Forest Ecology and Management 409 (2018) 552–563

Contents lists available at ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

Evaluation of alternative approaches for landscape-scale biomass estimation in a mixed-species northern forest

Coeli M. Hoover[a,*], Mark J. Ducey[b], R. Andy Colter[c], Mariko Yamasaki[a]

[a] USDA Forest Service, Northern Research Station, Durham, NH 03824, United States
[b] University of New Hampshire, Department of Natural Resources and the Environment, Durham, NH 03824, United States
[c] USDA Forest Service, White Mountain National Forest, Campton, NH 03223, United States

ARTICLE INFO

Keywords:
Aboveground biomass
LiDAR
Forest inventory

ABSTRACT

There is growing interest in estimating and mapping biomass and carbon content of forests across large landscapes. LiDAR-based inventory methods are increasingly common and have been successfully implemented in multiple forest types. Asner et al. (2011) developed a simple universal forest carbon estimation method for tropical forests that reduces the amount of required field measurements. We tested this approach, along with standard regression and Random Forest modeling techniques, in a northern hardwood-dominated watershed in the White Mountains of New Hampshire. Additional objectives included assessing the effects of different inventory plot designs and GPS accuracy. The universal model performed poorly in this forested landscape due to the lack of a clear relationship between canopy height and stand basal area. Simple regression modeling also produced poor model fits; the Random Forest models produced somewhat better biomass predictions than either the universal or regression models, and had low predictive power as measured by R^2 but root mean squared errors were comparable to those from other studies in complex forests. Effects of positional accuracy from survey vs. resource grade GPS units were slight, as were the effects of varying plot designs, although errors generally increased when larger basal area factors were used.

1. Introduction

Inventory and monitoring is an essential, but expensive, component of forest management. Inventory data are important for meeting multiple management objectives including timber production, wildlife habitat, forest health, and carbon sequestration (Kershaw et al., 2016). At regional to national scales, National Forest Inventory data meet some needs. For example, while the USDA Forest Service's Forest Inventory and Analysis Program (FIA) is a source of detailed forest inventory data (USDA Forest Service, 2017), the inventory is designed to be used at the state, regional, or national level, with one plot every 2428 ha (Bechtold and Patterson, 2005). As such, these data are generally not appropriate for landowners or managers due to the resolution of the sampling design. Because extensive field work is needed to collect inventory data at the level of stands or small landscapes, conducting routine forest inventories that meet an acceptable accuracy threshold is often quite expensive.

Airborne LiDAR, or light detection and ranging, employs a laser and high precision GPS to produce a three-dimensional representation of the ground beneath the aircraft's path; as the laser's energy hits a surface, it is reflected back to the instrument and recorded. Multiple returns are possible from each laser pulse. Airborne LiDAR has been in use for some time for terrain mapping; this product typically has a low return density (1–2 pulses per square meter, or ppsm) and is acquired when the forested portions of the landscape are in a leaf-off condition. Higher-density LiDAR data from full waveform and discrete return instruments have been used by researchers to assess various forest characteristics such as tree density, diameter, basal area (BA), and biomass (e.g. Lefsky et al., 1999, Beets et al., 2011, Hudak et al., 2006).

LiDAR studies of forest structure have occurred across a variety of biomes from tropical to boreal forests with variable model results; often with better results in conifer types or managed landscapes, where tree and forest structure is less complex and more regular. Zolkos et al. (2013) performed a meta-analysis on 70 studies reporting carbon or biomass across a variety of biomes to assess remote sensing approaches for measuring forest biomass. Biomass modeled from discrete return LiDAR data had an overall mean R^2 of 0.76, and a mean RMSE of 39.4 Mg/ha. In addition, they found that model error (in both absolute and relative terms) varied by forest type. Anderson and Bolstad (2013) estimated biomass in a Wisconsin forest by fitting LiDAR models by

* Corresponding author at: USDA Forest Service, Northern Research Station, 271 Mast Road, Durham, NH 03824, United States.
E-mail address: choover@fs.fed.us (C.M. Hoover).

https://doi.org/10.1016/j.foreco.2017.11.040
Received 28 September 2017; Received in revised form 14 November 2017; Accepted 18 November 2017
0378-1127/ Published by Elsevier B.V.

vegetation type, and report an R^2 of 0.74 and RMSE of 37.6 Mg/ha for coniferous stands, values of 0.71 and 42.8 Mg/ha for hardwood stands, while the mixed stands model had an R^2 of 0.44 with an RMSE of 48 Mg/ha. When all plots were included in the model, the R^2 was 0.55 with an RMSE of 43.5 Mg/ha.

While the cost of LiDAR data acquisition is dropping, use of these data for operational purposes requires that the field measurement component needed to model forest attributes be conducted efficiently. Because of the effort and cost associated with collecting information from a sufficient number of plots in each forest stratum, there have been efforts to generalize the modeling process to reduce the number of variables and/or plots measured. Lefsky et al. (2002) conducted an early test of a generalized model using waveform LiDAR at three sites in the US: temperate deciduous, temperate coniferous, and boreal coniferous (note that throughout this manuscript, deciduous refers to broadleaved species only). For both temperate coniferous and temperate deciduous sites, the R^2 values for the site specific and general (all sites combined) models were the same (0.87 and 0.65, respectively), while for the boreal conifer site the individual model outperformed the general (0.76 and 0.56, respectively). Use of a generalized model, if validated for a sufficient number of forests, would be one approach to reducing the field data collection burden. For tropical forests, Asner et al. (2011) developed a general approach to estimating aboveground biomass using mean canopy height from LiDAR data, and plot-level measurements of basal area and wood density weighted by basal area. Comparing the predicted and measured aboveground carbon for all 482 plots across four tropical study locations resulted in an R^2 of 0.95 with an RMSE of 15 MgC/ha. Substituting a regional wood density value produced an R^2 of 0.92. Asner and Mascaro (2014) tested a similar approach using hundreds of plots across 14 tropical ecoregions, and found that while LiDAR-derived canopy height accounted for 56% of the variation in aboveground carbon stock, a model that added basal area and wood density increased that value to 92%.

The intent of this study is to test if this type of generalized approach is feasible in the New England forested landscape, where deciduous, coniferous, and mixed stands are present. Use of a more generalized model with moderate resolution LiDAR data could provide an operationally feasible approach to LiDAR-based estimation of forest characteristics that would be practical for use by managers. We have four major objectives:

1. Test the Asner et al. (2011) approach for estimating aboveground biomass in a Northern hardwood forest.
2. Compare results from the Asner approach to those from conventional estimation methods.
3. Evaluate the suitability of moderate resolution LiDAR data for estimating common structural variables such as trees per hectare, basal area, and height.
4. Evaluate the impacts of changing plot design (including variable radius plot or prism sampling) and positional accuracy on the modeled outputs.

2. Materials and methods

2.1. Study area

The study was conducted in a small forested watershed on the Pemigewasset Ranger District of the White Mountain National Forest, located in Grafton County, New Hampshire, USA (Fig. 1). The study watershed is centered approximately at 44.0657°N, 71.8183°W and is 6885 ha in size. Elevation ranges from 328 to 1463 m, with slopes ranging from 0 to 85%. Annual precipitation averages about 1400 mm. The soils range from drainage classes of excessively drained to very poorly drained and have soil temperature regimes of frigid at the lower elevations to cryic at the higher elevations. The study area is predominantly of the soil order Spodosol and as soil parent materials of Lodgement and Ablation glacial tills, along with areas of Alluvium, Glaciofluvial and bedrock controlled outcrops. The vegetation is largely second growth and is a typical northern hardwood forest, consisting of sugar maple (*Acer saccharum*), American beech (*Fagus grandifolia*), yellow birch (*Betula alleghaniensis*), paper birch (*Betula papyrifera*), white ash (*Fraxinus Americana*), red oak (*Quercus rubra*), and red maple (*Acer rubrum*), with a conifer component of Eastern hemlock (*Tsuga canadensis*), balsam fir (*Abies balsamea*), white pine (*Pinus strobus*), and red spruce (*Picea rubens*).

2.2. Field data collection and processing

Plot locations were selected by stratified random sampling, with strata based on overlaying a conjectured soil group map (based on landform) with management zones defined by U.S. Forest Service regulations. A total of 176 plot locations were selected across the watershed. Field crews navigated to the specified coordinates for each plot using a recreational-grade GPS. Once the plot center was monumented and established, the plot was georeferenced using both a survey-grade (Trimble GeoXH, CE with Zephyr antenna) and resource-grade (Trimble GPS Pathfinder ProXH with Hurricane L1 antenna) GPS. Field data were collected in the summers of 2013 and 2014, after the final LiDAR acquisition was completed (the LiDAR data were used to inform plot selection).

All trees above 2.5 cm diameter at breast height (DBH) were measured using a mapped, nested-plot design. Trees from 2.5 to 12.6 cm DBH were measured on a 4.23 m radius fixed-radius plot, while trees from 12.7 cm to 30.0 cm DBH were measured on a 10 m radius fixed-radius plot. Trees 30.1 cm DBH and over were measured out to the limiting distance for a 2.25 m^2/ha basal area factor (BAF) variable radius plot. In addition to species, status as live or dead, and DBH (to the nearest 0.1 cm with a tape), the distance and bearing from the plot center to the pith of each tallied tree was recorded for all trees, allowing later simulation of sampling from a 10 m fixed-radius plot (with nested subplot for small trees), and variable radius plots with a range of BAF spanning and exceeding conventional inventory recommendations for the region (typically ranging from 3.5 to 4.6 m^2/ha, but with some practitioners using 2.3 m^2/ha; Wiant et al., 1984, Ducey, 2001), as well as the full original plot design. On all trees close enough to the plot center to be tallied using a 4 m^2/ha BAF variable radius plot, total height was measured using a Vertex hypsometer (Haglof, Inc.). The trees measured for height represent a size-weighted probability-based subsample of the full sample of those measured for DBH (Marshall et al., 2004; Kershaw et al., 2016, ch. 11).

To predict the heights of trees for which heights were not measured, we evaluated a series of regression equations using a mixed-effects modeling framework (Pinheiro and Bates, 2000), relying primarily on information-theoretic model selection using the Akaike Information Criterion (AIC) (Akaike, 1974, Burnham and Anderson, 2002) but with additional consideration of other regression diagnostics, including correlations between parameter estimates, error distribution and correlation, and Schwarz's Bayesian Information Criterion (BIC). The overall philosophy in model selection for this study was holistic, aiming at reliable predictions constructed from a model built on distributional assumptions that are satisfied by the data, rather than relying on automatic selection of a model by a single criterion (Claeskens and Hjort, 2008). We did not consider, and do not report here, p-values associated with regression coefficients: tree height is known to be correlated with tree diameter, and height-diameter relationships are known to depend on species and vary with site, both in general and in this region (e.g. Ducey, 2012), and since the null hypothesis of no relationship is not credible, such an approach would be a misuse of the null hypothesis testing paradigm (Anderson et al., 2000). We evaluated two primary model forms. The first followed Schumacher and Hall (1933) in log-transforming tree height H, and taking the reciprocal of DBH:

Fig. 1. Map showing location and topography of the study site on the White Mountain National Forest in New Hampshire, USA.

$$\ln(H_i) = (\beta_0 + \gamma_{0,s} + \phi_{0,p}) + (\beta_1 + \gamma_{1,s} + \phi_{1,p})/DBH_i + \epsilon_i \quad (1)$$

where β_0 and β_1 are fixed effects, $\gamma_{0,s}$ and $\gamma_{1,s}$ are random effects of species, $\phi_{0,p}$ and $\phi_{1,p}$ are random effects of plot location, and ϵ_i is the true (i.e., tree-level) residual. We also tested a log-log model, i.e.,

$$\ln(H_i) = (\beta_0 + \gamma_{0,s} + \phi_{0,p}) + (\beta_1 + \gamma_{1,s} + \phi_{1,p})\ln(DBH_i) + \epsilon_i \quad (2)$$

All regressions were fit in package lme4 (Bates et al., 2015) in the *R* statistical package (R Core Team, 2016). Models involving simplified random effects models (i.e. omitting $\gamma_{1,s}$, $\phi_{1,p}$, or both) were also fit; model fits were assessed using a combination of the Akaike Information Criterion (AIC), residual and quantile plots for the random effects and true residual (i.e. the residual at the individual observation level, not including random effects), and correlations within and between the fixed and random effects. Once a final equation form and error model was selected, heights were imputed for all trees without height measurements. Imputation included prediction of $\ln(H_i)$ using the fixed effects, as well as the modeled species effect and plot effect associated with the individual tree. For the very small number of trees of species that had no modeled species effect (because no trees of that species were measured for height), a random number from the distribution of the species effects was drawn. Finally, a normal deviate from the distribution of the true residuals was added, and then H_i was computed by taking the antilogarithm. Because error terms are assigned from their modeled distribution for the transformed variable (i.e. $\ln(H_i)$), then back-transformed, the need for bias-correction in logarithmic regression (e.g. Baskerville, 1972) is eliminated.

We calculated the biomass of individual trees using the allometric equations of Chojnacky et al. (2014), which represent a refinement of the generalized equations developed by Jenkins et al. (2003), and individual trees were assigned a wood specific gravity value based on Miles and Smith (2009). To allow comparison of the effects of sample design on LiDAR-based retrieval of stand-level parameters, we performed identical computations for the full sample (all trees tallied based on the field design described above, hereafter "full"), a fixed-radius plot sample (all trees above 12.6 cm DBH, including those 30.1 cm and above, restricted to those within 10 m of the plot center and horizontal point sampling with a basal area factor (BAF) of 2.25 m^2/ha and all integer values from 3 to 9 m^2/ha (including only those trees whose actual distance from the plot center was less than or equal to the corresponding limiting distance for a tree of that DBH under that BAF; Kershaw et al., 2016, ch. 11). Each of these sampling designs can be viewed as describing a function $r_d(DBH)$ that relates the limiting distance (the maximum distance a sample point may fall from the tree, and still include the tree in the sample) to the DBH for a given sample design *d*. For example, in the fixed-radius design, r_d is a step function; in horizontal point sampling, r_d is a simple linear function with the slope determined by the BAF. For each tree under each sampling design, we computed the tree factor as

$$TF_i = \frac{10{,}000}{\pi\,[r_d(DBH_i)]^2}$$

The tree factor determines the scaling for any and all contents of an

individual tree, from the sample point to a per-hectare value (Kershaw et al., 2016, ch. 9). Using these tree factors, we computed for each sample point and for each sample design, a series of common descriptors of stand structure. These included number of trees/ha, basal area (m^2/ha), stand quadratic mean diameter (cm), Lorey's height (the basal area-weighted mean height of trees in the stand, m), top height (the height of the tallest 100 trees/ha in the stand), and live tree biomass (Mg/ha). We also computed relative density using the mixed-species density measure of Ducey and Knapp (2010); relative density is an index of tree crowding that equals 0 for an unstocked stand, and 1 for a stand at "normal" or "A-line" stocking. The measure is a better measure of overall ecological packing than basal area, because it accounts for tree size as well as species functional traits. Finally, we computed the proportion of relative density in the stand contributed by conifer species, and the basal area-weighted mean wood specific gravity.

2.3. LiDAR data collection and processing

Discrete-return, small-footprint LiDAR data were acquired from a commercial vendor (Photo Science Geospatial Solutions, Lexington, KY) during November 2010 and April 2012 under snow-free leaf-off condition, since data were primarily acquired for a terrain-related project. The second flight was necessary to fill coverage holes that were found during the vendor's quality assessment process. These data are typical of data currently being collected over large areas in the region by public agencies for terrain mapping; much of that data is available free of charge to the public, and hence would be attractive if it could be used for forest inventory purposes as well. A Leica ALS70 instrument was used, with flying heights between 1524 and 1676 m. A scan rate of 41.5–45.6 Hz was used, field of view was 32 degrees, and sidelap was 20%, with a targeted point density of 3 points/m^2. Vendor supplied LAS files were processed using FUSION v. 3.42, which is available free of charge (McGaughey, 2014). The FUSION QA/QC process was run and indicated an average return density of 3.4/m^2, meeting the project specifications and providing an average of over 1000 returns within the footprint of a fixed-radius field plot. FUSION computes a large number of metrics based on the elevation distribution of LiDAR returns. We selected 42 of these, including all 6 of the metrics described as cover metrics by McGaughey (2014). Our selected metrics included the elevation minimum, maximum, mean, and mode; several statistics describing the distribution shape including the standard deviation, variance, coefficient of variation, interquartile range, skewness, kurtosis, average absolute deviation, mean absolute deviation from the median, and mean absolute deviation from the mode; the L_2, L_3, and L_4 norms of elevation (note the L_1 norm is identical to the mean); 15 quantiles of the elevation distribution ranging from 0.01 to 0.99; the canopy relief ratio; and root mean square elevation. In addition, the 6 cover metrics included the percentages of first returns and of all returns above 2 m, above the mean, and above the mode. For more details on these metrics, see McGaughey (2014). All metrics were computed using a cell size of 20 m (corresponding to the fixed-radius plot size and shape) and returns over a 2 m height cutoff. For each of the 176 field plots, the corresponding area was clipped from the LiDAR point cloud using the survey grade GPS coordinates and a variety of height, density, and other metrics were computed, for use in the modeling process. In addition, individual files for each metric were created for the entire project area.

2.4. Statistical analysis and modeling

To evaluate the Asner et al. (2011) approach in our study site, we fit a series of regression equations to the plot-level data. Because they corresponded most closely in spatial extent to the extraction of the LiDAR metrics, we used the fixed-radius plot design data for this analysis. All regressions were fit using the lm function in the base R package (R Core Team, 2016), and all regressions in which the dependent variable had been log-transformed were bias-corrected following Baskerville (1972).

Asner et al. (2011) begin by assuming a stand-level carbon or biomass allometry of the form

$$AGB = a \times BA^b \times H^c \times SG^d$$

where AGB is aboveground biomass (Mg/ha), BA is basal area/ha (m^2/ha), H is mean canopy height (m), and SG is basal area-weighted mean specific gravity. This motivates an initial regression of the form

$$\ln AGB = \beta_0 + \beta_1 \ln BA + \beta_2 \ln \overline{elev} + \beta_3 \ln \overline{SG}$$

where $\overline{elev}$ is mean elevation of the LiDAR returns, and $\overline{SG}$ is the basal area-weighted mean wood specific gravity of the stand. Next, Asner et al. (2011) use a ratio equation between H and BA to estimate stand-specific basal area. Following examination of a scatter plot to identify whether heteroscedasticity was present, we fit this relationship as a regression without an intercept using ordinary least squares:

$$BA = \beta_4 \overline{elev}$$

Last, Asner et al. (2011) substitute a regional mean wood specific gravity for the plot-specific mean to derive predictions. We evaluated these predictions, with and without the substitution of LiDAR-derived variables for individual field-measured variables, using additional regression analysis and equivalence testing (Robinson et al., 2005) to describe whether the predictions were close to the 1:1 line either individually or as an ensemble.

To evaluate the potential of multiple LiDAR metrics for predicting aboveground biomass and other structural characteristics in our study watershed, we employed the Random Forest algorithm (Breiman, 2001). Random Forest is a machine learning technique that has seen extensive application in remote sensing, including classification (reviewed by Belgiu and Drăguţ, 2016) as well as regression and imputation using LiDAR data (e.g. Hudak et al., 2008, Yu et al., 2011, Hayashi et al., 2014). In comparison with linear regression methods, Random Forest regression has several advantages: built-in safeguards against overfitting, no need for *a priori* variable selection, internal assessment of model performance on so-called "out of bag" samples (eliminating the need for separate training and validation data), and avoidance of distributional assumptions (Breiman, 2001). We fit separate Random Forest regression models for each of the stand structural variables, and for each simulated sampling design, using the 42 candidate LiDAR metrics described above, using package randomForest (Liaw and Wiener, 2002) in R (R Core Team, 2016). Each model was built with 500 decision trees, using 1/3 of the available variables for each tree. Models were evaluated using the R^2, root mean square error (RMSE), and relative root mean square error (RRMSE; $100 \times RMSE/\bar{y}$) as computed on the out-of-bag samples.

Finally, to assess the influence of geopositioning accuracy, we re-extracted all of the LiDAR metrics for the plots using the coordinates as measured using the resource grade GPS. We then refit all of the Random Forest regressions for every stand structural variable and sample design, using the new LiDAR metrics. Because the resource grade GPS did not obtain a fix for a small number of the field plots, we also refit the models using the metrics based on the survey-grade locations, but using only the subset of plots with a resource-grade fix. The influence of geopositioning accuracy was evaluated using the difference in RRMSE between the models based on resource-grade and survey-grade models.

3. Results

The study watershed is dominated by closed canopy forest that is largely mature; summary data from the field plots is presented in Table 1. Only 6 plots (less than 4% of the total) had live tree aboveground biomass less than 50 Mg/ha; only 10 plots (less than 6% of the total) had relative density values less than 0.5, indicating insufficient live tree stocking to support a closed canopy. Mean and median values

Table 1
Summary of plot-level variables for the Wild Upper Ammonoosuc study area, based on data from the fixed design.

	Minimum	Median	Mean	Maximum	Std. Deviation
Trees/ha	223	1805	2452	17880	2328
Basal Area (m^2/ha)	1.0	34.4	34.5	68.6	11.7
QMD[a] (cm)	4.3	14.9	15.8	33.2	5.8
Relative Density	0.03	0.93	0.94	1.90	0.30
Lorey's Height (m)	5.9	18.0	17.4	27.4	5.5
Top Height (m)	7.0	20.9	20.4	30.3	5.3
Live Tree AG[b] Biomass (Mg/ha)	3.0	200.5	200.4	418.8	78.2
Conifer Fraction	0.00	0.08	0.23	1.00	0.28
Basal Area-weighted Wood SG[c]	0.35	0.56	0.54	0.64	0.08

[a] Quadratic mean diameter at breast height.
[b] Aboveground.
[c] Specific gravity.

were similar for basal area (BA), relative density, height (Lorey's and top), wood specific gravity, and live tree aboveground biomass. The regression model selected for individual tree heights was

$$\ln(H_i) = (0.951(\pm 0.046) + \gamma_{0,s} + \phi_{0,p}) + 0.563(\pm 0.011)\ln(DBH_i)$$

where standard errors for the parameters are given in parentheses, with a marginal R^2 (ignoring species- and plot-level variation) of 0.65, and a conditional R^2 (including species- and plot-level effects) of 0.80, as calculated following Nakagawa and Schielzeth (2013). The variance associated with the species effect $\gamma_{0,s}$ was 0.022; that for the plot effect $\phi_{0,s}$ was 0.013, while the true residual ϵ_i had variance 0.044. This latter term translates into a standard error of prediction for individual tree heights of approximately 23%. Only 3 tree species of the 22 encountered in the full sample lacked height measurement trees, and hence an individual species effect (*Prunus pensylvanica*, *Ostrya virginiana*, and *Ulmus americana*); collectively, these 3 species represented 0.23% of trees in the full sample. The LiDAR-derived canopy height map (Fig. 2) shows the variability of canopy height across the watershed using a 10 m pixel and indicates a range of 0–39 m for top of canopy height, which is broadly consistent but slightly higher than the range for either Lorey's height or top height from the field plots.

3.1. Implementation of Asner et al. (2011) generalized approach

Asner et al. (2011) approach involves substituting mean wood density and a regional relationship between LiDAR mean canopy height and basal area for field measured basal area values to estimate aboveground biomass with minimal field work, as outlined above. Following Asner et al. (2011), we first fit the predictive relationship for aboveground live biomass (AGB) using LiDAR mean canopy height and plot-specific, field-measured values of basal area and wood density, yielding the relationship

$$\ln AGB = 0.926(\pm 0.112) + 1.062(\pm 0.020)\ln BA + 0.424(\pm 0.035)\ln\overline{elev} + 0.701(\pm 0.072)\ln\overline{SG}$$

All of the parameter values are significantly different from 0 ($p < .0001$), and $R^2 = 0.96$ (for log-transformed biomass). Following bias-correction (after Baskerville, 1972), the resulting predictive model is (Fig. 3a)

$$AGB = 2.544 BA^{1.062}\overline{elev}^{0.424}\overline{SG}^{0.701}$$

The next step in the Asner et al. (2011) approach is to substitute a LiDAR-predicted basal area:

$$\widehat{BA} = 2.688(\pm 0.092)\overline{elev}$$

Although the slope in this relationship is well-constrained, the result is an immediate and dramatic deterioration in the predictive relationship for biomass (Fig. 3b). By contrast, using the average $\overline{SG}$ for the plots (0.539 ± 0.006) but retaining field-measured basal area does not result in such deterioration (Fig. 3c). The combined effect of using LiDAR-predicted basal area and average $\overline{SG}$ can be seen in Fig. 3d; incorporating both substitutions leads to the predictive relationship

$$AGB = 4.714\overline{elev}^{1.486}$$

The variance of the prediction errors for this relationship (and that in Fig. 3b) is greater than the original variance of field-estimated biomass, leading to a negative pseudo-r^2: prediction errors would be lower using a constant prediction (the mean field-estimated biomass) instead of the Asner et al. (2011) method. The lack of predictive capacity is clearly associated with the inadequate relationship between *BA* and $\overline{elev}$ (Fig. 4). Although a conventional r^2 is difficult to interpret for regression-through-the-origin models (Eisenhauer 2003), the poor RMSE (exceeding the original standard deviation of *BA*) clearly indicates a poor fit.

By contrast, direct log-log regression of biomass on mean elevation gives potentially useful, if weak, prediction. The original regression gives

$$\ln AGB = 2.902(\pm 0.284) + 0.940(\pm 0.116)\ln\overline{elev}$$

Following bias-correction, the resulting predictive relationship is

$$AGB = 20.884\overline{elev}^{0.940}$$

which is a considerably more linear relationship than that implied by the Asner et al. (2011) approach. The original regression and back-transformed predictions are shown in Fig. 5. Unlike the Asner et al. (2011) result, the RMSE is less than the original standard deviation of biomass; however, the percentage of variance explained is low (16%).

3.2. Random Forest models

The fit of the Random Forest regression models, for each forest structural attribute and each sampling design, is described in terms of R^2 in Table 2, and in terms of RMSE and RRMSE in Table 3. Of immediate interest is the performance of the model for live tree aboveground biomass, under the fixed-radius plot design, as this is the same variable and design used to evaluate the Asner et al. (2011) approach. The out-of-bag R^2 and RMSE for the Random Forest model are better than the comparable statistics for direct modeling using $\overline{elev}$, but not dramatically so. A plot of observed vs. predicted values is shown in Fig. 6. This suggests that the LiDAR metrics as an ensemble contain limited information for estimating biomass at this scale in this forested landscape. The FUSION cover variables scored high in variable importance for this model (taking 5 of the top 6 positions as judged using percent increase in MSE); the importance of individual elevation distribution metrics appeared to have been diluted by the presence of a large number of highly correlated variables.

Examining the impact of sampling design on biomass prediction, performance was best (as judged by R^2) for the fixed-radius design, but best (as judged by RMSE) for the full design, which incorporated slightly larger inclusion zones for trees over 30 cm DBH. Broadly speaking, performance was similar for the full, fixed-radius, and variable radius designs with BAFs of 3 m^2/ha and less, with performance gradually declining as BAF exceeded 4 m^2/ha. This same general pattern can be observed for the other structural variables, with the exception of trees/ha and quadratic mean diameter (QMD), which one would expect would have a high sampling variance under the variable radius designs with their small inclusion zones for small-diameter trees. There were also clear differences among structural variables in their predictability from LiDAR. Lorey's height was consistently the easiest variable to predict, and remained predictable even with relatively large BAFs; top height, which is sensitive only to the tallest 3 trees on each plot in the fixed-radius plot design, was less predictable but similarly insensitive to sampling design. The fraction of relative density

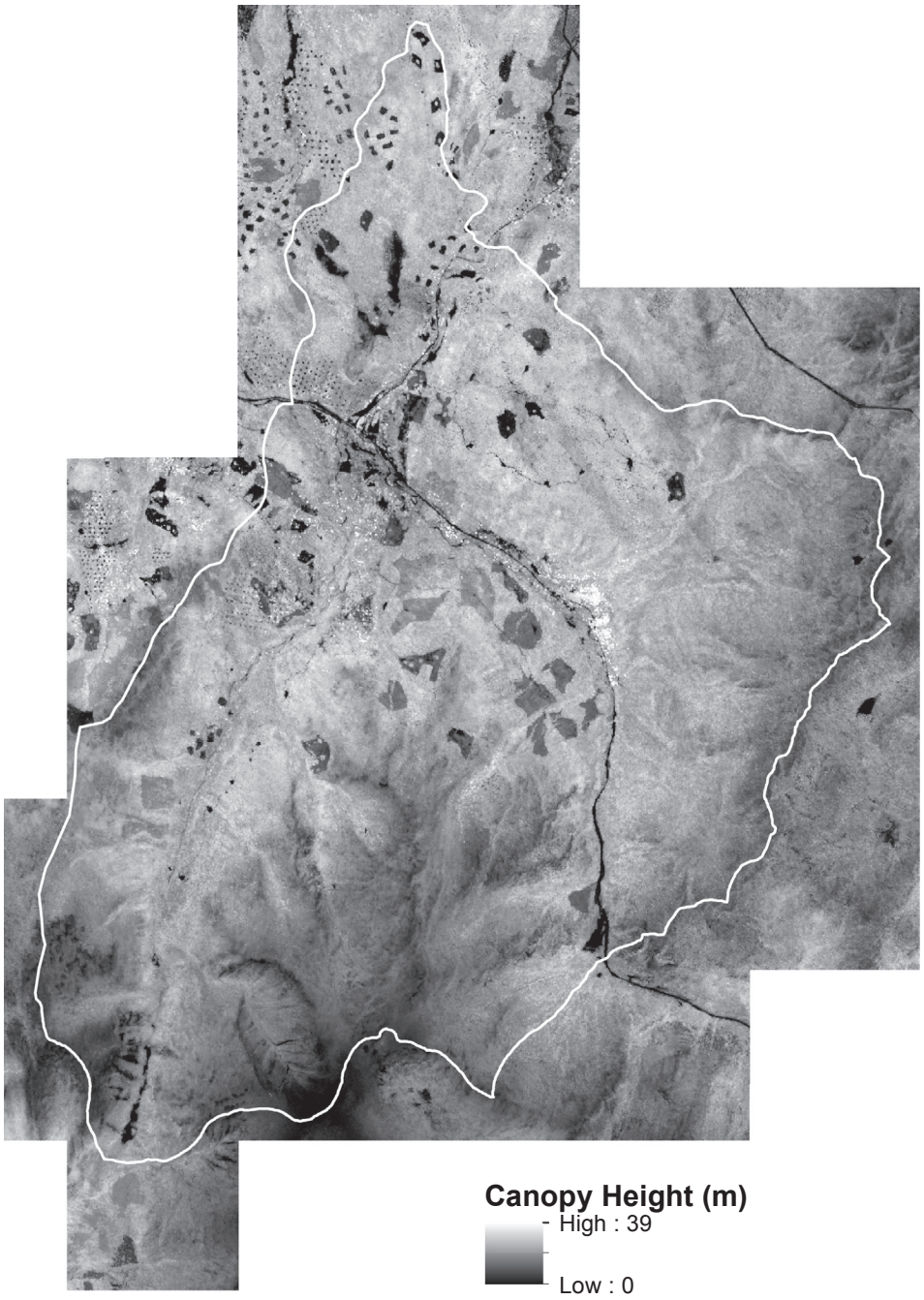

Fig. 2. Canopy height map of study area, 10 m pixel.

contributed by conifers, a simple metric of composition, was also relatively amenable to prediction across a range of sampling designs. These patterns are illustrated for the full, fixed-radius, and BAF 3 m^2/ha designs, for selected variables, in Fig. 7.

Differences between survey-grade and resource-grade positioning were substantial relative to the fixed-radius design plot. The root mean squared difference in position between the two GPS units was 2.59 m. However, this difference was heavily influenced by a small number of outliers: the median difference was 1.19 m, but the maximum was 15.88 m. Nonetheless, for the 165 plots for which both survey and resource-grade positions were obtained, the effects on Random Forest regression performance were relatively slight. In terms of RRMSE, the maximum increases were found with trees/ha, and the greatest of these for any design was 1.3% (compared to RRMSE typically over 100% for this variable; see Table 3). The fractional contribution of conifers to relative density, with a maximum increase of 0.9%, and QMD, with a maximum increase of 0.6%, were the only other variables to show increases over 0.5% for any design; the maximum increase for live tree biomass, for any design, was only 0.1%. In practical terms, these differences are negligible. This result can perhaps be better understood by considering the displacement between the survey-grade and resource-grade points in terms of area of overlap of circles corresponding to the 10 m radius of the fixed plot design. Using the formula for the area of overlap of two circles of radius r, with centers displaced by distance d (e.g. Gove et al., 1999),

$$a = 2r^2\cos^{-1}\left(\frac{d}{2r}\right) - \frac{d}{2}\sqrt{4r^2 - d^2}$$

we find that the median overlap between survey-grade and resource-grade circles is 92.4%, while the minimum is 10.9%. The average overlap, over the 165 plots for which both coordinates are available, is 89.6%, and 95% of plots have an overlap of 72.7% or higher.

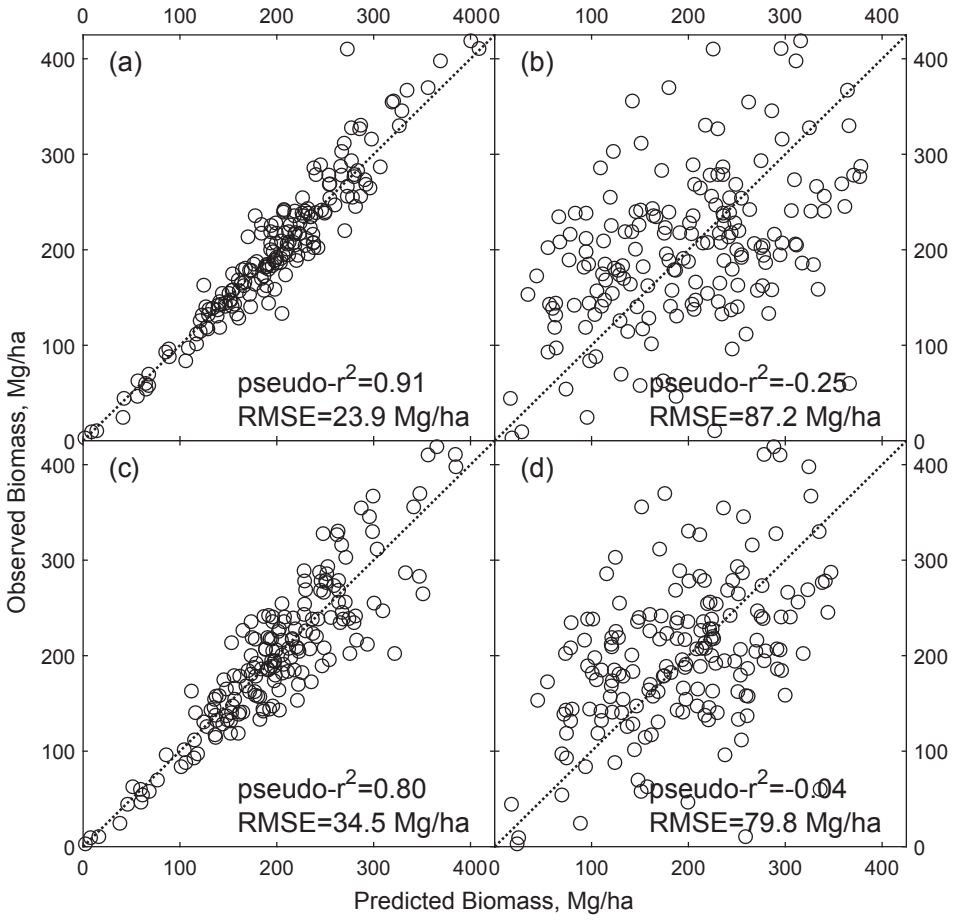

Fig. 3. Observed aboveground live biomass versus aboveground live biomass predicted using different elements of the Asner et al. (2011) approach: (a) model fit with LiDAR mean canopy height and plot-level field measured values of basal area and specific gravity; (b) model with LiDAR mean canopy height, LiDAR-predicted basal area, and plot level specific gravity; (c) model with LiDAR mean canopy height, field measured basal area, and average specific gravity; and (d) full implementation of Asner et al., universal model with LiDAR mean canopy height, LiDAR-predicted basal area and average specific gravity.

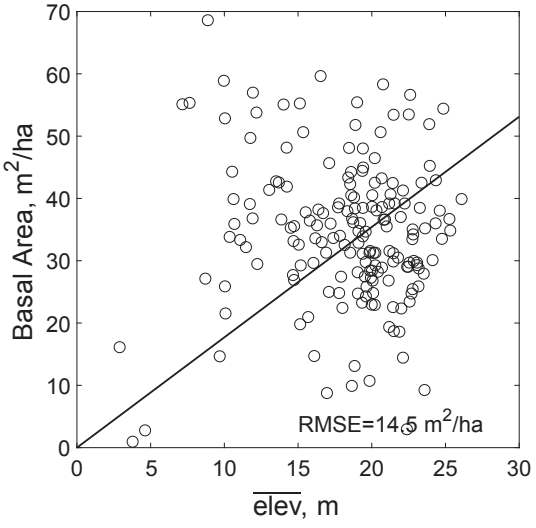

Fig. 4. Relationship between basal area and LiDAR mean canopy height.

Considering further that in this landscape, few plots are located near hard edges (such as abrupt transitions between unharvested stands and large gaps), the effects of displacement on LiDAR-derived metrics are relatively minor.

4. Discussion

The relatively mature condition of most forests in our study watershed likely influences the performance of LiDAR for biomass estimation, especially when that performance is assessed using metrics such as R^2 that are sensitive to the variance of the observed biomass values. The mean biomass of our sample plots (200 Mg/ha, Table 1) is substantially greater than regional means for New England (120 Mg/ha, Zheng et al., 2008), and corresponds closely to the mean for "mature" forests given by Keeton et al. (2011). For comparison, mean live-tree biomass of old-growth forests in the region has been reported to range from 220 to 270 Mg/ha (Keeton et al., 2011, Hoover et al., 2012, Gunn et al., 2014). Although portions of our study watershed are zoned for management by the U.S. Forest Service, much of the watershed is unmanaged and has seen no timber harvesting since the earliest decades of the 20 th century, with correspondingly greater recovery of biomass and other structural attributes than is typical of the region (Ducey et al., 2013). The spatial distribution of biomass over this portion of the watershed is likely controlled much more by environmental conditions than stand age, and does not reflect the distributional pattern that would be apparent in a landscape composed of a balanced age distribution of large, uniform harvest blocks (the idealized "regulated forest" of timber management; Bettinger et al., 2017).

The universal approach to estimating tree biomass described in Asner et al. (2011) was developed for use in tropical forests. Despite the differences in vegetation structure between tropical forests and our predominantly northern hardwood study site, the underlying relationship performed quite well when data from field plots were used; Fig. 3a compares well to the corresponding Fig. 6a in Asner et al. (2011). Fig. 3c, using field-measured BA but mean wood density, also compares well to the corresponding Fig. 6c in the Asner study. However, the approach fails to generate good predictions of biomass due to the lack of a relationship between LiDAR canopy height and BA; in Asner's study, R^2 values for the canopy height: BA relationship varied between 0.55 and 0.84 among the four ecoregions in the study (Fig. 3 in Asner

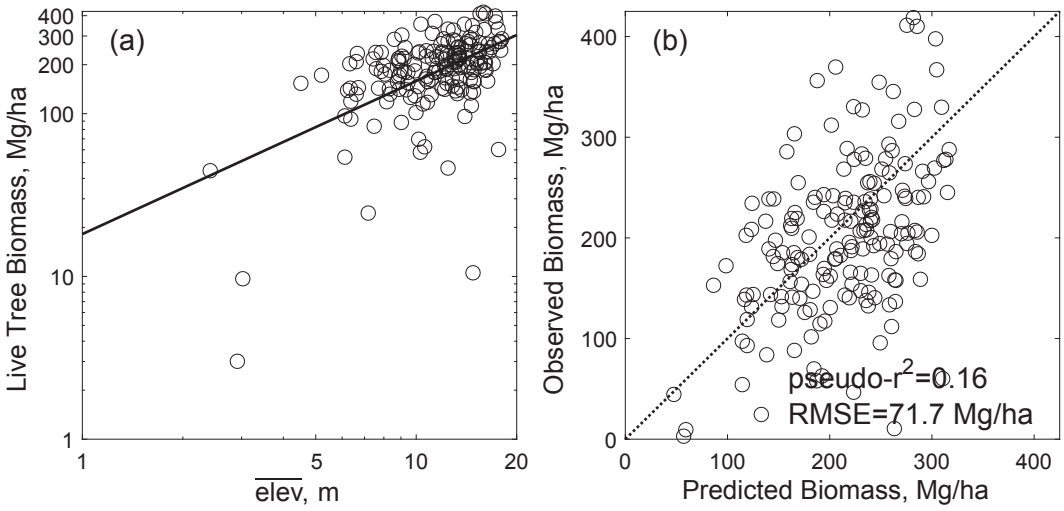

Fig. 5. Relationship between LiDAR mean canopy height and live aboveground biomass: (a) original regression; (b) observed versus predicted biomass (back transformed).

et al., 2011). However, Fig. 4 reveals the lack of a strong relationship between the two variables at our study site, resulting in a near total loss of predictive performance, compared to Asner et al.'s (2011) result of $R^2 = 0.80$ with an RMSE of 27.6 Mg/ha.

Much of the research on relationships between LiDAR variables and forest structural attributes in the United States has been conducted in western conifer-dominated forests, where relationships between LiDAR variables and structural attributes are generally quite strong (e.g. Hudak et al., 2006, Lefsky et al., 1999). Less work has been conducted in the mixed and deciduous forests of the Lake States and Northeast, and the outcomes have been variable. Lim et al. (2003) report strong relationships between LiDAR mean canopy height and basal area ($R^2 = 0.88$, residual SE = 0.39 for ln(BA)) in a forest dominated by sugar maple and yellow birch near Sault Ste. Marie, Ontario. Biomass predictions using mean canopy height were similarly strong, with an R^2 of 0.78 and residual SE = 0.55 for log-transformed biomass. However, those residual standard errors for log-transformed variables would correspond to relative prediction errors of approximately 54% and 73%, respectively. Applying those relative errors to the plot-level means for their study would yield prediction errors of 9.9 m^2/ha for basal area, and 104 Mg/ha for biomass. Anderson and Bolstad (2013) modeled biomass using LiDAR variables in several forest types in northern Wisconsin and achieved good fits in conifer and hardwood types ($R^2 = 0.74$ and 0.71; RMSE 37.6 and 42.8 Mg/ha, respectively), although model performance declined in mixed stands ($R^2 = 0.46$, RMSE = 48.1 Mg/ha). Deo et al. (2016) report similar results; using their best basal area factor and LiDAR extraction radius, they obtained a RMSE for stem volume of 35.2 m^3/ha across 47 plots in 6 stands. The mean plot volume in their study was 108.8 m^3/ha, implying a relative prediction error of approximately 32%. Given the close correlation between stem volume and AGB, one might expect similar results for biomass. Hawbaker et al. (2010) report results from a mixed hardwood forest in southern Wisconsin; a univariate model using height only explained 35% of the variability in basal area (RMSE = 6.02 m^2/ha as recalculated from their Table 2), while adding additional variables increased this value to 46%. Hawbaker et al. (2010) also found that height explained 43% of the variation in mean diameter (RMSE = 3.3 cm) but just 9% of variation in stem density (RMSE 112 stems/ha), while we report values of 31 and 46%, respectively (note that Hawbaker et al. (2010) predict mean diameter while we report quadratic mean diameter; LiDAR return density reported by Hawbaker et al. (2010) was also lower than that used in this study). Hayashi et al. (2014) studied the utility of LiDAR variables for predicting forest attributes in a Maine forest consisting of conifer-dominated and mixed hardwood-conifer stands; every model developed contained at least one height variable. The resulting QMD model had an R^2 of 0.489 and an RMSE of 3.68 cm; values for the BA model were 0.344 and 13.01 m^2/ha. Skowronski et al. (2007) used LiDAR data to predict forest structure in the pinelands of New Jersey and found that while biomass could be predicted reasonably well using LiDAR 80 th percentile height in pine-oak stands ($R^2 = 0.63$), model results for oak-pine and pine-scrub oak stands were less satisfactory ($R^2 = 0.282$ and 0.355, respectively). Skowronski et al. (2007) use a regression-through-the-origin model but do not report which of the (potentially conflicting) definitions of R^2 was

Table 2
R^2 values for RandomForest models of various stand characteristics, as a function of plot design (full or fixed) or basal area factor.

Variable	Full	Fixed	HPS Basal Area Factor(m^2/ha)							
			2.25	3.00	4.00	5.00	6.00	7.00	8.00	9.00
Trees/ha	0.465	0.473	0.115	0.157	0.155	0.070	0.058	0.080	0.046	0.097
Basal Area (m^2/ha)	0.415	0.341	0.416	0.359	0.297	0.195	0.175	0.187	0.211	0.177
QMD[a] (cm)	0.218	0.315	0.125	0.128	0.210	0.159	0.137	0.116	0.128	0.196
Relative Density	0.392	0.272	0.377	0.274	0.212	0.090	0.116	0.099	0.141	0.126
Lorey's Height (m)	0.729	0.711	0.730	0.712	0.673	0.659	0.624	0.608	0.588	0.596
Top Height (m)	0.472	0.585	0.457	0.509	0.529	0.525	0.499	0.526	0.506	0.530
Live Tree AG[b] Biomass (Mg/ha)	0.183	0.222	0.200	0.191	0.155	0.113	0.088	0.081	0.077	0.090
Conifer Fraction	0.531	0.518	0.520	0.513	0.563	0.518	0.529	0.506	0.433	0.448
Basal Area-weighted Wood SG[c]	0.447	0.424	0.314	0.301	0.321	0.270	0.207	0.223	0.359	0.296

[a] Quadratic mean diameter at breast height.
[b] Aboveground.
[c] Specific gravity.

Table 3
RMSE and relative RMSE (%, in parentheses) values for RandomForest models of various stand characteristics, as a function of plot design (full or fixed) or basal area factor.

	Full	Fixed	2.25	3.00	4.00	5.00	6.00	7.00	8.00	9.00
Trees/ha	1705	1685	2857	2923	3077	3253	3443	3345	3524	3116
	(70.3)	(68.7)	(119.8)	(122.0)	(128.7)	(140.8)	(150.8)	(147.1)	(155.0)	(145.4)
Basal Area (m^2/ha)	8.6	9.4	9.0	9.7	11.0	12.6	13.6	13.5	14.8	15.9
	(28.8)	(27.3)	(29.9)	(29.9)	(32.9)	(37.2)	(39.4)	(39.2)	(42.4)	(45.4)
QMD[a] (cm)	4.5	4.8	7.4	7.9	8.7	10.8	11.3	12.0	12.4	12.6
	(30.8)	(30.4)	(42.3)	(42.9)	(43.7)	(50.2)	(51.3)	(53.5)	(54.3)	(53.2)
Relative Density	0.24	0.25	0.27	0.30	0.34	0.39	0.39	0.41	0.44	0.46
	(28.4)	(26.6)	(31.3)	(32.9)	(36.5)	(42.1)	(42.0)	(43.3)	(46.5)	(48.7)
Lorey's Height (m)	2.5	2.9	2.6	2.8	3.3	3.5	3.9	4.0	4.3	4.4
	(15.0)	(16.9)	(15.5)	(16.3)	(19.1)	(19.8)	(22.0)	(22.7)	(24.4)	(25.0)
Top Height (m)	3.3	3.4	3.5	3.6	4.0	4.3	4.7	4.5	5.1	5.3
	(17.4)	(16.7)	(18.2)	(18.0)	(19.9)	(21.0)	(23.1)	(22.0)	(24.7)	(25.6)
Live Tree AG[b] Biomass (Mg/ha)	50.8	68.8	52.3	58.6	70.2	81.9	88.7	87.5	96.6	100.9
	(30.6)	(34.3)	(31.3)	(31.8)	(36.3)	(41.4)	(44.0)	(43.6)	(47.6)	(49.2)
Conifer Fraction	0.19	0.19	0.19	0.19	0.19	0.21	0.21	0.22	0.25	0.24
	(81.7)	(84.2)	(84.6)	(86.3)	(87.3)	(91.9)	(92.8)	(95.5)	(104.6)	(103.8)
Basal Area-weighted Wood SG[c]	0.06	0.06	0.08	0.08	0.09	0.09	0.11	0.10	0.10	0.11
	(11.2)	(11.5)	(14.2)	(14.3)	(16.9)	(17.9)	(20.1)	(19.9)	(19.5)	(21.4)

[a] Quadratic mean diameter (at breast height).
[b] Aboveground.
[c] Specific gravity.

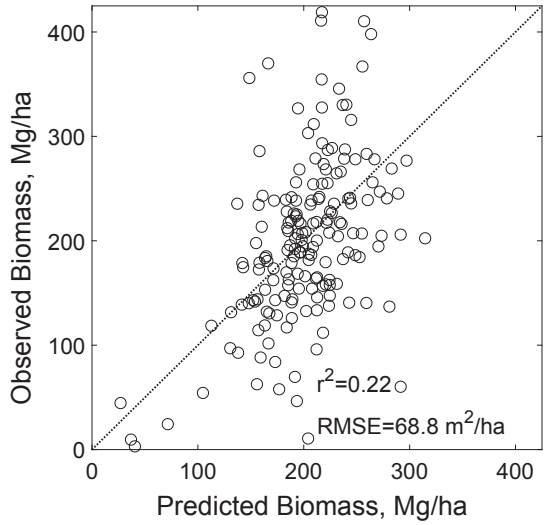

Fig. 6. Observed versus predicted live aboveground biomass values from Random Forest modeling approach.

used (Eisenhauer, 2003), so these values cannot necessarily be compared to those of other studies (especially those with intercepts in the regression) and RMSE cannot be recovered from the published results. They concluded that it is difficult to estimate tree biomass accurately from LiDAR variables when the relationship between height and BA is weak. They also note that in study areas where thinning had occurred, the relationship between height and biomass was the weakest, indicating that forest management activities may affect the ability to model overstory biomass from LiDAR variables.

Anderson et al. (2006) explored the ability of waveform LiDAR data to predict forest structure and biomass in a mixed conifer and deciduous forest in New Hampshire, similar to the composition of our study forest. They found that models developed using data from plots that matched the LiDAR footprint size, and were purposely selected to span a range of structural conditions, performed fairly well as judged by R^2 (R^2 = 0.54 and RMSE = 235.65 cm^2 for squared QMD, R^2 = 0.61 and RMSE = 58.0 Mg/ha for aboveground biomass). When these models were applied to data from a network of 409 systematically-collected inventory plots (0.1 ha) the R^2 values would suggest that results were poor (biomass, R^2 = 0.27, RMSE 56.1 Mg/ha; QMD, R^2 = 0.20, RMSE 3.28 cm), but note the change in dependent variable from QMD^2 which makes direct comparison of R^2 and RMSE impossible. Like Skowronski et al. (2007), they report that management treatments and stand composition affected model performance, as judged using R^2. However, it is notable that RMSE remains essentially unchanged, and when using LiDAR for mapping forest attributes, RMSE is a more direct and informative measure of model performance than R^2.

To understand why R^2 is an incomplete and potentially misleading indicator of fit (e.g. Barrett, 1974), consider a definition of R^2 that is applicable to most linear and nonlinear models that allow an intercept:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y}_i)^2}$$

where $\hat{y}_i$ is the model prediction. Note that the numerator is proportional to RMSE, while the denominator is proportional to the variance of the original observations. Thus, if a particular remote sensing method can predict a given forest attribute with a typical level of error (as described by RMSE), the value of R^2 will be sensitive to the underlying variance of the fitting data (or of the validation data if cross-validation is performed). As an example, the dramatic decline in R^2 for aboveground biomass reported by Anderson et al. (2006) between study-specific plots, and the systematic inventory plots, can easily be explained by the decline in the variance of biomass between the two sets of plots, as driven by plot selection (s^2 = 7815 for the former and s^2 = 4316 for the latter, based on their Table 1), rather than any similarly dramatic change in the relationship between the physical LiDAR retrieval and the plot design (since RMSE remains nearly constant). Likewise, the R^2 for aboveground biomass from our study is low compared to that reported in several of the studies cited above; but the RMSE is comparable to those reported for mixed stands. Were we to have inflated the variance of aboveground biomass in our data, either by oversampling the small number of patches of young, regenerating forest or by allocating more samples to krummholz near the altitudinal limit of woody vegetation in the watershed, our R^2 would be higher even though the accuracy of plot-level predictions (i.e. the RMSE) would not change appreciably.

With that said, RMSE for biomass and several other key stand descriptors remains disappointingly high in our study. (By comparison, Zheng et al., 2008 report an R^2 of 0.27 and RMSE of 46.5 Mg/ha for forests in New England using LANDSAT alone; one might have expected better performance from LiDAR.) It appears that developing satisfactory estimates of aboveground tree biomass using LiDAR metrics is often

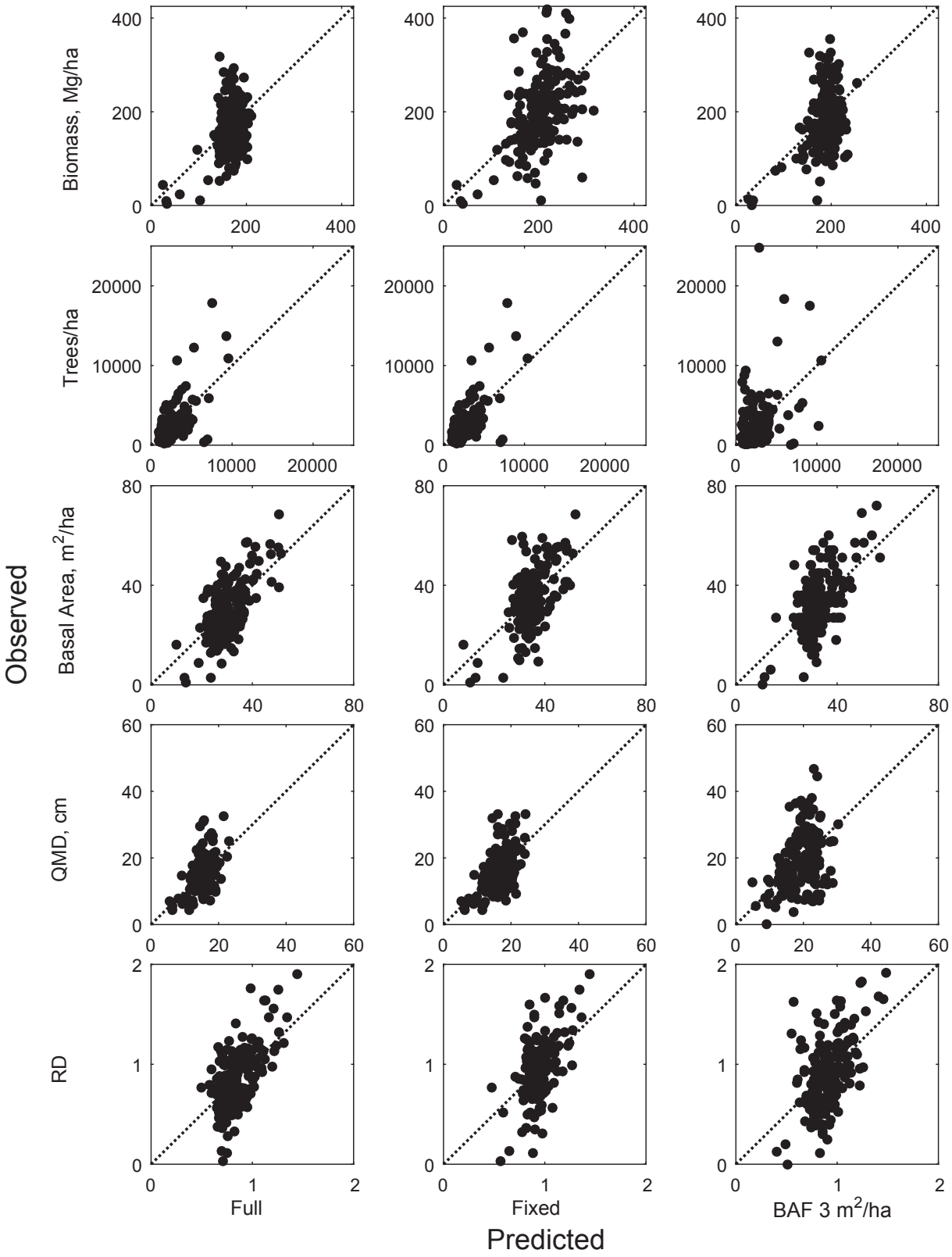

Fig. 7. Observed versus predicted values from Random Forest modeling for the full, fixed-radius, and BAF 3 m^2/ha plot designs for: (a) live aboveground biomass; (b) trees/ha; (c) basal area; (d) quadratic mean diameter; and (e) relative density.

difficult in northeastern deciduous or mixed forests, and that a major reason is the lack of a strong or predictable relationship between LiDAR canopy height metrics and plot level basal area. Our study, like others including Hayashi et al. (2014) and Hawbaker et al. (2010), utilized LiDAR data that were collected for another purpose. While LiDAR data collection designed for characterizing forest structural attributes would normally be collected during the leaf-on period, the data used in this study were collected during leaf-off conditions; this may have been a

factor in our inability to relate LiDAR canopy height to stand basal area. However, numerous investigations (including Hawbaker et al., 2010) have been conducted with leaf-off LiDAR data, and multiple studies explicitly address the performance of leaf-on and leaf-off LiDAR-derived variables in forest modeling. Anderson and Bolstad (2013) found small differences in model performance between leaf-on and leaf-off data, with leaf-on producing slightly better fits in some forest types, while the opposite was true in other types. They conclude that leaf-off data may be used to estimate biomass across a variety of Lake States forest types. Næsset (2005) examined the effects of using leaf-on and leaf-off data to estimate a variety of stand properties and reports that first return LiDAR data are less affected than last return data for most LiDAR metrics, and maximum canopy height is less sensitive to canopy conditions than mid-and low-canopy strata. In addition, the differences in LiDAR metrics between the two canopy conditions were greater in plots that were mainly deciduous. However, models predicting BA showed better fits when using leaf-off data (R^2 = 0.66 vs 0.62) although the errors were similar. The volume model also performed better with leaf-off data and showed a decline in error, with an RMSE of 0.29 m^3/ha for leaf-off vs. 0.35 m^3/ha using leaf-on data. Næsset (2005) concluded that in mixed forests, estimates of structural properties are not affected or are slightly improved when using leaf-off data. White et al. (2015) and Wasser et al. (2013) reached the same conclusion: leaf-off data produce acceptable models of various stand properties, with only small differences in performance from models constructed from leaf-on data. Both studies found that differences in performance were greater in deciduous forest types. It seems unlikely, then, that the use of leaf-off data was a major factor in the failure of either the Asner et al. (2011) or standard modeling approaches to predict stand biomass accurately in our study watershed.

Anderson et al. (2006, 2011) investigated the impact of various tree species on model performance in a forest located near, and similar to, our study site. They found that biomass and diameter model fits improved and error decreased when plots containing yellow birch or American beech (major species at our site) were excluded from the analysis; model performance also improved when white pine was present or red spruce represented more than 25% of the trees present. Restricting their analysis to plots dominated by sugar maple improved prediction of non-biomass attributes. Considering these results, and that many authors cited above report poorer model performance in areas dominated by deciduous trees, it is likely that the allometry of these species, including the frequent absence of a well-defined top, was a key factor in the lack of a distinct relationship between canopy height and basal area that resulted in our inability to adequately model aboveground biomass by either the Asner et al. (2011) or machine learning approaches.

In our study, plot design (fixed-radius plot, variable radius plot, or the full plot incorporating elements of both designs) did impact recovery of stand structural metrics from LiDAR, but the main influence appears not to have been so much the design itself as the relative size of the inclusion zones for dominant trees. It has often been assumed, based on reasonable geometric intuition, that the extraction area for LiDAR metrics needed to match the sample plot area as closely as possible, in terms of its area, shape, and location (e.g. Næsset, 2002). That intuition has been supported in simulation studies (e.g. Frazer et al., 2011), and has motivated the use of fixed-radius plot designs in conjunction with extremely precise georeferencing. However, evidence from the field suggests greater latitude in design choices. Hayashi et al. (2014, 2015) and Scrinzi et al. (2015) found that variable radius plots could serve as well as fixed-radius plots for LiDAR calibration. Hayashi et al. (2015) found the best agreement when LiDAR extraction radii were slightly larger than plot radii (for fixed-radius plots) or inclusion zone radii (for variable radius plots). Deo et al. (2016) report similar performance between fixed-radius plots and variable-radius plots for LiDAR calibration, with the best results at relatively low basal area factors (1.14–2.29 m^2/ha). Similarly, Tomppo et al. (2017) also found broad agreement between variable-radius plots and LiDAR, though the RMSE increased as BAF increased. In our study, errors also increased with increasing BAF, beginning noticeably at approximately the same BAF typically recommended for purely ground-based operational forest inventory in the region (4–5 m^2/ha; Wiant et al., 1984, Ducey 2001). Whether such a BAF is inherently too large for use with LiDAR data, or whether a better match between extraction area and inclusion area is needed, is unclear. The influence of extraction area on LiDAR calibration and mapping remains an active area of research (e.g. Deo et al., 2016, Hayashi et al., 2016), and general guidance on matching extraction areas to variable-radius plot designs has not yet emerged (Kirchhoefer et al., 2017).

5. Conclusions

1. A generalized approach for mapping biomass developed in tropical forests (Asner et al., 2011) did not perform well in our northern hardwood-dominated study area, due to a lack of a simple relationship between basal area and LiDAR height metrics. Simple regression of biomass on mean LiDAR elevation was more successful.
2. Using multiple LiDAR metrics within a Random Forest machine learning framework produced models of many structural variables that were disappointing in terms of R^2, but were not dissimilar to previously reported studies in terms of RMSE. This discrepancy can be understood in the context of the forest landscape characteristics of our study watershed. These models did outperform the Asner et al. (2011) and simple regression approach for biomass.
3. Moderate imprecision in georeferencing of plot centers had a negligible impact on model performance.
4. Fixed-radius plot, variable radius plot, and the hybrid full designs had similar performance across a range of stand structural metrics, though the performance of variable radius plots was best at basal area factors somewhat smaller than those usually recommended for operational inventory. Variable radius plots should not be ruled out for calibrating LiDAR relationships but additional research on best practices would be beneficial.
5. LiDAR did show sufficient relationship with biomass to support its use for stratification or as a covariate in sampling, but did not provide sufficient accuracy for high-resolution mapping of biomass. The use of alternative modeling approaches with higher-resolution LiDAR data, or fusion with other sources of remotely sensed data (e.g. to predict species composition), might improve the results.

Acknowledgements

This project was supported by the Northeastern States Research Cooperative through funding made available by the USDA Forest Service. Additional support came from agreement 11-CR-11092200-039, Terrestrial Ecological Units on the White Mountain National Forest, between the USDA Forest Service White Mountain National Forest and the University of New Hampshire. The conclusions and opinions in this paper are those of the authors and not of the NSRC, the Forest Service, or the USDA. The authors would like to thank Felicia Morrissette, Colleen Kenny, Nick Haskell, Sarah Janson, and Erica Roberts for field work, James E. Smith for constructing Fig. 1, and Ryan Hanavan for feedback on the manuscript.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automatic Control 19, 716–723.
Anderson, R.S., Bolstad, R.V., 2013. Estimating aboveground biomass and average annual wood biomass increment with airborne leaf-on and leaf-off LiDAR in Great Lakes forest types. North. J. Appl. For. 30 (1), 16–22.
Anderson, D.R., Burnham, K.P., Thompson, W.L., 2000. Null hypothesis testing: problems, prevalence, and an alternative. J. Wildl. Manage. 64 (4), 912–923.

Anderson, J., Martin, M.E., Smith, M.-L., Dubayah, R.O., Hofton, M.A., Hyde, P., Peterson, B.E., Blair, J.B., Knox, R.G., 2006. The use of waveform lidar to measure northern temperate mixed conifer and deciduous forest structure in New Hampshire. Remote Sens. Environ. 105, 248–261.
Anderson, J.E., Fast, A., Martin, M.E., Smith, M.-L., Plourde, L., Ducey, M.J., Lee, T.D., Dubayah, R.O., Hofton, M.A., Hyde, P., Peterson, B., Blair, J.B., 2011. Use of lidar and hyperspectral sensors to assess selected spatial, compositional, and structural patterns with recent and repeat disturbance and the abundance of sugar maple (*Acer saccharum* Marsh.) in a temperate mixed hardwood and conifer forest. J. Appl. Remote Sens. 5, 053504.
Asner, G.P., Mascaro, J., 2014. Mapping tropical forest carbon: calibrating plot estimates to a simple LiDAR metric. Remote Sens. Environ. 140, 614–624.
Asner, G.P., Mascaro, J., Muller-Landau, H.C., Vieilledent, G., Vaudry, R., Rasamoelina, M., Hall, J.S., van Breugel, M., 2011. A universal airborne LiDAR approach for tropical forest carbon mapping. Oecologia 168 (4), 1147–1160.
Barrett, J.P., 1974. The coefficient of determination—some limitations. Am. Stat. 28 (1), 19–20.
Baskerville, G.L., 1972. Use of logarithmic regression in the estimation of plant biomass. Can. J. For. Res. 2, 49–53.
Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67 (1), 1–48. http://dx.doi.org/10.18637/jss.v067.i01.
Bechtold, William A., Patterson, Paul L. (Eds.), 2005. The enhanced forest inventory and analysis program - national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC. 85 p.
Beets, P.N., Reutebuch, S., Kimberley, M.O., Oliver, G.R., Pearce, S.H., McGaughey, R.J., 2011. Leaf area index, biomass carbon and growth rate of radiata pine genetic types and relationships with LiDAR. Forests 2, 637–659.
Belgiu, M., Drăguț, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogram. Remote Sens. 114, 24–31.
Bettinger, P., Boston, K., Siry, J., Grebner, D., 2017. Forest Management and Planning, second ed. Academic Press, Cambridge, MA 362 pages.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer, New York.
Chojnacky, D.C., Heath, L.S., Jenkins, J.C., 2014. Updated generalized biomass equations for North American tree species. Forestry 87 (1), 129–151.
Claeskens, G., Hjort, N.L., 2008. Model Selection and Model Averaging. Cambridge University Press, Cambridge, UK.
Deo, R.K., Froese, R.E., Falkowski, M.J., Hudak, A.T., 2016. Optimizing variable radius plot size and lidar resolution to model standing volume in conifer forests. Can. J. Remote Sens. 42 (5), 428–442.
Ducey, M.J. 2001. Pre-cruise planning. In: Bennett, K. (Ed.), Workshop Proceedings: Forest Measurements for Natural Resource Professionals. II. Getting the Most from your Cruise, Caroline A. Fox Research and Demonstration Forest, Hillsborough, New Hampshire, October 19, 2001. pp. 26–35. U.N.H. Cooperative Extension, Durham, N. H. < https://extension.unh.edu/resources/files/Resource000398_Rep420.pdf > (last accessed 7/27/2017).
Ducey, M.J., 2012. Evergreenness and wood density predict height-diameter scaling in trees of the northeastern United States. For. Ecol. Manage. 279, 21–26.
Ducey, M.J., Knapp, R.A., 2010. A stand density index for complex mixed species forests in the northeastern United States. For. Ecol. Manage. 260, 1613–1622.
Ducey, M.J., Gunn, J.S., Whitman, A.A., 2013. Late-successional and old-growth forests in the northeastern United States: structure, dynamics, and prospects for restoration. Forests 4 (4), 1055–1086.
Eisenhauer, J.G., 2003. Regression through the origin. Teach. Stat. 25 (3), 76–80.
Frazer, G.W., Magnussen, S., Wulder, M.A., Niemann, K.O., 2011. Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. Remote Sens. Environ. 115 (2), 636–649.
Gove, J.H., Ringvall, A., Ståhl, G., Ducey, M.J., 1999. Point relascope sampling of downed coarse woody debris. Can. J. For. Res. 29, 1718–1726.
Gunn, J.S., Ducey, M.J., Whitman, A.A., 2014. Late-successional and old-growth forest carbon temporal dynamics in the northern forest (northeastern USA). For. Ecol. Manage. 312, 40–46.
Hawbaker, T.J., Gobakken, T., Lesak, A., Trømborg, E., Contrucci, K., Radeloff, V., 2010. Light detection and ranging-based measures of mixed hardwood forest structure. For. Sci. 56 (3), 313–326.
Hayashi, R., Weiskittel, A., Sader, S., 2014. Assessing the feasibility of low-density LiDAR for stand inventory attribute predictions in complex and managed forests of northern Maine, USA. Forests 5, 363–383.
Hayashi, R., Kershaw Jr., J.A., Weiskittel, A., 2015. Evaluation of alternative methods for using lidar to predict aboveground biomass in mixed species and structurally complex forests in northeastern North America. Math. Comput. Forest. Natural Resour. Sci. 7, 49–65.
Hayashi, R., Weiskittel, A., Kershaw Jr., J.A., 2016. Influence of prediction cell size on LiDAR-derived area-based estimates of total volume in mixed-species and multi-cohort forests in northeastern North America. Can. J. Remote Sens. 42, 473–488.
Hoover, C.M., Leak, W.B., Keel, B.G., 2012. Benchmark carbon stocks from old growth forests in northern New England, USA. For. Ecol. Manage. 266, 108–114.
Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M.S., Gessler, P.E., Morgan, P., 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. Can. J. Remote Sens. 32 (2), 126–138.
Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E., Falkowski, M.J., 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sens. Environ. 112, 2232–2245.
Jenkins, J.C., Chojnacky, D.C., Heath, L.S., Birdsey, R.A., 2003. National-scale biomass estimators for United States tree species. For. Sci. 49 (1), 12–35.
Keeton, W.S., Whitman, A.A., McGee, G.C., Goodale, C.L., 2011. Late-successional biomass development in northern hardwood-conifer forests of the northeastern United States. For. Sci. 57 (6), 489–505.
Kershaw Jr., J.A., Ducey, M.J., Beers, T.W., Husch, B., 2016. Forest Mensuration, fifth ed. Wiley, New York.
Kirchhoefer, M., Schumacher, J., Adler, P., Kändler, G., 2017. Considerations towards a novel approach for integrating angle-count sampling data in remote sensing based forest inventories. Forests 8 (7), 239. http://dx.doi.org/10.3390/f8070239.
Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A., Harding, D.J., 1999. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. Remote Sens. Environ. 70, 339–361.
Lefsky, M.A., Cohen, W.B., Harding, D.J., Parker, G.G., Acker, S.A., Gower, S.T., 2002. Lidar remote sensing of above-ground biomass in three biomes. Global Ecol. Biogeogr. 11, 393–399.
Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R News 2 (3), 18–22.
Lim, K., Treitz, P., Baldwin, K., Morrison, I., Green, J., 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. Can. J. Remote Sens. 29 (5), 658–678.
Marshall, D.D., Iles, K., Bell, J.F., 2004. Using a large-angle gauge to select trees for measurement in variable plot sampling. Can. J. For. Res. 34, 840–845.
McGaughey, R.J., 2014. FUSION/LDV: Software for LIDAR data analysis and visualization. USDA Forest Service, Pacific Northwest Research Station, Portland, OR. < http://forsys.cfr.washington.edu/fusion/fusionlatest.html > .
Miles, P.D., Smith, W.B., 2009. Specific gravity and other properties of wood and bark for 156 tree species found in North America. USDA For. Serv. Res. Note NRS-38, Newtown Square, PA.
Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Rem. Sens. Environ. 80, 88–99.
Næsset, E., 2005. Assessing sensor effects and effects of leaf-off and leaf-on canopy conditions on biophysical stand properties derived from small-footprint airborne laser data. Remote Sens. Env. 98, 356–370.
Nakagawa, S., Schielzeth, H., 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol. Evol. 4 (2), 133–142.
Pinheiro, J.C., Bates, D.M., 2000. Mixed-Effects Models in S and S-Plus. Springer, New York, pp. 528.
R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Robinson, A.P., Duursma, R.A., Marshall, J.D., 2005. A regression-based equivalence test for model validation: shifting the burden of proof. Tree Physiol. 25, 903–913.
Schumacher, F.X., Hall, F.D.S., 1933. Logarithmic expression of timber-tree volume. J. Agric. Res. 47, 719–734.
Scrinzi, G., Clementel, F., Floris, A., 2015. Angle count sampling reliability as ground truth for area-based LiDAR applications in forest inventories. Can. J. For. Res. 45 (4), 506–514.
Skowronski, N., Clark, K., Nelson, R., Hom, J., Patterson, M., 2007. Remotely sensed measurements of forest structure and fuel loads in the Pinelands of New Jersey. Remote Sens. Env. 108, 123–129.
Tomppo, E., Kuusinen, N., Mäkisara, K., Katila, M., McRoberts, R.E., 2017. Effects of field plot configurations on the uncertainties of ALS-assisted forest resource estimates. Scand. J. For. Res. 32 (6), 488–500.
USDA Forest Service, 2017. FIA Data and Tools. U.S. Department of Agriculture, Forest Service, Washington, DC. A website: < https://www.fia.fs.fed.us/tools-data/index.php > (accessed 8 August 2017).
Wasser, L., Day, R., Chasmer, L., Taylor, A., 2013. Influence of vegetation structure on lidar-derived canopy height and fractional cover in forested riparian buffers during leaf-off and leaf-on conditions. PLoS One 8 (1), e54776. http://dx.doi.org/10.1371/journal.pone.0054776.
White, J.C., Arenett, John T.T.R., Wulder, M.A., Tompalski, P., Coops, N.C., 2015. Evaluating the impact of leaf-on and leaf-off airborne laser scanning data on the estimation of forest inventory attributes with the area-based approach. Can. J. For. Res. 45, 1498–1513.
Wiant Jr., H.V., Yandle, D.O., Andreas, R., 1984. Is BAF 10 a good choice for point sampling? North. J. Appl. Forest. 2, 23–25.
Yu, X., Hyyppä, J., Vastaranta, M., Holopainen, M., Viitala, R., 2011. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. ISPRS J. Photogram. Remote Sens. 66, 28–37.
Zheng, D.L., Heath, L.S., Ducey, M.J., 2008. Spatial distribution of forest aboveground biomass estimated from remote sensing and forest inventory data in New England, USA. J. Appl. Remote Sens. 2, 021502.
Zolkos, S.G., Goetz, S.J., Dubayah, R., 2013. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. Remote Sens. Environ. 128, 289–298.