


Geospatial Technologies

An Accuracy Assessment of Field and Airborne Laser Scanning–Derived Individual Tree Inventories using Felled Tree Measurements and Log Scaling Data in a Mixed Conifer Forest

Aaron M. Sparks,^{1,*}  Mark V. Corrao,^{1,2} Robert F. Keefe,¹ Ryan Armstrong,² and Alistair M.S. Smith³

¹Department of Forest, Rangeland, and Fire Sciences, College of Natural Resources, University of Idaho, Moscow, Idaho, 83844, USA (asparks@uidaho.edu, mcorrao@uidaho.edu, robk@uidaho.edu).

²Northwest Management Incorporated, Moscow, Idaho, 83843, USA (mcorrao@uidaho.edu, armstrong@northwestmanagement.com).

³Department of Earth and Spatial Sciences, College of Science, University of Idaho, Moscow, Idaho, 83844, USA (alistair@uidaho.edu).

*Corresponding author email: asparks@uidaho.edu

Abstract

On-the-ground sample-based forest inventory methods have been the standard practice for more than a century, however, remote sensing technologies such as airborne laser scanning (ALS) are providing wall-to-wall inventories based on individual tree measurements. In this study, we assess the accuracy of individual tree height, diameter, and volume derived from field-cruising measurements and three ALS data-derived methods in a 1.1 ha stand using direct measurements acquired on felled trees and log-scale volume measurements. Results show that although height derived from indirect conventional field measurements and ALS were statistically equivalent to felled tree height measurements, ALS measured heights had lower root mean square error (RMSE) and bias. Individual tree diameters modeled using a height-to-diameter-at-breast-height model derived from local forest inventory data and the software ForestView had moderate RMSE (8.3–8.5 cm) and bias (-3.0 – -0.3 cm). The ALS-based methods underdetected trees but accounted for 78%–91% of the field reference harvested merchantable volume and 71%–99% of the merchantable volume scaled at the mill. The results also illustrate challenges of using mill-scaled volume estimates as validation data and highlight the need for more research in this area. Overall, the results provide key insights to forest managers on accuracies associated with conventional field-derived and ALS-derived individual tree inventories.

Study Implications: Forest inventory data provide critical information for operational decisions and forest product supply chain planning. Traditionally, forest inventories have used field sampling of stand conditions, which is time-intensive and cost-prohibitive to conduct at large spatial scales. Remote sensing technologies such as airborne laser scanning (ALS) provide wall-to-wall inventories based on individual tree measurements. This study advances our understanding of the accuracy of conventional field-derived and ALS-derived individual tree inventories by evaluating these inventories with felled tree and log scaling data. The results provide key insights to forest managers on errors associated with conventional field and ALS-derived individual tree measurements.

Keywords: airborne laser scanning, lidar, forest inventory, stem volume, felled tree, log scaling, validation

Forest inventory is a fundamental component of forest management and the forest products supply chain or the flow of wood products from the forest to the end user. Inventory data provides critical information for long-term forest planning, operational decisions, harvest scheduling, investment, and forest product supply chain planning (Maltamo et al. 2021; Tinkham et al. 2018). For more than a century, forest inventories have relied on field sampling of forest stand conditions (Frayer and Furnival 2000; Maltamo et al. 2021), which are time-intensive, accuracy-limited (Luoma et al. 2017) and cost-prohibitive to collect wall-to-wall (i.e., spatially complete) (Durrieu et al. 2015; Vauhkonen et al. 2014a). With advances in remote sensing, forest inventories

have shifted toward incorporating technologies such as airborne scanning light detection and ranging (LiDAR), also referred to as airborne laser scanning (ALS), as it can gather wall-to-wall, three-dimensional forest structural data. These technologies can provide wall-to-wall data at a lower cost per unit area than conventional field sampling, especially when applied over large spatial scales or cost-shared (Hudak et al. 2020; White et al. 2016).

Wall-to-wall forest inventories using ALS data are derived either through area-based or individual tree detection (ITD) approaches (Holopainen et al. 2014; Vauhkonen et al. 2014a; White et al. 2016). Area-based methods use gridded summaries of the ALS point cloud (e.g., height percentiles, height stratified

Received: October 31, 2023. Accepted: March 8, 2024.

© The Author(s) 2024. Published by Oxford University Press on behalf of the Society of American Foresters. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

Height and DBH Accuracy Assessment

Several statistical tests were used to evaluate the similarity of cruised and ALS-derived individual tree height and DBH to felled tree height and DBH. A Kolmogorov-Smirnov test was used to assess whether cruised and ALS-derived height distributions were statistically similar to the felled tree height distribution. This test further assessed the similarity of modeled and ALS-derived DBH distributions to the cruised DBH distribution. The Kolmogorov-Smirnov test identifies differences in the cumulative distribution function of two data distributions. The test statistic (D) is the maximum difference between the cumulative distribution functions, where D values close to 0 indicate significant overlap in the data distributions and values close to one indicate little to no overlap in the data distributions. For this test, the null hypothesis is that the test dataset distribution follows a reference dataset distribution, and the alternate hypothesis is that the distributions are not similar. The null hypothesis is rejected if $P < .05$. This test is particularly useful for quantifying the similarity of the shape of two distributions. For example, two distributions with the same mean but different shape will produce large D values.

The accuracy of cruised and ALS-derived individual tree height and DBH was assessed using regression-based equivalence tests (Robinson et al. 2005). For these tests, the null hypothesis is that the regression slope and intercept between paired sets of data are significantly different, and the alternate hypothesis is that they are not significantly different. A linear regression model is fit using the paired datasets and an upper and a lower one-sided 95% confidence interval for both the intercept and slope is computed using the standard error regression outputs. The null hypothesis of dissimilarity is rejected if the joint one-sided 95% confidence intervals are entirely contained within a user-defined region of equivalence (e.g., $\pm X\%$). Intercept equality implies the means of two datasets are not significantly different and slope equality implies that the regression slope is not significantly different than one. Many similar studies (Corrao et al. 2022; Falkowski et al. 2008; Robinson et al. 2005; Sparks and Smith 2021; Sparks et al. 2022) have used an arbitrarily selected region of equality ($\pm 25\%$) for the intercept and slope. Here, we follow the Robinson et al. (2005) suggestion of reporting the minimum region of equivalence that would still result in the rejection of the null hypothesis of dissimilarity. Equivalence was assessed separately between cruised tree height and felled tree height and ALS-derived tree height and felled tree height. Similarly, equivalence between modeled DBH (local and regional models) and cruised DBH and ALS-derived growth-adjusted DBH and cruised DBH were also assessed. The average RMSE (1) and mean bias (2) were calculated for all modeled and ALS-derived height and DBH as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{x} - x_i)^2}{n}} \quad (1)$$

$$Mean\ bias = \frac{\sum_{i=1}^n (\hat{x} - x_i)}{n} \quad (2)$$

where \hat{x} are the predicted values, x_i are the observed values, and n is the number of observations. All statistical analyses were conducted in R (R Core Team 2023), and we used the “equivalence” R package (Robinson 2016) to conduct the regression-based equivalence tests.

Comparison of Harvested Volume

We calculated gross harvested merchantable and pulp volume for all harvested trees within the cruise area and compared the total volume estimated by the cruise and ALS-derived methods. Specifically, we compared total volume estimates for (1) matched cruise and ALS-detected trees and (2) all trees within the cruise area (i.e., matched and unmatched trees). The cubic volume of each harvested tree was calculated using five sets of height and DBH measurements: (1) felled height and cruised DBH, (2) cruised height and DBH, (3) ALS-derived height and DBH modeled using the local forest inventory dataset, (4) ALS-derived height and DBH modeled using the regional forest inventory dataset, and (5) ForestView-derived height and DBH. For all sets of measurements, we estimated diameter inside bark using the Kozak (2004) taper equation from 0.3 m, assumed to be the stump height, to the treetop in 0.5 m segments. We used regionally derived taper equation parameters (Pancoast 2018; Poudel et al. 2018) and the Smalian formula (3) for calculating the cubic volume (V , m^3) of all segments of each harvested tree:

$$V = \frac{L}{2} (A_1 + A_2) \quad (3)$$

where L is the length of the log (m), A_1 is the area of the small end of the log (m^2), and A_2 is the area of the large end of the log (m^2). The total volume of each tree was calculated as the sum of merchantable volume, or volume of the stem where the diameter was greater than 15.24 cm, and pulp volume, or volume of the stem where the diameter was less than 15.24 cm but greater than 7.62 cm. All harvested individual tree volumes were summed to provide gross harvested merchantable and pulp volume for the cruise area. The volume derived using felled heights and cruised DBH was assumed to be the most accurate and served as the reference volume estimate.

We also compared gross harvested merchantable volume for the entire study stand to scaled merchantable volume conducted at the processing mill. Harvest of the study stand largely occurred on separate days from harvesting of the surrounding stand, so it was possible to assign harvested trees within the study area to the specific truckloads taken to the mill. One exception was a single harvesting day where logs from the study area and the surrounding stand were mixed. Ground personnel estimated that up to two truckloads out of the six for that day originated from the study area. Due to this complexity, we report scaled volume with uncertainty bounds ranging from zero truckloads (0% of harvested volume for that day) to two truckloads (~33% of harvested volume for that day).

All logs on ten out of the twelve log loads originating from the study stand were scaled at the processing mill using the Scribner decimal C log rule. This log rule estimates the number of one-inch-thick boards spaced one-quarter inch apart that can fit inside the circular area of a log’s smallest end. Board foot yield of this scaling cylinder can be calculated by summing the widths of each board, dividing by 12, and multiplying by the length of the log. Most logs are tapered and are not perfect cylinders, which means that this method ignores volume outside the scaling cylinder. Other cubic scales provide a more accurate estimate of total usable fiber (e.g., Newton and Smalian cubic volume formulations), however, the Scribner decimal C log rule

Height-to-DBH Models

Figure 4 shows the allometric relationships between height and DBH for the local forest inventory dataset (figure 4a) and the regional forest inventory dataset (figure 4b). The best-fit regression model for both the local and regional datasets used a power-law function (local results: power-law AIC: 350.7, logarithmic AIC: 11278.6, linear AIC: 10889.8; regional results: power-law AIC: 300.2, linear AIC: 12447.5, logarithmic AIC: 12714.5). The relationship between height and DBH for the local dataset had a high r^2 (0.80) and low residual standard error (0.3 cm). The relationship between height and DBH for the regional dataset had a lower r^2 (0.70) and low residual standard error (0.3 cm).

Individual Tree DBH Accuracy

Figure 5 shows individual tree DBH distributions and associated cumulative distribution functions derived from the local and regional height-to-DBH models and ForestView in comparison with the cruised DBH distribution. All models produced unimodal DBH distributions with a peak around 30 cm, whereas the cruised DBH distribution was bimodal with peaks around 25 cm and 40 cm. The Kolmogorov-Smirnov test statistic D was closer to zero than one for all modeled DBH versus cruised DBH tests, indicating significant overlap in the paired distributions. However, the null hypothesis that the modeled DBH distribution follows the cruised DBH distribution was rejected for both the

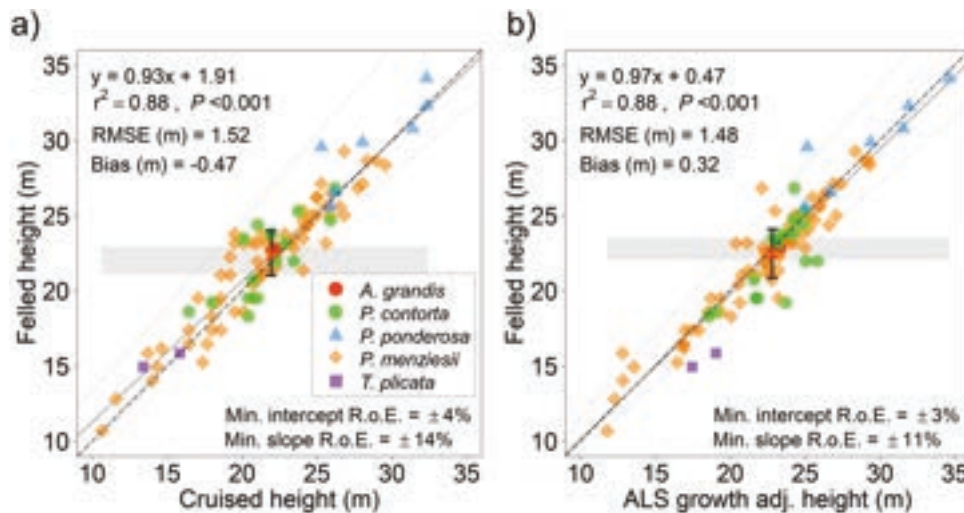


Figure 3 Regression-based equivalence test graphs for cruised height versus felled height (a) and ALS-derived growth-adjusted height versus felled height (b). Data for the ninety-nine matched and felled trees are shown. The minimum region of equivalence (R.o.E.) that would still lead to rejection of null hypothesis of dissimilarity is reported for both the intercept and slope. The grey polygon represents the minimum region of equivalence for the intercept. The cruised and ALS-derived mean heights are equivalent to the mean felled heights when the vertical red bar is completely within the grey polygon. The grey dashed lines represent the minimum region of equivalence for the slope. If the vertical black bar is within the grey dashed lines, then the regression slope is significantly similar to one. The solid black line represents the best-fit linear regression model, and the black dashed line represents the 1:1 line. The coefficient of determination (r^2) and associated P -value for the linear regression models are also presented. Individual trees are symbolized by field-classified species and are shown for illustrative purposes.

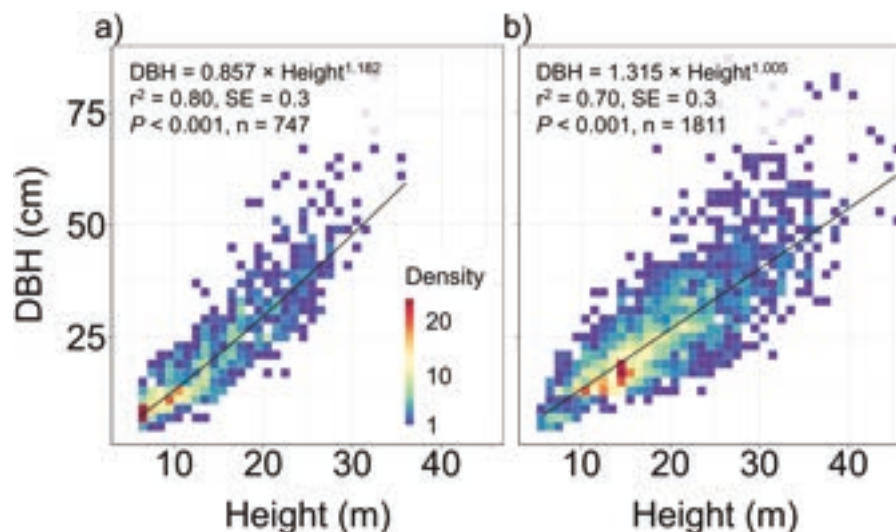


Figure 4 Relationships between height and DBH using the local forest inventory dataset (a) and the regional forest inventory dataset (b). The black line represents the best-fit regression model. The scatterplot point densities, calculated using a $\sim 2 \times 2$ quantization of the plot axes, are displayed with a rainbow color scale.

achieved lower DBH errors (RMSE: 3.8–5.9 cm) than those in this study (RMSE: 8.28–10 cm) (Heurich 2008; Persson et al. 2002; Popescu 2007; Salas et al. 2010). For example, Popescu (2007) observed an RMSE of 4.9 cm in *Pinus taeda* L.-dominated stands, Persson et al. (2002) observed RMSE of 3.8 in *Picea abies* L. Karst. and *Pinus sylvestris* L. stands, and Heurich (2008) observed RMSE of 4.6–5.9 in *Picea abies* and *Fagus sylvatica* stands. It is likely that incorporating a refined screening process in the local and regional height-to-DBH modeling to select plots with similar site conditions to our study stand would result in more accurate DBH estimates. However, including simple height-to-DBH models in this analysis is useful information in cases where a forest manager has limited ALS-derived data and processing capability (i.e., no available individual crown diameter data). Additionally, like the observed height error, RMSE in this study is likely higher than in other studies due to error introduced by using average diameter growth between the ALS acquisition in 2020 and felled tree measurements in 2022. Tree DBH modeled from UAS LiDAR and SfM data has been shown to have higher accuracy (RMSE: ~0.8–4.8 cm) than DBH modeled from ALS data (Kukkonen et al. 2022; Sun et al. 2022; Swayze et al. 2021), however, data acquisition using ALS is more efficient over large areas (i.e., >1,000 ha) compared with present day UAS (White et al. 2016).

Scaled volume provides a promising independent validation data source; however, as this study illustrates, there are many challenges and uncertainties associated with using this data. First, it should be noted that although we tried to minimize tree detection and matching errors using manual matching methods, the overall detection rate was 60%. This underdetection is likely the primary factor of the ~1%–29% merchantable volume underestimation of the top ALS methods. Second, uncertainty in log loads originating from the study area is a significant source of volume error. Tracking harvested trees is a challenge given that harvesting and log transport logistics sometimes necessitate mixed log loads that incorporate harvested trees from different parts of the stand or multiple stands. Third, as volume of two of the twelve log loads was estimated via a weight-to-volume conversion factor, underestimation or overestimation could occur due to weight-to-volume conversion factor error. Furthermore, the Scribner-scaled loads could be underestimated compared to volume calculated using the Smalian formula, as this rule assumes a quarter-inch kerf allowance. Uncertainty in ALS-estimated DBH and board feet to cubic volume conversion factors could contribute to the volume mismatch. The ALS-based methods tended to underpredict DBH (figure 6), which directly translates into lower individual tree volume estimates than cruised volume estimates (figure 7a). We could not find error estimates associated with board feet to cubic volume conversion factors, which is a research need, considering the widespread use of both board feet and cubic volume in the western United States (Saralecos et al. 2014; Spelter 2002, 2004) and the likely increasing use of scaled volume as ALS-derived forest inventory validation data. Other studies have shown moderate agreement between ALS-derived volume estimates and postharvest volume measures. White et al. (2014) found that volume derived from cover type–adjusted volume tables underestimated weight scale volumes by 19.8% whereas volume derived from ALS area-based analysis had much closer agreement to weight scale volumes (+0.6%) in coniferous boreal forest in Alberta, Canada. Likewise,

Woods et al. (2011) reported that mean area-based analysis estimates of stand volume from ALS data were within 10% of scaled volume in a coniferous boreal forest stand in Ontario, Canada. Other studies have shown that individual tree volume derived from ALS measurements was within ~4% of harvester-derived individual tree volume (Peuhkurinen et al. 2007).

A potential limitation of this study was the number of matched individual trees ($N = 99$), representing five conifer species from a single stand. Although this may appear to be a small sample size, it is comparable to the upper end of previous studies using direct measurements on felled trees or high-precision measurements acquired using terrestrial laser scanning (TLS). Although focused on a single stand, the species variability in this study is also greater than most earlier studies that mainly focused on only one or two species. For example, Tinkham et al. (2016) used measurements from 60 felled *Abies grandis* in the intermountain western United States, Sibona et al. (2017) used measurements from 100 felled conifers in the Italian Alps, Ganz et al. (2019) used measurements from 30 felled *Pseudotsuga menziesii* in Germany, and Corrao et al. (2022) used measurements from 139 felled *Pinus taeda* in the southern United States. Likewise, Andersen et al. (2006) compared ALS measurements to TLS measurements for fifty-nine total *Pseudotsuga menziesii* and *Pinus ponderosa* individuals. The primary reason for these small sample sizes is the time and person hours required to acquire felled-tree measurements. It is also often infeasible to fell trees or collect measurements on felled trees due to safety or logistical issues, leading most ITD validation studies to use field-collected indirect heights. However, studies that use direct measurement are critical for validation of remotely sensed measurements, and the relative lack of such studies indicates that more are needed to assess ALS-derived measurement accuracy across diverse species and stand conditions. Equally, these types of field measurements are essential for assessing the accuracy and utility of larger spatial scale projects and commercial applications.

Conclusions

Accurate forest inventories are necessary for forest management and forest products supply chain planning. This study advances our understanding of the accuracy of conventional field and ALS-derived individual tree inventories by evaluating these inventories with felled tree measurements and log scaling data in a coniferous forest stand with diverse species composition and structure. The results show that although ALS-derived and indirect field measurements of height are statistically equivalent to direct height measurements, ALS-derived height had lower RMSE (1.48 m) and bias (0.32 m) for this stand than field measurements (RMSE = 1.52 m, bias = -0.47 m). Our results also highlight the utility and uncertainty of using ALS-derived individual tree height to model DBH. In this stand, although both ForestView-derived DBH and DBH derived from local height-to-DBH models were statistically equivalent to field measured DBH, RMSE was moderate (8.3–8.5 cm). The results show that the largest error source in volume estimation was the underdetection of trees, followed by error in DBH and height. Like prior studies, this study shows that although there was an underdetection of trees (60% detection rate), the best ALS-derived volume estimates accounted for 78%–91% of the

