Scientific Research Publishing

Open Journal of Forestry, 2025, 15(3), 226-246
https://www.scirp.org/journal/ojf
ISSN Online: 2163-0437
ISSN Print: 2163-0429

Predictive Ability of Volume-Basal Area Ratios (VBARs) in Minnesota, USA Forests to Estimate Volume Per Hectare

Curtis L. VanderSchaaf

Central Mississippi Research and Extension Center, Mississippi State University, Raymond, USA
Email: clv127@msstate.edu

How to cite this paper: VanderSchaaf, C.L. (2025). Predictive Ability of Volume-Basal Area Ratios (VBARs) in Minnesota, USA Forests to Estimate Volume Per Hectare. *Open Journal of Forestry, 15,* 226-246.
https://doi.org/10.4236/ojf.2025.153013

Received: April 30, 2025
Accepted: July 6, 2025
Published: July 9, 2025

Abstract

Natural resource inventories often aim to acquire desired information in the least amount of time while minimizing costs. Within Minnesota, USA due to access issues and costs, in particular for management inventories, often few field-based sampling points can be established. Additionally, inventories conducted to establish timber sales are often of stands that contain low value timber and that consequently have low sale rates. Thus, in these cases, the cost of establishing enough field-based sampling points to meet some desired statistical level of precision is not justified. Therefore, alternative inventory methods, such as the use of remotely sensed data including LiDAR, are being examined. This study examined the ability of two methods to estimate VBAR, where volume and basal area are stand-level values. The first method is to use a constant VBAR across all conditions of a cover type. A second method is to estimate VBAR by cover type using various combinations of plot age, site index, and basal area per hectare. Data was obtained from national inventory plots as part of the US Department of Agriculture, Forest Service, Forest Inventory and Analysis program. Although LiDAR is not actually used during this assessment, results can be used to infer about the bias, precision, and accuracy associated with using LiDAR determined tree heights to predict diameter and then to estimate stand densities to ultimately estimate volume per hectare. Results showed that basal area per hectare is not a consistently useful variable to estimate VBARs. Site index and stand age are better predictors. Based on inference from this study, at the current time, it appears that the use of LiDAR to ultimately estimate volume per hectare looks most promising for those conditions that require less accuracy and precision in estimates, such as for management plan inventories and for appraisals of low value timber.

Keywords

Forest Inventory, Remote Sensing, *Abies*, *Acer*, *Betula*, *Fraxinus*, *Larix*, *Picea*, *Pinus*, *Populus*, *Quercus*, *Thuja*

1. Introduction

Natural resource inventories aim to acquire desired information in the least amount of time while minimizing costs. Many forest inventories and most timber sale appraisals in Minnesota, USA use on-the-ground variable-radius approaches (Burkhart et al., 2019, Chapter 12; Shiver & Borders, 1996, Chapter 4). However, in particular for inventories conducted for management planning purposes, often few sampling points can be established due to access issues and labor costs. Additionally, inventories conducted to establish timber sales are often in stands that contain low value timber and that consequently have low sale rates. Hence, in these cases, the cost of establishing enough sampling points to meet some desired statistical level of precision was not justified. For example, the Minnesota Department of Natural Resources Division of Forestry (DoF), a large government forestland holding agency (around 1.11 million ha of timberland), used a fixed minimum number of points for a particular sized stand in the late 2010's (**Table 1**), regardless of cover type, when conducting variable-radius management plan inventories using a 10 basal area factor (BAF) (2.296 BAF metric) prism. Beyond that, due to rising salaries and fringe benefit costs (e.g. health insurance and retirement benefits), the DoF was experiencing staffing issues.

Table 1. Minimum number of plots/points to install when using a 10 BAF (2.296 BAF metric) prism and measuring volume on every tree as implemented by the Minnesota (USA) Department of Natural Resources (DNR) Division of Forestry (DoF) during inventories conducted for management purposes.

Acres	Hectares	Number of Points
1 - 10	0.4047 - 4.0469	3
11 - 20	4.4515 - 8.0937	4
21 - 40	8.4984 - 16.1874	5
41 - 80	16.5921 - 32.3749	6
81 - 120	32.7796 - 48.5623	7
121 - 160	48.9670 - 64.7498	8
161 - 200	65.1544 - 80.9372	9
201 - 240	81.3419 - 97.1246	10
241+	97.5293	Add one plot for every additional 40-acre increase in stand size.

Therefore, the DoF was looking at inventory alternatives, such as using aerial photography and LiDAR to either supplement on-the-ground variable-radius approaches or replace them. The DoF used aerial photographs to conduct regeneration surveys of Aspen (*Populus* spp.) and Black Spruce (*Picea mariana* (Mill.) B.S.P.) stands. LiDAR looks like a promising remotely sensed method that will provide sufficient data to conduct management inventories and likely even provide enough detail to conduct appraisals of lower value timber (Hummel et al., 2011; Watt et al., 2013; St. Peter et al., 2021; Brown et al., 2022). For the purposes of management, management plan inventories can be of lower intensity than timber appraisal inventories, because less precise information is needed. Inventory estimated variables such as basal area per hectare and quadratic mean diameter can be used to gain some idea of what forested stands may need to be thinned over the next planning interval and those two variables plus average tree height can be used to gain some idea of habitat characteristics for wildlife management purposes. Lower value timber, for example, are the majority of Tamarack (*Larix laricina* (Du Roi) K. Koch) stands and sites of other more valuable species located far from likely markets, resulting in greater transportation costs by loggers and hence lower stumpage revenues. Currently, Aspen stands located in far western Minnesota, USA have low sale rates because of relatively lower productivity levels and high transportation costs to mills. The majority of these stands will be sold exclusively as pulpwood, perhaps with some relatively more valuable bolts. Thus, due to DoF staffing issues and high ground-based inventory costs relative to the purpose of the inventory and the value of the timber, the DoF did not want to allocate substantial financial and logistical resources to conduct inventories of these difficult access and lower saleable sites. However, as the state agency, they were tasked by law to manage and conduct timber sales on these sites. Thus, LiDAR estimates perhaps would be advantageous on these sites.

Some believe on-the-ground plots or points provide better inventories than remotely sensed data. However, in fact, in certain situations remotely sensed inventories may be superior (Hummel et al., 2011; Watt et al., 2013; Brown et al., 2022). For example, timber cruises conducted using plots or points only directly measure a portion of a population or area of interest (or stand), a sample cruise. LiDAR can theoretically measure every tree in a stand. Most field inventories have low levels of precision since few plots can be established because of costs and sometimes access issues, similarly mentioned by Brown et al. (2022). For stands where most, and if not all, timber is being sold for pulpwood, LiDAR data will likely be sufficient. However, for stands that contain a variety of product classes (e.g. pulpwood, bolt, veneer, poles, and sawlog), currently, LiDAR will likely be inferior compared to many field based inventories. At this time, LiDAR does not have the ability to merchandize trees, at least at a satisfactory level.

Several studies have found that larger spatial scale LiDAR-based forest inventories are at a minimum comparable in costs to ground-based, or on-the-ground, inventories (Hummel et al., 2011; Kepler, 2019; Arney & Corrao, 2021). This is often particularly true as the area of interest increases in size because of economies of scale (Hummel et al., 2011; White et al., 2016; Kepler, 2019). LiDAR-based inventories also have the advantage of providing information across the entire area of interest unlike individual ground-based stand exams that often only directly measure a small portion of the area of interest. Plus, ground-based inventories over large areas often require greater time lengths relative to LiDAR-based forest inventories (Kepler, 2019; Arney & Corrao, 2021). One disadvantage currently about LiDAR data, particularly in reference to conducting an up-to-date inventory of an individual stand, is that available data may be a few months old or even years old; whereas a ground-based inventory can more easily collect data today for that individual stand.

To conduct an inventory using LiDAR, individual trees would be identified using an algorithm and their heights (H) would be estimated using computer techniques. The diameter (D) of each identified tree would then be estimated using a H-D equation developed using field-collected independent data. Currently, the H-D equation would need to be generalized across species within a cover type because of the difficulty in determining individual tree species in a timely and efficient manner at a large spatial scale using LiDAR. The combination of the computer-generated H from LiDAR, satellite imagery, aerial photography, or another remotely sensed source of data, and the predicted tree D, would allow for dominant height and an average stand-level per hectare value of basal area to be calculated, where tree density per hectare (acre) would be estimated using the remotely sensed data. Dominant height if desired would ultimately lead to an estimate of site index, the species selected based on the cover type. If age was desired, hopefully, it could be obtained from previous inventory data. This stand-level per hectare information would then be used to estimate volume per hectare.

Currently, individual tree volume equations cannot be used adequately because of the difficulty in determining species when using LiDAR. Hence, average VBAR, or volume-basal area ratio (Burkhart et al., 2019: pp. 276-277; Shiver & Borders, 1996: pp. 99-103) relationships at the per hectare (acre) by cover type level is a viable alternative.

VBARs are the amount of a product (e.g. merchantable volume measured in cords) per square meter (foot) of basal area at some D along the tree stem, in this case 1.372 m (4.5 feet) above the ground, referred to as dbh, or diameter at breast height, in the USA. Based on predicted Ds using the H-D equation, basal area per hectare would be calculated from an estimate of tree density per hectare using the remotely sensed inventory data and then the previously developed VBAR from research data would be used to estimate volume per hectare. VBAR could be for individual trees or on a per hectare (acre) basis though. If for the stand as a whole, LiDAR would also be used to estimate trees per hectare (acre).

This study examines the ability of two methods to estimate VBAR, where volume and basal area are per hectare (acre) values, with the ultimate aim of predicting volume per hectare (acre). The first method (Method One) is to use a constant VBAR across all conditions of a cover type. A second method (Method Two) is to estimate VBARs using regression analysis within a cover type. Regressors include plot age as reported within the US Department of Agriculture (USDA), Forest Service, Forest Inventory and Analysis (FIA) database, site index (base age 50) as reported by FIA, and basal area per acre (hectare) as reported by FIA. Predicted VBARs along with observed basal areas per acre (hectare) will be used to predict volume per acre (hectare), which will then be converted to volume per hectare, of an independent validation dataset to see which of the two methods produces the best predictions. Although LiDAR is not actually used during this assessment, results can be used to infer about the bias, precision, and accuracy associated with using LiDAR determined tree heights (H) to predict diameter (D) and then to estimate stand densities to ultimately estimate volume per hectare.

2. Material and Methods

The data used in model development were obtained from USDA, Forest Service, Forest Inventory and Analysis (FIA) annual surveys completed between 2008 and 2012 (EVAL_GRP = 272012 within the FIA POP_EVAL_GRP table). Survey data were obtained from all regions of Minnesota, USA but only using those plots classified as being located on state lands (OWNCD = 31). Data were obtained from the FIA database website (O'Connell et al., 2013; USDA Forest Service, 2013).

The DoF has a database entitled Forest Inventory Module (FIM) that contains estimates of basal area per acre and cords per acre from inventories conducted by DoF personnel. However, estimates of volumes contained within trees is based upon visual estimates of merchantable stem length. If inventories were actually operationally conducted using LiDAR, the independent data source used to estimate the VBAR component would likely be FIA data, not FIM data. FIA protocol dictates that D and H of every tree within a plot be actually measured, not visually estimated.

2.1. FIA Sampling Protocol

Plots are clusters of four points arranged such that point 1 is central, with points 2 through 4 located 36.576 m (120 ft) from point 1 at azimuths of 0, 120, and 240 degrees (Bechtold & Scott, 2005). Each cluster point is surrounded by a 7.315 m (24.0 ft) fixed-radius subplot where trees 12.700 cm (5.0 in.) dbh and larger are measured. Combined, the four subplots total approximately 0.0672 ha (1/6th acre). Each subplot contains a 2.073 m (6.8 ft) fixed-radius microplot where saplings (2.540 to 12.446 cm dbh) (1.0 to 4.9 in.) are measured. The four microplots total approximately an area of 0.0054 ha (1/75th acre). Condition classes are assigned to differentiate conditions occurring on a plot and a subplot can have more than one condition class. A condition class differentiates stands based on variables that FIA monitors, such as cover type, ownership, stand density.

In most states FIA is collected at a single-intensity level and hence plots are located roughly every 2428.12 ha (6000 acres). However, Minnesota, USA is on a double-intensity data collection level and hence plots are located roughly every 1214.06 ha (3000 acres). FIA monitors which plots are single-intensity and which plots are double-intensity. This is useful because we can use one group as a model development dataset (double-intensity plots) and the second group as a model validation dataset (single-intensity).

2.2. Data Used

Merchantable volume as defined by FIA (VOLCFNET within the FIA Tree table, O'Connell et al., 2013) was used as the dependent variable. Essentially, this is the merchantable volume of trees with dbh's of 12.700 cm (5.0 in.) and greater, from a 0.305 m (1 ft) stump to a minimum 10.160 cm (4.0 in.) top diameter outside bark (DOB). These specifications are basically the same as those used by the DoF during management inventories. VOLCFNET from FIA was converted to cords by dividing VOLCFNET by 79 (Gevorkiantz & Olsen, 1955).

FIA forest type codes (FORTYPCD within the FIA Condition table, O'Connell et al., 2013) were mapped to currently used DoF cover types as found on the FIA Evalidator tool (Miles, 2013). Only living trees were included (STATUSCD = 1 in the FIA Tree table, O'Connell et al., 2013), all tree classes were included (hence TREECLCD within the FIA Tree table (O'Connell et al., 2013) was not used to filter the data), and FIA timberland was used (hence SITECLCD, O'Connell et al., 2013 from 1 to 6 in the FIA Condition table). Only those plots where all four subplots were defined as the same condition class (CONDID = 1 in the FIA Condition table) were used. **Table 2** presents summary data for the plots used to calculate VBAR (double-intensity, INTENSITY = 2 in the FIA Plotsnap table) and **Table 3** summarizes those plots used to test or validate predictive ability (single-intensity, INTENSITY = 1 in the FIA Plotsnap table) by cover type.

2.3. Estimating VBAR

As mentioned earlier, FIA plots within Minnesota, USA are collected at a double-intensity. Hence, rather than an FIA plot being located roughly every 2428.12 ha (6000 acres), a plot is located roughly every 1214.06 ha (3000 acres). For the purposes of this paper, the double-intensity plots can be used to estimate the VBAR (model development) while the single-intensity plots will provide an independent set of plots to predict volumes (model validation) using the estimated VBAR from the double-intensity plots.

The first method to estimate VBAR only estimates a single VBAR value of a cover type using all available plots (Method One). The second method (Method Two) uses equations to predict the VBAR as a function of plot age, site index (base age 50), and basal area per acre (hectare) by cover type. For this study, VBAR was defined simply as the standing merchantable volume in cords divided by the total basal area (square feet) of a plot. Basal area included all trees, but volume was only calculated using those trees considered merchantable. Currently, it will be nearly impossible for LiDAR to distinguish between non-merchantable and merchantable basal area. Additionally, when calculating VBAR all species were pooled within a FIA plot. Although theoretically possible to use species-specific VBARS within the same cover type, it will be difficult using current LiDAR methods to identify the species of trees to enough precision justifying the use of different VBARs by species within a plot.

Table 2. Mean, maximum, minimum, and standard deviation of the observations used to estimate Volume-Basal Area Ratios (VBAR) (FIA INTENSITY = 2, VBAR estimation dataset). VBARs are quantified in cords but are independent of the unit of area (whether hectares or acres). Where: LH is Lowland Hardwood, BP is Balsam Poplar, NH is Northern Hardwood, BF is Balsam Fir, BSL is Black Spruce Lowland, NWC is Northern White Cedar, and RPP is Red Pine Plantation.

VBAR estimation dataset													
Cover		Volume-basal area ratio (VBAR)				Cords Per Ha				Basal Area Per Ha (sq m)			
Type	*n*	Mean	Max	Min	Std dev	Mean	Max	Min	Std dev	Mean	Max	Min	Std dev
LH	33	0.1756	0.2592	0.0556	0.0501	36.7	77.0	2.9	22.2	19.1	32.5	1.5	9.5
Aspen	114	0.1371	0.3004	0.0012	0.0824	25.5	114.3	0.3	24.0	15.8	40.3	0.2	8.9
Birch	21	0.1411	0.2332	0.0466	0.0647	23.3	69.1	1.2	18.7	14.1	27.5	2.1	7.1
BP	17	0.1193	0.2195	0.0380	0.0569	17.9	59.1	0.3	17.0	13.1	27.7	0.4	8.5
NH	35	0.2037	0.2866	0.1005	0.0532	54.3	99.6	10.4	28.5	23.7	36.8	6.2	9.2
Oak	23	0.2061	0.3298	0.1154	0.0536	56.3	128.9	7.1	32.7	24.3	45.8	3.8	9.1
BF	7	0.1246	0.1936	0.0380	0.0524	26.8	86.0	7.8	27.4	19.7	51.7	6.9	15.0
BSL	58	0.0711	0.2286	0.0019	0.0595	14.7	101.0	0.3	18.9	16.3	41.1	0.9	8.9
Tamarack	67	0.1124	0.2396	0.0031	0.0633	19.7	48.3	0.3	15.3	16.0	37.8	0.5	9.0
NWC	34	0.1389	0.2300	0.0405	0.0476	41.0	120.4	4.1	24.9	27.7	69.2	3.1	13.2
RPP	7	0.2095	0.2431	0.1463	0.0328	63.1	117.8	19.8	29.9	27.1	46.2	12.6	10.5

VBAR estimation dataset									
Cover		Plot Age (years)				Site index (m, base age 50)			
Type	*n*	Mean	Max	Min	Std dev	Mean	Max	Min	Std dev
LH	33	80	128	35	27	16.4	23.5	10.7	3.7
Aspen	114	36	109	1	22	20.0	29.0	12.8	3.4
Birch	21	56	100	15	24	15.6	22.6	9.1	3.5
BP	17	39	107	6	27	18.7	26.8	13.4	3.5
NH	35	66	126	3	29	19.2	25.6	11.6	3.2
Oak	23	70	128	1	25	18.3	25.6	11.6	3.8
BF	7	78	171	1	65	15.3	21.6	6.7	6.3
BSL	58	74	128	34	25	10.1	18.3	6.1	3.0
Tamarack	67	73	181	4	35	12.2	23.8	6.4	3.9
NWC	34	101	195	35	38	9.3	15.2	4.9	3.1
RPP	7	45	109	15	31	20.2	24.7	16.2	3.2

Table 3. Mean, maximum, minimum, and standard deviation of the observations used to validate predictions (FIA INTENSITY = 1, Validation Dataset). VBARs are quantified in cords but are independent of the unit of area (whether hectares or acres). Where: LH is Lowland Hardwood, BP is Balsam Poplar, NH is Northern Hardwood, BF is Balsam Fir, BSL is Black Spruce Lowland, NWC is Northern White Cedar, and RPP is Red Pine Plantation.

Validation Dataset													
Cover		Volume-basal area ratio (VBAR)				Cords Per Ha				Basal Area Per Ha (sq m)			
Type	*n*	Mean	Max	Min	Std dev	Mean	Max	Min	Std dev	Mean	Max	Min	Std dev
LH	17	0.1607	0.3101	0.0264	0.0608	41.4	175.2	1.0	38.1	21.0	52.5	3.4	11.2
Aspen	111	0.1375	0.2840	0.0022	0.0744	26.5	109.2	0.3	25.3	16.0	39.3	0.2	9.3
Birch	23	0.1583	0.2677	0.0467	0.0581	30.4	63.9	1.2	18.2	16.8	31.0	2.2	8.4
BP	16	0.1438	0.2383	0.0451	0.0640	23.1	91.0	0.9	22.4	12.9	35.5	0.6	8.0
NH	24	0.1886	0.2671	0.0046	0.0597	46.2	84.2	1.0	26.7	22.1	33.1	1.6	9.3
Oak	23	0.2206	0.3099	0.0854	0.0559	66.8	149.0	11.1	40.8	26.3	45.6	6.5	11.1
BF	11	0.1365	0.2150	0.0724	0.0467	22.6	64.2	4.7	16.0	15.1	27.8	6.0	7.5
BSL	67	0.0819	0.2201	0.0038	0.0647	16.0	69.8	0.3	18.0	15.0	34.9	0.2	8.9
Tamarack	56	0.1253	0.2635	0.0097	0.0770	20.1	93.2	0.4	19.5	15.1	39.0	0.2	10.5
NWC	32	0.1351	0.2024	0.0643	0.0338	52.1	99.4	15.5	20.0	36.2	56.4	11.7	11.4
RPP	6	0.2065	0.2913	0.1192	0.0610	63.3	99.1	7.9	33.5	27.8	49.0	6.2	14.9

Validation Dataset									
Cover		Plot Age (years)				Site index (m, base age 50)			
Type	*n*	Mean	Max	Min	Std dev	Mean	Max	Min	Std dev
LH	17	75	143	1	30	15.1	24.1	11.0	3.5
Aspen	111	38	115	2	22	18.7	25.6	8.8	3.5
Birch	23	59	86	20	20	16.6	23.5	9.4	3.5
BP	16	38	75	1	22	18.0	31.1	12.8	4.4
NH	24	71	210	1	42	17.2	24.1	11.6	3.0
Oak	23	72	127	2	32	19.1	25.0	10.7	4.0
BF	11	56	112	20	29	14.0	19.8	10.7	3.2
BSL	67	77	152	17	29	10.8	18.3	6.1	3.0
Tamarack	56	73	170	19	34	12.5	24.4	6.7	3.7
NWC	32	100	147	52	26	8.8	17.7	4.6	3.1
RPP	6	42	76	24	19	16.5	20.4	12.5	3.5

For this analysis, the Lowland Hardwood, Aspen, Birch (*Betula papyrifera* var. papyrifera), Balsam Poplar (*Populus balsamifera* L.), Northern Hardwood, Oak (*Quercus* spp.), Balsam Fir (*Abies balsamea* (L.) Mill.), Black Spruce Lowland (BSL), Tamarack, Northern White Cedar (*Thuja occidentalis* L., NWC), and Red Pine Plantation (*Pinus resinosa* Aiton, RPP) cover types were examined. The Jack Pine (*Pinus banksiana* Lamb.), Red Pine Natural, White Pine (*Pinus strobus* L.), and White Spruce (*Picea glauca* (Moench) Voss) cover types did not have sufficient numbers of plots for analysis. For Jack Pine and White Spruce cover types, even when combining both natural and planted sources together, sample sizes were still low. Balsam Fir did not have a single equation with a positive Adjusted R^2 and hence was excluded for Method Two (e.g. regression equations). For RPP, only those plots considered plantations by FIA (STDORGCD = 1 in the FIA Condition table) were used, these plots have sufficient evidence to be classified as plantations. All statistical analyses were carried out using the SAS program, Version 9.4 (SAS Institute Inc., 2004).

2.4. Using Regression Analysis to Estimate VBAR

Equations (1) to (5) were used to estimate VBAR as a function of site variables that would be easily obtainable from most stand inventory databases:

$$\text{VBAR} = b_0 \text{Age}^{b_1} \text{SI}^{b_2} \text{BA}^{b_3} + \varepsilon \tag{1}$$

$$\text{VBAR} = b_0 \text{Age}^{b_1} \text{SI}^{b_2} + \varepsilon \tag{2}$$

$$\text{VBAR} = b_0 \text{Age}^{b_1} \text{BA}^{b_3} + \varepsilon \tag{3}$$

$$\text{VBAR} = b_0 \text{SI}^{b_2} \text{BA}^{b_3} + \varepsilon \tag{4}$$

$$\text{VBAR} = b_0 \text{BA}^{b_3} + \varepsilon \tag{5}$$

where:

VBAR—volume (cords) -basal area ratio,

Age—plot age,

SI—site index (feet, base age 50),

BA—square feet of basal area per acre,

b_0, b_1, b_2, b_3—parameters to be estimated, and

ε—random error where it is assumed $\varepsilon \sim N(0, \sigma^2 I)$.

Equations (1) to (5) were fit using English units. Table 4 presents parameter estimates and model fitting results for Equations (1) to (5). In many cases parameter estimates are not significant at the 0.10 level and Adj. R^2s are low. However, in some cases solid Adj. R^2s were observed.

Table 4. Parameter estimates of Equations (1) to (5) used to predict the Volume-Basal Area Ratio (VBAR) by cover type when using reported FIA diameter at breast height, dbh. dbh occurs at 4.5 feet above the ground (1.372 m). Bold parameter estimates are not significant at a 0.10 alpha level. Where: SI is site index (feet, base age 50), BA is basal area per acre (square feet), L is Lowland, N is Northern, BSL is Black Spruce Lowland, NWC is Northern White Cedar, and RPP is Red Pine Plantation. b_1 corresponds to age, b_2 corresponds to site index, and b_3 corresponds to basal area per acre. Parameter estimates are in English units.

Equation (1)						
			Age	SI	BA	
Cover Type	n	b_0	b_1	b_2	b_3	Adj. R^2
L Hardwood	33	0.0059090	0.3218131	0.5552649	**−0.0495698**	0.2489
Aspen	114	0.0003118	0.7128529	0.8658865	**−0.0182708**	0.3650
Birch	21	0.0004825	0.5943290	0.7772666	**0.0666207**	0.3992
Balsam Poplar	17	0.0063668	0.4962834	**0.1752198**	**0.1175991**	0.2401

Continued

N Hardwood	35	0.0054965	0.1360530	0.5816599	**0.1441795**	0.2211
Oak	23	0.0084692	**−0.0826879**	0.5578109	0.2711941	0.3576
BSL	58	0.0000134	0.4287372	1.6382199	0.2366949	0.3948
Tamarack	67	0.0025864	0.3448830	0.7630315	**−0.1154368**	0.1082
NWC	34	0.0445850	**0.0248504**	0.3997938	**−0.0702739**	0.0874
RPP	7	0.0133986	**0.1342196**	**0.4461077**	**0.0826445**	0.5234
Equation (2)						
Cover Type	n	b_0	b_1	b_2	b_3	Adj. R^2
LH	33	0.0064464	0.2921510	0.5133406	-	0.2549
Aspen	114	0.0002998	0.7122718	0.8569956	-	0.3704
Birch	21	0.0004153	0.6222646	0.8554330	-	0.4243
Balsam Poplar	17	0.0339744	0.4438009	**−0.0816194**	-	0.2007
NH	35	0.0070330	0.1676521	0.6506907	-	0.1929
Oak	23	0.0144051	**−0.0062102**	0.6569744	-	0.2201
BSL	58	0.0000330	0.3960857	1.7075334	-	0.3878
Tamarack	67	0.0028158	0.2663198	0.7021056	-	0.1047
NWC	34	0.0413945	**−0.0214659**	0.3870237	-	0.1078
RPP	7	0.0150956	0.1724377	**0.4774092**	-	0.6134
Equation (3)						
Cover Type	n	b_0	b_1	b_2	b_3	Adj. R^2
LH	33	0.0591165	0.2582616	-	**−0.0073560**	0.0535
Aspen	114	0.0147218	0.6020661	-	**0.0257976**	0.3116
Birch	21	0.0111580	0.4700085	-	**0.1673884**	0.2652
Balsam Poplar	17	0.0136870	0.4871792	-	**0.1132881**	0.2908
NH	35	0.0641597	**0.0553472**	-	0.2037628	0.1117
Oak	23	0.0582950	−0.1239666	-	0.3835598	0.1738
BSL	58	0.0037789	**−0.0635438**	-	0.7492343	0.0947
Tamarack	67	-	-	-	-	-
NWC	34	-	-	-	-	-
RPP	7	0.0700424	**0.0963454**	-	**0.1573628**	0.3796
Equation (4)						
Cover Type	n	b_0	b_1	b_2	b_3	Adj. R^2
LH	33	0.0266477	-	0.4809177	**−0.0060281**	0.0940
Aspen	114	0.0070031	-	0.5476394	0.1694879	0.0560
Birch	21	0.0111199	-	**0.3900418**	**0.2513429**	0.0589
Balsam Poplar	17	-	-	-	-	-

Continued

NH	35	0.0147023	-	**0.3920645**	0.2208525	0.1641
Oak	23	0.0072939	-	0.6211699	0.1743281	0.3284
BSL	58	0.0001974	-	1.4239636	**0.2128786**	0.3677
Tamarack	67	0.0189976	-	0.5374517	**−0.0477648**	0.0771
NWC	34	0.0487980	-	0.3933726	**−0.0608758**	0.1162
RPP	7	0.0150890	-	**0.3544346**	**0.2429083**	0.4339
Equation (5)						
Cover Type	n	b_0	b_1	b_2	b_3	Adj. R^2
LH	33	-	-	-	-	-
Aspen	114	0.0633146	-	-	0.1907877	0.0410
Birch	21	0.0494205	-	-	**0.2618250**	0.0643
Balsam Poplar	17	-	-	-	-	-
NH	35	0.0710560	-	-	0.2304540	0.1216
Oak	23	0.0718355	-	-	0.2290321	0.0888
BSL	58	0.0027308	-	-	0.7610744	0.1102
Tamarack	67	-	-	-	-	-
NWC	34	-	-	-	-	-
RPP	7	0.0578380	-	-	0.2726478	0.4172

2.5. Estimating Volume Per Hectare

Cords per hectare were estimated for each cover type using two VBAR estimation methods of either a Single VBAR (Method One) or VBAR as estimated using stand-level regressors which is Method Two. To ultimately produce estimates of cords per hectare, VBARs using one of the methods were used to estimate cords per acre which was then converted to cords per hectare (**Table 5**).

In operational practice, when using VBAR to estimate volume per hectare, tree H would be measured using the LiDAR data, an equation would then be used to estimate D (currently generalized across species since it is difficult to currently determine individual tree species at a large spatial scale using LiDAR), and LiDAR would be used to determine trees per hectare (acre), where the combination of the estimated Ds and trees per hectare (acre) would produce the estimated basal area per hectare (acre).

Here, for this analysis, the tree Hs are assumed known (in a sense simulating LiDAR estimates of individual tree Hs), but Ds are predicted using species-specific H-D equations fit using FIA data (VanderSchaaf, 2012; 2013). During this analysis, for each species, only the population-average regression equation was used. To account for the transformation bias, the procedure recommended by Baskerville (1972) was used. In this study, species-specific H-D equations are used—this is somewhat of an optimal situation because currently, operationally, cover type H-D equations would need to be used and not species-specific equations due to the difficulty of identifying individual tree species at a large spatial scale using LiDAR data.

Although species-specific H-D equations were used, unfortunately, equations were not developed for eastern white pine and a few rare or non-timber species such as eastern hophornbeam (*Ostrya virginiana* (Mill.) K. Koch); in these cases the FIA reported Ds were used. Regardless of species, FIA does not report Hs for trees with a D of 4.9 in. and less (12.446 cm). For these trees, the measured D as reported by FIA was used.

2.6. Validating Volume Per Hectare Estimates

The true volume per hectare values are assumed to be the reported volumes per hectare (acre) by FIA of the single-intensity FIA plots. The difference between the observed (V_{oi}) and estimated (V_{ei}) volumes per hectare were calculated for each individual plot (i) for each of the two methods of estimating VBAR and ultimately volume per hectare. For each VBAR method, the mean residual ($\bar{e}$) and the sample variance (v) of residuals were computed by cover type across all plots and considered to be estimates of bias and precision; respectively.

An estimate of mean square error (MSE) was obtained for both of the VBAR methods by combining the bias and precision measures using the following formula:

$$\text{MSE} = \bar{e}^2 + v \tag{6}$$

2.7. Biological Interpretation of Parameter Estimates

In regression analysis, parameter estimates are examined to make sure they are logical given biological theory. It should be remembered that Equations (1) to (5) are estimating VBAR and not standing volume. For instance, as age and site index increase you would expect standing volume to increase. However, for the same amount of basal area, should volume vary—stated differently, should VBAR change. Likely with age, as basal area remains constant, the VBAR should increase for total merchantable volume because trees should be increasing in height. However, for site index it is somewhat difficult to determine what is logical. If the parameter estimate is positive, this is saying that for some reason, as basal area is held constant, as site quality increases that volume becomes greater. It could be that stem form is related to site quality, such that more cylindrical trees with greater volumes per basal area are observed on higher quality sites, and perhaps for the same basal area better quality sites have taller trees.

3. Results and Discussion

VBARs differed across cover types (VBAR Mean in Table 2 and Table 3). These VBARs show there is variability across cover types suggesting that a single VBAR value across all cover types would not be applicable. For these forests, based on statistical significance levels, basal area does not appear to be a consistently useful variable to estimate VBARs (**Table 4**). Site index and stand age appear to be better predictors—Equation (2). Studies in other forest types, such as Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) plantations in New Zealand (Watt et al., 2013), have found high correlations between LiDAR height percentiles and stand-level basal area. However, studies of naturally regenerated longleaf pine (*Pinus palustris* Mill.) in southwestern Georgia (Silva et al., 2016) and mixed-species forests (both pine (*Pinus* spp.) and hardwood) of both natural regenerated and planted origins in southern Alabama (Brown et al., 2022), similar to this study, found that using H to predict D resulted in poor predictions of basal area. Silva et al. (2016) attributed the poor predictions to the loss of H-D allometry in trees with D's exceeding 25 cm due to asymptotic height behavior while Brown et al. (2022) had similar conclusions, believing the poor results due to weak relationships between a tree's H and its basal area.

Across all cover types and equations, the Adjusted R^2 values were generally poor (**Table 4**). Although basal area may be beneficial in certain cases, operationally it would be a predicted variable as a function of the LiDAR data. The error associated with the prediction of basal area would be "passed" along to the prediction of VBAR. Therefore, the gain from including basal area in the regression analyses does not appear substantial enough to justify its inclusion.

In this study, when estimating basal area per acre (hectare) using predicted D's obtained from the H-D equations, D was predicted simply as a function of H ignoring individual stand-level factors. Mixed-effects models can be calibrated if some Hs and Ds were actually field measured within the stand. However, that of course will require time and money to visit the stand which would likely defeat the purpose of using LiDAR, or other remotely sensed data. Alternatively, one could use a H-D equation that includes stand-level variables. This could potentially improve the D estimate; however, operationally, the stand-level variables would be obtained from LiDAR data, or other remotely sensed sources of data, and thus estimates of these stand-level variables may not be as good as desired. Future research should concentrate on this alternative.

Across all observations (n = 375) and thus all cover types (**Table 5**, **Figure 1** and **Figure 2**), a single cover type VBAR (Method One) performed better in terms of MSE when predicting cords per hectare than when using regression to estimate VBAR (Method Two). Similarly, for many cover types, a single cover type VBAR performed better in terms of MSE when predicting cords per hectare than when using regression to estimate VBAR (**Table 5**). Based on MSE, when using regression equations (Method Two), Equation (2), or only using SI and Age as regressors, appears to be sufficient for many cover types to predict VBAR and ultimately volume per hectare. Regression showed particularly poor results for MSE as compared to a single cover type VBAR for the Aspen and Oak cover types. Additionally, a single cover type VBAR was superior in terms of MSE relative to all equations for the Birch and RPP cover types.

Table 5. Predicted bias (e), variance (v), and Mean Square Error (MSE) values when predicting cords per hectare for the validation dataset (single-intensity, INTENSITY = 1 in the FIA Plotsnap table) when using either a single VBAR (Method One) by cover type or stand-level regressors to predict VBAR by cover type (Method Two) when diameter at breast height (dbh) is predicted using height-diameter equations (VanderSchaaf, 2012, 2013). n = 375. For the Single VBAR, n = 386 since the Balsam Fir cover type was also predicted. dbh occurs at 4.5 feet above the ground (1.372 m).

		Single VBAR			Equation (1)			Equation (2)		
Cover Type	n	e	v	MSE	e	v	MSE	e	v	MSE
LH	17	−9.1	366	450	−6.6	374	418	−8.5	368	441
Aspen	111	−2.8	216	223	−7.2	661	712	−7.6	719	777
Birch	23	2.7	114	122	−3.2	192	202	−2.2	158	163
Balsam Poplar	16	4.3	59	78	0.0	142	142	2.7	45	53
NH	24	−17.1	424	717	−19.0	717	1079	−14.8	466	684
Oak	23	−7.6	1343	1400	−24.4	3959	4554	−10.8	1714	1831
Balsam Fir	11	1.4	44	46		-			-	
BSL	67	2.9	150	158	−2.1	133	138	−0.5	117	117
Tamarack	56	−5.8	256	290	−6.1	212	249	−8.7	432	508
NWC	32	−11.4	223	354	−7.8	176	237	−10.4	231	339
RPP	6	0.2	965	965	0.8	1323	1324	3.3	995	1006
All—no BF	375	−3.8	322	336	−7.3	627	680	−6.5	505	547
All Cover Types	386	−3.6	315	328	-	-	-	-	-	-

		Equation (3)			Equation (4)			Equation (5)		
Cover Type	n	e	v	MSE	e	v	MSE	e	v	MSE
LH	17	−8.4	505	575	−8.9	263	343		–	
Aspen	111	−8.5	596	668	−7.2	579	630	−7.8	538	599
Birch	23	−2.1	174	178	−3.4	301	312	−2.4	271	277
Balsam Poplar	16	0.2	132	132		–			–	
NH	24	−24.2	853	1437	−20.8	800	1230	−24.2	874	1459
Oak	23	−26.9	4,429	5153	−20.4	3104	3521	−19.8	2911	3304
Balsam Fir	11		–			–			–	
BSL	67	−1.5	169	171	−1.0	136	137	−1.7	172	175
Tamarack	56		–		−6.0	261	297		–	
NWC	32		–		−8.0	182	246		–	
RPP	6	−5.0	1865	1891	−4.7	2120	2142	−8.8	2558	2635
All—no BF	375	−8.6	826	899	−7.5	593	649	−8.4	736	807
All Cover Types	386	–	–	–	–	–	–	–	–	–

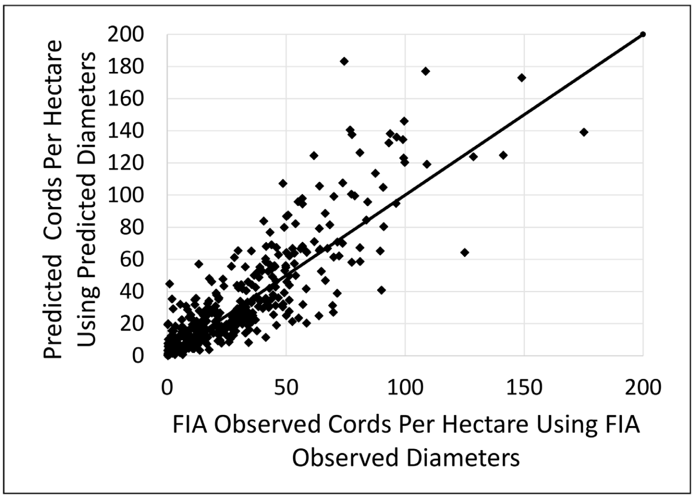

Figure 1. Predicted cords per hectare over observed cords per hectare (FIA observed diameters) when using a single VBAR (Method One) for a particular cover type. For the predicted cords, diameter at breast height (dbh) is predicted diameters using equations found in VanderSchaaf (2012) and VanderSchaaf (2013). $n = 386$. Average error across all $n = 386$ observations equals −3.608 cords per hectare with a variance of 314.707, and a mean square error (MSE) of 327.722.

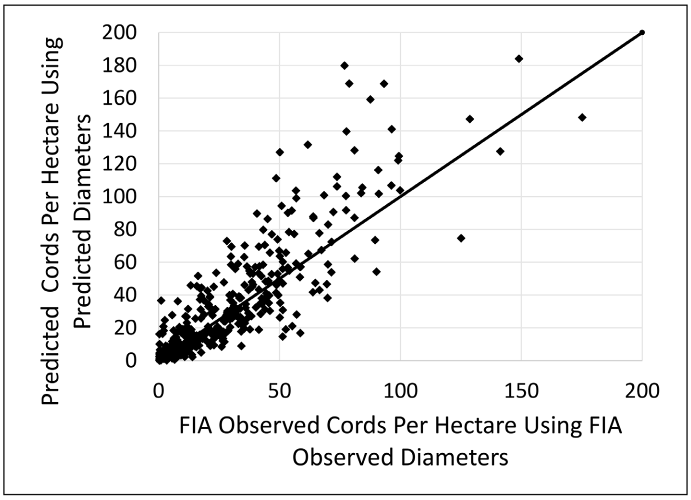

Figure 2. Predicted cords per hectare over observed cords per hectare (FIA observed diameters) when using Equation (2) for each particular cover type (Method Two), where stand age and site index are used to predict the VBAR for each individual stand. For the predicted cords, diameter at breast height (dbh) is predicted diameters using equations found in VanderSchaaf (2012) and VanderSchaaf (2013). $n = 375$. A sample size of 375 is less than **Figure 1** ($n = 386$) because no equation was fit for the Balsam Fir cover type. Average error across all $n = 375$ observations equals −6.483 cords per hectare with a variance of 505.339, and a mean square error (MSE) of 547.368.

Figure 3, **Figure 4**, and **Figure 5** present predicted cord per hectare estimates by cover type when using a single VBAR for a particular cover type (Method One). It is obvious that the use of predicted Ds, where the observed (e.g. reported by FIA) Ds can be assumed to be the correct cord per hectare value, has less impact for some cover types than others (Table 5 and Figure 3, Figure 4 and Figure 5), perhaps due to issues such as asymptotic H behavior in general and/or stand density that may impact taper rates and Hs through time. The use of predicted Ds generally resulted in overprediction (Table 5 and Figure 1, Figure 3, Figure 4 and Figure 5). Predictions, on average, were relatively precise for Balsam Poplar, Balsam Fir, BSL, and Birch (Table 5 and Figure 3 and Figure 4). For LH, NH, Oak, and NWC, based on the error term or bias (e), predicted Ds produced relatively sizable overpredictions on average (Table 5 and Figure 3, Figure 4 and Figure 5). Oak showed a relatively large amount of variability (v) as well. Red pine plantation (RPP) predictions were nearly unbiased but highly variable, using predicted Ds sometimes overpredicted volume and other times underpredicted volume relative to when using observed Ds.

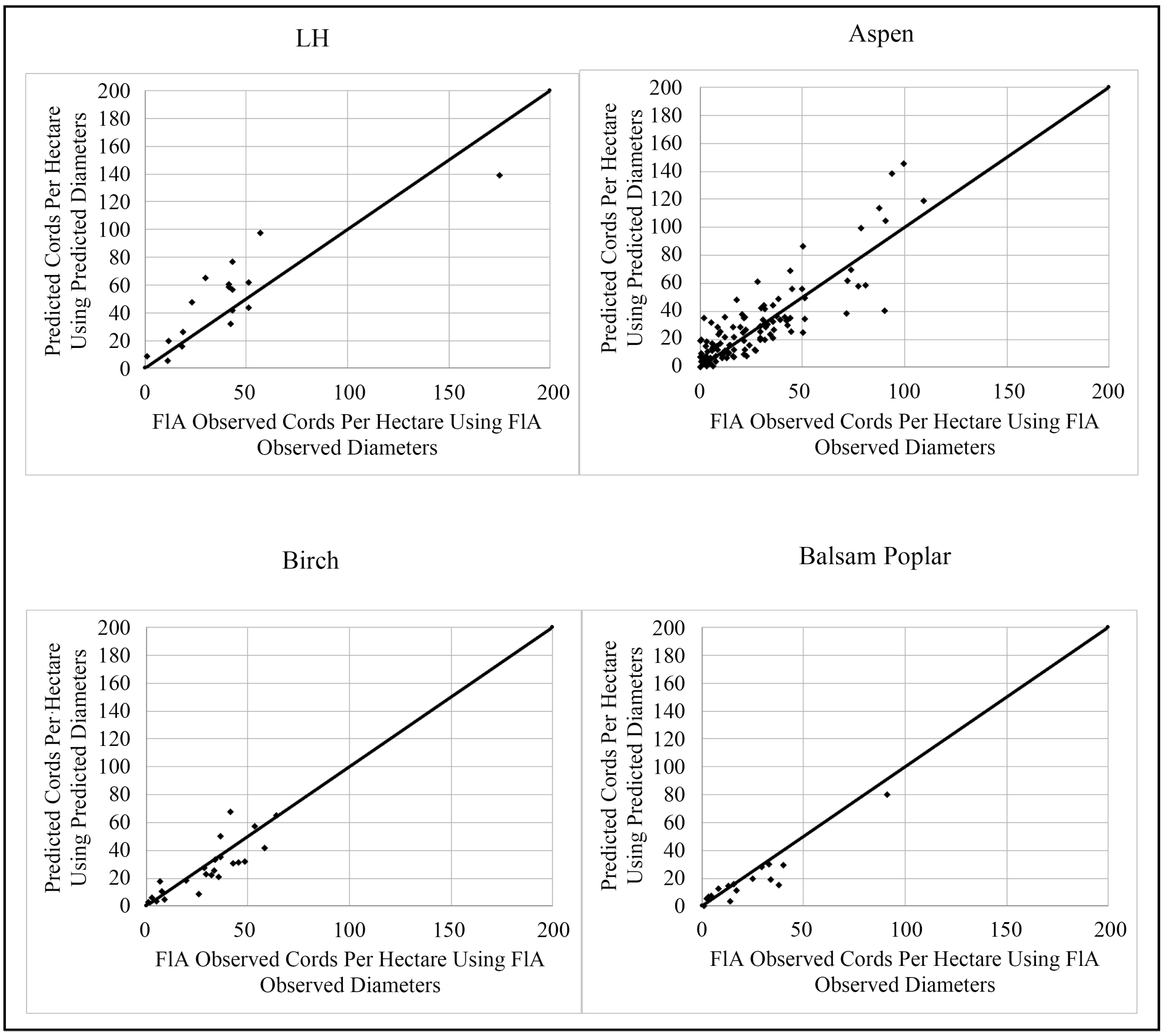

Figure 3. Predicted cords per hectare over observed cords per hectare (FIA observed diameters) by cover type when using a single VBAR (Method One) for a particular cover type. For the predicted cords, diameter at breast height (dbh) is predicted diameters using equations found in VanderSchaaf (2012) and VanderSchaaf (2013). Where: LH is Lowland Hardwood (n = 17), Aspen (n = 111), Birch (n = 23), and Balsam Poplar (n = 16).

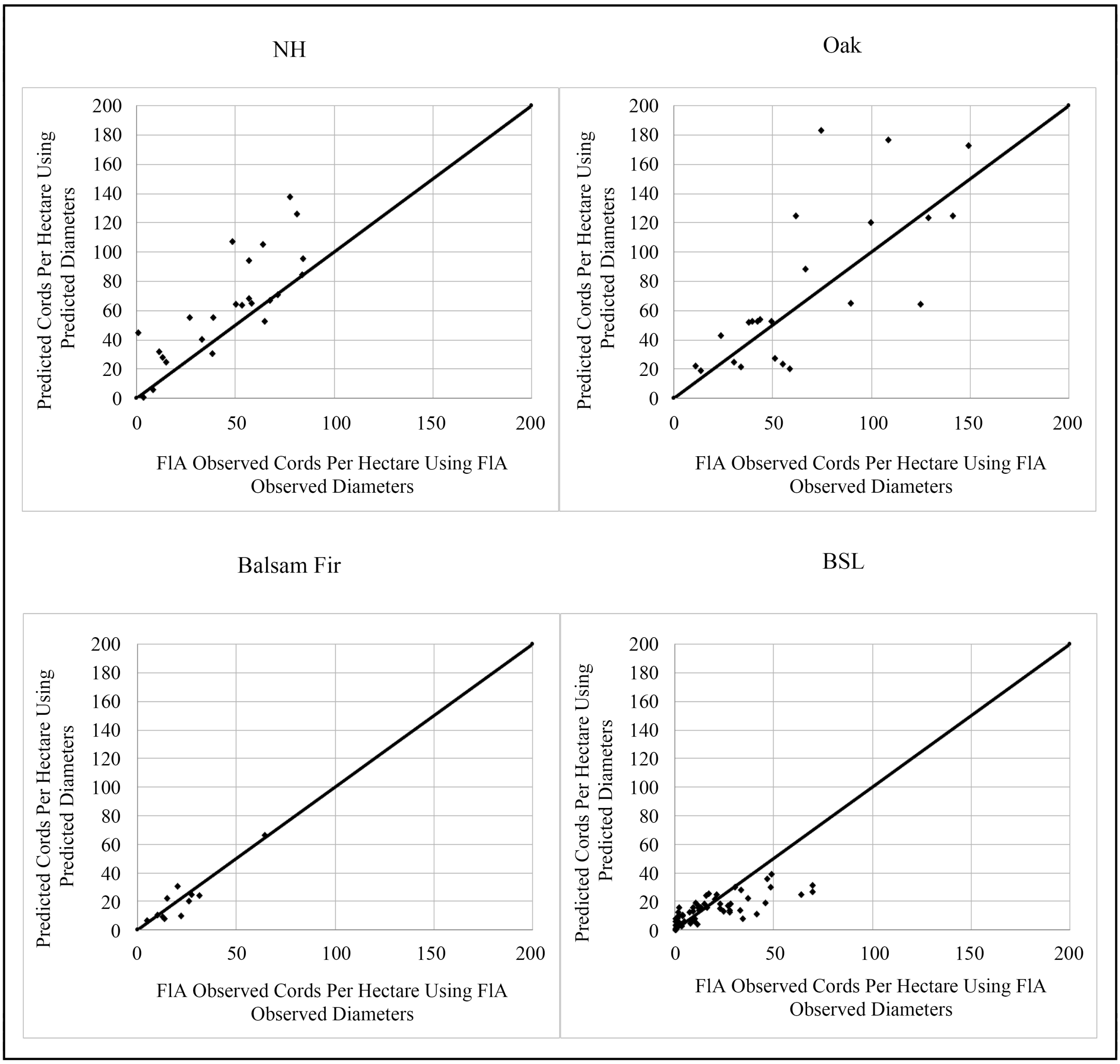

Figure 4. Predicted cords per hectare over observed cords per hectare (FIA observed diameters) by cover type when using a single VBAR (Method One) for a particular cover type. For the predicted cords, diameter at breast height (dbh) is predicted diameters using equations found in VanderSchaaf (2012) and VanderSchaaf (2013). Where: NH is Northern Hardwood (n = 24), Oak (n = 23), Balsam Fir (n = 11), and BSL is Black Spruce Lowland (n = 67).

Although LiDAR was not actually used here to determine H, based on this analysis where H is actually assumed to be known as obtained from the FIA data, we can infer that future research will need to concentrate on producing better predictions of D when determining H using LiDAR, or another source of remotely sensed data. At some point in the future, perhaps due to higher resolution imagery, reduced costs and wide availability of that higher resolution imagery, and continually more sophisticated image processing techniques, LiDAR may sufficiently measure D directly.

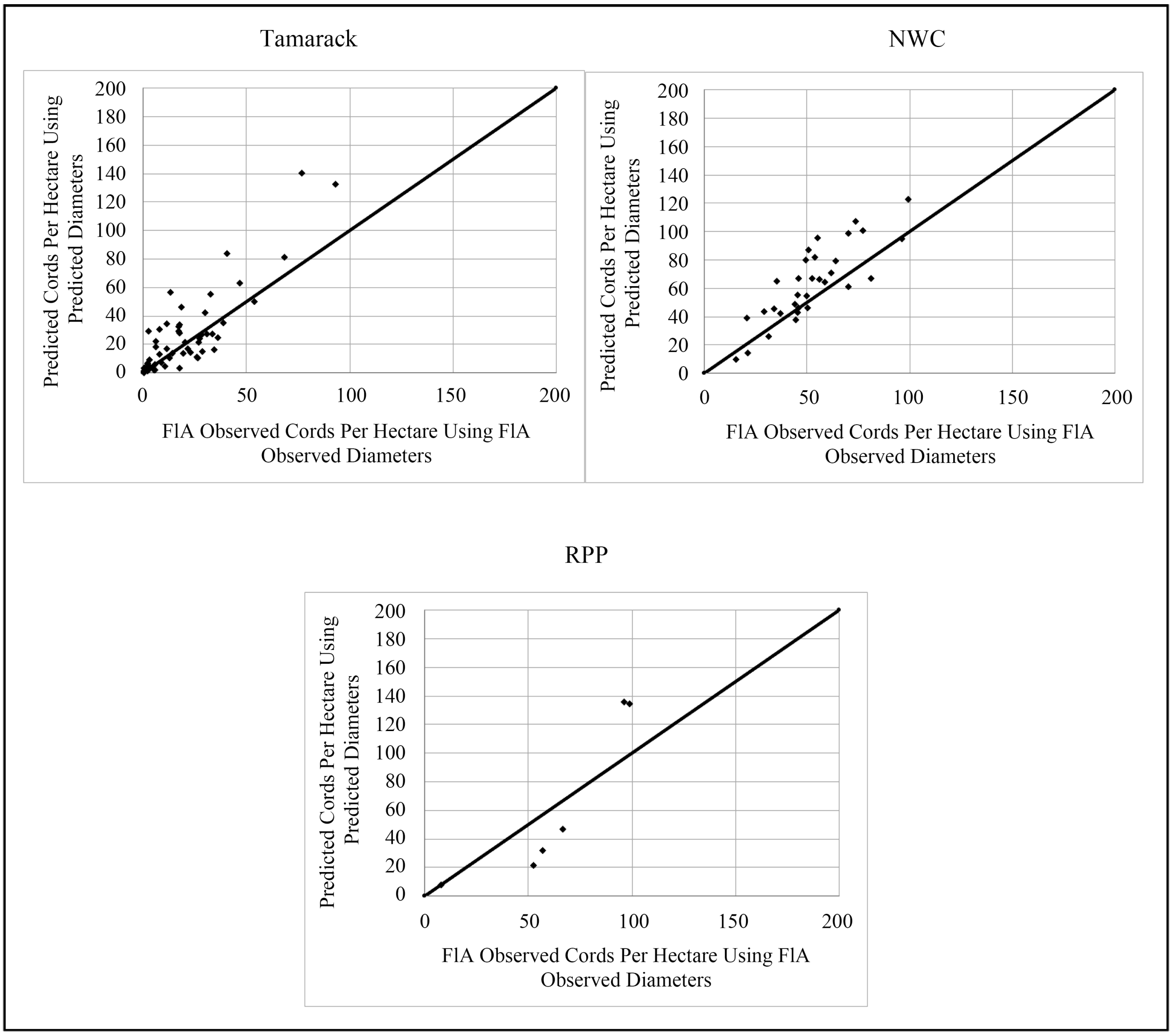

Figure 5. Predicted cords per hectare over observed cords per hectare (FIA observed diameters) by cover type when using a single VBAR (Method One) for a particular cover type. For the predicted cords, diameter at breast height (dbh) is predicted diameters using equations found in VanderSchaaf (2012) and VanderSchaaf (2013). Where: Tamarack (n = 56), NWC is Northern White Cedar (n = 32), and RPP is Red Pine Plantation (n = 6).

4. Conclusions

As expected, average VBARs differ substantially across cover types. At the current time at least, if the use of LiDAR is the preferred approach to ultimately estimate volume (cords) per hectare (acre), individual cover type VBARs will need to be used.

These data suggest basal area is not a favorable variable to predict VBARs. Even when knowing the "true" heights (since they were obtained from FIA) and predicting D as a function of species-specific H-D equations, predicted diameters using H-D equations often produced less than optimum basal area per acre (hectare) predictions. These poor results were observed despite the estimation process being nearly optimum since the "true" Hs and the species of each tree was known. Thus, since most likely at the current time, we will not be able to identify individual species from LiDAR data at an operational level, and that in practice, the Hs will be predicted using LiDAR metrics, and not directly measured in the field, operational predictions of basal area per hectare (acre) will be even poorer. Additionally, this study was conducted assuming trees per hectare (acre) is known, in practice trees per hectare (acre) would be estimated using the LiDAR data. Hence, when operationally using LiDAR, basal area per hectare (acre) predictions may have even more error associated with them.

However, the use of H-D equations that include stand-level variables may improve the estimate of D and hence basal area per hectare (acre) estimates. But as previously mentioned, operationally, the stand-level variables would be obtained from LiDAR data and thus estimates of these stand-level variables may not be as good as desired. These results suggest that site index and age will be better predictors of VBARs than basal area if LiDAR is used to estimate volume (cords) per hectare (acre).

In support for the use of LiDAR, it should be remembered though that field-based operational inventories use plots located within stands and that often only a small percentage of the total area is actually measured. Often, due to relatively high inventory costs largely due to poor stand access (e.g. tamarack swamp sites) and the costs of labor, a low number of plots are established, particularly during management plan inventories, and hence a relatively small amount of the stand is used to infer about the behavior across the entire stand. LiDAR has the advantage of "measuring" trees across the whole stand (Hummel et al., 2011; Watt et al., 2013; Brown et al., 2022). This current study doesn't quantify the errors associated with plot/point sampling versus LiDAR sampling, but this should be taken into consideration if one absolutely believes an entirely field-based sample is necessarily better than a LiDAR based sample.

An alternative approach is to predict volume per hectare (acre) directly as a function of basal area per hectare (acre) and to avoid VBARs. Future research should examine this alternative.

Acknowledgements

Helpful comments were received from a reviewer.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

Arney, J. D., & Corrao, M. V. (2021). *The Evolution of Forest Inventory (a Comparison of Satellite, Lidar, and Traditional Forest Inventory Methods)*. Forest Biometrics Research Institute.
https://forestbiometrics.org/references-articles/publications/the-evolution-of-forest-inventory/

Baskerville, G. L. (1972). Use of Logarithmic Regression in the Estimation of Plant Biomass. *Canadian Journal of Forest Research, 2,* 49-53. https://doi.org/10.1139/x72-009

Bechtold, W. A., & Scott, C. T. (2005). The Forest Inventory and Analysis Plot Design. In W. A. Bechtold, & P. L. Patterson (Eds.), *The Enhanced Forest Inventory and Analysis Program—National Sampling Design and Estimation Procedures* (pp. 27-42). U.S. Department of Agriculture, Forest Service, General Technical Report SRS-GTR-80.

Brown, S., Narine, L. L., & Gilbert, J. (2022). Using Airborne Lidar, Multispectral Imagery, and Field Inventory Data to Estimate Basal Area, Volume, and Aboveground Biomass in Heterogeneous Mixed Species Forests: A Case Study in Southern Alabama. *Remote Sensing, 14,* Article No. 2708. https://doi.org/10.3390/rs14112708

Burkhart, H. E., Avery, T. E., & Bullock, B. P. (2019). *Forest Measurements* (6th ed.). Waveland Press.

Gevorkiantz, S. R., & Olsen, L. P. (1955). *Composite Volume Tables for Timber and Their Application in the Lake States* (51 p.). U.S. Department of Agriculture Technical Bulletin 1104.

Hummel, S., Hudak, A. T., Uebler, E. H., Falkowski, M. J., & Megown, K. A. (2011). A Comparison of Accuracy and Cost of LiDAR versus Stand Exam Data for Landscape Management on the Malheur National Forest. *Journal of Forestry, 109,* 267-273. https://doi.org/10.1093/jof/109.5.267

Kepler, D. (2019). *Development of Innovative Cost-Saving Methodology for Forest Inventory* (42 p.). Environment and Natural Resources Trust Fund (ENRTF) M.L. 2016 Work Plan Final Report. https://www.lccmr.mn.gov/projects/2016/finals/2016_03o.pdf

Miles, P. D. (2013). *Thu Dec 05 10:21:57 CST 2013. Forest Inventory EVALIDator Web-Application version 1.5.1.05*. U.S. Department of Agriculture, Forest Service, Northern Research Station. http://apps.fs.fed.us/Evalidator/tmattribute.jsp

O'Connell, B. M., LaPoint, E. B., Turner, J. A., Ridley, T., Boyer, D., Wilson, A. M. et al. (2013). *The Forest Inventory and Analysis Database: Database Description and Users Manual Version 5.1.6 for Phase 2* (556 p.). U.S. Department of Agriculture, Forest Service.

SAS Institute Inc. (2004). *SAS/STAT® 9.4 User's Guide*. SAS Institute Inc.

Shiver, B. D., & Borders, B. E. (1996). *Sampling Techniques for Forest Resource Inventory* (356 p.). John Wiley & Sons, Inc.

Silva, C. A., Hudak, A. T., Vierling, L. A., Loudermilk, E. L., O'Brien, J. J., Hiers, J. K. et al. (2016). Imputation of Individual Longleaf Pine (*Pinus palustris* Mill.) Tree Attributes from Field and LiDAR Data. *Canadian Journal of Remote Sensing, 42,* 554-573. https://doi.org/10.1080/07038992.2016.1196582

St. Peter, J., Drake, J., Medley, P., & Ibeanusi, V. (2021). Forest Structural Estimates Derived Using a Practical, Open-Source Lidar-Processing Workflow. *Remote Sensing, 13,* Article No. 4763. https://doi.org/10.3390/rs13234763

USDA Forest Service (2013). *FIA DataMart 5.1*. https://research.fs.usda.gov/products/dataandtools/fia-datamart

VanderSchaaf, C. L. (2012). Mixed-Effects Height-Diameter Models for Commercially and Ecologically Important Conifers in Minnesota. *Northern Journal of Applied Forestry, 29,* 15-20. https://doi.org/10.5849/njaf.11-005

VanderSchaaf, C. L. (2013). Mixed-Effects Height-Diameter Models for Commercially and Ecologically Important Hardwoods in Minnesota. *Northern Journal of Applied Forestry, 30,* 37-42. https://doi.org/10.5849/njaf.12-018

Watt, M. S., Meredith, A., Watt, P., & Gunn, A. (2013). Use of LiDAR to Estimate Stand

Characteristics for Thinning Operations in Young Douglas-Fir Plantations. *New Zealand Journal of Forestry Science, 43,* 18. https://doi.org/10.1186/1179-5395-43-18

White, J. C., Coops, N. C., Wulder, M. A., Vastaranta, M., Hilker, T., & Tompalski, P. (2016). Remote Sensing Technologies for Enhancing Forest Inventories: A Review. *Canadian Journal of Remote Sensing, 42,* 619-641. https://doi.org/10.1080/07038992.2016.1207484