Taylor & Francis
Taylor & Francis Group

Canadian Journal of Remote Sensing
Journal canadien de télédétection

ISSN: 0703-8992 (Print) 1712-7971 (Online) Journal homepage: http://www.tandfonline.com/loi/ujrs20

Optimizing Variable Radius Plot Size and LiDAR Resolution to Model Standing Volume in Conifer Forests

Ram Kumar Deo, Robert E. Froese, Michael J. Falkowski & Andrew T. Hudak

To cite this article: Ram Kumar Deo, Robert E. Froese, Michael J. Falkowski & Andrew T. Hudak (2016) Optimizing Variable Radius Plot Size and LiDAR Resolution to Model Standing Volume in Conifer Forests, Canadian Journal of Remote Sensing, 42:5, 428-442, DOI: 10.1080/07038992.2016.1220826

To link to this article: http://dx.doi.org/10.1080/07038992.2016.1220826

Accepted author version posted online: 11 Aug 2016.
Published online: 11 Aug 2016.

Canadian Journal of Remote Sensing, 42:428–442, 2016

ISSN: 0703-8992 print / 1712-7971 online
DOI: 10.1080/07038992.2016.1220826

Taylor & Francis
Taylor & Francis Group

Optimizing Variable Radius Plot Size and LiDAR Resolution to Model Standing Volume in Conifer Forests

Ram Kumar Deo[1], Robert E. Froese[1,*], Michael J. Falkowski[2], and Andrew T. Hudak[3]

[1]School of Forest Resources and Environmental Science, Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931, USA
[2]Department of Ecosystem Science and Sustainability, Colorado State University, A204 NESB - Campus Delivery 1476, Fort Collins, CO 80523, USA
[3]USDA Forest Service, Rocky Mountain Research Station, 1221 South Main Street, Moscow, ID 83843, USA

Abstract. The conventional approach to LiDAR-based forest inventory modeling depends on field sample data from fixed-radius plots (FRP). Because FRP sampling is cost intensive, combining variable-radius plot (VRP) sampling and LiDAR data has the potential to improve inventory efficiency. The overarching goal of this study was to evaluate the integration of LiDAR and VRP data. FRP and VRP plots using different basal-area factors (BAF) were colocated in 6 conifer stands near Alberta, Michigan, in the United States. A suite of LiDAR metrics was developed for 24 different resolutions at each plot location, and a number of nonparametric prediction models were evaluated to identify an optimal LiDAR resolution and an optimal scale of VRP to spatially extend the data. An FRP-based model had root mean square error (RMSE) of 31.8 m^3 ha^{-1}, whereas the top VRP-based models were somewhat less precise, with RMSE of 38.0 m^3 ha^{-1} and 45.8 m^3 ha^{-1} using BAF 2.06 m^2 ha^{-1} and BAF 2.29 m^2 ha^{-1}, respectively. The optimal LiDAR resolution for the VRP data was found to be 18 m for the selected stands, and plot-level estimates based on a model using BAF 2.29 m^2 ha^{-1} were statistically equivalent to the FRP measurements. The use of VRP data shows promise and can substitute for FRP measurements to improve efficiency.

Résumé. L'approche conventionnelle de modélisation de l'inventaire forestier à base de LiDAR dépend des données d'échantillonnage sur le terrain par placettes à rayon fixe « variable-radius plot » (VRP). Comme l'échantillonnage par FRP est couteux, la combinaison de l'échantillonnage par placettes à rayon variable « variable-radius plot » (VRP) et des données LiDAR a le potentiel d'améliorer l'efficacité de l'inventaire. L'objectif global de cette étude était d'évaluer l'intégration des données LiDAR et par VRP. Les placettes de FRP et de VRP qui utilisent différents facteurs de surface terrière « basal-area factors » (BAF) ont été colocalisées dans 6 peuplements de conifères à proximité de l'Alberta, dans l'État du Michigan aux États-Unis. Une série de mesures LiDAR a été développée pour 24 résolutions différentes à chaque emplacement de placette et un certain nombre de modèles prédictifs non paramétriques ont été évalués pour identifier une résolution LiDAR optimale et une échelle optimale de PRV pour étendre spatialement les données. Un modèle à base de FRPF avait une « root mean square error » (RMSE) de 31,8 m^3 ha^{-1} tandis que les meilleurs modèles à base de PRV étaient un peu moins précis, avec une RMSE de 38,0 m^3 ha^{-1} et 45,8 m^3 ha^{-1} en utilisant un BAF de 2,06 m^2 ha^{-1} et un BAF de 2,29 m^2 ha^{-1}, respectivement. La résolution LiDAR optimale obtenue pour les données de VRP était de 18 m pour les peuplements sélectionnés, et les estimations au niveau de la placette sur la base d'un modèle en utilisant un BAF de 2,29 m^2 ha^{-1} étaient statistiquement équivalentes aux mesures de FRP. L'utilisation des données de VRP semble prometteuse et peut remplacer les mesures de FRP pour améliorer l'efficacité.

INTRODUCTION

Spatially explicit mapping of forest structural attributes such as standing tree volume and biomass has gained increased attention over the past few decades, especially in support of strategic and operational forest planning (Kellndorfer et al. 2010; Saatchi et al. 2011; Shugart et al. 2010). Because conventional forest inventories exclusively based on field sample plots are time and cost intensive, recent research endeavors have focused on developing and improving methods to spatially extend forest inventories across large extents by integrating sparse networks of field plot measurements with remote sensing or other geospatial data (Brosofske et al. 2014; Ohmann et al. 2014). Although many different types of remotely sensed data can be used, data from active sensors such as LiDAR (aka Light Detection and Ranging or Laser Altimetry) are popular, given their high sensitivity to 3-dimensional vegetation properties (Lefsky et al. 2002; Mitchell

Received 15 October 2015. Accepted 24 July 2016.
*Corresponding author e-mail: froese@mtu.edu.

et al. 2011; Sun et al. 2011). The applications of LiDAR data are expanding for multiple reasons, including the wider availability and coverage of acquisitions (Hudak et al. 2009), as well as the potential improvement in inventory cost-efficiency (Hummel et al. 2011). Indeed, LiDAR-derived metrics characterizing the vertical and horizontal distribution of vegetation (e.g., mean height, maximum height, canopy cover, etc.) have been used extensively in statistical modeling frameworks to characterize and estimate numerous structural attributes, including forest biomass and volume, among others (Dubayah et al. 2010; Latifi et al. 2010; Takagi et al. 2015).

LiDAR-based forest inventories often use inventory parameters collected at fixed-radius plots (FRP), where all trees within a specified plot radius are sampled and measured with equal (and constant) probability. FRP data are ideal when working with remotely sensed data, because the plot size can be approximately matched to the spatial resolution of the remotely sensed data, minimizing geometric inconsistencies. This is particularly true of LiDAR data, which provide 1–2 orders of magnitude more measurements for a given land area compared to passive optical imagery. For instance, a typical FRP can contain hundreds to thousands of discrete LiDAR returns (depending on data density and plot size) but only a small number of optical image pixels, depending on plot size and image spatial resolution (e.g., ~1 pixel–4 pixels for moderate-resolution imagery such as 30-m Landsat to ~5 pixels–10 pixels for high-resolution imagery such as WorldView or Quickbird). Hudak et al. (2008) demonstrated the importance of matching FRP size to LiDAR metric bin sizes. Specifically, Hudak et al. (2008) calculated LiDAR metrics from returns within 0.04 ha inventory plots (~400 m^2) to predict plot-level basal area by species, and subsequently mapped the predictions at a comparable 20-m × 20-m (400 m^2) resolution. Metrics derived from LiDAR returns binned within the plot footprint more accurately represented the forest structure within the FRP than metrics derived from 20-m × 20-m map cells (akin to image pixels) that only approximated the FRP footprint. Although plot-based modeling approaches generally produce more accurate results compared to pixel-based approaches, the collection of FRP data is costly, especially when collecting a sufficiently large sample of inventory plots that fully represent the range of forest structural and compositional variability across the area of interest.

Because the cost of an inventory is directly related to the size, number, and distribution of plots, forest managers often prefer employing inventory protocols that have high sampling efficiency. A common strategy to increase both statistical efficiency and cost-effectiveness is to employ a variable-radius plot (VRP) or point sampling scheme (Avery and Burkhart 2002; Husch et al. 2003). The VRP sampling scheme selects sample trees with probability proportional to basal area, and basal area is usually determined from the measurement of tree diameter at breast height (DBH) level (~1.33 m above ground). Because volume, which is an attribute of particular interest in most forest inventories, is proportional to a squared power of DBH, and large trees usually occur less frequently than small trees, this approach greatly improves sampling efficiency and produces unbiased inventory estimates (Packard and Radtke 2007). Because VRP selection probabilities are proportional to basal area, a single inventory plot has an unknown shape and an unknown area. This is problematic when performing LiDAR or other remote-sensing-based inventories because there are substantial geometric inconsistencies between the VRP data and the remotely sensed data (Figure 1). These geometric inconsistencies are almost certainly a major source of error in the statistical inventory models and final inventory maps derived from the integration of VRP and remote sensing data. The uncertainty of models is further increased when tree selection probabilities (i.e., basal area factor, BAF) are varied during VRP sampling. In practice, BAF for a stand inventory is typically optimized so that an efficient number of trees (usually 4–8) are selected for tally and measurement (Reed and Mroz 1997). In other words, if too many trees are encountered per plot, then inventory foresters typically use a larger BAF to reduce the sample set (i.e., select fewer trees) and enhance operational efficiency. As a result, the appropriate BAF to use can sometimes vary considerably among stands, especially in areas where there is substantial variability in forest structure. This adds an additional layer of spatial complexity and increases the geometric inconsistency between plot and remote sensing measures.

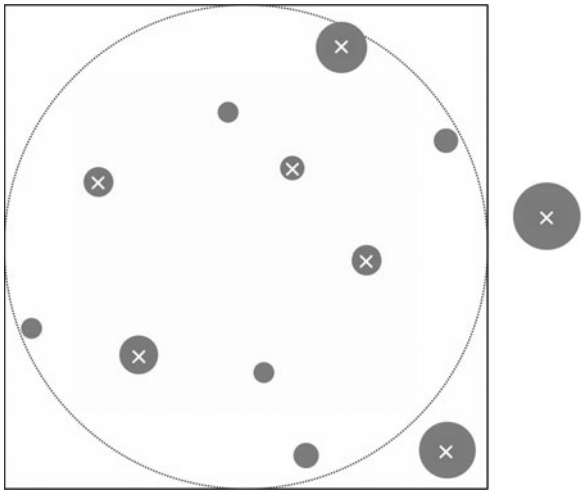

FIG. 1. An example layout of a fixed-radius and variable-radius plot with the same center. The black-square outline is a LiDAR bin or image pixel, the dotted gray circle is a fixed-radius plot (FRP), and the solid gray dots represent trees. All trees within the dotted gray circle are counted in FRP sampling but only the trees marked with a cross (×) are counted in variable-radius sampling at a given BAF.

A key challenge for improving the integration of VRP data with LiDAR data is to reduce the geometric inconsistency between the datasets. This often has involved optimizing the spatial resolution of the LiDAR-metrics to match the area sampled by

VRP data as close as possible (Golinkoff et al. 2011; Hollaus et al. 2007; Jochem et al. 2011). For example, Hollaus et al. (2009) attempted to reduce geometric inconsistency by arbitrarily choosing 4 different areas (16, 20, 24, and 28 m diameter) for calculating LiDAR metrics when working with VRP data collected with a BAF of 4 m^2 ha^{-1} $tree^{-1}$. Although they found very consistent precision across the 4 sizes, the best accuracy ($R^2 = 0.86$; root mean square error [RMSE] = 92.3 m^3 ha^{-1}) was attained with LiDAR metrics derived from a 20-m diameter area for volume prediction. In another study, Kronseder et al. (2012) calculated LiDAR-metrics in large 1-ha circular areas centered on each VRP when working with VRP data collected with a single BAF (4 m^2 ha^{-1} $tree^{-1}$). They reasoned that calculating the metrics in 1-ha circular areas would produce the best results because the VRP method employed provided inventory attribute estimates on a per-hectare basis. They also found high coefficient of determination ($R^2 = 0.83$) and a low error (RMSE = 96.74 Mg ha^{-1}) for biomass in lowland tropical forests. In an earlier study using the same field data, Hollaus et al. (2007) used 5 different resolutions of LiDAR predictors (at plots of 18, 20, 22, 24, and 26 m diameter) to develop the best relationship with VRP inventory and found that a diameter of 24 m led to the highest R^2 (= 0.84). Van Aardt et al. (2006) employed a stand-level mapping approach, opposed to pixel-level, by coupling LiDAR metrics at the stand level with stand-level VRP summaries based on a BAF = 2.29 m^2 ha^{-1} $tree^{-1}$ (i.e., 10 ft^2 ac^{-1} $tree^{-1}$) field inventory. When modeling and mapping volume and biomass, they reported moderate performance of the models, where adjusted R^2 ranged from 0.59–0.74 and 0.58–0.79 for volume and biomass models, respectively. Hudak et al. (2014) aggregated 30-m × 30-m binned LiDAR metrics sampled from within stands at the estimated VRP locations to relate to the stand-level inventory, based on VRP sampling. This predictive model was then applied to 30-m LiDAR grid metrics to generate maps of the stand structure attributes to represent within-stand structural variation in a spatially explicit manner, as conventional stand-level inventory data alone cannot do. Although these studies suggest high correlations between VRP and LiDAR data, they are limited in that they applied a single VRP sampling probability (i.e., BAF) in the field inventory. Further, the accuracy of models based on LiDAR data extracted for 1 ha at sample plot locations (e.g., Kronseder et al. 2012) likely results from the smoothing of derived predictors across such a large area, and this coarse spatial scale is not sufficient for most operational forest inventory needs. Indeed, comprehensive studies that integrate LiDAR data with VRP data collected at different sampling probabilities (i.e., different BAFs) in different forest types are lacking.

In an effort to improve the efficiency of LiDAR-based forest inventory, the primary goal of this study is to assess the efficacy of performing an inventory of multiple stands for standing volume via the integration of LiDAR and VRP data. Specifically, we conduct a comprehensive analysis of the influence of VRP sampling probability and LiDAR metric resolution on model performance by addressing the following questions: (*i*) what is the optimum LiDAR resolution to employ when working with VRP data collected with different sampling probabilities, (*ii*) conversely, what is the optimal BAF to use when conducting a LiDAR-based inventory, and (*iii*) how does model accuracy change if inventory attributes are modeled from LiDAR data by leveraging different proportions of FRP and VRP data? To achieve this, we employ LiDAR and VRP data to develop spatial models of standing volume and subsequently compare inventory predictions with independent field measurements.

METHODS

Study Area

The study was carried out in 6 conifer stands (Figure 2) at Michigan Technological University's Ford Center and Forest (FCF), located near the village of Alberta in the western Upper Peninsula of Michigan, in the United States (latitude 46° 37′ N, longitude 88° 29′ W). The core area of the FCF covers approximately 1400 ha and is roughly split between jack pine (*Pinus banksiana*) and hemlock–northern hardwood cover types, but also contains small patches of quaking aspen (*Populus tremuloides*) and natural red pine (*Pinus resinosa*) of fire origin. The coniferous cover types were the focus of this study and elevation of the sampled stands ranged from 359 m–425 m above sea level. Soils in the area originated from glacial outwash and belong primarily to the orders Spodosol and Entisol. The target stands differ in composition and structural complexity and include 1 old protected stand dominated by jack pine, 4 young-stage, even-aged pure jack pine stands, and 1 old-stage mixed uneven-aged dominant red pine stand (Table 1). These stands have not been harvested since 2007;however, thinning did occur in 1 stand in 1991 and in another in 2007.

Overview of Approach

The VRP-based modeling and mapping approach employed 4 major steps. First, VRP sampling was carried out in the sample stands using a prism of BAF 1.15 m^2 ha^{-1} (i.e., 5 ft^2 ac^{-1}), denoted hereafter as BAF 5. Tree DBH, species, as well as distance of each tally tree from plot center were recorded. Next, the VRP data were processed to derive new datasets corresponding to 6 larger BAF sizes (1.60 m^2 ha^{-1}, 2.06 m^2 ha^{-1}, 2.29 m^2 ha^{-1}, 2.75 m^2 ha^{-1}, 3.21 m^2 ha^{-1}, and 3.44 m^2 ha^{-1}, equivalent to 7 ft^2 ac^{-1}, 9 ft^2 ac^{-1}, 10 ft^2 ac^{-1}, 12 ft^2 ac^{-1}, 14 ft^2 ac^{-1}, and 15 ft^2 ac^{-1}, respectively) by counting the trees as "in" if the measured tree distance was less than or equal to the calculated limiting distance (Equation 1). In the third step, statistical models relating plot-level volume to LiDAR structural metrics were fit to LiDAR data extracted across a range of radii (24 different radii ranging from 7 m to 38 m; Figure 3). Finally, LiDAR metrics at the resolution of an optimum radius (for the most suitable BAF) were calculated across the entire study area and the statistical models were extended spatially to produce maps of forest

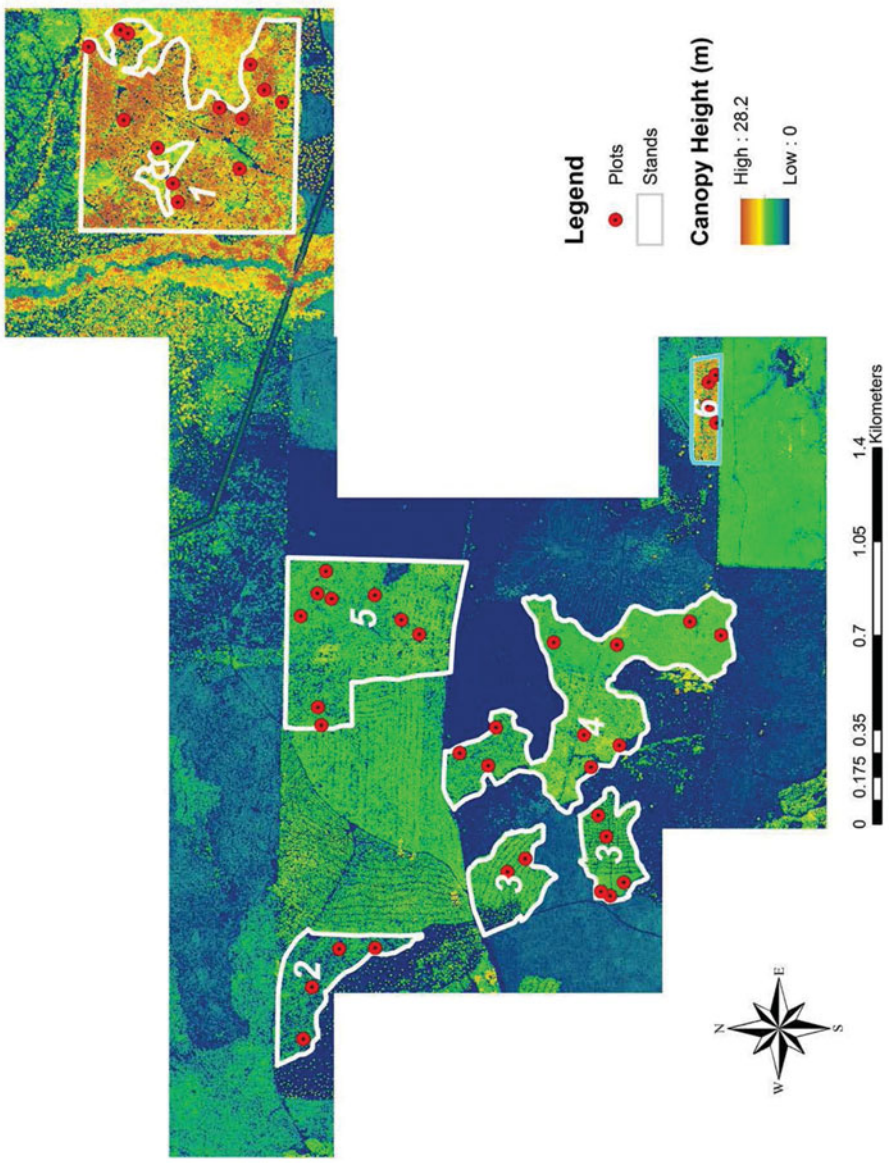

FIG. 2. The 6 target conifer stands across a range of development stages in the western Upper Peninsula of Michigan, USA. Stand boundaries and sample plot locations are overlaid on a LiDAR-derived canopy height model.

TABLE 1
Characteristics of the sample stands based on forest inventory measurements

Stand	No. of Plots	Area (ha)	Max. DBH (cm)	QMD (cm)*	Trees per ha	BAWHT (m)†	Mean Volume‡ (m^3 ha^{-1}) FRP	Mean Volume‡ (m^3 ha^{-1}) VRP BAF9	Dominant Species	Age Class
1	13	50.3	82.0	26.9	515	18.7	204.8(2394.6)	200.1(3649.9)	red pine	Old
2	4	10.1	23.8	17.8	412	12.4	46.7(33.8)	42.8(133.9)	jack pine	Young
3	7	14.2	23.6	14.8	803	11.1	40.6(297.3)	66.4(942.8)	jack pine	Young
4	10	33.1	32.2	15.9	840	11.8	65.1(164.1)	71.8(630.6)	jack pine	Young
5	9	32.1	27.6	15.3	961	11.4	68.2(178.5)	65.9(709.7)	jack pine	Young
6	4	4.1	35.5	22.6	729	14.9	160.5(684.3)	141.3(2602.6)	jack pine	Old

*QMD: quadratic mean diameter; † BAWHT: basal area weighted canopy height; ‡ numbers in the parentheses represent sample variance.

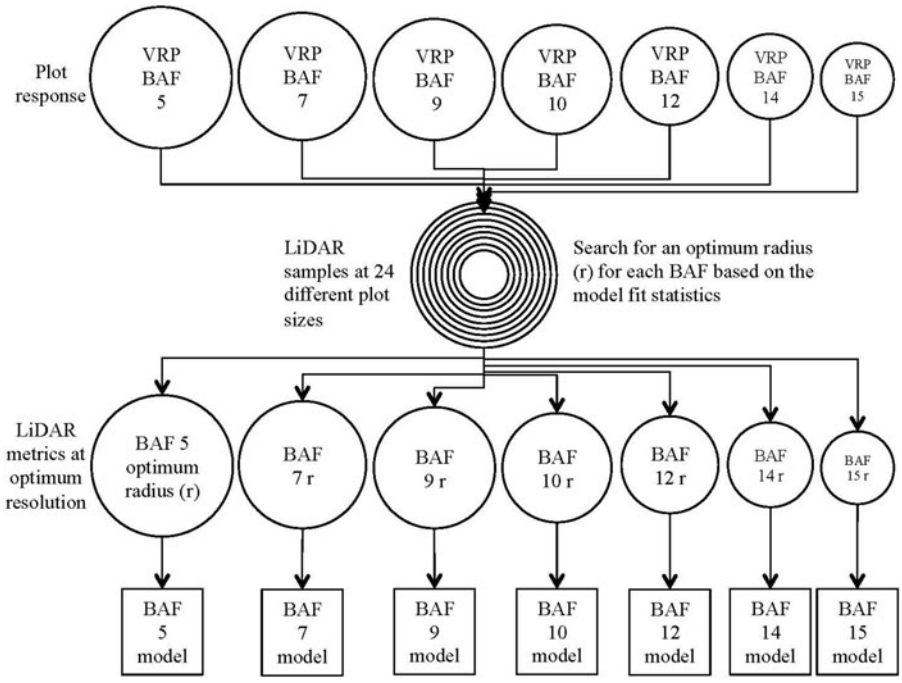

FIG. 3. The VRP-based modeling approach.

volume. These steps are described in more detail in the subsequent sections.

Field Inventory Data

The field inventory was carried out in the summer of 2012 over a network of permanent FRPs (0.04 ha, each) that were originally established by using a stratified random sampling design. Plot-center coordinates were obtained using a Trimble GeoXH 6000 GPS with differential correction postprocessing (via Trimble Pathfinder Office software) that resulted in an average horizontal precision of 0.80 m. Field data from both fixed-radius and variable-radius sampling schemes were obtained for a total of 47 plots across the 6 sampled stands. The number of plots per stand ranged from 4 to 13, depending on stand size, tree density, and stand heterogeneity, with a minimum sampling intensity of 1 plot per 3.8 ha (based on a maximum sampling error objective of 20%; Table 1). The FRP inventory was supplemented with a VRP inventory in September 2013, using a BAF of 1.15 $m^2\ ha^{-1}$ (BAF 5) at the existing plot locations. For the FRPs, tree species and DBH (for trees $\geq$ 10 cm) were recorded, and the total height of the smallest and largest trees of each species within the plot was measured using a Haglof Vertex Laser VL400 Hypsometer. For the VRPs, DBH and species were recorded for every tree selected with a BAF 5 prism, and the horizontal distance of each tally tree from the plot center (at breast height level) was also measured, using the laser hypsometer fixed on a tripod directly over plot center.

Additional VRP data corresponding to larger BAFs (i.e., smaller sampling probabilities) were derived from the BAF 5 data, using the limiting distance Equation 1. These larger BAF sizes corresponded to sampling probabilities of 1.60 $m^2\ ha^{-1}$, 2.06 $m^2\ ha^{-1}$, 2.29 $m^2\ ha^{-1}$, 2.75 $m^2\ ha^{-1}$, 3.21 $m^2\ ha^{-1}$ and 3.44 $m^2\ ha^{-1}$ (hereafter denoted as their imperial unit equivalents of BAF 7, BAF 9, BAF 10, BAF 12, BAF 14, and BAF 15, respectively). The derived inventory data for larger BAFs were based on the comparison of the measured horizontal distance of a tally tree from plot center to the calculated limiting distance, using Equation 1. The limiting distance is the maximum horizontal distance from plot center to the face of a tree, with a given DBH, so that the tree would still be considered a tally tree according to the sampling probability (i.e., BAF). If the measured distance was less than or equal to the limiting distance, then the subject tree was considered a tally tree at the selected BAF. The limiting distance (R) was calculated as

$$R = \frac{8.6962}{\sqrt{BAF}} \times DBH, \quad [1]$$

where R is in meters, BAF is in $m^2\ ha^{-1}\ tree^{-1}$, and DBH is in cm.

Individual tree volumes were calculated using regional species-specific equations adopted by the United States Forest Service's Forest Inventory and Analysis program (O'Connell et al. 2013), details of which can be found in Miles and Hill (2010) and Woodall et al. (2010). Plot-level volume was calculated as the volume of individual trees summed then multiplied by the appropriate sampling weight (Avery and Burkhart 2002).

LiDAR Data and Processing

LiDAR data for the area were collected in June 2011 by Aerometric, Inc. (Sheboygan, WI, USA) using a RIEGL LMS-Q680i airborne laser scanner on board a helicopter flown at an altitude of approximately 457 m and a ground speed of 111 km per hour. The LiDAR system, which operated at 1550 nm with pulse frequency of 400 kHz and scan angle of ±30° from nadir, generated an average point density of 18 pulses per square meter. The sensor captured up to 9 returns per pulse. The dataset was processed using FUSION software (McGaughey 2014) to produce information quantifying forest structure as well as elevation of the bare-earth surface. The "ground filter" tool in FUSION was used to separate ground and nonground returns, using default coefficients for the weight function (described in McGaughey 2014) and a tolerance value of 0.03 m after 10 iterations. A high-resolution (1.5 m) digital elevation model (DEM) was then created from the filtered ground returns and applied to normalize the raw LiDAR point cloud so that the remaining LiDAR returns represented the elevation of returns. Plot-level forest structure metrics were calculated by clipping the nonground returns for the size of the FRP (11.3 m radius) and other 23 different radii ranging from 7 m to 38 m ("LiDAR extracted areas"; Table 2). These metrics were based on the returns above 1.3 m from the ground surface (approximately equal to the height at which DBH is measured), and included the suite of metrics described in Hudak et al. (2008), Falkowski et al. (2010), and McGaughey (2014). Altogether, 57 metrics representative of canopy cover, height distributional statistics, and relative vegetation density by height class (i.e., percentage returns by height strata) were derived (see Appendix for the comprehensive list and definition of predictor metrics used in the models).

Modeling Standing Volume

The optimal LiDAR extracted area (i.e., clip radius) to couple VRP data with colocated plot-level LiDAR metrics was determined across the range of VRP BAFs, using a model selection procedure (Murphy et al. 2010) dependent on the Random Forest algorithm (RF; Breiman 2001; Cutler et al. 2007; Liaw and Wiener 2002). The model selection procedure was run in the R statistical software package (R Core Team 2013) after deriving a suite of aforementioned LiDAR metrics at each plot location at different resolutions, based on Equation 1. Specifically, Equation 1 was used to calculate a range of limiting distances from which to vary the clipping radii of the LiDAR samples (Table 2). This approach was intended to minimize the exclusion of tally trees from the LiDAR sample and improve the geometric consistency between the VRP data and the LiDAR sample.

For each model, a standard variable screening process, based on the QR-decomposition algorithm (Cížková and Cížek 2012; Falkowski et al. 2009) was employed to remove multicollinear variables from the entire set of 57 LiDAR metrics. Although the RF algorithm can handle datasets with high dimensionality, previous research has demonstrated that the presence of multicollinear variables negatively impact RF model performance (Evans and Hudak 2007). Following multicollinear variable screening, the RF model selection procedure (Evans and Murphy 2015) was employed to obtain RF models with an optimal set of LiDAR metrics for predicting plot-level volume across each set of BAFs and LiDAR extracted areas. The model selection procedure is designed to derive the most parsimonious model by selecting a suite of predictor variables (i.e., LiDAR metrics) that optimize the RF model improvement ratio and variation explained by the individual model (Murphy et al. 2010). The impact of LiDAR metrics of varying resolution on the response variable (standing volume) from VRP was evaluated to identify the optimal LiDAR extracted area for each BAF. The LiDAR-extracted area leading to the least RMSE was taken as the optimal, and the BAF that minimized both bias and RMSE was preferred for the VRP model. The optimal extracted area for a preferred BAF was eventually adopted for development of wall-to-wall LiDAR grid metrics to spatially extend the VRP data over the area of interest.

The impact of incorporating a mixture of FRP and VRP data on model accuracy was also assessed in an effort to improve inventory efficiency and accuracy. Specifically, different ratios of plot numbers from both FRP and VRP sampling were used for training the volume models; the ratios of FRP to VRP numbers tested were 64:36, 36:64, 32:68, 68:32, and 50:50, so that the sample size was 47 plots in each model-training set. For this analysis, the LiDAR data were clipped and summarized into structural metrics at resolutions corresponding to the FRP size (11.3 m radius) and the optimal LiDAR extracted area for the preferred VRP BAF, after which the 2 datasets were pooled.

Mapping Standing Volume and Accuracy Assessment

After completing the modeling procedure, spatially explicit maps of standing volume were generated via a *k* nearest neighbor (*k*-NN) imputation approach based on the top two performing VRP models. The spatial resolution of predictor LiDAR grid metrics was set equivalent to the optimal size of VRP for the 2 most suitable BAFs. Specifically, the yaImpute R package was used to implement a *k*NN imputation approach based on the RF algorithm (Crookston and Finley 2008; Falkowski et al. 2010; Hudak et al. 2008). The RF-*k*NN imputation approach leverages the RF proximity matrix (quantifying statistical proximity of observations) to determine nearest neighbors (see Falkowski et al. 2009) for a detailed description of this approach). Standing volume was also mapped, using the same approach solely based on FRP data and gridded LiDAR metrics at the corresponding spatial resolution (22.6 m)

The accuracy of VRP-based volume prediction was evaluated at the plot-level via a goodness-of-fit (correlation, RMSE, and bias) comparison with the FRP field measurements. Following Robinson et al. (2005) and Robinson and Froese (2004), equivalence tests were conducted to assess the statistical equiv-

TABLE 2
Summary of VRP sampling statistics across the range of BAFs at the 6 x target stands

BAF	Total Tally Trees in all 47 Plots	Min. Tally Trees per Plot	Max. Tally Trees per Plot	No. of Plots with Less than 4 tally trees	Average DBH (cm)	Max. DBH (cm)	Average Limiting Distance (m)	Max. Limiting Distance (m)
5	840	6	33	0	23.9	81.6	11.2	38.1
7	627	5	27	0	24.3	81.6	9.6	32.2
9	479	2	22	1	24.3	81.6	8.5	28.4
10	441	2	21	1	24.5	81.6	8.1	27.0
12	367	2	18	5	24.7	81.6	7.4	24.6
14	317	2	14	9	24.6	81.6	6.9	22.8
15	291	2	14	11	24.5	81.6	6.6	22.0

alence of VRP-based predictions to FRP field measurements (the region of equivalence for the intercept and slope were both set to 25%, which is default in the R equivalence package). The equivalence test relies on a subjective choice of a region in which difference between predicted and observed values are considered negligible. A region of equivalence (or indifference) set to ± 25% of the standard deviation around the mean imply that predicted and observed values are equivalent if the absolute value of the mean of the differences is less than 25% of the standard deviation. In addition to plot-level comparison, stand-level volume estimates from the 2 VRP models were also compared with stand-level estimates derived from the FRP model within each of the 6 sampled stands. However, an equivalence test could not be performed for the stand-level comparisons, because the sample size was too small.

RESULTS

Comparisons of the plot-level volume estimates, with the 7 different intensities of VRP at each sample location, revealed that VRP sampling with BAF 9 provides the most similar inventory (in terms of bias) to the FRP sampling at the same locations (Table 3a). RMSE decreased with decreasing BAF until becoming relatively stable at BAF 9. Hence, BAF 9 was identified as the optimal or preferred among the other VRP scales. When the sample plots were split into 2 different groups based on development stage (young and old), analysis of residuals revealed that the BAF 9 performs better (lower bias and error) in young stands, whereas BAF 10 performs better in the old stands (Tables 3b and 3c). It was also evident that the VRP-based inventory had a larger bias in the young stands compared to the old stands, and the estimates were negatively biased (i.e., underestimated) with smaller BAFs in high biomass areas (i.e., old stands; Table 3c). The VRP-based volume estimates demonstrated that larger BAFs overestimated volume in low biomass areas, whereas smaller BAFs underestimated volume in high biomass areas. In addition, the variance of VRP estimates increased with increasing mean volume (of FRPs) across all BAFs tested.

The FRP model had an RMSE of 31.8 m^3 ha^{-1}, which, as expected, was the least across all models that were based on the entire set of 47 plots (Table 4). The FRP model was, therefore, taken as the reference model. Integrating the VRP data with LiDAR data at a fixed sample radius (11.35 m) resulted in lower amounts of variation explained and higher RMSE compared to models that integrated an optimized radius for LiDAR samples for each BAF (Table 4). The optimal resolution of LiDAR extracted area declined with increasing BAF; the optimum was 9 m radius for the preferred BAF 9 or 10. The combination of inventory data from FRP and VRP, in different ratios from both young and old stands, and LiDAR metrics at respective resolutions

TABLE 3a
Residual errors and correlation statistics between VRP-based and FRP-based volume estimates across all plots

Statistics	BAF5	BAF7	BAF9	BAF10	BAF12	BAF14	BAF15
Bias (m^3 ha^{-1})	−2.5	3.8	1.5	4.5	4.8	5.7	3.5
Relative bias (%)	−2.3	3.4	1.4	4.1	4.3	5.0	3.2
RMSE (m^3 ha^{-1})	35.7	35.4	36.2	38.6	41.0	45.1	44.7
Relative RMSE (%)	34.19	31.9	33.2	34.5	36.6	39.9	40.4
Corr. coef.	0.9	0.9	0.9	0.9	0.7	0.8	0.8

TABLE 3b
Residual errors and correlation statistics between VRP-based and FRP-based volume estimates across plots in the young stand development class

Statistics	BAF5	BAF7	BAF9	BAF10	BAF12	BAF14	BAF15
Bias (m^3 ha^{-1})	11.0	11.1	7.11	7.7	6.9	7.9	8.0
Relative bias (%)	15.9	16.1	10.9	11.7	10.7	12.0	12.2
RMSE (m^3 ha^{-1})	21.5	24.8	22.6	26.8	27.1	29.4	30.9
Relative RMSE (%)	31.2	36.0	34.9	40.9	41.8	44.7	47.0
Corr. coef.	0.6	0.6	0.5	0.4	0.4	0.3	0.3

revealed that inclusion of a higher proportion of FRPs from older stands produced better models in terms of RMSE. Similar results were obtained when the training data contained a mixture of VRP data from the 2 top-performing levels of BAF (9 and 10) and FRP data (i.e., sampling designs were mixed within and between stands) and models were developed using LiDAR metrics at respective resolution (see the last 5 rows in Table 4). This implies that VRP data from younger stands and FRP data from older stands, or only VRP data from all stands, can be combined to formulate a generalized model with some compromise in accuracy.

The selected LiDAR metrics in the 2 VRP and FRP models were relatively stable between all models (Figure 4). The relative importance ranking of predictors (% increase in mean square error) show that Stratum5 (see definition in Appendix) was the most important LiDAR metric in all 3 models, and the suite of LiDAR metrics for each model included metrics related to both vegetation height distribution and cover.

Comparisons of the plot-level volume estimates of the BAF 9 and BAF 10 imputation models against the FRP measurements are presented via equivalence plots in Figure 5. The equivalence test uses the null hypothesis of dissimilarity of 2 target datasets being compared (Robinson et al. 2005; Robinson and Froese 2004). According to this test, if 2 one-sided confidence intervals (at a given alpha level) for the slope and intercept of the line of best fit each lie within a specified region of similarity, the 2 datasets are statistically equivalent. Because the line of best fit for the VRP BAF 10 model predictions in Figure 5 lies within the region of similarity (i.e., 25% for both slope and intercept), and the confidence intervals for slopes and intercepts lie within the respective regions of similarity, the VRP BAF 10 model-based estimates are statistically equivalent to the FRP measurements. However, the null hypothesis of dissimilarity was not rejected for VRP-BAF-9-model estimates at 25% region of similarity because the confidence interval for the slope extends partly outside the region.

Scatterplots of observed against predicted volume for the 3 sampling strategies (FRP, VRP BAF 9, and VRP BAF 10) reveal strong, linear relationships (Figure 6). Observed correlations were 0.95, 0.92, and 0.92 for the FRP, VRP BAF 9, and VRP BAF 10 models, respectively. However, when the data were divided into young and old stands, the relationship was found to be weaker, especially for young stands. For the young stands only, the correlation coefficients were 0.66, 0.52, and 0.75 for the FRP, VRP BAF 9, and VRP BAF 10 models, respectively.

Comparison of the stand-level mean volume estimates obtained from the standing volume maps corresponding to the FRP and VRP models (i.e., raster outputs at 18-m resolution for the VRP models, and at 22.6-m resolution for the FRP model; figures not shown) revealed that the BAF 10 model is least biased. Because the volume estimates from the FRP model are assumed to be the best (i.e., closest to measured growing stock), the VRP model estimates were compared to the FRP model estimates. The average biases were found to be -9.8 and -0.4 m^3 ha^{-1} for the BAF 9, and BAF 10 models, respectively. The (pixel level) means of the volume maps obtained using the FRP and

TABLE 3c
Residual errors and correlation statistics between VRP-based and FRP-based volume estimates for plots in the old stand development class

Statistics	BAF5	BAF7	BAF9	BAF10	BAF12	BAF14	BAF15
Bias (m^3 ha^{-1})	−26.2	−9.1	−8.2	−1.1	0.9	1.7	−4.4
Relative bias (%)	−15.6	−4.9	−4.4	−0.6	0.5	0.9	−2.3
RMSE (m^3 ha^{-1})	52.1	48.7	49.4	53.3	57.9	63.9	61.9
Relative RMSE (%)	31.0	26.3	26.5	27.6	29.6	32.6	32.6
Corr. coef.	0.6	0.6	0.6	0.6	0.7	0.6	0.6

TABLE 4
A summary of RF-based models from the different combinations of field plot inventory data and LiDAR derived predictors. Lowest RMSE per sampling strategy is indicated in boldface

Description of Model Inputs	% Variance Explained	RMSE (m^3 ha^{-1})	LiDAR Extract Radius (m)
Based on FRP samples only (volume from FRP; LiDAR from 11.33 m radius)			
All 47 FRP samples	81.1	**31.8**	11.3
30 FRPs (in young stands only)	6.4	17.3	11.3
17 FRPs (in old stands only)	23.1	42.5	11.3
Based on VRP samples only (volume from VRP; LiDAR from an optimum radius)			
30 VRPs @BAF 5 (in young stands only)	26.1	19.2	15
30 VRPs @BAF 9 (in young stands only)	35.3	20.7	9
30 VRPs @ BAF 10 (in young stands only)	28.9	22.6	9
17 VRPs @ BAF 5 (in old stands only)	28.4	44.3	15
17 VRPs @ BAF 9 (in old stands only)	24.5	54.1	9
17 VRPs @ BAF 10 (in old stands only)	23.9	59.1	9
47 VRPs @ BAF 5	65.6	**35.2**	15
47 VRPs @ BAF 7	61.6	43.8	11
47 VRPs @ BAF 9	72.9	37.9	9
47 VRPs @ BAF 10	65.6	43.7	9
47 VRPs @ BAF 12	61.5	50.5	8
47 VRPs @ BAF 14	61.3	50.9	8
47 VRPs @ BAF 15	61.2	49.3	8
Based on VRP samples (volume from VRP; LiDAR from 11.33 m radius)			
47 VRPs @ BAF 5	61.7	**37.2**	11.3
47 VRPs @ BAF 7	61.8	43.7	11.3
47 VRPs @ BAF 9	68.2	41.2	11.3
47 VRPs @ BAF 10	65.6	45.7	11.3
47 VRPs @ BAF 12	60.0	51.5	11.3
47 VRPs @ BAF 14	57.3	53.5	11.3
47 VRPs @ BAF 15	55.7	52.7	11.3
Based on mixed FRP and BAF 9 samples*			
30 FRPs in young + 17 VRPs in old stands	74.3	37.5	11.3; 9
30 VRPs in young + 17 FRP in old stands	78.7	**33.3**	11.3; 9
Less FRPs and more VRPs (32:68) for all stands	70.5	40.6	11.3; 9
More FRPs and less VRPs (68:32) for all stands	76.5	34.6	11.3; 9
50% FRPs & 50% VRPs per stand	72.7	37.7	11.3; 9
Based on mixed FRP and BAF 10 samples*			
30 FRPs in young + 17 VRPs in old stands	69.6	43.6	11.3; 9
30 VRPs in young + 17 FRP in old stands	76.8	**34.7**	11.3; 9
Less FRPs and more VRPs (32:68) for all stands	68.1	44.9	11.3; 9
More FRPs and less VRPs (68:32) for all stands	76.5	34.8	11.3; 9
50% FRPs & 50% VRPs per stand	69.6	40.5	11.3; 9
Based on mixed FRP and BAF9 and BAF10 samples†			
30 FRPs in young + 17 VRPs in old stands	72.3	41.0	11.3; 9
30 VRPs in young + 17 FRP in old stands	77.5	**34.4**	11.3; 9
Less FRPs and more VRPs (32:68) for all stands	68.9	42.6	11.3; 9
More FRPs and less VRPs (68:32) for all stands	77.1	**34.4**	11.3; 9
50% FRPs & 50% VRPs per stand	71.5	38.9	11.3; 9

*volume from respective plots and LiDAR from 11.33 m radius for FRP & 9 m radius for VRP; †volume and LiDAR from respective plots; 50% VRP data from BAF9, and 50% from BAF 10.

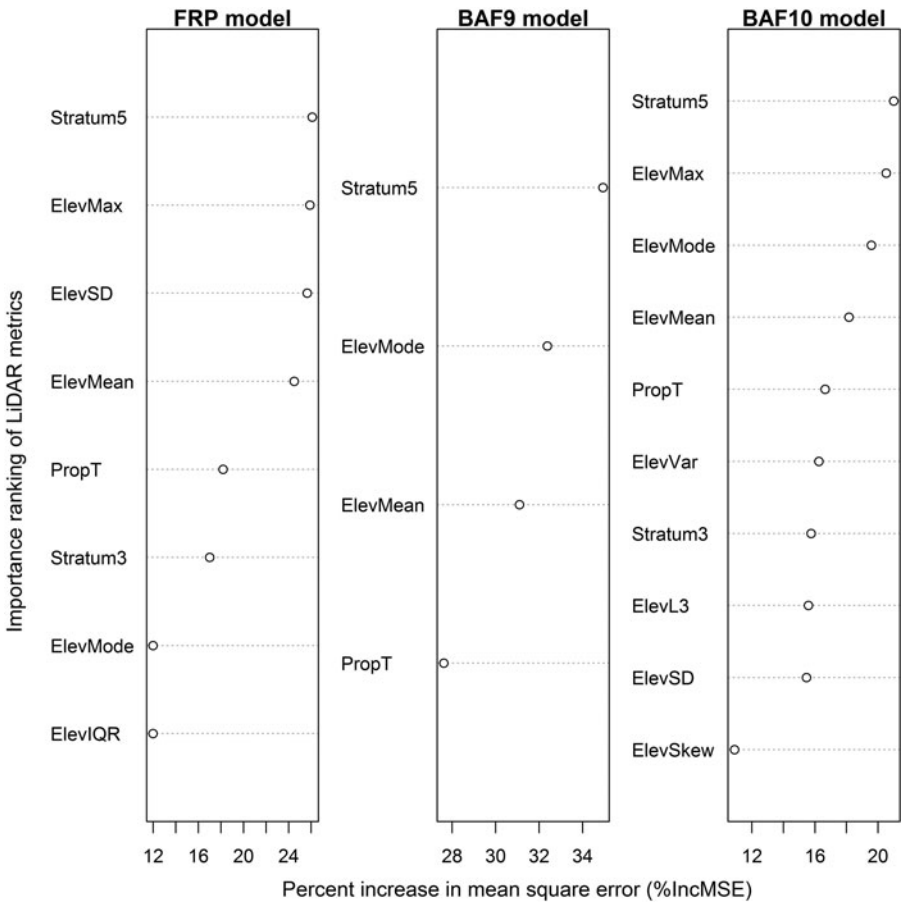

FIG. 4. Importance ranking of LiDAR metrics (see definitions in Appendix) selected in the FRP and VRP models. Criterion for the ranking of variables was the percentage increase in mean square error when individual predictors are sequentially dropped and substituted with random numbers.

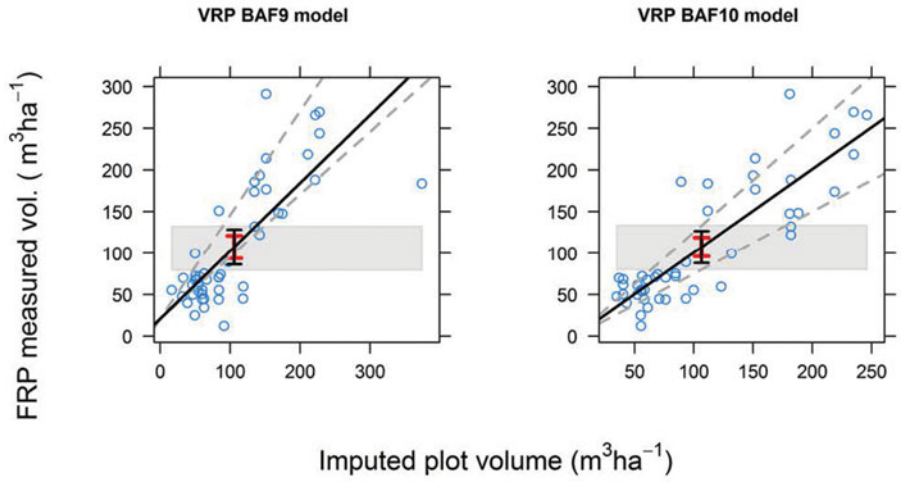

FIG. 5. Equivalence plot of the measured and imputed plot-level volumes by the 2 VRP models. The black line represents the line of best fit, the dashed gray lines represent the 25% region of similarity for the slope, the shaded gray polygon represents the 25% region of similarity for the intercept. The black vertical bar represents a confidence interval ($\alpha = 0.05$) for the slope, and red vertical bar depicts the confidence interval for the intercept.

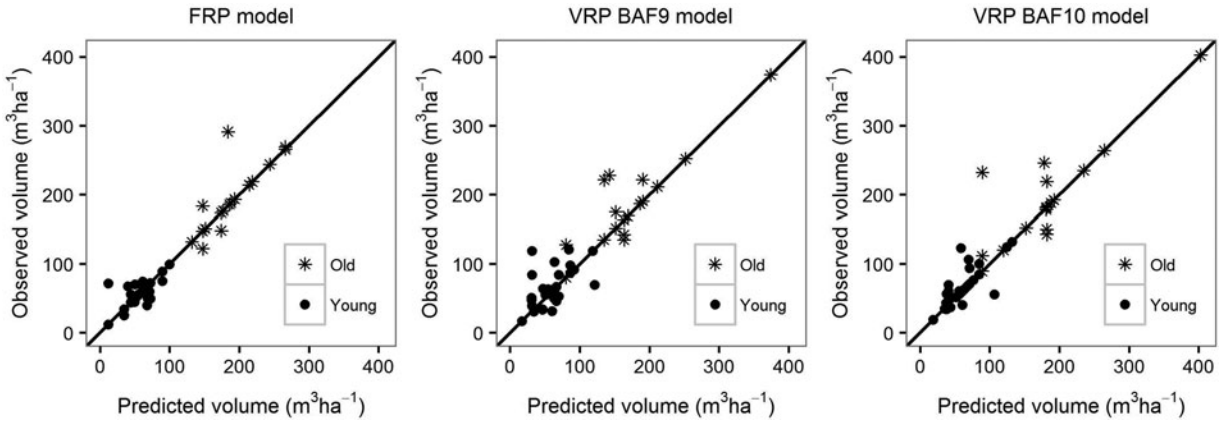

FIG. 6. Scatterplots of field-observed plot volumes by the FRP, VRP BAF 9, and VRP BAF 10 sampling strategies against the predicted values by age class.

VRP BAF 9 and BAF 10 models were found to be 78.3, 65.4, and 71.9 m^3 ha^{-1}, respectively, which also implies that the BAF 10 model had the least bias. The (pixel-level) standard deviations of the maps were found to be 56.9, 50.6, and 56.6 m^3 ha^{-1} for the FRP and VRP BAF 9 and BAF 10 models, respectively, which suggests that the FRP model best describes forest structural variability, followed by the BAF 10 model.

DISCUSSION

The overarching goal of this study was to conduct a detailed assessment of the integration of LiDAR and VRP data, with the ultimate aim of improving the efficiency of LiDAR-based inventories. To meet this objective, we first explored 2 questions focused on understanding (*i*) the optimum LiDAR resolution to employ when working with VRP data collected with different sampling probabilities, and (*ii*) how model accuracy is affected when using VRP data rather than the more traditional approach of using FRP data. Our results reveal that the optimum size of LiDAR metrics for the VRP-based inventory is dependent on forest structural complexity (both development stage and forest composition) as well as the sampling intensity (BAF); hence, choosing the right value for both is important in order to maximize efficiency. A comparison of design-based volume estimates from VRP with FRP revealed that BAF 9 performed best followed by BAF 10 (Tables 3a,b,c). LiDAR-based models trained on the VRP datasets similarly revealed that BAF 9 and BAF 10 models performed best when paired with LiDAR metrics calculated at a spatial resolution of 18 m diameter, evaluated using explained variance and RMSE (Table 4). As expected, LiDAR models trained on FRP data performed best when LiDAR metrics were derived at a spatial resolution approximating FRP diameter (22.6 m). This finding reflects the importance of spatially matching the location, size, and shape of field plots with the location, size, and shape (i.e., resolution) of LiDAR metrics.

Equivalence tests comparing plot-level volume predictions (m^3 ha^{-1}) by VRP BAF 9 and BAF 10 models against the FRP field measurements show that the BAF 10 model, but not the BAF 9 model, produces estimates statistically equivalent to FRP-based estimates. Similarly, the comparisons of stand-level total volume estimates by the 2 VRP-based models against the FRP model suggest that the BAF 10 model is least biased. Avery and Newton (1965) reported that BAF 10 VRPs are roughly equivalent to 0.04 ha and 0.02 ha FRPs, in terms of tree tallies, in stands with average DBH of 34.2 cm and 24.1 cm, respectively. It seems plausible that BAF 10 sampling approximates 0.04 ha FRPs in our study area, since the optimal resolution of LiDAR metrics for the VRP BAF 10 model was found to be 18 m diameter (equivalent to 0.032 ha) and the average DBH of the stands ranged from 14.4 cm to 24.1 cm. This is further supported by the low RMSE for models based on LiDAR samples from FRP and inventory data collected at the VRPs with BAF 9 and BAF 10 (41.2 and 45.7 m^2 ha^{-1}, respectively; Table 4).

The observation that BAF 10 performed best in our study area is reassuring for multiple reasons. In the given structural conditions, BAF 10 VRP sampling included the optimal numbers of trees per plot (at least 4) across all stands in our study area; according to Avery and Burkhart (2002), Schreuder et al. (1993), and NRIS (2014), at least 4 trees should be selected per VRP to obtain reliable estimates. Although an optimal scale of VRP depends on forest structural variability, our result of equivalent volume estimates by BAF 10 VRP and FRP models conforms to Avery and Newton (1965), who reported similar volume estimates using BAF 10 and 0.04 ha FRP field sampling. Further, angle gauges with BAFs in multiples of 5 ft^2 ac^{-1} are commonly available.

The integration of VRP data with colocated LiDAR metrics from fixed-radius samples (11.35 m) resulted in higher RMSE compared to models that integrated an optimized radius for LiDAR samples for each BAF. Our results suggest that the optimal LiDAR extracted area decreases with increasing BAF. This can

be explained by the observation that increasing BAF increases sighting angle and diminishes apparent spatial coverage of a VRP (i.e., for a tree of given DBH to be included in the tally, it must lie closer to the plot center with increasing BAF). Including sample plots in a LiDAR-dependent inventory that are too small or too large also involve drawbacks due to edge effects (White et al. 2013) or smoothing of LiDAR predictors with expanding grid size. The edge effect for smaller plots might be significant, particularly in situations where tree stems along plot borders that are not contained within the field sample have crown components included or contained within the colocated LiDAR sample. A reverse case is also possible when trees are actually within the field sample but portions of crown elements are not within a colocated LiDAR sample (e.g., due to leaning trees). This error due to edge effect increases as plot size (and more particularly the perimeter/area ratio) decreases. White et al. (2013) suggested that a LiDAR metric grid size greater than 15 m produces acceptable results in most forest types.

The results presented herein suggest that data collected using VRP sampling can provide inventory estimates comparable to that based on FRP sampling as long as each VRP contains > 4 tally trees and the identified optimum radius for the VRP approximates the standard size of FRP (i.e., 0.04 ha). Combining a suite of LiDAR metrics with data from fixed-area plots against variable-radius plots in the older stands improved model accuracy by a small amount (i.e., from 74.2% to 78.6% for BAF 9, and from 69.6% to 76.8% for BAF 10). In other words, substituting FRPs for VRPs in the young stands did not improve accuracy. However, because the young stands contained low-standing volumes, it could be that VRP can work well in older stands that have large volumes. Scott (1990) reported that VRP samples perform better than FRP in older stands with larger trees and a wide range of diameters. Note that FRP sampling requires more sample trees per plot than VRP sampling in order to yield volume estimates with the same level of precision (Matern 1972).

Application of plot inventory data from mixed sampling schemes (FRP and VRP, or only VRP at different levels of BAF) at separate locations in model building also seems to work well in terms of goodness-of-fit statistics (variance explained ranged from 68.93% to 77.52%). However, the use of multiple BAFs in LiDAR-based multistand inventory is work intensive because multiple LiDAR grids at different resolutions corresponding to different BAFs need to be prepared in order to extend the model spatially in order to produce maps.

Because volume is a 3-dimensional metric, we used all non-ground LiDAR returns from various canopy height strata for the inventory estimation (Wulder et al. 2008). The removal of collinear predictors and selection of an optimum suite of LiDAR metrics in the models showed that canopy height distribution, percent cover, and vertical strata density metrics were the most influential in model training. The selected LiDAR metrics were relatively stable among all models. Vertical strata density (e.g., Stratum5) and canopy height distribution metrics were particularly influential predictors because volume in conifer forests is more related to height than crown size.

The errors in the LiDAR-based models can be attributed to several factors. A prominent factor is accuracy of plot-center coordinates that influences spatial matching of LiDAR data and field inventory (Gleason and Im 2011). The average horizontal precision of the plot coordinates from differentially corrected GPS data in our study was about 0.8 m. Nonetheless, it is possible that the extracted LiDAR samples for the field plots could exclude portions of canopy of sampled trees and include canopy elements of trees that are outside the plot. A tree just outside the plot or a leaning tree can contribute a large number of returns in the LiDAR samples. The use of allometric equations for volume calculation also introduces error, both because the equations might be dated and unsuitable for the contemporary forest, and because they contain intrinsic prediction error that is otherwise not accounted for in our analysis.

Overall, the apparent reduction in precision when choosing the best VRP design over the FRP design for our data (VRP9 with a 9 m LiDAR extract area) was about 19% (i.e., RMSE of 31.8 m^3 ha^{-1} for FRP vs 37.9 m^3 ha^{-1} for VRP9; see Table 4). The model developed from FRP data within young stands ($n = 30$) had a much lower RMSE than the model from FRP data within old stands ($n = 17$), and the same pattern was reflected in models developed from VRP data, although the RMSE was larger than for FRP models for the same age class (Table 4). These results suggest that LiDAR had greater sensitivity to volume in younger stands, regardless of inventory design. However, the relative difference in precision between VRP and FRP models was much larger for the old stand age class. For example, for young stands, the RMSE for the VRP9 model was 19.6% larger than that for the corresponding FRP model, but for old stands the RMSE for the VRP9 model was 27.3% larger. Thus, the magnitude of the loss of accuracy depends on the stand structure, and modelers who elect to use VRP designs should expect a larger negative effect in higher-volume stands. Notably, the combination of inventory from FRP and VRP, in different ratios from both young and old stands, and LiDAR metrics at respective resolutions revealed that including a higher proportion of FRPs from older stands produced a better model. This might indicate that the primary cause for the reduction in precision is increasing spatial mismatch between VRP field data and LiDAR metrics. It might be that the reduction in precision in such circumstances could be compensated for by increasing the size of the training dataset, by collecting additional plots, and this should be investigated in future research.

CONCLUSIONS

Although the FRP-based LiDAR inventory produced the most accurate results, VRP sampling was shown to be a useful means to enhance inventory efficiency through integration with a LiDAR dataset in our forest types, including young stands. However, selection of an optimal BAF and corresponding LiDAR-extracted area had notable influence on accuracy, and the

optimum varied with stand structure and complexity. VRP sampling of younger stands in the study resulted in a larger relative bias compared to the older stands. VRP sampling using BAF 10 was the most suitable in our area because it tallied at least 4 trees per plot and overcame practical difficulties associated with the use of several BAFs. As expected, integration of FRP data with LiDAR data provided the best model. Nevertheless, VRP data have the potential to be leveraged with LiDAR data for inventory prediction, with some compromises in accuracy. VRP data can also be combined with FRP data (or substituted for FRP data) for spatial inventory with LiDAR derived metrics at an optimum grid resolution. Fixed-area plots in the older stands against variable plots improved model accuracy by a small amount, which implies that VRP can work well in older stands that have large volumes. A combination of VRP data from younger stands and FRP data from older stands, or only VRP data from all stands, can be used to formulate a generalized model with accuracy comparable to an FRP-only model.

FUNDING

Funding for this work was provided from the NASA Carbon Monitoring Systems Program (14- CMS14-0026) to Andrew Hudak at the USFS Rocky Mountain Research Station, through a Joint Venture Agreement (08-JV-11221633-159) and NASA New Investigator Program grant (NNX14AC26G) to Michael Falkowski at the University of Minnesota, and by the School of Forest Resources and Environmental Science at Michigan Technological University. We would also like to thank Michigan Technological University's Research Excellence Fund for an award that was used to fund LiDAR data acquisition.

REFERENCES

Avery, T.E., and Burkhart, H.E. 2002. *Forest Measurements (5th ed.)*. New York, NY: McGraw-Hill.

Avery, G., and Newton, R. 1965. "Plot sizes for timber cruising in Georgia." *Journal of Forestry*. Vol. 63(No. 12): pp. 930–932.

Brosofske, K., Froese, R.E., Falkowski, M.J., and Banskota, A. 2014. "A review of methods for mapping and prediction of inventory attributes for operational forest management." *Forest Science*, Vol. 60(No. 2): pp. 1–24.

Breiman, L. 2001. "Random forest." *Machine Learning*. Vol. 45(No. 1): pp. 5–32.

Cížková, L., and Cížek, P. 2012. "Numerical linear algebra." In *Handbook of Computational Statistics: Concepts and Methods*, edited by J.E. Gentle, W.K. Hardle, and Y. Mori, pp. 105–137. Heidelberg: Springer.

Crookston, N.L., and Finley, A.O. 2008. "yaImpute: an R package for *k*NN imputation." *Journal of Statistical Software*, Vol. 23(No. 10): pp. 1–16.

Cutler, D.R., Edwards, T., Beard, K.H., Cutler, A., Hell, K.T., Gibson, J., and Lawler, J. 2007. "Random forests for classification in ecology." *Ecology*, Vol. 88(No. 11): pp. 2783–2792.

Dubayah, R.O., Sheldon, S.L., Clark, D.B., Hofton, M.A., Blair, J.B., Hurtt, G.C., and Chazdon, R.L. 2010. "Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica." *Journal of Geophysical Research*, Vol. 115(No. G2): pp. 1–17.

Evans, J.S., and Hudak, A.T. 2007. "A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments." *IEEE Transactions on Geoscience and Remote Sensing of Environment*, Vol. 45(No. 4): pp. 1029–1038.

Evans, J.S., and Murphy, M.A. 2015. *rfUtilities: Random Forests Model Selection and Performance Evaluation*. R package version 1.0-2, accessed November 10, 2015, http://cran.r-project.org/pack-age = rfUtilities.

Falkowski, M.J., Evans, J.S., Martinuzzi, S., Gessler, P.E., and Hudak, A.T. 2009. "Characterizing forest succession with LiDAR data: an evaluation for the inland northwest, USA." *Remote Sensing of Environment*, Vol. 113(No. 5): pp. 946–956.

Falkowski, M.J., Hudak, A.T., Crookston, N.L., Gessler, P.E., Uebler, E.H., and Smith, M.S. 2010. "Landscape-scale parameterization of a tree-level forest growth model: a k-nearest neighbor imputation approach incorporating LiDAR data." *Canadian Journal of Forest Research*, Vol. 40(No. 2): pp. 184–199.

Gleason, C.J., and Im, J. 2011. "A review of remote sensing of forest biomass and biofuel: options for small-area applications." *GIScience & Remote Sensing*, Vol. 48(No. 2): pp. 141–170.

Golinkoff, J., Hanus, M., and Carah, J. 2011. "The use of airborne laser scanning to develop a pixel-based stratification for a verified carbon offset project." *Carbon Balance and Management*, Vol. 6(No. 9): pp. 1–17.

Hollaus, M., Wagner, W., Maier, B., and Schadauer, K. 2007. "Airborne laser scanning of forest stem volume in a mountainous environment." *Sensors*, Vol. 7(No. 8) pp. 1559–1577.

Hollaus, M., Wagner, W., Schadauer, K., Maier, B., and Gabler, K. 2009. "Growing stock estimation for alpine forests in Austria: a robust LiDAR-based approach." *Canadian Journal of Forest Research*, Vol. 39(No. 7): pp. 1387–1400.

Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E., and Falkowski, M.J. 2008. "Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data." *Remote Sensing of Environment*, Vol. 112(No. 5): pp. 2232–2245.

Hudak, A.T., Evans, J.S., and Smith, A.M.S. 2009. "LiDAR utility for natural resource managers." *Remote Sensing*, Vol. 1: pp. 934–951.

Hudak, A.T., Haren, A.T., Crookston, N.L., Liebermann, R.J., and Ohmann, J.L. 2014. "Imputing forest structure attributes from stand inventory and remotely sensed data in western Oregon, USA." *Forest Science*, Vol. 60(No. 2): pp. 253–269.

Hummel, S., Hudak, A.T., Uebler, E.H., Falkowski, M.J., and Megown, K.A. 2011. "A comparision of accuracy and cost of LiDAR versus stand exam data for landscape management of the Malheur national forest." *Journal of Forestry*, Vol. July/August: pp. 267–273.

Husch, B., Beers, T.W., and Kershaw, J.A. 2003. *Forest Mensuration (4th ed.)*. Hoboken, NJ: John Wiley & Sons.

Jochem, A., Hollaus, M., Rutzinger, M., and Hofle, B. 2011. "Estimation of aboveground biomass in alpine forests: a semi-empirical approach considering canopy transparency derived from airborne LiDAR data." *Sensors*, Vol. 11(No.1): pp. 278–295.

Kellndorfer, J.M., Walker, W.S., LaPoint, E., Kirsch, K., Bishop, J., and Fiske, G. 2010. "Statistical fusion of LiDAR, InSAR, and optical remote sensing data for forest stand height characterization: a regional-scale method based on LVIS, SRTM, Landsat ETM+, and ancillary data sets." *Journal of Geophysical Research*, Vol. 115(No. G2) pp. 1–10.

Kronseder, K., Ballhorn, U., Bohm, V., and Siegert, F. 2012. "Above ground biomass estimation across forest types at different degradation levels in Central Kalimantan using LiDAR data." *International Journal of Applied Earth Observation and Geoinformation*, Vol. 18: pp. 37–48.

Latifi, H., Nothdurft, A., and Koch, B. 2010. "Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors." *Forestry*, Vol. 83(No. 4): pp. 395–407.

Lefsky, M.A., Cohen, W.B., Parker, G.G., and Harding, D.J. 2002. "LiDAR remote sensing for ecosystem studies." *BioScience*, Vol. 52(No. 1): pp. 19–30.

Liaw, L.A., and Wiener, M. 2002. "Classification and regression by random forest." *R News*, Vol. 2(No. 3): pp. 18–22.

Matern, B. 1972. "The precision of basal area estimates." *Forest Science*, Vol. 18(No. 2): pp. 123–125.

McGaughey, R.J. 2014. *FUSION/LDV: Software for LiDAR Data Analysis and Visualization, Version 3.21*. Seattle: USDA Forest Service/Pacific Northwest Research Station/University of Washington.

Miles, P.D., and Hill, A.D. 2010. *Volume Equations for the Northern Research Station's Forest Inventory and Analysis Program as of 2010*. General Technical Report NRS-74, Northern Research Station. USDA Forest Service, Newtown Square, PA.

Mitchell, J.J., Glenn, N.F., Sankey, T.T., Derryberry, D.R., Anderson, M.O., and Hruska, R.C. 2011. "Small-footprint LiDAR estimations of sagebrush canopy characteristics." *Photogrammetric Engineering and Remote Sensing*, Vol. 77(No. 5): pp. 1–10.

Murphy, M.A., Evans, J.S., and Storfer, A. 2010. "Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics." *Ecology*, Vol. 91(No. 1): pp. 252–261.

NRIS (2014). *Common stand exam user guide: Chapter 2- preparation and design*. US Forest Service Natural Resource Information System, Washington, DC, accessed April 1, 2014, http://www.fs.fed.us/nrm/fsveg/index.shtml.

O'Connell, B., LaPoint, E., Turner, J., Ridley, T., Boyer, D., Wilson, A., Waddell, K., and Conkling, B. 2013. *The Forest Inventory and Analysis Database: Database Description and Users' Manual Version 5.1.5 for Phase 2*. General Technical Report, Rocky Mountain Research Station. Moscow, ID: USDA Forest Service.

Ohmann, J.L., Gregory, M.J., and Roberts, H.M. 2014. "Scale considerations for integrating forest inventory plot data and satellite image data for regional forest mapping." *Remote Sensing of Environment*, Vol. 151 (No. 2012 ForestSAT): pp. 3–15.

Packard, K.C., and Radtke, P.J. 2007. "Forest sampling combining fixed- and variable-radius sample plots." *Canadian Journal of Forest Research*, Vol. 37(No. 8): pp. 1460–1471.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, accessed October 15, 2013, http://www.R-project.org/.

Reed, D.R., and Mroz, G.D. 1997. *Resource Assessment in Forested Landscapes*. New York: John Wiley & Sons.

Robinson, A.P., Duursma, R.A., and Marshall, J.D. 2005. "A regression-based equivalence test for model validation: shifting the burden of proof." *Tree Physiology*, Vol. 25(No. 7): pp. 903–913.

Robinson, A.P., and Froese, R.E. 2004. "Model validation using equivalence tests." *Ecological Modelling*, Vol. 176(No. 3–4): pp. 349–358.

Saatchi, S.S., Harris, N.L., Brown, S., Lefsky, M.A., Mitchard, E.T., Salas, W., Zutta, B.R., Beurmann, W., Lewis, S.L., Hagen, S., Petrova, S., White, L., Silman, M., and Morel, A. 2011. "Benchmark map of forest carbon stocks in tropical regions across three continents." In *Proceedings of the National Academy of Sciences of USA*, edited by S.E. Trumbore, Vol. 108 (No. 24): pp. 9899–9904. Irvine, CA: University of California.

Schreuder, H.T., Gregoire, T.G., and Wood, G.B. 1993. *Sampling Methods for Multisource Forest Inventory*. New York: John Wiley and Sons.

Scott, C.T. 1990. "An overview of fixed versus variable-radius plots for successive inventories." In *State-of-the-Art Methodology of Forest Inventory, Jul 30–Aug 5, 1989*. General Technical Report PNW-GTR-263, Pacific Northwest Research Station, edited by V. LaBau and T. Cunia, pp. 94–104. Syracuse, NY: USDA Forest Service.

Shugart, H.H., Saatchi, S.S., and Hall, F.G. 2010. "Importance of structure and its measurement in quantifying function of forest ecosystems." *Journal of Geophysical Research*, Vol. 115(No. G2): pp. 1–16.

Sun, G., Ranon, K.J., Guo, Z., Zhang, Z., Montesano, P., and Kimes, D. 2011. "Forest biomass mapping from LiDAR and radar synergies." *Remote Sensing of Environment*, Vol. 115(No. 11): pp. 2906–2916.

Takagi, K., Yone, Y., Takahashi, H., Sakai, R., Hojyo, H., Kamiura, T., Nomura, M., et al. 2015. "Forest biomass and volume estimation using airborne LiDAR in a cool-temperate forest of northern Hokkaido, Japan." *Ecological Informatics*, Vol. 26(No. 3): pp. 54–60.

Van Aardt, J.A.N., Wynne, R.H., and Oderwald, R.G. 2006. "Forest volume and biomass estimation using small-footprint LiDAR-distributional parameters on a per-segment basis." *Forest Science*, Vol. 52(No. 6): pp. 636–649.

White, J.C., Wulder, M.A., Varhola, A., Vastaranta, M., Coops, N.C., Cook, B.D., Pitt, D., and Woods, M. 2013. *A Best Practice Guide for Generating Forest Inventory Attributes from Airborne Laser Scanning Data Using an Area-Based Approach (Version 2.0)*. Victoria, BC: Natural Resources Canada/Canadian Forest Service/Canadian Wood Fibre Centre

Woodall, C.W., Heath, L.S., Domke, G.M. and Nichols, M.C. 2010. *Methods and Equations for Estimating Aboveground Volume, Biomas, and Carbon for Trees in the US Forest Inventory*. General Technical Report NRS-88, Northern Research Station. USDA Forest Service. Please provide city of publisher or Northern Research Station for Woodall et al. 2010

Wulder, M.A., Bater, C.W., Coops, N.C., Hilker, T. and White, J.C. 2008. "The role of LiDAR in sustainable forest management." *The Forestry Chronicle*, Vol. 84(No. 6): pp. 807–826.

APPENDIX

Description of 90 different LiDAR metrics used in this study

Predictor	Description
PropT	proportion of total return > 1.5 m (total returns > 1.5 m/total returns)
Prop1	proportion of first return > 1.5 m (first returns > 1.5 m/total returns > 1.5 m)
Prop2	proportion of second return > 1.5 m (second returns > 1.5 m/total returns > 1.5 m)
Prop3	proportion of third return > 1.5 m (third returns > 1.5 m/total returns > 1.5 m)
Prop4	proportion of fourth return > 1.5 m (fourth returns > 1.5 m/total return count > 1.5 m)
Prop5	proportion of fifth return > 1.5 m (fifth returns > 1.5 m/total return count > 1.5 m)
ElevMin	Elevations minimum
ElevMax	Elevations maximum
ElevMean	Elevations mean
ElevMode	Elevations mode
ElevSD	Elevations standard deviation
ElevVar	Elevations variance
ElevCV	Elevations coefficient of variation
ElevIQR	Elevations interquartile range
ElevSkew	Elevations skewness
ElevKurt	Elevations kurtosis
ElevAAD	Elevations average absolute deviation
EMADmed	Median of the absolute deviations from the overall median of elevations
EMADmod	Mode of the absolute deviations from the overall mode of elevations
ElevL1	Elevations first L-moment
ElevL2	Elevations second L-moment
ElevL3	Elevations third L-moment
ElevL4	Elevations fourth L-moment
ElevLCV	Elevations L-moment coefficient of variation
ElevLskew	Elevation L-moment skewness
ElevLkurt	Elevation L-moment kurtosis
ElevP-*i*	Elevations *i*th percentile, where $i = 1, 5, 10, 20, 25, 30, 40, 50, 60, 70, 75, 80, 90, 95,$ and 99
CRR	Canopy relief ratio ((mean-min)/(max-min))
EQM	Elevation quadratic mean
ECM	Elevation cubic mean
Density1	overstory canopy density as% of first return > 3m (1st returns > 3m/total 1st returns × 100)
Density2	overstory canopy density as% of all return > 3m (all returns > 3m/total all returns × 100)
Density3	Percentage first returns above mean
Density4	Percentage first returns above mode
Density5	Percentage all returns above mean
Density6	Percentage all returns above mode
Stratum0	proportion of ground return
Stratum1	proportion of aboveground returns below 1.5 m
Stratum2	proportion of vegetation returns above 1.5 m and below 6 m
Stratum3	proportion of vegetation returns above 6 m and below 10.6 m
Stratum4	proportion of vegetation returns above 10.6 m and below 15.2 m
Stratum5	proportion of vegetation returns above 15.2 m and below 19.8 m
Stratum6	proportion of vegetation returns above 19.8 m