Single Photon LiDAR
Petawawa Research Forest Implementation

PREPARD FOR:
Joanne White
Canadian Forest Service

PREPARED BY:
Murray Woods

Margaret Penner
Forest Analysis Ltd.

April 6, 2020

Summary

A Single Photon LiDAR (SPL) enhanced forest inventory was developed for the Petawawa Research Forest (PRF). The SPL and field calibration data used are described herein, as is the validation data used to assess the results. A number of options were investigated and are summarized in the appendices to this report. These options included investigating the effect of reducing the number of predictors used in model development, altering the height threshold used in SPL metric calculation (from 1.3 m to 0 m), and altering which SPL returns were included in the SPL metric calculation (i.e. all returns or just vegetation returns), as well as different options for constraining gross total volume of merchantable trees to be less than or equal to gross total volume.

The final inventory was produced using the full set of SPL metrics as predictors, with metrics calculated using all returns and no height threshold, and using volume/basal area ratios in the volume constraints for calculating merchantable volumes. The final models were validated on 27 independent and intensively sampled stands. For gross total volume of merchantable stems, the average overall bias was 1 m^3/ha ± 7 m^3/ha (standard error), with the greatest overestimation for managed white pine stands and the greatest underestimation for red pine plantations. The results for merchantable volume were similar to those of gross total volume of merchantable stems (2 m^3/ha ± 7 m^3/ha). Basal area (BA) of merchantable stems was underestimated by about 2% (0.5 m^2/ha ± 0.7 m^2/ha), and quadratic mean Dbh of merchantable stems was underestimated by 1% overall (0.2 cm ± 0.6 cm).

Table of Contents

1. Introduction

A Single Photon LiDAR (SPL) Inventory was developed for the Petawawa Research Forest (PRF). This report summarizes the final results including evaluation of the predictions against an independent validation dataset.

2. Data

2.1 Single Photon Lidar

SPL was flown for the PRF and Canadian Nuclear Laboratories (CNL) site, near Chalk River Ontario July 2, 2018. There are many options for generating predictors from the SPL point cloud. After testing many alternatives (summarized in the appendices), this study used LAStools[1] predictors using the full point cloud with no threshold (Table 1). There are some dependencies between the predictors. For instance, a_d0_2 = 100 – a_dns_2m.

Table 1. The SPL predictors were generated using LAStools using the full point cloud with all returns. No threshold was applied in metric calculation (with the exceptions noted in the table).

Metric	Height threshold used in metric calculation (m)	Description
a_std_95	0	STD_Trimmed @95%
a_ske_95	0	Skewness_Trimmed @95%
a_kur_95	0	Kurtosis_Trimmed @95%
a_avg	0	avgHt
a_qav	0	average_Square_Ht
a_p01	0	1st Percentile Height
a_p05	0	5th Percentile Height
a_p10	0	10th Percentile Height
a_p20	0	20th Percentile Height
⋮		
a_p90	0	90th Percentile Height
a_p95	0	95th Percentile Height
a_p99	0	99th Percentile Height
a_d0_2	0	Number of returns from 0-2m/All returns
a_d2_4	0	Number of returns from 2-4m/All returns
⋮		
a_d44_46	0	Number of returns from 44-46m/All returns
a_d46_48	0	Number of returns from 46-48m/All returns
a_b10	0	decile 10% of points between 0 and 99% height
a_b20	0	decile 20% of points between 0 and 99% height
⋮		
a_b80	0	decile 80% of points between 0 and 99% height
a_b90	0	decile 90% of points between 0 and 99% height
a_dns_2m	2	Density_Percentage of All Returns 2m-49m/All Returns
a_dns_4m	4	Density_Percentage of All Returns 4m-49m/All Returns
a_dns_5m	5	Density_Percentage of All Returns 5m-49m/All Returns
a_dns_6m	6	Density_Percentage of All Returns 6m-49m/All Returns
a_dns_10m	10	Density_Percentage of All Returns 10m-49m/All Returns
a_dns_12m	12	Density_Percentage of All Returns 12m-49m/All Returns
a_dns_14m	14	Density_Percentage of All Returns 14m-49m/All Returns
a_dns_15m	15	Density_Percentage of All Returns 15m-49m/All Returns
a_dns_16m	16	Density_Percentage of All Returns 16m-49m/All Returns
a_dns_18m	18	Density_Percentage of All Returns 18m-49m/All Returns
a_dns_20m	20	Density_Percentage of All Returns 20m-49m/All Returns
a_dns_25m	25	Density_Percentage of All Returns 25m-49m/All Returns
a_vci_1mbin	0	Vertical Complexity Index with a 1 m bin
a_vci_0.5bin	0	Vertical Complexity Index with a 0.5 m bin

[1] Martin Isenburg, LAStools - efficient tools for LiDAR processing. Version 190604, http://lastools.org.

2.2 Calibration data

249 Ground calibration plots were measured on the PRF during the summer of 2018 and an additional 20 plots were measured on the CNL lands in the summer of 2019 (the impact of the additional 20 CNL plots is analyzed in Section 8 – Appendix C). The majority of the plots were previously established for a 2012 LiDAR inventory. The locations chosen for the 2012 field plots followed a LiDAR PCA matrix developed to ensure full coverage of the LiDAR conditions. In 2018/2019, additional plots were added to the 2012 plot network to expand the number of observations and ensure coverage of the main forest types of the PRF. Live and Dead trees with a Diameter at breast height (Dbh) > 9.0 cm were measured on 625 m^2 (14.1m radius) ground plots. Trees with 2.5 < Dbh ≤ 9 cm were measured on 50 m^2 (3.99m radius) ground plots, centered on the larger plot. Dbh was measured on all trees and a subsample of trees were measured for height.

Height-Dbh curves were fit at the plot level, all species combined, and used to estimate heights of the trees without measured heights. The average dominant/codominant height was calculated as the average height of the live dominant/codominant/emergent trees with measured heights. Only trees on the large tree plot had crown status recorded. The Sharma (2016) equation was used.

(1) $$Ht = 1.3 + \alpha \cdot SHt^{\delta} \cdot \left(1 - e^{-\beta \cdot (TPH/BA)^{\cdot \varphi} Dbh}\right)^{\gamma}$$

Where:

Ht is total tree height (m)
Dbh is Diameter at breast height (cm)
BA is stand basal area (m^2/ha)
TPH is stand density (trees/ha)
SHt is stand height (average of the dominant/codominant height of the plot)
α, β. γ, δ and φ are parameters to be estimated.

Stem volumes were estimated using the Zakrzewski & Penner (2013) and Sharma & Parton (2009) models. Biomass was estimated using the equations of Lambert et al (2005) using height and Dbh.

Table 2. The ground plots cover a range of forest types. The means are followed in brackets by the range. The maximum values on a 625 m^2 plot are expected to be larger than the maximum values at the polygon level.

Strata	N	Gross total volume (m^3/ha)	Basal area (m^2/ha)	Trees/ha	Dom/codom height (m)	Quadratic mean Dbh (cm)
Intolerant Hwd	15	402 (143 - 802)	36.3 (15.8 - 55.3)	1905 (392 - 4200)	25.3 (14.8 - 33.1)	17.1 (8.1 - 23.5)
Lowland Con	4	233 (79 - 370)	33.2 (14 - 46.7)	2902 (1312 - 6224)	16.9 (12.6 - 20.4)	13.3 (9.2 - 17.4)
MIXED hardwood	28	207 (35 - 350)	25.2 (9.2 - 38.2)	2171 (552 - 5656)	19.2 (11.8 - 24.8)	13.6 (4.6 - 21.5)
MIXED conifer	13	163 (58 - 310)	22.5 (8.5 - 39.3)	2014 (272 - 5512)	16.7 (11.1 - 21.2)	14.3 (7.2 - 20)
Mid tolerant Hwd	13	173 (3 - 299)	20.6 (0.5 - 32)	1726 (16 - 6280)	19.7 (13.8 - 27)	16.1 (4.9 - 34.3)
Pine Oak	14	211 (31 - 352)	27.1 (5.7 - 39.2)	1643 (48 - 2824)	18 (11.5 - 24.5)	16.3 (5.7 - 38.8)
Pj Plant	10	181 (115 - 264)	20.4 (15.8 - 26.3)	1377 (480 - 2976)	19.8 (16.3 - 22.9)	15.1 (9.7 - 21.5)
Pr Plant	23	431 (116 - 1000)	40.5 (24.4 - 70.2)	1695 (408 - 3312)	22.5 (9.2 - 33.9)	19 (10.5 - 28.6)
Pw Plant	7	195 (14 - 435)	23.5 (6.7 - 44.4)	1239 (256 - 2456)	19.8 (5.6 - 37.2)	17.4 (8.9 - 34.3)
Pw Pr	93	371 (20 - 1067)	33.8 (2.9 - 68.4)	2082 (96 - 15392)	26 (8.2 - 43)	17.9 (6 - 39.9)
Sb	14	156 (38 - 250)	21 (10.6 - 34)	1738 (592 - 2848)	16.3 (9.7 - 19.2)	13.2 (6.9 - 20.5)
Spruce Plant	12	237 (110 - 459)	29.8 (12.3 - 52.5)	1705 (424 - 2968)	19 (13.8 - 26.5)	16.1 (10.7 - 26.4)
Tolerant Hwd	23	273 (45 - 493)	29 (7.8 - 44)	1507 (360 - 4400)	23.9 (12.3 - 32.2)	16.9 (8.3 - 30.4)
All	269	293 (3 - 1067)	30 (0.5 - 70.2)	1885 (16 - 15392)	22.2 (5.6 - 43)	16.6 (4.6 - 39.9)

2.3 Validation data

For validation, 27 forest stands were selected and intensively sampled using BAF2 prism sweeps on a 50m grid. The stands selected covered a range of mature stand types found at the PRF.

The species and Dbh of all live "in" trees with Dbh > 9.0 were recorded. The largest Dbh tree on each prism sweep was measured for height. Crown class was not recorded. The height trees were averaged by stand and this was used as the dom/codom height. Equation (1) was fit at the stand level using this dom/codom height. However, only 3 height trees had a Dbh < 20 cm (of the 1,000+ height trees) and there are about 3,500 trees with Dbh < 20cm that required the prediction of height.

Stand 465 (Figure 1) illustrates the problem. There is almost no relationship between Dbh and height and when equation (1) was fit at the stand level, it predicted tall heights for small Dbh trees for some stands. The reverse also happened in some stands – height was underestimated for smaller Dbh trees). For some plots there were no observations to anchor the left side of the curve.

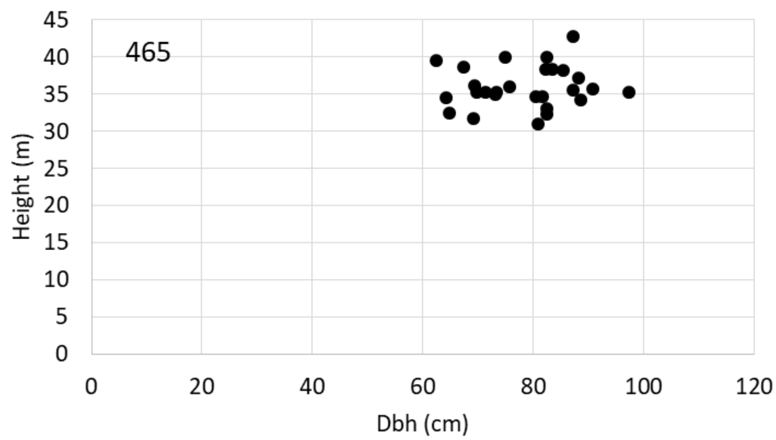

Figure 1. The trees measured for height in stand 465 have almost no relationship between Dbh and height, making hard to calibrate a ht-Dbh curve.

Because there were so few height measurements for small trees, fitting at the population level (with stand level localization using dom/codom height and stand BA) wasn't much better. The issue was the lack of small trees. Options to incorporate the calibration plot ht-Dbh data were investigated. This led to the search for a common method to calculate dom/codom height on the calibration and validation plots.

The largest Dbh tree on the prism plots was measured for height. The average BA was approximately 26 m^2/ha so approximately 13 trees were measured/sweep. The calibration plots had an average of 688 merchantable stems/ha or approximately 43 trees/plot. To get an average height for the calibration plots that was roughly comparable to the validation plot height, the 3 thickest trees measured for height on each plot were selected. This was called "BigHt" and treated the same as the average of the measured heights on the validation plots at the stand level. Then the measured heights from the calibration and validation plots were pooled and used to calibrate species level ht-Dbh curves with the following exceptions.

There were some species with not enough ht-DBH measurements. Balsam poplar trees were combined with trembling aspen to calibrate a poplar ht-Dbh curve. Elm, ironwood and black cherry were pooled to estimate a single ht-Dbh curve that was used for all three species.

Four of the validation stands were planted. Stands 186, 455 and 192 were red pine and stand 191 was planted jack pine. All the red pine in the first three stands was considered planted and all the jack pine in the last stand was considered planted. All other stems were assumed to be natural origin. Ht-Dbh curves were fit to natural origin stems by species and for planted red pine and planted jack pine.

Top height was taken as the average height of the trees measured for height (the thickest tree on each point sample). Lorey's height was calculated for each prism sweep. The stand Lorey height is the arithmetic average of the Lorey height of the point samples.

The 27 polygons were selected so that 3 stands were sampled for each of 7 forest types. These "planned" forest types were the basis for most of the analyses in the appendices. Based on the field samples, the forest types were revised and stands manually assigned to their appropriate forest types. These are the "actual" forest types reported here (Table 3).

Table 3. The validation polygons are summarized by forest type. The mean is followed by the range (at the polygon level). As expected, the maximum values at the polygon level are smaller than the maximum values at the plot level (Table 2).

Actual Forest Type	N	Stations/ polygon	Gross total Volume (m^3/ha) (Dbh > 9.0 cm)	Gross merchantable Volume (m^3/ha)	Number of sample points	Area (ha)
BlackSpruce	2	22 (21 - 22)	185 (178 - 192)	152 (145 - 159)	22 (21 - 22)	7 (7 - 8)
JackPine	2	28 (20 - 36)	206 (198 - 214)	183 (178 - 188)	28 (20 - 36)	9 (6 - 12)
LowlandConifer	1	47 (47 - 47)	150 (150 - 150)	125 (125 - 125)	47 (47 - 47)	15 (15 - 15)
Mixedwood	4	53 (48 - 60)	221 (205 - 229)	168 (151 - 176)	53 (48 - 60)	18 (16 - 19)
Oak	2	55 (55 - 55)	279 (247 - 310)	229 (202 - 256)	55 (55 - 55)	19 (19 - 20)
Poplar	2	18 (11 - 24)	255 (223 - 286)	210 (187 - 232)	18 (11 - 24)	7 (4 - 9)
Pr Plantation	4	16 (12 - 25)	330 (231 - 429)	295 (216 - 396)	16 (12 - 25)	5 (4 - 9)
PwManaged	3	36 (26 - 48)	174 (103 - 213)	161 (95 - 201)	36 (26 - 48)	12 (9 - 16)
PwNatural	4	40 (24 - 60)	308 (292 - 316)	266 (249 - 285)	40 (24 - 60)	13 (10 - 19)
TolerantHwd	3	56 (43 - 63)	233 (206 - 249)	178 (149 - 195)	56 (43 - 63)	18 (14 - 21)
All	27	37 (11 - 63)	246 (103 - 429)	208 (95 - 396)	37 (11 - 63)	13 (4 - 21)

2.4 Forest inventory attributes to be predicted

The attributes to be predicted directly are given in Table 4. Several options were explored for predicting TVOL_merch and MVOL (see Section 9, Appendix D). Only the recommended options are included here.

Table 4. The inventory attributes that were predicted directly are defined. Merchantable attributes are only predicted when p99 > 5m.

Attribute	Definitions	Size
topht	top height (average height of thickest 6 trees/plot) (m)	All
HL_all	Lorey height, all stems (average height weighted by BA) (m)	All
CD_ht	average height of dominant/codominant trees (m)	All
DQ_all	Quadratic mean Dbh, all stems (cm)	All
BA_all	Basal area, all stems (m^2/ha)	All
TVOL_all	gross total volume, all stems (m^3/ha)	All
BIO_all	aboveground biomass, all stems (kg/ha)	All
DQ_merch	Quadratic mean Dbh, merchantable stems (cm)	Dbh > 9
Vbar_TVOL_ratio	Vbar_TVOL_merch/vbar_tvol	Dbh > 9
Vbar_mvol_ratio	Vbar_mvol/vbar_TVOL_merch	Dbh > 9
BA_merch_ratio	BA_merch/BA_all	Dbh > 9
HL_merch_ratio	HL_merch/HL_all	Dbh > 9
BIO_merch_ratio	BIO_merch/ BIO _all	Dbh > 9

The attributes in Table 5 are calculated from the predictions in Table 4.

Table 5. The forest inventory attributes that were not predicted directly are defined.

Attribute	Definitions	Size	Calculation (from predicted values)

Attribute	Definitions	Size	Calculation (from predicted values)
TPH_all	stems/ha, all stems	All	BA_all/ (DQ_all*DQ_all*0.00007854)
HL_merch	Lorey height, merchantable stems (m)	Dbh > 9	HL_all*HL_merch_ratio
BA_merch	Basal area, merchantable stems (m^2/ha)	Dbh > 9	BA_all* BA_merch_ratio
tph_merch	stems/ha, merchantable stems	Dbh > 9	BA_merch/ (DQ_merch*DQ_merch*0.00007854)
TVOL_merch[2]	gross total volume, merchantable stems (m^3/ha)	Dbh > 9	Tvol_all*vbar_TVOL_ratio*ba_merch_ratio
BIO_merch	aboveground biomass, merchantable stems (kg/ha)	Dbh > 9	BIO_all*BIO_merch_ratio
MVOL	merchantable stem volume (m^3/ha)	Dbh > 9	TVOL_merch*vbar_mvol_ratio

Volume, basal area and DQ were predicted by the Dbh classes in Table 6.

Table 6. The size classes are defined by Dbh range.

Size	Size class	Dbh range (cm)
Large	Large sawlog	Dbh > 49
Med	Medium sawlog	37 < Dbh ≤ 49
Small	Small sawlog	25 < Dbh ≤ 37
Pole	Polewood	9 < Dbh ≤ 25

The size class attributes predicted directly are given in Table 7.

Table 7. The size class attributes that were predicted are given.

Attribute	Definitions	Size	Calculation
baLarge_frac	Basal area (m^2/ha)	large	BALarge/BA_merch
baMedium_frac	Basal area (m^2/ha)	medium	BAMedium/(BA_merch – BALarge)
baSmall_frac	Basal area (m^2/ha)	small	BASmall/(BA_merch – BALarge – BAMedium)
mvolLarge_frac	gross merchantable volume	large	mvolLarge/mvol
mvolMedium_frac	gross merchantable volume	medium	mvolMedium/(mvol – mvolLarge)
mvolSmall_frac	gross merchantable volume	small	mvolSmall/(mvol – mvolLarge – mvolMedium)
tvolLarge_frac	gross total volume	large	tvolLarge /TVOL_merch
tvolMedium_frac	gross total volume	medium	tvolMedium/(TVOL_merch – tvolLarge)
tvolSmall_frac	gross total volume	small	tvolSmall/(tvol_mvol – tvolLarge – tvolMedium)
bioLarge_frac	aboveground biomass	large	BioLarge/BIO_merch
bioMedium_frac	aboveground biomass	medium	BioMedium/(BIO_merch – BioLarge)
bioSmall_frac	aboveground biomass	small	BiomSmall/ (BIO_merch - BioLarge – BioMedium)
DQ_Poles	Quadratic mean Dbh (cm)	poles	
DQ_Large	Quadratic mean Dbh (cm)	large	
DQ_Medium	Quadratic mean Dbh (cm)	medium	
DQ_Small	Quadratic mean Dbh (cm)	small	

The size class attributes calculated from the predicted attributes in Table 7 are given in Table 8

Table 8. The calculation of the size class attributes for BA and TPH are given. The calculations for mvol, tvol and biomass are similar to those for BA.

Attribute	Definitions	Size	Calculation (from predicted attributes)
BALarge	Basal area	large	BA_merch*baLarge_frac
BAMedium	Basal area	medium	(BA_merch – BALarge)* baMedium _frac
BASmall	Basal area	small	(BA_merch – BALarge – BAMedium)* baSmall _frac
BAPoles	Basal area	poles	BA_merch - BALarge - BAMedium - BASmall
TPH_Poles	stems/ha	poles	BA_Poles/(DQ_Poles DQ_Poles*0.00007854)
TPH_Large	stems/ha	large	BA_Large/(DQ_Large*DQ_Large*0.00007854)
TPH_Medium	stems/ha	medium	BA_Medium/(DQ_Medium*DQ_Medium*0.00007854)
TPH_Small	stems/ha	small	BA_Small/(DQ_Small*DQ_Small*0.00007854)

[2] Refer to Section 9 – Appendix D

3. Methodology

Prior to generating models, the calibration data were examined to ensure there were no outstanding issues.

3.1 Data cleaning

There were some ground samples where the Lorey height was taller than the top height (Table 9). This generally occurred when the overstory consisted of a few large trees with an understory of small trees as in the case of pine shelterwood or seed tree management. Top height calculations in stands with a sparse overstorey layer or with scattered veteran trees can be problematic. For a plot size of 625 m^2, the heights of the 6 thickest trees (approximately 100 trees per ha) are averaged for Top Height. A pine stand, managed with uniform shelterwood and a 50% crown spacing could have an overstorey of 32+m tall white and red pine, but may have only 1, 2, 3 or maybe 4 overstory trees on a 625 m^2 plot. The result is that smaller/shorter trees are included in the calculation of Top Height. This discrepancy is captured in Figure 2 for the PwPr forest type. The trees included in the calculation of top height should be restricted to the main canopy and trees from secondary layers excluded; however trees were not assigned to layers in the field surveys.

Table 9. The samples where (Lorey height – top height) > 3m are given.

Plot	Species composition	Trees/ha (Dbh > 9 cm)	Lorey height (m)	Top height (m)	Dom/codom height (m)	
PRF043	Pw65Sw14Bf11Mr10	32	16.5	8.3	21.9	
PRF046	Pw99Po1	32	33.6	23.0	33.9	
PRF064	Pw81Mr18Sw1	160	29.4	25.7	34.1	three large Pw
PRF066	Pw60Sw16Or10Mr10Po2Bw1Pr1	880	23.6	17.6	11.4	one large Sw
PRF080	Pw50Pr41Or4Bf4Sw1	176	24.3	19.3	31.0	three large pine
PRF330	Pr52Pw45Mr3	96	32.3	26.8	32.5	four large pine

There were a number of plots where the top height was significantly greater than p95 (Table 10). This can occur when the canopy is dense and there are few ground returns (Figure 3). It can also occur when there is only one tall tree in the plot and there is a large difference between p99 and top height (Figure 2). It can also occur when tall trees are leaning out of the plot and are included in the ground plot summary but not in the SPL summaries. It may also occur if the ground plot and the clipped SPL cloud don't line up exactly. Another potential reason is that these often extremely tall trees may be difficult to measure accurately in the field. All of these conditions are expected to occur in the population and, although they add to the unexplained variation in the predictions, these conditions are not expected to cause bias.

Table 10. The samples where (top height – p95) > 4m are given. Note the species composition for some plots is quite mixed.

Plot	Species composition	Trees/ha (Dbh > 9 cm)	Lorey height (m)	Top height (m)	Dom/codom height (m)	P95	P99
PRF017	Pj57Pw36Po3Bw2Pr1Mr1	608	18.8	25.4	21.6	20.1	22.4
PRF041	Or100	16	13.8	13.8	13.8	3.1	10.7
PRF043	Pw65Sw14Bf11Mr10	32	16.5	8.3	21.9	3.5	26.9
PRF193	Ms65Be35Iw0	432	29.8	32.9	30.7	28.1	28.7
PRF307	Sb58La35Bf7Mr0	816	15.7	21.2	17.6	16.8	19.1
PRF317	Pw91Mr8Bf1	320	39.5	42.3	42.3	38.1	39.5
PRF350	Po62Mr21Pw13Iw4	400	16.0	22.5	22.5	17.5	25.4
PRF356	Po83Mr6Or5Ms3Iw3	816	28.4	33.5	30.7	28.8	29.4
PRF362	Mr49Be14Bf14Sw11Pw8Po4	192	8.8	13.7	12.6	9.2	14.3
PRF366	Po69Ms29Bf2	832	28.3	35.1	31.7	30.6	31.4

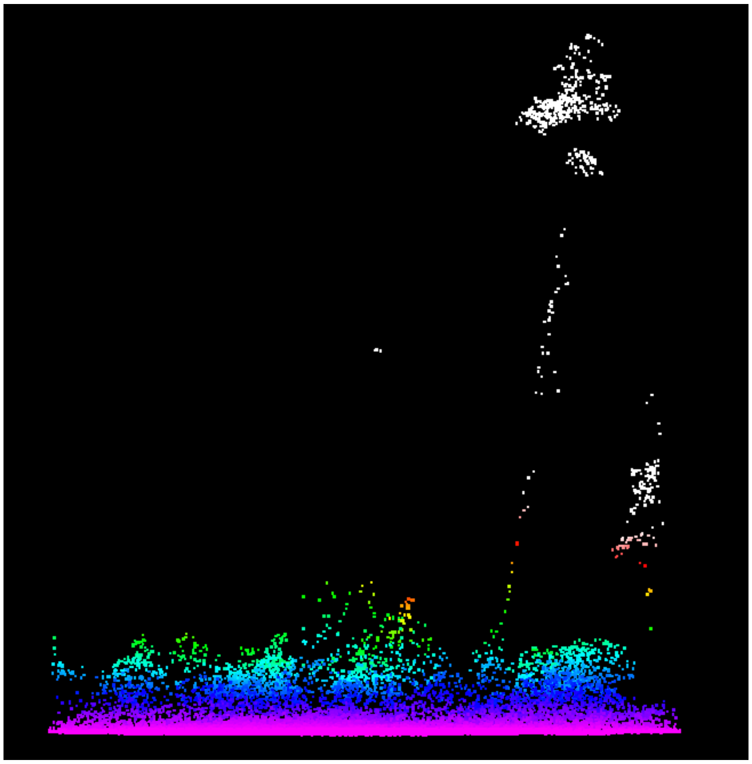

Figure 2. Plot PRF043. This is an example of a white pine shelterwood, after the final removal of overstorey trees.

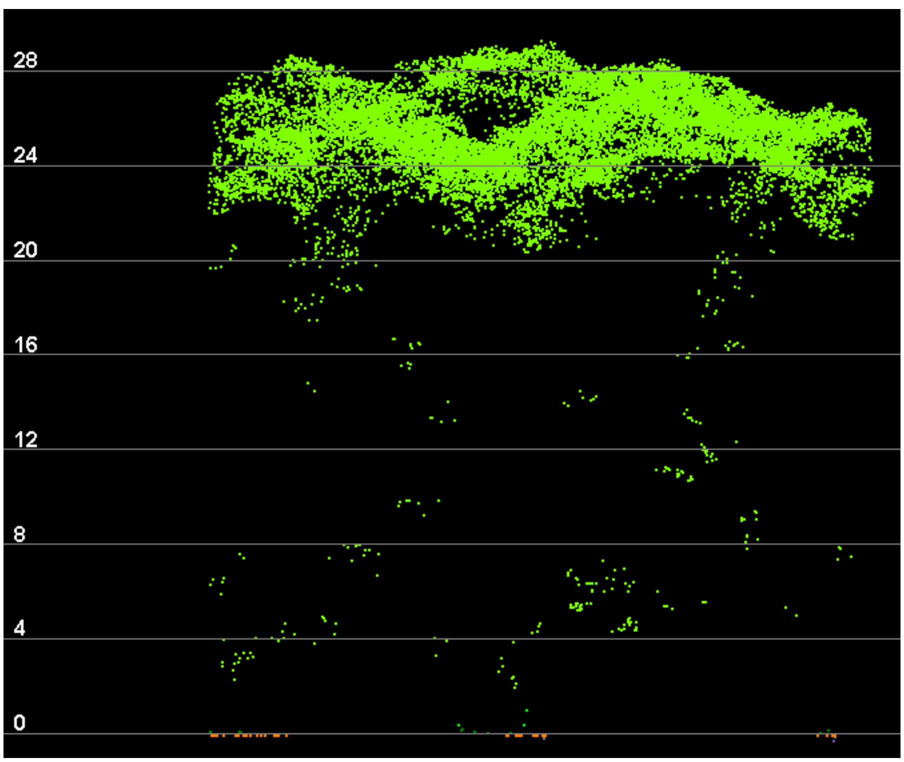

Figure 3. The top height on plot PRF193 is taller than the maximum SPL return height. The plot has a very dense maple canopy with very few ground returns.

3.2 Stratification

There was no stratification for model development. There was a wide range of species compositions and types of management. Parametric predictions would likely require stratification by species and possibly management activities. It was assumed that nonparametric approaches will not require stratification.

3.3 RandomForest

The R package RandomForest (RF - Liaw and Wiener 2002) was used for the nonparametric modeling. Various options for reducing the number of predictors were examined (Section 7 – Appendix B). The differences between using all the predictors and subsets of predictors were minor and the differences in using mtry of p/2 and p/3 were minor. The consensus was to use all predictors and mtry = p/3.

4. Results

The bias is the difference between the observed Y_{obs} and predicted Y_{pred} attribute and the average bias was calculated as follows.

$$\overline{bias} = \frac{\sum(Y_{obs} - Y_{pred})}{n} = \frac{\sum bias_i}{n}$$

The standard error (SE) of the bias was calculated as follows. It is a measure of how consistent the bias is. When the bias is reported (e.g., Table 14 and Figure 11b), the standard error is reported.

$$SE_{bias} = \sqrt{\frac{\sum(bias_i - \overline{bias})^2}{n} \Big/ n} = \sigma_{bias} / \sqrt{n}$$

The root mean squared error (RMSE) is another measure of the goodness of the predictions and was calculated as follows.

$$RMSE = \sqrt{\frac{\sum(Y_{obs} - Y_{pred})^2}{n}}$$

4.1 Summary of calibration results

Results are given here for CD_ht, TVOL_merch and BA_merch. The full set of predictions for the calibration plots is available in a separate file.The “predict” function in the RF package was used to obtain the out of bag (OOB) predictions, a form of internal cross validation (White et al. 2017).

4.1.1 CD_ht

Predictions of Dom/codom height on the calibration data are reported in Table 11, and illustrated in Figure 4 and Figure 5. Generally bias was ≤ 3%, except for mixed conifers where bias was -6%. Relative RMSE varied by forest type, ranging from a low of 4% in lowland conifer and red pine plantations, to a high of 19% for managed pine stands (PwPr).

Table 11. The Dom/codom height results are given. SE = standard error of the mean. RMSE = Root Mean Squared Error. The bias and RMSE are given as a percent of the observed mean.

Forest Type	N	Observed (m)	Predicted (m)	Bias + SE (m)	% bias	RMSE (m)	% RMSE
IntHwd	15	25.3	24.4	0.8 ± 0.6	3%	2.5	10%
LowCon	4	16.9	17.2	-0.3 ± 0.4	-2%	0.7	4%
MIXEDH	28	19.2	19.5	-0.4 ± 0.5	-2%	2.6	14%
MIXEDC	13	16.7	17.7	-0.9 ± 0.7	-6%	2.7	16%
MidHwd	13	19.7	19.4	0.3 ± 0.5	2%	1.7	9%
PineOak	14	18	18.2	-0.2 ± 0.6	-1%	2.1	12%
PjPlant	10	19.8	19.2	0.6 ± 0.3	3%	1.2	6%
PrPlant	23	22.5	22.3	0.2 ± 0.2	1%	0.9	4%
PwPlant	7	19.8	19.1	0.7 ± 1.2	4%	2.9	15%
PwPr	93	26	25.7	0.3 ± 0.5	1%	5	19%
Sb	14	16.3	16.6	-0.3 ± 0.2	-2%	0.8	5%
Spruce Plant	12	19	18.3	0.7 ± 0.4	4%	1.5	8%
TolHwd	23	23.9	24.1	-0.2 ± 0.4	-1%	1.9	8%
All	269	22.2	22.1	0.1 ± 0.2	0%	3.3	15%

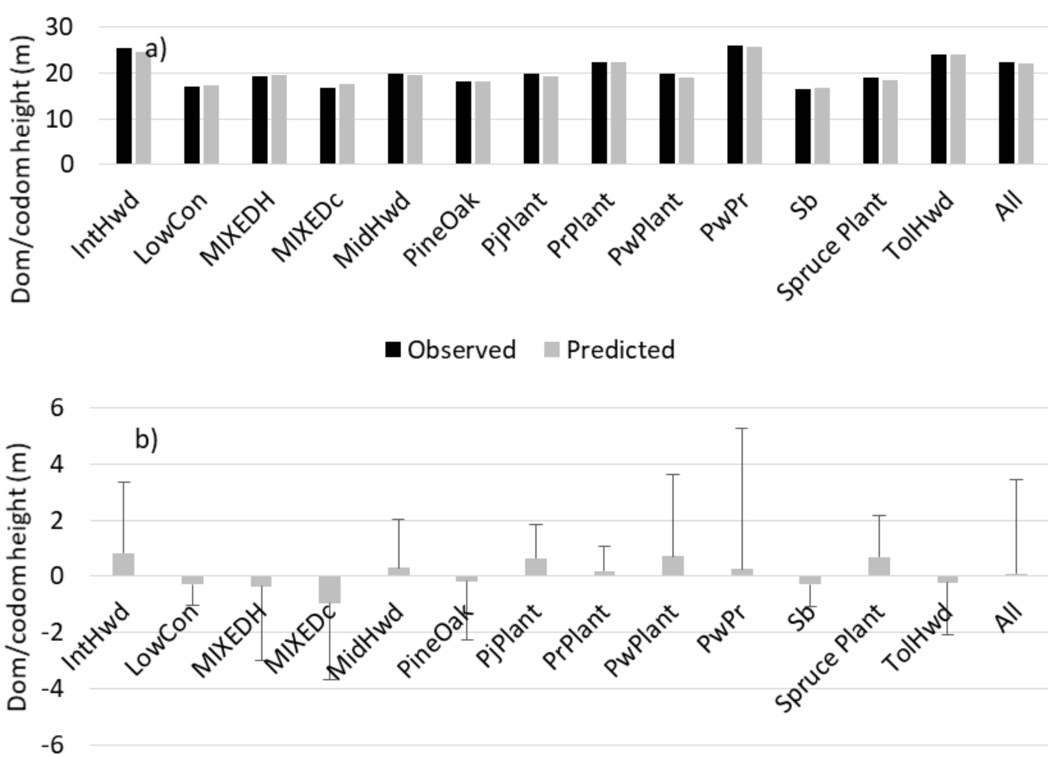

Figure 4. The observed and predicted Dom/codom height (a) and the bias and standard error bars (b) are given by forest type.

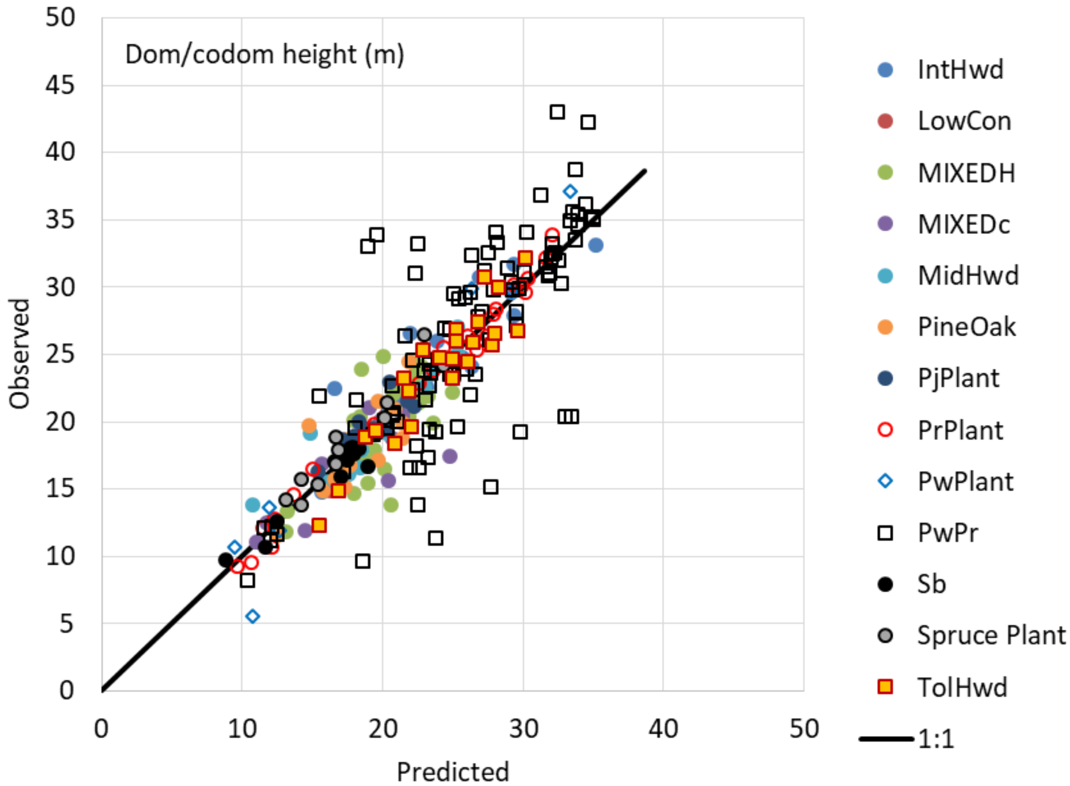

Figure 5. The observed and predicted Dom/codom height are plotted.

4.1.2 TVOL_merch

Overall, the bias for TVOL_merch (gross total volume of merchantable trees) was small but it varied with strata. There was one sample (PRF208, PwPlant) where p99 < 5m and no merchantable attributes were predicted (Table 12, Figure 6 and Figure 7).

Table 12. The TVOL_merch results are given. SE = standard error of the mean. RMSE = Root Mean Squared Error. The bias and RMSE are given as a percent of the observed mean.

Forest type	N	Observed (m^3/ha)	Predicted (m^3/ha)	Bias + SE (m^3/ha)	% bias	RMSE (m^3/ha)	% RMSE
IntHwd	15	391	406	-15 ± 16	-4%	63	16%
LowCon	4	219	183	36 ± 29	16%	62	28%
MIXEDH	28	196	212	-16 ± 8	-8%	44	22%
MIXEDC	13	152	170	-18 ± 10	-12%	39	26%
MidHwd	13	166	237	-71 ± 12	-43%	81	49%
PineOak	14	203	212	-9 ± 15	-4%	54	26%
PjPlant	10	177	187	-9 ± 10	-5%	31	18%
PrPlant	23	426	332	94 ± 24	22%	145	34%
PwPlant	6	222	231	-10 ± 16	-4%	38	17%
PwPr	93	360	350	10 ± 11	3%	102	28%
Sb	14	149	129	20 ± 7	13%	32	22%
Spruce Plant	12	230	196	34 ± 18	15%	67	29%
TolHwd	23	266	335	-69 ± 16	-26%	101	38%
All	268	286	285	1 ± 5	0%	87	31%

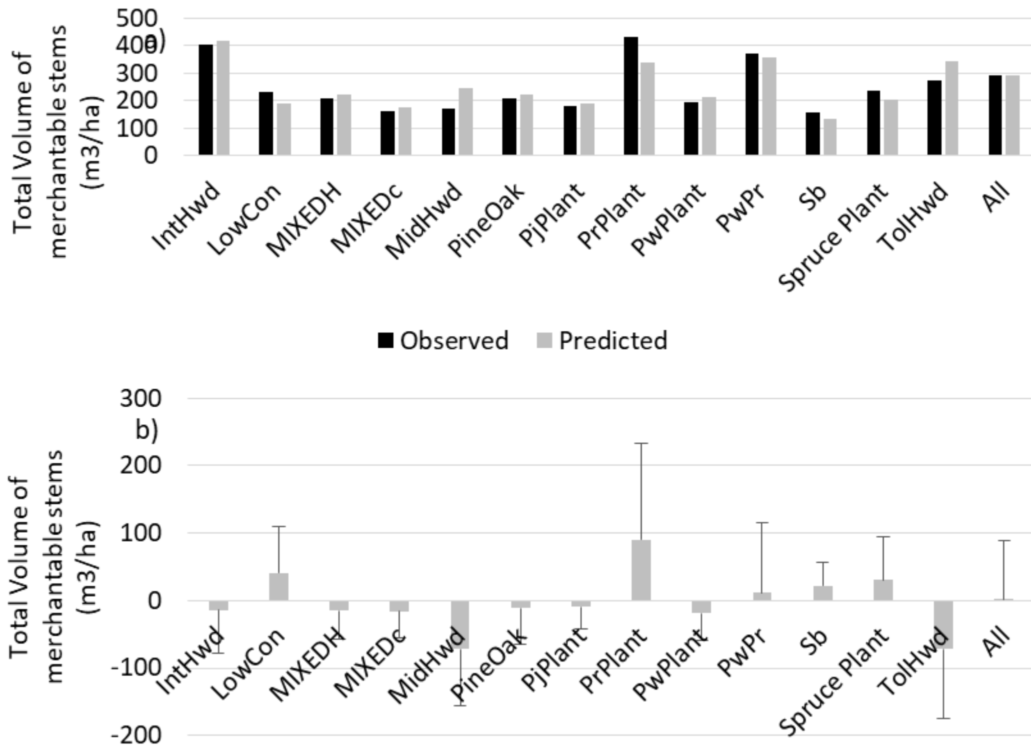

Figure 6. The observed and predicted TVOL_merch (a) and the bias and standard error bars (b) are given by strata

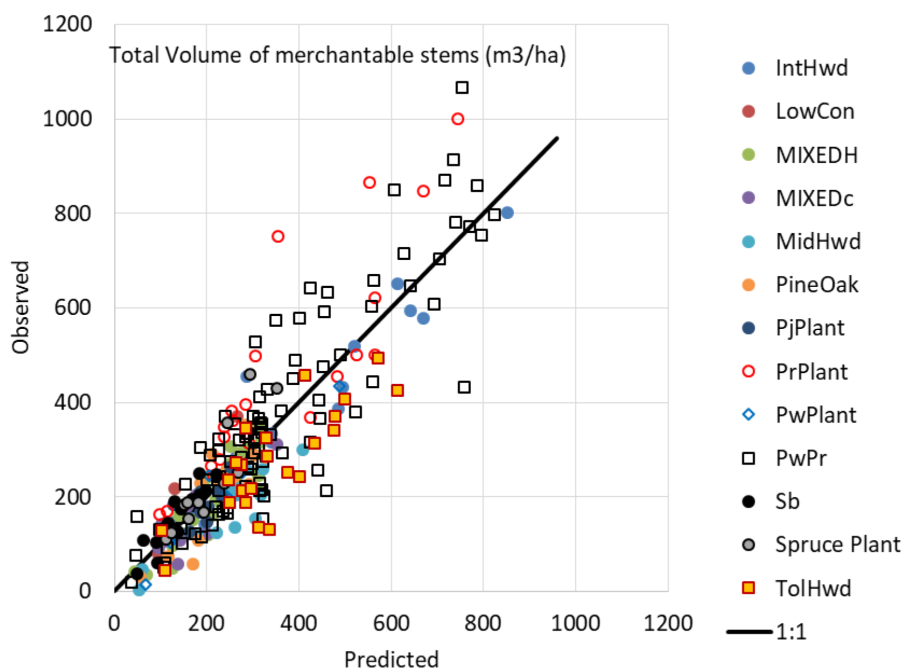

Figure 7. The observed and predicted TVOL_merch are plotted.

4.1.3 BA_merch

The basal area of merchantable trees (BA_merch) was overestimated by approximately 1% (Table 13, Figure 8 and Figure 9).

Table 13. The BA_merch results are given. SE = standard error of the mean. RMSE = Root Mean Squared Error. The bias and RMSE are given as a percent of the observed mean.

Strata	N	Observed (m^2/ha)	Predicted (m^2/ha)	Bias + SE (m^2/ha)	% bias	RMSE (m^2/ha)	% RMSE
IntHwd	15	33.9	35.1	-1.1 ± 1.5	-3%	5.8	17%
LowCon	4	29	22.5	6.5 ± 4.1	23%	9.7	33%
MIXEDH	28	22.4	24.1	-1.8 ± 0.9	-8%	5.1	23%
MIXEDC	13	20.1	20.9	-0.8 ± 1	-4%	3.5	17%
MidHwd	13	18.6	25.6	-7 ± 1	-38%	7.9	42%
PineOak	14	25	25.8	-0.9 ± 1.5	-3%	5.4	22%
PjPlant	10	19.6	21.7	-2.2 ± 1.2	-11%	4.2	21%
PrPlant	23	39.2	31.2	7.9 ± 1.4	20%	10.4	27%
PwPlant	6	25.2	22.1	1.3 ± 2.8	5%	6.4	25%
PwPr	93	31	29.9	1.1 ± 0.8	3%	8.2	26%
Sb	14	19.3	17	2.3 ± 1	12%	4.3	22%
Spruce Plant	12	28.3	25.2	3.2 ± 1.6	11%	6.1	21%
TolHwd	23	27.3	31.7	-4.5 ± 1.6	-16%	8.7	32%
All	268	27.9	27.5	0.3 ± 0.5	1%	7.4	26%

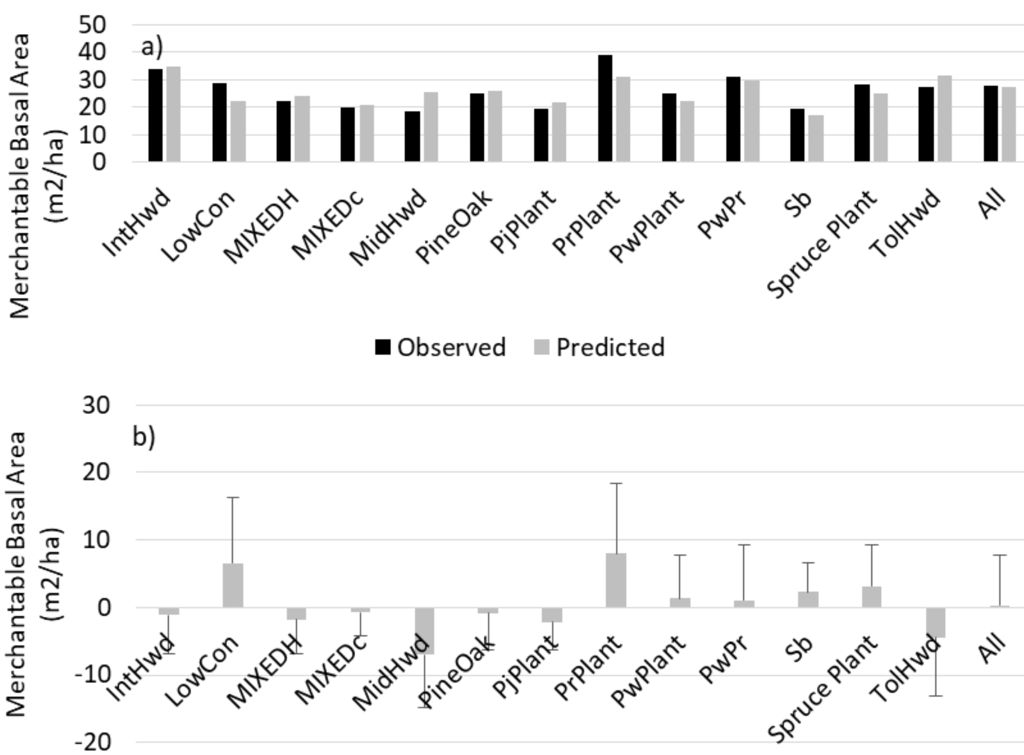

Figure 8. The observed and predicted BA_merch (a) and the bias and standard error bars (b) are given by strata

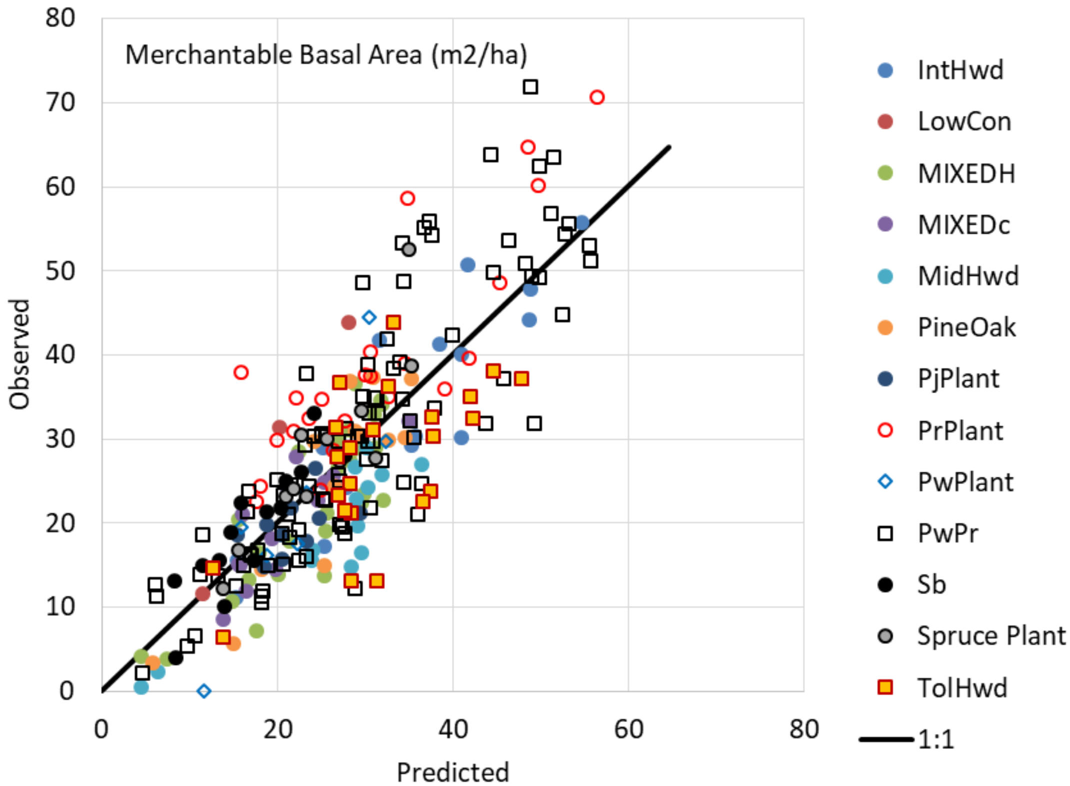

Figure 9. The observed and predicted BA_merch are plotted.

4.2 Summary of Validation results

The validation polygons were selected based on 2000 imagery and photo interpretation. The validation stand boundaries were updated for roads using the 2018 imagery and associated digital terrain model. Stand boundaries were modified to exclude unharvested areas in managed stands and to exclude swamps. An inner 12.5 m buffer was applied to the boundaries and plots with centres within the buffer were dropped from the stand summaries. Zonal summary statistics were calculated for raster cells within the stand modified polygon. A 20 m buffer was also investigated but excluded too many field plots in small stands

Results are given here for Top_ht, TVOL_merch, mvol, BA_merch and DQ_Merch. The full set of predictions for the validation stands is available in a separate file. The *observed* attribute for each polygon is the arithmetic average of the attribute from the field samples, and therefore the validation data are stand-level estimates of the attributes of interest. The *predicted* attribute for each polygon is the arithmetic average of the predicted values for each pixel. The predictions for the validation data differ from the calibration data in a number of respects. The validation data are an independent dataset that were not used for model development or calibration. Top heights were measured differently on the validation field plots.

4.2.1 Top_ht

For the validation stands, top height was generally underpredicted (Table 14, Figure 10 and Figure 11). As noted in section 2.3, only the largest tree on each validation plot was measured for height (while the 6 largest trees were measured on the calibration plots). The large biases for top height are due largely to differences in field protocols between the calibration and validation data. Had the field protocols been the same for calibration and validation, it is expected the top height bias would have been negligible.

Table 14. The Top height results are given. SE = standard error of the mean. RMSE = Root Mean Squared Error. The bias and RMSE are given as a percent of the observed mean.

Strata	N	Observed (m)	Predicted (m)	Bias + SE (m)	% bias	RMSE (m)	% RMSE
BlackSpruce	2	25.6	22	3.6 ± 0.1	14%	3.6	14%
JackPine	2	22.2	22.7	-0.4 ± 0.3	-2%	0.5	2%
LowlandConifer	1	19.1	18.2	0.9 ± NA	4%	NA	NA
Mixedwood	4	25	22.3	2.7 ± 0.4	11%	2.8	11%
Oak	2	26.3	24.7	1.7 ± 1.2	6%	2.1	8%
Poplar	2	27.8	25.2	2.7 ± 2.4	10%	3.6	13%
Pr Plantation	4	25.3	24.6	0.7 ± 1	3%	1.9	8%
PwManaged	3	32	25.2	6.8 ± 1.6	21%	7.2	22%
PwNatural	4	28.7	25.4	3.3 ± 0.1	12%	3.3	12%
TolerantHwd	3	26.9	23.3	3.6 ± 0.6	14%	3.7	14%
All	27	26.5	23.8	2.7 ± 0.5	10%	3.6	14%

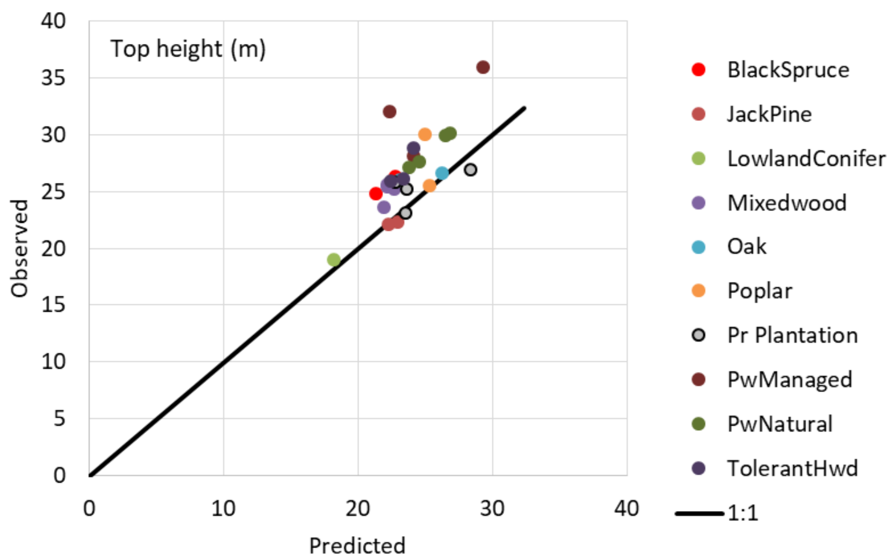

Figure 10. The observed and predicted Top height are plotted.

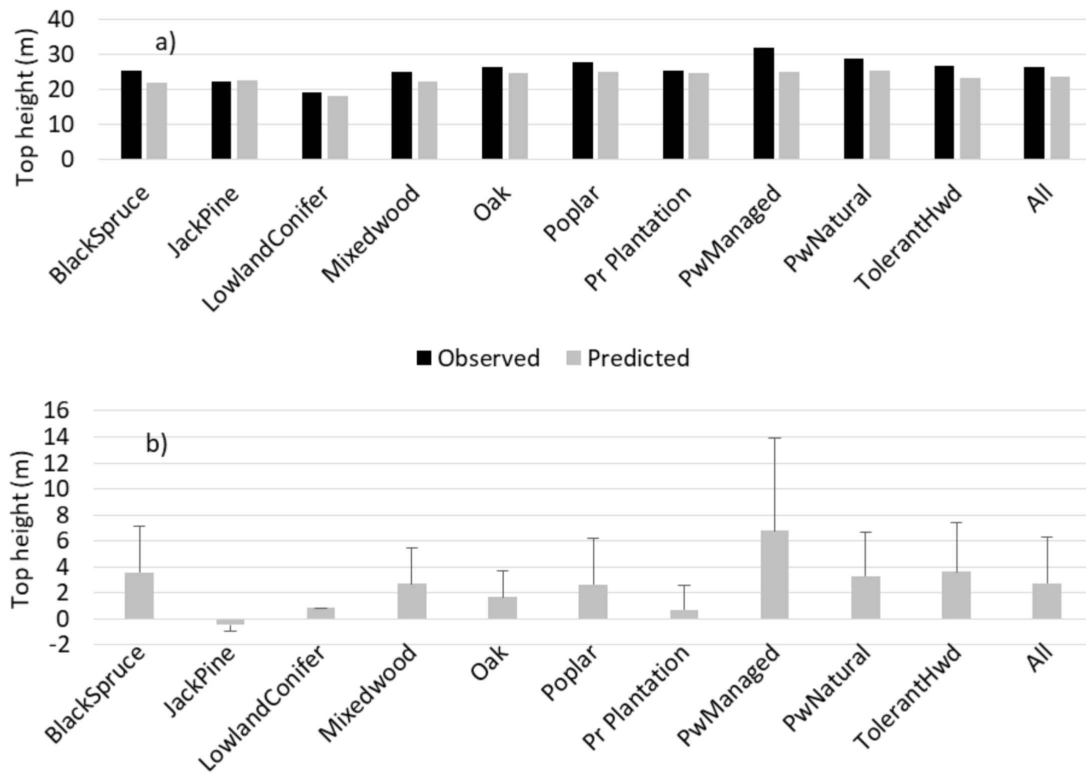

Figure 11. The observed and predicted Top height (a) and the bias and standard error bars (b) are given by strata.

4.2.2 TVOL_merch

TVOL_merch (gross total volume of merchantable stems) predictions were relatively unbiased with the largest underestimation for the Pr Plantation strata and largest overestimation for Pw managed (Table 15, Figure 12 and Figure 13).

PrPlant had a large bias (underestimation) due in part to having the highest volumes. RF uses approximately 63% of the calibration data (the default used here) to generate each classification and regression tree and does not extrapolate. As a result, RF has a tendency to underpredict high values and overpredict low values. The calibration plots cover a much larger range of volumes than the validation polygons so this is not expected to have a large effect at the polygon level.

The PwManaged plots are in tall stands that had some of the overstory removed to encourage growth of the residual stems and establish regeneration. They have a relatively low volume/height ratio compared to the rest of the strata, leading to an overestimation of volume. Poplar had the next highest bias (overestimation). This may be less of a prediction issue and more of a data compilation issue. The ground estimates of volumes use individual tree taper models. Taper models generally work well for trees with single straight stems. Hardwood trees often have heavy branching and it is difficult to define and measure the main stem volume, and also difficult to predict the main stem volume. In addition, for the poplar validation stands, the crew noted mortality in the poplar overstory.

Table 15. Validation results for TVOL_merch (gross total volume of merchantable stems). Bias and RMSE are given as a percent of the observed mean.

Strata	N	Observed (m^3/ha)	Predicted (m^3/ha)	Bias + SE (m^3/ha)	% bias	RMSE (m^3/ha)	% RMSE

Strata	N	Observed (m^3/ha)	Predicted (m^3/ha)	Bias + SE (m^3/ha)	% bias	RMSE (m^3/ha)	% RMSE
BlackSpruce	2	185	199	-14 ± 10	-8%	17.2	9%
JackPine	2	206	217	-11 ± 3	-5%	11.5	6%
LowlandConifer	1	150	125	25 ± NA	16%	NA	NA
Mixedwood	4	221	227	-6 ± 6	-3%	11.2	5%
Oak	2	279	287	-8 ± 1	-3%	8.2	3%
Poplar	2	255	288	-34 ± 1	-13%	33.8	13%
Pr Plantation	4	330	271	59 ± 21	18%	68.8	21%
PwManaged	3	174	215	-41 ± 19	-24%	49.6	28%
PwNatural	4	308	283	25 ± 4	8%	26.1	8%
TolerantHwd	3	233	250	-17 ± 11	-7%	23.2	10%
All	27	246	245	1 ± 7	0%	36	15%

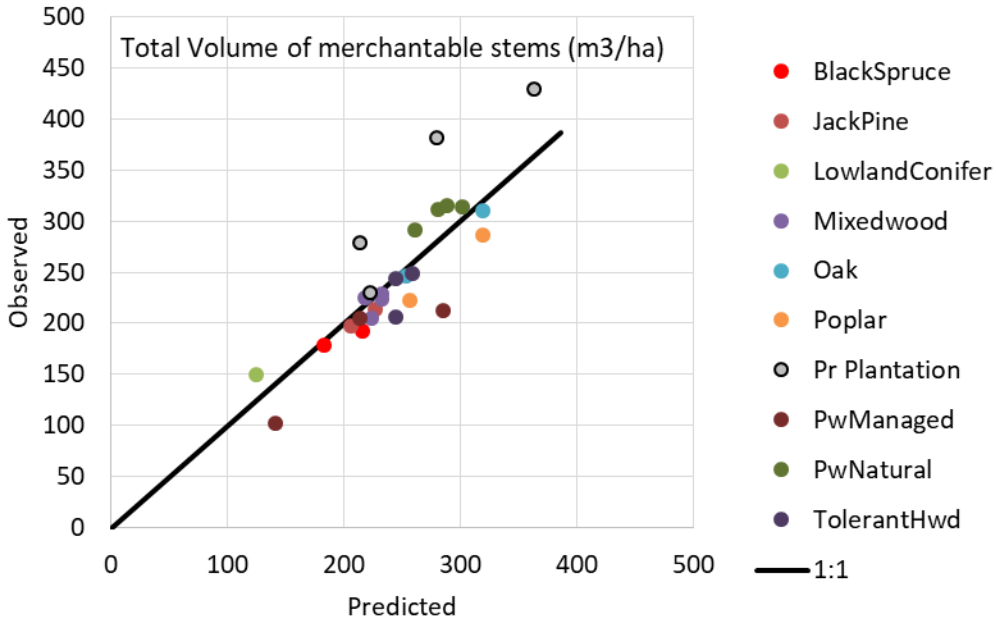

Figure 12. The observed and predicted total volume of merchantable stems are plotted.

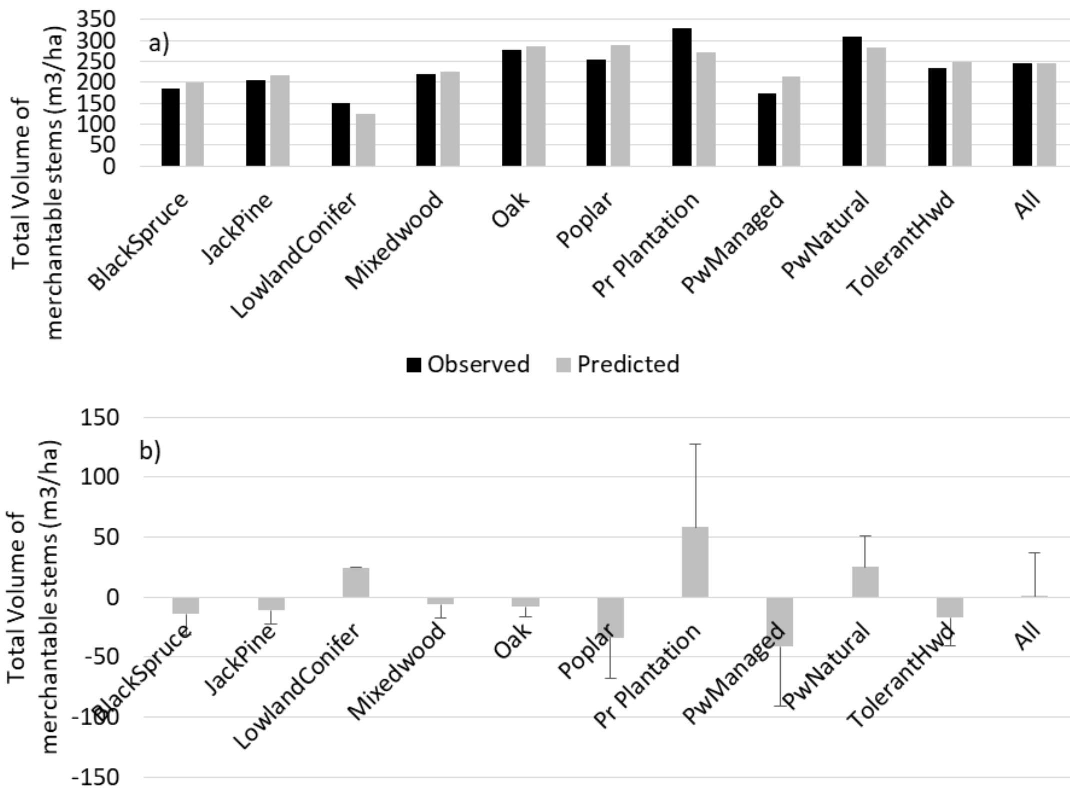

Figure 13. The observed and predicted total volume of merchantable stems (a) and the bias and standard error bars (b) are given by strata.

4.2.3 Mvol

The results for merchantable stem volume (Table 16, Figure 14 and Figure 15) are similar to those for TVOL_merch.

Table 16. Validation results for merchantable stem volume. SE = standard error of the mean. The bias and RMSE are given as a percent of the observed mean.

Strata	N	Observed (m^3/ha)	Predicted (m^3/ha)	Bias + SE (m^3/ha)	% bias	RMSE (m^3/ha)	% RMSE
BlackSpruce	2	152	167	-15 ± 11	-10%	18.8	12%
JackPine	2	183	187	-4 ± 6	-2%	7.5	4%
LowlandConifer	1	125	97	27 ± NA	22%	NA	NA
Mixedwood	4	168	175	-7 ± 6	-4%	13	8%
Oak	2	229	227	2 ± 5	1%	5.5	2%
Poplar	2	210	245	-36 ± 10	-17%	37.1	18%
Pr Plantation	4	295	239	56 ± 15	19%	61.3	21%
PwManaged	3	161	195	-34 ± 19	-21%	43.6	27%
PwNatural	4	266	241	26 ± 5	10%	27.3	10%
TolerantHwd	3	178	194	-17 ± 14	-9%	25.3	14%
All	27	208	205	2 ± 7	1%	33.7	16%

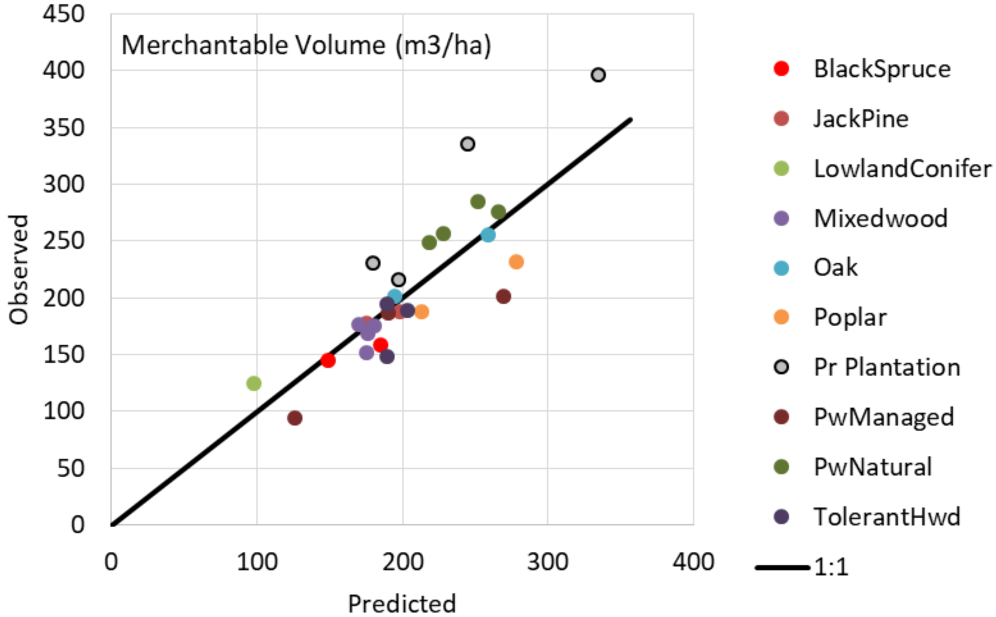

Figure 14. The observed and predicted merchantable volume are plotted.

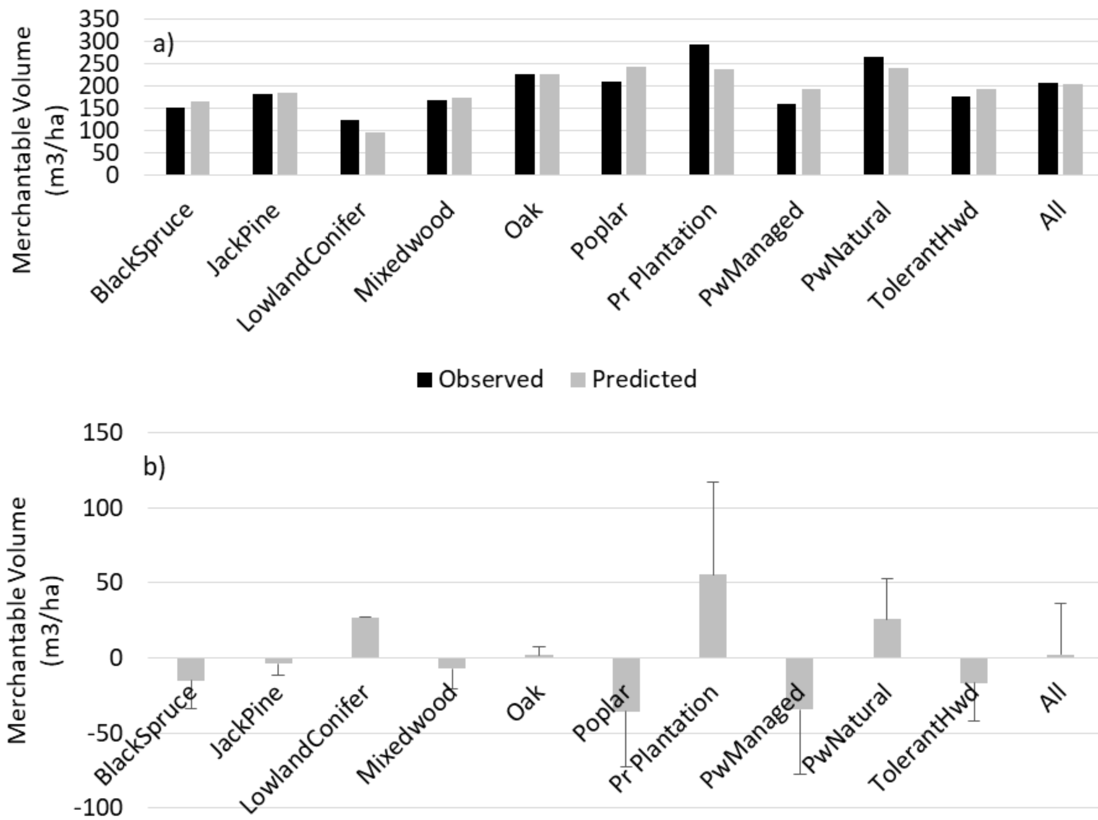

Figure 15. The observed and predicted merchantable volume (a) and the bias and standard error bars (b) are given by strata.

4.2.4 BA_merch

Merchantable basal area results are presented in Table 17, Figure 16 and Figure 17. As was seen in the validation of volumes, there is a trend to overestimate BA_merch for managed pine and poplar stands and underestimate mature red pine plantations. Possible explanations for this were given in section 4.2.2.

Table 17. Validation results for basal area for merchantable stems. SE = standard error of the mean. The bias and RMSE are given as a percent of the observed mean.

Strata	N	Observed (m^2/ha)	Predicted (m^2/ha)	Bias + SE (m^2/ha)	% bias	RMSE (m^2/ha)	% RMSE
BlackSpruce	2	22.4	23.5	-1.1 ± 0.8	-5%	1.4	6%
JackPine	2	22.3	22.8	-0.5 ± 0.5	-2%	0.7	3%
LowlandConifer	1	23.4	16.5	6.9 ± NA	29%	NA	NA
Mixedwood	4	26.8	27.3	-0.5 ± 0.8	-2%	1.5	5%
Oak	2	29.6	30.5	-0.9 ± 1	-3%	1.3	4%
Poplar	2	24.7	29.1	-4.4 ± 0.6	-18%	4.4	18%
Pr Plantation	4	31.6	25.3	6.3 ± 1.9	20%	7.1	22%
PwManaged	3	15	18.4	-3.4 ± 1.2	-23%	3.8	25%
PwNatural	4	31.5	29.5	2 ± 1.1	6%	2.8	9%
TolerantHwd	3	27.5	28.2	-0.7 ± 1.1	-3%	1.7	6%
All	27	26.2	25.8	0.5 ± 0.7	2%	3.8	14%

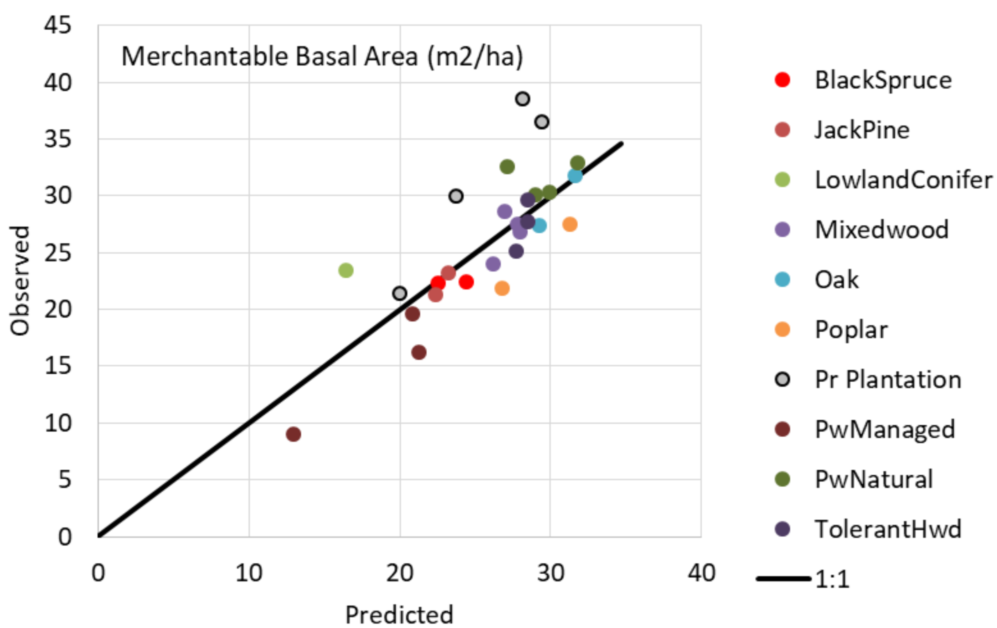

Figure 16. The observed and predicted BA_merch are plotted.

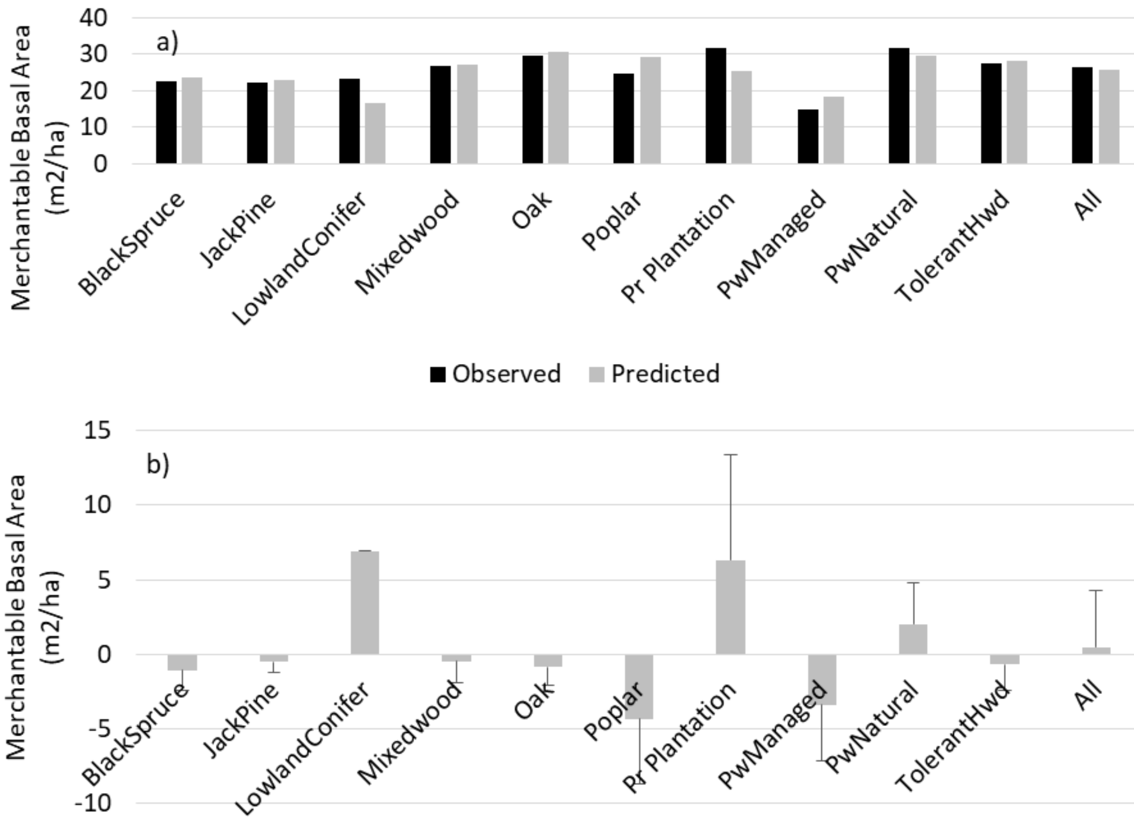

Figure 17. The observed and predicted BA_merch (a) and the bias and standard error bars (b) are given by strata.

4.2.5 DQ_merch

The predicted DQ_merch for the validation stands was calculated as the arithmetic average of the predicted DQ_merch of the pixels within the polygon. The predictions are relatively unbiased (Table 18, Figure 18 and Figure 19).

Table 18. Validation results for DQ_merch. SE = standard error of the mean. The bias and RMSE are given as a percent of the observed mean.

Strata	N	Observed (cm)	Predicted (cm)	Bias + SE (cm)	% bias	RMSE (cm)	% RMSE
BlackSpruce	2	19.7	20.5	-0.8 ± 0.3	-4%	0.9	4%
JackPine	2	19.4	20.8	-1.4 ± 1.5	-7%	2	10%
LowlandConifer	1	19.2	18.3	0.9 ± NA	5%	NA	NA
Mixedwood	4	20.3	20.8	-0.5 ± 0.3	-3%	0.7	4%
Oak	2	24.7	23.2	1.5 ± 1.1	6%	1.9	7%
Poplar	2	20.5	24.1	-3.6 ± 2.4	-17%	4.3	21%
Pr Plantation	4	26.3	25.5	0.7 ± 2.4	3%	4.3	16%
PwManaged	3	35.7	31	4.7 ± 1.8	13%	5.4	15%
PwNatural	4	24.2	23.9	0.4 ± 0.8	1%	1.4	6%
TolerantHwd	3	21	21.8	-0.8 ± 0.6	-4%	1.1	5%
All	27	23.7	23.5	0.2 ± 0.6	1%	2.9	12%

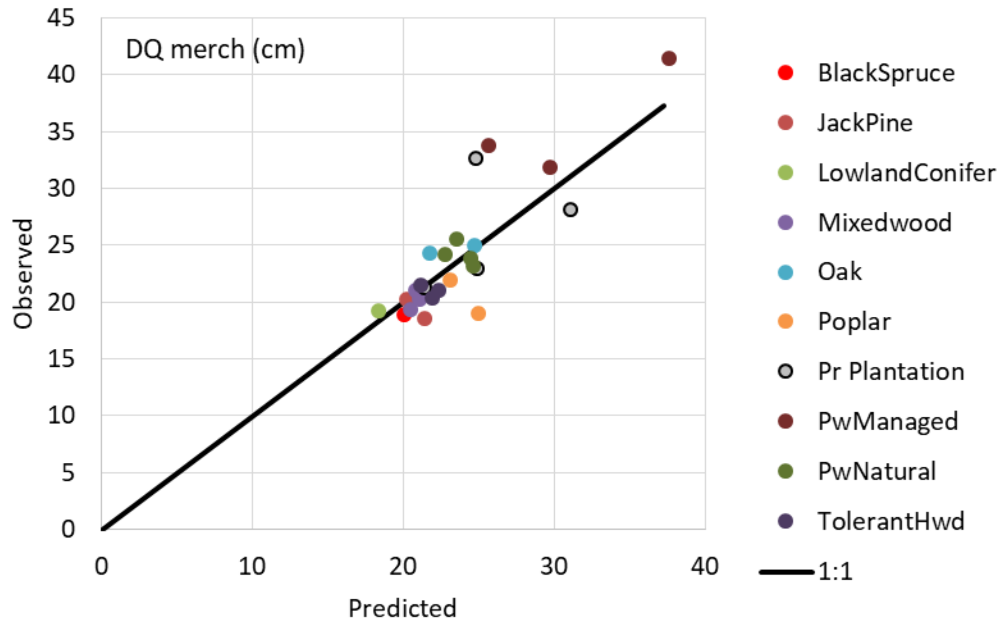

Figure 18. The observed and predicted DQ_merch are plotted.

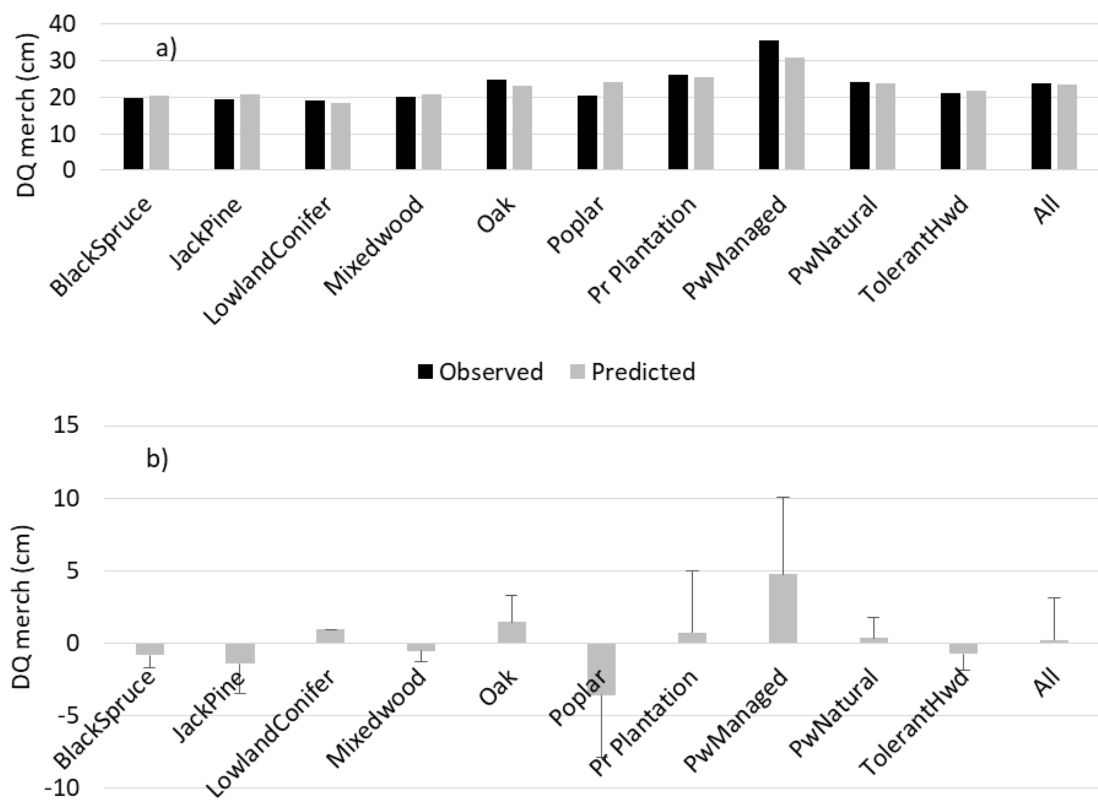

Figure 19. The observed and predicted DQ_merch (a) and the bias and standard error bars (b) are given by strata.

5. Literature Cited

Lambert, M.-C., Ung, C.-H and Raulier, F. 2005. Canadian national tree aboveground biomass equations. Can. J. For. Res. 35:1996-2018.

Liaw, A. and M. Wiener (2002). Classification and Regression by randomForest. R News 2(3), 18--22.

Sharma, M. 2016. Comparing height-diameter relationships of boreal tree species grown in plantations and natural stands. For. Sci. 62:70-77.

Sharma, M. and J. Parton. 2009. Modeling stand density effects on taper for jack pine and black spruce plantations using dimensional analysis. For. Sci. 55:268-282.

White, J.C., M.A. Wulder, A. Varhola, M. Vastaranta, N.C. Coops, B.C. Cook, D. Pitt and M. Woods. 2013. A best practices guide for generating forest inventory attributes from airborne laser scanning data using and area-based approach. Canadian Forest Service, Canadian Wood Fibre Centre Information Report FI-X-010 33p + app.

White, J.C., P. Tompalski, M. Vastaranta, M.A. Wulder, N. Saarinen, C. Stepper and N.C. Coops. 2017. A model development and application guide for generating an enhanced forest inventory using airborne laser scanning data and an area-based approach. Canadian Forest Service, Canadian Wood Fibre Centre Information Report FI-X-018 38p.

Yang, Y., Monserud, R.A. and Huang, S. 2004. An evaluation of diagnostic test and their role in validating forest biometric models. Can. J. For. Res. 34:619-629.

Zakrzewski, W.T and M. Penner. 2013. A comparison of tree stem taper models for use in Ontario. Ont. For. Res. Inst., Queen's Printer for Ontario. Forest Research Report No. 176. 26p.

6. Appendix A –SPL metrics - Thresholds & Returns

Three different subsets of the SPL point cloud were compared for metric calculation and area-based attribute modelling. Subsetting and subsequent analysis were undertaken at the plot level, using the 249 plots established in the summer of 2018 at PRF. The three subsets were as follows:

FPC_T0 – SPL metrics were computed using the full point cloud (FPC) and a threshold of 0 m (T0).

Veg_T0 – SPL metrics were computed using only returns classified as vegetation (i.e., returns 3, 4, 5; Veg) and a threshold of 0 m (T0).

Veg_T130 – SPL metrics were computed using only returns classified as vegetation (i.e., returns 3, 4, 5; Veg) and a threshold of 1.3 m (T130).

Plot level models were developed and applied to the validation stands at the grid-cell level. A 20 m buffer was applied to the validation stand boundary to remove grid cells at the stand edge, and thereby avoid mixing forest conditions from adjacent stands. Stand-level predictions were generated by taking the average of the grid-cell level predictions.

The estimates from the three different sets of predictors can be thought of as repeated measures of the validation stands. The effect of the predictor subsets was tested using the Wilk's lambda. Estimates generated using FPC_T0 predictors were compared to predictions generated using Veg_T0 and Veg_T130 for top height, gross volume, DQ, basal area, and Lorey's height, all for merchantable stems (Table 19).

Table 19. The results of the repeated measures ANOVA, the Wilk's Lambda, by attribute. Wilk's Lambda is a test of the H_0: no effect of calibration dataset (FPO_T0 vs. Veg_T0 vs. Veg_T130).

Source	Degrees of freedom Numerator	Denominator	Top height (m)	Tvol merch (m^3/ha)	Tvol merch (m^3)	DQ merch (cm)	BA merch (m^2/ha)	Lorey height (m)
Mean	2	17	0.1505	0.7098	0.4476	0.1291	<.0001	0.7697
Forest type	16	34	0.0123	0.1315	0.0589	0.6526	0.0192	0.0832

The effect of the different SPL predictor subsets was not statistically significant for any attribute except merchantable basal area. This result is difficult to explain as the results for BA_merch and TVOL_merch seem similar (Figure 20 and Figure 21) and there is no statistically significant effect of SPL subset on TVOL_merch. It may be that for BA_merch, the results are more consistent across forest type. For TVOL_merch, the differences between the SPL subsets are larger for Pr plantations and Pw managed and minor for the rest of the forest types. For BA_merch, the differences are large for Pr plantations and Pw managed but also for black spruce and jack pine.

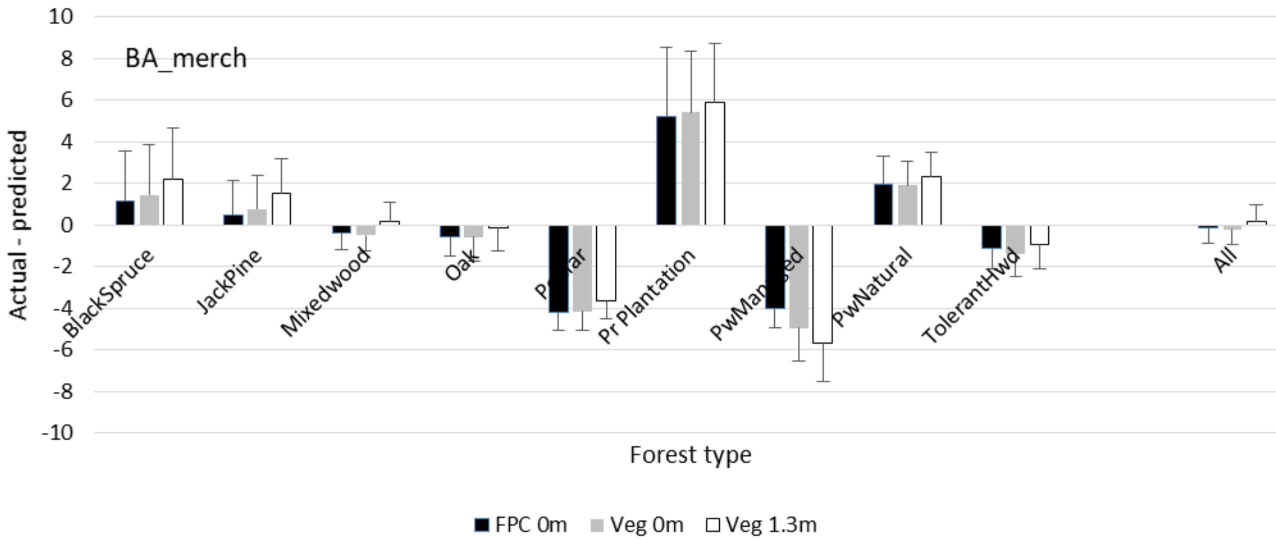

Figure 20. The average BA_merch bias is given by forest type, along with standard error bars, for the validation stands. There are 3 stands/forest type. These are the intended forest types.

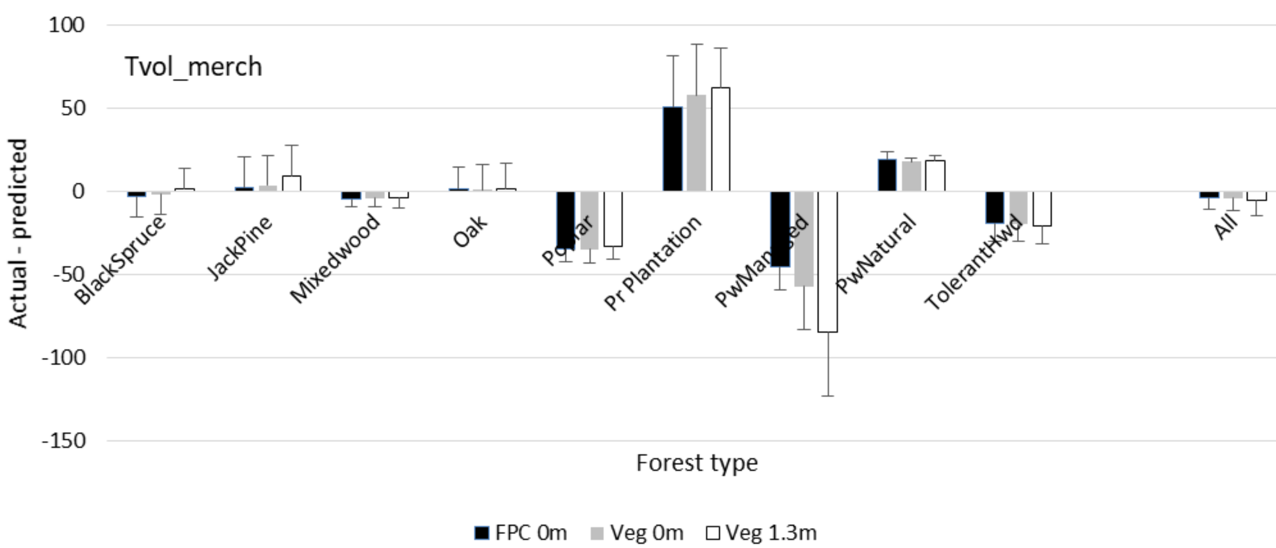

Figure 21. The average TVOL_merch bias is given by forest type, along with standard error bars, for the validation stands. There are 3 stands/forest type. These are the intended forest types.

For BA_merch, the differences between FPC_T0 and Veg_T0 are not statistically significant but the differences between Veg_T0 and Veg_T130 are statistically different.

The significant effect of the height threshold on BA_merch is surprising but, of all the attributes considered, merchantable basal area is where one might expect there to be an effect: all of the other attributes (height, volume and DQ) are highly correlated with height (which SPL measures accurately). In most forests, BA_merch will also be highly correlated with height. However, forests like the PRF, with shelterwood silviculture, thinnings and spacing, will have some stands with tall trees and low BA (e.g. Pw Managed). For those types of stands, no threshold results in improved prediction accuracy.

Predictions were also evaluated by regressing the observed value (from the validation data described in Section2.3) on the predicted value. Good predictions will have a regression with an intercept of 0 and a slope of 1. The null hypothesis H0: intercept b_0 = 0 and slope b_1 = 1 was evaluated using the simultaneous F-test (Yang et al. 2004).

$$F_{2,n-2} = \frac{n(b_0 - 0)^2 + 2 \cdot \sum \hat{y}_i \cdot (b_0 - 0) \cdot (b_1 - 1) + \sum \hat{y}_i \cdot (b_1 - 1)^2}{2 \cdot \sum (y_i - \hat{y}_i) \Big/ (n-2)}$$

None of the predictions showed any departures from the assumption of an intercept of 0 and slope of 1 (Table 20) and the null hypothesis could not be rejected.

Table 20. The results of the simultaneous F-test which tests the H_0: b_0 = 0 and b_1 = 1 for the regression of observed on predicted attribute. If the F-test was statistically significant, the regression was examined to see whether it was as a result of the intercept or slope (or both). FPC_T0 is the same as PRF in Table 28.

Attribute		FPC_T0	Veg_T0	Veg_T130
Top height	$P(F > F_{obs})$	<.0001	<.0001	<.0001
	Significant term	neither	neither	neither
	Intercept	1.93548	2.01385	1.57184
	Slope	1.02864	1.0226	1.03689
TVOL_merch (m^3/ha)	$P(F > F_{obs})$	<.0001	<.0001	<.0001
	Significant term	neither	neither	neither
	Intercept	-51.5864	-31.5818	8.81039
	Slope	1.19265	1.11043	0.94321
TVOL_merch (m^3)	$P(F > F_{obs})$	<.0001	<.0001	<.0001
	Significant term	neither	neither	neither
	Intercept	-36.5018	-27.051	-5.34601
	Slope	0.98527	0.97924	0.96567
DQ merch (cm)	$P(F > F_{obs})$	<.0001	<.0001	<.0001
	Significant term	neither	neither	neither
	Intercept	-3.48915	-4.0561	-4.62836
	Slope	1.15731	1.18859	1.2109
BA merch (m^2/ha)	$P(F > F_{obs})$	<.0001	<.0001	<.0001
	Significant term	neither	neither	neither
	Intercept	-3.9007	-3.16387	-2.52663
	Slope	1.14145	1.11104	1.10402
Lorey height (m)	$P(F > F_{obs})$	<.0001	<.0001	<.0001
	Significant term	neither	neither	neither
	Intercept	-3.75039	-3.72989	-3.26602
	Slope	1.18318	1.18149	1.15951

7. Appendix B - Feature Selection

7.1 Feature Selection alternatives

There are many SPL predictors (Table 1) and some are highly correlated. This led to an investigation in to whether reducing the number of predictors would affect the predictions of forest attributes. A number of different algorithms were used to reduce the number of predictors. The predictions using the various subsets of predictors were compared to predictions using all predictors. All tests were conducted using the "train" function in the "caret" package to fit random Forests (RFs) with the following control options. Note that this analysis used the 249 PRF plots and the Veg_T130 SPL summaries.

RScript 1. Setting the dependent variable (in this case to CD_ht), removing records with missing values and setting Caret training options. "Combined" is the master dataset with the ground and SPL summaries and one record for each plot.

```
Yname <- "CD_ht"
# include this step if predicting merchantable attributes
Combined_merch <- Combined[Combined[,"a_p95"]>5, ]
Combined_noNA <- Combined_merch[!is.na(Combined_merch[ ,Yname]), ]
train_control <- trainControl(method="repeatedcv", number = 10, repeats = 5, savePredictions=TRUE)
set.seed(1212)
```

Mtry is the number of variables randomly selected as candidates for decision rules (splitting rules) at each split or node. The caret function compares 3 values of mtry and selects the one with the lowest RMSE. Three values of mtry are used – the minimum mtry = 2, the maximum mtry = all the potential predictors and the midpoint between the minimum and maximum.

7.1.1 All Predictor

Predictions were generated with all standard LAStool canopy metrics predictors. This is the base case.

RScript 2. The code using all the LAStool predictors.

```
set.seed(1212)
caret_model1 <- train(x = Combined_noNA[ ,LiDARPredictors], y=Combined_noNA[,Yname],
trControl=train_control, method="rf", ntree=1000, importance=TRUE)
```

7.1.2 findCorrelation

The "findCorrelation" function in the "caret" package was used to reduce the number of predictors, keeping predictors with low pair-wise correlations.

RScript 3. The code using "findCorrelation".

```
source("./findCorrelation_fix.R")
assignInNamespace("findCorrelation",findCorrelation, ns = "caret")
drop <- findCorrelation(correlationMatrix, cutoff = 0.9, verbose=FALSE, names=TRUE, exact=TRUE)
keep <- colnames(LiDARsubset[ ,!colnames(LiDARsubset) %in% drop])
set.seed(1212)
caret_model2 <- train(x = Combined_noNA [ ,keep], y=Combined_noNA [,Yname], trControl=train_control,
method="rf", ntree=1000, importance=TRUE)
```

7.1.3 Top30

The top 30 predictors were selected using the eleaps package. Note the time limit is set to 500 seconds. The default is 15 seconds. The results for 300 seconds and 500 seconds were identical

RScript 4. The code using the top 100 subset of predictors.

```
drop1 <- trim.matrix(correlationMatrix)
drop2 <- eleaps(drop1$trimmedmat,kmax=30, nsol=1, timelimit=500)
set.seed(1212)
keeptop30 <- colnames(LiDARsubset)[drop2$subsets[ ,,"Card.30"]]set.seed(1212)
caret_model2 <- train(x = Combined_noNA[ ,keeptop30], y=Combined_noNA[,Yname], trControl=train_control,
method="rf", ntree=1000, importance=TRUE)
```

7.1.4 Boruta

Boruta is an all relevant feature selection wrapper algorithm, capable of working with any classification method that outputs variable importance measure (VIM). Boruta uses Random Forest by default. Boruta performs a top-down search for relevant features by comparing original attributes' importance with importance achievable at random, estimated using their permuted copies, and progressively eliminating irrelevant features to stabilize that test.

RScript 5. The code using the subset selected using Boruta.

```
test <- Boruta(x = Combined_noNA[ ,LiDARPredictors], y=Combined_noNA[,Yname ])
xvars <- getSelectedAttributes(test)
set.seed(1212)
caret_model3 <- train(x = Combined_noNA[ ,xvars], y=Combined_noNA[,Yname], trControl=train_control,
method="rf", ntree=1000, importance=TRUE)
```

7.1.5 Boruta - Node size = 1

The Boruta subset of predictors from the previous section was used but the training used a nodesize = 1 (rather than the default of 5). This is not recommended but was investigated to see the effect on reducing the underprediction at high values (e.g. Pr plantations) and overprediction at low values (e.g. Pw managed) of the dependent variable.

RScript 6. The same as RScript 5 except using a node size of 1.

```
set.seed(1212)
caret_model4 <- train(x = Combined_noNA[ ,xvars], y=Combined_noNA[,Yname], trControl=train_control,
method="rf", ntree=1000, importance=TRUE, nodesize = 1)
```

This did not have the expected effect. The results for nodesize = 5 and nodesize = 1 were virtually identical. The nodesize is not the number of observations in a terminal node. According to the documentation, it is the minimum number of observations in a terminal node. So, with nodesize = 1, all the terminal nodes could still have 5 observations. However, searching on the net, https://stackoverflow.com/questions/28417826/nodesize-parameter-ignored-in-randomforest-package, it seems that the nodesize is the minimum number of observations that must exist in a node in order for a split to be attempted. So nodesize is the minimum node size for the next to terminal split. With a nodesize = 5, the terminal node could have a single observation. This option was not pursued further.

7.2 Results

The caret package uses three values of mtry and selects the one with the lowest root mean squared error (RMSE). The differences between different values of mtry were small (Figure 22). The differences between different methods of selecting predictor variables in terms of RMSE and bias were small (Table 21).

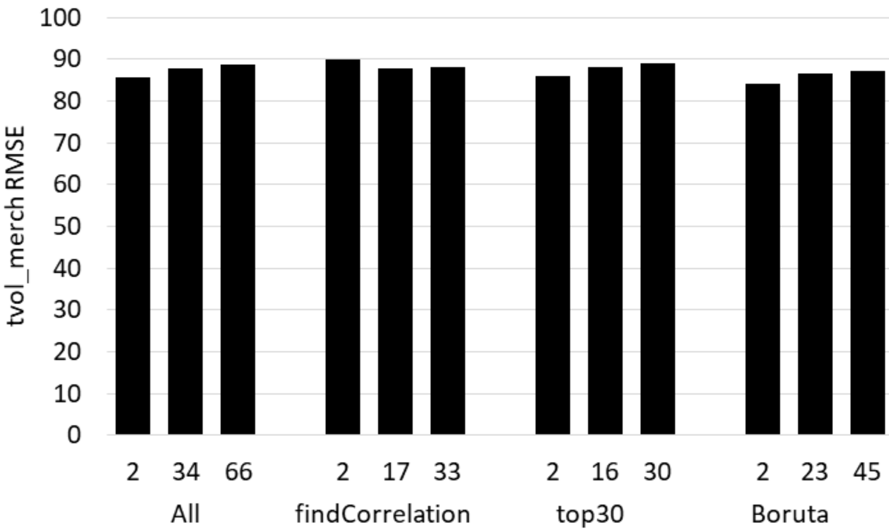

Figure 22. RMSE is given for each method of selecting predictor variables and for three values of mtry for each method. The results are for TVOL_merch.

Results are given for gross total volume for trees with Dbh > 9 cm (TVOL_merch), dominant/codominant height (CD_ht), basal area of trees with Dbh > 9 cm (BA_merch) and quadratic mean Dbh of trees with Dbh > 9 cm (Dbhq_merch). The results for the various fitting methods in terms of RMSE and mean absolute deviation (MAD) were very similar. "findCorrelation" generally had a slightly higher RMSE.

Table 21. Summary statistics are given for each dependent variable and fitting method. Mtry is the number of variables randomly selected as candidates for decision rules (splitting rules) at each split or node. The root mean squared error (RMSE), mean absolute deviation (MAD) and bias are given. The results are given for the mtry with the lowest RMSE.

Dependent variable	Method used to select predictors	Number of Predictors	mtry	Absolute RMSE	Absolute MAD	Absolute bias	Relative RMSE	Relative MAD	Relative bias
TVOL_merch (m^3/ha)	All predictors	66	2	85.54	0.78	-0.06	31%	0%	0%
	findCorrelation	33	17	87.79	0.76	0.93	32%	0%	0%
	Top 30	30	2	85.91	0.77	-1.30	31%	0%	0%
	Boruta	45	2	84.26	0.78	-0.28	31%	0%	0%
CD_ht (m)	All predictors	66	34	3.12	0.80	0.05	14%	4%	0%
	findCorrelation	33	17	3.28	0.78	0.07	15%	4%	0%
	Top 30	30	16	3.11	0.80	0.05	14%	4%	0%
	Boruta	35	18	3.12	0.80	0.04	14%	4%	0%
BA_merch (m^2/ha)	All predictors	66	2	7.50	0.63	-0.03	27%	2%	0%
	findCorrelation	33	17	7.70	0.61	-0.01	28%	2%	0%
	Top 30	30	2	7.46	0.64	-0.01	27%	2%	0%
	Boruta	48	2	7.46	0.64	-0.02	27%	2%	0%
Dbhq_merch (cm)	All predictors	66	66	4.66	0.73	0.10	19%	3%	0%
	findCorrelation	33	17	4.72	0.73	0.10	19%	3%	0%
	Top 30	30	30	4.74	0.72	0.14	19%	3%	1%
	Boruta	44	44	4.63	0.73	0.11	19%	3%	0%

Predictions were relatively insensitive to the method used to select predictors and to different mtry values, and no method or value of mtry was signficiantly better or worse than another. Nevertheless, it seems like "findCorrelation" is generally the poorest method to use (Figure 23 and Figure 24). Of the remaining parameter selection methods, the optimal mtry was 2 for the area-based attributes (e.g. BA and tvol) and higher for attributes that are tree-based (e.g. CD_ht and Dbhq).

The "predict" function uses all calibration data and gives much better results than the out of bag predictions. Rather than the "predict" function, the average of the predictions from the 5 repeats (RScript 1) was used.

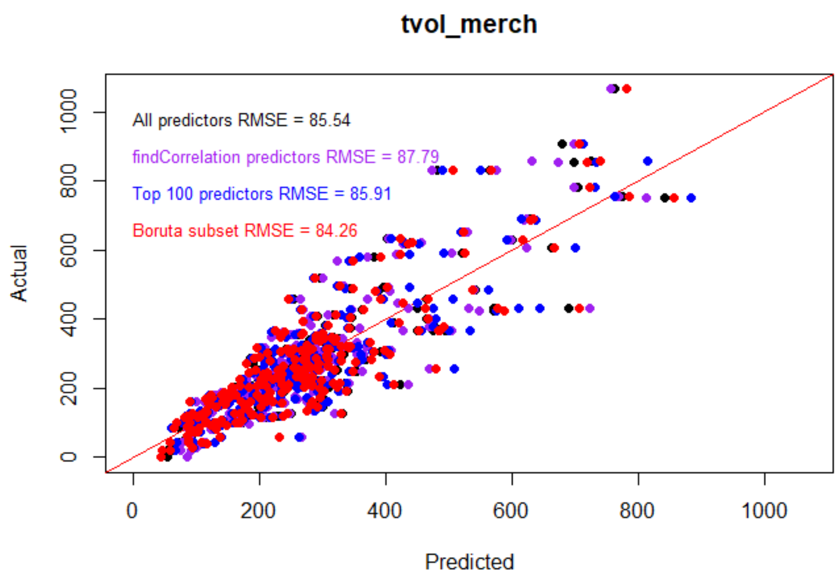

Figure 23. The predictions of total volume (Dbh > 9 cm) are compared.

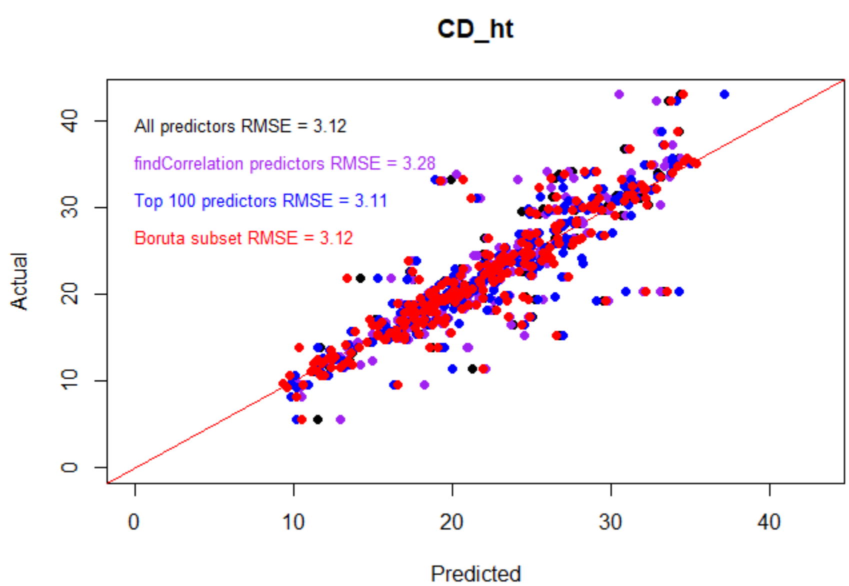

Figure 24. The predictions of CD_ht are compared.

Depending on the prediction algorithm, different subsets of predictors were used (Table 22). The "findcorrelation" algorithm tended to have fewer p-values and density percentages and more slice data, whereas the "top 30" had fewer slice and density predictors than the other algorithms.

Table 22. The predictors selected for each prediction option are given. The "findcorrelation" and top30 options do not depend on the attribute to be predicted. The predictors with green shading were common to all selected predictor subsets. There is more agreement between top30 and Boruta, partly because "findcorrelation" did not include many of the percentile (p) or density slice (dns) predictors.

			Boruta			
Predictor	findcorrelation	Top30	CD_ht	TVOL_merch	BA_merch	Dbhq_merch
a_std_95	Yes	Yes	Yes	Yes	Yes	Yes
a_ske_95	Yes	Yes			Yes	
a_kur_95	Yes	Yes				
a_avg			Yes	Yes	Yes	Yes
a_qav		Yes	Yes	Yes	Yes	Yes
a_p01	Yes	Yes			Yes	
a_p05	Yes	Yes		Yes	Yes	
a_p10	Yes	Yes		Yes	Yes	Yes
a_p20	Yes	Yes		Yes	Yes	Yes
a_p30		Yes	Yes	Yes	Yes	Yes
a_p40		Yes	Yes	Yes	Yes	Yes
a_p50		Yes	Yes	Yes	Yes	Yes
a_p60		Yes	Yes	Yes	Yes	Yes
a_p70		Yes	Yes	Yes	Yes	Yes
a_p80		Yes	Yes	Yes	Yes	Yes
a_p90		Yes	Yes	Yes	Yes	Yes
a_p95		Yes	Yes	Yes	Yes	Yes
a_p99	Yes	Yes	Yes	Yes	Yes	Yes
a_d0_2					Yes	Yes
a_d2_4	Yes				Yes	
a_d4_6	Yes			Yes	Yes	
a_d6_8	Yes		Yes			Yes
a_d8_10	Yes		Yes			Yes
a_d10_12	Yes		Yes			Yes
a_d12_14	Yes		Yes			Yes
a_d14_16	Yes		Yes			Yes
a_d16_18	Yes		Yes	Yes		Yes
a_d18_20	Yes		Yes			Yes
a_d20_22	Yes		Yes	Yes		Yes
a_d22_24	Yes		Yes	Yes		Yes
a_d24_26	Yes		Yes	Yes	Yes	Yes
a_d26_28	Yes		Yes	Yes	Yes	Yes
a_d28_30	Yes		Yes	Yes	Yes	Yes
a_d30_32	Yes		Yes	Yes	Yes	Yes
a_d32_34	Yes		Yes	Yes	Yes	
a_d34_36	Yes			Yes	Yes	
a_d36_38	Yes	Yes		Yes		
a_d38_40						
a_d40_42						
a_d42_44	Yes					
a_d44_46						
a_d46_48						
a_b10					Yes	Yes
a_b20				Yes	Yes	Yes
a_b30				Yes	Yes	Yes
a_b40		Yes	Yes	Yes	Yes	

			Boruta			
Predictor	findcorrelation	Top30	CD_ht	TVOL_merch	BA_merch	Dbhq_merch
a_b50		Yes		Yes	Yes	
a_b60	Yes	Yes		Yes	Yes	
a_b70		Yes		Yes	Yes	
a_b80				Yes	Yes	
a_b90	Yes				Yes	
a_dns_2m	Yes				Yes	Yes
a_dns_4m				Yes	Yes	Yes
a_dns_5m				Yes	Yes	Yes
a_dns_6m				Yes	Yes	Yes
a_dns_8m				Yes	Yes	Yes
a_dns_10m			Yes	Yes	Yes	Yes
a_dns_12m			Yes	Yes	Yes	Yes
a_dns_14m			Yes	Yes	Yes	Yes
a_dns_15m			Yes	Yes	Yes	Yes
a_dns_16m			Yes	Yes	Yes	Yes
a_dns_18m			Yes	Yes	Yes	Yes
a_dns_20m	Yes	Yes	Yes	Yes	Yes	Yes
a_dns_25m	Yes	Yes	Yes	Yes	Yes	Yes
a_vci_1mbin	Yes					
a_vc1_0.5bin						

Variable importance is a measure of the value of a predictor to the model and it is calculated using the prediction error (MSE) of the out-of-bag portion for each tree. Then the RMSE is calculated after permuting each predictor variable. The variable importance score is the average difference between the two RMSEs, normalized by the standard deviation of the differences.

For CD_ht, the most important predictors are p70, p80, and d14_16 (Table 23 and Figure 25).

Table 23. The predictors with the top 20 variable importance scores are given for CD_ht. The predictors shaded green were in the top 5 for at least three of the prediction methods. The predictors shaded orange were within the top 10 for at least three of the prediction methods.

Rank	All		findCorrelation		Top 30		Boruta	
1	a_p70	100	a_dns_20m	100	a_p70	100	a_p70	100
2	a_p80	95.36	a_d14_16	73.88	a_p80	90.02	a_p80	93.47
3	a_d14_16	78.36	a_d18_20	68.63	a_p90	71.31	a_d14_16	77.62
4	a_d22_24	64.91	a_d16_18	65.17	a_p60	71.01	a_d20_22	62.87
5	a_d20_22	64.53	a_dns_25m	58.03	a_p50	64.87	a_p90	56.72
6	a_d18_20	63.7	a_d20_22	53.73	a_p40	63.75	a_d12_14	53.48
7	a_p90	62.23	a_p99	48.51	a_std_95	58.88	a_d10_12	53.22
8	a_d12_14	58.04	a_d22_24	48.11	a_p95	57.45	a_d18_20	49.44
9	a_d16_18	56.84	a_d12_14	47.79	a_dns_20m	52.1	a_d16_18	49.01
10	a_dns_20m	56.83	a_std_95	44.33	a_b40	43.2	a_dns_20m	46.67
11	a_p60	56.7	a_d10_12	37.33	a_qav	42.15	a_d22_24	43.74
12	a_d10_12	50.97	a_d26_28	37.28	a_p30	38.65	a_p60	41.79
13	a_dns_15m	46.95	a_d32_34	32.71	a_p99	36.22	a_p95	38.51
14	a_p95	46.87	a_d24_26	32.25	a_b30	34.04	a_p50	31.52
15	a_p40	45.92	a_d28_30	31.45	a_dns_25m	33.87	a_p40	30.6
16	a_p50	45.77	a_d6_8	30.56	a_ske_95	24.03	a_b40	30.25
17	a_d26_28	43.09	a_d30_32	29.49	a_kur_95	23.79	a_dns_15m	29.79
18	a_qav	42.95	a_d8_10	26.76	a_p20	23.65	a_d30_32	28.21
19	a_dns_12m	42.49	a_p30	22.39	a_b60	21.66	a_d26_28	28.02
20	a_d30_32	40.67	a_vci_1mbin	20.68	a_b20	20.91	a_qav	26.46

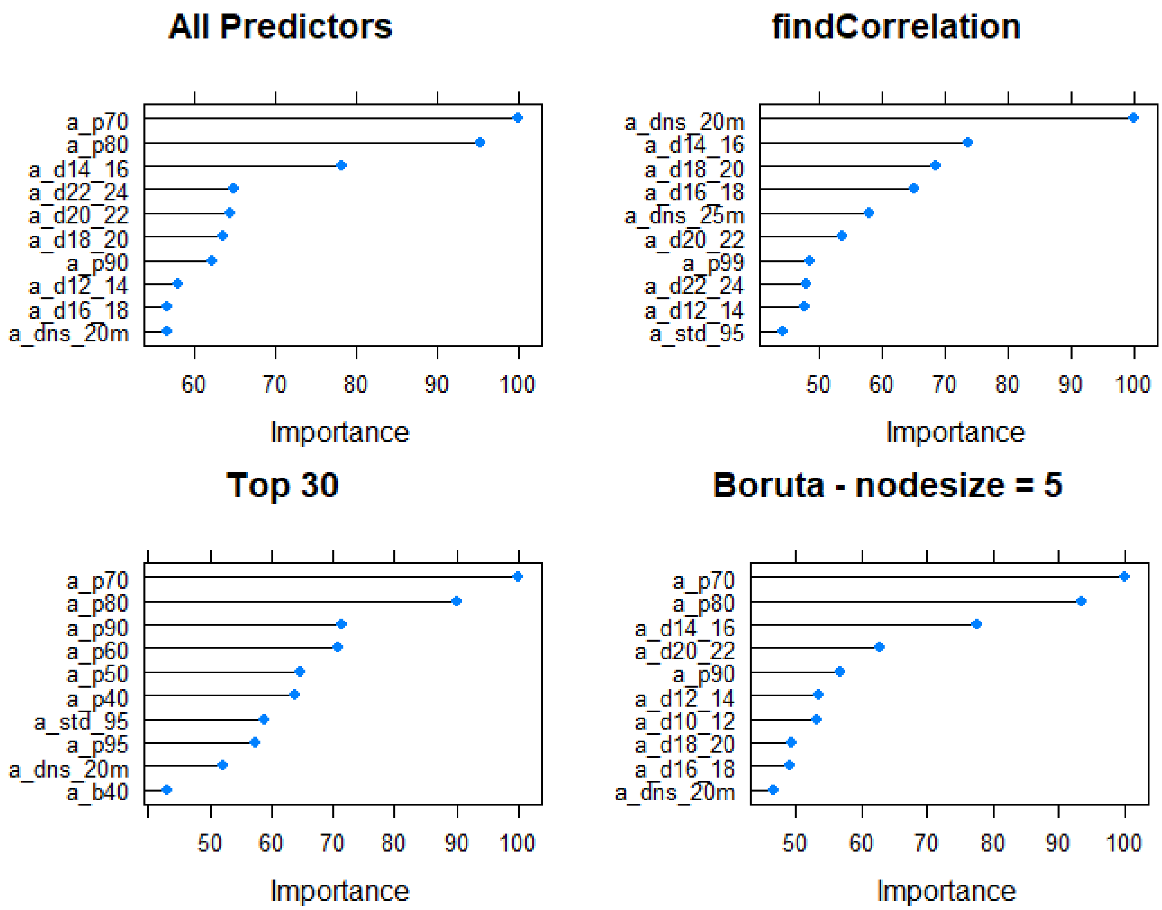

Figure 25. The variable importance graphs are given for CD_ht for the top 10 predictors.

For merchantable tvol, the most common top predictors were dns_25m and dns_20m (Table 24 and Figure 26).

Table 24. The predictors with the top 20 variable importance scores are given for TVOL_merch. The predictors shaded green were in the top 5 for at least three of the prediction methods. The predictors shaded orange were within the top 10 for at least three of the prediction methods.

Rank	All		findCorrelation		Top 30		Boruta	
1	a_p70	100	a_dns_25m	100	a_dns_25m	100	a_dns_15m	100
2	a_dns_20m	96.62	a_dns_20m	89.25	a_dns_20m	97.75	a_avg	96.36
3	a_p95	94.97	a_p20	64.21	a_p70	95.71	a_dns_10m	95.44
4	a_p90	93.62	a_d28_30	47.48	a_p90	92.16	a_dns_25m	91.36
5	a_p50	92.76	a_d34_36	46.19	a_p50	92	a_qav	87.1
6	a_dns_25m	92.43	a_d16_18	45.98	a_p80	91.8	a_dns_16m	83.67
7	a_dns_12m	91.34	a_d26_28	43.74	a_p60	91.79	a_p60	82.99
8	a_dns_18m	90.38	a_p30	42.7	a_qav	87	a_p50	82.59
9	a_d24_26	89.82	a_d24_26	42.33	a_p95	79.03	a_p90	80.37
10	a_p80	88.8	a_d4_6	40.9	a_p30	78.63	a_dns_18m	79.82
11	a_qav	86.95	a_d22_24	38.44	a_p40	74.69	a_dns_20m	77.51
12	a_dns_16m	85.08	a_d30_32	36.71	a_b70	73.52	a_d26_28	77.01
13	a_p60	84.97	a_p99	33.81	a_p99	73.33	a_p70	76.71
14	a_avg	84.3	a_d32_34	33.47	a_b40	71.72	a_dns_14m	68.02

Rank	All		findCorrelation		Top 30		Boruta	
15	a_dns_15m	81.58	a_d20_22	29.86	a_b30	70.16	a_d28_30	62.78
16	a_p99	80.91	a_d10_12	29.82	a_p20	69.06	a_p20	60.17
17	a_d30_32	79.63	a_std_95	29.33	a_b60	64.19	a_p95	60.14
18	a_d28_30	79.07	a_b90	28.72	a_std_95	63.38	a_p80	59.7
19	a_p40	75.01	a_dns_2m	28.28	a_b50	61.26	a_d30_32	58.69
20	a_d26_28	74.57	a_b60	27.81	a_b20	58.42	a_p30	56.4

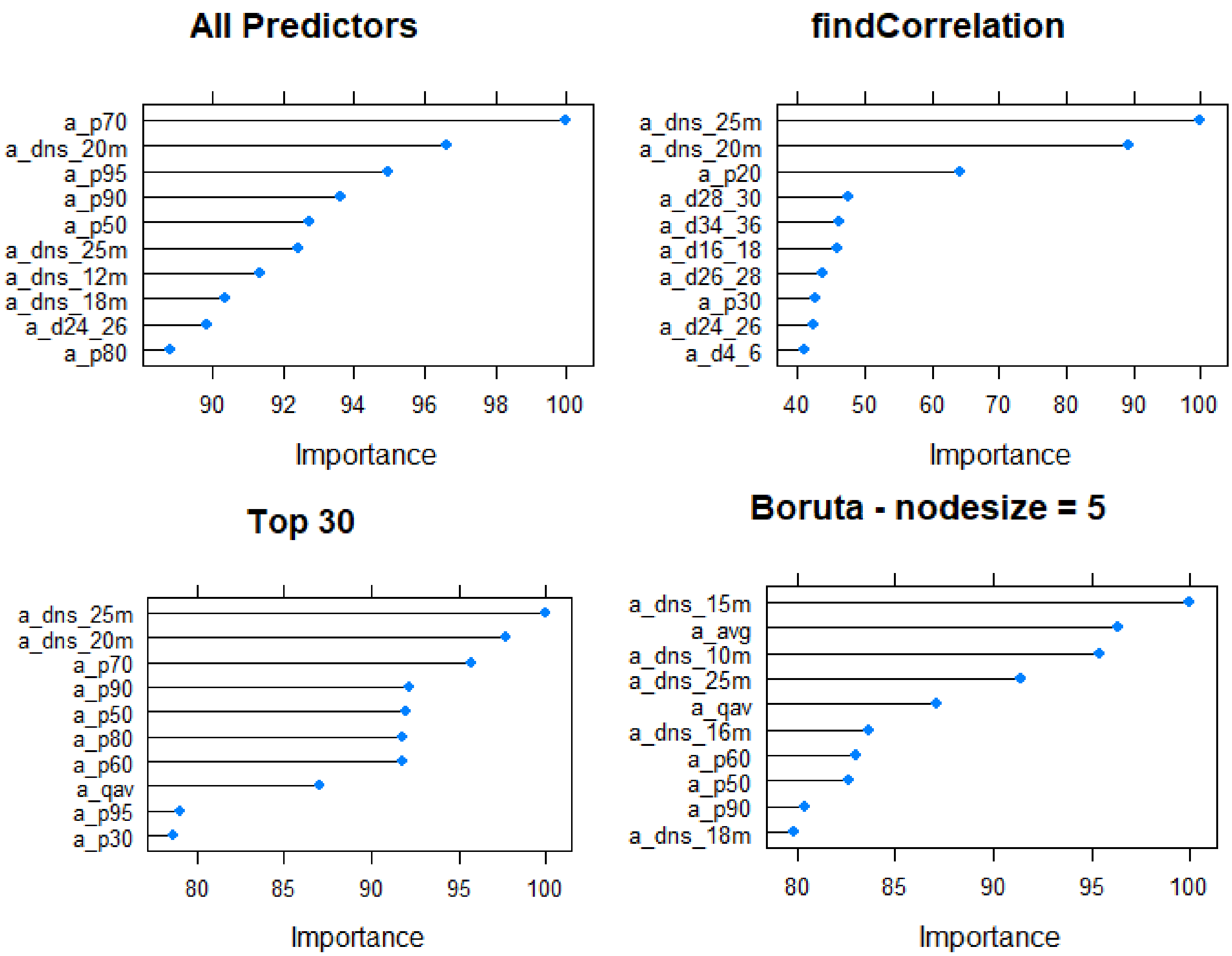

Figure 26. The variable importance diagrams are given for TVOL_merch for the top 10 predictors.

For merchantable BA, the top predictor was p20 followed by p30, p50, dns_20m and dns_25m (Table 25 and Figure 27).

Table 25. The predictors with the top 20 variable importance scores are given for merchantable BA. The predictors shaded green were in the top 5 for at least three of the prediction methods. The predictors shaded orange were within the top 10 for at least three of the prediction methods.

Rank	All		findCorrelation		Top 30		Boruta	
1	a_p40	100	a_p20	100	a_dns_25m	100	a_p30	100
2	a_dns_10m	97.48	a_dns_20m	84.26	a_p20	94.86	a_p50	99.08
3	a_dns_14m	94.35	a_dns_2m	79.98	a_b40	92.68	a_dns_12m	98.17
4	a_b30	92.25	a_dns_25m	69.66	a_p50	90.58	a_b50	93.84
5	a_p20	91.64	a_b60	64.54	a_p30	90.44	a_p20	92.26
6	a_p30	91.4	a_p30	52.45	a_dns_20m	89.38	a_dns_16m	87.4
7	a_dns_12m	91.28	a_d34_36	43.81	a_p60	86.52	a_p40	86.12
8	a_dns_20m	90.99	a_p05	39.3	a_b60	86.45	a_dns_14m	80.75
9	a_dns_25m	89.45	a_d30_32	38.71	a_p80	86.19	a_p10	77.63
10	a_p50	89.34	a_d24_26	38.69	a_b30	86.18	a_b30	77.47
11	a_b40	88.97	a_d28_30	37.69	a_p90	82.5	a_b40	77.1
12	a_p70	88.34	a_d26_28	36.52	a_p70	82.38	a_avg	76.32
13	a_b50	87.93	a_d16_18	36.26	a_b50	82.17	a_dns_15m	75.67
14	a_dns_6m	87.03	a_p99	34.36	a_p95	78.28	a_p90	72.58
15	a_dns_16m	85.99	a_d4_6	33.87	a_b70	77.41	a_b60	70.15
16	a_qav	85.48	a_ske_95	33.27	a_b20	77.17	a_dns_10m	68.61
17	a_p80	85.13	a_d2_4	31.36	a_qav	76.34	a_dns_25m	66.09
18	a_p60	84.24	a_p01	30.98	a_p40	75.36	a_p60	65.96
19	a_dns_15m	80.29	a_d10_12	30.42	a_p99	74.53	a_qav	65.07
20	a_d4_6	79.8	a_b90	29.3	a_p10	74.14	a_p70	60.79

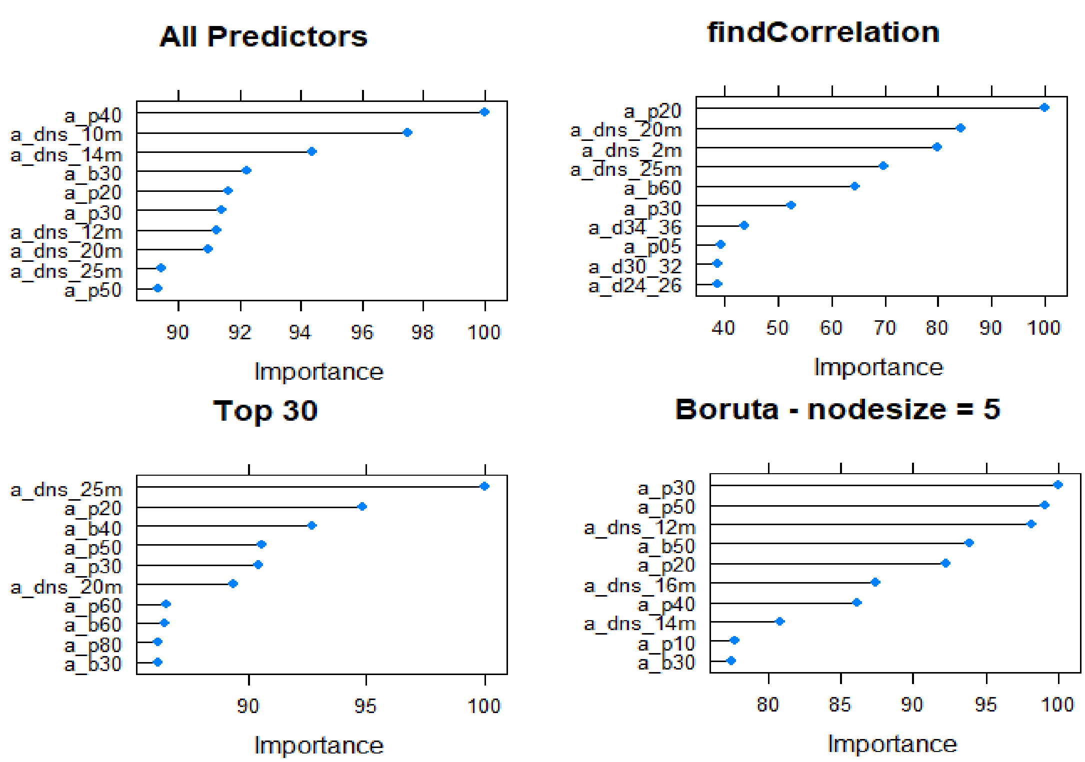

Figure 27. The variable importance diagrams are given for BA_merch for the top 10 predictors.

For merchantable Dbhq, the top predictors are p70, std_95, d10_12 and a_qav (Table 26 and Figure 28).

Table 26. The predictors with the top 20 variable importance scores are given for merchantable Dbhq. The predictors shaded green were in the top 5 for at least three of the prediction methods. The predictors shaded orange were within the top 10 for at least three of the prediction methods.

Rank	All		findCorrelation		Top 30		Boruta	
1	a_p70	100	a_std_95	100	a_p70	100	a_std_95	100
2	a_std_95	95.02	a_d10_12	92.3	a_std_95	79.96	a_p70	93.84
3	a_d10_12	68.87	a_dns_20m	80.36	a_qav	61.38	a_d10_12	68.93
4	a_qav	65.78	a_d14_16	76.97	a_p40	48.16	a_d12_14	53.91
5	a_d12_14	57.74	a_d16_18	71.77	a_b10	47.42	a_qav	53.21
6	a_d14_16	55.46	a_d12_14	70.68	a_p99	42.49	a_d14_16	46.48
7	a_p90	49.65	a_p99	70.42	a_p80	41.5	a_p90	46.03
8	a_dns_15m	48.94	a_dns_25m	68.98	a_b20	40.5	a_d16_18	40.75
9	a_p40	46.96	a_p30	57.81	a_p90	39.73	a_p80	37.94
10	a_p80	44.82	a_dns_2m	55.84	a_b30	38.72	a_p60	33.66
11	a_p99	42.31	a_d26_28	52.98	a_p60	38.21	a_dns_15m	32.94
12	a_p60	41.45	a_d8_10	45.89	a_p01	37.95	a_p99	32.22
13	a_dns_8m	40.24	a_d24_26	45.01	a_p50	34.39	a_d20_22	28.69
14	a_d28_30	39.19	a_d28_30	44.55	a_b70	32.9	a_p95	27.67
15	a_d16_18	38.91	a_d6_8	43.36	a_p30	32.29	a_p40	26.63
16	a_dns_16m	36.98	a_d22_24	36.42	a_p20	32.15	a_dns_8m	25.8
17	a_p95	36.47	a_d18_20	36.4	a_dns_20m	32.02	a_d6_8	25.13
18	a_d6_8	35.88	a_d20_22	33.47	a_p05	30.23	a_b30	20.83
19	a_b20	33.03	a_p20	30.55	a_p95	29.86	a_d28_30	20.37
20	a_b90	32.58	a_d30_32	30.29	a_p10	27.54	a_dns_20m	19.22

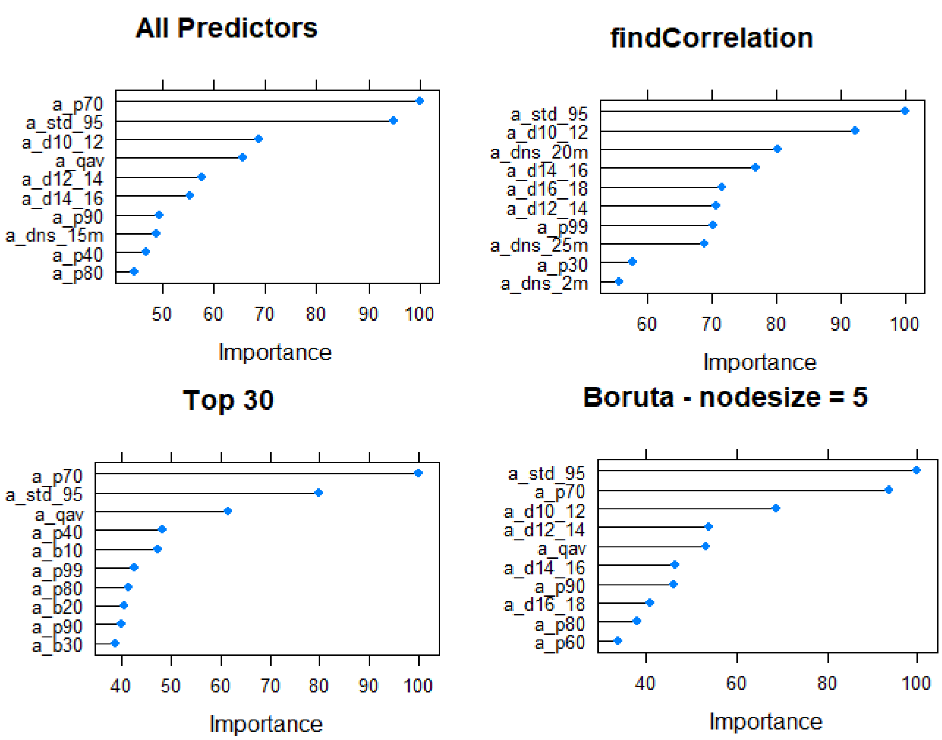

Figure 28. The variable importance diagrams are given for dbhq_merch for the top 10 predictors.

7.3 Discussion

7.3.1 Which subset of predictors?

The differences between the methods of selecting subsets of predictors were minor. There were 66 potential predictors, many of which were highly correlated. For "findCorrelation" and "top30", the subset of predictors was common across all dependent variables. This is convenient as it has the potential to reduce the number of SPL layers that need to be produced and possibly reduce processing time. The disadvantage is that only a subset of the dependent variables, all of which are overall size attributes, were examined. These tended to have top predictors like the upper p-values and upper slice data. It's possible that there are some dependent variables (e.g., biomass of poles) that might have some of the lower density slice data as important predictors.

The Boruta algorithm selects the subset of predictors based on the dependent variable. For the attributes examined, it reduced the number of predictors from 66 to between 35 and 48 (average 43), a reduction of about a third. There was some concern that including too many predictors, particularly if they are highly correlated, may introduce bias, particularly at the extremes. There was no evidence of this in the current study.

Processing time considerations are relatively minor and should not carry too much weight in deciding which algorithm to use in selecting a subset of predictors. The best subset of predictors did change with dependent variable. If there is concern about having too many predictors, using Boruta appears to be the best option. The differences in predictions between the different methods of selecting predictors were not compelling, and this may be a function of a small number of overall predictors to begin with. Going from 66 to 39 predictors did not result in a huge gain, especially when the results were so similar.

7.3.2 Mtry

The differences for the various values of mtry were minor. The default was p/3. The caret algorithm defaults to 3 values – the minimum mtry = 2, the maximum mtry = all variables, and the midpoint between the minimum and maximum. Setting mtry = 2 seemed a bit small. Using all the predictors isn't recommended because that eliminates one of the advantage of using RF: the inherent randomness of the algorithm. The consensus was to use the default of p/3, as the differences between p/3 and p/2, in terms of prediction outcomes were minor.

7.3.3 VarImp

For the area attributes (TVOL_merch and BA_merch), "All", "top30" and "Boruta", the top 5 predictors all had variable importance estimates of over 85%. For the tree attributes (CD_ht and Dbhq_merch), for "All", "top30" and "Boruta", only the top 2 predictors had variable importance estimates > 85% and all the rest were less than 80%. For the area attributes, more predictor variables had high variable importances while for the tree attributes, there are a few predictors with high variable importances.

7.4 Conclusion

The differences in the predictions using the full set of predictors and the Boruta subset were minor. The Boruta subset of predictors varies with each dependent variable so it is likely all predictors would need to be calculated and any reductions in processing time would be minor.

The differences between and mtry of p/2 and p/3 were minor so the default (p/3) was used.

Therefore, in this analysis the conclusion was to use the full set of predictors and an mtry value of p/3.

8. Appendix C – Additional plots (PRF vs. PRF + CNL)

The calibration data consisted of fixed-area field plots covering a range of forest types and development stages. The impact of including additional CNL plots on model development were tested. Two sets of calibration data were used here:

PRF – Field plots were only those measured at the Petawawa Research Forest (PRF) in the summer of 2018. There were 249 field plots.

PRF&CNL – Field plots included the 249 field plots along with the additional 20 plots from the Canadian Nuclear Laboratory (CNL) lands measured in the summer of 2019, for a total of 269 field plots.

The addition of the 20 CNL plots was tested by comparing predictions calibrated using the 249 PRF plots and the 269 PRF&CNL plots. FPC_T0 SPL metrics were used (see Appendix A).

The two calibration datasets above were compared using a similar approach to the different SPL subsets and metrics (as described in Appendix A). The estimates from the two calibration datasets can be thought of as repeated measures on the validation stands. The effect of sample size was tested using the Wilk's lambda (Table 27). Top height was the only attribute examined that had a significant dataset effect. As noted above, field measurement protocols for top height differed between the calibration and validation data and this confounds these types of comparisons. In the validation dataset, there were 35 trees with heights > 40m and 5 trees with heights > 45m. One height of the thickest stems counted was measured on each field plot so these were all top heights. On the calibration plots, two field plots had a top height > 40m, one with top height > 45.The average top height on the CNL plots was 26.7m while the average top height on the 249 plots was 24.3m. The CNL plots had more tall trees which likely helped the top height predictions of pixels with tall trees. Thus, the significant result for top height needs to be considered within this context and should not be considered definitive. None of the other attributes considered in the analysis saw a significant effect of the different calibration datasets.

Table 27. The results of the repeated measures ANOVA, the Wilk's Lambda, are given by attribute. Wilk's Lambda is a test of the H_0: no effect of calibration dataset.

	Degrees of	freedom	Top height	Tvol merch	Tvol merch	DQ merch	BA merch	Lorey height
Source	Numerator	Denominator	(m)	(m^3/ha)	(m^3)	(cm)	(m^2/ha)	(m)
Mean	1	18	0.0332	0.507	0.2669	0.9328	0.5999	0.1439
Forest type	8	18	0.6335	0.1259	0.3636	0.3616	0.2649	0.4557

Predictions of forest attributes were also evaluated by regressing the observed value on the predicted value. Good predictions would have a regression with an intercept of 0 and a slope of 1. The null hypothesis H0: intercept b_0 = 0 and slope b_1 = 1 was evaluated using the simultaneous F-test (Table 28).

None of the predictions showed any departures from the assumption of an intercept of 0 and slope of 1. The results for top height suggests that there may be bias in the estimates; however this confirms the aforementioned limitations of the validation data for top height. The range of observed top height was approximately 20–40m, therefore although the regression of observed on predicted had an intercept of approximately 2m, the standard error associated with that estimate was fairly large, leading to the conclusion that the intercept is not statistically different from zero.

Table 28. The results of the simultaneous F-test which tests the H_0: $b_0 = 0$ and $b_1 = 1$ for the regression of observed on predicted attribute. If the F-test was statistically significant, the regression was examined to see whether it was as a result of the intercept or slope (or both). The significant term is given.

Attribute		PRF	PRF&CNL
Top height	$P(F > F_{obs})$	<.0001	<.0001
	Significant term	neither	neither
	Intercept	1.93548	1.96309
	Slope	1.02864	1.02683
TVOL_merch	$P(F > F_{obs})$	<.0001	<.0001
(m^3/ha)	Significant term	neither	neither
	Intercept	-51.5864	-34.7657
	Slope	1.19265	1.12245
TVOL_merch	$P(F > F_{obs})$	<.0001	<.0001
(m^3)	Significant term	neither	neither
	Intercept	-36.5018	-49.2984
	Slope	0.98527	0.98821
DQ merch	$P(F > F_{obs})$	<.0001	<.0001
(cm)	Significant term	neither	neither
	Intercept	-3.48915	-5.18102
	Slope	1.15731	1.22949
BA merch	$P(F > F_{obs})$	<.0001	<.0001
(m^2/ha)	Significant term	neither	neither
	Intercept	-3.9007	-3.57522
	Slope	1.14145	1.12781
Lorey height	$P(F > F_{obs})$	<.0001	<.0001
(m)	Significant term	neither	neither
	Intercept	-3.75039	-4.34341
	Slope	1.18318	1.20858

9. Appendix D – predicting merchantable volume

Two options were explored for predicting merchantable volumes (TVOL_merch and mvol):
Option 1 - Predicting the volume ratios directly:

TVOL_merch_ratio = TVOL_merch/tvol_all and
mvol_ratio = mvol/TVOL_merch

Then calculating:
pred_TVOL_merch = pred_TVOL_merch_ratio * pred_tvol_all and,
pred_mvol = pred_mvol_ratio * pred_TVOL_merch.

Option 2 - Predicting the vbar ratios directly:

vbar_tvol_ratio = vbar_TVOL_merch/vbar_tvol,
vbar_mvol_ratio = vbar_mvol/vbar_TVOL_merch and,
ba_merch_ratio = ba_merch/ba_all

Then calculating:
pred_TVOL_merch = pred_tvol_all * pred_vbar_tvol_ratio * pred_ba_merch_ratio and,
pred_mvol = pred_TVOL_merch * pred_vbar_mvol_ratio.

The options were compared for the 27 validation stands and the results for TVOL_merch were very similar, with slightly larger bias for Option 1 (Table 29 and Figure 29).

Table 29. The predictions using options 1 and 2 for predicting TVOL_merch are compared. The two options give very similar results.

	Area		TVOL_merch (m^3/ha)			Bias ± standard error	
Forest type	(ha)	Polygons	Observed	Option 1	Option 2	Option 1	Option 2
BlackSpruce	13.7	2	185	199	199	-13.6 ± 9.7	-14.1 ± 9.7
JackPine	16.4	2	206	216	217	-10.7 ± 2.9	-11.2 ± 2.8
LowlandConifer	15.1	1	150	124	125	26 ± 0	24.7 ± 0
Mixedwood	69.5	4	221	226	227	-5.4 ± 5.5	-5.8 ± 5.5
Oak	37.5	2	279	286	287	-7.6 ± 1	-8.1 ± 1.2
Poplar	12.4	2	255	288	288	-33.2 ± 0.9	-33.8 ± 1.4
Pr Plantation	16.7	4	330	269	271	60.8 ± 19.3	58.7 ± 20.8
PwManaged	35.8	3	174	213	215	-39.7 ± 18.7	-41.5 ± 19.2
PwNatural	52.7	4	308	283	283	25.5 ± 4.3	25 ± 4.3
TolerantHwd	54	3	233	249	250	-16.6 ± 11.2	-17 ± 11.1
All	323.8	27	246	245	245	1.9 ± 7	1 ± 7.1

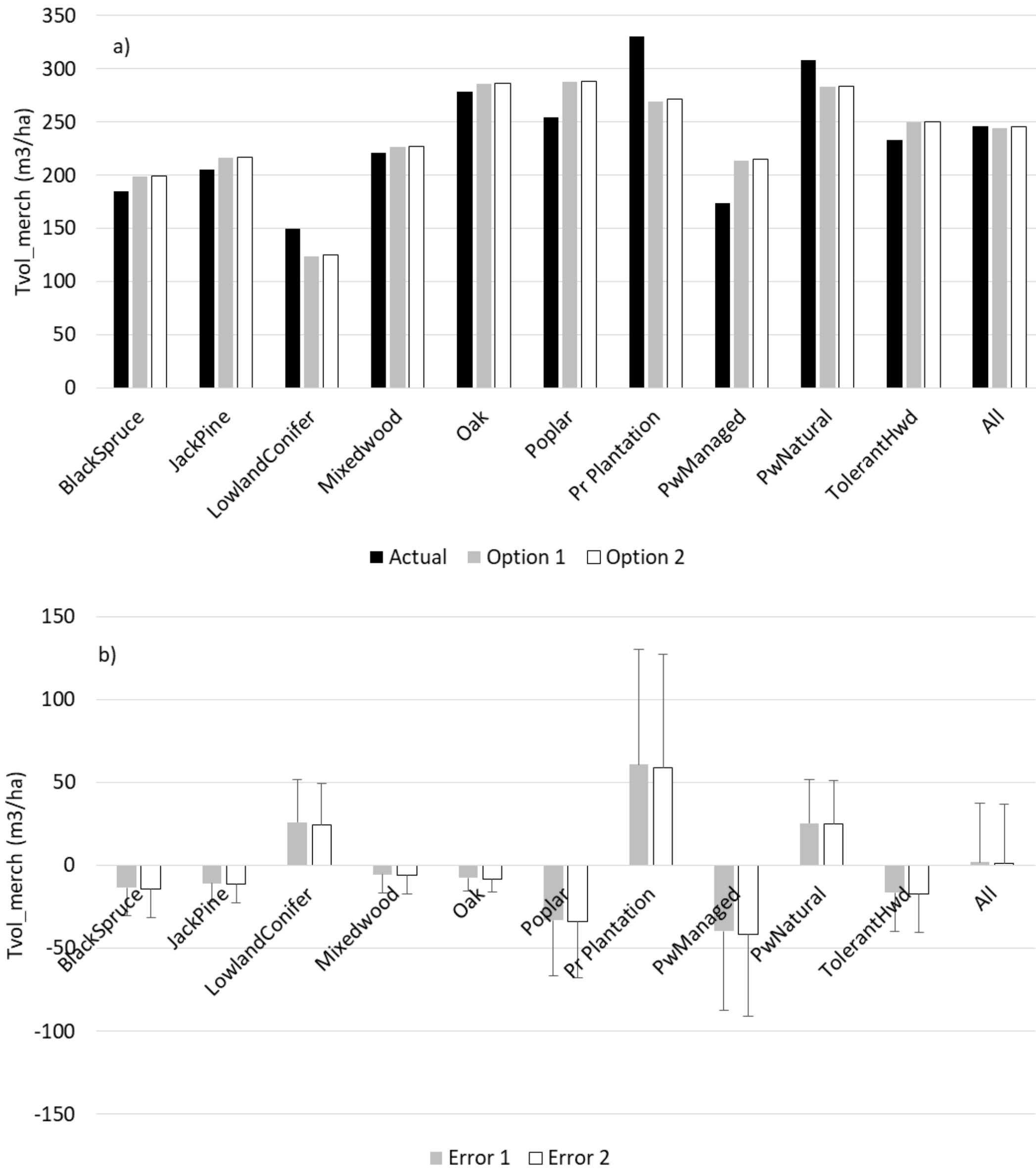

Figure 29. The predictions using options 1 and 2 for predicting TVOL_merch are compared. The two options giver very similar results.

Similar results were obtained when predicting mvol with the two options, again with bias being slightly greater for Option 1 (Table 30 and Figure 30). Based on the results of this analysis for TVOL_merch and mvol, Option 2 was selected and implemented for the final model.

Table 30. The predictions using options 1 and 2 for predicting merchantable volume are compared. The two options give very similar results.

	Area		Mvol		(m^3/ha)	Bias ± standard error	
Forest type	(ha)	Polygons	Observed	Option 1	Option 2	Option 1	Options 2
BlackSpruce	13.7	2	152	167	167	-14.5 ± 11.2	-15 ± 11.4
JackPine	16.4	2	183	187	187	-3.8 ± 6.4	-3.9 ± 6.4
LowlandConifer	15.1	1	125	97	97	28 ± 0	27.2 ± 0
Mixedwood	69.5	4	168	175	175	-7.2 ± 6.2	-7.3 ± 6.2
Oak	37.5	2	229	227	227	2 ± 4.9	1.9 ± 5.2
Poplar	12.4	2	210	245	245	-34.9 ± 10.5	-35.6 ± 10.2
Pr Plantation	16.7	4	295	237	239	57.2 ± 13.9	55.6 ± 14.9
PwManaged	35.8	3	161	193	195	-32.5 ± 18.6	-34.3 ± 19
PwNatural	52.7	4	266	240	241	26.2 ± 5.4	25.6 ± 5.3
TolerantHwd	54	3	178	194	194	-16.5 ± 13.6	-16.6 ± 13.6
All	323.8	27	208	204	205	3.1 ± 6.6	2.4 ± 6.6

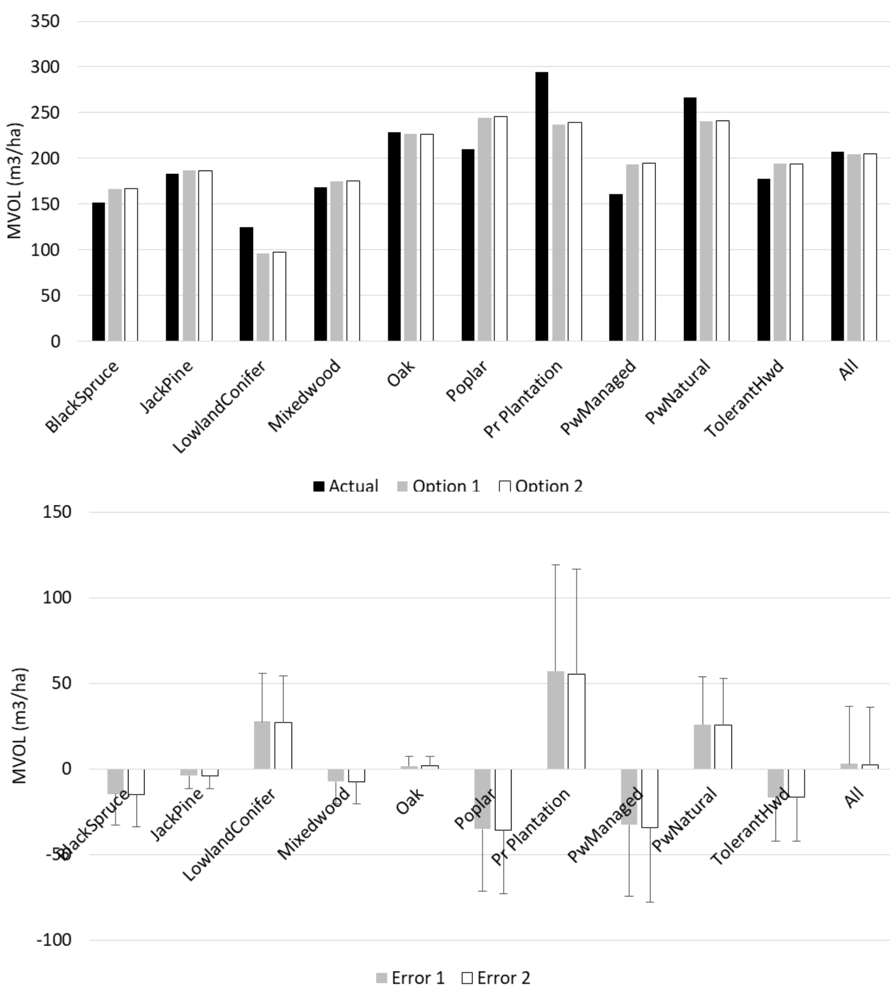

Figure 30. The predictions using options 1 and 2 for predicting merchantable volume are compared. The two options giver very similar results.

10. Appendix E - The effect of the number of ground returns

One concern with SPL is whether there are sufficient ground returns under dense forest canopies to support area-based forest attribute modelling. The area-based predictions generated in this study were examined to see if predictions were affected when there were few ground returns. The predictions of CD_ht (Figure 31) did not appear to be affected by the number of ground returns. A separate study is examining the effect of canopy cover on the digital elevation model (DEM) but the results here appear to indicate the DEM is adequate.

Likewise, the predictions of volume (TVOL_merch) did not appear to be affected by having few ground returns (Figure 32). The plot with the highest volume had few ground returns and was underpredicted. The out of bag predictions for the plot with the highest volume are based on plots with lower volumes, leading to the underprediction, and this underprediction is likely more related to the inability of RF to extrapolate these extreme values of the TVOL_merch range.

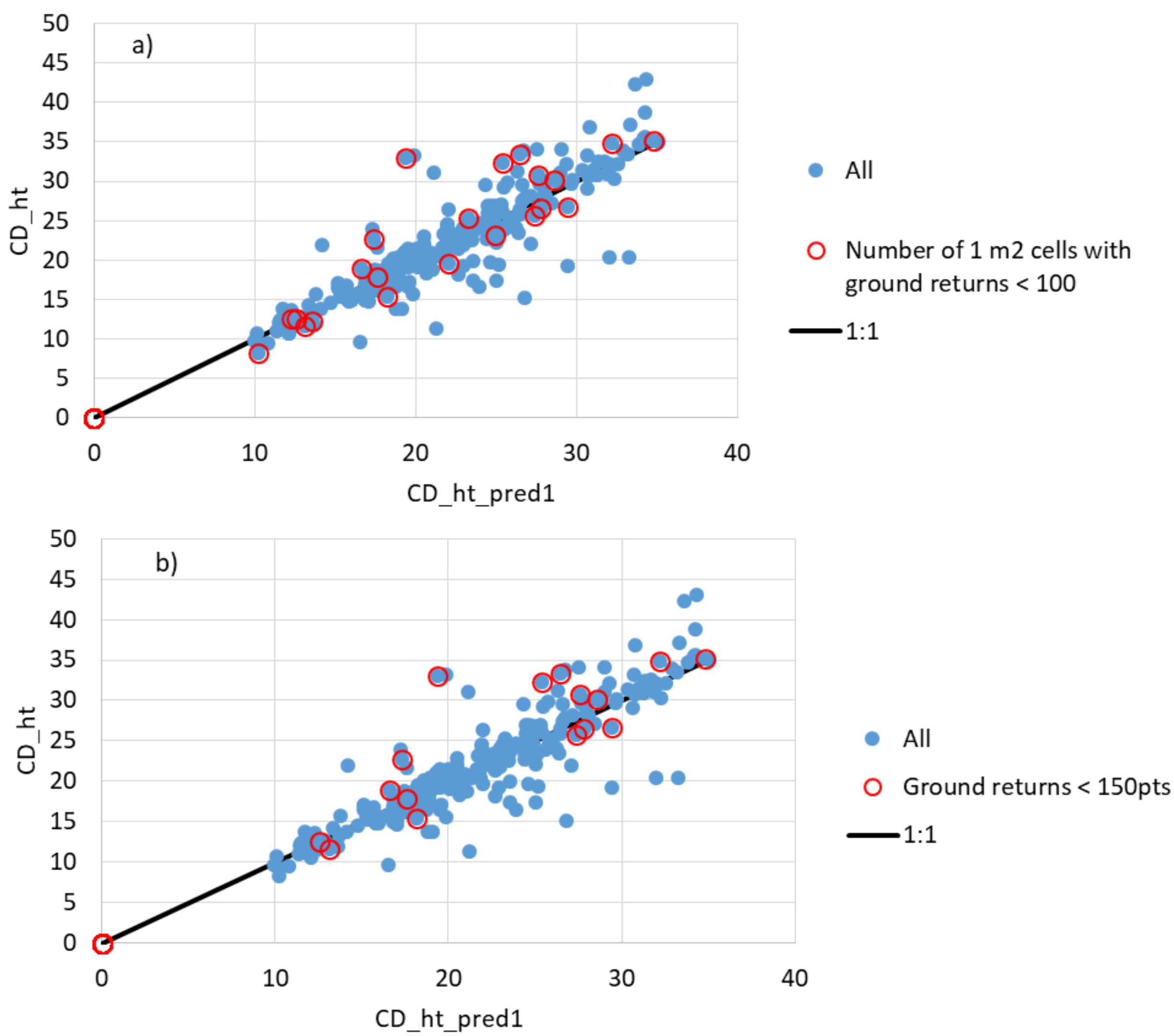

Figure 31. The predictions of CD_ht (using "findCorrelation" to select predictors) do not appear to be affected when the less than 100 of the 1 m^2 cells within plot have ground returns (a) or when the total number of ground returns is low (b).

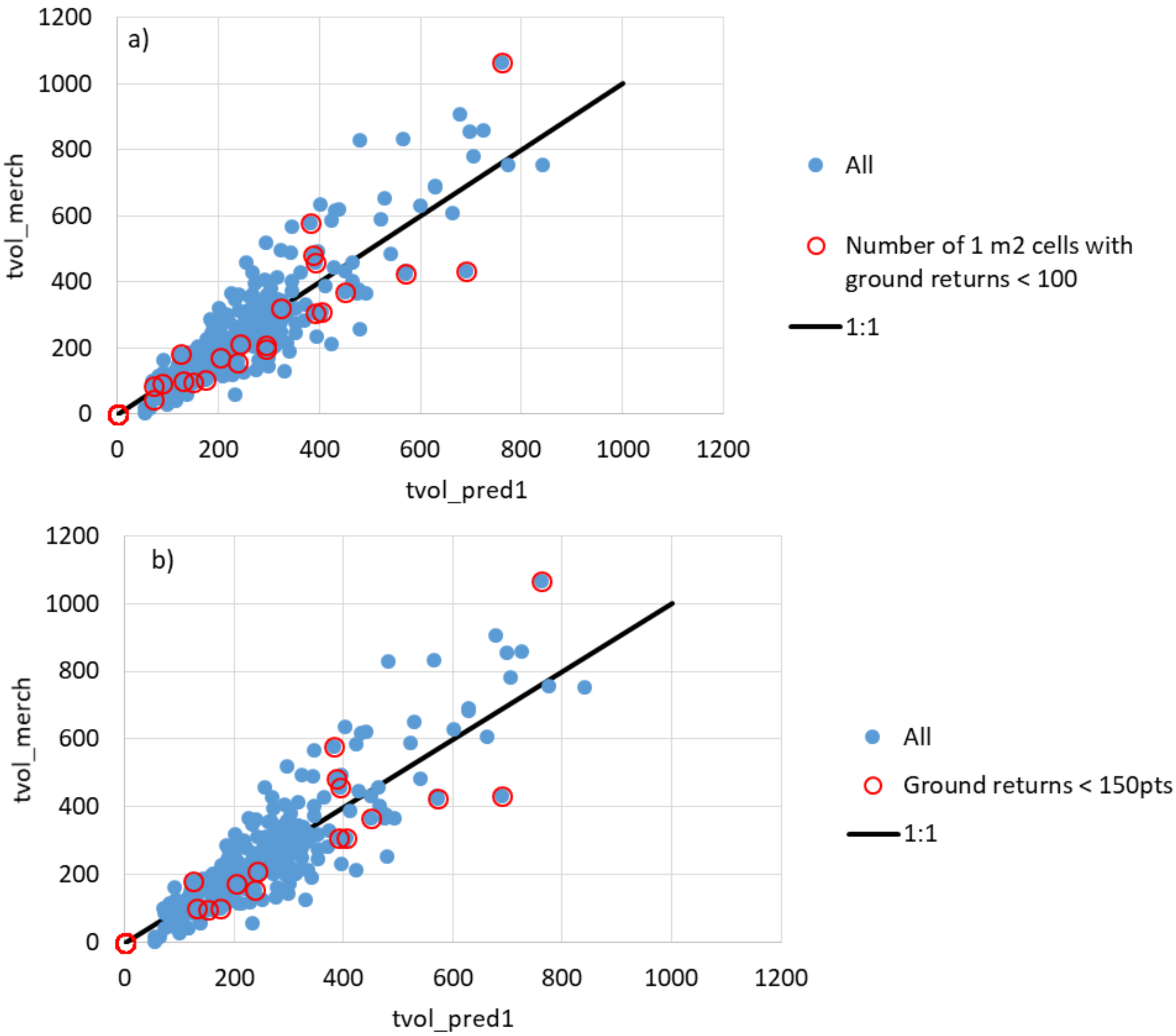

Figure 32. The predictions of TVOL_merch (using "findCorrelation" to select predictors) do not appear to be affected when the less than 100 of the 1 m^2 cells within plot have ground returns (a) or when the total number of ground returns is low (b). The ground plot with the highest volume has few ground returns and is underpredicted. However, RF has a tendency to underpredict high values and may not be related to the low number of ground returns.