LiDAR Derived T2 Inventory for the Algonquin Park Forest

KTTD Project # 20B-2021

July 2022

Report prepared for: Forestry Futures Trust

Report prepared by: Murray Woods & Margaret Penner

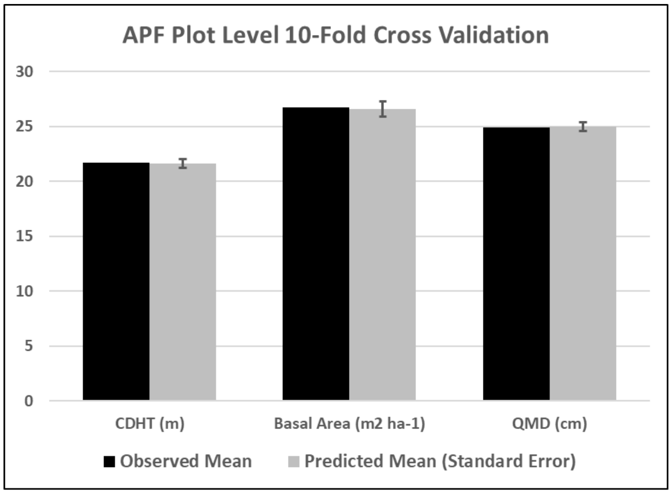
Table of Contents

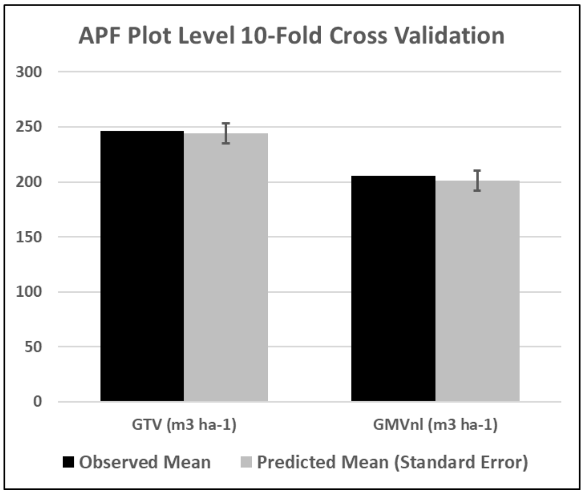
Executive Summary

Single Photon LiDAR (SPL) was acquired over the Algonquin Park Forest (APF) during the summer of 2019. A total of 225 LiDAR calibration plots (400m^2 – 11.28m radius) were established on the APF and measured between July 7, 2020 through to August 24, 2020. These plots were used to derive an inventory update ("T2") based on LiDAR models for Height (Dominant/Codominant, Lorey, Top Height), Basal Area (BA), Basal Area merchantable (BAmerch), Volumes (Gross Total (GTV), Gross Merchantable (GMV_NL and GMV_WL)), Quadratic Mean Diameter (QMD), Total Above Ground Biomass (Biomass), Stems, and Basal Area and Gross Merchantable volume by four-size classes. Merchantable volume predictions used the provincial scaling specifications for upper diameter limits along with a 30cm stump height. An additional predicted volume raster was produced for the Algonquin Forest Authority (AFA) for specific red/jack pine utility poles.

Plot level Model Validation

A 10-Fold Cross Validation (CV) of plot level (400m^2) predictions were calculated as a measure of model performance. Root Mean Square Error (RMSE) of models for height were 12.0% and 7.6% for Dominant/Codominant and Top height respectively. BA had a 23.1% RMSE while volumes (GTV, GMV_NL, GMV_WL) had 25.1%, 29.1% and 30.1 % respectively. QMD reported an RMSE of 20.5% and Biomass 23.2%. Stems resulted in an RMSE of 48.0%. Examples of mean observed and model predictions (along with standard error) of inventory attributes from cross validation are provided below.





Stand level Model Validation

Additional validation of the LiDAR predictions for 18 cruised stands was conducted. A stand (or harvest block) represents the scale inventory estimates will be used to support management decisions. Two validation stands were identified as outliers, and these were excluded in a second comparative summation of the results. The majority of inventory attribute RMSE's declined at the stand level from that reported via CV at the plot scale. Height attributes are not significantly impacted by scale. However, attributes such as ones expressed per area (i.e., volume) are. CDht RMSE for the validation stands (with outliers N=18/without outliers N=16) was 18%/18%. RMSE for QMD, BA, GTV, GMV, and Biomass were reduced to 12%/13%, 18%/13%, 23%/17%, 28%/20% and 16%/15%.

T2 Polygon updating

Raster (20 x 20m) surfaces of the LiDAR predictions were created for the forest polygons. Polygon layers were created from the raster surfaces using the T1 (OPI) polygon layer. The polygon attributes were calculated as the mean of the raster predictions within the polygon **where age ≥ 20 years**. Stand level QMD was calculated from polygon BA and Stems. These polygon-based estimates, were used in conjunction with T1 polygon age and species composition to calculate the following additional T2 inventory attributes:

- Site Index
- Stocking
- Cull Fraction
- Net Merchantable Volume (NMV).

Objective

The objective of this Forestry Futures Trust Knowledge, Transfer & Tool Development (KTTD) project is to develop open source (OS) software code for processing Ontario's Single Photon (SPL) Light Detection and Ranging (LiDAR) and to produce a raster-based product suite and an update for a new T2 polygon Forest Resources Inventory (FRI) for the Algonquin Park Forest (APF).

Study Site

The APF Forest has a total area of 763.009 ha (Figure 1) and is located in the Great Lakes - St. Lawrence Forest Region. 63.9% of this area is under Forest Management, 25.7% is Parks (Wilderness, Nature Reserves, etc.) 10.1% is Water, and 0.3% is Patent and Other. The forest contains two major forest complexes – the tolerant hardwoods/hemlock communities (western portion) and white and red pine communities found on the eastern portions (AFA Forest management Plan 2021-2031 FMP Online (gov.on.ca). A detailed breakdown of the AFA Forest Units is presented in Figure 2.

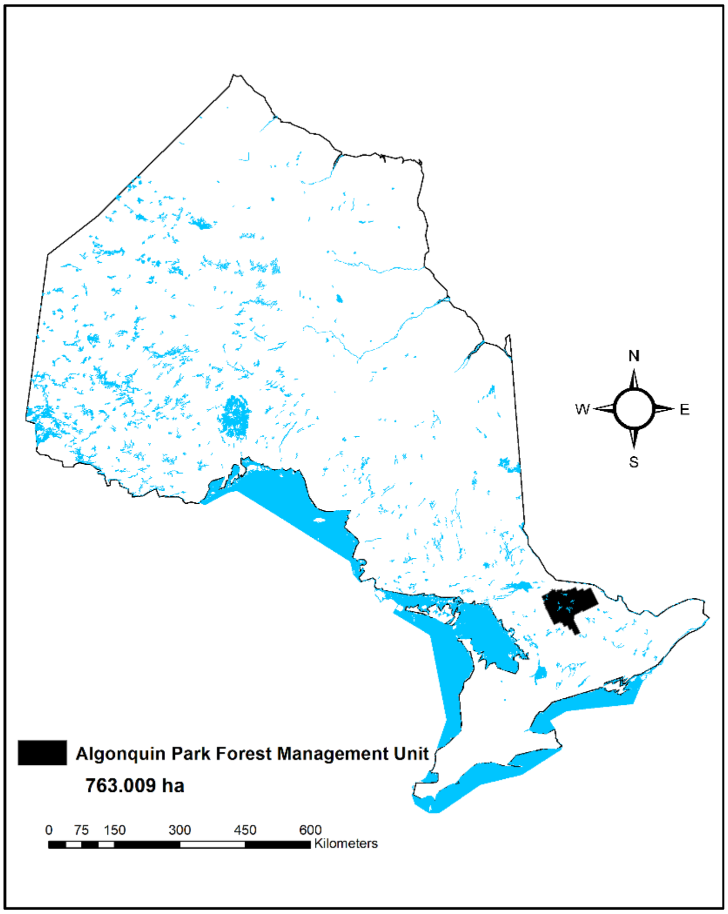


Figure 1 – Algonquin Park Forest Management Unit Location

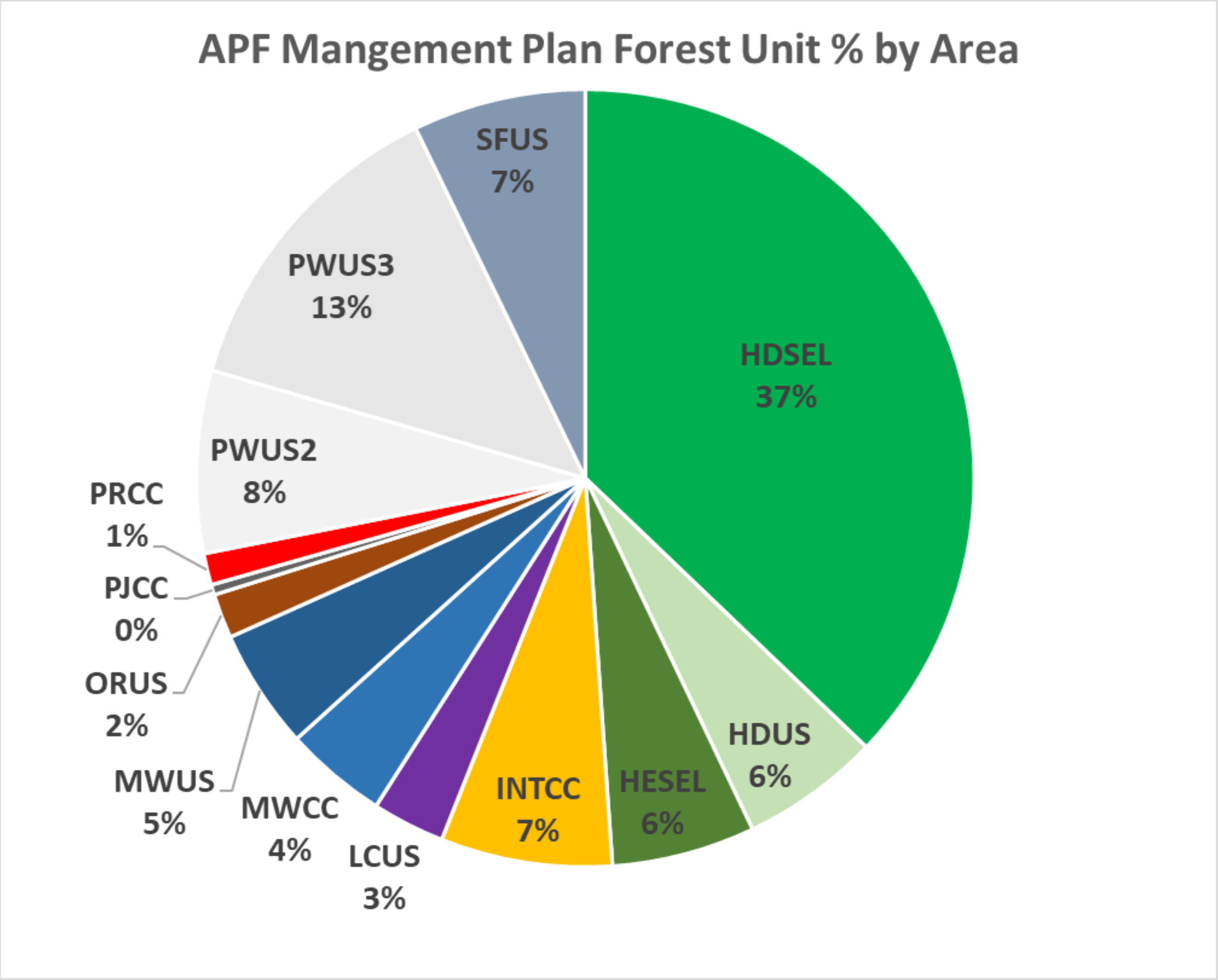


Figure 2 - Percent area by Plan Forest Unit for the APF.

Data

Airborne LIDAR data

Single Photon LiDAR (SPL) was acquired over the APF during the summer of 2019. The SPL100 sensor was flown aboard a Piper–PA–31–350 at an average altitude of 3760m. More details of acquisition parameters are provided in Table 1.

Table 1 - LiDAR acquisition specifications for 2019–SPL mission

Parameter	2019 – SPL Description
Pulse repetition rate	6000 KHz
Frequency	21Hz
Scan Angle	+/– 15 Degrees
FOV	30 Degrees
Swath Width	2000m

LiDAR Model Calibration Data

Calibration ground sample measurements followed the province of Ontario's Vegetation Sampling Network Protocol document *(Science and Research Technical Manual TM)*. The Vegetation Sampling Network (VSN) protocol consists of 3 potential plot measurement methodologies. "*A* modules" provide a base set of attributes for all plots. They include a range of stand attributes, tree attributes, and site and substrate attributes. "*B* modules" add in protocols for stem mapping and crown delineations and for assessing a smaller tree and shrub subplot, both of which support LiDAR diagnostics and development. When applied to the permanent subset of VSN plots, the smaller tree and shrub subplot module also supports tracking recruitment and succession. "*C* modules" apply only to the permanent plot subset and add some focus on understory vegetation (understory vegetation subplot) and down woody debris, as well as tree deformities and evidence of wildlife use. The A plot measurement thresholds, common to all protocols, were used to include as many plots as possible in this project.

A total of 225 LiDAR calibration plots (400m^2 – 11.28m radius) were established and measured between July 7, 2020 and September 24, 2020. Calibration plots were selected using a "structurally guided" approach. LiDAR structure measurements for the population were used to determine the full range of structural conditions. Calibration plots were then selected to sample the range of conditions. Where possible, existing provincial permanent sample plots were incorporated into the sampling framework where they met required structural conditions. These plots become the link between ground attributes (i.e., heights, volumes, etc.) and the LiDAR point cloud.

Data Quality

Initial data screening steps quickly identified some field measurement quality issues on the calibration plots established on the APF. Total tree height measured by the field crew was often found to be higher than the maximum LiDAR return acquired for that plot. In some cases, the differences were extreme. There were some situations where the opposite was true too, field measured heights being substantially shorter than the maximum LiDAR return. Figure 3 provides some examples where field heights were found to exceed maximum LiDAR returns.

Taking quality height measurements, especially during leaf-on periods of the year on these tall tree species requires extra time and care. Possible reasons for the field measured height issues:

- These are tall trees. For accurate height measurement, it is recommended the heights be measured from a distance as least as great as the height.
- In some cases, the canopy cover is dense, particularly in tolerant hardwoods when leaves are on. It may be difficult to see the top of the tree to get a good measurement. It may also be difficult to identify the top of the tree in tolerant hardwoods.
- There were issues with the height measurements. It appears the height measurements corresponded to the height above 1.3m, not the height above the ground. 1.3m was added to every height.

Additional complications and challenges of working with the APF data set:

- There were no field audits, so data and measurement issues were not identified and corrected.
- There may be GPS errors leading to the field plots not lining up exactly with the LiDAR Point cloud.

- Due to the GPS technology being used by the field crews to locate the target plot location, plots may not have been established at the target location and may not sample the intended structural class.
- Some plots don't have many ground returns which can impact the LiDAR normalization (this is more likely seen in dense tolerant hardwood plots).
- Sometimes trees lean in or out of the plot.
    - For leaning trees, it's not clear whether the crew measured the height of the tree tip above the ground (which seems to be the field manual procedure) or the length of the bole. There is a tree with a degree of lean = 90 (Dbh = 7.9 cm, height = 4.1). One tree has Dbh = 29.1, height = 29.0 and a degree of lean of 80. No adjustments were made to height based on degree of lean.
- Some plot point clouds contain returns from crowns of trees outside the plot. These trees were not noted as leaning into the plot and therefore, have no mensuration information (Figure 4)
- Also, in some cases there can be tall dead trees captured in the extracted point cloud (Figure 5)

Figure 6 provides a comparison on maximum LiDAR return versus maximum field height measured on each plot. The 1:1 line indicates identical measurements. It is clearly evident that a significant proportion of calibration plots over-estimated the largest tree height (Figure 6) and the average of the 2 largest field measured heights (Figure 7).

The issues with height measurement quality also raised suspicion on the care undertaken on using the GPS to identify the target plot location, DBH measurement or determination of what trees were within the 11.28m radius plot boundary. Unfortunately, there is no way with the LiDAR returns to evaluate the data quality.

Calibration Plot Data Adjustment

Adjustments to field data measurements is something that should never be required. However, it was clear that the field measurements seem to be higher than LiDAR measurements. Possible explanations for these differences have been discussed. It is critical that future field crews understand that the quality of the field data collection impacts the quality of the derived inventory product suite.

However, because tree height has a large impact on the calculation of tree volume, we felt that an adjustment was required for this dataset (and likely any dataset where these field crews collected height measurements).

A decision was made (in consultation and approval of NDMRNF FRI staff) to adjust the field heights using the relationship between the height of the tallest tree on the plot (MaxHt) and the maximum LiDAR return (zmax). A "plot level ratio" adjustment was made to each plot for the APF. Where no suitable height trees were available to make a plot level ratio adjustment, the population level adjustment was be used.

Examples of the adjustments to heights and the impact on gross total volume are presented below (Figure 8 and Figure 9).

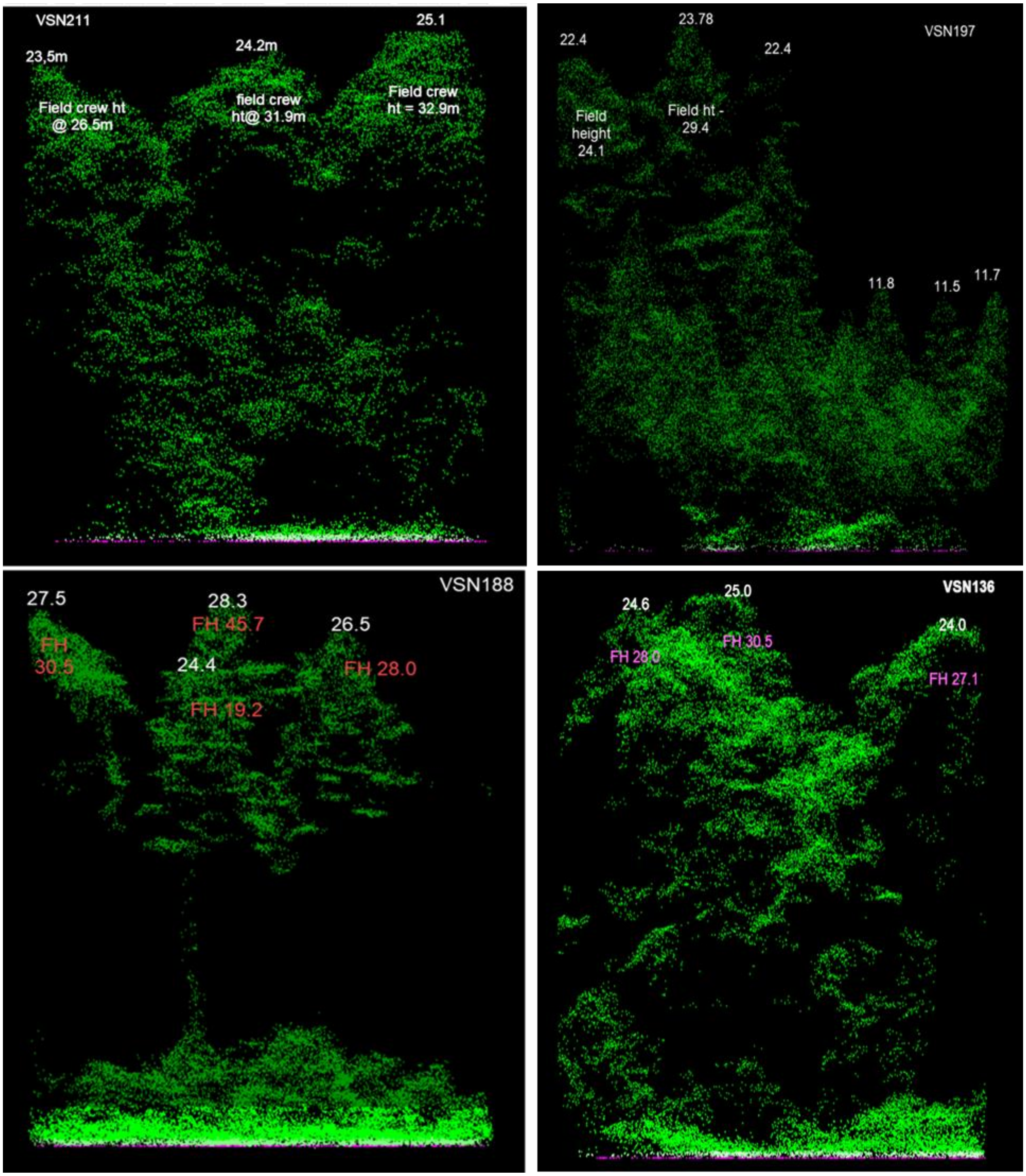


Figure 3 - Examples of Field Height (FH) overestimating Maximum LiDAR return. Field heights assigned to trees from tallest to shortest.

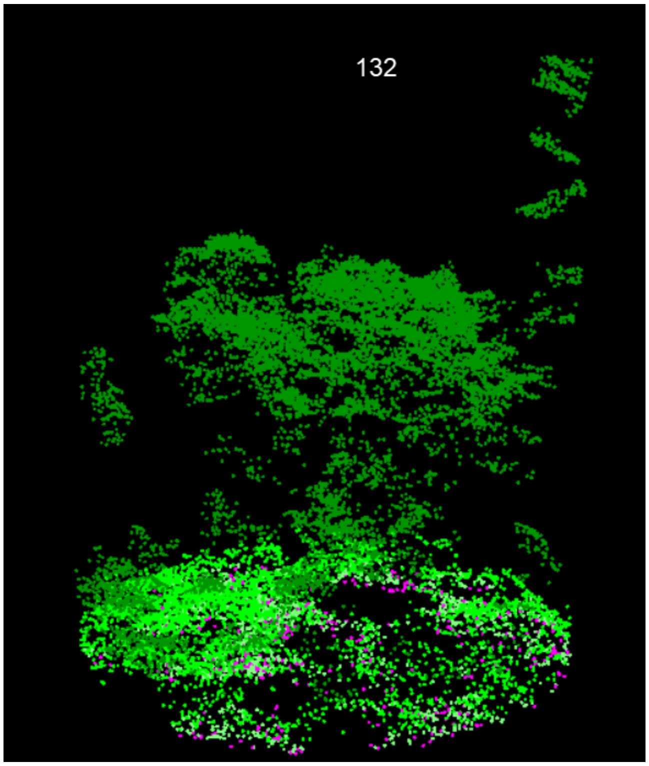


Figure 4 - Example of a plot with crown returns from a tree outside and above the calibration plot.

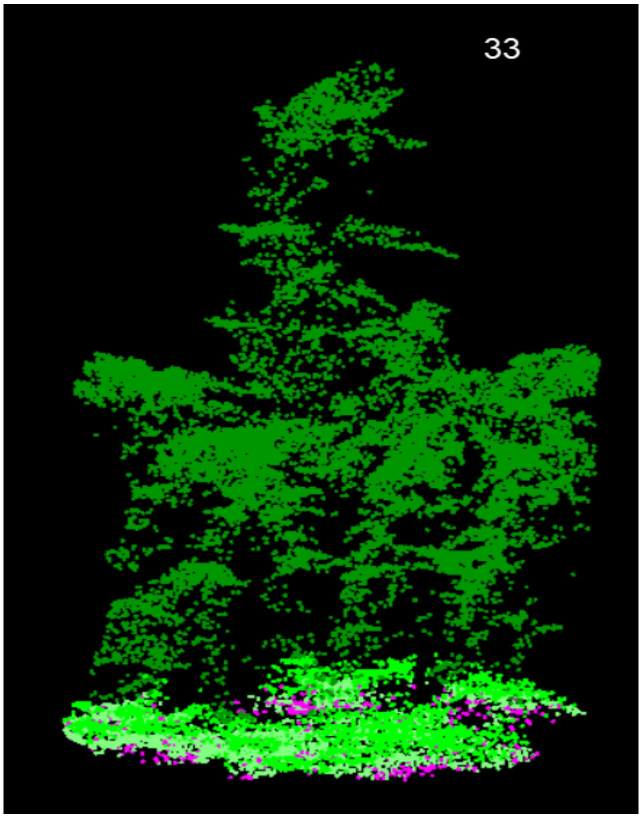


Figure 5 - Example of a tall dead tree capturing lidar returns. This tree is not summarized as part of the plot conditions

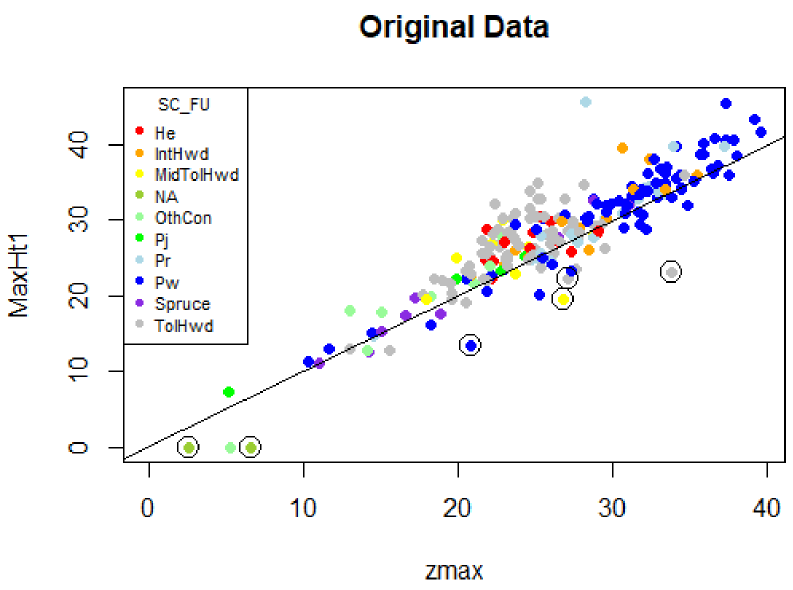


Figure 6 - The height of the tallest live tree on the plot (MaxHt1) is plotted against the maximum LiDAR return (zmax). The 1:1 line is given. The SC_FU is based on the leading species. The circled plots are ones where tree crowns outside the plot tally are impacting the LiDAR zmax or where the tallest tree was dead. In particular there were a number of plots where the tallest tree was quite a bit taller than the highest LiDAR return.

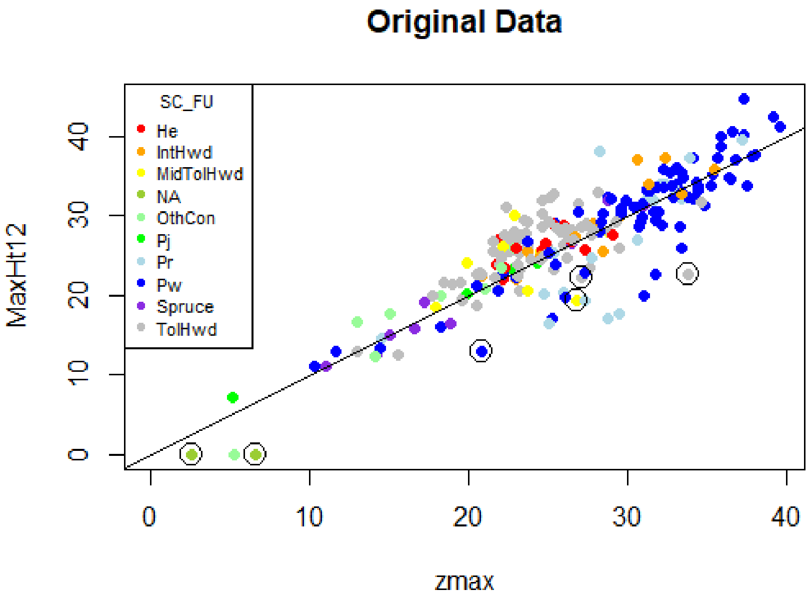


Figure 7 - The same as Error! Reference source not found. **except the average of the two tallest trees is plotted on the y-axis. The average difference between MaxHt1 and MaxHt12 was 1.34 m**

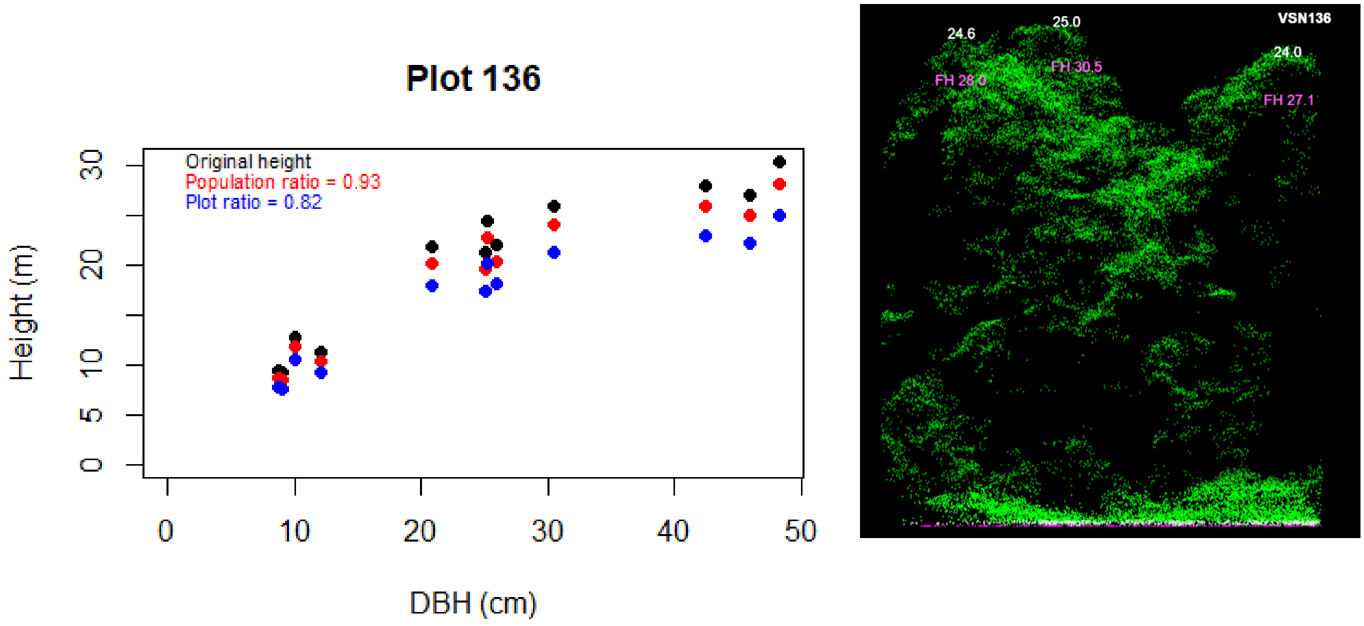


Figure 8 - For this plot, both the population and plot adjustments are downward. Unadjusted GTV 193.5 m3/ha, adjusted GTV = 164.3 m3/ha.

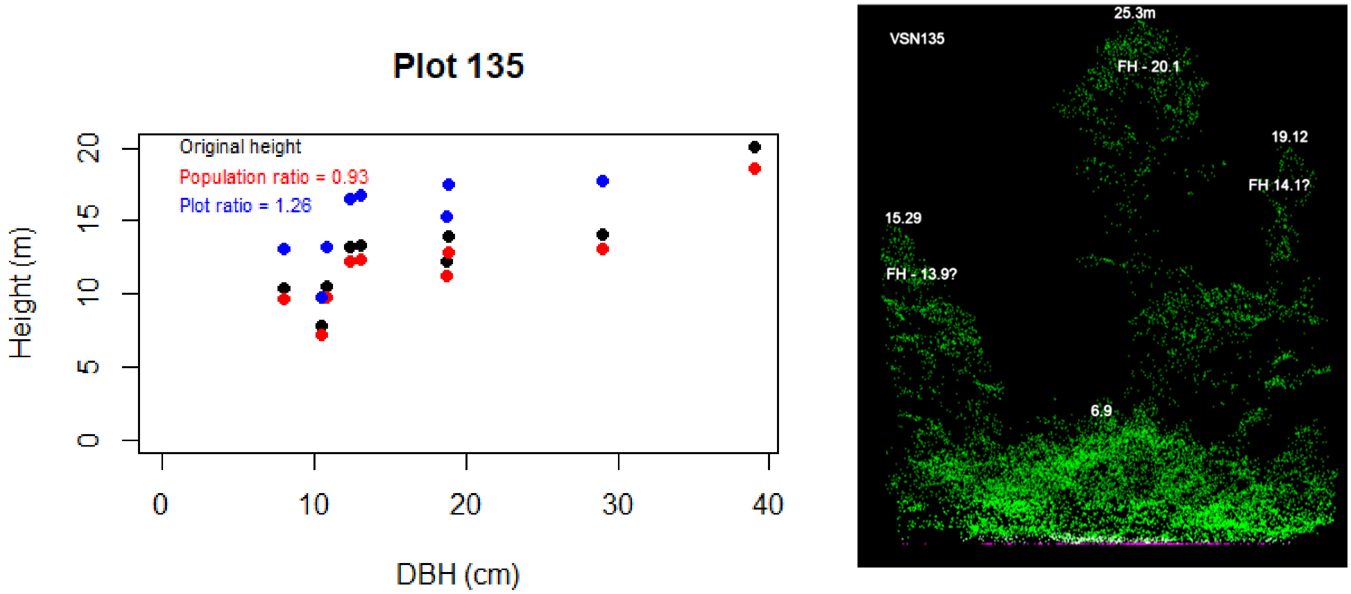


Figure 9 - Here is an example where the field crew underestimated the heights. The population ratio adjusts height down while the plot ratio adjusts height higher. Unadjusted GTV = 49.3, Adjusted GTV = 61.6 m3/ha

Plot Compilation

For all live trees with DBH $\geq$ 7.1cm (common minimum DBH threshold for all VSN plot types) species, origin, Dbh, height, vigour and crown class were recorded. On some plots ages were recorded for a sample of trees. For dead trees ≥ 10cm (and > 2m), species, Dbh, height, vigour and decay class were recorded. Trees that had crowns leaning in or out of the plot were noted as were broken top trees.

Plots were summarized to per hectare values for all live trees ≥ 7.1cm. Dead trees were also summarized for their informational value in explaining potential differences noted between modeling results and plot summaries. However, dead trees were not used to calibrate the LiDAR models.

An approved provincial standard set of inventory attributes were summarized for model prediction. In addition to these, staff managing the APF requested an additional volume summarization (based on a red pine utility pole specification) of the calibration data and subsequent modeling products. Table 2 lists the inventory attributes that were summarized for modeling (live trees with DBH ≥ 7.1cm unless noted) on the APF. Individual tree volumes were calculated using Zakrzewski and Penner (2014) taper models developed for Ontario. No height estimation was required for the APF dataset as each tree had a measured height. In the case of the APF dataset, the “adjusted” height was used.

Individual tree total above ground biomass was calculated by species using the equations published in Lambert et al. (2005). Individual species equations were used when available. When no species coefficients existed, broader “hardwood” or “softwood” model coefficients were used.

Table 2 - Inventory attributes summarized from calibration plots and predicted from LiDAR. Volume estimates came from Zakrzewski and Penner 1983. Biomass estimates came from Lambert et al. 2005.

Inventory Attribute	Units	Description
Stems	Stems ha^{-1}	Number of live trees
BA	m^2 ha^{-1}	Basal Area (Dbh ≥ 7.1cm)
BAmerch	m^2 ha^{-1}	Basal Area Merchantable (Dbh ≥ 9.1cm)
CDht	m	Average CoDominant-Dominant height
LoreyHeight	m	Lorey Height. Mean height weighted by basal area
TopHt	m	Top Height defined as height of the 100 largest DBH trees per hectare (irrespective of species)
QMD	cm	Quadratic mean diameter
GTV	m^3 ha^{-1}	Gross Total Volume (includes stump and top)
GMV_NL	m^3 ha^{-1}	Gross Merchantable Volume with no minimum piece length requirement. Stump height 30cm and upper diameter as per Ontario Scaling Manual (Table 3)
GMV_WL	m^3 ha^{-1}	Gross Merchantable Volume in 2.54 m log lengths • Stump height 30cm and upper diameter as per Ontario Scaling Manual (Table 3)
BA_Poles	m^2 ha^{-1}	Basal Area for the Pole size class. [9 < Dbh ≤ 25 cm]
BA_SmS	m^2 ha^{-1}	Basal Area for the Small Sawlog size class [25 < Dbh ≤ 37 cm]
BA_MedS	m^2 ha^{-1}	Basal Area for the Medium Sawlog size class. [37 < Dbh ≤ 49 cm]
BA_LgS	m^2 ha^{-1}	Basal Area for the Large Sawlog size class. [Dbh > 49 cm]
GMV_Poles	m^3 ha^{-1}	Gross Merchantable Volume (_NL) with no minimum piece length requirement for the Pole size class. [9 < Dbh ≤ 25 cm] • Stump height 30cm and upper diameter as per Ontario Scaling Manual (Table 3
GMV_SmS	m^3 ha^{-1}	Gross Merchantable Volume ((_NL)) with no minimum piece length requirement for the Small Sawlog size class. [25 < Dbh ≤ 37 cm] • Stump height 30cm and upper diameter as per Ontario Scaling Manual (Table 3)
GMV_MedS	m^3 ha^{-1}	Gross Merchantable Volume ((_NL)) with no minimum piece length requirement for the Medium Sawlog size class. [37 < Dbh ≤ 49 cm • Stump height 30cm and upper diameter as per Ontario Scaling Manual (Table 3)
GMV_LgS	m^3 ha^{-1}	Gross Merchantable Volume ((_NL)) with no minimum piece length requirement for the Large Sawlog size class. [Dbh > 49 cm] • Stump height 30cm and upper diameter as per Ontario Scaling Manual (Table 3)
Biomass	Tonnes ha^{-1}	Total above ground biomass (wood + bark + branches + foliage)
GMV_Util	m^3 ha^{-1}	Gross merchantable volume to APF specifications for red pine utility poles

Table 3 - Minimum upper diameter limits for merchantable volume calculation by species group

Species	Minimum Diameter Outside Bark (DOB)
Hardwoods (except poplar/white birch)	18cm class, 17.1 cm
Conifers (except White and Red Pine, Hemlock	10cm class, 9.1 cm
White & Red Pine Hemlock	14cm class, 13.1 cm
Poplar, White Birch	14cm class, 13.1 cm
Source: OMNRF. 2020. Scaling Manual, Toronto. Queen's Printer for Ontario. 169 pp ISBN:978-1-4868-4495-1	

Calibration Plot Spatial Positioning

Once target plot locations were identified, all established plots were spatially located with a survey grade GNSS system. Data was post–processed to meet required sub–metre positional requirements.

Exclusion of Calibration Plots

As noted earlier, LiDAR was acquired for the bulk of the APF forest during the summer of 2019 and plot measurements were initiated in July 2020 through to the end of September, 2020. The intent of the calibration plots is to capture vegetation conditions that match the LiDAR measurements. However, some calibration plots sampled structural conditions made up of trees too small (minimum Dbh threshold of 7.1cm or < 5m in height) to provide opportunity for summarization and inclusion in the modeling. Table 4 identifies the 2 plots excluded from the calibration of the LiDAR and their reason for removal. A total of 223 calibration plots remained to produce the LiDAR inventory. Further filtering of calibration plots for model construction is discussed later.

Table 4 - APF calibration plots excluded from analysis

Plot Number	Reason for Exclusion
VSN451061	Young stand with no Live trees >= 7.1cm
VSN451097	Young stand with no Live trees >= 7.1cm

A summary of the calibration plots by APF (FUs) (Assignment SQL provided in Appendix E) is provided in Table 5. Of note is the number of calibration plots per FU. Some conditions seem under sampled while others appear oversampled. This disparity in sample size by FU is a function of the structural sampling approach adopted by the province of Ontario. Forest conditions with a wide range of vertical structures (i.e., pine shelterwoods) were sampled more than more "simple" structures often found in conditions like pure red pine plantations.

LiDAR Data Processing

Raw classified LiDAR LAS datasets were provided to the province by the vendor. Standard American Society for Photogrammetry and Remote Sensing (ASPRS) classification coding standards were used by the vendor. Classification codes (2) ground , (3) low vegetation , (4) medium vegetation and (5) high vegetation return data only were processed. LAStools (LAStools, 2021) was used to "normalize" the

Table 5 – Statistics – Mean (range) of calibration plots by APF Forest Units[1] on the APF used for LiDAR modeling

APF-Forest Unit	No Plots	Breast Height Age (yrs)[2]	TopHt (m)	CDHT (m)	Lorey Ht (m)	Stems (ha)	Basal Area (m^2 ha^{-1})	QMD (cm)	GTV (m^3 ha^{-1})	GMV_NL (m^3 ha^{-1})	GMV_WL (m^3 ha^{-1})	Biomass (Tonnes ha^{-1})
HDSEL	70	73 (N=61) (25 - 145)	21.3 (11.4 - 34.6)	19.1 (10.5 - 25.6)	19.3 (10.8 - 25.6)	651 (275 - 1375)	24.9 (4.4 - 46.2)	22.7 (9.7 - 33.3)	193 (24.2 - 409.5)	140 (0 - 343)	131 (0 - 332)	168 (21 - 357)
HDUS	6	66 (N=5) (39 - 91)	20.3 (16.2 - 28.8)	17.7 (14.8 - 23.1)	18.0 (15 - 23.7)	788 (275 - 1475)	25.8 (14.6 - 30.6)	22.4 (16.2 - 37.6)	186 (90.3 - 263)	127 (48 - 236)	116 (37 - 231)	154 (69 - 203)
HeSel	13	111 (N=12) (46 - 176)	22.1 (18.2 - 26.3)	20.2 (14.2 - 25)	20.9 (15.7 - 25.5)	548 (250 - 1050)	33.4 (21.4 - 51.5)	29.5 (17.2 - 42.1)	247 (131 - 395)	206 (99 - 365)	199 (96 - 357)	187 (94 - 281)
INTCC	5	94 (55 - 120)	27.9 (19.7 - 33.3)	25.1 (20 - 30.9)	24.4 (19.4 - 30)	835 (675 - 1075)	45.4 (34.7 - 64.8)	26.4 (20.8 - 29.9)	491 (313 - 708)	442 (260 - 643)	428 (244 - 624)	276 (181 - 409)
LCUS	9	74 (38 - 102)	13.3 (N=8) (9.7 - 17.9)	11.4 (N=8) (8.1 - 16.2)	11.8 (N=8) (8.4 - 16.5)	1028 (25 - 2000)	17.3 (0.1 - 37.2)	13.4 (7.3 - 24.7)	92 (0 - 236)	66 (0 - 206)	60 (0 - 199)	54 (0 - 109)
MWCC	3	69 (54 - 87)	26.8 (18.5 - 33.4)	22.6 (16.7 - 27)	22.4 (16.2 - 26.7)	692 (650 - 750)	32.3 (14.6 - 43.5)	23.7 (16.9 - 28.7)	326 (101 - 495)	266 (59 - 432)	254 (49 - 422)	206 (83 - 285)
MWUS	2	76 (58 - 94)	23.2 (21.3 - 25.1)	19.1 (17.2 - 21)	19.4 (16.2 - 22.6)	1025 (650 - 1400)	27.3 (24.3 - 30.4)	19.2 (16.6 - 21.8)	227 (213 - 241)	180 (159 - 201)	166 (144 - 189)	134 (130 - 138)
OrUS	1	56 (56 - 56)	15.4 (15.4 - 15.4)	14.9 (14.9 - 14.9)	13.9 (13.9 - 13.9)	925 (925 - 925)	23.7 (23.7 - 23.7)	18.1 (18.1 - 18.1)	147 (147 - 147)	102 (102 - 102)	95 (95 - 95)	118 (118 - 118)
PjCC	3	46 (16 - 74)	15.5 (4.2 - 22.4)	13.2 (4 - 19.7)	13.0 (4.1 - 20.7)	517 (300 - 925)	11.7 (5.2 - 18.3)	19.2 (8.5 - 26.8)	81 (8 - 165)	72 (0 - 155)	71 (0 - 152)	47 (9 - 87)
PrCC	21	110 (N=20) (46 - 136)	27 (14.1 - 35.6)	26.5 (11.9 - 33.6)	25.2 (12.9 - 32.7)	451 (50 - 1425)	24.6 (2.9 - 58.2)	28.1 (16.9 - 38.5)	293 (32 - 843)	270 (30 - 788)	264 (29 - 778)	141 (15 - 403)
PwUS	82	90 (N=68) (9 - 155)	28.3 (8.6 - 37.8)	24.7 (6.3 - 35.4)	24.7 (7.1 - 34.9)	589 (25 - 2075)	27.0 (3.0 - 66.5)	27.3 (11 - 59.7)	283 (16 - 881)	252 (6 - 825)	246 (4 - 815)	148 (10 - 439)
SFUS	8	70 (34 - 115)	20.3 (13.9 - 29.3)	16.9 (11.1 - 23.6)	17.4 (11.9 - 24.2)	1488 (450 - 3525)	31.1 (12.4 - 50.5)	18.4 (11.5 - 27)	225 (65 - 450)	173 (46 - 412)	165 (39 - 406)	138 (43 - 241)
All	**223**	**85 (N=197) (9 - 176)**	**24.3 (N=222) (4.2 - 37.8)**	**21.6 (N=222) (4.0 - 35.4)**	**21.6 (N=222) (4.1 - 34.9)**	**660 (25 - 3525)**	**26.5 (0.1 - 66.5)**	**24.8 (7.3 - 59.7)**	**243 (0 - 881)**	**203 (0 - 825)**	**195 (0 - 815)**	**154 (0 - 439)**

---

[1] APF Forest Unit syntax was used to assign FU. However, some information like Site Class (used in some FU assignment) or Stage of Management was not available at the plot level so was not used. As a result, HDSEL and HDUS may have some confusion and Pine Shelterwoods (3-cut and 2-cut) were combined.

[2] Breast height age is the average breast height age of dominant/codominant trees with measured ages. Trees were not measured for age on all plots and the sample sizes for age are less than the number of plots.

LiDAR returns to the terrain (converting “z” height from elevation to height above ground). An additional script was implemented to compress the LAS formatted files to a space efficient LAZ format.

A modeling predictor set on a 20m x 20m grid was created for the 2018 LiDAR data set using the lidR (Roussel and Auty 2020, Roussel et al. 2020) software package in R (R development Core Team 2020). A total of 112 potential LiDAR predictors were derived from structural statistical queries of all-return, normalized point cloud data. Following testing of predictive model performance from thresholding the returns at 0 m and 2.0 m, a decision was made to use all returns greater than 0 m for modeling inventory attributes on the APF. This choice of threshold was also documented in other studies in Ontario (White *et al.* 2021, Woods *et al.* 2011). Data “z” spikes were removed by dropping any returns > 48m. A complete list and description of the LiDAR predictors created is provided in Appendix A. Predictors that were selected for use in the predictive models are indicated.

LiDAR Model Development

A non-parametric Random Forest (RF) model (Liaw and Wiener 2002) solution via the statistical package R (R development Core Team 2020) was used for the prediction of inventory attributes. All model predictions were made at the plot scale and at a 20 m raster cell (matching the 400 m^2 plot size) with the model mtry parameter set to the default (number of predictors/3) and the parameter ntree (number of trees to construct) set to 1000. Only calibration plots with zq99 > 5m were used in the prediction of stand level metrics to better align with the calibration plot minimum DBH of 7.1 cm. This filter resulted in the dropping of an additional calibration plot from the modeling but ensured that only plots with at least some merchantable sized trees were utilized in the models and the predictions made at the landscape level. In the prediction of merchantable volume attributes, calibration plots with Zq99 > 9m were used as plots with Zq99 ≤ 9m had little or no merchantable volume.

Investigation of the initial modeling of specific inventory attributes of (BA, BA_merch, and QMD), identified that calibration plots consisting of tolerant and mid-tolerant hardwoods (> 50% hardwoods) were being generally overpredicted by a single un-stratified RF model intended to model all forest species conditions. The desire to utilize a nonparametric modeling approach like RF for the derivation of a LiDAR inventory is to eliminate the requirement for species information, usually only interpreted and provided at the polygon scale. In most situations, a dynamic RF modeling solution of matching point-cloud distribution statistical measurements at a pixel level (20m x 20m) and desired inventory attribute summaries, without any a priori knowledge of species, has resulted in flexible models (i.e., White et al. 2021) capable of predicting attributes a range of species conditions. However, it became clear for the APF forest and this SPL dataset, that creating a stratified, 2 RF model solution resulted in better predictions for some of the inventory attributes (a basal area comparison of a single-strata vs 2-strata is presented in Figure 10). The list of inventory attributes predicted by a single or stratified RF model approach and modeling strata description is presented in Table 6.

LiDAR predictions for each attribute were made independently. In most cases (e.g., DomCodom height, Top Height, Lorey Height) this works well. However, to ensure some logic and biological consistency in predictions, some attributes were predicted as a fraction of other attributes. An example

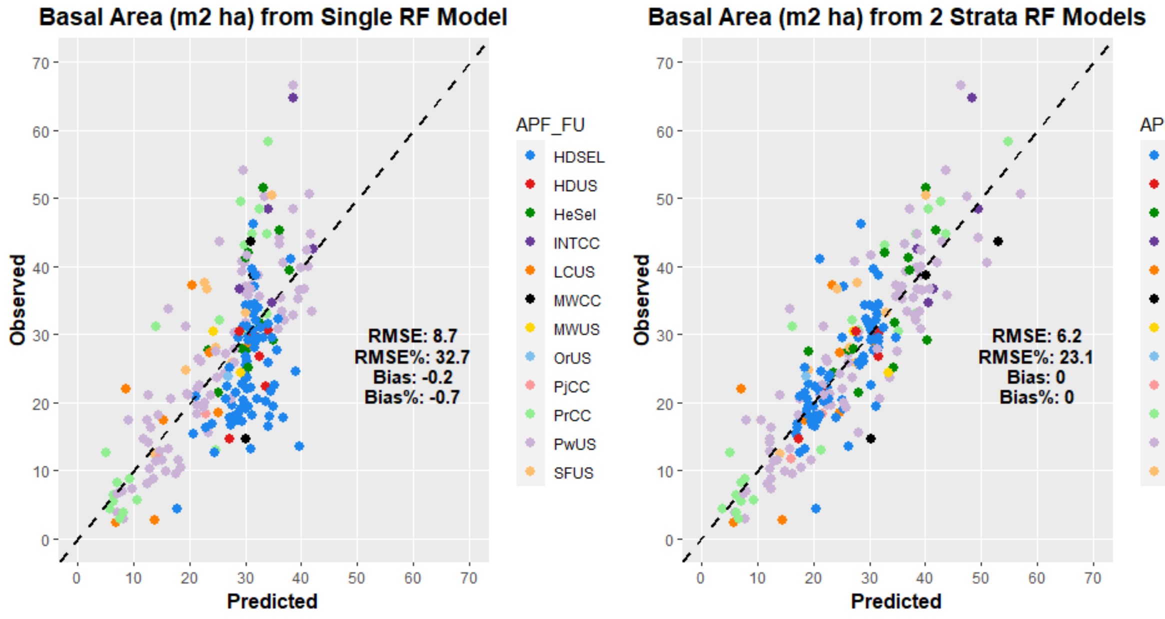


Figure 10 - Comparison of a single RF model solution versus a stratified RF model solution for basal area on the APF (units are m2 ha-1)

Table 6 - Modeling method adopted for the APF.

Inventory Attribute	**Modeling Approach**
Topht	Single Strata Model
CDht	Single Strata Model
LoreyHt	Single Strata Model
Vbar_gtv	Single Strata Model
Biomass	Single Strata Model
gmvnl_ratio	Single Strata Model
gmvwl_ratio	Single Strata Model
UtilPole_ratio	Single Strata Model
Ba by Size Class	Single Strata Model
GMV by Size Class	Single Strata Model
Basal Area	Two Strata Model -Tolerant Hardwood >50% , Conifer + Intolerant Hardwoods
Basal Area Merch	Two Strata Model -Tolerant Hardwood >50% , Conifer + Intolerant Hardwoods
QMD	Two Strata Model -Tolerant Hardwood >50% , Conifer + Intolerant Hardwoods

of such an attribute is gross merchantable volume (GMV). Actual GMV is never larger than gross total volume (GTV). To constrain the prediction of GMV, the fraction of GMV/GTV was predicted. Different constraining approaches were tested and the rationale for the method chosen for the various volume predictions is described below.

Gross Total Volume (GTV)

Rather than predicting GTV directly, it was predicted as a function of basal area (BA) and the volume to basal area ratio (vbar). Both options were tested and resulted in similar RMSEs and biases. The vbar option to estimate GTV was chosen as it may help preserve a bit of the relationship between BA and GTV by ensuring the predicted vbar is always within the range observed in the calibration data.

1. BA is predicted directly.
2. vbar_GTV = GTV/BA is predicted directly.
3. GTV is calculated as predicted BA x predicted vbar_GTV

Gross Merchantable Volume (GMV)

All merchantable volumes are constrained to be less than or equation to the predicted GTV. This is accomplished through predicting the ratio GMV/GTV.

1. Predict GTV using the method described above
2. Predict ratio GMV = GMV/GTV directly
3. Calculate GMV as GTV x ratio GMV

This is mathematically equivalent to constraining the vbar_GMV to be less than or equal to vbar_GTV.

$$ratio_{GMV} = \frac{GMV}{GTV} = \frac{vbar\_GMV}{vbar\_GTV} = \frac{GMV/_{BA}}{GTV/_{BA}}$$

All merchantable volumes (GMV_NL, GMV_WL and GMV_UtiilPoles) were constrained against GTV. Merchantable volumes (i.e., GMV_NL and GMV_WL) were not constrained to be greater or equal to each other.

Size Class estimates of Merchantable Basal Area and GMV_NL

Size class estimates of merchantable volume and basal area were constrained to always sum to either predicted GMV_NL or Basal Area_merch. An example of the method for basal area by size class is described below. First, the merchantable BA was split into BAMedLg and BASmPole. Then BAMedLg was split into BAmedium and BAlarge and BASmPole was split into BAsmall and BApole. Similar splits were made for GMV_NL

- Calculate BAmerch = BaPoles + BAsmall + BAmedium + BAlarge from calibration plot data
- Calculate a BAmedium + BAlarge fraction of plot BAmerch from calibration plot data

  BAMedLg_frac <- (BAlarge + BAmedium)/BAmerch

- Calculate fraction of Large BA in Medium and Large Sawlogs from calibration plot data

  BALg_ratio <- BAlarge/(BAlarge + BAmedium)

- Calculate fraction of Small BA in Small sawlogs and Poles from calibration plot data

  BASm_ratio <- BAsmall/(BAsmall + BApoles)

- Develop RF models for: BAmerch, BALg_ratio, BASm_ratio
- Calculate basal area in medium and large sawlogs (BAMedLg)

  BA_MedLg <- (BA_MedLgSawFrac * BAmerch)

- Calculate the proportion of the predicted BA_MedLg is Large sawlog where P99 $\geq$ 20 else be set to 0 & resulting in value moves to the Medium sawlog

  BA_LgS <- ifel(zq99 >= 20, (BA_MedLg * BA_LgRatio), 0)

- Calculate the BA_MedLg sawlog where P99 $\geq$ 15 else be set to 0 & resulting in value moves to the small sawlog and pole basal area

  BA_MS <- ifel(zq99 > 15, (BA_MedLg - BA_LgS),0)

- Calculate the Basal area for SmallSawlog & Poles

  BA_SmPl <- (BAmerch - BA_LgS - BA_MS)

- Calculate the BA for Small Sawlogs

  BA_SmS <- (BA_SmPl * BA_SmRatio)

- Calculate Pole BA as the difference between predicted BAmerch and predicted Large, Medium and Small sawlog basal areas

  BA_Pole <- (BAmerch - BA_LgS - BA_MS - BA_SmS)

Table 7 indicates which attributes were predicted directly from the statistical predictor summaries of the raw LiDAR point cloud. Table 7 indicates which inventory attributes are calculated as a fraction of another one to help ensure logical predictions.

Table 7- Inventory attributes predicted directly from the point cloud predictors.

Inventory Attribute
TopHt
CDht
LoreyHeight
BasalArea
QMD
Biomass

Table 8 - Description of attributes and their calculations predicted indirectly. All attributes are summarized from > 7cm unless noted (P_ = Predicted)

Inventory Attribute	Calculation
Stems	Stems = (P_BasalArea / P_QMD^2) / 0.00007854
GTV	GTV = P_BasalArea * P_VBAR_GTV
GMV_NL	GMV_NL = P_GTV * P_GMV_NL_ratio
GMV_WL	GMV_WL = P_GTV * P_GMV_WL_ratio
BA_Poles	As described above [9 < Dbh ≤ 25 cm]
BA_SmS	As described above [25 < Dbh ≤ 37 cm]
BA_MedS	As described above [37 < Dbh ≤ 49 cm]
BA_LgS	As described above [Dbh > 49 cm]
GMV_Poles	As described above [9 < Dbh ≤ 25 cm]
GMV_SmS	As described above [25 < Dbh ≤ 37 cm]
GMV_MedS	As described above [37 < Dbh ≤ 49 cm]
GMV_LgS	As described above [Dbh > 49 cm]
APF_Util	APF_Util = P_GTV * APF_Utilpole_ratio * red pine/jack pine percent

LiDAR Model Results

All LiDAR predictions are based on the LiDAR structure statistics and the field plot measurement summaries only[3]. Figure 11 illustrates the observed versus the predicted estimate for each LiDAR model. The diagonal dashed line indicates a perfect match between the measured plot summary and the prediction.

Plot level Validation

All calibration plots were used in model training and prediction. As a result, no independent plots were available to test model prediction error with. A "Cross Validation" (CV) can be used to estimate prediction error at the plot scale (20m x 20m) in the absence of an independent validation data set. V-fold CV error is generated by dividing the data set randomly into *V* equal parts. Training for the model is done on *V-1* parts and testing is done on the remaining part. This is repeated many times (10 times in this study) and the error rate estimate is an average of the results.

RMSE and Bias were calculated using the following equations:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2}{n}},$$

$$\mathrm{RELATIVE\ RMSE} = \frac{\mathrm{RMSE}}{\bar{Y}},$$

$$\mathrm{BIAS} = \frac{\sum_{i=1}^{n}(\hat{Y}_i - Y_i)}{n},$$

$$\mathrm{RELATIVE\ BIAS} = \frac{\mathrm{BIAS}}{\bar{Y}}.$$

Plot level 10-fold CV comparisons of root mean square error (RMSE) and bias are presented by inventory attribute in Table 9. CV RMSE (%) AND bias (%) are graphically presented in Figure 12. These results reflect modeling of all species/silviculture/origin based solely on LiDAR point cloud structure and at the plot or 20 x 20m pixel scale. The RMSE is a measure of how well the model performed. It is the square root of the average squared distance between the predicted values and the observed values in the dataset. The lower the RMSE, the better the modeling results. Bias is the difference between the average prediction and the correct value. Similarly, a lower bias is always preferred.

---

[3] The field measurement summaries include species composition and age. However, they were not used in modeling except to use a tolerant hardwood model vs. other as noted in table 6.

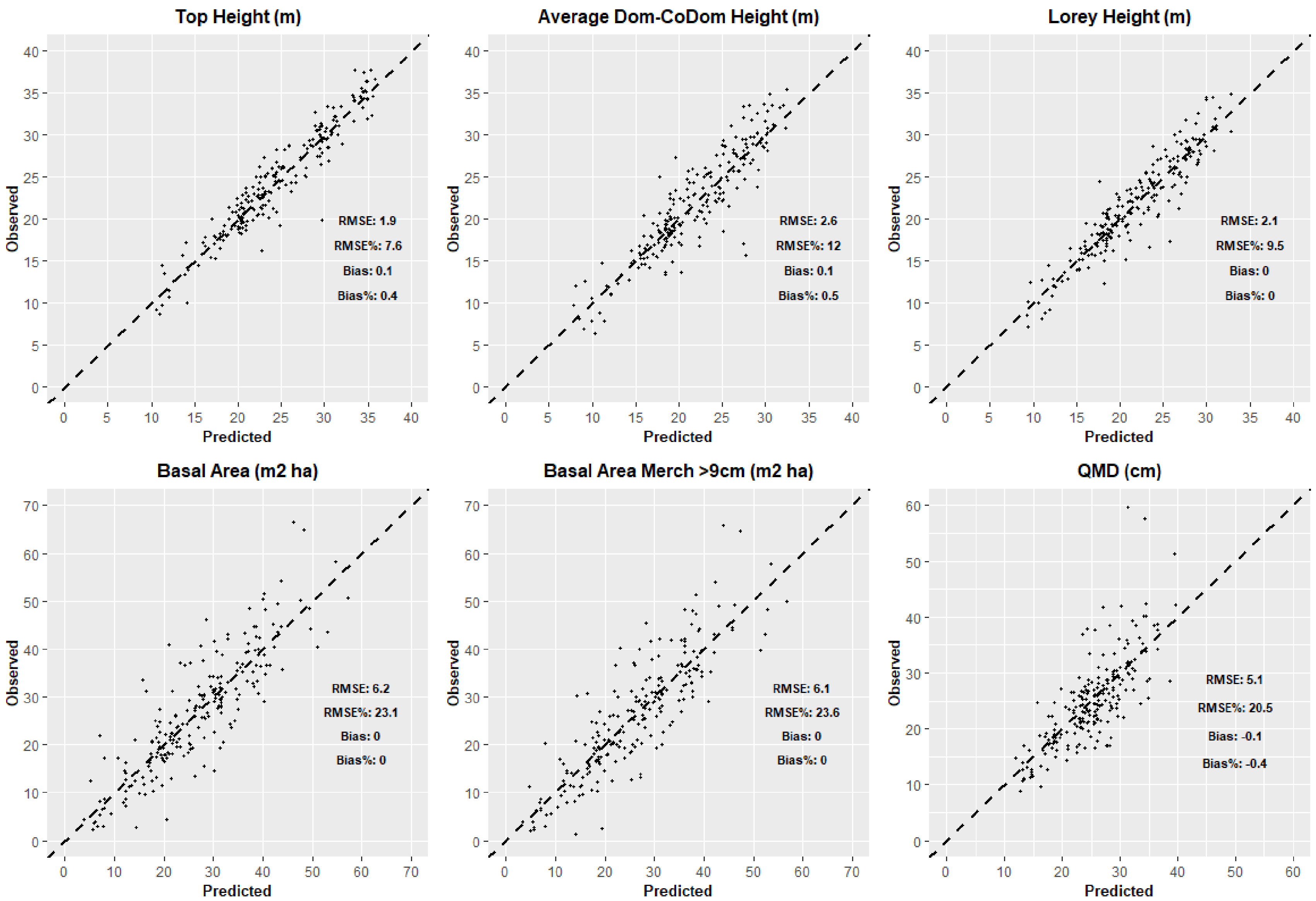
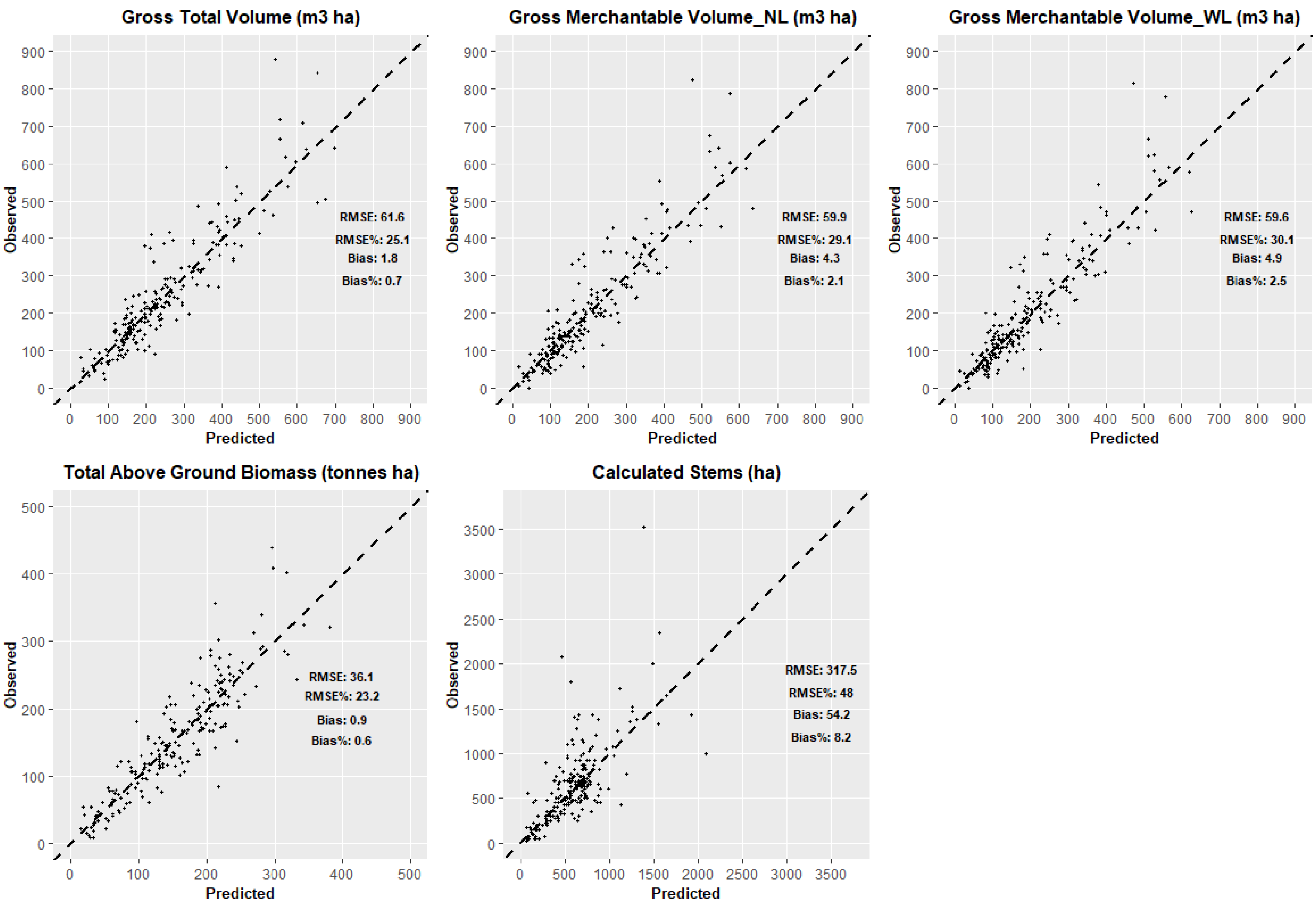


Figure 11 - Modeling results of Observed versus Predicted for selected inventory attributes on the APF. Error statistics are based on a 10-fold Cross Validation sample.



Gross Total Volume (m3 ha)

Gross Merchantable Volume_NL (m3 ha)

Gross Merchantable Volume_WL (m3 ha)

Total Above Ground Biomass (tonnes ha)

Calculated Stems (ha)

Figure 11 continued - Modeling results of Observed versus Predicted for selected inventory attributes on the APF. Error statistics are based on a 10-fold Cross Validation sample.

Table 9 - Plot level validation statistics using a10-fold Cross Validation methods for the APF

	Observed				10-Fold Cross Validation (CV)					
Inventory Metric	**N**	**Mean**	**Min**	**Max**	**P_Mean**	**P_SE**	**RMSE**	**% RMSE**	**BIAS**	**% BIAS**
CDHT m	221	21.7	6.3	35.4	21.6	0.4	2.6	12.0	0.1	0.6
TOPHT m	221	24.4	8.6	37.8	24.3	0.4	1.9	7.6	0.0	0.2
LoreyHt m	221	21.7	7.1	34.9	21.7	0.4	2.1	9.5	0.0	0.0
QMDht m	221	22.7	7.2	36.1	22.7	0.4	2.3	10.0	0.0	0.0
BA m^2 ha^{-1}	221	26.7	2.3	66.5	26.6	0.7	6.2	23.1	0.0	0.0
BA merch m^2 ha^{-1}	221	26.0	1.3	65.9	26.0	0.7	6.1	23.6	0.0	0.0
QMD cm	221	24.9	9.0	59.7	25.0	0.4	5.1	20.5	-0.1	-0.4
GTV m^3 ha^{-1}	220	246	13.2	880.7	244.2	9.4	61.6	25.1	1.8	0.7
GMV_NL m^3 ha^{-1}	220	205.7	0.0	825.1	201.3	9.0	59.9	29.1	4.3	2.1
GMV_WL m^3 ha^{-1}	220	197.9	0.0	815.4	193.1	8.9	59.6	30.1	4.9	2.5
Biomass T ha^{-1}	221	155.4	8.3	439.4	154.6	4.9	36.1	23.2	0.9	0.6
Stems ha^{-1}	221	661	25	3525	607	22	317	48.0	54.2	8.2

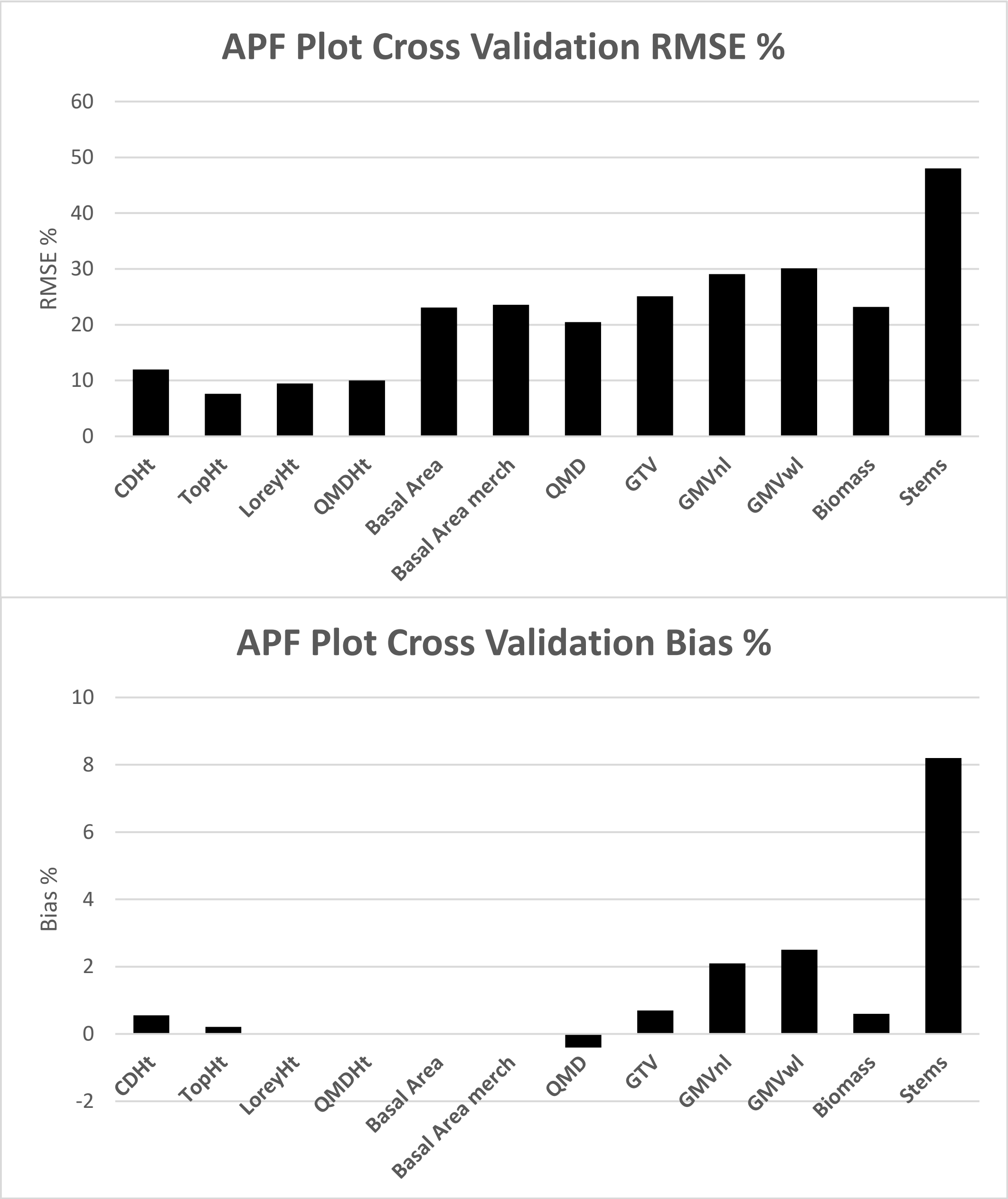


Figure 12 - RMSE (%) and Bias (%) for inventory attribute validation using OOB and a 10-fold Cross Validation.

Although the LiDAR models were not fit by forest type, the results can be presented in that manner to get a sense at the pixel scale how a model is performing overall. Figure 13 provides CV comparisons of RMSE (%) by FU and by inventory attribute. **Note, the number of plots by forest type varies and the results should be viewed in that light.** In some cases, the number of plots per FU was small or difficult to split out by a FU query with the information in hand. Accordingly, all PWUS3 and PWUS2 were combined to PWUS. HDSEL, HDUS and ORUS were combined to HDSEL_US. MWCC and MWUS were combined to MW. Appendix B provides a tabular summary of CV plot level predictions by forest types on the APF forest.

LiDAR Prediction Raster Surface Adjustments

Predicted raster products were modified to align pixel predictions with the limitations of the calibration plot network (DBH > 7.1 cm). Table 10 identifies the 99th percentile LiDAR height that was used as a threshold. Pixels with a Zq99 < 5m were not expected to have trees with DBH ≥ 7.1 cm. Pixels with a Zq99 < 9m were not expected to have merchantable sized trees.

Table 10 - Adjustments to LiDAR raster predictions based on zq99 thresholds.

Raster Surface	**Zq99 Threshold**	**Adjustment of Raster Predictions**
CDHT	5 m	CDHT predictions replaced with zq99 value where zq99 < 5 m
TOPHT	5 m	TopHt predictions set to **NULL** where zq99 < 5 m
LoreyHt	5 m	LoreyHt predictions set to **NULL** where zq99 < 5 m
Basal Area	5 m	Basal Area predictions set to 0 where zq99 < 5 m
QMD	5 m	QMD predictions set to **NULL** where zq99 < 5 m
GTV	5 m	GTV predictions set to 0 where zq99 < 5 m
Biomass	5 m	Biomass predictions set to 0 where zq99 < 5 m
Stems	5 m	Stems calculation set to 0 where zq99 < 5 m
Basal Area merch	9m	BA_merch predictions set to 0 where zq99 < 9m
BA_SmPoles	9 m	BA_SmPoles predictions set to 0 where zq99 < 9m
BA_LgPoles	9 m	BA_LgPoles predictions set to 0 where zq99 < 9 m
BA_SmSaw	9 m	BA_SmSaw predictions set to 0 where zq99 < 9 m
BA_LgSaw	9 m	BA_LgSaw predictions set to 0 where zq99 < 9 m
GMV_NL	9 m	GMV_NL predictions set to 0 where zq99 < 9 m
GMV_WL	9 m	GMV_WL predictions set to 0 where zq99 < 9 m
GMV_Util	9 m	GMV_SFL predictions set to 0 where zq99 < 9 m
GMV_SmPoles	9 m	GMV_NL_SmPoles predictions set to 0 where zq99 < 9 m
GMV_LgPoles	9 m	GMV_NL_LgPoles predictions set to 0 where zq99 < 9 m
GMV_SmSaw	9 m	GMV_NL_SmSaw predictions set to 0 where zq99 < 9 m
GMV_LgSaw	9 m	GMV_NL_LgSaw predictions set to 0 where zq99 < 9 m

The LiDAR derived CDHT raster for the APF is provided in (Figure 14). Additional examples of derived inventory raster outputs are provided in Appendix C.

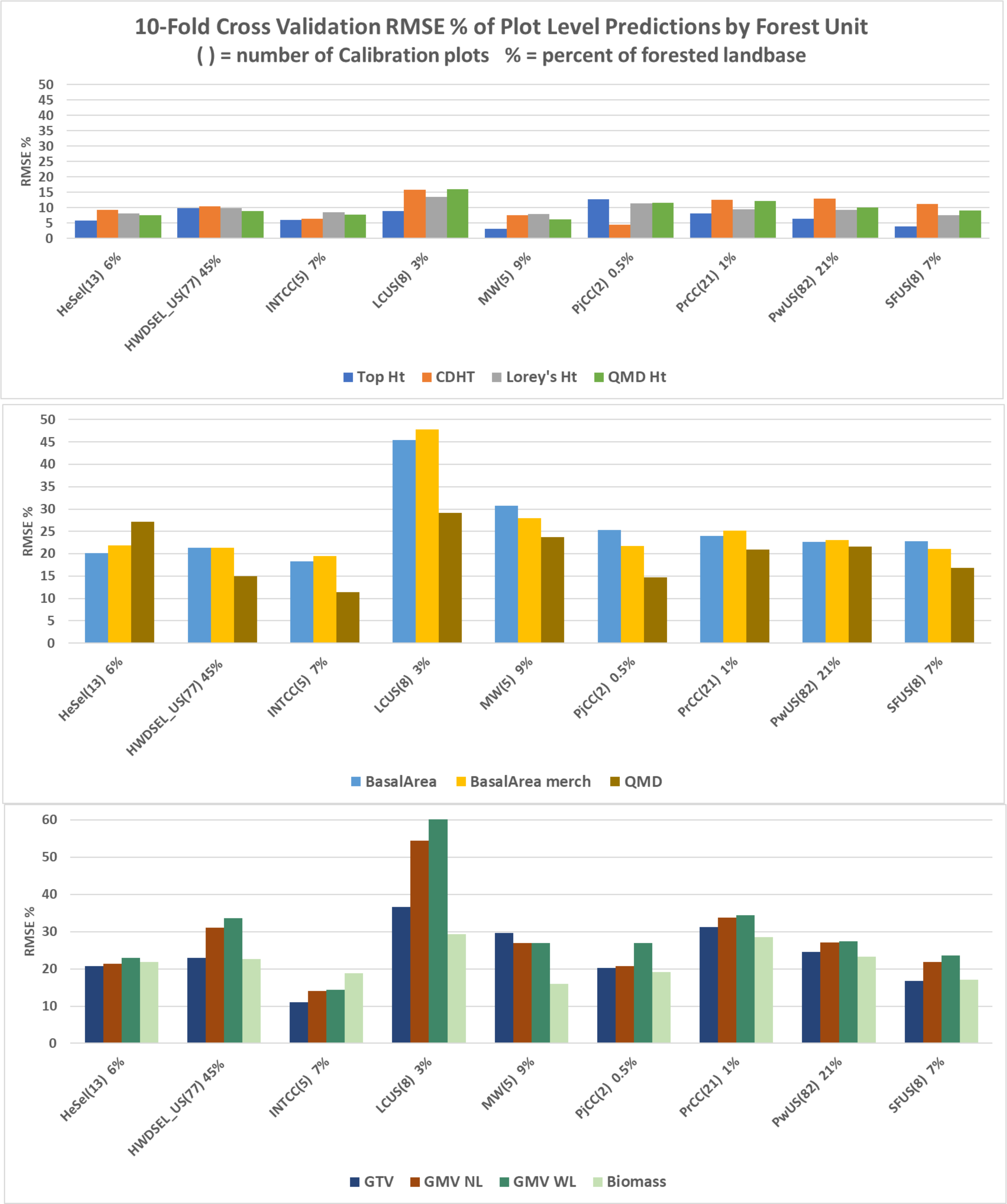


Figure 13 - 10-Fold cross validation RMSE (%) results of plot level predictions by AFA Forest Unit. Included are the number of calibration plots (in brackets) and % of forested landscape area the FU occupies.

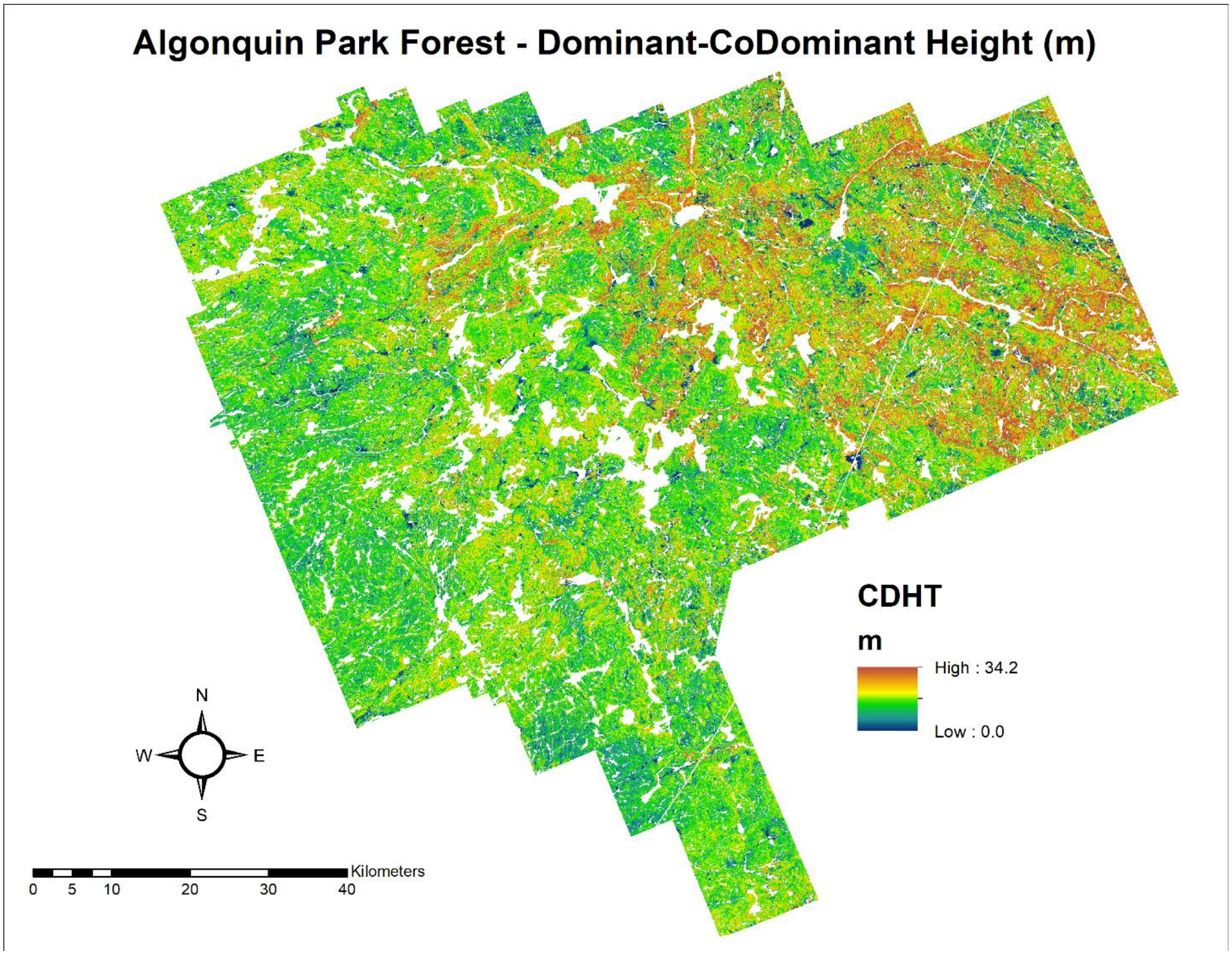


Figure 14 - LiDAR derived APF Dominant/CoDominant Height raster

Stand Level Validation

Most forest management decisions are not made at a raster pixel (20 m x 20 m) scale. Usually, decisions are made on an aggregation of pixels within a forest stand or harvest block. Eighteen forest stands that had been cruised by Algonquin Forest Authority (AFA) staff were available to provide a measure of model performance at the scale decisions are usually made. The eighteen stands were linked to another ongoing KKTD study looking at the automation of vertical structure characterization and as such, were intentionally chosen to represent a range for forest types and vertical structures. As a result, these validation stands may not represent the most common conditions on the APF forest.

Validation Sampling

A minimum of 10 stations spaced on a 75m or 100m grid covering the entire polygon was targeted depending on the stand size and shape. Ideally this would be about 1 plot/ha on a 100m x 100m grid. The stand polygons were also buffered by –20m to ensure that plot centres are at least 20m from a stand boundary (Figure 15). At each station, a BAF2 prism was used to determine “in” trees > 7cm. Each “in” tree was assessed for species, dbh, crown status (superstory, overstory, understory) and

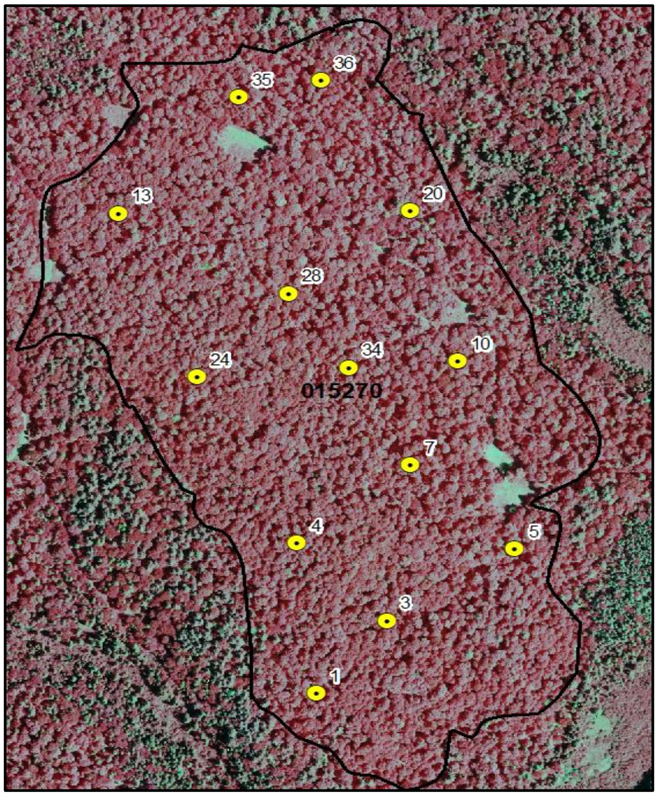


Figure 15 - Example of sampling stations established in Polygon 015270

measured for height. Some stations had only every other tree measured for height if the prism identified a high number of trees. Table 11 provides a description of the 18 stands cruised.

Validation Results

Figure 16 graphically displays the average validation stand prediction results (N=18). Figure 17 provides a more detailed comparison of individual validation stands observations and predictions for selected inventory attribute using a 1:1 line to represent agreement.

RMSE and Bias results for the 18 cruised polygons are presented in Table 12. Seven key inventory attributes were compared. Figure 18 provides a comparison of observed and predicted attributes polygon mean (with standard error).

Two stands (074133-LCUS and 089916-PRCC) were identified as outliers. Very few calibration plots sampled those conditions and as a result the stand level predictions are poor. A total of 9 LCUS calibration plots were measured. Of those 9, 6 were Sb and 3 were Cedar leading. Only 1 of the 21 PRCC FU calibration plots were in a plantation, and as a result, was underestimated. Both PRCC and LCUS represent a total of ~2% of the forested APF area. Table 13 presents the overall RMSE and BIAS statistics with those 2 stands removed.

Table 11 - Description of validation stands and number of BAF2 stations sampled along with a pseudo sampling intensity (based on 0.04 ha plots)

Polygon	Cruised Species Composition	Reporting FU	Area (ha)	Stations	~ Sampling Intensity
15270	Mh 84 By8 Be4 He2 Ce1 Sw1	HDSEL_US	32.7	13	2%
25404	Be 36 Mh33 By16 He5 Sw4 Cb4 Mr1 Sr1	HDSEL_US	17.9	13	3%
26147	He 24 Mh19 By15 Sb11 Bf9 Mr9 Sw8 Ce3 Be1	HeSel	28.1	13	2%
26847	Mh 72 By8 Be4 Mr4 Iw3 Sw3 Bf2 He2 Cb1 Ab1	HDSEL_US	23.6	12	2%
27172	Mh 44 By17 Be14 He11 Iw4 Mr3 Aw2 Ab2 Bw2	HDSEL_US	21.3	12	2%
68460	Pw 61 Sw27 Mr8 Bf1 Ce1 Po1 Pr1	PwUS	9.1	11	5%
70809	Pw 68 Mr15 Po7 Pr3 Sw3 Or2 Be1 Bw1	PwUS	12.8	11	3%
74025	Pw 52 Mr13 Po12 Pr8 Or6 Sw5 Bw2 Bf1 Mh1	PwUS	16.7	11	3%
74133	Ce 68 Sb19 La9 Ab1 Pw1 Bf1 Bw1	LCUS	16.9	11	3%
74329	Pw 51 Po23 Sw7 Pr5 Or5 Mr3 Bf3 Mh2 Be1	PwUS	22.5	12	2%
74445	Po 28 Pw24 Mh11 Or10 Mr9 Sw6 Bf5 Ab3 Be2	MW	13.5	11	3%
74732	Pw 47 Bf24 Sw9 Mr6 Or6 Bw2 Mh2 Po2 Pr2	PwUS	17.0	11	3%
88543	Pr 63 Po19 Or6 Pj6 Pw6	PrCC	23.9	9	2%
89916	Pr 68 Pj31 Pw1	PrCC	14.6	11	3%
90147	Po 26 Pr24 Pw14 Bf13 Pj8 Mr7 Or3 Sw3 Bw2	PwUS	16.5	11	3%
90863	Be 48 Pw20 Or9 Mh8 Po8 By2 Iw2 Mr2 Pr1	HDSEL_US	14.2	11	3%
92778	Pr 36 Pw27 Mr8 Bf8 Ce6 Po6 Ab5 Sw2 Bw1 By1	PwUS	31.4	11	1%
100453	Pw 50 Or32 Mr8 Be4 Pr3 Bw1 Iw1 Po1	PwUS	15.6	11	3%
Average		**All**	**19.4**	**11.4**	**2.4%**

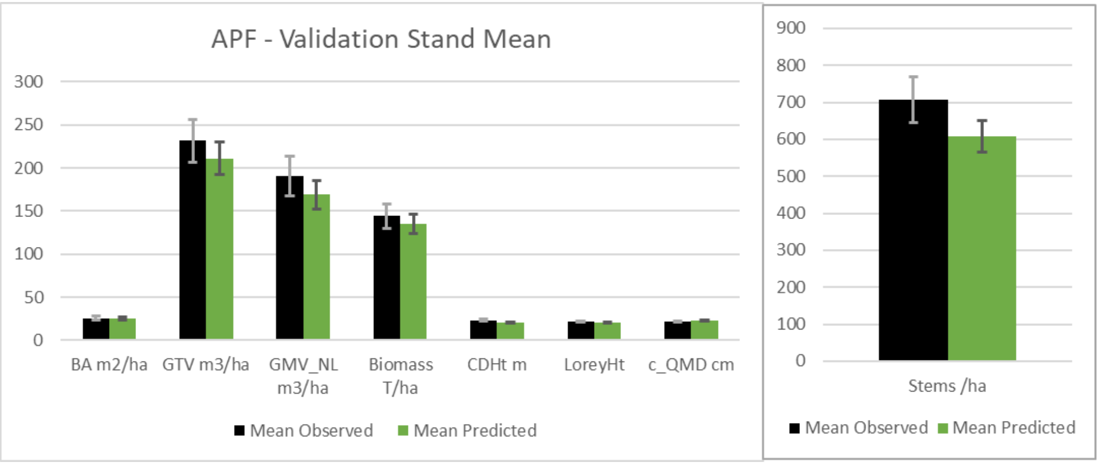


Figure 16 - Validation stand mean stand observed conditions and predictions. Standard error of observed and prediction is provided.

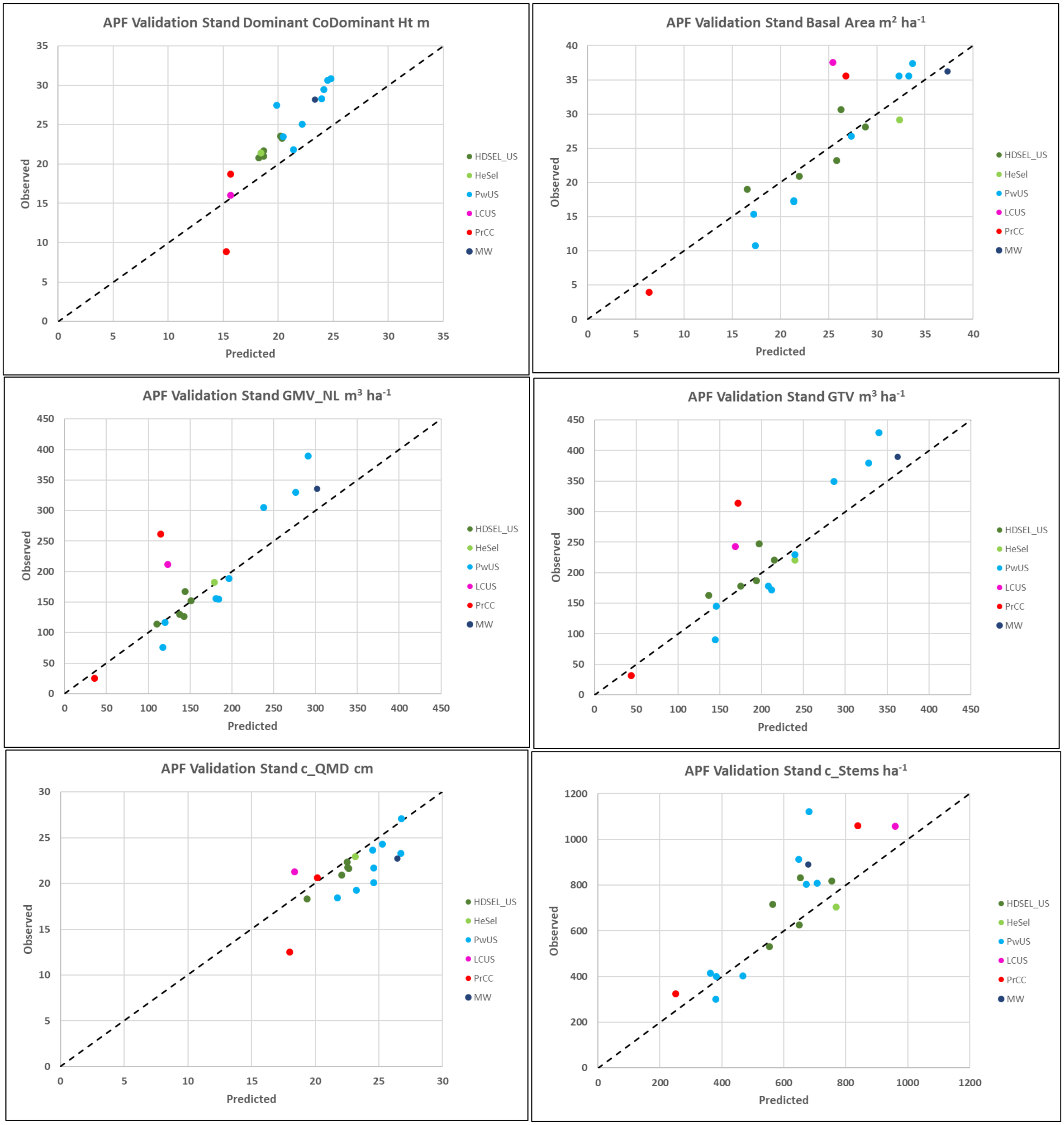


Figure 17 - Validation comparison by stand of inventory observed and predicted attributes.

Table 12 - Validation RMSE and Bias results for the 18 cruised polygons.

	CDHT	BA	QMD[4]	GTV	GMVnl	Biomass	Stems
RMSE	4.2	4.6	2.6	53.1	54.0	23.3	164.3
RMSE %	18.1%	18.1%	12.3%	22.9%	28.3%	16.1%	23.2%
MeanBias	3.1	0.5	-1.6	20.6	21.8	8.9	99.1
Bias %	13.2%	2.1%	-7.6%	8.9%	11.4%	6.2%	14.0%
N	18	18	18	18	18	18	18

Table 13 - Validation RMSE and Bias results for 16 of the 18 cruised polygons
(2 under sampled outliers removed)

	CDHT	BA	QMD[5]	GTV	GMVnl	Biomass	Stems
RMSE	4.4	3.2	2.7	39.4	37.7	21.3	163.2
RMSE %	18.4%	13.1%	12.6%	17.4%	20.4%	14.7%	24.6%
MeanBias	3.2	-0.7	-2.0	9.7	9.7	6.1	91.3
Bias %	13.4%	-3.0%	-9.6%	4.3%	5.2%	4.2%	13.7%
N	16	16	16	16	16	16	16

An additional comparison of average predicted BA and GMV by the typical Great Lakes St Lawrence four management size classes are presented for the stands of the HDSEL_US and PWUS Forest Units (Figure 18). Individual validation stand graphs are presented in Figure 19 - Figure 22 - Observed and predicted basal area and gross merchantable volume by size class for the HeSel, LCUS, PRCC and MW validation Stands.

T2 Inventory Updating

LiDAR Raster updating

The T2 inventory polygon update began with the Operational Planning Inventory (OPI) provided by the Algonquin Park Authority. This was updated to 2021. The T1 polygon boundaries were used and mean raster values by T1 polygon are calculated and provided for the following attributes:

- **Heights** - TopHt, CDHT, LoreyHt
- Basal Area,
- Stems

---

[4] QMD = Calculated QMD from predicted stand basal area and predicted stems.

[5] QMD = Calculated QMD from predicted stand basal area and predicted stems.

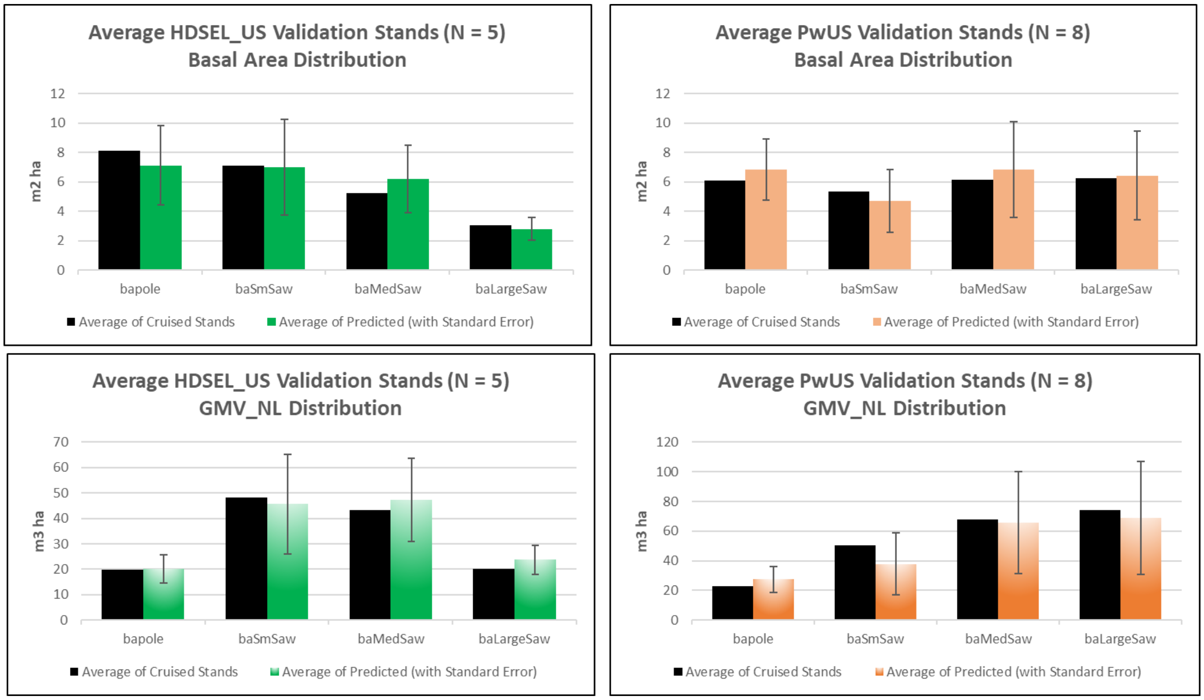


Figure 18 - Observed and predicted average basal area and average gross merchantable volume by size class for validation stands by Forest Unit. Standard error of predictions are indicated.

- **Volumes** – GMV_NL, GMV_WL, GMV_NL
- **By Size Class** – Basal Area, GMV_NL
- QMD is calculated for each polygon based on mean stand Basal Area and Stems

Stand Level GMV_NL Quantiles

To provide a measure of stand level volume variation, the 15th and 85th quantiles of gross merchantable (NL) volume were also provided. They are provided as:

- GMV_NL_15 and GMV_NL_85.

An example of a raster prediction for GMV_NL and the corresponding mean polygon information are presented in Figure 23. Note how within stand variation of GMV_NL predictions are lost when the rasters are summarized for their mean value by polygon. The addition of Q15 and Q85 values allows the users of the inventory to also know that 70% of the GMV_NL pixels are between the Q15 and Q85 values for the polygon.

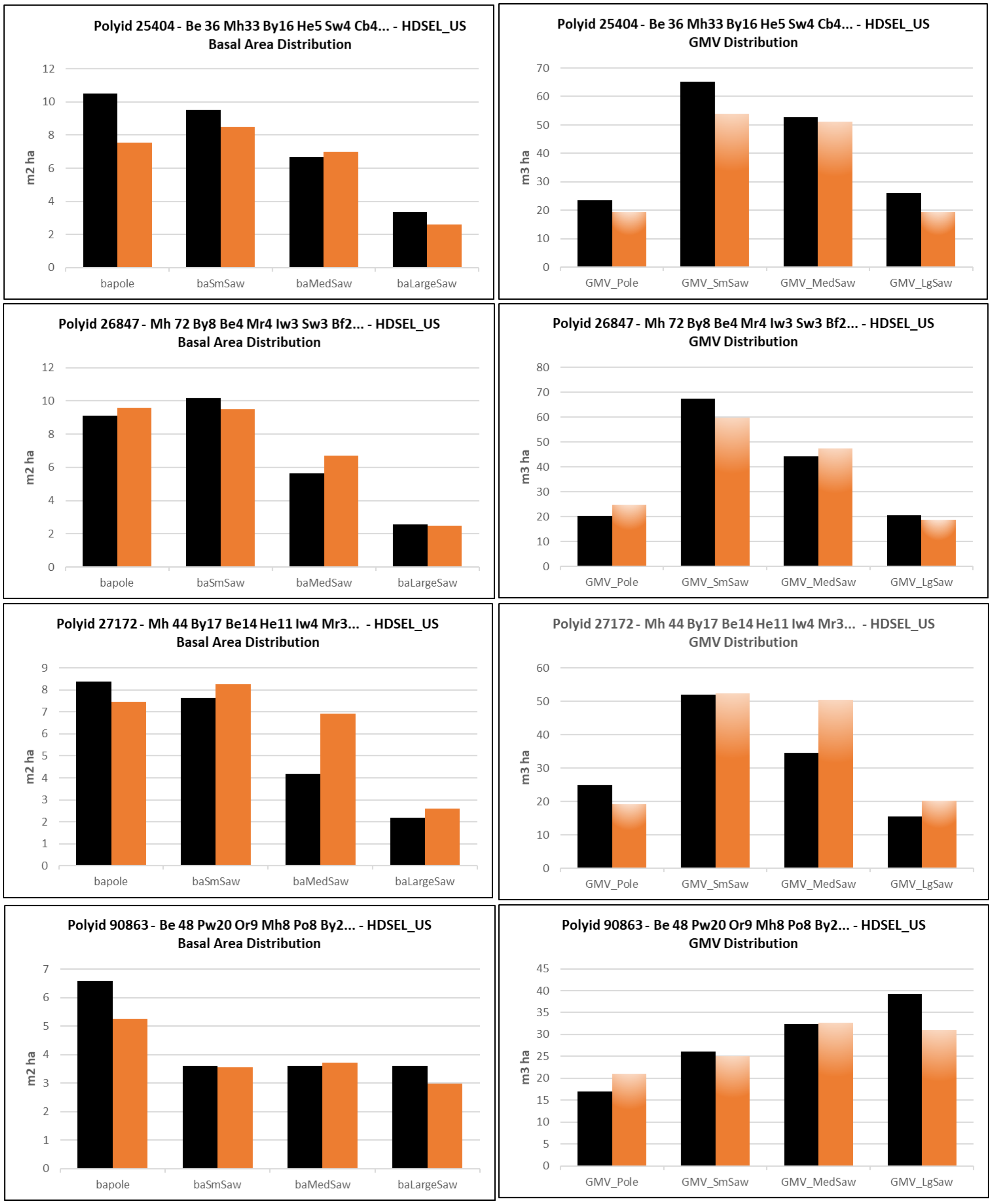


Figure 19 - Observed and predicted basal area and gross merchantable volume by size class for the HDSEL_US validation Stands.

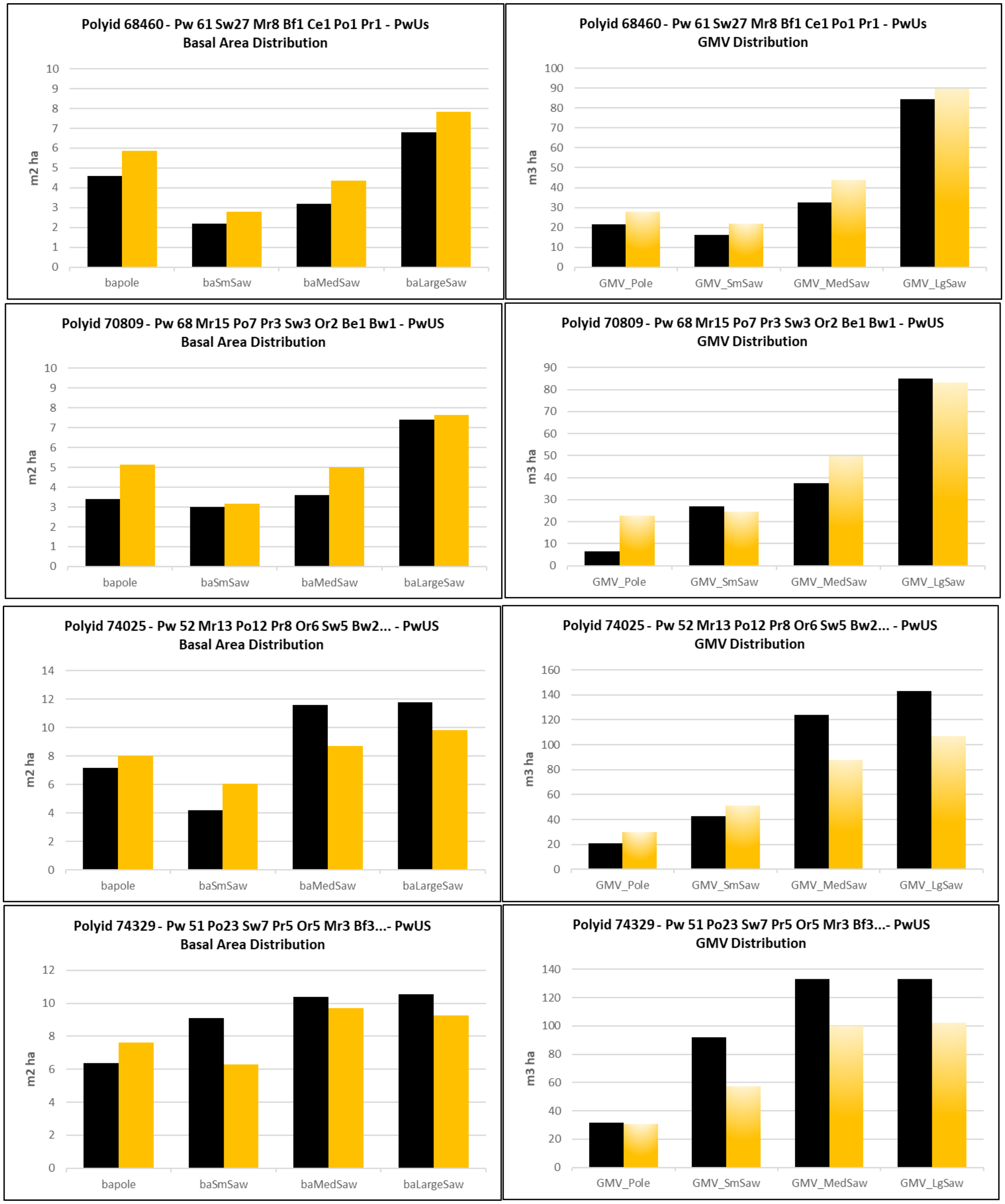


Figure 20 - Observed and predicted basal area and gross merchantable volume by size class for the PwUS validation Stands.

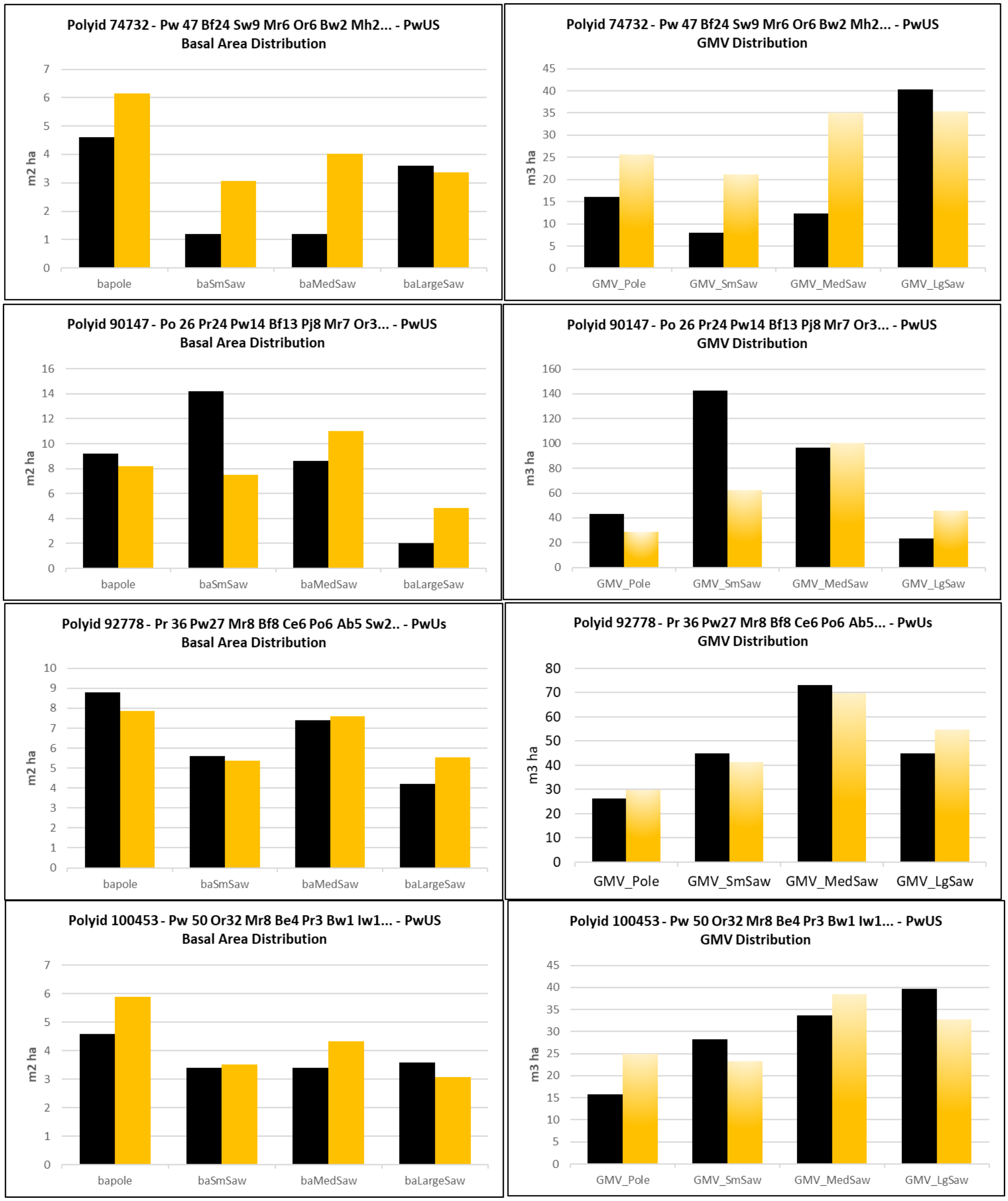
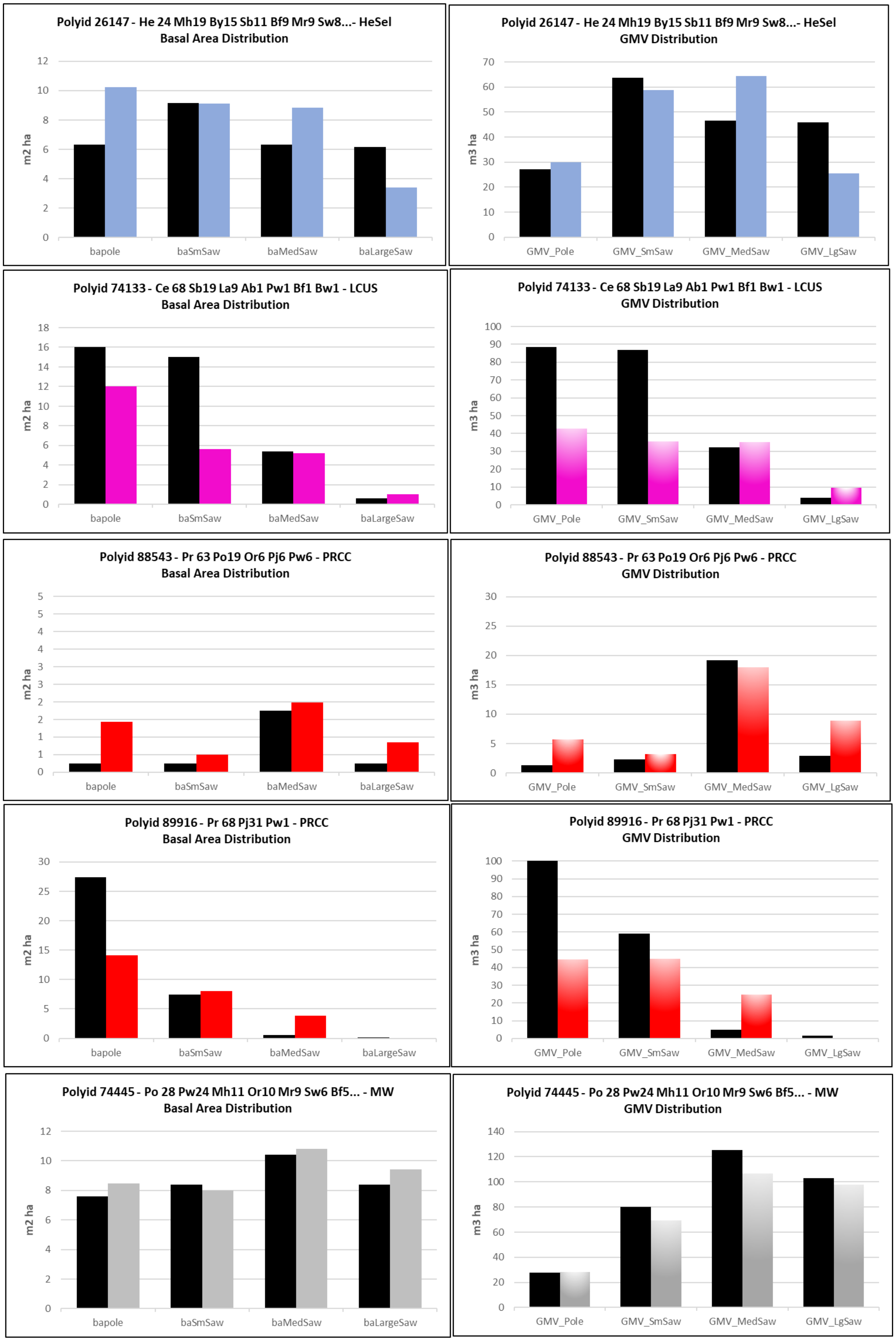


Figure 21 - Observed and predicted basal area and gross merchantable volume by size class for the PwUS validation Stands (continued).


Polyid 26147 - He 24 Mh19 By15 Sb11 Bf9 Mr9 Sw8...- HeSel
Basal Area Distribution

Polyid 26147 - He 24 Mh19 By15 Sb11 Bf9 Mr9 Sw8...- HeSel
GMV Distribution

Polyid 74133 - Ce 68 Sb19 La9 Ab1 Pw1 Bf1 Bw1 - LCUS
Basal Area Distribution

Polyid 74133 - Ce 68 Sb19 La9 Ab1 Pw1 Bf1 Bw1 - LCUS
GMV Distribution

Polyid 88543 - Pr 63 Po19 Or6 Pj6 Pw6 - PRCC
Basal Area Distribution

Polyid 88543 - Pr 63 Po19 Or6 Pj6 Pw6 - PRCC
GMV Distribution

Polyid 89916 - Pr 68 Pj31 Pw1 - PRCC
Basal Area Distribution

Polyid 89916 - Pr 68 Pj31 Pw1 - PRCC
GMV Distribution

Polyid 74445 - Po 28 Pw24 Mh11 Or10 Mr9 Sw6 Bf5... - MW
Basal Area Distribution

Polyid 74445 - Po 28 Pw24 Mh11 Or10 Mr9 Sw6 Bf5... - MW
GMV Distribution

Figure 22 - Observed and predicted basal area and gross merchantable volume by size class for the HeSel, LCUS, PRCC and MW validation Stands.

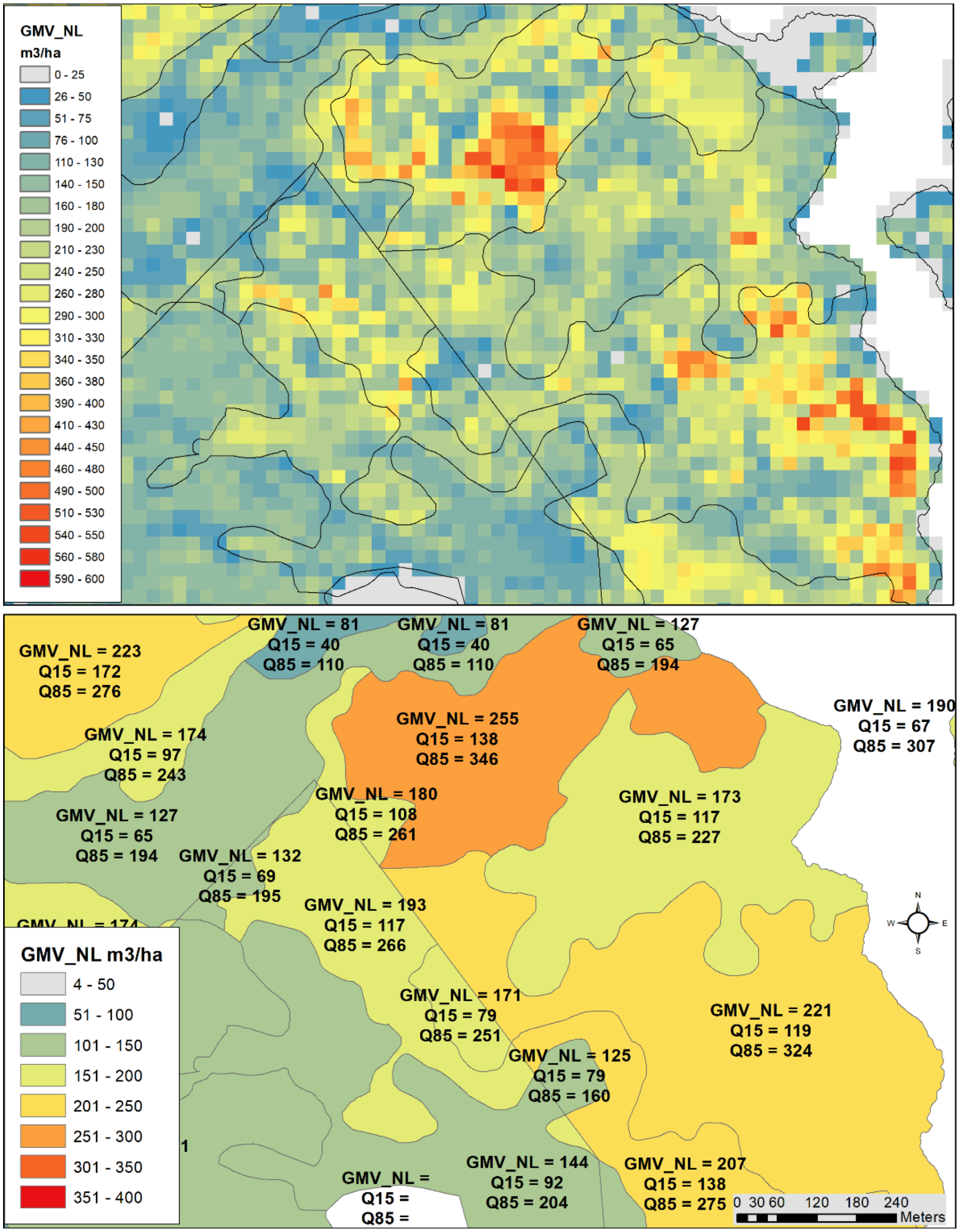


Figure 23 - Example of a GMV_NL raster prediction and mean T2 Polygon summary. Mean GMV_NL is labeled in each polygon along with the 15th and 85th quantile value.

Utility Pole Volume

A request was made by the AFA for an additional volume calculations for potential utility poles. Volumes were calculated for stands containing red pine and jack pine based on T1 species composition information. No adjustments were made for any potential stem quality issues (i.e., knot size, distance between knots, etc.) assessment. Table 14 presents the specifications used.

Table 14 - AFA specific volumes adjustments by T1 polygon species composition for Utility Poles

Volume	**Stump height**	**Diameter at 30cm + 6' (2.13m)**	**Top Diameter**	**Minimum Length**	**Multiply by**
GMV_UtilPoles (m^3 ha^{-1})	30 cm	30cm	14cm (5")	10.668 m (35')	(Pr% + Pj %)/100

Additional Attributes Calculated for T2 Inventory

To provide further value to the T2 update of the inventory, polygon-based summation (mean) of LiDAR attributes, were used in conjunction with T1 polygon age and species composition to calculate the following additional T2 inventory attributes:

- Site Index
- Stocking
- Cull Fraction
- Net Merchantable Volume (NMV).

Refer to Table 15 for a list of attributes and their source.

Table 15 - Additional T2 calculated inventory attributes and their source.

Attribute	**LiDAR Derived**	**Calculated**	**T1 Polygon Information**	**Literature Source**
Site Index	CDht		Age, Leading Species	Various (refer to Appendix D)
Stocking	Basal Area	Site Index	Age, Leading Species	Plonski 1974
Cull Fraction		VBAR, Site Index	Age, Species Composition	Basham 1991
Net Merchantable Volume	Basal Area , GMV_NL	Cull Fraction Species VBAR[6]	Age, species composition	

---

[6] Species vbar are calculated from a combination of calibration plots for the SFL and provincial monitoring plots

Site Index

Site index is calculated using the leading species from the T1 species composition and the age from the T1 inventory updated to 2021 and the predicted LiDAR CDht. **For polygons with p99 < 5m, SI and stocking are not estimated.**

Most SI equations use breast height age. For young stands, small change in age result in large changes in SI. The SI estimates for young ages are unstable (Figure 24). The inventory age, particularly for young stands, may come from supplementary information and may not correspond to the LiDAR heights. This issue is illustrated for the APF Forest.

Based on Figure 24, the SI for ages < 20 was set to missing and the SI for ages >= 20 was capped at 35m. Figure 24 identifies issues with the available set of SI curves. The trend of SI with age is likely partly an artefact of the SI curves and partly an issue of the ages for older polygons not corresponding to the height. For older stands, the age is likely the age since disturbance and the heights are likely from younger trees.

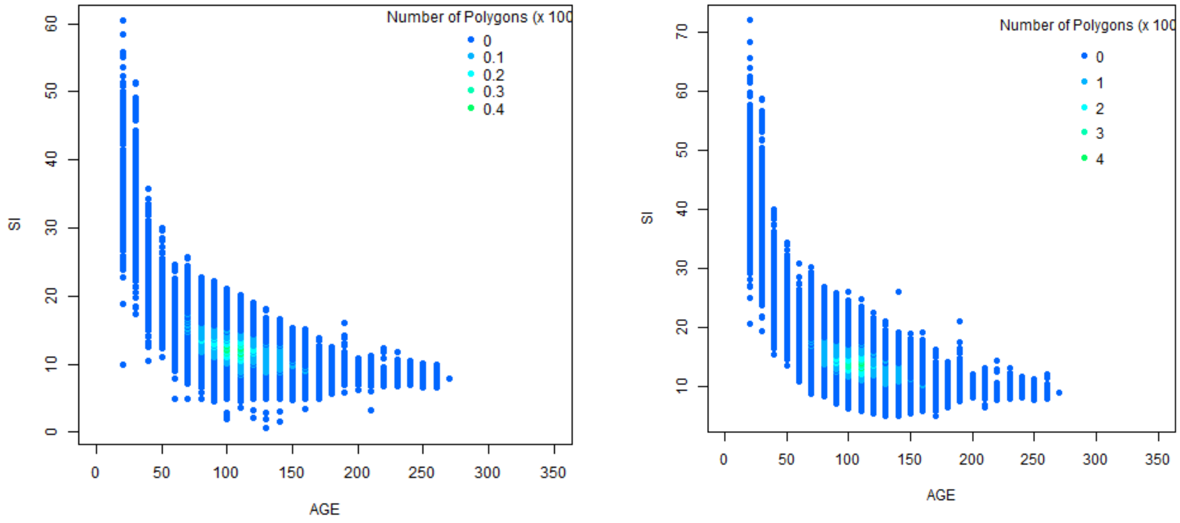


Figure 24 - Site index is plotted against age for ages 10+ and for ages 20+ for the APF. Note the minimum SI is set to 5m

Stocking

Stocking was calculated from predicted LiDAR basal area and the T1 polygon age and leading species. Stocking is in reference to Plonski's Normal Yield Table (Plonski 1974). Stocking is also a challenge for young stands. Stocking requires SI and SI was set to missing for stands < 20 years old so stocking is also not calculated when age is < 20. Stocking was capped at 2. Figure 25 provides a graphic of the number of APF polygons by stocking and age. Note that stands less than 20 years old are not presented.

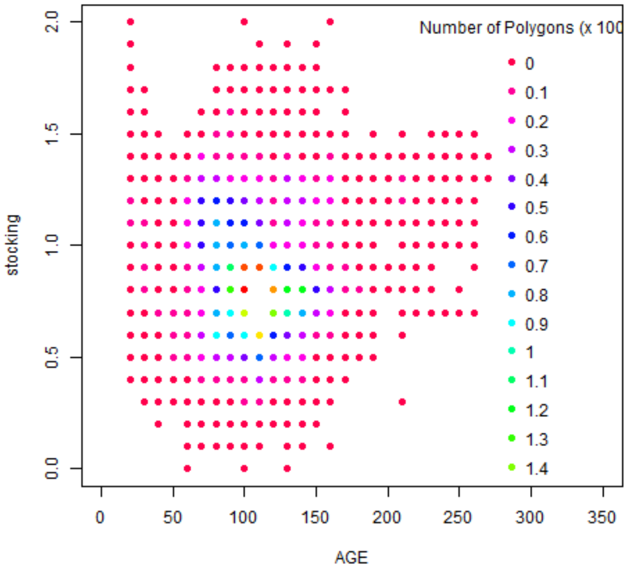


Figure 25 - Calculated Plonski stocking by polygon for the APF. Note: no stocking estimates for stands < 20 years old.

Cull and Cull Fraction

Cull as estimated following the procedure implemented in MIST. Gross merchantable volume is estimated without respect to species. However, Net merchantable volume (NMV) requires estimates of cull. Basham (1991) provides estimates of cull by species and age.

First, a cull model ((1)) was fit, by species, using published data (see Table 16). The model predicts the cull fraction increases as a sigmoidal function of age.

$$(1) \qquad \widehat{cull} = \left(1 - e^{-d_0 \cdot age}\right)^{d_1}$$

Where, $\widehat{cull}$ is the estimate of cull as a percentage of tree volume at a given age.

To apply this to GMV, the GMV by species was estimated by fitting a volume to basal area ratio (vbar) prediction model ((2)) by species using the provincial PSP/PGP database (gyPSPPGP_2021_10_04.bak).

$$(2) \qquad vbar = (v_0 + v_1 \cdot SI) \cdot \left(1 - e^{-v_2 \cdot age^{v_3}}\right)$$

Where, *vbar* is the volume to basal area ratio for a species, *SI* is the site index, *age* is the Plot age and v_0, v_1, v_2, and v_3 are coefficients.

The vbar estimate was used to estimate the relative GMV by species.

$$(3) \qquad mvol\ frac_i = \frac{species\ fraction_i \cdot vbar_i}{\sum species\ fraction_i \cdot vbar_i}$$

Table 16 - The sources for the cull estimates are given. The table references are from Basham (1991) except for red pine.

Species	Table	Comment
Hemlock	Table 7	
Sugar Maple	Table 13	
Yellow birch	Table 12	
Red pine		Source unknown. Basham (1991) reports an average of 1% for the 141-160 age class.
White pine	Table 1	
Cedar	Table 8	
White birch	Table 11	
Trembling aspen	Table 9	
Ironwood	Table 20	
Basswood	Table 16	
Balsam fir	Table 6	
White elm	Table 19	
Red oak	Table 18	
Black ash	Table 17	
Beech	Table 15	
Red maple	Table 14	
White spruce	Table 5	Note that the data for age 170 was taken from Table 6 of OMNR (1978)
Jack pine	Table 2	
Black spruce	Table 4	Note that data from ages 200+ were not used

Then the weighted cull estimate, all species combined, is estimated as follows.

$$(4) \qquad cull = \sum mvol\ frac_i \cdot spp\ cull\ est_i$$

Sample calculations are given in Table 17. An example of vbar estimates by age and species is presented in Figure 26.

Table 17 - Vbar and cull calculations are given for sample conditions. The age = 100 and SI = 20m. Poplar has a slightly higher vbar, giving slightly more weight to the poplar cull estimate.

	Spp	Vbar coefficient					Cull coefficient			Mvol	weighted
Spp	frac	V0	V1	V2	V3	Vbar	D0	D1	cull	frac	cull
Pj	0.8	2.36509	0.54016	0.018021	1.01063	11.2	-0.01264	8.3752	0.062	0.79	0.049
Po	0.2	2.99849	0.50008	0.006109	1.30665	11.9	-0.00521	1.4052	0.282	0.21	0.059
All											0.108

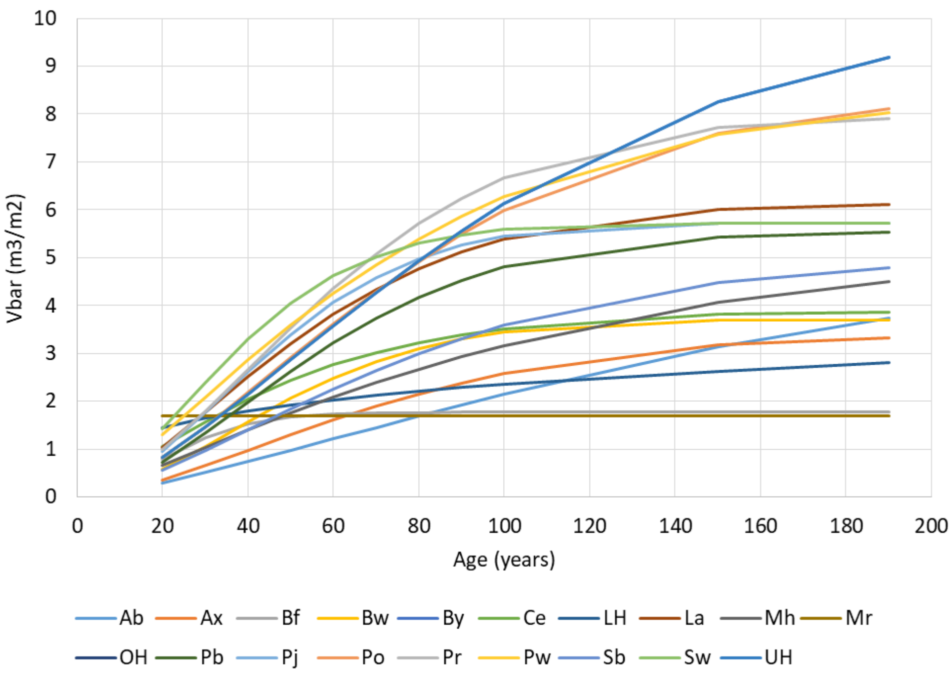


Figure 26 - The vbar estimates are given by age and species, for SI = 20

Net Merchantable Volume

For the T1 polygons, cull was estimated at using the T1 age and species composition.

Net merchantable volume (NMV) is calculated as the GMV minus cull.

$$(5) \qquad NMV = \mathrm{GMV} \cdot (1 - cull)$$

Constraint of T2 Inventory Updates

Only trees $\geq$ 7.1 cm were measured on all the calibration plots. As a result, shorter (and young) stands do not have any measured trees to support defensible LiDAR predictions. **Stands < 20 years are not being updated with LiDAR derived predictions.** In addition, different polygon CDHT thresholds were used to constrain provided inventory attributes (Table 18). Crown Closure (CC2m) was retained for all stands. If stand CDHT <5 m, zq99 replaced estimated CDHT value.

Table 18 - T2 polygon inventory attributes and instituted constraints for all stands with age ≥ 20 years

	Polygon CDHT <5m	**Polygon 5m > CDHT <9m**	**Polygon CDHT >9m**
CC2m	🟩	🟩	🟩
TOPHT	NULL	🟩	🟩
CDHT	Zq99	🟩	🟩
LoreyHT	NULL	🟩	🟩
BA	0	🟩	🟩
BAmerch	0	0	🟩
Stems	0	🟩	🟩
QMD	NULL	🟩	🟩
GTV	0	0	🟩
GMV_NL	0	0	🟩
GMV_WL	0	0	🟩
GMV_Util	0	0	🟩
NMV_NL	0	0	🟩
NMV_WL	0	0	🟩
NMV_Util	0	0	🟩
Biomass	0	0	🟩
BA_Poles	0	0	🟩
BA_SmSaw	0	0	🟩
BA_MedSaw	0	0	🟩
BA_LgSaw	0	0	🟩
GMV_Poles	0	0	🟩
GMV_SmSaw	0	0	🟩
GMV_MedSaw	0	0	🟩
GMV_LgSaw	0	0	🟩
Site Index	NULL	🟩	🟩
Stocking	NULL	🟩	🟩
Cull Fraction	NULL	NULL	🟩

Discussion

Calibration Plot Data Quality

Concern about calibration plot data quality has been expressed earlier. Although an attempt was undertaken to adjust tree heights due to observed field measurement errors, other potential sources of measurement error (i.e., target plot location not being achieved, DBH measurement, missed trees, etc.) could not be evaluated or adjusted for in compilation. The assumption had to be made that the other tree attributes were done well. However, a level of concern exists on the unknown impact of the field plot data quality and the results reported. **FRI quality is directly linked to field plot quality. Audits of field plots should occur as soon as data collection begins in order to identify and correct any data collection issues as quickly as possible.**

Plot Level Model Validation (CV)

Overall, the APF pixel level predictions are similar to those reported in other studies in Ontario. White et al. (2021) reported results for similar forest types and SPL. In their work, larger calibration plots (625m2) were used and a lower Dbh measurement threshold (2.5cm vs. 7.1cm used in this study) were chosen. They reported RF Out-of-Bag (OOB) RMSE errors which are comparable to the CV RMSE error statistics reported here. White et al. (2021) reported RMSE of 15% for CDHT for all forest types. We report a similar RMSE of 12%. For volumes, White et al. (2021) reported 25.4% and 29.6% for basal area and GTV. This study reported 23% and 25%. In White et al.'s (2021) work they reported a similar total above ground biomass RMSE of 25% vs our reported 23%.

Where possible to broadly compare forest units (criteria differ but leading species is similar) we found the following. White et al. (2021) reported Tolerant Hardwood stand RMSEs for CDHT, Basal Area, GTV and biomass of 8%, 31%, 38% and 27% respectively. This study found similar or better results (likely partly due to a higher minimum Dbh threshold) in the HDSEL_US FU of 11%, 21%, 23% and 23%. White pine was reported as managed/natural stands in the work of White et al. (2021). Thy reported 12%/19%, 20%/26%, 26%/ 27%, and 21%/26% for CDHT, Basal Area, GTV and biomass respectively. This study reported on a combined managed and natural PWUS stand grouping at 13%, 23%, 25% and 23.2%. Interestingly, the results of White et al. (2021) were from a single RF model, while the ones reported here were from a stratified (2) set of RF models; (1) tolerant hardwoods >50% and "conifer + Intolerant hardwoods. The RMSEs at the plot/pixel scale are very similar.

Stand/Block Level Model Validation

As has been demonstrated in other published LiDAR inventory projects (White et al. 2021), validation of LiDAR predictions is more appropriately evaluated at the scale at which most management decisions are based. In Ontario, this is generally the harvest block or stand level.

Although a validation sample of 18 stands is not large, it can provide a sense of expected model performance for an inventory at that scale. It should be mentioned that the field sampling of the validation stands is also an estimate. Calculating the sampling intensity is challenging for variable radius plots but if the plots were 0.04ha, the approximate sampling intensity was 2.4% (range of 1% -5%) or 1.7 ha being sampled by each plot (Table 11). Overall, RMSEs (except for CDht) for the stand level predictions were less than the plot level RMSEs (Figure 27), representing an improvement, for most inventory attributes. CDht error is not impacted greatly by scale, but more from stand structure. The validation stands, chosen to sample a wide range of forest vertical structure for another project and do not necessarily represent the population condition. Basal Area, GTV, GMV_NL and Biomass exhibit the expected trend of less error as noted in the study of White et al. (2021). Figure 27 includes two RMSE summaries: one with 18 validation stands and one with 16 stands. The 16-stand summation has dropped the PRCC (89916) and LCUS (74133) stands. These two validation stands were predicted poorly and identified the fact that very few calibration plots were established to support modeling of those conditions. The fact that these FUs represent a small fraction (~2%) of the APF forested landbase is the reason they were not initially sampled heavily. Since they do represent such a small fraction of the forest, these two stands were excluded, and a comparison provided.

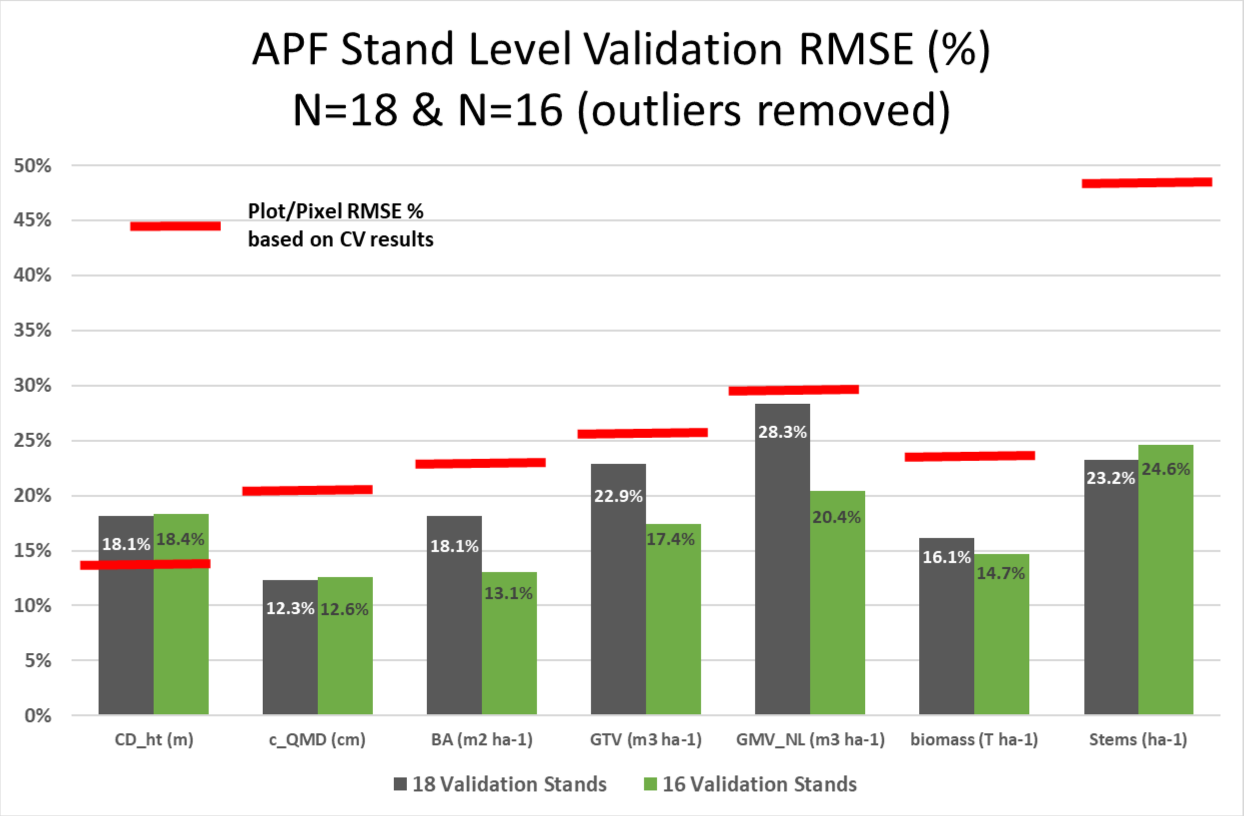


Figure 27 - Comparison of Stand level validation RMSE with Cross Validation at the plot/pixel scale.

Challenges with aligning and summarizing vector data and raster data

T1 information in the inventory is polygon based, including species composition and forest classification (forest vs. non forest). LiDAR derived information in pixel based. An issue arises when aligning the two sources of information. T1 polygon boundaries do not follow raster edges and, as a result, bisect pixels. Since, currently in Ontario, forests are managed at the polygon level, approaches to summarizing raster values within polygons was explored.

Two main approaches investigated for operational inventory production are discussed here.

1. Centroid based zonal summation
2. Area-weighted based summation

Some tools provide polygon summaries from raster layers by only selecting raster pixels with centroids within the polygon. This can result in edge raster pixels being excluded if they border linear features such as roads/rivers, water bodies (Figure 28) and the centroid is in that feature. In addition, where polygons bisect raster pixels, only one polygon is assigned the value of the raster pixel (Figure 29). The issue is particularly problematic for small polygons (< 1 ha).

In an area-weighted approach, the pixel's contribution to a polygon is weighted by the portion of the pixel falling within a polygon. This means a pixel can potentially be part of more than one polygon.

Pixels that fall entirely within the polygon will have a weight of one. If half of a pixels falls within a polygon, the pixel will be given a weight of 0.5.

The decision to implement the area-weighted approach to generating T2 polygon raster summaries was selected. This method ensured that each polygon benefits from an appropriately weighted proportion of each raster pixel covered by the polygon.

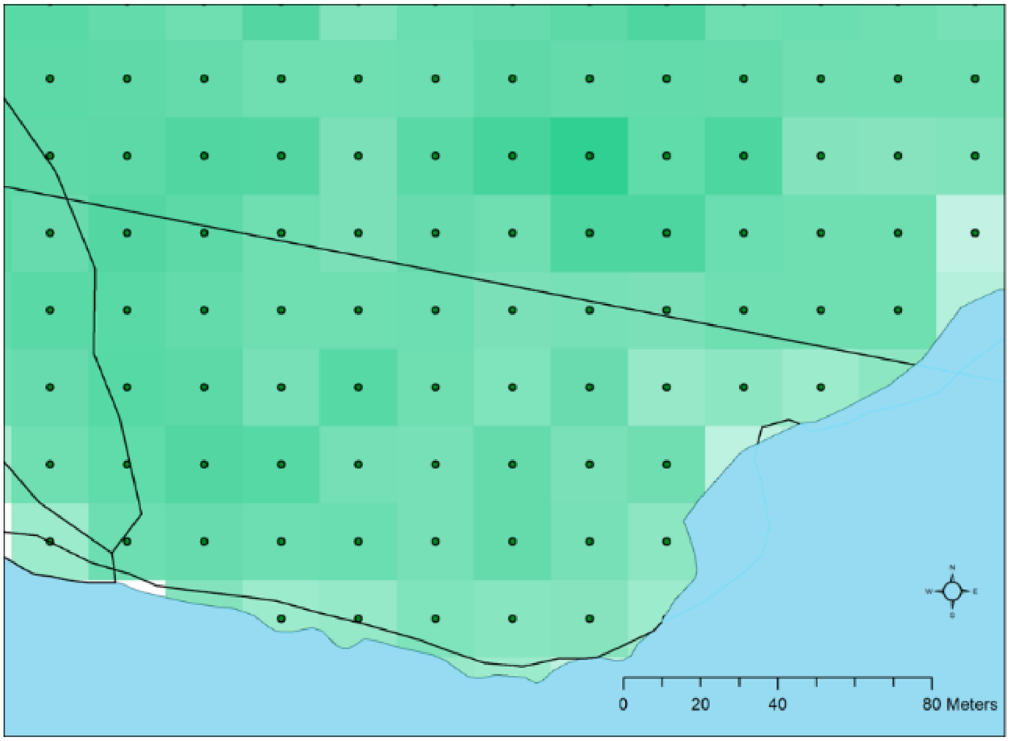


Figure 28 - Example of centroid selection or raster cells excluding raster values for narrow polygons along waterbodies.

Literature Cited

Basham, J.T. 1991. Stem decay in living trees in Ontario's forests: A users' compendium and guide. CFS Inf. Rep. O-X-408.

Buckmann, R.E., B. Bishaw, T.J. Hanson, and F.A. Benford. 2006. Growth and yield of red pine in the lake states. USDA FS Gen. Techn. Rep. NC-271, 114p.

Buda, N. and J. Wang. 2006. Suitability of two methods of evaluating site quality for sugar maple in central Ontario. For. Chron. 82:733 – 744.

Figure 29 - Example of a raster pixel being bisected into 4 by polygon boundaries with only one polygon including the centroid value.

Carmean, W.H., J. T. Hahn, and R.D. Jacobs. 1989. Site index curves for forest tree species in the eastern United States. USDA Forest Service North Central For. Expt. Stn. Gen Tech. Rep. NC-128. 142p,

Carmean, W.H. 1996. Site-quality evaluation, site-quality maintenance, and site-specific management for forest land in northwest Ontario. Ontario Ministry of Natural Resources, NWST Technical Report TR-105, Thunder Bay, ON. 121p.

Carmean, W. and J. Thrower. 1995. Early height growth and site index for planted red pine in north central Ontario. N. J. Appl. For. 12:23 – 29.

Carmean, W., G. Niznowski & G. Hazenberg. 2001. Polymorphic site index curves for jack pine in northern Ontario. For. Chron. 77:141 – 150.

Carmean, W., G. Hazenberg and K. Deschamps. 2006. Polymorphic site index curves for black spruce and trembling aspen in northwest Ontario. For. Chron.

Carmean, W., J. Hahn., R. McRoberts, D. Kaisershot. 2013. Site index comparisons for forest species in the upper Great Lakes area of the United States and Canada. USDA FS Gen Techn. Rep. NRS – 113.

Goelz, J. and T. Burk. 1998. Long-term trends in height growth of jack pine in north central Ontario. For .Sci. 44:158 – 164.

Guo, J. and J. Wang. 2006. Comparison of height growth and growth intercept models of jack pine plantations and natural stands in northern Ontario. Can. J. For. Res. 36: 2179–2188.

Lambert, M.–C., Ung, C.–H and Raulier, F. (2005). Canadian national aboveground biomass equations. Canadian Journal of Forest Research, 35: 1996–2018.

LAStools, "Efficient LiDAR Processing Software" (commercial) obtained from http://rapidlasso.com/LAStools

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

Parresol, B.R. and J.S. Vissage. 1998. White pine site index for the southern forest survey. USDA FS Res. Pap. SRS-10, 8p.

Penner, M., D.G. Pitt, and M. E. Woods. (2013). Parametric vs nonparametric LiDAR models for operational forest inventory in boreal Ontario. Can. J. Rem Sensing, Vol. 39, No. 5., pp. 426-443.

Plonski, W. L. (1974) Normal yield tables (metric) for major forest species of Ontario.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R–project.org/.

Roussel, J.R , Auty D, Coops NC, Tompalski P, Goodbody TR, Meador AS, Bourdon J, de Boissieu F, Achim A (2020). “lidR: An R package for analysis of Airborne Laser Scanning (ALS) data.” Remote Sensing of Environment, 251, 112061. ISSN 0034–4257,doi: 10.1016/j.rse.2020.112061, https://www.sciencedirect.com/science/article/pii/S0034425720304314.

Sharma, M., N. Subedi, M. Ter-Mikaelian and J. Parton. 2015. Modeling climatic effects on stand height/site index of plantation-grown jack pine and black spruce trees. For. Sci. 61:25 – 34.

Sharma, M. and J. Parton. 2018a. Analyzing and modelling effects of climate on site productivity of white spruce plantations. For. Chron. 93 173 – 182.

Sharma, M. and J. Parton. 2018. Climate effects on site productivity of red pine plantations. For. Sc. doi: 10.1093/forsci/fxy013

Sharma, M. and D. Reid. 2018. Stand height/site index equations for jack pine and black spruce trees grown in natural stands. For. Sci. 64:33 – 40.

Sharma, M and J. Parton. 2019. Modelling the effects of climate on site productivity of white pine plantations. Can. J. For. Res. 49: 1289–1297.

Sharma, M. 2021. Climate effects on jack pine and black spruce productivity in natural origin mixed stands and site index conversion equations. Trees, Forests and People https://doi.org/10.1016/j.tfp.2021.100089

Subedi, N. and M. Sharma. 2010. Evaluating height-age determination methods for jack pine and black spruce plantations using stem analysis data. NORTH. J. APPL. FOR. 27(2) 2010

Westfall, J., M. Hatfield, P. Sowers and B. O'Connell. Site index models for tree species in the northeastern United States. For. Sci. 63(3):283–290

White, J.C., Penner, M., Woods, M. 2021. Assessing single photon LiDAR for operational implementation of an enhanced forest inventory in diverse mixedwood forests. The Forestry Chronicle, 97(1): 78-96. DOI: 10.5558/tfc2021-009

Woods, M. D. Pitt, M. Penner, K. Lim, D. Nesbitt, D. Etheridge and P. Treitz. 2011. Operational implementation of a LiDAR inventory in Boreal Ontario. Forestry Chronicle. 87(4). 512-528.

Van Ewijk, K.Y., Treitz, P.M., and Scott, N.A. 2011. Characterizing forest succession in central Ontario using LiDAR-derived indices. Photogrammetric Engineering and Remote Sensing, Vol. 77, No. 3, pp. 261_269.

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

Appendix A - LiDAR predictors for APF- SPL–2018

Full point cloud predictor suite derived from LidR software scripts from a threshold height > 0 m unless specified. Predictors selected for use in Random Forest modeling of inventory attributes are indicated.

Attribute	Threshold	Description	Model Predictor
zmax	>0m	max height of z	
zmean	>0m	mean height of z	▨
zsd	>0m	standard deviation of z	
zskew	>0m	skewness of z	
zkurt	>0m	kurtosis of z	
zentropy	>0m	entropy of height distribution (z)	▨
pzabovezmean	>0m	percentage of returns above zmean	▨
pzabove0	>0m	percentage of returns above threshold	
zq5	>0m	height of the 5th percentile of z	▨
zq10	>0m	height of the 10th percentile of z	▨
zq15	>0m	height of the 15th percentile of z	▨
zq20	>0m	height of the 20th percentile of z	▨
zq25	>0m	height of the 25th percentile of z	▨
zq30	>0m	height of the 30th percentile of z	▨
zq35	>0m	height of the 35th percentile of z	▨
zq40	>0m	height of the 40th percentile of z	▨
zq45	>0m	height of the 45th percentile of z	▨
zq50	>0m	height of the 50th percentile of z	▨
zq55	>0m	height of the 55th percentile of z	▨
zq60	>0m	height of the 60th percentile of z	▨
zq65	>0m	height of the 65th percentile of z	▨
zq70	>0m	height of the 70th percentile of z	▨
zq75	>0m	height of the 75th percentile of z	▨
zq80	>0m	height of the 80th percentile of z	▨
zq85	>0m	height of the 85th percentile of z	▨
zq90	>0m	height of the 90th percentile of z	▨
zq95	>0m	height of the9 5th percentile of z	▨
zq99	>0m	height of the 99th percentile of z	▨
zpcum1	>0m	percent of z returns below the 1st decile	▨
zpcum2	>0m	percent of z returns below the 2nd decile	▨
zpcum3	>0m	percent of z returns below the 3rd decile	▨
zpcum4	>0m	percent of z returns below the 4th decile	▨
zpcum5	>0m	percent of z returns below the 5th decile	▨
zpcum6	>0m	percent of z returns below the 6th decile	▨
zpcum7	>0m	percent of z returns below the 7th decile	▨
zpcum8	>0m	percent of z returns below the 8th decile	▨
zpcum9	>0m	percent of z returns below the 9th decile	▨
zsd95	>0m	standard deviation of z trimmed to 95%	▨
zskew95	>0m	skewness of z trimmed to 95%	▨
zkurt95	>0m	kurtosis of z trimmed to 95%	▨
zmin	>0m	minimum height of z returns	
allpts	>=0m	count of all points > Threshold (2,3,4,5)	
allptsAT	>0m	count of all points (2,3,4,5)	
vegcnt	>=0m	count of vegetation points (3,4,5)	
firstveg	>=0m	count of first return points of vegetation (3,4,5)	
firstcnt	>=0m	count of first returns	
firstonlycnt	>=0m	count of first and ONLY return points of vegetation (3,4,5)	
groundcnt	>=0m	count of f=ground returns (2)	
vegratio	>=0m	vegetation ratio (vegetation points (vegcnt) / all points (allpts))	
da	>=0m	percentage of First Returns / all returns (firstcnt / allpts) *100	
db	>=0m	percentage of "First & Only" Returns / all returns (firstonlycnt	
dv	>=0m	percentage of "Vegetation & Only" Returns / all returns	
vdr	>0m	Vertical Distribution Ratio (max-median)/max	▨
cv	>0m	coefficient of variation of z returns	▨

vci_1m	>0m	vegetation complexity index - 1m bins (Van Ewijk 2011)	
cov_2m	NA	canopy cover % above 2m (number of first returns above 2m / number of first returns) * 100	
cov_4m	NA	canopy cover % above 4m (number of first returns above 4m / number of first returns) * 100	
cov_6m	NA	canopy cover % above 6m (number of first returns above 6m / number of first returns) * 100	
cov_8m	NA	canopy cover % above 8m (number of first returns above 8m / number of first returns) * 100	
cov_10m	NA	canopy cover % above 10m (number of first returns above 10m / number of first returns) * 100	
cov_12m	NA	canopy cover % above 12m (number of first returns above 12m / number of first returns) * 100	
cov_14m	NA	canopy cover % above 14m (number of first returns above 14m / number of first returns) * 100	
cov_16m	NA	canopy cover % above 16m (number of first returns above 16m / number of first returns) * 100	
cov_18m	NA	canopy cover % above 18m (number of first returns above 18m / number of first returns) * 100	
cov_20m	NA	canopy cover % above 20m (number of first returns above 20m / number of first returns) * 100	
cov_22m	NA	canopy cover % above 22m (number of first returns above 22m / number of first returns) * 100	
cov_24m	NA	canopy cover % above 24m (number of first returns above 24m / number of first returns) * 100	
cov_26m	NA	canopy cover % above 26m (number of first returns above 26m / number of first returns) * 100	
cov_28m	NA	canopy cover % above 28m (number of first returns above 28m / number of first returns) * 100	
cov_30m	NA	canopy cover % above 30m (number of first returns above 30m / number of first returns) * 100	
dns_2m	NA	canopy cover % above 2m (number of all returns above 2m / number of all returns) * 100	
dns_4m	NA	canopy cover % above 4m (number of all returns above 4m / number of all returns) * 100	
dns_6m	NA	canopy cover % above 6m (number of all returns above 6m / number of all returns) * 100	
dns_8m	NA	canopy cover % above 8m (number of all returns above 8m / number of all returns) * 100	
dns_10m	NA	canopy cover % above 10m (number of all returns above 10m / number of all returns) * 100	
dns_12m	NA	canopy cover % above 12m (number of all returns above 12m / number of all returns) * 100	
dns_14m	NA	canopy cover % above 14m (number of all returns above 14m / number of all returns) * 100	
dns_16m	NA	canopy cover % above 16m (number of all returns above 16m / number of all returns) * 100	
dns_18m	NA	canopy cover % above 18m (number of all returns above 18m / number of all returns) * 100	

dns_20m	NA	canopy cover % above 20m (number of all returns above 18m / number of all returns) * 100	▨
dns_22m	NA	canopy cover % above 22m (number of all returns above 18m / number of all returns) * 100	▨
dns_24m	NA	canopy cover % above 24m (number of all returns above 24m / number of all returns) * 100	▨
dns_26m	NA	canopy cover % above 26m (number of all returns above 26m / number of all returns) * 100	▨
dns_28m	NA	canopy cover % above 28m (number of all returns above 28m / number of all returns) * 100	▨
dns_30m	NA	canopy cover % above 30m (number of all returns above 30m / number of all returns) * 100	▨
vegden_0_2	>=0m	Percent vegetation returns between 0 and 2m	▨
vegden_2_4	>=0m	Percent vegetation returns between 2 and 4m	▨
vegden_4_6	>=0m	Percent vegetation returns between 4 and 6m	▨
vegden_6_8	>=0m	Percent vegetation returns between 6 and 8m	▨
vegden_8_10	>=0m	Percent vegetation returns between 10 and 10m	▨
vegden_10_12	>=0m	Percent vegetation returns between 10 and 12m	▨
vegden_12_14	>=0m	Percent vegetation returns between 12 and 14m	▨
vegden_14_16	>=0m	Percent vegetation returns between 14 and 16m	▨
vegden_16_18	>=0m	Percent vegetation returns between 16 and 18m	▨
vegden_18_20	>=0m	Percent vegetation returns between 18 and 20m	▨
vegden_20_22	>=0m	Percent vegetation returns between 20 and 22m	▨
vegden_22_24	>=0m	Percent vegetation returns between 22 and 24m	▨
vegden_24_26	>=0m	Percent vegetation returns between 24 and 26m	▨
vegden_26_28	>=0m	Percent vegetation returns between 26 and 28m	▨
vegden_28_30	>=0m	Percent vegetation returns between 28 and 30m	▨
L1	NA	L1 moment of vegetation points (3,4,5)	
L2	NA	L2 moment of vegetation points (3,4,5)	▨
L3	NA	L3 moment of vegetation points (3,4,5)	▨
L4	NA	L4 moment of vegetation points (3,4,5)	▨
Lskew	NA	L Skewness of vegetation points (3,4,5)	▨
Lkurt	NA	L Kurtosis of vegetation points (3,4,5)	▨
Lcoefvar	NA	L Coeficient of Variation of vegetation points (3,4,5)	▨
ngrcnt	-0.15	count of all points (2,3,4,5) between -0.15 and 0.15 for LPI	
allptscnt_ngr	-0.15	Count of all points (2,3,4,5) between -0.15 and 48m for LPI	
lpi	-0.15	LiDAR penetration index - count of returns between (-0.15 - 0.15)/all points (-.15 to 30m) * 100 [Uses Class 2,3,4,5]	▨
ri_pts	NA	rumple index based on LiDAR points - 1m DSM	▨

Appendix B – Plot level validation statistics by CV methods

Ten-Fold Cross Validation Plot level model statistics by Forest Unit

Top Ht		**Observed**			**Prediction**					
m	**N**	**Mean**	**Min**	**Max**	**Mean**	**StdErr**	**RMSE**	**RMSE%**	**BIAS**	**BIAS%**
HeSel	13	22.1	18.2	26.3	22.0	0.7	1.3	5.9	0.1	0.5
HWDSEL_US	77	21.2	11.4	34.6	21.5	0.4	2.1	9.9	-0.4	-1.9
INTCC	5	27.9	19.7	33.3	28.9	2.2	1.7	6.1	-0.9	-3.2
LCUS	8	13.3	9.7	17.9	13.3	0.7	1.2	9.0	0.0	0.0
MW	5	25.3	18.5	33.4	24.8	2.4	0.8	3.2	0.5	2.0
PjCC	2	21.2	20.0	22.4	18.4	1.4	2.7	12.7	2.7	12.7
PrCC	21	27.0	14.1	35.6	26.6	1.0	2.2	8.1	0.4	1.5
PwUS	82	28.3	8.6	37.8	28.0	0.7	1.8	6.4	0.3	1.1
SFUS	8	20.3	13.9	29.3	20.4	2.0	0.8	3.9	-0.1	-0.5

CDHT		**Observed**			**Prediction**					
m	**N**	**Mean**	**Min**	**Max**	**Mean**	**StdErr**	**RMSE**	**RMSE%**	**BIAS**	**BIAS%**
HeSel	13	20.2	14.2	25.0	19.6	0.7	1.9	9.4	0.5	2.5
HWDSEL_US	77	19.0	10.5	25.6	19.2	0.3	2.0	10.5	-0.3	-1.6
INTCC	5	25.1	20.0	30.9	25.9	1.9	1.6	6.4	-0.8	-3.2
LCUS	8	11.4	8.1	16.2	10.8	0.9	1.8	15.8	0.6	5.3
MW	5	21.2	16.7	27.0	22.0	2.2	1.6	7.5	-0.8	-3.8
PjCC	2	17.8	15.8	19.7	17.1	1.5	0.8	4.5	0.7	3.9
PrCC	21	26.5	11.9	33.6	24.9	1.0	3.3	12.5	1.6	6.0
PwUS	82	24.7	6.3	35.4	24.5	0.6	3.2	13.0	0.2	0.8
SFUS	8	16.9	11.1	23.6	18.0	1.9	1.9	11.2	-1.1	-6.5

Lorey's Ht		**Observed**			**Prediction**					
m	**N**	**Mean**	**Min**	**Max**	**Mean**	**StdErr**	**RMSE**	**RMSE%**	**BIAS**	**BIAS%**
HeSel	13	20.9	15.7	25.5	19.8	0.6	1.7	8.1	1.1	5.3
HWDSEL_US	77	19.1	10.8	25.6	19.4	0.3	1.9	9.9	-0.3	-1.6
INTCC	5	24.4	19.4	30.0	25.8	1.8	2.1	8.6	-1.4	-5.7
LCUS	8	11.8	8.4	16.5	11.6	0.7	1.6	13.6	0.2	1.7
MW	5	21.2	16.2	26.7	22.0	2.2	1.7	8.0	-0.8	-3.8
PjCC	2	17.4	14.1	20.7	16.7	1.4	2.0	11.5	0.7	4.0
PrCC	21	25.2	12.9	32.7	24.5	1.0	2.4	9.5	0.8	3.2
PwUS	82	24.7	7.1	34.9	24.5	0.6	2.3	9.3	0.2	0.8
SFUS	8	17.4	11.9	24.2	18.0	1.8	1.3	7.5	-0.6	-3.4

QMD Ht		**Observed**			**Prediction**					
m	**N**	**Mean**	**Min**	**Max**	**Mean**	**StdErr**	**RMSE**	**RMSE%**	**BIAS**	**BIAS%**
HeSel	13	21.4	16.0	26.2	20.8	0.7	1.6	7.5	0.7	3.3
HWDSEL_US	77	19.9	11.6	26.0	20.3	0.4	1.8	9.0	-0.4	-2.0
INTCC	5	25.9	20.6	32.0	26.9	1.9	2.0	7.7	-1.0	-3.9
LCUS	8	11.8	7.3	16.9	11.8	0.7	1.9	16.1	0.1	0.8
MW	5	22.2	17.0	28.0	22.6	2.0	1.4	6.3	-0.3	-1.4
PjCC	2	18.0	14.5	21.4	17.6	1.4	2.1	11.7	0.4	2.2
PrCC	21	26.3	13.7	33.8	25.6	1.1	3.2	12.2	0.7	2.7
PwUS	82	26.0	7.2	36.1	25.8	0.6	2.6	10.0	0.2	0.8
SFUS	8	18.4	12.0	26.9	18.8	1.9	1.7	9.2	-0.4	-2.2

BasalArea		**Observed**			**Prediction**					
m^2 ha^{-1}	**N**	**Mean**	**Min**	**Max**	**Mean**	**StdErr**	**RMSE**	**RMSE%**	**BIAS**	**BIAS%**
HeSel	13	33.4	21.4	51.5	32.5	2.0	6.7	20.1	0.9	2.7
HWDSEL_US	77	24.9	4.4	46.2	25.1	0.6	5.3	21.3	-0.2	-0.8
INTCC	5	45.4	34.7	64.8	43.7	2.2	8.3	18.3	1.8	4.0
LCUS	8	19.4	2.3	37.2	18.7	3.2	8.8	45.4	0.7	3.6
MW	5	30.3	14.6	43.5	36.8	4.6	9.3	30.7	-6.5	-21.5
PjCC	2	15.0	11.7	18.3	18.8	2.8	3.8	25.3	-3.8	-25.3
PrCC	21	24.6	2.9	58.2	22.8	3.6	5.9	24.0	1.8	7.3
PwUS	82	27.0	3.0	66.5	27.5	1.4	6.1	22.6	-0.4	-1.5
SFUS	8	31.1	12.4	50.5	26.6	2.8	7.1	22.8	4.5	14.5

BasalArea merch		**Observed**			**Prediction**					
m^2 ha^{-1}	**N**	**Mean**	**Min**	**Max**	**Mean**	**StdErr**	**RMSE**	**RMSE%**	**BIAS**	**BIAS%**
HeSel	13	32.9	20.4	51.3	31.6	1.9	7.2	21.9	1.2	3.6
HWDSEL_US	77	24.3	2.4	45.5	24.5	0.6	5.2	21.4	-0.2	-0.8
INTCC	5	45.1	34.3	64.7	42.8	3.2	8.8	19.5	2.3	5.1
LCUS	8	18.0	1.3	36.7	17.3	2.8	8.6	47.8	0.7	3.9
MW	5	29.7	13.9	43.2	35.6	4.7	8.3	27.9	-5.9	-19.9
PjCC	2	14.7	11.3	18.2	17.9	2.7	3.2	21.8	-3.1	-21.1
PrCC	21	24.2	2.8	57.9	22.4	3.6	6.1	25.2	1.9	7.9
PwUS	82	26.4	2.3	65.9	27.0	1.3	6.1	23.1	-0.6	-2.3
SFUS	8	28.9	11.6	47.2	25.3	2.7	6.1	21.1	3.6	12.5

GTV		**Observed**			**Prediction**					
m^3 ha^{-1}	**N**	**Mean**	**Min**	**Max**	**Mean**	**StdErr**	**RMSE**	**RMSE%**	**BIAS**	**BIAS%**
HeSel	13	246.9	130.7	394.8	259.9	22.2	51.3	20.8	-13.0	-5.3
HWDSEL_US	77	192.2	24.2	409.5	195.4	6.1	44.3	23.0	-3.2	-1.7
INTCC	5	490.8	312.9	707.7	480.7	59.5	54.7	11.1	10.2	2.1
LCUS	7	116.8	13.2	236.1	106.3	22.9	42.7	36.6	10.5	9.0
MW	5	286.3	100.5	494.7	343.6	84.7	85.0	29.7	-57.3	-20.0
PjCC	2	117.6	70.6	164.7	129.0	26.2	23.8	20.2	-11.4	-9.7
PrCC	21	293.3	31.9	842.7	251.7	45.1	91.6	31.2	41.6	14.2
PwUS	82	283.2	16.4	880.7	283.5	17.3	69.7	24.6	-0.3	-0.1
SFUS	8	224.6	65.5	449.6	206.1	38.0	37.4	16.7	18.5	8.2

GMV NL		**Observed**			**Prediction**					
m^3 ha^{-1}	**N**	**Mean**	**Min**	**Max**	**Mean**	**StdErr**	**RMSE**	**RMSE%**	**BIAS**	**BIAS%**
HeSel	13	206.2	99.2	365.3	198.9	20.6	44.0	21.3	7.4	3.6
HWDSEL_US	77	138.8	0.0	342.8	145.1	5.5	43.1	31.1	-6.4	-4.6
INTCC	5	442.0	260.2	643.3	422.7	62.0	62.1	14.0	19.3	4.4
LCUS	7	84.7	7.3	205.6	66.9	15.4	46.1	54.4	17.8	21.0
MW	5	231.7	58.7	432.5	278.0	75.8	62.3	26.9	-46.2	-19.9
PjCC	2	108.7	62.1	155.2	96.5	27.5	22.6	20.8	12.1	11.1
PrCC	21	270.3	30.1	788.5	226.7	41.7	91.3	33.8	43.5	16.1
PwUS	82	251.7	5.6	825.1	248.3	16.3	68.2	27.1	3.4	1.4
SFUS	8	173.5	46.4	412.4	156.2	38.8	38.0	21.9	17.2	9.9

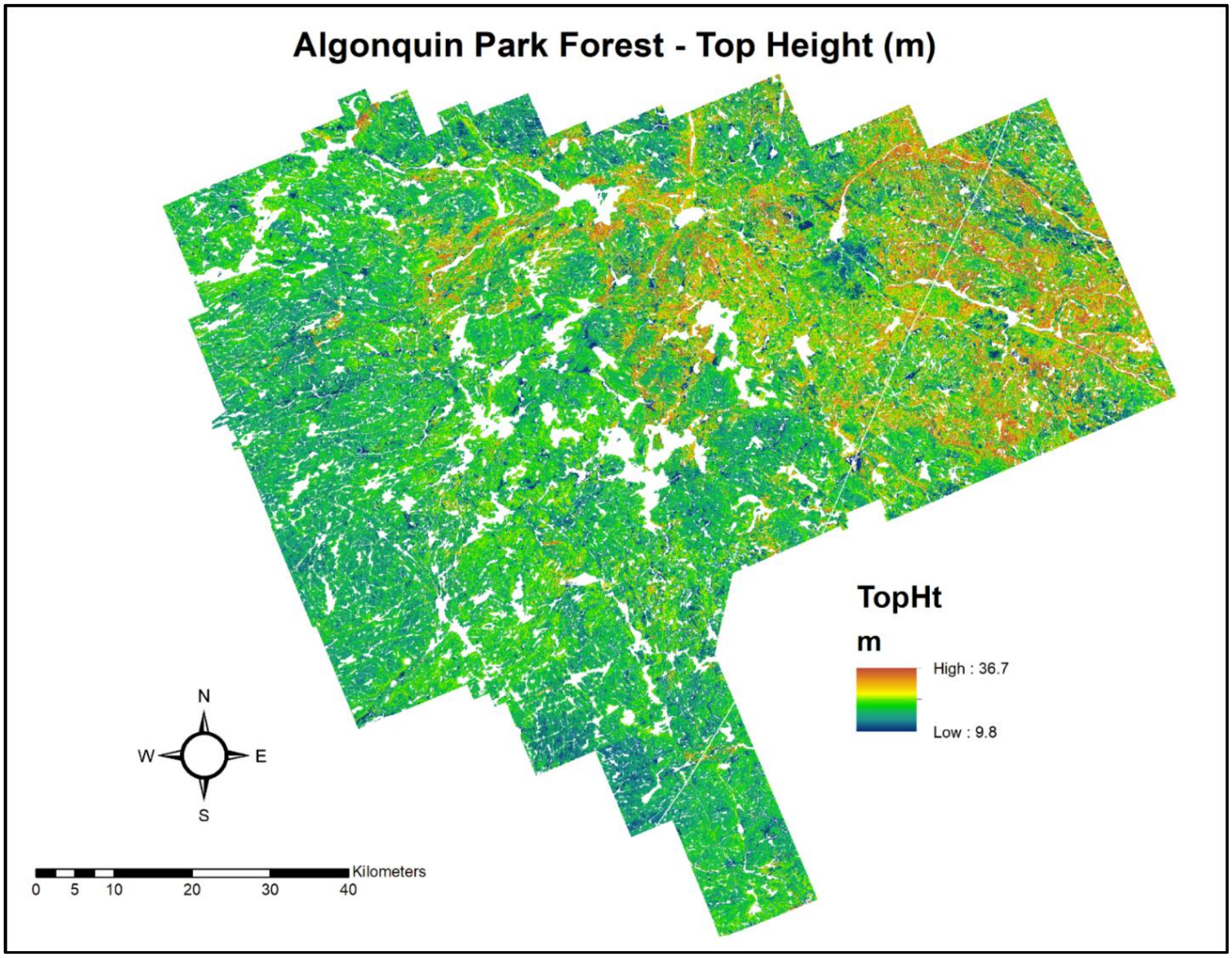
GMV WL		**Observed**			**Prediction**					
m^3 ha^{-1}	**N**	**Mean**	**Min**	**Max**	**Mean**	**StdErr**	**RMSE**	**RMSE%**	**BIAS**	**BIAS%**
HeSel	13	199.0	95.7	357.1	188.7	19.9	45.8	23.0	10.4	5.2
HWDSEL_US	77	129.0	0.0	331.9	135.3	5.2	43.3	33.6	-6.4	-5.0
INTCC	5	428.4	244.5	623.6	410.5	62.2	61.5	14.4	17.9	4.2
LCUS	7	76.7	6.0	199.2	58.2	13.6	46.6	60.8	18.5	24.1
MW	5	218.9	48.7	422.4	266.4	74.0	59.2	27.0	-47.5	-21.7
PjCC	2	106.1	60.3	151.9	87.6	24.1	28.6	27.0	18.5	17.4
PrCC	21	264.4	28.8	778.3	220.6	40.9	91.0	34.4	43.8	16.6
PwUS	82	246.0	4.2	815.4	241.7	16.2	67.3	27.4	4.3	1.7
SFUS	8	164.9	39.0	406.1	147.7	39.1	38.9	23.6	17.2	10.4

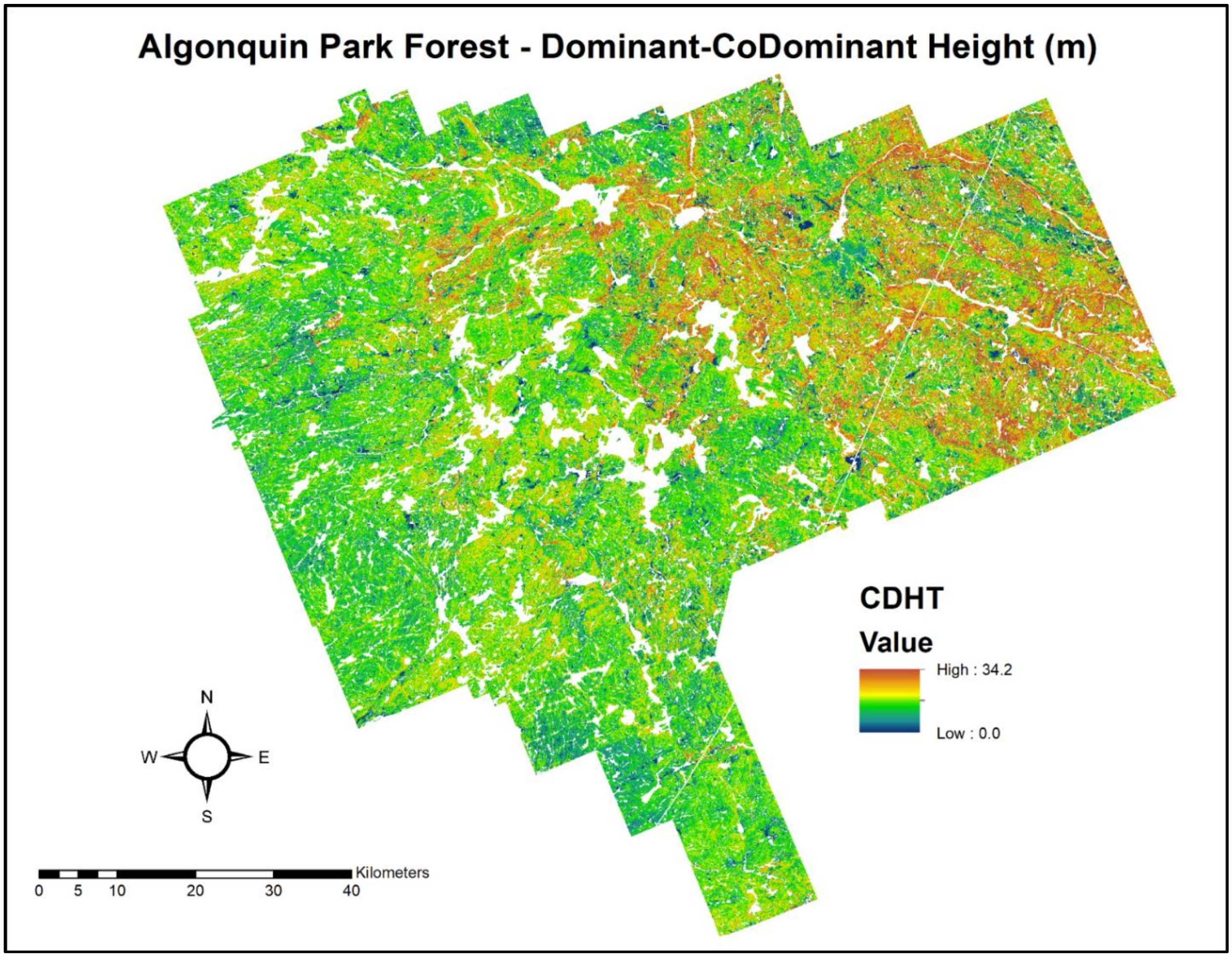
QMD		**Observed**			**Prediction**					
cm	**N**	**Mean**	**Min**	**Max**	**Mean**	**StdErr**	**RMSE**	**RMSE%**	**BIAS**	**BIAS%**
HeSel	13	29.5	17.2	42.1	24.3	0.9	8.0	27.1	5.3	18.0
HWDSEL_US	77	22.6	9.7	37.6	22.8	0.4	3.4	15.0	-0.2	-0.9
INTCC	5	26.4	20.8	29.9	27.3	1.0	3.0	11.4	-0.9	-3.4
LCUS	8	14.1	9.0	24.7	14.1	0.5	4.1	29.1	0.0	0.0
MW	5	21.9	16.6	28.7	26.9	2.4	5.2	23.7	-5.0	-22.8
PjCC	2	24.5	22.2	26.8	21.1	3.5	3.6	14.7	3.4	13.9
PrCC	21	28.1	16.9	38.5	28.9	1.3	5.9	21.0	-0.8	-2.8
PwUS	82	27.3	11.0	59.7	27.6	0.6	5.9	21.6	-0.4	-1.5
SFUS	8	18.4	11.5	27.0	19.8	1.9	3.1	16.8	-1.5	-8.2

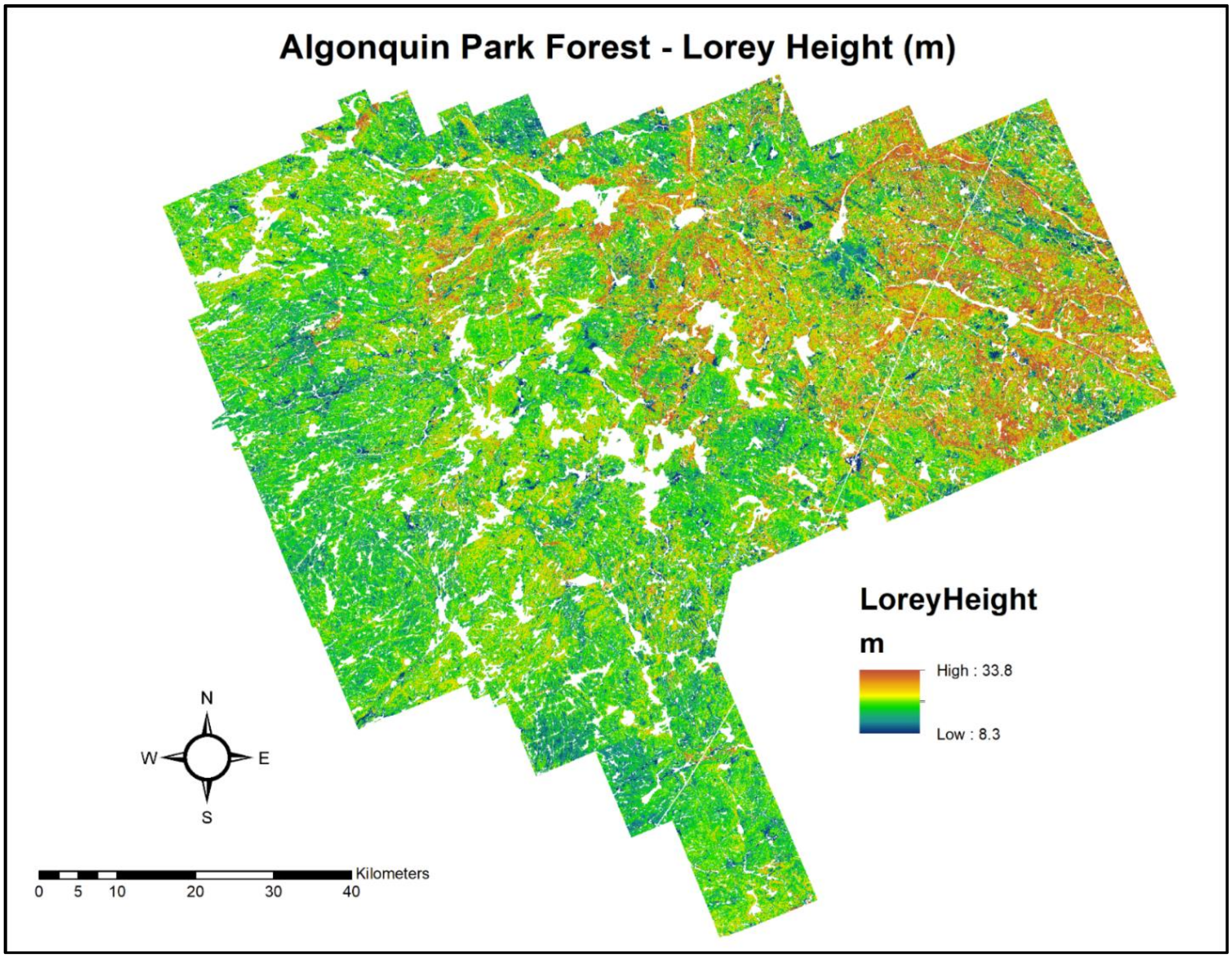
Biomass		**Observed**			**Prediction**					
T ha^{-1}	**N**	**Mean**	**Min**	**Max**	**Mean**	**StdErr**	**RMSE**	**RMSE%**	**BIAS**	**BIAS%**
HeSel	13	186.8	93.7	281.0	165.5	11.0	40.7	21.8	21.3	11.4
HWDSEL_US	77	166.7	21.2	357.3	172.4	5.3	37.9	22.7	-5.7	-3.4
INTCC	5	275.6	180.8	409.4	251.1	30.3	52.1	18.9	24.5	8.9
LCUS	8	60.5	8.3	109.0	58.0	11.3	17.7	29.3	2.5	4.1
MW	5	177.4	82.7	285.5	187.5	36.0	28.4	16.0	-10.1	-5.7
PjCC	2	65.7	44.2	87.1	75.7	13.8	12.6	19.2	-10.0	-15.2
PrCC	21	141.2	15.3	402.8	124.0	21.2	40.3	28.5	17.2	12.2
PwUS	82	148.0	10.5	439.4	150.5	9.1	34.4	23.2	-2.5	-1.7
SFUS	8	138.0	42.7	241.3	122.3	19.8	23.4	17.0	15.7	11.4

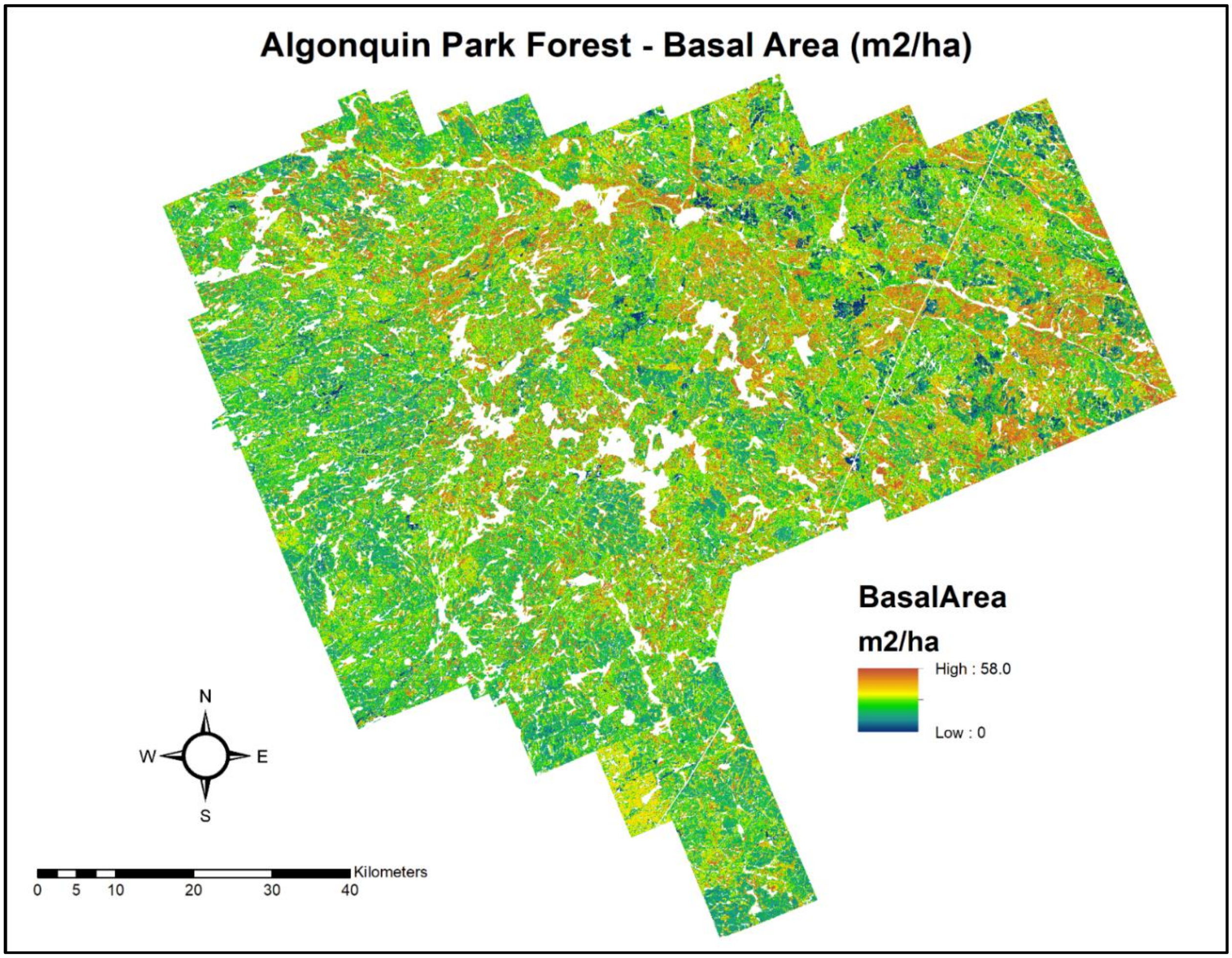
Density		**Observed**			**Prediction**					
Stems ha^{-1}	**N**	**Mean**	**Min**	**Max**	**Mean**	**StdErr**	**RMSE**	**RMSE%**	**BIAS**	**BIAS%**
HeSel	13	548	250	1050	713	41.1	291	53.2	-165	-30.0
HWDSEL_US	77	665	275	1475	646	23.9	194	29.1	19	2.8
INTCC	5	835	675	1075	748	22.4	156	18.7	87	10.4
LCUS	8	1153	250	2000	1128	145.5	477	41.4	25	2.2
MW	5	825	650	1400	656	56.9	379	45.9	170	20.5
PjCC	2	313	300	325	556	99.9	268	85.7	-243	-77.8
PrCC	21	451	50	1425	401	91.7	227	50.4	50	11.2
PwUS	82	589	25	2075	511	34.1	319	54.1	78	13.2
SFUS	8	1488	450	3525	960	145.8	871	58.5	528	35.5

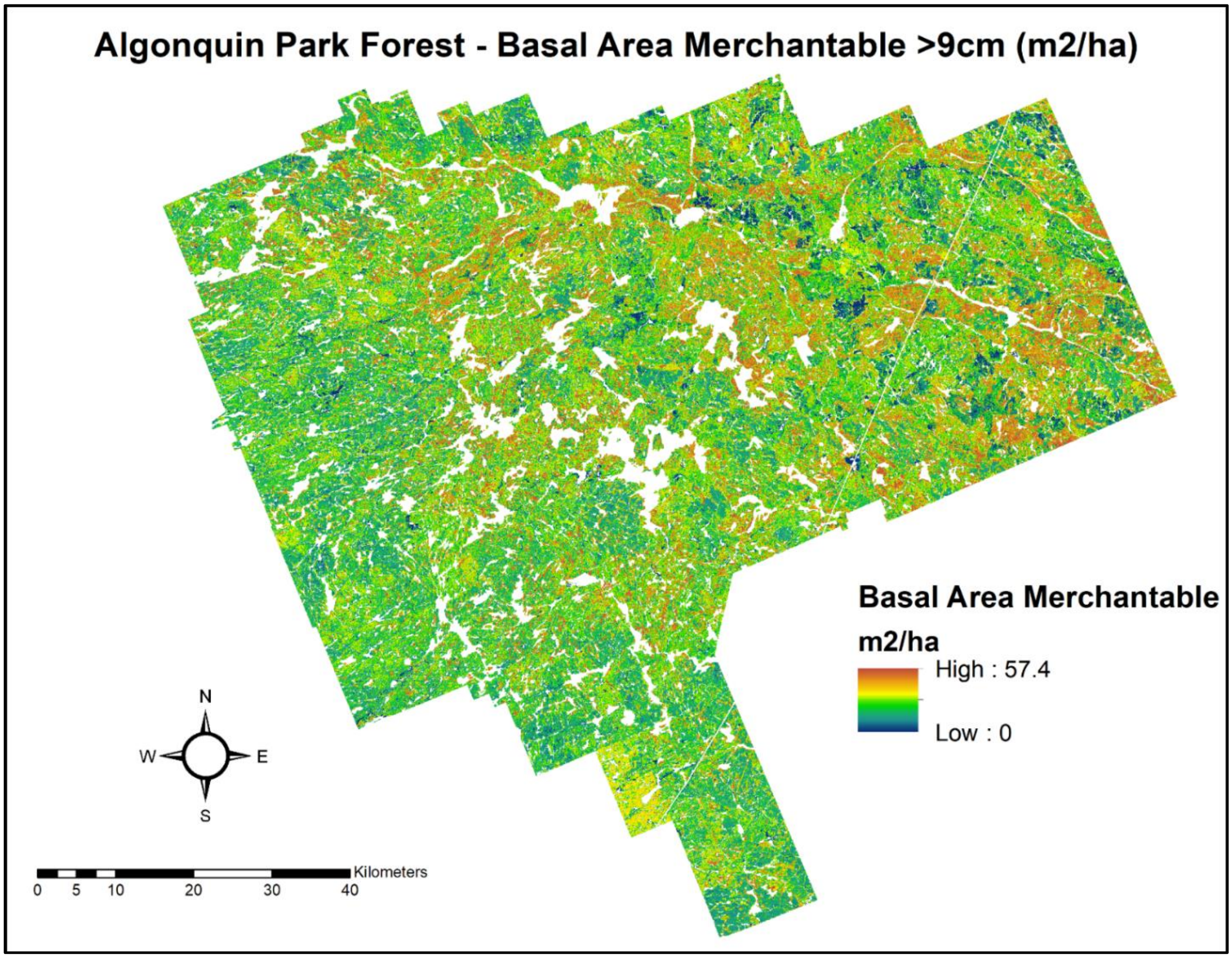
Appendix C – APF Inventory Rasters

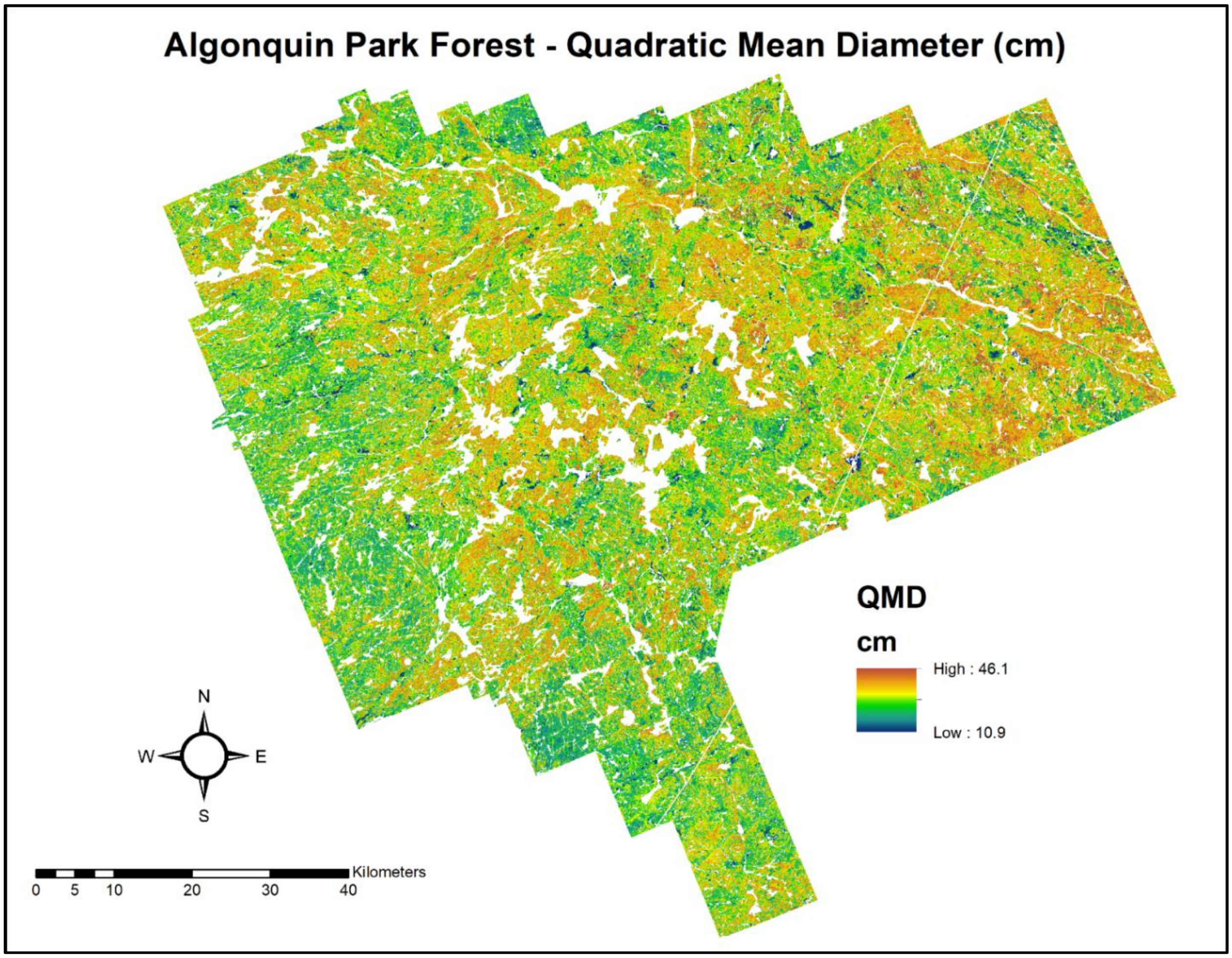


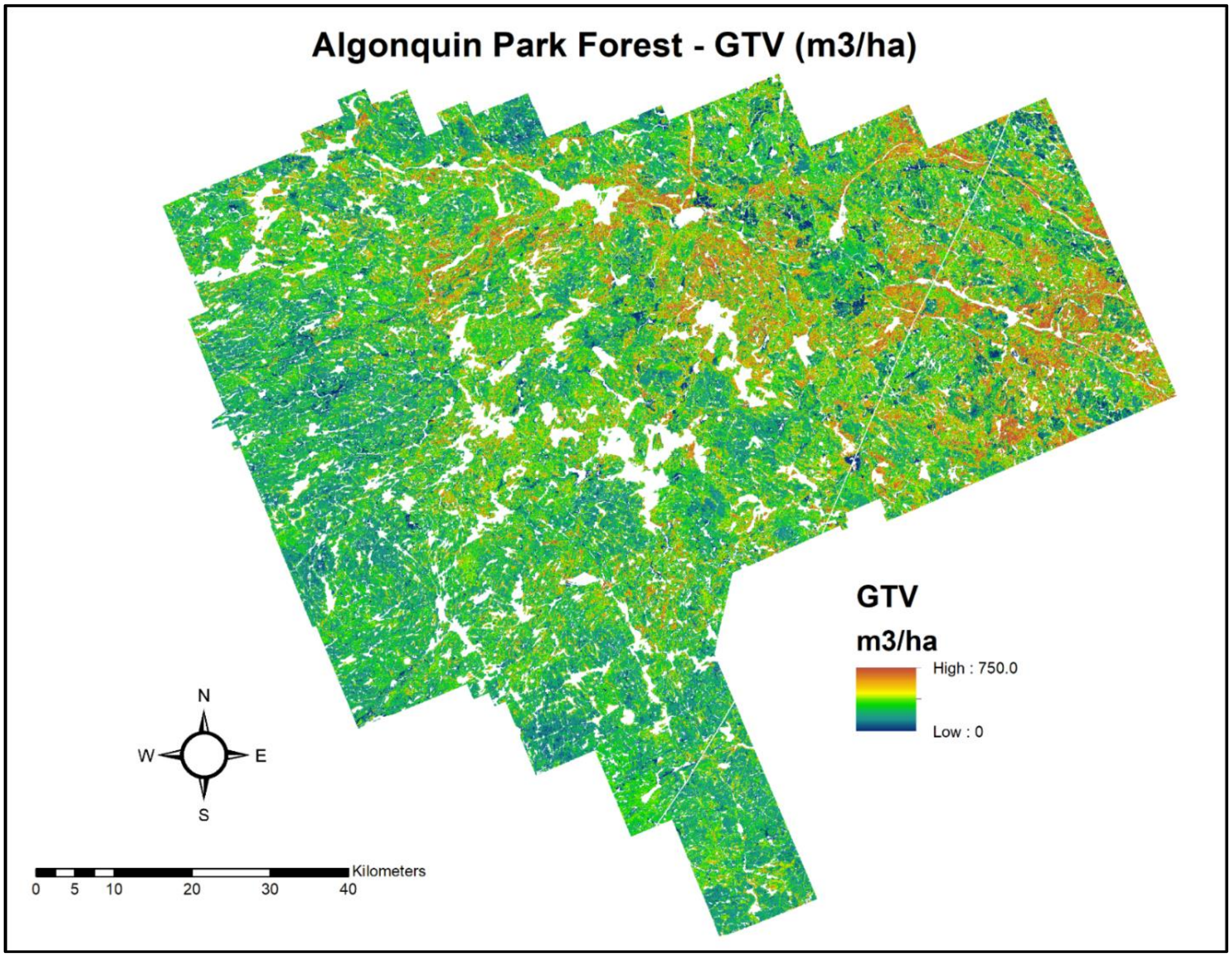

Algonquin Park Forest - Dominant-CoDominant Height (m)
CDHT
Value
High : 34.2
Low : 0.0
N
W
E
S
0
5
10
20
30
40
Kilometers

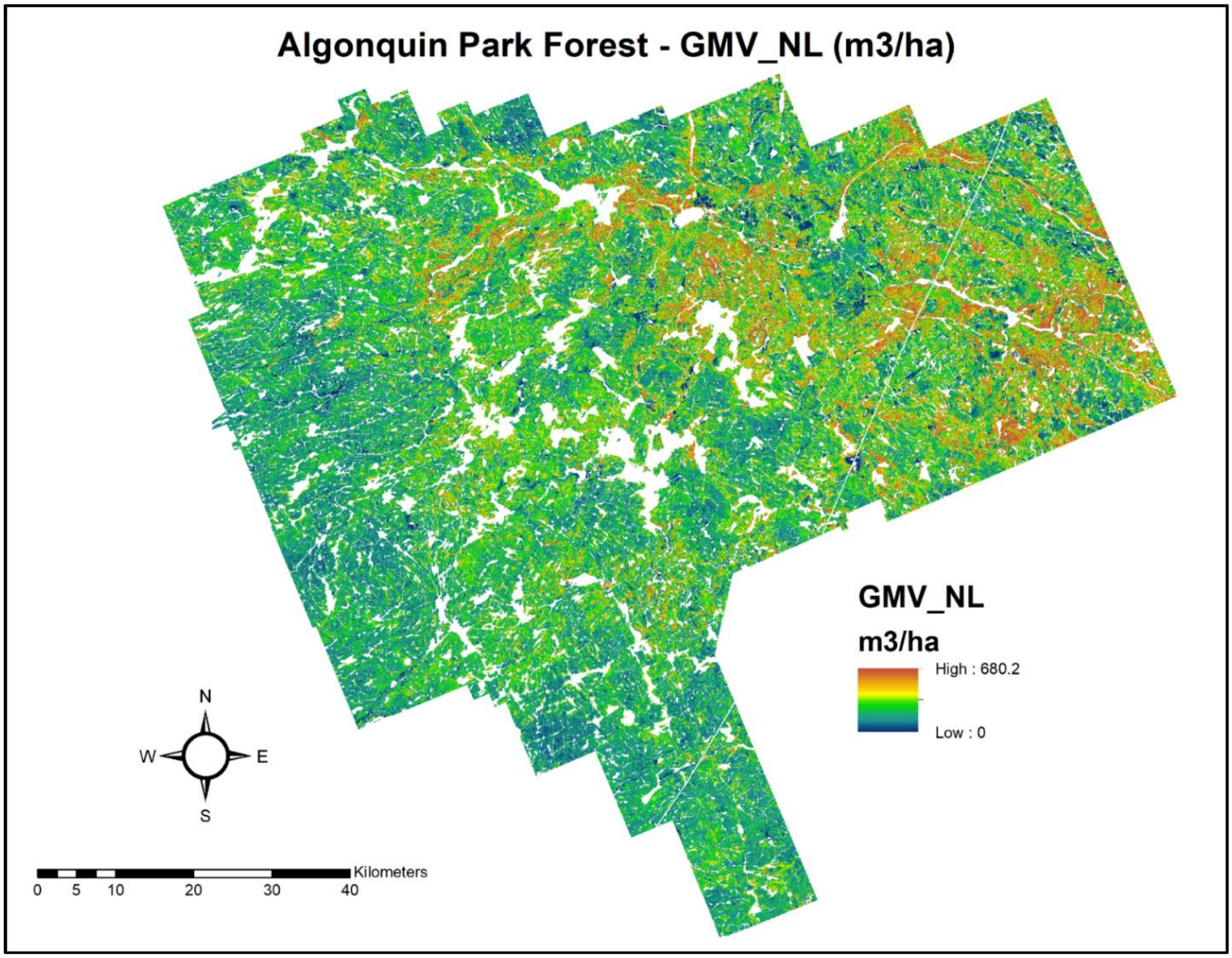

Algonquin Park Forest - Lorey Height (m)
LoreyHeight
m
High : 33.8
Low : 8.3
N
W
E
S
0
5
10
20
30
40
Kilometers

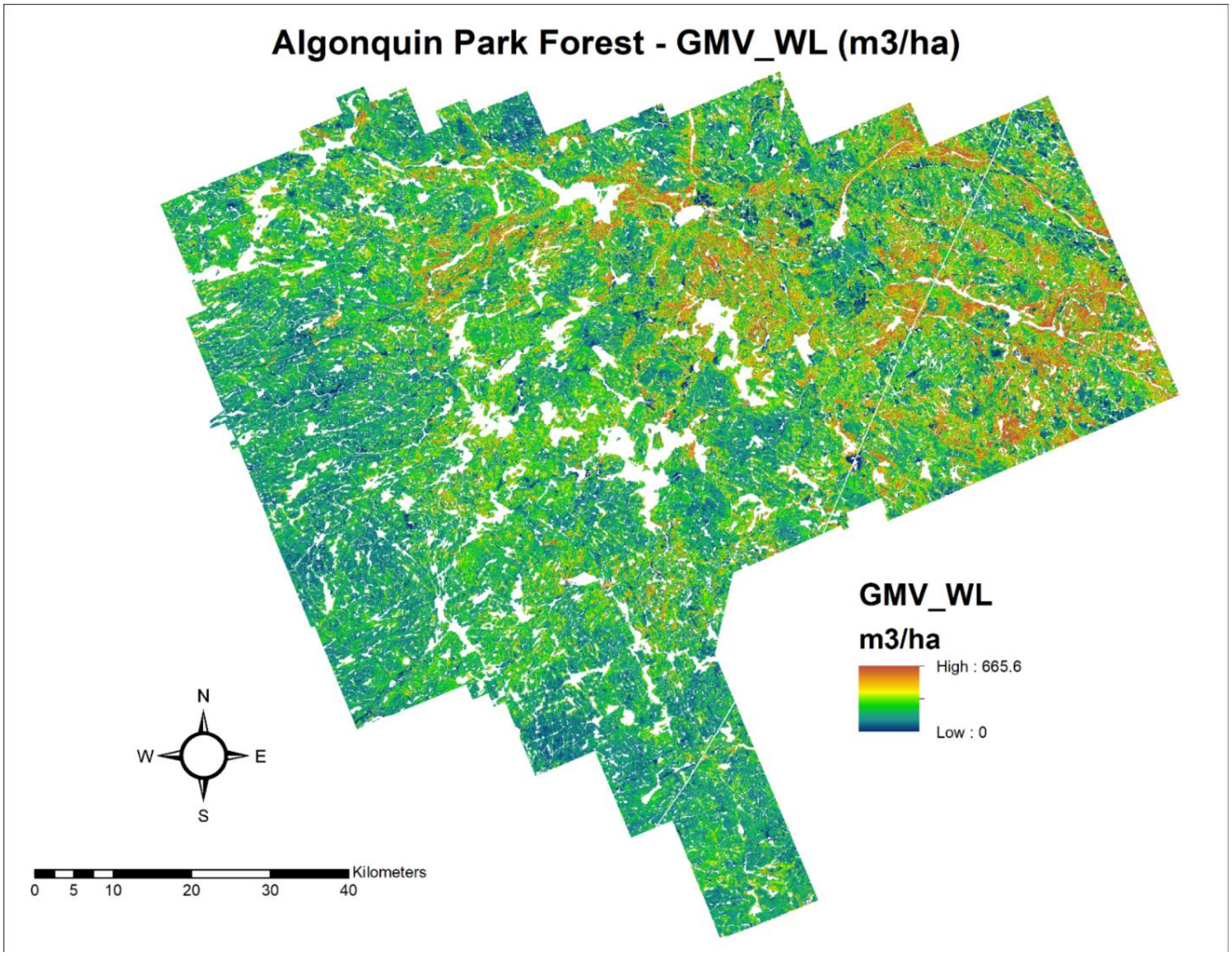

Algonquin Park Forest - Basal Area (m2/ha)
BasalArea
m2/ha
High : 58.0
Low : 0
N
W
E
S
Kilometers
0
5
10
20
30
40

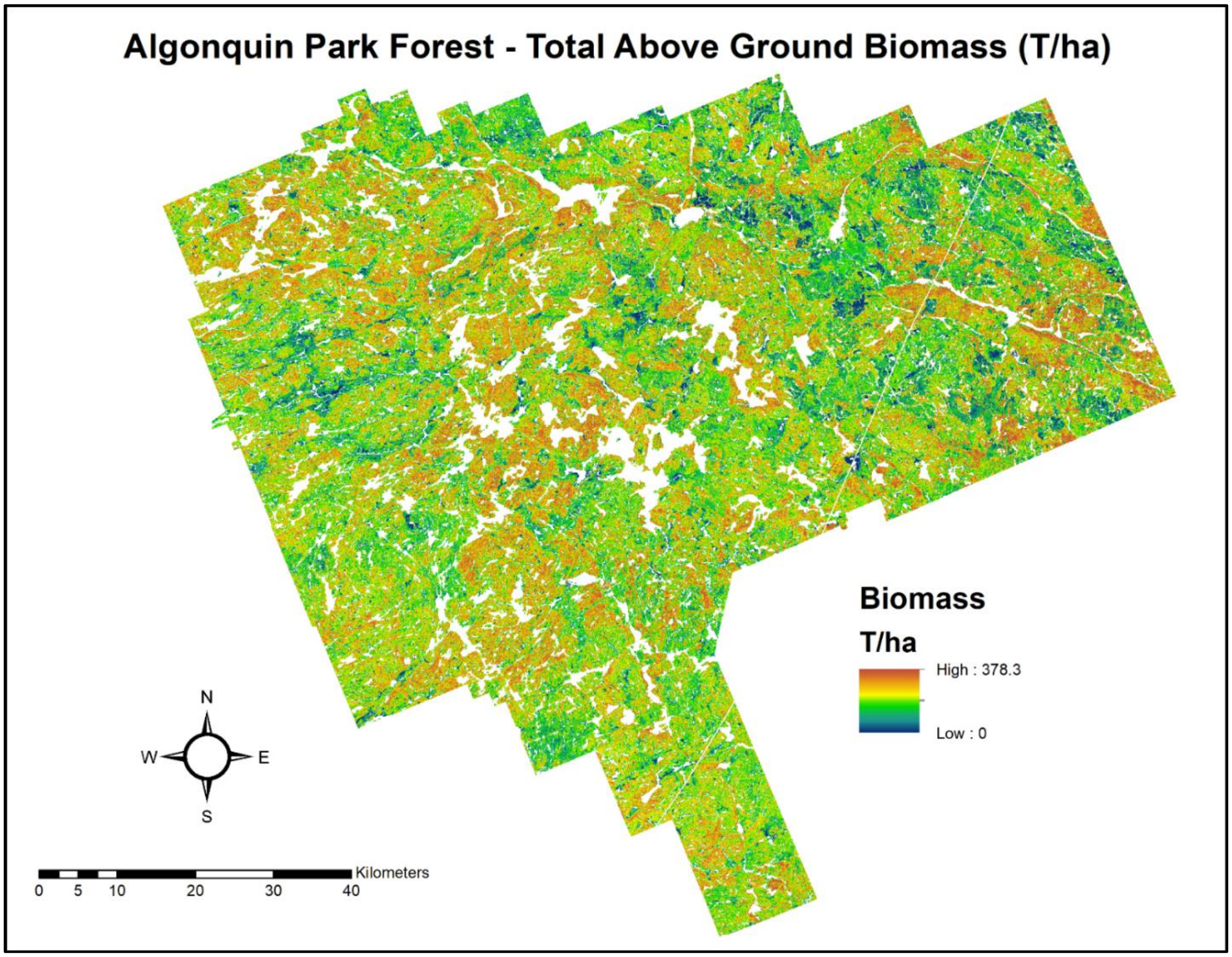

Algonquin Park Forest - Basal Area Merchantable >9cm (m2/ha)
Basal Area Merchantable
m2/ha
High : 57.4
Low : 0
N
W
E
S
0
5
10
20
30
40
Kilometers

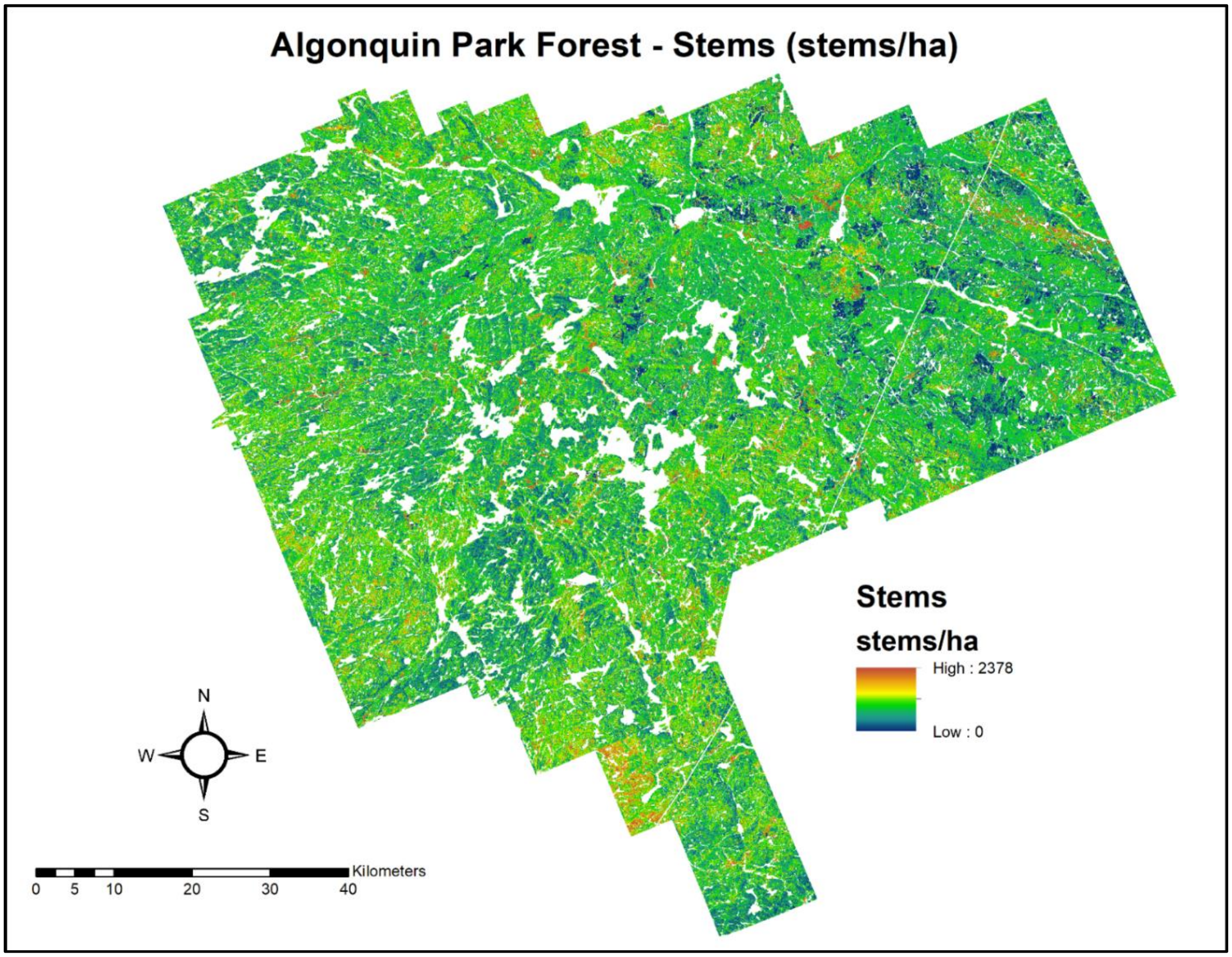

Algonquin Park Forest - Quadratic Mean Diameter (cm)
QMD
cm
High : 46.1
Low : 10.9
N
W
E
S
0
5
10
20
30
40
Kilometers


Algonquin Park Forest - GTV (m3/ha)
GTV
m3/ha
High : 750.0
Low : 0
N
W
E
S
Kilometers
0
5
10
20
30
40

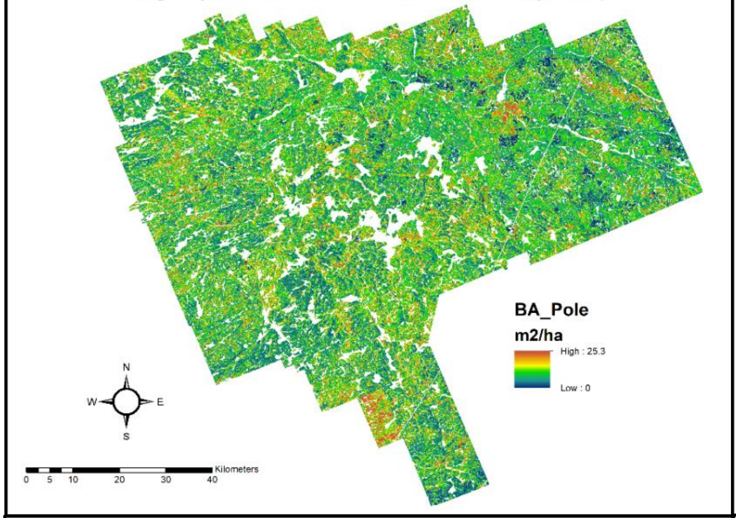

Algonquin Park Forest - GMV_NL (m3/ha)
GMV_NL
m3/ha
High : 680.2
Low : 0
N
W
E
S
Kilometers
0
5
10
20
30
40


Algonquin Park Forest - GMV_WL (m3/ha)
GMV_WL
m3/ha
High : 665.6
Low : 0
N
W
E
S
Kilometers
0
5
10
20
30
40

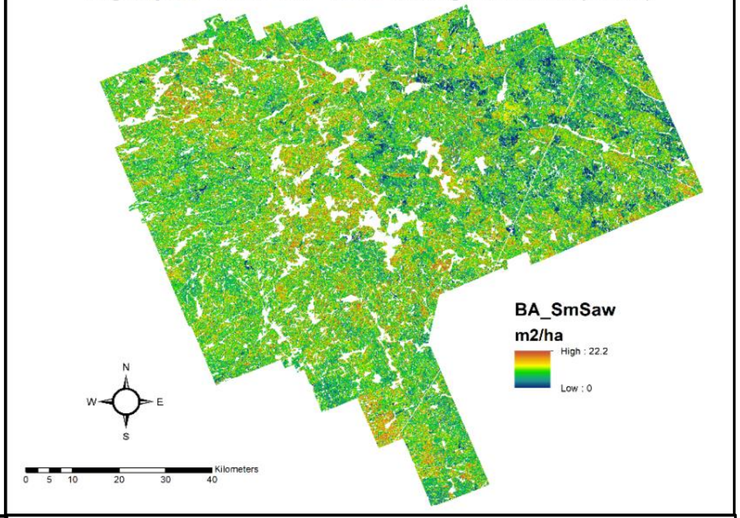

Algonquin Park Forest - Total Above Ground Biomass (T/ha)
Biomass
T/ha
High : 378.3
Low : 0
N
W
E
S
0
5
10
20
30
40
Kilometers


Algonquin Park Forest - Stems (stems/ha)
Stems
stems/ha
High : 2378
Low : 0
N
W
E
S
Kilometers
0
5
10
20
30
40

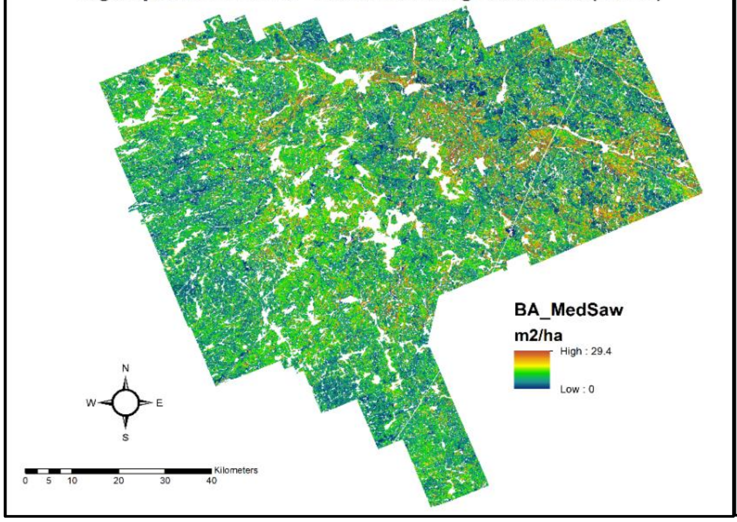
Algonquin Park Forest - Pole Basal Area (m2/ha)



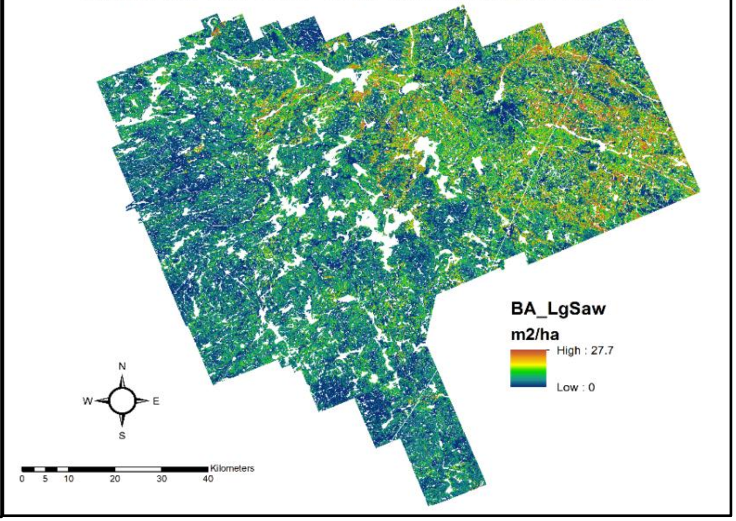
Algonquin Park Forest - Small Sawlog Basal Area (m2/ha)

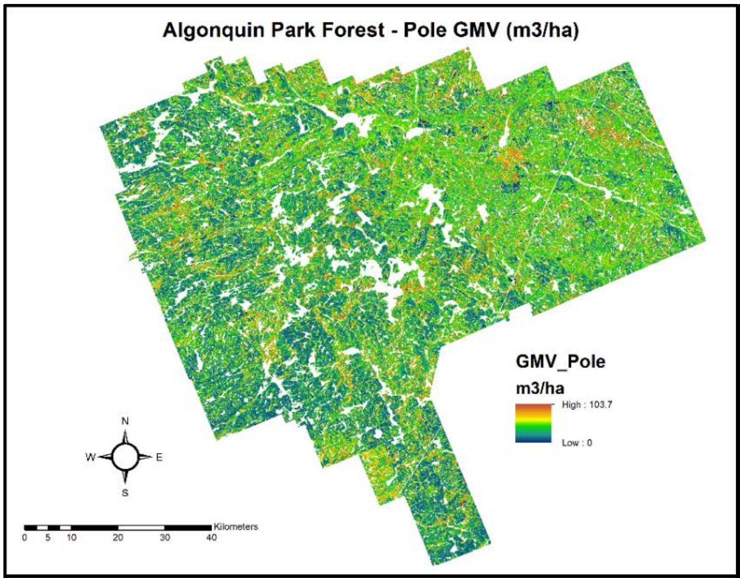


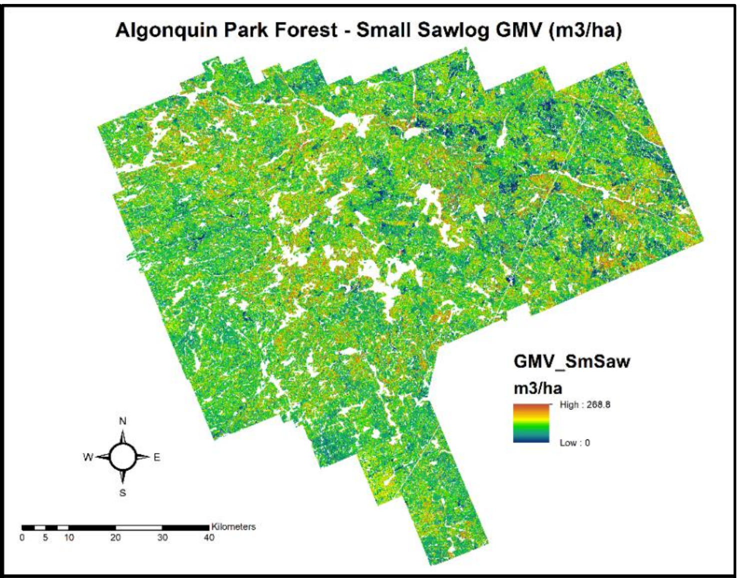
Algonquin Park Forest - Medium Sawlog Basal Area (m2/ha)

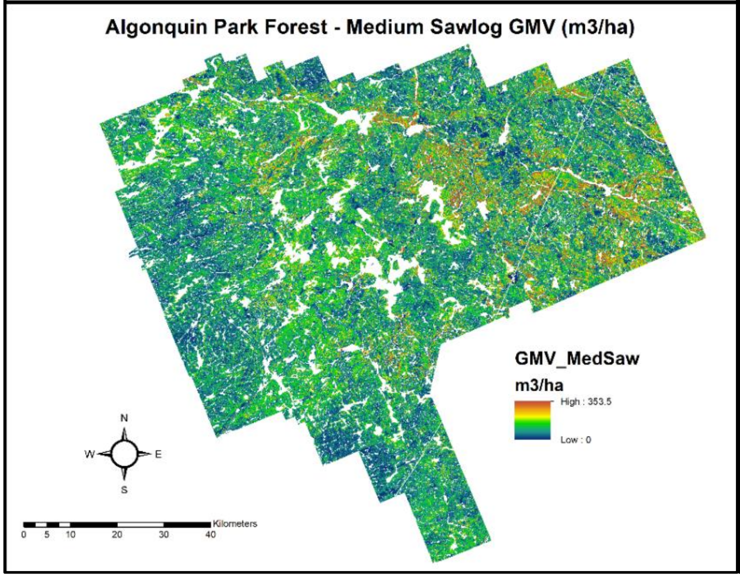


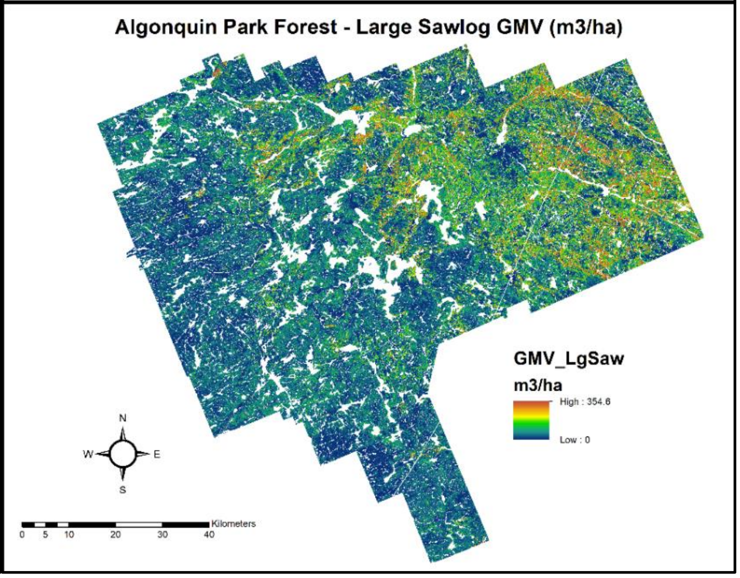
Algonquin Park Forest - Large Sawlog Basal Area (m2/ha)




Algonquin Park Forest - Pole GMV (m3/ha)


Algonquin Park Forest - Small Sawlog GMV (m3/ha)


Algonquin Park Forest - Medium Sawlog GMV (m3/ha)


Algonquin Park Forest - Large Sawlog GMV (m3/ha)

Appendix D – Site Index Curve Sources

Sharma and Reid (2018) recommend that height and age be estimated from at least five independent sample within a stand and for trees that have at least 6 years of growth beyond breast height age.

Table 1. The available site index curves are listed by species and origin. The recommended equations are **bolded**. If there is only one reference, it is the curve used.

Species	Planted	Natural
White pine	**Sharma & Parton (2019) equation 1, table 2 no climate**	Parresol & Vissage (1998)
Red pine	**Sharma & Parton (2018b) equation 1, table 4 (no climate)**, Carmean & Thrower (1995)	**Buckmann et al. (2006)**
Jack pine	**Sharma et al. (2015) equation 1 (no climate)**, Guo and Wang (2006), Subedi & Sharma (2010)	**Sharma & Reid (2018), equation 3, table 4** Sharma (2021), Carmean et al. (2001) Goelz and Burk. (1998), Guo and Wang (2006)
White spruce	**Sharma & Parton (2018a) equation 1, table 2 (no climate)**	
Black spruce	**Sharma et al. (2015) equation 1 (no climate)**, Subedi & Sharma (2010)	**Sharma & Reid (2018), equation 3, table 4** Sharma (2021), Carmean et al. (2006)
Hemlock		Carmean et al. (1989) figure 127
Balsam fir		Carmean (1996) figure 18
Tamarack		Carmean (1996) figure 16
cedar		Carmean et al. (1989) figure 57
Sugar maple		Buda & Wang (2006)
Red maple		Carmean et al. (1989) figure 1
Yellow birch		Carmean et al. (1989) figure 6
White birch		Carmean (1996) figure 14
Poplar (all including Aspen, largetooth and balsam poplar)		Carmean et al. (2006), Sharma working on Po/Pj for Dec. 2021
White ash		Carmean et al. (1989) figure 13
Black ash		Carmean et al. (1989) figure 14
Red oak		Carmean et al. (1989) figure 48
Elm		Carmean et al. (1989) figure 53
Basswood		Carmean et al. (1989) figure 51
Beech		Carmean et al. (1989) figure 11
Black cherry		Carmean et al. (1989) figure 34
SI conversion	Carmean et al. (2013), Sharma (2021), working on Po,Pj for Dec 2021	
Northeastern US	Westfall et al. (2017)	

Appendix E – Implemented APF Forest Unit SQL

Order	FU	APF SQL Query	Implemented SQL Query[7]
1	PrCC	Pr>=.6	Pr >= 55
2	PwUS	WG=Pw or (WG=Pr and Pr <=.5) or Pr+Pw>=.4	WG=Pw or (WG=Pr and Pr < 55) or Pr+Pw>=35
3	PjCC	WG=Pj	WG=Pj
4	INTCC	(WG=Po or Bw) and Bw+Po+Pt+Pl+Pb>=.7	(WG=Po or Bw) and Bw+Po+Pt+Pl+Pb>=65
5	SbCC	WG=Sb and Sb>=.7 and Po+Bw+Pt+Pl+Pb<=.3 and Pw+Pr<=.3	WG=Sb and Sb>=65 and Po+Bw+Pt+Pl+Pb<35 and Pw+Pr<=35
6	LCUS	WG=Ce or La or Sb	WG=Ce or La or Sb
7	SFUS	WG=Bf or Sw	WG=Bf or Sw
8	OrUS	WG=Or	WG=Or
9	HeSEL	WG=He and He>=.4	WG=He and He>=35
10	HDUS	WG=(By or Ms or Ax) or (WG=Mh and (age<80 or STG<.4 or SC=3)) or (WG=OH and (age<80 or STG<.4 or SC=3))	WG=(By or Ms or Ax or Ab or Aw or Ag)
11	HDSEL	WG=Mh or OH	WG=Mh or OH
12	MWUS	remainder	(SW+PW+PR+CE+MH+BY+AW+CH+[OR]+OW+IW+BE+HE+IW)>=30

---

[7] Calibration plot information did not contain Site Class or stocking and was dropped from query