ISSN 1946-7664. MCFNS 2022
Vol. 14, Issue 1, pp. 1–14
*Mathematical and Computational Forestry & Natural-Resource Sciences*

AVAILABLE ONLINE AT HTTP://MCFNS.COM

Submitted: Nov. 26, 2021
Accepted: Mar. 28, 2022
Published: Apr. 30, 2022
Last Correction: Mar. 30, 2022

# ESTIMATING DIAMETER AND HEIGHT DISTRIBUTIONS FROM AIRBORNE LIDAR VIA COPULAS

Ting-Ru Yang, John A. Kershaw, Jr.*

*University of New Brunswick, The Faculty of Forestry and Environmental Management, Fredericton, NB, Canada*
**Corresponding Author*

Abstract. Estimates of quantitative variables in forest stands often are required. Light Detection and Ranging (LiDAR) can create three-dimensional point clouds of forest structures and ground surface elevation maps. These features are useful for quantifying forest stand parameters such as volume and canopy height at broad scales. This study explores the potential of applying copulas and LiDAR metrics to obtain diameter and height estimates. Predicted values were compared with field measurements. Diameter and height distributions were obtained using moment–based parameter recovery and prediction of moments using nonlinear least squares from LiDAR attributes. We then used copula methods to link the diameter and height distributions. Using the diameter and height distributions, other attributes such as volume or carbon content can be estimated and summed to obtain area-based estimates.

**Keywords:** stand structure, height-diameter relationships, bivariate distributions, copulas, moment-based parameter recovery

## 1 Introduction

Estimates of quantitative variables in forest stands are required by forest managers to evaluate forest resources and to schedule future silviculture treatments (Clutter et al., 1983). Because measurements of tree height and diameter are expensive and laborious, it is desirable to predict these variables in many situations. Light Detection and Ranging (LiDAR) is an increasingly popular method of remote sensing because it provides three-dimensional point clouds to develop detailed canopy surface and ground surface elevation maps and has potential to estimate stand or individual tree attributes across large geographic areas (Asner et al., 2011; Hummel et al., 2011; Reutebuch et al., 2005; Wulder et al., 2008). Most efforts in analyzing LiDAR data focus on point estimation or individual tree estimation which requires exact georeferencing from field-based training data to LiDAR data (Kaartinen et al., 2012; Maltamo et al., 2009; White et al., 2013). While point estimates are useful in making management decisions, a better use of LiDAR datasets might be found by predicting distributions of forest inventory attributes.

Developing methods for estimating diameter and height distributions using probability distribution functions can result in biological- and economical-based predictions that are more informative to forest managers. Various probability distributions and methods for estimating parameters of those distributions have been used to model diameter distributions and characterize the height-diameter relationships via airborne LiDAR scan (ALS) data (Arias-Rodil et al., 2018; Gobakken and Næsset, 2004; Mehtätalo et al., 2007; Thomas et al., 2008). One of the most popular methods, Weibull distributions with moment-based parameter recovery, shows a potential link in developing landscape-level approaches to LiDAR prediction (Arias-Rodil et al., 2018; Mehtätalo et al., 2007). Due to the ability of obtaining good height estimates from ALS, there is a potential to build relationships between diameter and height with lower inventory costs and associated errors.

A copula, with an emphasis on modelling population distributions, is a special class of multivariate distributions where the marginal distributions are all uniform distributions [0,1] (Genest and MacKay, 1986; Nelson, 2006). The uniform distribution can be replaced with any probability distribution and stripped from the copula through the statistical process of translation (Genest and MacKay, 1986; Nelson, 2006). By translating into any mix of distributions (Nelson, 2006), copulas become a flexible and powerful tool for analyzing depen-

dent processes arising from a number of different underlying factors (Genest and MacKay, 1986; Wang, 1998). Because of the strong relationship between diameter and height, copulas provide a possible framework for developing LiDAR-based analyses that may have broader application domains than the region from which they were developed (MacPhee et al., 2018). While copulas have been widely applied in many fields (Frees and Valdez, 1998; Genest and MacKay, 1986; Nelson, 2006; Wang, 1998; Yan, 2007), they have only recently been applied to forest spatial structures (Kershaw et al., 2010) and individual tree height-diameter relationships (MacPhee et al., 2018; Wang et al., 2008, 2010). The use of parametric methods in estimating diameter and height distributions from ALS via copula is still largely unexplored. LiDAR presents many opportunities for individual tree analyses and attribute estimation (e.g., Ayrey et al., 2017; Brandtberg et al., 2003; Culvenor, 2002; Li et al., 2012). Copula-based diameter-height models may provide a method to improve individual tree attribute prediction from LiDAR data.

This study utilizes an ALS dataset to estimate copula-based diameter and height distributions in mixed species Acadian Forests in New Brunswick, Canada. The primary objectives were: 1) to develop diameter-height copulas based on LiDAR-derived distribution moments and moment-based parameter recovery; and 2) explore the accuracy of LiDAR-derived diameter-height copulas.

## 2 Materials and Methods

The whole procedure of LiDAR-derived diameter-height copula development and model comparisons is shown in Figure 1.

### 2.1 Study Site

The Noonan Research Forest (NRF, $N45°59'12''$, $W66°25'15''$), located 30km northeast of Fredericton, New Brunswick, Canada, is approximately 1500ha and is composed of a diversity of stand structures and species compositions typical of the Acadian Forest (Loo and Ives, 2003).

*Field Data*—Eighty-three 0.04ha fixed-area permanent sample plots were established on a 100m by 100m sample grid across the NRF's Femelschlag Research Area. All live trees $\geq$ 6.0cm DBH were identified by species, and DBH (nearest 0.1cm) and total height (HT; nearest 0.1m) were measured. All plots were measured in 2014. The mean number of species per Femelschlag plot was about 5 (range was 2 to 8). The other forest characteristics including DBH, quadratic mean diameter, height, stand density and volume for the Femelschlag plots are shown in Table 1.

Table 1: The average, range and standard deviation (Std. Dev.) of plot-level mean diameter at breast height (DBH), quadratic mean diameter at breast height (Dq), height (HT), quadratic mean height (HTq), stand density (TPH) and correlation between DBH and HT (DHcor) for the 83 Femelschlag plots.

| Plot Attribute | Mean | Min. | Max. | Std. Dev. |
|---|---|---|---|---|
| DBH (cm) | 15.5 | 9.7 | 24.5 | 3.1 |
| Dq (cm) | 17.4 | 10.5 | 26.8 | 3.4 |
| HT (m) | 12.4 | 8.8 | 17.4 | 1.8 |
| HTq (m) | 13.2 | 9.0 | 17.6 | 1.9 |
| TPH (# $ha^{-1}$) | 1883 | 575 | 4375 | 807 |
| DHcor | 0.836 | 0.668 | 0.938 | 0.064 |

*Airborne Laser Scanning Data*—A LiDAR scan was obtained on August 2 to September 28, 2015, during leaf-on conditions, using an airborne Riegl Q780i scanner mounted on a plane. The mean flying altitude above sea level was approximately 1,000m. The sensor pulse repetition frequency was between 300 and 400KHz, and the laser wavelength was 1,550nm with a scan angle between $-34°$ to $36°$ from the nadir. The mean swath pulse density was 3 pulses per $m^2$ (this was flown at 50% overlap, providing a final density of 6 pulses per $m^2$) with a footprint size of $0.35m^2$ and the sensor collected up to 10 returns per pulse.

### 2.2 LiDAR Processing

Various LiDAR metrics were derived from LiDAR point clouds. A ground surface model was derived from the last returns and interpolated across the regions of interest using the lasground() function in the lidR package (Roussel et al., 2020). LiDAR returns were then classified as either ground or non-ground points and the ground points removed. Elevations of LiDAR returns (Z) were converted to heights above ground (H) by subtracting interpolated ground elevation (G) from Z using the lasnormalize() function in the lidR package. All LiDAR manipulations were conducted in the R statistical software (R Development Core Team, 2021).

*LiDAR Metrics*—As recommended by Hayashi et al. (2015), 20m by 20m LiDAR cells centered on the Femelschlag plot centers were used to extract metrics, and then used to fit relationships between the Femelschlag plot data and the LiDAR metrics. Nineteen unique LiDAR metrics (Table 2) were extracted from the ground-normalized LiDAR HT distributions. Canopy height surface models established by the first returns and point heights for given distributional quantiles were used to obtain height-based metrics. Different return numbers per $m^3$ of plot space from ground, canopy

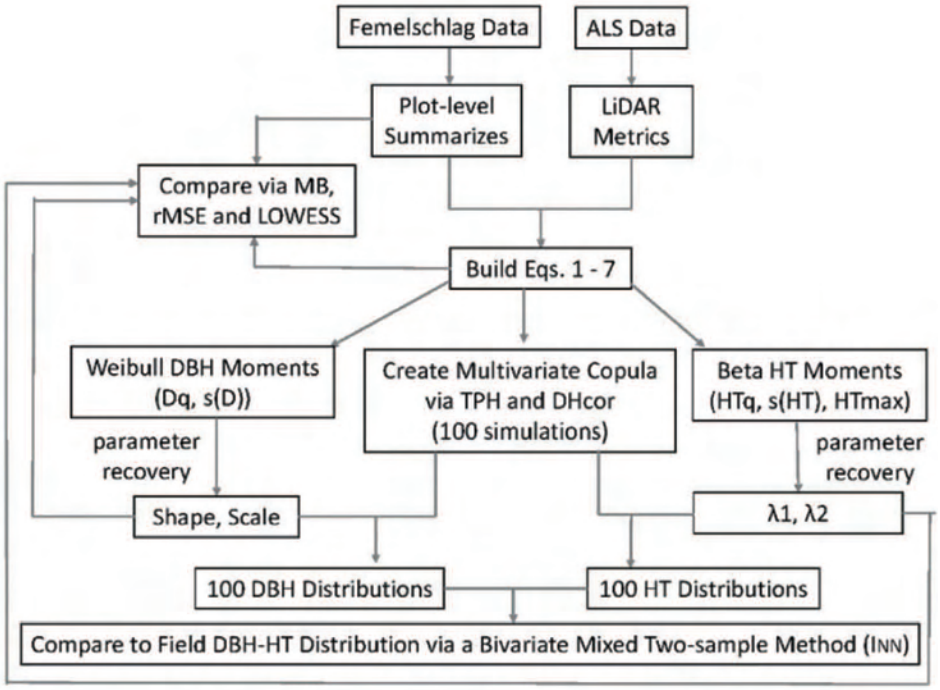


Figure 1: Procedure of approach establishment, simulation design and model comparison.

and the total LiDAR points were used to determine the density-based metrics. The variation in horizontal and vertical distributions were represented by crown rugosity and kurtosis. Lacunarity (Plotnick et al., 1993), as measure of gap size, was used to characterize horizontal heterogeneity.

### 2.3 Diameter-Height Modelling

*Diameter Marginal Distribution Modeling*—Because the truncated Weibull function consistently has shown to be more accurate for diameter distribution modeling than beta or Johnson's SB functions (Bailey and Dell, 1973; Borders and Patterson, 1990; Little, 1983; Palahí, 2007; Zutter et al., 1982), the two-parameter, left-truncated Weibull distribution (Bury, 1975; Zutter et al., 1986) was used to model diameter distributions. The algorithm (moment-based parameter recovery) proposed by Burk and Newberry (1984) was used to recover the parameters from the predicted moments. Prediction systems using NLS estimation and LiDAR metrics as predictor variables were developed. A custom R function based on Burk and Newberry's (1984) algorithm was used to recover parameters from the predicted moments.

The two-parameter, left-truncated Weibull distribution includes a scale parameter and a shape parameter. The minimum DBH measured in the field data was 6.0cm and was used as the truncation point. To recover the Weibull shape and scale parameters, the first two moments of the distribution were required (Burk and Newbury, 1984). The first two moments are defined by the arithmetic mean $\left(\overline{D}\right)$ and the quadratic mean $D_q$ diameters. Either (1) both moments need to be estimated, or (2) one moment estimated directly and the difference $\left(\Delta = D_q - \overline{D}\right)$, standard deviation $(SD)$, or coefficient of variation $(CV)$ estimated and the other moment calculated. Commonly, the second approach is used predicting $D_q$ and $SD$ because it assures that $D_q \geq \overline{D}$.

After an exhaustive search using a combination of graphical methods, boosted regression methods, and preliminary model screening, the following equation form was adopted for predicting $D_q$:

$$D_q = b_0 \cdot q15LiDAR^{b_1} \cdot q45LiDAR^{b_2}, \qquad (1)$$

where $q15LiDAR =$ the $15^{th}$ quantile of the LiDAR HT distribution and $q45LiDAR =$ the $45^{th}$ quantile of the LiDAR HT distribution. Standard deviation of the DBH distributions $(s(D))$ was predicted using:

$$s\left(D\right) = b_0 \cdot q75CHT^{b_1}, \qquad (2)$$

where $q75CHT =$ the $75^{th}$ quantile of the canopy surface model.

*Height Marginal Distribution Modeling*—Because of the presence of emergent trees on some plots, the four-parameter Beta distribution (Bury, 1975) was used to

Table 2: LiDAR metrics and associated statistics (mean ± stand deviation) for use in modeling forest attributes on the Femelschlag plots.

| No. | Metric | Definition | Statistics |
|---|---|---|---|
| 1 | MaxCHT | Maximum canopy surface height | 30.2 ± 10.4 |
| 2 | MeanCHT | Mean canopy surface height | 14.2 ± 1.3 |
| 3 | q25CHT | 25th percentile of canopy surface height | 12.1 ± 1.9 |
| 4 | q50CHT | 50th percentile of canopy surface height | 15.0 ± 1.4 |
| 5 | q75CHT | 75th percentile of canopy surface height | 17.1 ± 1.2 |
| 6 | q15LiDAR | Height of 15th percentile of point clouds | 5.8 ± 1.4 |
| 7 | q25LiDAR | Height of 25th percentile of point clouds | 8.5 ± 1.3 |
| 8 | q35LiDAR | Height of 35th percentile of point clouds | 10.4 ± 1.2 |
| 9 | q45LiDAR | Height of 45th percentile of point clouds | 11.9 ± 1.2 |
| 10 | q55LiDAR | Height of 55th percentile of point clouds | 13.2 ± 1.2 |
| 11 | q65LiDAR | Height of 65th percentile of point clouds | 14.4 ± 1.2 |
| 12 | q75LiDAR | Height of 75th percentile of point clouds | 15.5 ± 1.1 |
| 13 | q85LiDAR | Height of 85th percentile of point clouds | 16.8 ± 1.1 |
| 14 | q95LiDAR | Height of 95th percentile of point clouds | 18.7 ± 1.0 |
| 15 | Kurtosis | Kurtosis of return heights | 2.8 ± 0.4 |
| 16 | Rugosity | Roughness of canopy surface | 6.8 ± 1.5 |
| 17 | Rumple | Ratio of the canopy surface area to plot area | 5.8 ± 1.4 |
| 18 | LAD | Leaf area density | 0.5 ± 0.1 |
| 19 | Lacunarity | Gap size of the canopy surface | 1.1 ± 0.2 |

model the height marginal distribution. The four-parameter Beta distribution is a variant of the Beta distribution that transforms the $[0, 1]$ distribution to a $[Min, Max]$ distribution. Thus, the four parameters were Minimum HT ($HT_{min}$), Maximum HT ($HT_{max}$), and the two shape parameters $\lambda_1$ and $\lambda_2$ (Bury, 1975). The $HT_{min}$ was 1.3m corresponding to breast height. Maximum height was predicted from LiDAR using:

$$HT_{\max} = b_0 \cdot MeanCHT^{b_1} \cdot q75LiDAR^{b_2}, \quad (3)$$

where $q75LiDAR$ = the $75^{th}$ quantile of the LiDAR HT distribution. Similar to the procedure for Weibull parameter recovery, the two shape parameters can be recovered from estimates of the first two moments. As with diameter, we predicted the quadratic mean heights ($HT_q$) and SDs ($s(HT)$). The shape parameters ($\lambda_1$ and $\lambda_2$) were recovered using an approach similar to Burk and Newberry (1984) but modified to reflect the Beta distribution.

The following equation was used to predict $HT_q$ from LiDAR:

$$HT_q = b_0 \cdot q75CHT^{b_1} \cdot q15LiDAR^{b_2} \cdot q45LiDAR^{b_3}, \quad (4)$$

where $q75CHT$ = the $75^{th}$ quantile of the canopy surface model; $q15LiDAR$ = the $15^{th}$ quantile of the LiDAR HT distribution; and $q45LiDAR$ = the $45^{th}$ quantile of the LiDAR HT distribution. Standard deviations of the HT distributions ($s(HT)$) were predicted from LiDAR using:

$$s\,(HT) = b_0 \cdot q25CHT^{b_1} \cdot q75LiDAR^{b_2}, \quad (5)$$

where $q25CHT$ = the $25^{th}$ quantile of the canopy surface model and $q45LiDAR$ = the $45^{th}$ quantile of the LiDAR HT distribution. $HT_q$ and $s(HT)$ were calculated for each Femelschlag plot from the individual tree data and combined with the LiDAR metrics. All equations were estimated using the nls() function in R (R Development Core Team, 2021).

*Number of Trees per Hectare*—Before using copulas to link predicted DBH and HT distributions, the number of standard normal variates is required (i.e., the number of trees to simulate). We directly estimate the number of trees per hectare ($TPH$, stems/ha) using LiDAR metrics. The equation for predicting $TPH$ was:

$$TPH = b_0 \cdot q15LiDAR^{b_1} \cdot MeanCHT^{b_2} \cdot Kurt^{b_3}, \quad (6)$$

where $q15LiDAR$ = the $15^{th}$ quantile of the LiDAR HT distribution; $MeanCHT$ = the mean canopy HT estimated from the canopy surface model; and $Kurt$ = kurtosis of return heights.

*DBH-HT Correlation*—The copula required an estimate of the correlation between DBH and HT ($DH_{cor}$). Be-

cause $0 \leq DH_{cor} \leq 1$, a logistic equation was adopted:

$$DH_{cor} = \frac{e^{(b_0+b_1 \cdot MeanCHT+b_2 \cdot q25LiDAR)}}{1+e^{(b_0+b_1 \cdot MeanCHT+b_2 \cdot q25LiDAR)}}, \quad (7)$$

where $MeanCHT$ = the mean canopy HT estimated from the canopy surface model; and $q25LiDAR$ = the $25^{th}$ quantile of the LiDAR HT distribution.

*Copula Model*—MacPhee et al. (2018) compared nonlinear mixed effects (NLME), random forest imputation (Breiman, 2001; Temesgen and Ver Hoef, 2015), and copula methods for predicting heights from diameters. In this study, we estimate stand density from LiDAR metrics first, and then the goal is to predict the joint distribution of DBH and HT via copula methods. Nonlinear least squares (NLS) was used in this study instead of NLME because we did not stratify the Femelschlag plots by forest type due to the limited numbers of plots in conditions other than a mixed species, multicohort state. The field measured data and attributes from the Femelschlag plots were used to parameterize copula models.

A normal copula was used to model and simulate the relationship between DBH and HT (Genest and MacKay, 1986; Kershaw et al., 2010; MacPhee et al., 2018). The approach used in this study followed that presented by Kershaw et al. (2010) and MacPhee et al. (2018). Diameters and heights were predicted for each tree based on random sampling from a normal copula. The procedure was as follows:

1. The number of standard normal variates for diameter ($S_D$) and height (($S_H$) distributions were randomly generated based on the estimated number of trees per plot (eq. 6). Both $S_D$ and $S_H$ are Normal distribution variates with a mean of 0 and standard deviation of 1.

2. $S_H$ and $S_D$ were column-bound to form a matrix called $[S]$. Both $S_H$ and $S_D$ columns were correlated by multiplying a symmetric positive-definite square matrix established by the correlation between DBH and HT ($DH_{cor}$, eq. 7):

$$[C] = [S] \cdot chol\left(\begin{bmatrix} 1 & DH_{cor} \\ DH_{cor} & 1 \end{bmatrix}\right), \quad (8)$$

   where $[C]$ is the correlated standard Normal variates; $chol\,()$ is the Choleski decomposition of the symmetric correlation matrix.

3. The inverse (cumulative) normal distribution function was used to strip off the normal distribution. This matrix of uniform correlated marginals ($[U]$) is the multivariate copula.

4. The left-truncated, two-parameter Weibull distribution was used to model the DBH distribution (eqs. 1–2) and was applied to the second column of $[U]$ to obtain DBH estimates. Similarly, the four-parameter Beta Distribution was used to model the HT distribution (eqs. 3–5) and applied to the first column of $[U]$.

The copula links the predicted HT and DBH distributions. Both DBH and HT were estimated by randomly sampling from the copula. We initially explored using 1, 5 and 10 random samples but, as found by MacPhee et al. (2018), the mean DBH was unaffected by sample size while the SD slightly decreased with increasing copula samples. Therefore, we only present results for the average of 5 random samples and 100 replicates. DBH–HT distributions were simulated and compared to the field measured distributions.

### 2.4 Statistical Analyses

Differences between field measured values and the predictions derived from ALS data were examined by 1:1 plots and locally weighted scatterplot smoothing (LOWESS, Cleveland, 1979; Kutner et al., 2004) to highlight differences over the range of observed and predicted values. Two point-wise measures of goodness-of-fit, mean bias ($MB$), and root mean squared error ($rMSE$) were calculated:

$$MB = \frac{\sum_{i=1}^{n} (Observed_i - Predicted_i)}{n} \quad (9)$$

and

$$rMSE = \sqrt{\frac{\sum_{i=1}^{n} (Observed_i - Predicted_i)^2}{n}}, \quad (10)$$

where $Observed_i$ was a parameter calculated from field measurements on the $i^{th}$ Femelschlag plot, and $Predicted_i$ was a parameter derived from ALS data, and $n$ was the total number of Femelschlag plots.

In this study, the focus is on DBH and HT distributions rather than individual tree DBH and HT predictions. Point-wise goodness of fit statistics for individual trees were not evaluated. Instead, we used the bivariate mixed two-sample method to compare the HT-DBH distributions derived here based on Euclidian distance measures (MacPhee et al., 2018; Narsky, 2003a; Schilling, 1986). Briefly, the bivariate mixed two-sample method selects a random sample from distribution A and calculates the Euclidian distance from all other points in A and all points in distribution B, and then repeats the process selecting a random sample from B and calculates

the Euclidian distance from all other points in B and all points in distribution A.

The set of $k$ nearest neighbors is determined and the $k$ nearest neighbors classified as either belonging to AA or BB ($I_{NN} = 1$) or AB or BA ($I_{NN} = 0$) If the two distributions are identically distributed, the average of $I_{NN}$ should be 0.5 since there would be an equal probability that the nearest neighbor would be from distribution A or B (Narsky, 2003b). Because single nearest neighbor statistics can be quite variable (Ripley, 2004), this study followed the recommendation from MacPhee et al. (2018) and used k=3 nearest neighbors to calculate $I_{NN}$.

## 3 Results

### 3.1 Marginal Distributions

Table 3 shows the parameter estimates, standard errors, and associated NLS regression statistics for Eqs. 1–7. In general, the ALS data provided metrics that explained lower proportions of the variation in the diameter moments and resulted in higher root mean square errors ($rMSE$s) than the height moments. For predicting $D_q$, the model accounted for 34% of the variation and had an $rMSE$ of 2.8cm, while the $HT_q$ model explained 47% of the variation with a $rMSE$ of 1.3m. The model predicting $s\,(D)$ accounted for 23% of the variation with a $rMSE$ of 1.7cm. Around 49% of the variation in s(HT) was explained with an rMSE of 0.7m. Compared to the other two height moments, only 28% of variation was explained in $HT_{max}$ with a $rMSE$ of 2.5m. The model predicting $TPH$ accounted for 50% of the variation and had a $rMSE$ of 569 trees/ha. Less than 20% of the variation in $DH_{cor}$ was accounted for with a $rMSE$ of 0.06.

The observed and predicted parameters for $DBH$, $HT$, $TPH$ and $DH_{cor}$ are presented in Figure 2. No matter which parameter is considered, small observations were always overestimated and large observations were always underestimated. Mean bias ($MB$) and $rMSE$ for parameters of $DBH$, $HT$ and $DH_{cor}$ are presented in Table 4. Similar to the fit results, models for height moments predicted much better than models for diameter moments. Bias for $TPH$ model was $< 1\%$ and errors were typically $< 30\%$. The model predicting $DH_{cor}$ had about -0.02% bias and about 7% error.

### 3.2 Copula-based HT-DBH Distributions

Table 5 shows the differences in the mixed two-sample index of goodness of fit ($I_{NN}$) between field HT-DBH distribution and the 100 simulated copula-based HT-DBH distributions. While differences in marginal distributions were observed, the resulting copulas produced similar ranges and shapes for the estimated HT-DBH distributions. Around 60% of the plots had average $I_{NN}$ values between 0.60 and 0.75. and 30% of the plots fall in the interval of $I_{NN}$ values between 0.75 and 0.90. Only 10% of the plots had $I_{NN}$ values close to 0.50.

Figure 3 shows the HT-DBH distributions for field measurements and copulas-based models for 9 field plots. Most copula-based models predicted the HT-DBH distributions well in the range of 10–30cm DBH and 5–20m HT. Both smaller (DBH < 10cm and/or HT < 5m) and larger trees (DBH > 35cm and/or HT > 20m) were not frequently predicted, especially the larger trees (Fig. 3).

Bimodality in DBH distributions impacted $I_{NN}$ (Figs. 3D–3F). Plots with bimodal distributions tended to have large values of $I_{NN}$ indicating a higher probability of a nearest neighbor being a copula sample rather than a field measure. The copulas simulated here were continuous distributions. If field measures were discontinuous, then it was more likely to have a copula simulated value nearer than a field value.

Variation in HT also had greater influence on $I_{NN}$ than did variation in DBH (Figs. 3G–3I). Plots with large variation in HTs for a given diameter often had $I_{NN}$ values > .80. Large values of $I_{NN}$ indicate that field observations have a higher probability of having a nearest neighbor that belongs to the copula-based HT-DBH distribution than to the field distribution (i.e., the two distributions are distinctly different). For these plots, the DBHs tended to be evenly dispersed between 6 and 30cm rather than concentrated in large or small size classes. While DBHs were evenly distributed, HTs exhibited large levels of variation within DBH classes. Plots with the better performance were ones which had well-defined modal distributions of DBH with narrower HT distributions (Figs. 3A–3C).

## 4 Discussion

While LiDAR provides a rich 3D data structure, the most intuitive metrics utilized in many studies are those related to individual tree heights (Ayrey et al., 2019). Although LiDAR cannot directly extract individual tree diameters in the same manner as heights are extracted, we have shown that this important size variable can be estimated from LiDAR using moment-based parameter recovery methods and copulas. Using the Femelschlag data and LiDAR attributes, a set of predictors for the required moments (Table 3) were developed and the resulting estimates generally fit the observed data well in terms of root mean square errors but much variation remained unexplained (Table 3). The goodness-of-fits of the resulting HT-DBH copulas were more varied (Table 5). Plots with bimodal DBH distributions generally had the poorest performance (Fig. 3). Arias-Rodil et al.'s (2018) model for estimating $D_q$ explained

Table 3: Parameter estimates, standard errors (in parentheses), pseudo-$R^2$ for estimation of Weibull moments for DBH distributions, Beta moments for HT distributions, number of trees per ha (TPH) and correlation between DBH and height (DHcor) using LiDAR data.

| Attribute | Parameter Estimate | | | | Regression Summary | |
|---|---|---|---|---|---|---|
| | $b_0$ | $b_1$ | $b_2$ | $b_3$ | $rMSE$ | Pseudo-$R^2$ |
| $D_q$ | 0.7021 (0.3571) | -0.2015 (0.0656) | 1.4345 (0.2291) | - | 2.7529 | 0.3437 |
| $s(D)$ | 0.0394 (0.0429) | 1.854 (0.3813) | - | - | 1.6975 | 0.2292 |
| $HT_q$ | 2.3182 (1.3005) | -0.7066 (0.3843) | -0.1804 (0.0494) | 1.6375 (0.3079) | 1.3347 | 0.4734 |
| $s(HT)$ | 0.0115 (0.0081) | -0.1958 (0.0533) | 2.3559 (0.2745) | - | 0.7032 | 0.4914 |
| $HT_{max}$ | 1.5102 (0.7354) | -0.5401 (0.2777) | 1.5101 (0.3593) | - | 2.5058 | 0.2813 |
| $TPH$ | 3122279.5 (3142066) | 0.6305 (0.192) | -3.4797 (0.418) | 0.6688 (0.2886) | 568.712 | 0.4972 |
| $DH_{cor}$ | -0.3206 (0.5242) | 0.2795 (0.0653) | -0.2353 (0.0696) | - | 0.0582 | 0.1758 |

Table 4: Mean bias (observed – predicted), root mean square error (rMSE) for parameters of diameter, height, density and correlation for LiDAR data.

| Attribute | Parameter | bias | $rMSE$ |
|---|---|---|---|
| Diameter | QuadMean ($D_q$) | 0.0032 | 2.7529 |
| | Std Dev ($s(D)$) | 0.0025 | 1.6975 |
| | Shape | -0.1847 | 0.8607 |
| | Scale | -2.2715 | 6.5327 |
| Height | QuadMean ($H_q$) | 0.0013 | 1.3348 |
| | Std Dev ($s(H)$) | 0.0005 | 0.7032 |
| | Maximum ($H_{max}$) | 0.0012 | 2.5058 |
| | $\lambda_1$ | 0.0873 | 0.9900 |
| | $\lambda_2$ | 0.1686 | 1.3393 |
| Density | $TPH$ | 10.1205 | 550.2932 |
| Correlation | $HD_{cor}$ | -0.0001 | 0.0582 |

Table 5: Numbers and percentages for mean INN from 100 simulations per plot (total = 83) and for each $I_{NN}$ from each simulation (total = 83*100) by different classes of $I_{NN}$

| Class | Mean $I_{NN}$ | | Individual $I_{NN}$ | |
|---|---|---|---|---|
| | Num. | Percent | Num. | Percent |
| 0.85 - 0.90 | 4 | 5% | 565 | 7% |
| 0.80 - 0.85 | 9 | 11% | 845 | 10% |
| 0.75 - 0.80 | 13 | 16% | 1255 | 15% |
| 0.70 - 0.75 | 15 | 18% | 1482 | 18% |
| 0.65 - 0.70 | 22 | 27% | 2102 | 25% |
| 0.60 - 0.65 | 13 | 16% | 1321 | 16% |
| 0.55 - 0.60 | 6 | 7% | 508 | 6% |
| 0.50 - 0.55 | 0 | 0% | 132 | 2% |
| 0.45 - 0.50 | 1 | 1% | 90 | 1% |
| Total | 83 | 100% | 8300 | 100% |

80% of the observed variability with a $rMSE$ value of 3.42cm. Although less variation was explained in our results, our $rMSE$ value for the $D_q$ equations were smaller (2.75cm; Table 3). These discrepancies may be partially explained by differences in stand structure and individual tree size. The mature forest stands of *P. radiata*, a shade-intolerant conifer species, in Galicia were the main data used for Arias-Rodil et al.'s (2018) study. Our forest structures were more complex leading to lower percent variation explained; however, when the $rMSE$s are expressed on a per cent of mean $D_q$, the two studies have almost equivalent errors (13.8% for *P. radiata* versus 15.8% for the trees used in this study). Treitz et al. (2012) studied a broad range of forest types and conditions across Ontario and reported that $rMSE$s can range from 0.76 to 4.3cm for $D_q$, and our results fall well within this range.

Terrestrial LiDAR scanning (TLS) is an alternative to ALS for estimating tree diameters. Some TLS studies report diameter errors of less than 1cm (Henning and Radtke, 2006; Liang et al., 2014; Olofsson and Holmgren, 2016; Pitkänen et al., 2019), other studies more commonly report errors in the 1–4cm range (de Conto et al., 2017; Maas et al., 2008; Pueschel et al., 2013; Srinivasan, 2015). TLS remains an expensive technology, though mobile TLS technology is quickly evolving

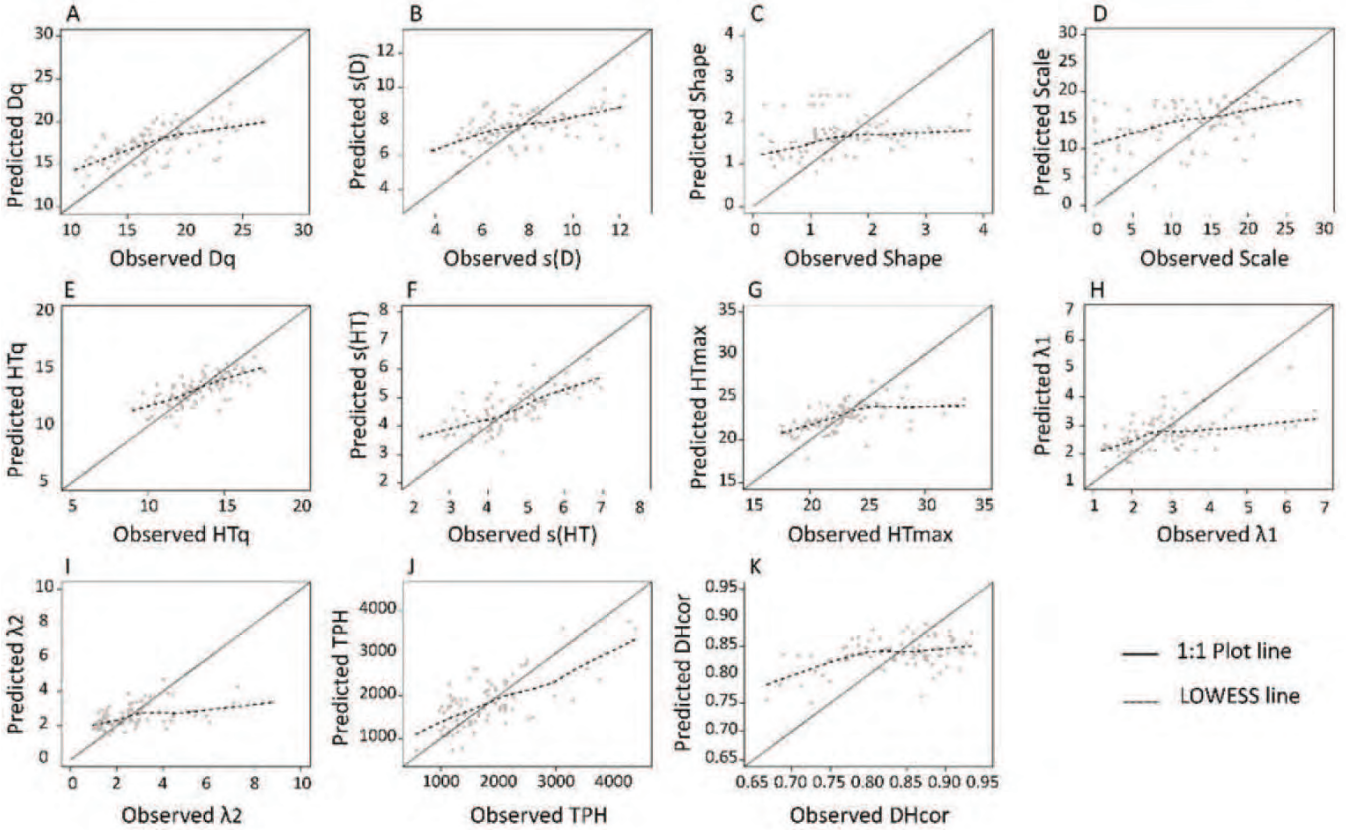


Figure 2: Predicted versus observed parameters: A) quadratic mean diameter ($D_q$, cm); B) diameter standard deviation ($s\,(D)$); C) shape; D) scale; E) quadratic mean height ($HT_q$, m); F) height standard deviation ($s\,(HT)$); G) maximum height ($HT_{max}$, m); H) $\lambda_1$; I) $\lambda_2$; J) trees per hectare ($TPH$, $\#\ ha^{-1}$); K) correlation between diameter and height, $DH_{cor}$. Solid line is the 1:1 line and the dashed line is the LOWESS smooth trend line.

as are the algorithms to detect tree stems and fit functions to tree stems. However, though this technology has great promise, complex forest stand structures are likely to continue to pose challenges for accurate diameter and stem profile estimation.

Because the models fitted the Femelschlag plots marginally (Table 3), substantial differences ($MB$ and $rMSE$) between observed and predicted moments and recovered distribution parameters for DBH and HT were observed (Table 4). Most discrepancies between observed and predicted values were associated with the bigger trees (Fig. 2). In general, trees with DBHs > 30cm and/or taller than 20m were not predicted leading to underestimation of mean DBH and HT. For some plots, smaller trees (DBH < 10cm) and shorter trees (HT<8m) also were not well predicted (Fig. 2). Other studies also have observed that ALS tends to underestimate HT (Maltamo et al., 2004) in both coniferous and deciduous forests because many ALS pulses do not hit the accrual tree top (Næsset and Økland, 2002) and these pulses penetrate into the canopy before a first echo is detected (Gaveau and Hill, 2003). Mehtätalo et al. (2007) obtained accuracies for estimation of mean tree DBH and HT on a stand level with a rMSE of 2.35cm and 1.22m, similar to what was achieved in this study. Overall, reasonable approximations of mean DBH and HT have generally been obtained using ALS data. The quality of predictions are dependent upon the quality of ground data (Hayashi et al., 2015; Hsu et al., 2020; Yang et al., 2019). Since large amounts of ground data are time consuming and expensive to collect, field-measured DBH and HT for fitting are often only available from a small number of sample plots. Large errors can be created during the process of extrapolation from plots to LiDAR cells without adequate field measurements with sound underlying sampling designs (Hayashi et al., 2015, 2016). Our data represents a very complete and intensive dataset for LiDAR analysis (Ver Planck et al., 2018).

Despite differences in estimated and mean values of the population and distribution parameters, the HT-DBH copulas, overall, performed satisfactorily for many of the stand structures encountered in the data used in

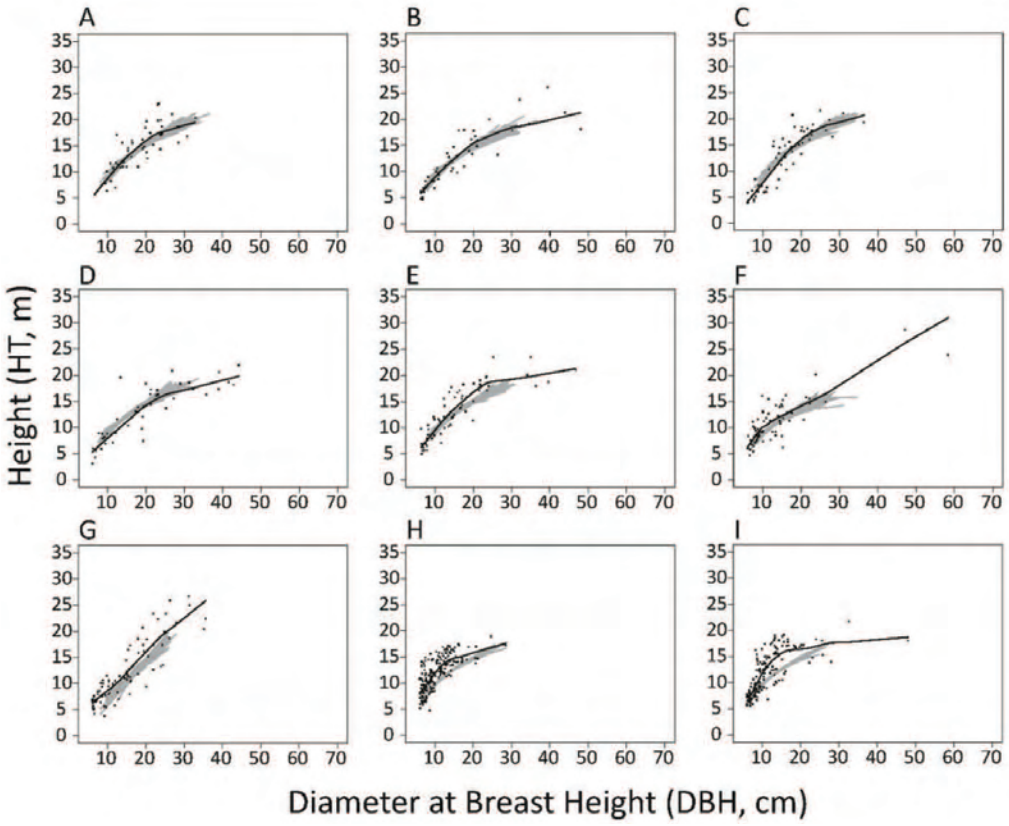


Figure 3: Comparison of copula-based distributions (gray lines represent individual simulations) for a selected plot from each class of $I_{NN}$ value relative to the observed values (black dots) and distribution (black line). The interval of $I_{NN}$ for each class was: A) 0.45 - 0.50; B) 0.50 - 0.55; C) 0.55 - 0.60; D) 0.60 - 0.65; E) 0.65 - 0.70; F) 0.70 - 0.75; G) 0.75 - 0.80; H) 0.80 - 0.85; and I) 0.85 – 0.90.

this study (Fig. 3 and Table 5). Approximately 50% of the plots had copula distributions that were more identical to the field measured HT-DBH distributions as measured by the mixed two-sample goodness-of-fit ($I_{NN} <$ 0.70; Table 5). These results were better than those obtained by Arias-Rodil et al. (2018) where the predicted diameter distributions performed very poorly (less than 40 % of their plot distributions were acceptable based on KS tests) even though their models accounted for a greater proportion of the variation than the models developed here.

Several plots having much higher $I_{NN}$ values were associated with plots that had high proportions of big or small trees, high variation in heights, and discontinuous distributions (especially DBH distributions, Fig. 3). Xu et al. (2014) encountered similar problems and developed ratio calibrations to correct LiDAR-derived heights using field measures. Recently, Xu et al. (2019) applied empirical copulas to estimate HT-DBH relationships. They divided the heights into several ranges and obtained empirical copulas within each range. Their approach produced better fits than obtained here, but relied on identification of nearest neighbors from which to build copulas (cf., Kershaw et al., 2017; McGarrigle et al., 2013) rather than modeling and predicting copulas as done here. For those plots with much higher $I_{NN}$ values, two situations consistently were observed: presence of several tall trees and multistoried stands. In this study we attempted to account for emergent trees in our estimation of height distribution parameters by using an indicator variable for their presence. This approach did not significantly ($p > .05$) improve our predictions and was subsequently dropped. This study also did not account for species composition. Both Hayashi et al. (2015) and Ver Planck et al. (2018) fitted forest-type-specific models to the Noonan data and found improvements in goodness-of-fit and model predictions. Because we used a moment-based parameter recovery approach, we had hoped to avoid the need for recognition of different stand types; however, the different stand types at Noonan, while similar in area-based estimates, have very different DBH and HT distributions.

Many of the Femelschlag plots were not pure stands so that both DBH and HT distributions can be multimodal,

disjointed, or highly variable in other ways (Fig. 3). Unimodal distribution functions cannot adequately capture these distributions (Liu et al., 2002). Those multimodal distributions might have been better fitted using mixture distributions (Liu et al., 2002) or one of the Johnson's translational distributions (Hafley and Schreuder, 1977). Despite these limitations, our models did fit a majority of the plots in this study well (Table 5). Improvement in fits to the other plots might be accomplished by using more flexible distribution functions or by broadening the types of LiDAR metrics to include additional measures of horizontal heterogeneity in LiDAR point clouds (Ayrey et al., 2019). Measures of horizontal heterogeneity may require examining the point clouds beyond the bounds of the relatively small plots used in this study (accepting the errors associated with spatially extracting beyond the plot boundaries) or increasing plot size (accepting the errors associated with reduced sample sizes). Regardless of which approach is selected, it will be important and required to ensure that the data represent a valid sample which covers the ranges of conditions well over which the LiDAR-derived models are going to be applied (Hayashi et al., 2015; Yang et al., 2019).

The use of copulas, as presented here, has potential for wide spatial scale prediction of tree-level attributes and represents, for the time being, a potentially more promising approach than trying to extract individual tree dimensions directly from point cloud attributes. The moment-based parameter recovery approach used here is robust since the estimation of mean tree dimensions is generally easier than the direct estimation of distribution parameters which may not have geometric interpretation (Burk and Newbury, 1984; Gove, 2003; Kershaw and Maguire, 1996). An additional potential advantage of the methods developed here is that careful coregistration of LiDAR point clouds and individual trees may not be required, since the emphasis is on population distributions rather than point estimates of individual trees. Copulas, as shown here, are more flexible than traditional bivariate distributions (Schreuder and Hafley, 1977), regression equations (MacPhee et al., 2018), and imputation methods (MacPhee et al., 2018). However, the results do indicate the need for adequate sample designs that cover the full range of conditions over which the models are applied and distribution functions that are flexible enough to capture the full range of structures being modelled. So far, the required moments of DBH and HT distributions were estimated independently from LiDAR data. Applying a compatible systems approach (Yang et al., 2021) would provide high consistency among moments and might improve the DBH-HT copulas models.

With improved HT-DBH distributions, other tree attributes such as volume or carbon content can be estimated and summed to obtain improved area-based estimates. In addition to obtaining better quality of field data across a wider range of field plots, terrestrial LiDAR scanning systems may provide additional potential to improve individual tree estimation because these systems quickly gather more detailed information from the understory perspective which may, in turn, prove useful for improving LiDAR-derived copulas such as those developed here. The combination of terrestrial LiDAR systems and airborne LiDAR-based analyses such as developed in this study is a relatively unexplored approach and holds much potential for improved inventory estimates.

## Acknowledgements

We are grateful to the dedicated work of the 2014 summer field crew: Mike Hutchinson, Shawn Donovan, and Dan Kilham. Funding for this project was provided, in part, by grants from the Natural Sciences and Engineering Research Council of Canada (Discovery Grant program) and the New Brunswick Innovation Foundation Research Assistantship Initiative. The first author also thanks the University of New Brunswick for the President's Doctoral Tuition Award and the Fundy Community Foundation for their support through the Dorothy and Kenneth Langmaid Forest Soils and Tree Growth Scholarship.

## References

Arias-Rodil, M., Diéguez-Aranda, U., Álvarez-González, J.G., Pérez-Cruzado, C., Castedo-Dorado, F., and González-Ferreiro, E. 2018. Modeling diameter distributions in radiata pine plantations in Spain with existing countrywide LiDAR data. Ann. For. Sci. 75: 36. doi: https://doi.org/10.1007/s13595-018-0712-z

Asner, G.P., Hughes, R.F., Mascaro, J., Uowolo, A.L., Knapp, D.E., Jacobson, J., Kennedy-Bowdoin, T., and Clark, J.K. 2011. High-resolution carbon mapping on the million-hectare Island of Hawaii. Front. Ecol. Environ. (e-View): 28. doi: https://doi.org/10.1890/100179

Ayrey, E., Fraver, S., Kershaw, J.A., Jr., Kenefic, L.S., Hayes, Weiskittel, A.R., and Roth. 2017. Layer stacking: A novel algorithm for individual forest tree segmentation from LiDAR point clouds. Can. J. Remote Sens. 43(1): 16–27. doi: https://doi.org/10.1080/07038992.2017.1252907

Ayrey, E., Hayes, D.J., Fraver, S., Kershaw, J.A., Jr., and Weiskittel, A.R. 2019. Ecologically-Based Metrics for Assessing Structure in Developing Area-Based,

Enhanced Forest Inventories from LiDAR. Can. J. Remote Sens.: 45(1): 88-112. doi: https://doi.org/10.1080/07038992.2019.1612738

Bailey, R.L., and Dell, T.R. 1973. Quantifying diameter distributions with the Weibull function. For. Sci. 19(2): 97–104. doi: https://doi.org/10.1080/07038992.2019.1612738

Borders, B.E., and Patterson, W.D. 1990. Projecting stand tables: A comparison of the Weibull diameter distribution method, a percentile-based projection method, and a basal area growth projection method. For. Sci. 36(2): 413–424. doi: https://doi.org/10.1093/forestscience/36.2.413

Brandtberg, T., Warner, T.A., Landenberger, R.E., and McGraw, J.B. 2003. Detection and analysis of individual leaf-off tree crowns in small footprint, high ampling density lidar data from the eastern deciduous forest in North America. Remote Sens. Environ. 85: 290–303. doi: https://doi.org/10.1016/S0034-4257(03)00008-7

Breiman, L. 2001. Random forests. Mach. Learn. 45(1): 5–32.

Burk, T.E., and Newberry, J.D. 1984. A simple algorithm for moment-based recovery of Weibull distribution parameters. For. Sci. 30(2): 329–332. doi: https://doi.org/10.1093/forestscience/30.2.329

Bury, K.V. 1975. Statistical Models in Applied Science. Robert E. Krieger Publishing Company, INC., Malabar, Florida.

Cleveland, W.S. 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. J. Am. Stat. Assoc. 74(368): 829–836.

Clutter, J.L., Fortson, J.C., Pienaar, L.V., Brister, G.H., and Bailey, R.L. 1983. Timber management. A quantitative approach. First. John Wiley & Sons, New York.

Culvenor, D.S. 2002. TIDA: an algorithm for the delineation of tree crowns in high spatial resolution remotely sensed imagery. Comput. Geosci. 28(1): 33–44. doi: https://doi.org/10.1016/S0098-3004(00)00110-2

de Conto, T., Olofsson, K., Görgens, E.B., Rodriguez, L.C.E., Almeida, G. 2017.Performance of stem denoising and stem modelling algorithms on single tree point clouds from terrestrial laser scanning. Comput. Electron. Agric., 143:165-176. doi: https://doi.org/10.1016/S0098-3004(00)00110-2

Frees, E.W., and Valdez, E. 1998. Understanding relationships using copulas. North Am. Actuar. J. 2(1): 1–25. doi: https://doi.org/10.1080/10920277.1998.10595667

Gaveau, D.L.A., and Hill, R.A. 2003. Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. Can. J. Remote Sens. 29(5): 650–657. doi: https://doi.org/10.5589/m03-023

Genest, C., and MacKay, J. 1986. The Joy of Copulas: Bivariate distributions with uniform marginals. Am. Stat. 40: 280–283. doi: https://doi.org/10.1080/00031305.1986.10475414

Gobakken, T., and Næsset, E. 2004. Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. Scand. J. For. Res. 19(6): 529–542. doi: https://doi.org/10.1080/02827580410019454

Gove, J.H. 2003. Moment and maximum likelihood estimators for Weibull distributions under length- and area-biased sampling. Envvironmental Ecol. Stat. 10(4): 455–467. doi: https://doi.org/10.1023/A:1026000505636

Hafley, W.L., and Schreuder, H.T. 1977. Statistical distributions for fitting diameter and height data in even-aged stands. Can. J. For. Res. 7(3): 481–487. doi: https://doi.org/10.1139/x77-062

Hayashi, R., Kershaw, J.A., Jr., and Weiskittel, A.R. 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. For. Nat. Resour. Sci. 7(2): 49–65. available from: https://mcfns.net/index.php/Journal/article/view/MCFNS7.2_2 [accessed Nov. 23, 2021]

Hayashi, R., Weiskittel, A.R., and Kershaw, J.A., Jr. 2016. Influence of prediction cell size on LiDAR-derived area-based estimates of total volume in mixed-species and multicohort forests in northeastern North America. Can. J. Remote Sens. 42(5): 473–488. doi: https://doi.org/10.1080/07038992.2016.1229597

Henning, J.G. and Radtke, P.J. 2006. Detailed stem measurements of standing trees from ground-based scanning lidar. For. Sci., 52(1):67-80, doi: https://doi.org/10.1093/forestscience/52.1.67

Hsu, Y.-H., Yang, T.-R., Chen, Y., and Kershaw, J.A., Jr. 2020. Sample strategies for bias correction of regional LiDAR-assisted forest inventory estimates on small woodlots. Ann. For. Sci. 77:75, 12p. doi: https://doi.org/10.1007/s13595-020-00976-8

Hummel, S., Hudak, A.T., Uebler, E.H., Falkowski, M.J., and Megown, K.A. 2011. A comparison of accuracy and cost of LiDAR versus stand exam data for landscape management on the Malheur National Forest. J. For. 109(5): 267–273. doi: https://doi.org/10.1093/jof/109.5.267

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B.M., and Wu, J.-C. 2012. An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. Remote Sens. 4(4): 950–974. doi:https://doi.org/10.3390/rs4040950.

Kershaw, J.A., Jr., and Maguire, D.A. 1996. Crown structure in western hemlock, Douglas-fir, and grand fir in western Washington: horizontal distribution of foliage within branches. Can. J. For. Res. 26(1): 128–142. doi: https://doi.org/10.1139/x26-014

Kershaw, J.A., Jr., Richards, E.W., McCarter, J.B., and Oborn, S. 2010. Spatially correlated forest stand structures: A simulation approach using copulas. Comput. Electron. Agric. 74(1): 120–128. doi: https://doi-org.proxy.hil.unb.ca/10.1016/j.compag.2010.07.005

Kershaw, J.A., Jr., Weiskittel, A.R., Lavigne, M.B., and McGarrigle, E. 2017. An imputation/copula-based stochastic individual tree growth model for mixed species Acadian forests: a case study using the Nova Scotia permanent sample plot network. For. Ecosyst. 4: 15. doi: https://doi.org/10.1186/s40663-017-0102-2

Kutner, M., Nachtsheim, C., and Neter, J. 2004. Applied linear regression models. Fourth Ed. McGraw-Hill, New York.

Li, W., Guo, Q., Jakubowski, M., and Kelly, M. 2012. A new method for segmenting individual trees from the lidar point cloud. Photogramm. Eng. Remote Sens. 78(1): 75–84. doi: https://doi.org/10.14358/PERS.78.1.75

Liang, X., Kankare, V., Yu, X., Hyyppä, Y., and Holopainen, M. 2014. Automated Stem Curve Measurement Using Terrestrial Laser Scanning. IEEE Trans. Geosci. Remote Sens., 52(3):1739-1748. doi: https://doi.org/10.1109/TGRS.2013.2253783

Little, S.N. 1983. Weibull diameter distributions for mixed stands of western conifers. Can. J. For. Res. 13(1): 85–88. doi: https://doi.org/10.1139/x83-012

Liu, C., Zhang, L., Davis, C.J., Solomon, D.S., and Gove, J.H. 2002. A finite mixture model for characterizing the diameter distributions of mixed-species forest stands. For. Sci. 48(4): 653–661. doi: https://doi.org/10.1093/forestscience/48.4.653

Loo, J., and Ives, N. 2003. The Acadian Forest: Historical condition and human impacts. For. Chron. 79(3): 462–474. doi: https://doi.org/10.5558/tfc79462-3

Maas, H., Bienert, A., Scheller, S., and Keane, E. 2008. Automatic forest inventory parameter determination from terrestrial laser scanner data. Int. J. Remote Sens., 29(5):1579-1593. doi: https//doi.org/10.1080/01431160701736406

MacPhee, C., Kershaw, J.A., Jr., Weiskittel, A.R., Golding, J., and Lavigne, M.B. 2018. Comparison of approaches for estimating individual tree height–diameter relationships in the Acadian forest region. Forestry 92: 132–146. doi: https://doi.org/10.1093/forestry/cpx039

Maltamo, M., Mustonen, K., Hyyppä, J., Pitkänen, J., and Yu, X. 2004. The accuracy of estimating individual tree variables with airborne laser scanning in a boreal nature reserve. Can. J. For. Res. 34: 1791–1801. doi: https://doi.org/10.1139/x04-055

Maltamo, M., Peuhkurinen, J., Malinen, J., Vauhkonen, J., Packalén, P., and Tokola, T. 2009. Predicting tree attributes and quality characteristics of Scots pine using airborne laser scanning data. Silva Fenn. 43(3). doi: https://doi.org/10.14214/sf.203.

McGarrigle, E., Kershaw, J.A., Jr., Ducey, M.J., and Lavigne, M.B. 2013. A new approach to modeling stand-level dynamics based on informed random walks: Influence of bandwidth and sample size. Forestry 86(3): 377–389. doi: https://doi.org/10.1093/forestry/cpt008

Mehtätalo, L., Maltamo, M., and Packalén, P. 2007. Recovering plot-specific diameter distribution and height-diameter curve using als based stand characteristics. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XXXVI: 288–293.

Næsset, E., and Økland, T. 2002. Estimating tree height and tree crown properties using airborne scanning

laser in a boreal nature reserve. Remote Sens. Environ. 79: 105–115. doi: https://doi.org/10.1016/S0034-4257(01)00243-7

Narsky, I. 2003a. Goodness of fit: What do we really want to know? In Proceedings. Physics Statistics 2003. pp. 70–74. available from: http://www.slac.stanford.edu/econf/C030908/papers/MOCT004.pdf [accessed Nov. 23, 2021]

Narsky, I. 2003b. Estimation of goodness-of-fit in multidimensional analysis using distance to nearest neighbor. High Energy Physics, California Technical University. available from: https://arxiv.org/pdf/physics/0306171 [accessed Nov. 23, 2021]

Nelsen, R.B. 2006. An Introduction to Copulas. 2nd Ed. Springer, New York.

Olofsson, K. and Holmgren, J. 2016. Single tree stem profile detection using terrestrial laser scanner data, flatness saliency features and curvature properties. Forests, 7(9): 207. doi: https://doi.org/10.3390/f7090207

Palahí, M., Pukkala, T., Blasco, E., and Trasobares, A. 2007. Comparison of beta, Johnson's SB, Weibull and truncated Weibull functions for modeling the diameter distribution of forest stands in Catalonia (northeast of Spain). Eur. J. For. Res. 126(4): 563–571. doi: https://doi.org/10.1007/s10342-007-0177-3

Pitkänen, T.P., Raumonen, P. and Kangas, A. 2019. Measuring stem diameters with TLS in boreal forests by complementary fitting procedure. ISPRS Journal of Photogrammetry and Remote Sensing. 147 :294-306. doi: https://doi.org/10.1016/j.isprsjprs.2018.11.027

Plotnick, R.E., Gardner, R.H., and O'Neill, R.V.. 1993. Lacunarity indices as measures of landscape texture. Landsc. Ecol. 8(3):201–211. doi: https://doi.org/10.1007/BF00125351

Pueschel, P., Newnham, G., Rock, G., Udelhoven, T., Werner, W., and Hill, J. 2013. The influence of scan mode and circle fitting on tree stem detection, stem diameter and volume extraction from terrestrial laser scans. ISPRS J. Photogramm. Remote Sens., 77:44-56. doi: https://doi.org/10.1016/j.isprsjprs.2012.12.001

R Development Core Team. 2021. R: A Language and Environment for Statistical Computing. Available from http://www.R-project.org [accessed Nov. 23, 2021]

Reutebuch, S.E., Andersen, H.-E., and McGaughey, R.J. 2005. Light Detection and Ranging (LIDAR): An Emerging Tool for Multiple Resource Inventory. J. For. 103: 286–292.

Ripley, B.D. 2004. Spatial statistics. Wiley, Hobokin, NJ.

Roussel, J.-R., Auty, D., Boissieu, F.D., Meador, A.S., and Jean-François, B. 2020. Airborne LiDAR Data Manipulation and Visualization for Forestry Applications. available from: https://github.com/Jean-Romain/lidR [accessed Nov. 23, 2021]

Schilling, M.F. 1986. Multivariate two-sample tests based on nearest neighbors. J. Am. Stat. Assoc. 81(395): 799–806. doi: https://doi.org/10.1080/01621459.1986.10478337

Schreuder, H.T., and Hafley, W.L. 1977. A useful bivariate distribution for describing stand structure of tree heights and diameters. Biometrics 33(3): 471–478. doi:https://doi.org/10.2307/2529361

Srinivasan, S., Popescu, C.S., Eriksson, M., Sheridan, D.R., and Ku, N. 2015. Terrestrial laser scanning as an effective tool to retrieve tree level height, crown width, and stem diameter. Remote Sens., 7(2):1877-1896. doi: https//doi.org/10.3390/rs70201877

Temesgen, H., and Ver Hoef, J.M. 2015. Evaluation of the spatial linear model, random forest and gradient nearest-neighbour methods for imputing potential productivity and biomass of the Pacific Northwest forests. Forestry 88(1): 131–142. doi: https://doi.org/10.1093/forestry/cpu036

Thomas, V., Oliver, R.D., and Woods, M. 2008. LiDAR and Weibull modeling of diameter and basal area. For. Chron. 84(6): 866–875. doi: https://doi.org/10.5558/tfc84866-6

Treitz, P.; Lim, K.; Woods, M.; Pitt, D.; Nesbitt, D.; Etheridge, D. 2012. LiDAR sampling density for forest resource inventories in Ontario, Canada. Remote Sensing 4:830:848. doi: https//doi.org/10.3390/rs4040830

Ver Planck, N.R., Finley, A.O., Kershaw, J.A., Jr., Weiskittel, A.R., and Kress, M.C. 2018. Hierarchical Bayesian models for small area estimation of forest variables using LiDAR. Remote Sens. Environ. 264: 287–295. doi: http://dx.doi.org/10.1016/j.rse.2017.10.024

Wang, M., Rennolls, K., and Tang, S. 2008. Bivariate distribution modeling of tree diameters and heights: Dependency modeling using copulas. For. Sci.

54(3): 284–293. doi: https://doi.org/10.1093/forestscience/54.3.284

Wang, M., Upadhyay, A., and Zhang, L. 2010. Trivariate distribution modeling of tree diameter, height, and volume. For. Sci. 56(3): 290–300. doi: https://doi.org/10.1093/forestscience/56.3.290

Wang, S.S. 1998. Aggregation of correlated risk portfolios: Models & algorithms. Proc. Casualty Actuar. Soc. 85(163): 848–937. available from: https://www.casact.org/sites/default/files/database/proceed_proceed98_1998.pdf [accessed Nov. 23, 2021]

White, J.C., Wulder, M.A., and Varhola, A. 2013. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. Natural Resources Canada, Canadian Forest Service, Canadian Wood Fibre Centre, Victoria, BC, Canada. available from: https://cfs.nrcan.gc.ca/publications/download-pdf/34887 [accessed Nov. 23, 2021]

Wulder, M.A., Bater, C.W., Coops, N.C., Hilker, T., and White, J.C. 2008. The role of LiDAR in sustainable forest management. For. Chron. 84(6): 807–826. doi: https://doi.org/10.5558/tfc84807-6

Xu, Q., Hou, Z., Maltamo, M., and Tokola, T. 2014. Retrieving suppressed trees from model-based height distribution by combining high and low density airborne laser scanning data. Can. J. Remote Sens. 40(3): 233–242. doi: http://dx.doi.org/10.1080/07038992.2014.935933

Xu, Q., Li, B., Maltamo, M., Tokola, T., and Hou, Z. 2019. Predicting tree diameter using allometry described by non-parametric locally-estimated copulas from tree dimensions derived from airborne laser scanning. For. Ecol. Manag. 434: 205–212. doi: https://doi.org/10.1016/j.foreco.2018.12.020

Yan, J. 2007. Enjoy the joy of copulas: With a package copula. J. Stat. Softw. 21(4): 1–21. doi: https://doi.org/10.18637/jss.v021.i04

Yang, T.-R., Kershaw, J.A., Weiskittel, A.R., Lam, T.Y., and McGarrigle, E. 2019. Influence of sample selection method and estimation technique on sample size requirements for wall-to-wall estimation of volume using airborne LiDAR. Forestry 92(3): 311–323. doi: https//doi.org/10.1093/forestry/cpz014

Yang, T.-R., Kershaw, J.A., Jr., and Ducey, M.J., 2021. Development of allometric systems of Equations for Compatible Area-Based LiDAR-Assisted Estimation. Forestry. 94(1):36-53. doi: https://doi.org/10.1093/forestry/cpaa019

Zutter, B.R., Oderwald, R.G., Farrar, R.M., Jr., and Murphy, P.A. 1982. WEIBUL: A program to estimate parameters of forms of the Weibull distribution using complete, censored, and truncated data. Publication, Virginia Polytechnic Institute and State University, Blacksburg, VA, School of Forestry and Wildlife Resources. available from: https://vtechworks.lib.vt.edu/bitstream/handle/10919/93556/FWS-3-82.pdf [accessed Nov. 23, 2021]

Zutter, B.R., Oderwald, R.G., Murphy, P.A., and Farrar, R.M. 1986. Characterizing diameter distributions with modified data types and forms of the Weibull distribution. For. Sci. 32(1): 37–48. doi: https://doi.org/10.1093/forestscience/32.1.37