# Relating Forest Attributes with Area- and Tree-Based Light Detection and Ranging Metrics for Western Oregon

**Michael E. Goerndt, Vincente J. Monleon, and Hailemariam Temesgen**

ABSTRACT

Three sets of linear models were developed to predict several forest attributes, using stand-level and single-tree remote sensing (STRS) light detection and ranging (LiDAR) metrics as predictor variables. The first used only area-level metrics (ALM) associated with first-return height distribution, percentage of cover, and canopy transparency. The second alternative included metrics of first-return LiDAR intensity. The third alternative used area-level variables derived from STRS LiDAR metrics. The ALM model for Lorey's height did not change with inclusion of intensity and yielded the best results in terms of both model fit (adjusted $R^2 = 0.93$) and cross-validated relative root mean squared error (RRMSE = 8.1%). The ALM model for density (stems per hectare) had the poorest precision initially (RRMSE = 39.3%), but it improved dramatically (RRMSE = 27.2%) when intensity metrics were included. The resulting RRMSE values of the ALM models excluding intensity for basal area, quadratic mean diameter, cubic stem volume, and average crown width were 20.7, 19.9, 30.7, and 17.1%, respectively. The STRS model for Lorey's height showed a 3% improvement in RRMSE over the ALM models. The STRS basal area and density models significantly underperformed compared with the ALM models, with RRMSE values of 31.6 and 47.2%, respectively. The performance of STRS models for crown width, volume, and quadratic mean diameter was comparable to that of the ALM models.

**Keywords:** area-level metrics, single-tree remote sensing, LiDAR intensity, georeference

The use of light detection and ranging (LiDAR) technology to estimate forest attributes has advanced dramatically in recent years. Specifically, the use of small-footprint LiDAR metrics as explanatory variables has become prominent, as technology for the acquisition, processing, and extraction of LiDAR metrics has improved (Means et al. 2000, Næsset et al. 2004, Gobakken and Næsset 2004, Falkowski et al. 2006, Lim et al. 2008). Current LiDAR research is important not only for direct estimation of selected forest attributes but also for determining the strength of correlations between LiDAR metrics and forest variables for use in forest inventory and assessment. This is particularly relevant for regions such as western Oregon, where acquisition of small-footprint LiDAR data is becoming much more common and less expensive for both the state and federal forest ownerships.

There are two primary approaches for estimating or predicting forest variables using LiDAR metrics. The first approach relates LiDAR metrics to ground-measured variables for individual trees through single-tree remote sensing (STRS). A considerable amount of work has been done to develop methods to detect and delineate individual trees on a landscape using only geographic information systems (GIS) visualization, ranging from work done by Avery (1958) to that done by Korpela (2004). Typically, the two individual tree variables assessed directly from aerial LiDAR are total tree height and crown width (or crown area). The acquisition of these variables from raw LiDAR data requires the use of an algorithm to detect individual trees on a LiDAR canopy height model (CHM) either by identifying gradient changes in canopy height or by using variable window technology (Popescu et al. 2003a, Chen et al. 2006). Other tree-level variables, such as dbh and volume, are estimated from the LiDAR-derived individual tree height and crown width. Theoretically, inference for forest variables, such as volume and standing biomass, can be made at the stand level or even the regional level using STRS for LiDAR. However, this approach can oftentimes be difficult to implement on either a small scale or a large scale because of errors of omission and inclusion of individual trees. This form of error is generally caused by one of three factors: (1) LiDAR pixel size (precision of LiDAR surface), (2) forest structure (density and percentage of dominant/codominant trees), and (3) type of tree segmentation algorithm used (Popescu et al. 2003a, 2003b, Chen et al. 2006, Popescu 2007, Anderson 2009)

The second approach to the estimation and prediction of forest attributes using LiDAR metrics does not attempt to identify individual trees as is done with STRS, but uses area-level metrics from the LiDAR point cloud and area-level estimates from ground data.

Manuscript received June 3, 2009; accepted November 11, 2009.

*Michael E. Goerndt (michael.goerndt@oregonstate.edu), Department of Forest Engineering, Resources, and Management, College of Forestry, Oregon State University, 204 Peavy Hall, Corvallis, OR 97331. Vicente J. Monleon, Forest Inventory Analysis, US Forest Service, Pacific Northwest Research Station, Corvallis, OR 97331; Hailemariam Temesgen, Department of Forest Engineering, Resources, and Management, College of Forestry, Oregon State University, Corvallis, OR 97331.* *We are indebted to both Jim Kiser and Debra Johnson of the Oregon State University College of Forestry for the use of the equipment needed to perform the ground work for this project. We convey our sincerest gratitude to Matt Trappe for his superb forest inventory skills, which helped to expedite the ground sampling phase of the project. We also thank Dr. Michael Wing for planning and facilitating the collection of the LiDAR data used for this project.*

This article uses metric units; the applicable conversion factors are: centimeters (cm): 1 cm = 0.39 in.; meters (m): 1 m = 3.3 ft; square meters ($m^2$): 1 $m^2$ = 10.8 $ft^2$; cubic meters ($m^3$): 1 $m^3$ = 35.3 $ft^3$; hectares (ha): 1 ha = 2.47 ac.

From height profiling to volume estimation, this type of LiDAR-based inference has advanced dramatically over the last 15 years (Næsset 1997a, 1997b, Magnussen and Boudewyn 1998, Means et al. 1999, 2000, Næsset and Bjerknes 2001, Breidenbach et al. 2008). Næsset (1997b) and Magnussen and Boudewyn (1998) found correlation between forest attributes and LiDAR-based metrics. Traditionally, forest LiDAR studies have used only metrics associated with canopy elevations of individual returns based on spatial coordinates. However, LiDAR intensity has been found to be highly correlated with changes in forest cover, specifically with regard to species or species-group classification in forests (Luzum et al. 2004, Donoghue et al. 2007, Korpela 2008, Morsdorf et al. 2008). The incorporation of LiDAR intensity metrics in this study stems from two factors: (1) there has been some past difficulty correlating LiDAR height and cover metrics with basal area (BA) and density; and (2) the models created in this analysis incorporated a mixture of conifers and hardwoods, although conifers are dominant.

Development of regression models for estimating basal area and tree density using both LiDAR metrics and Landsat imagery was recently assessed by Hudak et al. (2006). The primary focus in that study was to assess differences in model fit and precision between a variety of different model-selection methods, such as subset and stepwise regression using height, cover, and intensity LiDAR metrics. This study differs in that the primary focus is on assessing the model fit and precision for six different forest attributes with and without the use of LiDAR intensity metrics, as well as comparing these with models created through STRS.

The objectives of this study are to (1) develop empirical models to relate forest attributes of interest and LiDAR metrics, (2) examine the contribution of LiDAR intensity metrics in estimating selected forest attributes, and (3) compare the resulting precision of the area-based models with precision of area-level estimates through use of STRS. The attributes of interest for this study are total stem volume ($m^3$ $ha^{-1}$), Lorey's mean stand height (m), quadratic mean diameter (QDBH) (cm), BA ($m^2$ $ha^{-1}$), density (live stems $ha^{-1}$), and average crown width (CW) (m) for all live trees.

## Methods

### Study Area

The study was conducted in McDonald-Dunn Research Forest, located just north-northwest of the town of Corvallis in western Oregon. The forest covers approximately 4,553 ha, with an elevation range of approximately 60–500 m above sea level. The main tree species are conifers, including Douglas fir (*Pseudotsuga menziesii*) and grand fir (*Abies grandis*) as apex species and a small percentage of western hemlock (*Tsuga heterophylla*) and western redcedar (*Thuja plicata*). The primary deciduous species is bigleaf maple (*Acer macrophyllum*). Although the species composition does not vary significantly throughout the forest, individual stands vary widely by age, density, and management history.

### Field Data

Using stand-level inventory data collected in 2008, 29 square plots measuring 30 × 30 m (900 $m^2$) were stratified by 16 subcategories representing the combination of four age ranges (20–40, 40–60, 60–80, and 80+ years) and four Curtis relative density ranges (0.01–0.2, 0.2–0.5, 0.5–0.8, 0.8–1.1). This made it possible to capture the range of diversity in forest structure. Relative density is defined as the product of the square root of QDBH and BA (Curtis 1970). The location of each plot was predetermined using ARCMAP, and all four corners of each plot were located in the field using waypoints in a differential global positioning systems (GPS). A criterion laser was used to verify the positions of the plot corners, using the southeast corner as a reference. To georeference exact plot position, the GPS receiver was mounted on a leveled stand at the location of the southeast corner to collect a minimum of 1,000 positions, with a logging interval of 1 second. Each plot was stem mapped using reference points that were also georeferenced with the GPS and a criterion laser. A summary of per-plot information for relevant attributes can be seen in Table 1.

**Table 1. Summary of selected forest attributes in McDonald-Dunn Forest ($n$ = 29).**

| | Minimum | Maximum | Mean | Standard deviation |
|---|---|---|---|---|
| Density (trees/ha) | 44.4 | 866.6 | 469.3 | 285.5 |
| QDBH (cm) | 23.2 | 83.1 | 49.9 | 15.1 |
| BA ($m^2$/ha) | 13.9 | 286.1 | 78.1 | 52.6 |
| CW (m) | 7.3 | 20.6 | 13.5 | 3.4 |
| Lorey's height (m) | 20.3 | 59.7 | 37.1 | 8.8 |
| Volume ($m^3$/ha) | 193.8 | 2422.1 | 1006.6 | 583.0 |

QDBH, quadratic mean diameter; BA, basal area; CW, crown width.

For every living tree within the plots that had a dbh of greater than 11.4 cm, dbh, species, height, height to crown base, and crown width were measured. Crown widths were determined by measuring two crown radii, one to the tip of the outmost branch on the longest side of the crown and another 90° (perpendicular) from the first. For excessively leaning bigleaf maple and Pacific madrone (*Arbutus menziesii*), CW was estimated using an equation from Hann (1997). Tree-level volume estimates were calculated using the US Forest Service National Volume Estimator Library (US Forest Service 2000). Lorey's mean stand height (Lorey's height) for each plot was calculated as plot-level means of basal area weighted tree heights (Husch et al. 2003).

### LiDAR Data

#### *Area-Based LiDAR Metrics*

The LiDAR data were collected in May of 2008 using a Leica ALS50 II laser system. The sensor scan angle was ±14° from nadir (the point on the ground directly below the aircraft), with a pulse rate designed to yield an average number of pulses of ≥8 points per square meter over terrestrial surfaces. Classification of ground and vegetation points was performed by TerraScan version 7.012, as well as spatial interpolation of ground classified points to create the digital terrain model. The data were collected using opposing parallel flight lines with a ≥50% overlap, producing average ground-point and first-return densities of 1.12 and 10.0 points per meter, respectively. All area-based LiDAR metrics used in this study were extracted from the raw point data using LiDAR FUSION (McGaughey 2008). A summary of the LiDAR metrics with corresponding descriptions can be seen in Table 2.

All LiDAR metrics were extracted using only first returns above a height of 3 m off the ground, with the exception of the cover metrics and canopy transparency, which used a variety of predetermined height thresholds, because a high number of first returns from the ground and low-lying vegetation may introduce confounding noise in the LiDAR metrics (Kraus and Pfeifer 1998, Næsset and Bjerknes 2001, Strunk 2008). Aside from raw reflectance values, there are other factors that can affect LiDAR intensity, such as humidity,

**Table 2. Summary of light detection and ranging (LiDAR) metrics computed from LiDAR FUSION.**

| Metric | Description |
|---|---|
| Height | Distribution of all first return heights >3 m |
| Percentile height (e.g., 5th, 10th, 20th, . . . , 95th) | Height distribution by deciles of first returns > 3 |
| Intensity | Distribution of all first return intensities >3 m |
| Percentile intensity | Intensity distribution by deciles of first returns >3 m |
| Canopy cover (Cover_3, Cover_6, . . . , Cover_24) | Percentage (0–100%) of first returns equal to or greater than a specified height (3, 6, . . . , 24 m) above the ground |
| Canopy transparencies | Percentage (0–100%) of first returns above a specified height after the removal of returns below a lower specified height |

weather patterns, ground-elevation changes, and scanning angle. In many LiDAR data sets, including the one for this study, the intensity values for each return have already undergone a form of normalization to adjust for some of these factors. The raw LiDAR intensities extracted in a particular scan were passed through a proprietary algorithm by the LiDAR vendor, accounting for several variables, such as localized trends in intensity values, scanning angle, and target distance. Through this process, many weak-intensity readings were adjusted to correspond with the distribution of intensity values in the surrounding area on the basis of the aforementioned variables, outputting a final product consisting of intensity values per return that were calibrated to an 8-bit value with a range of 0–255. Although this technique helps to filter out much of the backscatter noise, there are additional normalizations focusing on target distance and scan angle that have been studied (Luzum et al. 2004, Donoghue et al. 2007). Unfortunately, there was not sufficient information available for this LiDAR data set to perform this type of normalization.

### *STRS LiDAR Metrics*

Over the past several years, a few different types of algorithm have been developed to identify individual tree tops and crowns in LiDAR. For this study, individual tree crown identification was performed using a watershed segmentation algorithm for each plot. This was done by creating an inverse of the CHM and applying morphological watershed segmentation, resulting in segmentation of the CHM into polygons associated with individual tree crowns (Hyyppa et al. 2001, Anderson 2009). A tree height was defined as the highest return value located within the crown polygon. This is the most widely used method of detecting individual trees in conifer forests, as conifers typically have very distinguishable crowns, unlike many deciduous species (Persson et al. 2002, Popescu et al. 2003a, Anderson 2009). To classify as many tree crowns as possible, intensity values were assessed for each plot to aid in differentiating between conifers and hardwoods. This was made possible by the fact that the LiDAR data acquisition was leaf-off, causing deciduous crowns to typically have lower signal reflectance and therefore lower intensity. Because of the high density of some of the plots, many of the intermediate and suppressed trees were not detected by the algorithm. In most cases, these trees were completely overtopped by surrounding dominant and codominant trees. Of all ground-measured trees, about 60% were discretely detected in LiDAR. This coincides well with past studies in that trees assessed through STRS from LiDAR are typically indicative of dominant and codominant tree characteristics for a given stand, except in the case of low stand density (Popescu et al. 2003a, Anderson 2009). There was omission of only about 15% for dominant and codominant trees, most of which was caused by high canopy density in many of the plots. Individual tree variables extracted from the STRS of the LiDAR coverage were tree height and crown area (CA).

## Statistical Analysis

### *Estimating Tree-Level Attributes from STRS*

Prior to developing area-level models for forest attributes using tree-level variables from STRS, it was necessary to assess the correlation between LiDAR-derived values of single-tree height and CW, and ground measured values of height, CW, and dbh. dbh is a variable that cannot be directly obtained from LiDAR STRS, but it has been shown to be highly correlated with LiDAR tree height and CW for conifers (Popescu et al. 2003a, Popescu 2007). It was necessary to quantify the relationship between both LiDAR and ground-based tree height and CW because of tendencies in LiDAR tree height to have a slight negative bias and the fact that LiDAR CW is based on detected CA. Linear regression was applied to the 700 trees detected by the segmentation algorithm for height, CW, and dbh using LiDAR-derived height and CW as explanatory variables. No transformation was necessary for height, but log transformation was applied to both CW and dbh to correct for non-normality and heteroscedasticity. All three models were validated using leave-one-out cross-validation for root mean squared error (RMSE). The resulting tree-level estimates were used to calculate plot-level minimum, maximum, mean, and SD of tree height, CW, and dbh. QDBH and BA were also calculated on the basis of single-tree dbh of dominant and codominant trees detected through STRS per plot. All of the aforementioned STRS variables were used as explanatory variables in area-level models for all six forest attributes of interest.

### *Area-Level Metrics and STRS Models*

Before the process of linear model selection began, the ground attributes were assessed for several transformations. This was done using residual plots, Shapiro-Wilk tests, and quantile-quantile plots to test for normality of the data. It was determined that natural log transformation was the most beneficial for all of the forest attribute response variables. This coincided well with similar past studies, as this transformation is often used on forest variables to correct for both non-normality and heteroscedasticity (Næsset 1997a, 1997b, Næsset and Bjerknes 2001, Woods et al. 2008). Because the response variables were log transformed for the analysis, the predicted values were bias-corrected using a correction factor of 0.5 times the mean squared error before back transformation (Baskerville 1972, Woods et al. 2008).

For all three sets of models, the actual model selection was performed using a subset regression technique that identifies the independent variables that create the best fitting linear regression models according to Bayesian information criteria (BIC) using an exhaustive search. This was performed using the *regsubsets( )* tool available

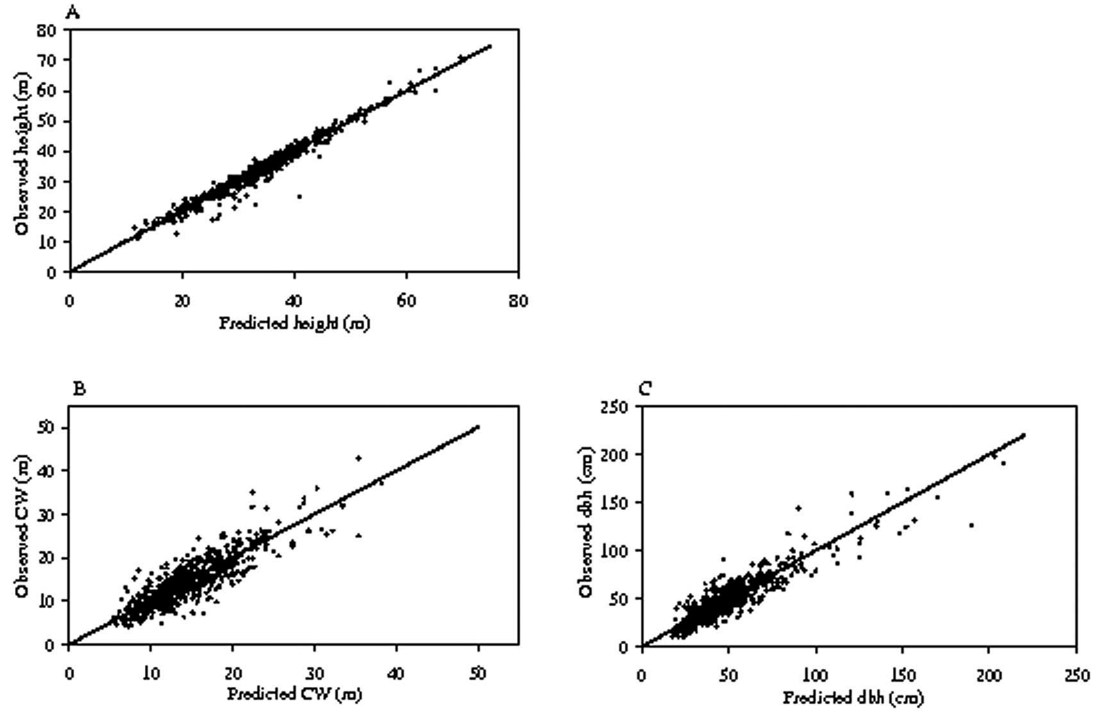


**Figure 1. A, Observed versus predicted height. B, Observed versus predicted crown width (CW). C, Observed versus predicted dbh.**

in the *leaps( )* package for *R* (R Development Core Team 2008, Lumley 2008). For each response variable, the program was set up to output the best 10 models for one to seven variables allowed per model using the BIC. The resulting output contained the information for a total of 70 models.

Several factors needed to be accounted for in determining the final models, including heteroscedasticity and multicollinearity. To expedite this process, an algorithm was created to automatically rank the complete list of models output by *regsubsets*( ) on the basis of BIC. At the same time, this algorithm calculated the variance inflation factors (VIF) for the predictors in each model, performed both a Breusch-Pagan test and a modified Levene's test for heteroscedasticity, and calculated both nontransformed RMSE and adjusted $R^2$ for each model. Any model that had a VIF score greater than 9.5 was automatically dropped from the final list. Validation for each selected model was performed using leave-one-out cross-validation for nontransformed RMSE, as was done by Næsset and Bjerknes (2001), Guillemette et al. (2008), Wulder et al. (2008), and Woods et al. (2008). RMSE and relative root mean squared error (RRMSE) were calculated for each model in original scale as

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n}},$$

where $Y_i$ is the observed value for plot $i$, $\hat{Y}_i$ is the predicted value for plot $i$, $n$ is the sample size, and RRMSE is RMSE divided by the mean of the observed values.

## Results and Discussion

### Tree-Level Models

The linear regression models created to estimate tree-level attributes from LiDAR all included height and CW as significant variables. The cross-validated RMSE and RRMSE values for height, CW, and dbh were 1.87 (5.5%), 2.97 (20.6%), and 11.67 (24%), respectively. The RRMSE for CW was slightly higher than expected based on results from past studies, such as those of Popescu et al. (2003a). This was most likely caused by the high density of some of the sampled stands, leading to underestimation of CA for some dominant and codominant trees. Figure 1 shows LiDAR field-measured values versus LiDAR predicted values for height, CW, and dbh.

### Area-Level Models

The final area-level metrics (ALM) models excluding LiDAR intensity used at least one variable from each of the other types of LiDAR metric, including return heights, cover, and canopy transparency. Each of the ALM models that included intensity metrics recognized at least one intensity metric as a significant variable, with the exception of Lorey's height, which was not altered by the inclusion of intensity metrics. The final STRS models used at least one variable from each type of area-level variable created by the single-tree analysis. With the exception of some intercept values, all of the metrics included in the final models were significant at least at the 0.05 level. Within the set excluding intensity, the original model for BA yielded an RRMSE of 55%, which is uncharacteristically high even for this particular attribute, which can have RRMSE values in a range of 25–30% (Drake et al. 2002, Holmgren 2003, Woods et al. 2008). Residual plots and Cook's distance suggested that one of the plot values was a significantly high outlier for BA. Closer examination of this plot provided evidence that many of its characteristics were not consistent with other plots with similar density range and age category. A combination of extremely dense dominant and codominant canopy, zero mortality, and a complete lack of both intermediate stems and hardwoods contributed to a severely inflated BA estimate for the plot. Because of the statistical and logistic implications of this, the plot was dropped from the analysis of BA. Significant predictor information and performance statistics for

**Table 3. Regression coefficients for area-level metrics models excluding intensity, with performance statistics in original scale.**

| Coefficient | ln(BA), ln ($m^2$ $ha^{-1}$) | ln(Lorey's), ln (m) | ln(Vol), ln ($m^3$ $ha^{-1}$) | ln(Density), ln (stems $ha^{-1}$) | ln(QDBH), ln (cm) | ln(CW), ln (m) |
|---|---|---|---|---|---|---|
| Intercept | 1.28 (0.27) | 2.56 (0.055) | 7.19 (1.89) | −7.24 (2.92) | 6.22 (1.19) | 1 (1.62) |
| HtMin | | | | 1.51 (0.73) | −0.827 (0.393) | −0.706 (0.335) |
| HtMax | | | 0.032 (0.005) | | | |
| HtMedian | | | | | 0.018 (0.006) | |
| HtKurtosis | | 0.017 (0.005) | | | | |
| HtInterqDist | 0.051 (0.011) | | | | | |
| Htp05 | 0.019 (0.008) | | | | | |
| Htp10 | | | | −0.035 (0.012) | | |
| Htp25 | | 0.024 (0.001) | | | | |
| Cover_3 | 0.025 (0.003) | | 0.026 (0.003) | | | |
| Cover_15 | | | | 0.073 (0.007) | −0.017 (0.004) | −0.023 (0.003) |
| Cover_21 | | | | −0.039 (0.006) | 0.014 (0.004) | 0.019 (0.002) |
| Trans3_12 | | | | | | −0.035 (0.014) |
| Trans6_12 | | | −0.04 (0.019) | −0.048 (0.021) | | 0.077 (0.027) |
| $R^2$(adj), % | 82.3 | 93 | 82.6 | 87.7 | 74.7 | 73 |
| $RMSE_{cross}$ | 14.7 | 2.99 | 309.7 | 184.5 | 9.95 | 2.3 |
| $RRMSE_{cross}$ | 20.8 | 8.1 | 30.7 | 39.3 | 19.9 | 17.1 |

Standard errors of coefficients are given in parentheses.
BA, basal area; Vol, volume; QDBH, quadratic mean diameter; CW, crown width; Ht, height; Min, minimum; Max, maximum; Trans, canopy transparency; $RMSE_{cross}$, cross-validated root mean squared error; $RRMSE_{cross}$, cross-validated relative root mean squared error.

**Table 4. Regression coefficients for area-level metrics models including intensity, with performance statistics in original scale.**

| Coefficient | ln(BA), ln ($m^2$ $ha^{-1}$) | ln(Lorey's), ln (m) | ln(Vol), ln ($m^3$ $ha^{-1}$) | ln(Density), ln (stems $ha^{-1}$) | ln(QDBH), ln (cm) | ln(CW), ln (m) |
|---|---|---|---|---|---|---|
| Intercept | −0.53 (0.422) | 2.56 (0.055) | 3.72 (1.61) | −1.31 (1.03) | 8.07 (1.12) | 5.22 (0.397) |
| HtMin | | | | | −0.875 (0.336) | |
| HtMax | | | | −0.013 (0.005) | 0.012 (0.004) | |
| HtKurt | | 0.017 (0.005) | | | | |
| Htp25 | | 0.024 (0.001) | | | | |
| Htp75 | 0.031 (0.005) | | 0.036 (0.005) | | | |
| Cover_3 | | | | | −0.007 (0.002) | −0.01 (0.002) |
| Cover_9 | | | | 0.036 (0.003) | | |
| Cover_15 | 0.023 (0.002) | | | | | |
| Cover_18 | | | 0.025 (0.002) | | | |
| Cover_21 | | | | | | 0.009 (0.001) |
| Cover_24 | | | | | 0.008 (0.002) | |
| Trans3_9 | | | 0.049 (0.021) | | | |
| Trans6_12 | | | −0.069 (0.024) | | | |
| Intp75 | 0.035 (0.004) | | 0.042 (0.005) | 0.036 (0.006) | | |
| IntMedian | −0.034 (0.004) | | −0.042 (0.005) | −0.035 (0.005) | | |
| IntMax | | | | 0.015 (0.005) | −0.008 (0.002) | −0.01 (0.002) |
| $R^2$(adj), % | 89.7 | 93 | 90.4 | 91.9 | 82.1 | 79.9 |
| $RMSE_{cross}$ | 15.9 | 2.99 | 349.7 | 127.8 | 8.17 | 1.89 |
| $RRMSE_{cross}$ | 22.6 | 8.1 | 34.7 | 27.2 | 16.4 | 14 |

Standard errors of coefficients are given in parentheses.
BA, basal area; Vol, volume; QDBH, quadratic mean diameter; CW, crown width; Ht, height; Min, minimum; Max, maximum; Trans, canopy transparency; Int, intensity; $RMSE_{cross}$, cross-validated root mean squared error; $RRMSE_{cross}$, cross-validated relative root mean squared error.

ALM models excluding intensity, ALM models including intensity, and STRS models are given in Tables 3, 4, and 5, respectively.

### Lorey's Height

The ALM model for Lorey's height, which remained unchanged with the inclusion of intensity metrics, ranked highest in the models for both ALM sets, with an adjusted $R^2$ value of 0.93. In addition, this model performed the best in terms of precision of prediction, with an RRMSE value of 8.1%. This result is within the range indicated by several other studies using either maximum height or mean height, including Næsset and Bjerkness (2001), Holmgren and Jonsson (2004), and Woods et al. (2008). Not surprisingly, the model for Lorey's height also ranked highest for the STRS models, with an increase of 1% for adjusted $R^2$ and a decrease of almost 4% for RRMSE. This is indicative of the high correlation between Lorey's height and mean and maximum heights of dominant and codominant stems, as is perceived by LiDAR.

### Stocking (BA and Density)

Although the adjusted $R^2$ value of the ALM predictive model for BA increased from 0.82 to 0.90 with the inclusion of intensity metrics, the RRMSE actually increased from 20.7 to 22.6%. Both of these RRMSE values are reasonable with regard to this study and the aforementioned previous studies, which demonstrates how the improvement in model fit introduced by intensity metrics did not really improve the precision of prediction for BA. Although there is insufficient evidence to make an assertion, this result may imply that the slight transition between conifers and hardwoods observed in the plots does not have a significant impact on the prediction of BA. The STRS model for BA was inferior to both ALM models with regard to both adjusted $R^2$ (80.1) and RRMSE (31.6%).

The ALM predictive model for density experienced a similar increase in adjusted $R^2$ to the model for BA. However, the decrease in RRMSE from 39.3 to 27.2% observed when intensity metrics were included was extremely significant for this attribute. Given that

Table 5. Regression coefficients for single-tree remote sensing models with performance statistics in original scale.

| Coefficient | ln(BA), ln ($m^2$ $ha^{-1}$) | ln(Lorey's), ln (m) | ln(Vol), ln ($m^3$ $ha^{-1}$) | ln(Density), ln (stems $ha^{-1}$) | ln(QDBH), ln (cm) | ln(CW), ln (m) |
|---|---|---|---|---|---|---|
| Intercept | 4.14 (0.297) | 2.4 (0.053) | 5.85 (0.159) | 7.69 (0.239) | 2.38 (0.173) | 2.23 (0.139) |
| Density | | | | | | −0.0007 (0.0002) |
| Max Ht | | 0.009 (0.002) | | | 0.022 (0.003) | |
| Mean Ht | −0.021 (0.009) | 0.022 (0.002) | | | | |
| StdHt | 0.046 (0.014) | | 0.054 (0.012) | | | |
| MinCW | | | | −0.117 (0.017) | 0.055 (0.009) | |
| StdCW | −0.03 (0.029) | 0.021 (0.0004) | −0.144 (0.027) | | | |
| Min dbh | | | −0.006 (0.004) | | | 0.009 (0.002) |
| Max dbh | | | | | | |
| Mean dbh | | | | −0.031 (0.004) | | 0.004 (0.001) |
| Std dbh | | | | −0.025 (0.005) | | |
| BA | 0.21 (0.024) | | 0.244 (0.019) | 0.257 (0.025) | | |
| $R^2$(adj), % | 80.1 | 96.4 | 86.4 | 87.7 | 73 | 78.6 |
| $RMSE_{cross}$ | 22.93 | 2.06 | 313.27 | 221.52 | 8.56 | 1.77 |
| $RRMSE_{cross}$ | 31.6 | 4.3 | 30.1 | 47.2 | 17.2 | 13.1 |

Standard errors of the coefficients are given in parentheses.
BA, basal area; Vol, volume; QDBH, quadratic mean diameter; CW, crown width; Ht, height; Min, minimum; Max, maximum; $RMSE_{cross}$, cross-validated root mean squared error; $RRMSE_{cross}$, cross-validated relative root mean squared error.

the 39.3% RRMSE value is consistent with similar past studies, such as those of Holmgren (2003) and Woods et al. (2008), it is unlikely that this difference is caused by a better classification of species composition brought on by the intensity metrics. It was more likely caused by the ability of the intensity metrics to describe overall canopy density, which can be highly correlated with the number of tree crowns (and hence tree stems) present in the area of interest. As with BA, the STRS predictive model was greatly inferior to both of the ALM models. This indicates that the profile of dominant and codominant tree composition provided by STRS in LiDAR was not able to provide sufficient information for predicting density with a high degree of precision. If the area-level analysis was based solely on dominant and codominant trees, it is very likely that the STRS model would outperform the ALM models because of a high correlation between the number of STRS-detected trees and the number of dominant and codominant trees per plot.

### Volume

Similar to BA, the ALM predictive model for cubic stem volume increased by 7% in adjusted $R^2$ but saw an increase in RRMSE from 30.7 to 34.7% when intensity metrics were introduced. This implies that although the model fit for volume improved with the inclusion of intensity metrics, the original model is better in terms of precision. The performance of the STRS predictive model for volume was almost a midpoint between the performances of the ALM volume models. The $R^2$ value for the STRS volume model was lower than that for the ALM model including intensity and higher than that for the ALM model excluding intensity and visa versa for RRMSE. It could be said that as a whole, the STRS volume model performed just as well as the ALM volume models.

### QDBH

The RRMSE values of 19.1 and 16.4% for the models excluding intensity and including intensity, respectively, for QDBH were both encouraging. The decrease in RRMSE when the intensity metrics were included was less than 3%, implying that inclusion of intensity metrics does not significantly improve precision of prediction for QDBH. These results combined, with the results from the BA model, provide evidence that area-level LiDAR metrics can be used effectively to predict forest attributes beyond the realm of simple height profiling. As with volume, the performance of the STRS model for QDBH was comparable to that of the ALM models, with an $R^2$ value of 73% and RRMSE of 17.2%.

### Crown Width

Crown width (CW) is an attribute that has traditionally been reserved for single-tree LiDAR analyses, as it represents one of the few variables that can be delineated for individual tree crowns using aerial LiDAR. Assessment of this attribute stemmed from a desire to assess the predictability of an important factor in the derivation of other forest attributes, such as crown competition factors. The results of both ALM models for CW in terms of RRMSE were extremely encouraging, with values of 17.1 and 14%, respectively. This indicates that the precision of prediction for CW, at least in the context of this study, was just as good as the precision observed for QDBH. The STRS model showed only a slight improvement in precision, with an RRMSE of 13.1%.

## Conclusion

The effective use of LiDAR for describing and predicting forest attributes has seen dramatic advancement over the last decade, for both single-tree and area-based inference. The analyses that have been described in this study are only a small part of the overall potential for forest attribute prediction using area-based LiDAR metrics. This study has provided valuable insights regarding the strength of correlation between selected forest attributes and LiDAR metrics.

Including LiDAR intensity metrics generally improved model fit for the forest attributes of interest. In the case of density, this coincided with a dramatic improvement in the precision of prediction for the model. However, it was also discovered that for many of the forest attributes, the use of intensity metrics either only slightly improved precision or slightly decreased the precision. This implies that for some forest attributes, such as BA and volume, the inclusion of intensity metrics is not required to improve precision.

Model fit and precision were generally not higher for the STRS models except in the case of Lorey's height. In the case of BA and density, the STRS models did not perform as well as the ALM models. However, it was observed that both model fit and precision were comparable between the STRS models and ALM models for volume, QDBH, and CW. This, combined with the result from the

Lorey's height model, supports the assertion that attributes that are highly correlated with vertical canopy structure can be estimated with fairly high precision using both area-level and tree-level LiDAR metrics. This study has shown that precise predictions of stand-level forest attributes, such as BA, volume, density, QDBH, Lorey's height, and CW, can be made across a wide range of forest compositions found in western Oregon.

## Literature Cited


ANDERSON, H. 2009. Using airborne light detection and ranging (LiDAR) to characterize forest stand condition on the Kenai Peninsula of Alaska. *West. J. Appl. For.* 24(2):95–102.

AVERY, G. 1958. Helicopter stereo-photography of forest plots. *Photogramm. Eng. Remote Sens.* 24:617–624.

BASKERVILLE, G.L. 1972. Use of logarithmic regression in the estimation of plant biomass. *Can. J. For. Res.* 2:49–53.

BREIDENBACH, J., C. GLASER, AND M. SCHMIDT. 2008. Estimation of diameter distributions by means of airborne laser scanner data. *Can. J. For. Res.* 38: 1611–1620.

CHEN, Q., D. BALDOCCHI, P. GONG, AND K. MAGGI. 2006. Isolating individual trees in a savanna woodland using small footprint LIDAR data. *Photogramm. Eng. Remote Sens.* 72(8):923–932.

CURTIS, R.O. 1970. Stand density measures: An interpretation. *For. Sci.* 16(4): 403–414.

DRAKE, J.B., R.O. DUBAYAH, D.B. CLARK, R.G. KNOX, J.B. BLAIR, M.A. HOFTON, R.L. CHAZDON, J.F. WEISHAMPEL, AND S.D. PRINCE. 2002. Estimation of tropical forest structural characteristics using large-footprint LiDAR. *Remote Sens. Environ.* 79(2):305–319.

DONOGHUE, D.N.M., P.J. WATT, N.J. COX, AND J. WILSON. 2007. Remote sensing of species mixtures in conifer plantations using LiDAR height and intensity data. *Remote Sens. Environ.* 110:509–522.

FALKOWSKI, M.J., A.M.S. SMITH, A.T. HUDAK, P.E. GESSLER, L.A. VIERLING, AND N.L. CROOKSTON. 2006. Automated estimation of individual conifer tree height and crown diameter via two-dimensional spatial wavelet analysis of LiDAR data. *Can. J. Remote Sens.* 32:153–161.

GOBAKKEN, T., AND E. NÆSSET. 2004. Estimation of diameter and basal area distributions in coniferous forests by means of airborne laser scanner data. *Scand. J. For. Res.* 19(6):529–542.

GUILLEMETTE, F., S. BÉDARD, AND M. FORTIN. 2008. Evaluation of a tree classification system in relation to mortality risk in Quebec northern hardwoods. *For. Chron.* 84(6):886–899.

HANN, D.W. 1997. *Equations for predicting the largest crown width of stand-grown trees in western Oregon.* Research Contribution 17. Forest Research Laboratory, Oregon State University, Corvallis, Oregon. 14 p.

HOLMGREN, J. 2003. *Estimation of forest variables using airborne laser scanning.* PhD dissertation, Swedish Univ. of Agricultural Sciences, Uppsala, Sweden. 41 p.

HOLMGREN, J., AND T. JONSSON. 2004. Large scale airborne laser scanning of forest resources in Sweden. *Int. Arch. Photogram. Remote Sens. Spatial Inf. Sci.* 36(8/W 2):157–160.

HUDAK, A.T., N.L. CROOKSTON, J.S. EVANS, M.K. FALKOWSKI, A.M.S. SMITH, P.E. GESSLER, AND P. MORGAN. 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return LiDAR and multispectral data. *Can. J. Remote Sens.* 32(2):126–138.

HUSCH, B., T.W. BEERS, AND J.A. KERSHAW, JR. 2003. *Forest mensuration.* John Wiley and Sons Inc. Hoboken, NJ.

HYYPPA, J., O. KELLE, M. LEHKIKOINEN, AND M. INKINEN. 2001. A segmentation-based method to retrieve stem volume estimates from 3-D height models produced by laser scanners. *IEEE Trans. Geosci. Remote Sens.* 39(5): 969–975.

KORPELA, I. 2004. Individual tree measurements by means of digital aerial photogrammetry. *Silva Fenn.* 3:1–93.

KORPELA, I. 2008. Mapping of understory lichens with airborne discrete-return LiDAR data. *Remote Sens. Environ.* 112(10):3891–3897.

KRAUS, K., AND N. PFEIFER. 1998. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS J. Photogramm. Remote Sens.* 53(4): 193–203.

LIM, K., C. HOPKINSON, AND P. TREITZ. 2008. Examining the effects of sampling point densities on laser canopy height and density metrics. *For. Chron.* 84(6):876–885.

LUMLEY, T. 2008. *Leaps: Regression subset selection.* R package, version 2.7.

LUZUM, B., M. STAREK, AND K.C. SLATTON. 2004. *Normalizing ALSM intensities.* GEM Center Rep. No. Rep_2004-07-001. University of Florida, Gainesville, FL.

MAGNUSSEN, S., AND P. BOUDEWYN. 1998. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Can. J. For. Res.* 28:1016–1031.

MCGAUGHEY, R. 2008. *FUSION/LDV: Software for LiDAR data analysis and visualization, FUSION Version 2.65.* Available online at forsys.cfr.washington.edu/fusion.html; last accessed Feb. 10, 2009.

MEANS, J.A., S.A. ACKER, D.J. HARDING, J.B. BLAIR, M.A. LEFSKY, W.B. COHEN, M.E. HARMON, AND W.A. MCKEE. 1999. Use of large footprint scanning airborne LiDAR to estimate forest stand characteristics in the western Cascades of Oregon. *Remote Sens. Environ.* 67:298–308.

MEANS, J.E., S.A. ACKER, B.J. FITT, M. RENSLOW, L. EMERSON, AND C.J. HENDRIX. 2000. Predicting forest stand characteristics with airborne scanning LiDAR. *Photogramm. Eng. Remote Sens.* 66(11):1367–1371.

MORSDORF, F., O. FREY, E. MEIER, K. ITTEN, AND B. ALLGÖWER. 2008. Assessment of the influence of flying height and scan angle on biophysical vegetation products derived from airborne laser scanning. *Int. J. Remote Sens.* 29(5): 1387–1406.

NÆSSET, E. 1997a. Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS J. Photogramm. Remote Sens.* 52(2):49–56.

NÆSSET, E. 1997b. Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sens. Environ.* 61(2):246–253.

NÆSSET, E., AND K.O. BJERKNES. 2001. Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. *Remote. Sens. Environ.* 78(3):328–340.

NÆSSET, E., T. GOBAKKEN, J. HOLMGREN, H. HYYPPA, J. HYYPPA, M. MALTAMO, M. NILSSON, H. OLSSON, A. PERSSON, AND U. SODERMAN. 2004. Laser scanning of forest resources: The Nordic experience. *Scand. J. of For. Res.* 19(6):482–499.

PERSSON, A., J. HOLMGREN, AND U. SODERMANN. 2002. Detecting and measuring individual trees using an airborne laser scanner, Photogramm. *Eng. Remote Sens.* 68:925–932.

POPESCU, S.C., R.H. WYNNE, AND R.E. NELSON. 2003a. Measuring individual tree crown diameter with LiDAR and assessing its influence on estimating forest volume and biomass. *Can. J. Remote Sens.* 29(5):564–577.

POPESCU, S.C., P.J. RADTKE, AND R.H. WYNNE. 2003b. Forest measurements with airborne and ground-based laser scanning. *Geospatial Solutions* 13(8):18.

POPESCU, S.C. 2007. Estimating biomass of individual pine trees using airborne LiDAR. *Biomass Bioenergy* 31:646–655.

R DEVELOPMENT CORE TEAM. 2008. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

STRUNK, J.L. 2008. *Two-stage forest inventory with LIDAR on the Fort Lewis military installation.* M.S. thesis, Univ. of Washington, Seattle, WA.

US FOREST SERVICE. 2000. *Western Oregon 1995–1997, August 15, 2000.* CD available online from www.fs.fed.us/pnw/fia; last accessed May 15, 2009.

WOODS, M., K. LIM, AND P. TREITZ. 2008. Predicting forest stand variables from LiDAR data in the Great Lakes-St. Lawrence forest of Ontario. *For. Chron.* 88(6):827–838.

WULDER, M.A., C.W. BATER, N.C. COOPS, T. HILKER, AND J.C. WHITE. 2008. The role of LiDAR in forest management. *For. Chron.* 84(6):807–826.