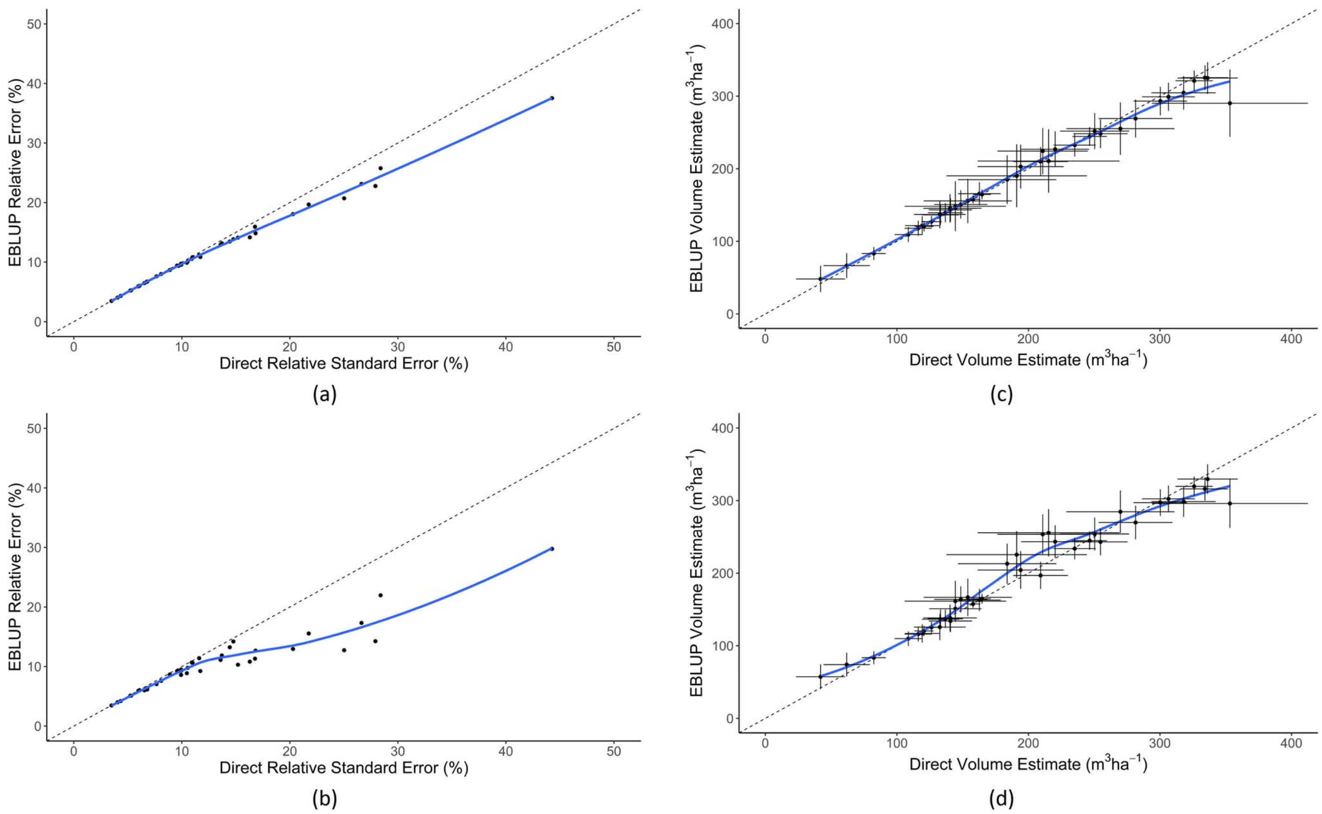

Figure 5 Area-level SAE results. Model with lidar 80th height percentile as only auxiliary information relative error comparison (a) model with lidar 80th height percentile and thinning status as auxiliary information relative error comparison (b). Estimate comparison for model with lidar 80th height percentile as only auxiliary information (c) and estimate comparison for model with lidar 80th height percentile and thinning status as auxiliary information (d). Smoothing lines are for visual interpretation only and are not representative of the SAE model fit. Error bars represent one standard error in the *x* direction and the root mean squared error in the y direction.

Discussion

The results of this research have confirmed the potential for reducing volume estimate uncertainty using SAE techniques. Both area and unit-level approaches reduced the relative error of the estimate for small areas (stands) compared to less precise direct estimates. Much like Goerndt *et al.* (2011), Magnussen *et al.* (2017) and Mauro *et al.* (2017), lidar was shown to be an effective source of auxiliary information that could be leveraged for use in SAE models. An interesting aspect in this work is that despite the lidar data having been acquired 4–5 years prior to collection of field inventory data, the linear relationship between total planted volume and the auxiliary data was still strong. This follows with McRoberts *et al.* (2018) in which lidar was found to remain useful in model-assisted estimation many years after collection. If more recent lidar collections were available, additional gains in precision could possibly be achieved, assuming a stronger linear relationship would be found with close temporal matching of lidar and field data collection. In cases with significant disturbances such as windthrow or stress-induced mortality, the relationship between the auxiliary data and the stand conditions may be less useful. In these situations, a separate inventory or lidar collection may be necessary.

Due to their ability to use finer-scale data, unit-level SAE models can offer greater gains in precision compared with area-level models (Molina and Marhuenda, 2015); however, they require precise co-registration of ground sample plots and the associated lidar. High accuracy GPS units are not always readily available, limiting the potential for unit-level estimation in some situations. Variable radius plots pose additional challenges to unit-level methods due to the sample unit's lack of a defined spatial area. Further, some auxiliary information cannot always be summarized at the unit-level. Thinning status is often a stand-level attribute and cannot easily be incorporated into the unit-level framework unless recorded at the plot level as it was in this study. Further, in some cases, thinning status may not be obvious at each plot location. Despite these limitations, the unit-level models did provide improvements in precision in this work when compared with area-level analysis for some stands, especially those that exhibited large direct estimate variances. The incorporation of both lidar height percentiles and thinning status did offer improvements in precision to both the area- and unit-level analysis. Unit-level model estimates, on average, generally had higher variance than direct or area-based estimates and produced estimates that exceeded both the direct and area-level estimates. While unit-level models are unbiased, the EBLUP estimator requires the true population values for the covariates to be known. In this analysis, there were discrepancies between the stand estimates derived from the area-level gridded lidar metrics and the averages from the unit-level cloud metrics. The

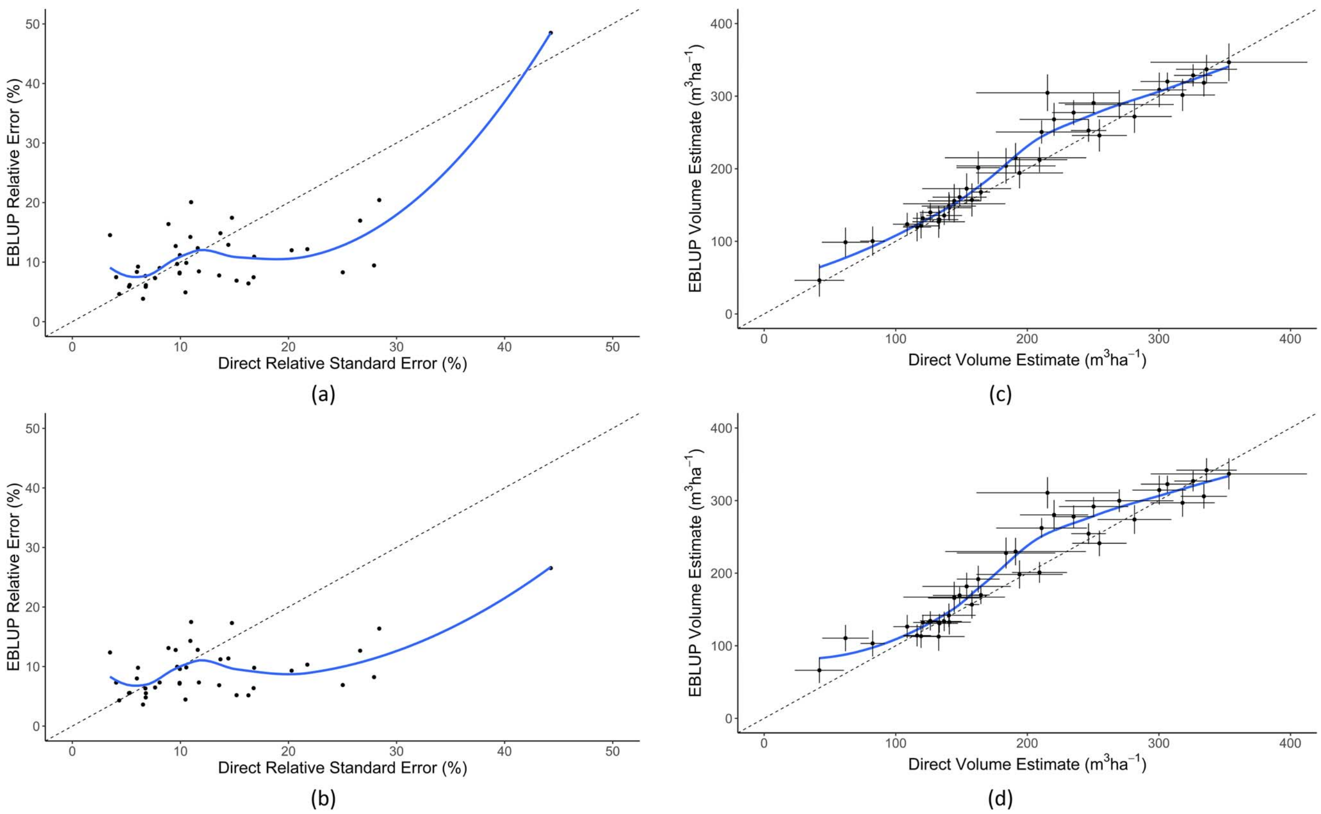

Figure 6 Unit-level SAE results. Model with lidar 80th height percentile as only auxiliary information relative error (RE) comparison (a) best performing model with lidar 80th height percentile and thinning status as auxiliary information RE comparison (b). Estimate comparison for model with lidar 80th height percentile as only auxiliary information (c) and estimate comparison for best performing model with lidar 80th height percentile and thinning status as auxiliary information (d). Smoothing lines are for visual interpretation only and are not representative of the SAE model fit. Error bars represent one standard error in the *x* direction and the root mean squared error in the *y* direction.

Table 5 Comparison of error ratio means of SAE *vs* direct estimates of area- and unit-level models.

Model fit	Area-level error ratio	Unit-level error ratio	Ratio of stands where area-level ratio ≤ unit-level ratio
TV_Planted ~ P80	0.95	1.09	24/40
TV_Planted ~ P90	0.95	1.14	24/40
TV_Planted ~ P95	0.95	1.16	24/40
TV_Planted ~ P99	0.95	1.18	26/40
TV_Planted ~ P80 + thin_status	0.89	0.97	21/40
TV_Planted ~ P90 + thin_status	0.89	1.01	25/40
TV_Planted ~ P95 + thin_status	0.89	1.04	25/40
TV_Planted ~ P99 + thin_status	0.89	1.07	26/40

unit-level 80th percentile height estimates were, on average, ~0.5 m lower than the gridded area-level heights. This is likely due to a combination of spatial smoothing and edge effects from overlapping neighboring tree crowns, which were not completely accounted for by filtering gridded area-level outliers. Due to the possible lack of a sufficient number of plots capturing edge conditions, the average unit-level lidar heights tended to be lower than the corresponding area-level height estimate. Additionally, differences may have arisen due to the two methods used to summarize the lidar point clouds. Finally, despite using a GPS capable of submetre accuracy, location error likely resulted in auxiliary information being summarized for locations different than the areas measured. Future investigations should focus on refining methods to remediate this disparity.

An important assumption when applying SAE techniques is a linear relationship between the auxiliary data and the variable of interest (Rao and Molina, 2015). For this work, we chose to consider a limited set of auxiliary variables that have a theoretical basis for their relationships with our variable of interest total volume. The 80th, 90th, 95th and 99th height percentiles are

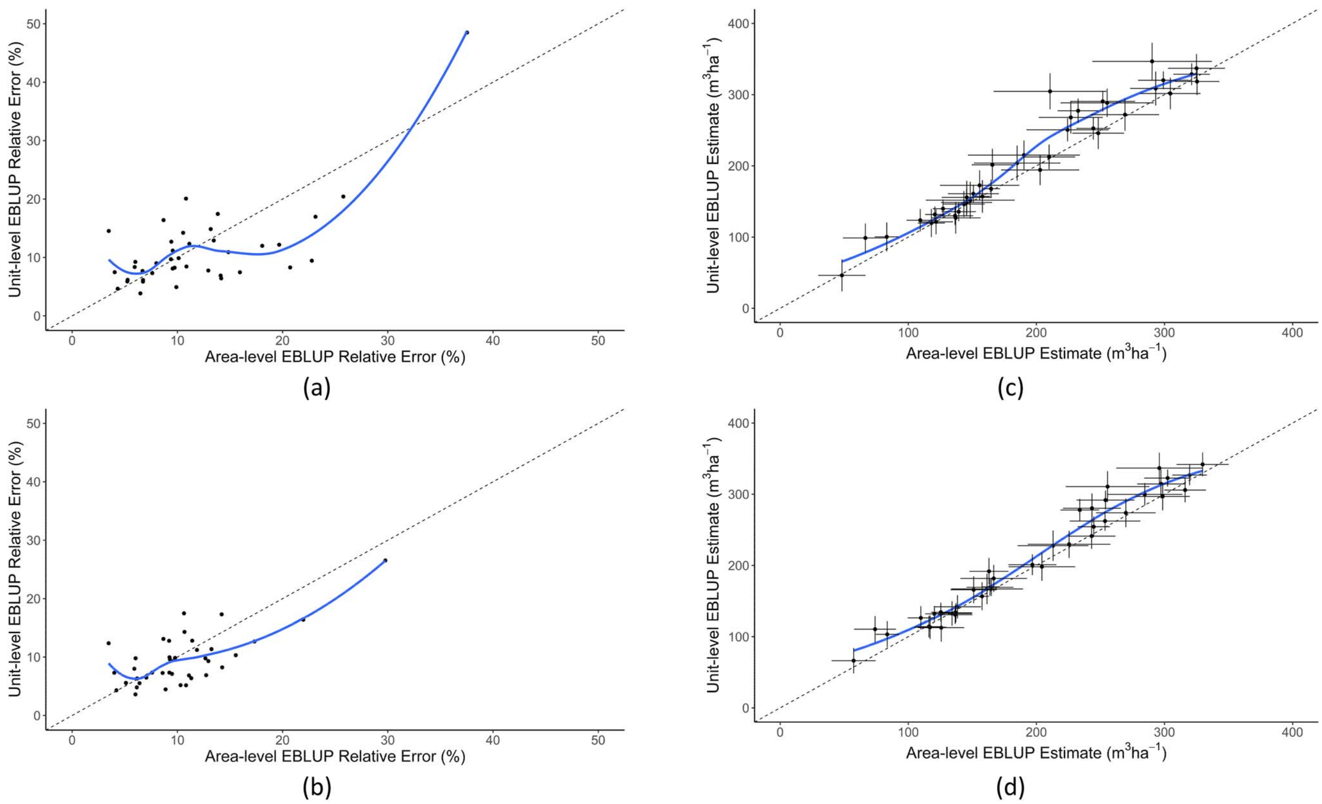

Figure 7 Comparison of area- and unit-level SAE models. Relative error (RE) comparison between area- and unit-level models using the lidar 80th percentile height as the only auxiliary variable (a), RE comparison between area- and unit-level models where the area-level model includes both the lidar 80th percentile height and thinning status as auxiliary information and the unit-level only includes the 80th percentile height as the auxiliary variable (b), estimate comparison between area- and unit-level models using the lidar 80th percentile height as the only auxiliary variable (c), and estimate comparison between area- and unit-level models where the models include both the lidar 80th percentile height and thinning status as auxiliary information (d). Smoothing lines are for visual interpretation only and are not a representative of the SAE model fit. Error bars represent model RMSE in the *x* and *y* directions.

of the most interest due to the linear relationship with height and total volume at similar basal areas (Burkhart *et al.* 2019) and in closed canopy stands (Yanli *et al.,* 2019). While height growth has been found to be significantly affected by density in some studies, within the ranges of densities commonly planted (741–2223 trees per hectare (tph)) differences in height due to density have been found to be either non-significant (Zhao *et al.,* 2011) or of relatively minor differences (e.g. Sharma *et al.,* 2002; Antón-Fernández *et al.,* 2011). Despite the relationship between height and total volume, density is important to consider. The inclusion of thinning status as auxiliary information, in effect, helped ‘adjust’ the volume estimate rather than using lidar-derived height metrics alone. As evident in Figure 4, much of the additional unexplained variation in the linear relationship between the auxiliary data and total planted volume is due to the lidar and thinning status not fully taking stand density into account. In most cases, unthinned planted volume that was underestimated by the linear relationship had higher estimated planted stem density than those overestimated by the linear relationship. The unit-level model including the lidar 80th percentile and thinning status explained similar amounts of variation in planted volume (0.64) compared to the area-level approach (0.67). While lidar metrics have been used to successfully predict stem density, the relationships are generally not as strong as other stand characteristics (Næsset and Bjerknes, 2001; Næsset, 2002; Noordermeer *et al.,* 2019). For this reason, we chose to not include lidar estimates of stand density in this work.

While estimates were generally similar, the unit-level estimates were consistently higher in most cases compared with the area-level estimates without thinning status included (Figure 7c) and when thinning status is incorporated (Figure 7d).

SAE was shown to reliably improve estimate precision in this study; however, these models have limitations. The SAE methods evaluated are not applicable when a direct estimate is not available (Goerndt *et al.,* 2011). In these cases, a model-based, synthetic estimate would be required. Further, it is assumed that sample variances ($\hat{\Psi}_i$) are known without error. While this is often an improbable assumption, it is required for area-level models (Magnussen *et al.,* 2017). Modified variants of the Fay-Herriot, area-level model have been proposed (Wang and Fuller, 2003) which take the uncertainty of the sample variance into account and have been successfully implemented in forest inventory applications (Magnussen *et al.,* 2017). These were not evaluated in this study; however, future investigations should consider their use in plantation forest inventory. Additionally, lidar adds minimal precision when the direct estimates are reliable. The resources spent collecting, processing and incorporating the data may prove unneeded if high-quality ground samples are available. Unit-level models are generally more restrictive and not applicable in cases where fine-scale, sample unit data are not available (Magnussen *et al.,* 2017).

Conclusions

This work has successfully demonstrated the potential for incorporating SAE techniques into operational forest inventory in loblolly pine plantations. Using both lidar and thinning status, the uncertainty of total planted volume estimates was reduced in many cases. To specifically answer our research questions outlined: (1) area-level SAE methods improved the precision compared with direct estimates for all lidar height percentiles evaluated, (2) unit-level SAE methods improved the precision of some estimates for all lidar height percentiles evaluated, particularly when the direct estimates exhibited high variability, (3) unit-level models demonstrated increased precision in some cases compared with area-level methods; however, the average error ratios were lower for area-level methods and (4) the incorporation of additional auxiliary information, in this case thinning status, improved estimation precision in both model formulations. The results of this study should be of interest to forest inventory managers who regularly conduct forest inventory in southern pine plantations. With the increased importance of enhancing and monitoring the productivity of loblolly pine for both commercial and ecological interests such as carbon sequestration (Zhao *et al.,* 2016), accurate and precise estimates of stand volume are essential. Using the methods outlined in this study, the precision of inventory estimates can be improved leading to more confidence when making management and planning decisions. The area-level SAE methods evaluated are broadly applicable to many cases in which linearly related covariates are available. Other sources of auxiliary data such as photogrammetrically derived point clouds and optical satellite platforms may prove useful with these techniques. Additional reductions in uncertainty can be realized if ancillary information can be coupled with data from fixed-area plots with highly accurate center locations under the unit-level SAE framework; however, these methods did not result in precision increases in all stands. Auxiliary data are more limited for unit-level analysis due to the scaling and locational issues.

Acknowledgements

The Forest Modeling Research Cooperative, the Department of Forest Resources and Environmental Conservation, Virginia Tech and the USDA McIntire-Stennis program (Project No. VA-136630) are all gratefully acknowledged for supporting this project. The authors would like to acknowledge John Peterson, Sheng-I Yang, Ralph Amateis and the Virginia Department of Forestry for their invaluable help with field work. Valerie Thomas and Randy Wynne are acknowledged for their valuable insight. Finally, the authors would like to acknowledge the two anonymous reviewers who provided helpful suggestions for improving the manuscript.

Conflict of interest statement

None declared.

Funding

The Forest Modeling Research Cooperative, the Department of Forest Resources and Environmental Conservation, Virginia Tech and the United States Department of Agriculture McIntire-Stennis program (Project No. VA-136630) are all gratefully acknowledged for supporting this project.

References

Andreu, M.G., Zobrist, K. and Hinckley, T. 2008 *Management Practices to Support Increased Biodiversity in Managed Loblolly Pine Plantations*. University of Florida IFAS Extension, Pub. FOR183, p. 8.

Antón-Fernández, C., Burkhart, H.E., Strub, M., Amateis, R.L. 2011 Effects of initial spacing on height development of loblolly pine. *For. Sci.* **57**(3), 201–211.

Battese, G.E., Harter, R.M., Fuller, W.A. 1988 An error-components model for prediction of county crop areas using survey and satellite data. *J. Am. Stat. Assoc.* **83**(401), 28–36.

Bivand, R., Keitt, T., Rowlingson, B. 2019 *rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.4-3*. https://CRAN.R-project.org/package=rgdal Accessed 9 May 2019.

Bivand, R.S., Pebesma, E. and Gomez-Rubio, V. 2013 *Applied Spatial Data Analysis with R*. Second edn. Springer, p. 405.

Breidenbach, J., Astrup, R. 2012 Small area estimation of forest attributes in the Norwegian National Forest Inventory. *Eur. J. For. Res.* **131**(4), 1255–1267.

Breidenbach, J., Magnussen, S., Rahlf, J. and Astrup, R. 2018 Unit-level and area-level small area estimation under heteroscedasticity using digital aerial photogrammetry data. *Remote Sens. Environ.* **212**, 199–211.

Burkhart, H.E., Avery, T.E. and Bullock, B.P. 2019 *Forest measurements*. 6th edn. Waveland Press, p. 434.

Campbell, J.B. and Wynne, R.H. 2011 *Introduction to Remote Sensing*. 5th edn. Guilford Press, p. 667.

Clark, A. III, Phillips, D.R. and Frederick, D.J. 1986 Weight, Volume, and Physical Properties of Major Hardwood Species in the Piedmont. Res. Pap. SE-255. U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station, Asheville, NC, p. 84.

Datta, G.S. and Lahiri, P. 2000 A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems. *Stat. Sin.* **10**, 613–627.

Dragulescu, A. A., Arendt, C., 2018. xlsx: *Read, Write, Format Excel 2007 and Excel 97/2000/XP/2003 Files. R package version 0.6.1*. https://CRAN.R-project.org/package=xlsx Accessed 17 January 2019.

Fay III, R.E., Herriot, R.A. 1979 Estimates of income for small places: An application of James-stein procedures to census data. *J. Am. Stat. Assoc.* **74**(366a), 269–277.

Fox, T.R., Jokela, E.J., Allen, H.L. 2007 The development of pine plantation silviculture in the southern United States. *J. For.* **105**(7), 337–347.

Goerndt, M.E., Monleon, V.J. and Temesgen, H. 2011 A comparison of small-area estimation techniques to estimate selected stand attributes using LiDAR-derived auxiliary variables. *Can. J. For. Res.* **41**, 1189–1201.

Goerndt, M.E., Monleon, V.J., Temesgen, H. 2013 Small-area estimation of county-level forest attributes using ground data and remote sensed auxiliary information. *For. Sci.* **59**(5), 536–548.

González-Manteiga, W., Lombardía, M.J., Molina, I., Morales, D., Santamaría, L. 2008 Bootstrap mean squared error of a small-area EBLUP. *J. Stat. Comput. Simul.* **78**(5), 443–462.

Gyawali, N. and Burkhart, H.E. 2015 General response functions to silvicultural treatments in loblolly pine plantations. *Can. J. For. Res.* **45**, 252–265.

Hijmans, R.J. 2019 *raster: Geographic Data Analysis and Modeling. R package version 2.8-19*. https://CRAN.R-project.org/package=raster Accessed 19 February 2019.

Holmgren, J. 2004 Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. *Scand. J. For. Res.* **19**, 543–553.

Johnsen, K.H., Wear, D., Oren, R., Teskey, R.O., Sanchez, F., Will, R., *et al*. 2001 Meeting global policy commitments: carbon sequestration and southern pine forests. *J. For.* **99**(4), 14–21.

LAStools. 2018 *"Efficient LiDAR Processing Software" (version 180620, academic)*. http://rapidlasso.com/LAStools Accessed 15 May 2019.

Magnussen, S., Mauro, F., Breidenbach, J., Lanz, A. and Kändler, G. 2017 Area-level analysis of forest inventory variables. *Eur. J. For. Res.* **136**, 839–855.

Maltamo, M., Næsset, E. and Vauhkonen, J. 2014 Forestry applications of airborne laser scanning: concepts and case studies. *Manag. For. Ecosys* **27**, 460.

Mauro, F., Monleon, V.J., Temesgen, H., Ford, K.R. 2017 Analysis of area level and unit level models for small area estimation in forest inventories assisted with LiDAR auxiliary information. *PLoS One*, **12**(12), p. e0189401.

McCombs, J.W., Roberts, S.D., Evans, D.L. 2003 Influence of fusing lidar and multispectral imagery on remotely sensed estimates of stand density and mean tree height in a managed loblolly pine plantation. *For. Sci.* **49**(3), 457–466.

McGaughey, R. 2018 FUSION/LDV: Software for lidar data analysis and visualization. FUSION version 3.80. http://forsys.cfr.washington.edu/fusion/fusion_overview.html Accessed 29 January 2019.

McRoberts, R.E., Chen, Q., Gormanson, D.D. and Walters, B.F. 2018 The shelf-life of airborne laser scanning data for enhancing forest inventory inferences. *Remote Sens. Environ* **206**, 254–259.

Means, J., Acker, S., Fitt, B., Renslow, M., Emerson, L., Hendrix C.J. 2000 Predicting forest stand characteristics with airborne scanning lidar. *Photo. Eng. Rem. Sens.* **66**(11), 1367–1371.

Militino, A.F., Ugarte, M.D., Goicoa, T., González-Audícana, M. 2006 Using small area models to estimate the total area occupied by olive trees. *J Agr. Biol. Environ. Stat.* **11**(3), 450–461.

Molina I., Marhuenda Y. 2015 SAE: An R Package for small area estimation. *The R Journal* **7**(1), 81–98. https://journal.r-project.org/archive/2015/RJ-2015-007/RJ-2015-007.pdf.

Næsset, E. 1997a Determination of mean tree height of forest stands using airborne laser scanning data. *ISPRS J. Photo. Rem. Sens.* **52**, 49–56.

Næsset, E. 1997b Estimating timber volumes of forest stands using airborne laser scanner data. *Remote Sens. Environ.* **61**, 246–253.

Næsset, E. and Bjerknes, K.-O. 2001 Estimating tree heights and number of stems in young forest stands using airborne laser scanning data. *Remote Sens. Environ.* **78**, 328–340.

Næsset, E. 2002 Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* **80**, 88–99.

Næsset, E. 2004a Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scand. J. For. Res.* **19**, 164–179.

Næsset, E. 2004b Accuracy of forest inventory using airborne laser scanning: evaluating the first Nordic full-scale operational project. *Scand. J. For. Res.* **19**, 554–557.

Næsset, E. 2007 Airborne laser scanning as a method in operational forest inventory: status of accuracy assessments accomplished in Scandinavia. *Scand. J. For. Res.* **22**, 433–442.

Nelson, R., Krabill, W. and Maclean, G. 1984 Determining forest canopy characteristics using airborne laser data. *Remote Sens. Environ.* **15**, 201–212.

Nelson, R., Krabill, W. and Tonelli, J. 1988 Estimating forest biomass and volume using airborne laser data. *Remote Sens. Environ.* **24**, 247–267.

Noordermeer, L., Bollandsås, O.M., Økra, H.O., Næsset, E. and Gobakken, T. 2019 Comparing the accuracies of forest attributes predicted from airborne laser scanning and digital aerial photogrammetry in operational forest inventories. *Remote Sens. Environ* **226**, 26–37.

Pebesma, E.J., R.S. Bivand. 2005 Classes and methods for spatial data in R. *R News* **5**(2), https://cran.r-project.org/doc/Rnews/

Popescu, S.C., Wynne, R.H. 2003 Seeing the trees in the forest: using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photo. Eng. Rem. Sens.* **70**(5), 589–604.

Popescu, S.C., Wynne, R.H., Nelson, R.F. 2003 Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Can. J. Rem. Sens.* **29**(5), 564–577.

QGIS Development Team 2019 QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org Accessed latest version used 26 August 2019.

R Core Team, 2018 R: a language and environment for statistical computing [online]. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Rao, J.N.K. and Molina, I. 2015 *Small Area Estimation*. Wiley, p. 441.

Sharma, M., Burkhart, H.E., Amateis, R.L. 2002 Modeling the effect of density on the growth of loblolly pine trees. *South. J. Appl. For.* **26**(3), 124–133.

Schultz, R.P. 1999 Loblolly–the pine for the twenty-first century. *New For.* **17**(1–3), 71–88.

Tasissa, G., Burkhart, H.E., Amateis, R.L. 1997 Volume and taper equations for thinned and unthinned loblolly pine trees in cutover, site-prepared plantations. *South. J. Appl. For.* **21**(3), 146–152.

Tukey, J.W. 1977 *Exploratory data analysis*. Addison-Wesley.

U.S. Geological Survey. 2017a 1 meter Digital Elevation Models (DEMs) - USGS National Map 3DEP Downloadable Data Collection, U.S. Geological Survey

U.S. Geological Survey. 2017b lidar Point Cloud - USGS National Map 3DEP Downloadable Data Collection, U.S. Geological Survey.

van Aardt, J.A., Wynne, R.H., Oderwald, R.G. 2006 Forest volume and biomass estimation using small-footprint lidar-distributional parameters on a per-segment basis. *For. Sci.* **52**(6), 636–649.

Wang, J., Fuller, W.A. 2003. The mean squared error of small area predictors constructed with estimated area variances. *J. Am. Stat. Assoc.* **98**(463), 716–723.

Warner, J.R. and Goebel, N.B. 1963 Total and bark volume tables for small diameter Loblolly, Shortleaf, and Virginia Pine in the upper South Carolina Piedmont. Department of Forestry, South Carolina Agricultural Experiment Station, Forest Research Series No. 9, Clemson, South Carolina.

Wickham, H. 2007 Reshaping data with the reshape package. *J. Stat. Softw.* **21**(12), 1–20. URL http://www.jstatsoft.org/v21/i12/.

Wickham, H. 2016 *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, p. 260.

Yanli, X., Chao, L., Zhichao, S., Lichun, J. and Jingyun, F. 2019 Tree height explains stand volume of closed-canopy stands: Evidence from forest inventory data of China. *For. Ecol. Manage.* **438**, 51–56.

Yu, X., Hyyppä, J., Holopainen, M. and Vastaranta, M. 2010 Comparison of area-based and individual tree-based methods for predicting plot-level forest attributes. *Rem. Sens.* **2**, 1481–1495.

Zhao, D., Kane, M., Borders, B.E. 2011 Growth responses to planting density and management intensity in loblolly pine plantations in the southeastern USA Lower Coastal Plain. *Ann. For. Sci.* **68**(3), 625–635.

Zhao, D., Kane, M., Teskey, R., Fox, T.R., Albaugh, T.J., Allen, H.L. *et al.* 2016 Maximum response of loblolly pine plantations to silvicultural management in the southern United States. *For. Ecol. Manage.* **375**, 105–111.