Spring 5-5-2023

Developing an Enhanced Forest Inventory in Maine Using Airborne Laser Scanning: The Role of Calibration Plot Design and Data Quality

Stephanie Willsey
University of Maine, stephaniewillsey90@gmail.com

DEVELOPING AN ENHANCED FOREST INVENTORY IN MAINE USING AIRBORNE LASER SCANNING: THE ROLES OF CALIBRATION PLOT DESIGN AND DATA QUALITY

By

Stephanie Kristin Willsey

B.A. Indiana University, 2017

M.F. University of Maine, 2021

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(in Forest Resources)

The Graduate School

The University of Maine

May 2023

Advisory Committee:

Daniel Hayes, Associate Professor of Remote Sensing and Spatial Analysis, Advisor

Aaron Weiskittel, Professor of Forest Biometrics and Modeling

Ivan Fernandez, Professor of Soil Science and Forest Resources

Shawn Fraver, Associate Professor of Forest Ecology

Ian Prior, Senior Inventory and Planning Manager at Seven Islands Land Company

DEVELOPING AN ENHANCED FOREST INVENTORY IN MAINE USING AIRBORNE LASER SCANNING: THE ROLES OF CALIBRATION PLOT DESIGN AND DATA QUALITY

By Stephanie Kristin Willsey

Thesis Advisor: Dr. Daniel Hayes

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Master of Science
(in Forest Resources)
May 2023

Forests provide essential ecosystem services such as carbon sequestration, clean water, and lumber, among others. Foresters need to know what is in the forest to manage the economic benefits and sustainability of the resource and for ecological soundness to prevent further degradation of these ecosystem services. Foresters must be able to conduct accurate and efficient forest inventories. Introducing remote sensing products in a traditional forest inventory presents novel opportunities to enhance forest characteristic predictions. These enhanced forest inventories (EFIs) can be done at varying resolutions based on plot size creating wall-to-wall raster data; thus, complete spatial knowledge of these estimates can be determined. Airborne Laser Scanning (ALS) has become a promising and commonly used remote sensing product that researchers and foresters apply to EFI workflows. Although best practice guides have been created in other countries, research on the best plot type and design has not been done for Maine's structurally diverse and intensively managed forests.

This study investigated a range of forest designs to determine the best ground-based calibration plot specifications for developing EFI models from ALS data in Maine. We conducted a study that compared fixed- versus variable-radius plots, sampling size and intensity, and sample design with ALS data to map EFI variables, including percent softwood, volume, BA, and tree count. These data were collected from the Penobscot Experimental Forest (PEF) in the

summer of 2022 using two plot types, sample sizes and sampling intensities, and two different sample designs. In addition, data from other study sites were provided to us by our partners that only included one plot type, sample size and intensity, and sample design. These datasets were used to cover a range of metrics not considered in the PEF (i.e., a grid sample design) and represent Maine's working forests in the North. For validation, we used data collected in the Demeritt Forest in the summer of 2022. We assessed model performance across forest inventories to investigate the best calibration plot design for ALS-based EFIs in our study areas. We determined that a principal component analysis for plot placement gave better model results than randomly placed plots. Also, fixed radius plots (FRPs) and a smaller sample size generated better evaluation statistics when predicting the percent softwood, volume, and tree count in the PEF. In contrast, variable radius plots (VRPs) with a smaller sample size provided better model outcomes when predicting basal area (BA). Once the best calibration plot design was identified and validated, we applied it to the PEF to estimate aboveground biomass.

We obtained mixed results, which may be due to different forest types. There was better model performance in spruce-fir forest types than in other forest types like oak-pine. Therefore more research must be done to ensure our potential recommendations are appropriate. Our results provide insights into an optimal approach for specific conditions and underscore the importance of future research to assist decision-making on plot type and sample design for the broad range of conditions on forested landscapes in Maine.

ACKNOWLEDGEMENTS

Thank you to everyone in the Wheatland Geospatial Lab: Dr. Daniel Hayes, my advisor, for his continuous support and guidance, David Sandilands for his patient teaching of R and collaboration, and Anthony Guay for his GIS expertise and sense of humor. I would also like to thank my committee, Dr. Aaron Weiskittel, Dr. Ivan Fernandez, Dr. Shawn Fraver, and Ian Prior, for their time and knowledge related to this research. I am grateful for my field collection crew Rissa Currie, Carly Fredericks, and David Ludwig as well as the University Forest Office crew who helped me collect inventory data for this project. Lastly, I would like to thank my friends and family who supported me unconditionally for the last several years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3DEP: 3D Elevation Program

ABA: area-based approach

ALS: airborne laser scanning

AGB: aboveground biomass

AOI: area of interest

BA: basal area

BAF: basal area factor

CDS/ac: cords per acre

CFI: continuous forest inventory

CHM: canopy height model

CRM: component ratio method

CV: coefficient of variation

DBH: diameter at breast height

DEM: digital elevation model

EFI: enhanced forest inventory

exts1: variable name for normalized difference vegetation index data used in our model

exts7: variable name for the near infrared band of the Sentinel-2 data used in our model

extsCV: variable name for the canopy cover data used in our model

FRP: fixed radius plot

G-LiHT: Goddard's LiDAR, Hyperspectral, and Thermal Airborne Imager

GPS: global positioning system

ikurt: intensity kurtosis (ALS metric)

isd: intensity standard deviation (ALS metric)

ITD: Individual Tree Detection

NASA: National Aeronautics and Space Administration

NDVI: normalized difference vegetation index

NLCD: national land cover database

OSR: overstory removal

PCA: principal component analysis

PEF: Penobscot Experimental Forest

PSW: percent softwood

NRMSE: root mean square error

SILC: Seven Islands Land Company

SRS: simple random sampling

SYS: systematic sampling

TPA: trees per acre

USGS: United States Geological Survey

VRP: variable radius plot

zmean: elevation mean (ALS metric)

zpcum3, zpcum8: cumulative percentage of returns located in the lower 30% and 80% of maximum elevation (ALS metrics)

zsd: elevation standard deviation (ALS metric)

zp5, zp20, zp25, zp45, zp85, zp90: elevation percentiles (ALS metrics)

CHAPTER 1: INTRODUCTION

1.1 Importance of ALS Data in Forestry

Forests are essential as they provide many ecosystem services, such as climate regulation, biomass fuel, fiber, and water cycling (Daily, 1997). Forest resource managers must exercise sustainable forestry practices so that society can continue to benefit from these valuable ecosystem services. Enhanced forest inventories (EFIs) potentially provide more accurate and precise estimates of forest characteristics than a traditional forest inventory that may help foresters conserve ecosystem services while also maintaining economic feasibility when managing a land base for timber.

A forest inventory can be completed in many ways, but these have historically occurred solely in the field (Nelson, 2014; Liang et al., 2016). Remote sensing is revolutionizing how forests are inventoried so that they can be completed more efficiently over time and space (Chen et al., 2019). The goal of using Airborne Laser Scanning (ALS) data in conjunction with forest inventories is to create accurate and updated maps of forest attributes so that foresters can potentially manage large areas of woodlands more sustainably (Schumacher et al., 2019; Ganz, 2019; Cheng & Wang, 2019). This process is accomplished by the time and cost efficiencies of cutting down on the number of plots used and using data at finer resolutions that can assist in many facets of forest management. Numerous researchers and forester practitioners in different countries have successfully conducted forest inventories using ALS. Some critical studies about operational ALS for EFI purposes are White et al., 2013 and 2017, Kangas et al., 2018, and Næsset et al., 2002.

Forest resource managers require field inventories to calibrate data collected from ALS to create models of forest attributes (Bolton et al., 2018). An EFI is a product of collecting field data and incorporating it with ALS, and other remote sensing estimates from which attributes such as tree height, volume, and basal area (BA) can be modeled (White et al., 2017). However, the study by White et al. (2013) recommends that forest resource managers consider implementing an area-based approach (ABA) for EFI techniques to be effective. Field data and ALS data must be collected for the ABA to model the known x variables (observed values) where ALS data exist. This allows for complete spatial knowledge of an area of interest as x and y (predicted) variables exist within the pixels of the raster output.

1.2 Forest Inventory Metrics

In recent decades, ALS data sets for the state of Maine have been made available to the public allowing forest landowners to utilize them in creating sustainable management plans. Even though these data are accessible to the public, some forest resource managers in Maine may be using a specific forest inventory design that does not provide the best results to suit their needs. In-situ data must be collected as accurately as possible to ensure proper correlation to ALS data; thus, acceptable forest attributes predictions. Research is needed to determine the optimal calibration plot design for generating the most accurate modeling results with an efficient data collection approach in Maine.

ALS technology is continually improving and becoming more available at a greater pace (Goetz & Dubayah, 2011; Alberdi, 2021). There is an increasing demand for improved methods of field inventory practices to complement the advancements in ALS technology (White et al., 2016). Many aspects of field-based inventory designs must be compared to determine which techniques are more effective for calibrating ALS data (White et al., 2013). These design specifications include decisions on fixed radius plots (FRP) versus variable radius plots (VRP), the appropriate sample size and intensity, and the sampling method.

It is impossible to have a complete census of timber resources in a forest because forest resource managers would need to sample every tree within the scope of their management plan. Time and monetary constraints mean that forest resource managers must decide the appropriate sample size and intensity to collect the most acceptable information for their needs. The sample size decision refers to the number of observations chosen from the population, whereas the sample intensity is the ratio of observations to the total population estimated (Kershaw et al., 2017). Increasing the sample size and intensity allows for complete information, but the tradeoffs include increasing time and money spent collecting data. Even though forest inventory efficiency can increase with the use of ALS, it can potentially plateau. It even decreases if the optimal sample size and intensity are not used (Lisańczuk et al., 2020). Maintaining high efficiency while not sacrificing time, money, and resources is a common goal shared by all forest resource managers (Kershaw et al., 2017).

Forest resource managers typically use FRPs to calibrate ALS data (Deo et al., 2016; Hayashi et al., 2014). This is the ideal case where the consistent areas of the plots can be coordinated with the spatial resolution of the ALS data and thus provide more accurate results (Deo et al., 2016). However, FRPs are expensive and time-consuming to implement, so foresters must make tradeoffs if using this plot type. Very few researchers have investigated using VRPs as calibration plots, even though there is a need to do so (Deo et al., 2016). Of the researchers who have explored this topic, some do not promote the use of VRPs as a type of calibration plot, citing a lack of research (White et al. 2017), while others have found that VRPs can be used under certain circumstances with specific BA factor plots (Deo et al., 2016).

Using a sampling design that fails to capture most of the variability in the forest may return unacceptable model predictions. Most forest resource managers tend to gravitate toward simple random sampling because it presents an unbiased estimate of the population mean (Kershaw et al., 2017). An additional technique employed for plot placement is the principal component analysis (PCA). PCA is a dimensionality reduction technique that condenses large data sets and highlights the variability within the data. PCA provides efficient differentiation in the variation of forest structure by identifying patterns in the point cloud data and emphasizing the similarities and differences (Zavyalov et al., 2009). By sampling across the major axis of the PCA, plot locations can then be assigned based on a targeted stratification of the overall variation in a multivariate data set, which will capture the range of forest parameters (Jolliffe & Cadima, 2016).

1.3 Applications in Maine

In North America, Canadian researchers in British Columbia have been at the forefront of these metric design questions and have attempted to explore how to create efficient calibration plots for their needs. From this research, they developed a set of recommendations for the area-based approach to EFI as a “best practices guide” (White et al., 2013). The guide aimed to direct foresters and others through gathering and processing ALS data, creating the best calibration plots, and modeling forest attributes to accurately estimate forest characteristics cost- and time-effectively over large areas (White et al., 2017). The best practices guide for Canada may have the potential to be applied to Maine. However, as the state has different forest types and conditions (Flatebo et al., 1999), further investigation is needed regarding its transferability to other forest regions. Maine sits in a transitional zone between the temperate forests to the south and the boreal forests to the north. This creates three unique ecotypes: the northern coniferous forests, the northern mixed-hardwood forests, and the oak/pine forests (McCaskill et al., 2016). Questions about the appropriate plot type, sampling size and intensity, and sample design are considered in this research to create a best practices guide for Maine.

1.4 Rationale and Significance

According to the Maine Forest Action Plan 2020, there are 18 million acres of forested land where small woodland owners manage approximately 5 million acres, approximately 10 million are managed by industry, and 3 million acres are permanently protected (Maine.gov, web). Both industrial and private landowners may see improved forest management by utilizing EFIs. As of 2019, ALS data have been publicly accessible for the entire state of Maine, which allows all forest landowners to utilize these valuable data to assist in creating sustainable management plans (CFRU, 2019). Foresters and landowners in Maine use remotely sensed data (i.e., ALS) to generate EFIs to guide their management decisions and have seen acceptable results. However, by determining best practices regarding plot type, sample size, and sample design, forest landowners may see improved accuracy in EFI predictions. Since Maine does not have a best practices guide for EFIs, forest inventory designs must be explored to determine the appropriate plot types for forest landowners to calibrate ALS data.

CHAPTER 2: DEVELOPING AN ENHANCED FOREST INVENTORY IN MAINE USING AIRBORNE LASER SCANNING: THE ROLES OF CALIBRATION PLOT DESIGN AND DATA QUALITY

2.1. Introduction

Forests are valuable resources because they provide various ecosystem services such as climate regulation, lumber, nutrient recycling, and other provisioning, regulating, or cultural supporting services (Pan et al., 2011). Disruptions of ecosystem services can occur because of several factors, including climate change-related impacts (Soucy et al., 2020), insects and pathogens (Busby & Canham, 2011), and land use conversion (Mauldin et al., 1999). To monitor changes and take stock of timber, foresters need to conduct inventories. These inventories help to evaluate forest characteristics such as tree species composition, growth, and mortality (Ma et al., 2021; Tewari, 2015; Hoover et al., 2020).

There are many forms a forest inventory can take, which hinge on time and monetary constraints (Köhl & Magnussen, 2016), the inventory's purpose, and the forest's size (Scott & Gove, 2002). Historically, these inventories have taken place on the ground, including measurements of physical characteristics and timber assessments (Krug & dos Santos, 2004). A cost- and time-effective plot type typically employed is a variable radius plot (VRP). Each tree has its own plot radius factor which determines whether it is tallied. This is a quicker way to collect data, while still maintaining the efficacy of the inventory, compared to measuring trees with a fixed radius plot (FRP) where all trees within the plot radius are tallied. There are also different sample designs that can be utilized. Biases are associated with each design, where the forester must determine which design is most time and cost-efficient for the inventory. While forest inventories are a valuable tool, limitations can hinder reliability. Didion et al. (2009) found that conducting inventories in intensively managed stands can negatively impact forest metric predictions because of the structural variability that is not sufficiently captured. There are also limitations related to a geographical extent. Not all forests are the same size, and larger forest parcels require more plots to cover large expanses of land, which can be costly and time expensive. Lastly, repeat inventories are usually conducted every 5-10 years or more (Yu et al., 2022). Tree growth and mortality that occur during this period may prove a previous forest inventory inaccurate leading to ill-informed management. Many foresters and researchers are turning to remote sensing to fill the gaps in a traditional forest inventory.

Traditional forest inventories, when coupled with remote sensing products, can potentially enhance estimates of forest metrics like BA, stem density, volume, and aboveground biomass (Sheridan et al., 2015; Iqbal et al., 2019). Interest in enhanced forest inventories (EFIs) continues to grow as remote sensing datasets become more reliable (Goodbody et al., 2019). Remote sensing products can generate wall-to-wall forest inventories of an area of interest. This is accomplished by feeding a model in-situ and remote sensing data to predict forest metrics in areas without ground plot data (Waser et al., 2015). Additionally, remote sensing products can cover larger geographical areas than traditional forest inventories. Lastly, the temporal resolution is finer than most traditional forest inventories, as some satellites have a return interval of 16 days (USGS Landsat Missions, web). While most satellite imagery provides adequate detail about an area of interest for landowners and some researchers, a limitation is its need for more ability to characterize the 3D structure of the forest.

Airborne laser scanning (ALS) is a remote sensing product that can measure the 3D structure of an object or surface, including forest canopies (White et al., 2016). A sensor attached to an aircraft measures the distance to objects using an emitted laser pulse as the aircraft is flown over an area of interest. The returns from the laser, along with the associated distance measurements, are then used to create a 3D point cloud. A point cloud of a forested area can be used for individual tree measurement (Weiser et al., 2022). Individual tree detection (ITD) is a useful tool that allows a user to extract tree height information without measuring each tree in the field (Weinstein et al., 2021). Due to the high computational requirements of ITD, this technique is most useful when applied to a small region of interest (Sparks & Smith, 2022; Jeronimo et al., 2018), whereas the area-based approach (ABA) may be applied across landscapes (White et al., 2017). After collecting an inventory, missing information about forest characteristics can be predicted from ABA models that use calibrated ALS data (Brosofske et al., 2014).

The ABA tessellates the area of interest and uses in-situ and ALS data wherein predictive equations are applied in each grid cell that aids in the creation of wall-to-wall predictions. This is the recommended approach to predicting forest characteristics for several reasons (White et al., 2013; Næsset et al., 2002; Woods et al., 2011). The most obvious advantage is the ability of users to have complete spatial knowledge of predictions, unlike traditional forest inventories that are not spatially explicit (White et al., 2013). Second, the ABA has been shown to offer more precise predictions (Coomes et al., 2017). Lastly, users can calculate confidence intervals for forest characteristic estimates (Woods et al., 2011).

The ABA for generating EFIs has shown promising results in different forest types in different countries (Table 2.1). White et al. (2013) suggested using the ABA for attribute estimation in Canada using a best practices guide. In Norway, Næsset (2005) found that the ABA provided sufficient forest characteristic estimations in an operational setting in a mixed wood forest. Frank et al. (2020) demonstrated that the ABA approach proved superior to semi-individual tree crown models in a mixed species-environment. In Southeast Asia, Coomes et al. (2017) determined that the ABA models perform better than tree-centric approaches in the lowland tropical rainforest. The ABA approach works well in different forest types, and mixed-species stands. However, the question remains about how well it performs in Maine's mixed-species, structurally complex, and intensively managed forests.

Table 2.1. Summary of studies done in four different countries using the area-based approach to estimate forest attributes and the respective results.

Study	Country	Forest Type	Results
White et al. (2013)	Canada	Boreal	Following a best practices guide created for Canada's forests, the ABA approach has the potential to provide accurate forest attribute estimations.
Næsset (2005)	Norway	Boreal	Under operational conditions, the ABA offered sufficiently better results than a traditional forest inventory.
Frank et al. (2020)	United States	Temperate Deciduous	ABA models had smaller residual variance and smaller random effect variance estimates than semi-individual tree crown models
Coomes et al. (2017)	Southeast Asia	Lowland Tropical Rainforest	ABA provided better results than a tree-centric approach when estimating forest characteristic predictions.

This research aimed to compare two plot types, various sample sizes and sampling intensities, and sample designs for predicting forest attributes. The two plot types considered were fixed radius plots (FRPs) and variable radius plots (VRPs). A typical plot type used with remotely sensed data is FRPs, where all trees (usually above a specified height threshold) are measured within a particular radius. FRPs cost- and time- inefficient. However, VRPs are a satisfactory solution to overcome those obstacles. This plot type requires that only a fraction of the trees be measured based on the basal area of the trees. Because VRPs are not circular (with an unknown shape and area), errors can be introduced into the EFI models (Deo et al., 2016). While research exists where VRPs are considered for use with remotely sensed products (Deo et al., 2016; Hayashi et al., 2014), more research should be done using VRPs for the ABA. We predict that FRPs will provide the best model performance since these plots tend to mitigate errors associated with edge effects.

The second consideration of this study was the sample size and sampling intensity. Sample size refers to the number of plots in the study site that is measured, whereas the sampling intensity is the proportion of the plots' area to the study site's area. Since VRPs have an unknown shape and area, the sampling intensity cannot be calculated, so sampling intensity in this study is only available for FRPs. Usually, a higher sample size captures more population variability, meaning better model performance. However, foresters face tradeoffs when considering their area of interest's sample size and sampling intensity. Plot implementation costs time and money, but foresters must be as accurate as possible (Fassnacht et al., 2014). Knowing that foresters face time and monetary constraints, our study tested different sample sizes and intensities to determine if smaller sample sizes return accurate and precise predictions. Because we used random forest as our predictive model (Breiman, 2001), we predict that smaller sample sizes and sampling intensities will provide better model results.

Lastly, we created models using three different sample designs – a simple random sample (SRS) design, a systematic sample (SYS) design, and a principal component analysis (PCA) for plot placement. An SRS is when plots are placed randomly, whereas an SYS places plots in a grid pattern. A PCA for plot placement uses ALS data to determine variability in the forest structure. More plots were placed in areas with more variability, and fewer plots were placed in areas with less variability. White et al. (2013) suggest using a structurally guided sample design, like a PCA for plot placement, but offer no comparisons to designs most used by foresters like an SRS and SYS. Although these sample designs were tested for comparison, we predict that a PCA for plot placement will return better results than the other two sample designs.

ALS is revolutionizing the way foresters conduct forest inventories. Research on in-situ data type and ALS data quality has been extensively conducted in areas like Norway (Næsset et al., 2004 & 2007) and Canada (White et al., 2013 & 2017) but may have limited applicability in Maine's forests. This study aimed to identify and demonstrate best practices concerning calibration plot design and data quality for foresters using ALS-based EFIs in Maine.

2.2. Methodology

2.2.1 Study Areas

Our study sites are in various regions across Maine (Figure 2.1) and include the Penobscot Experimental Forest (PEF) and the Demeritt Forest, as well as properties owned or managed by Baskahegan Company and Seven Islands Land Company (SILC). Maine consists of three distinct forest ecotypes including the (1) northern coniferous forests dominated by spruce-fir, (2) northern mixed-hardwood forests dominated by beech-birch-maple, and (3) oak-pine forests (McCaskill et al., 2016). Most of our study sites lie in the spruce-fir forest type, although the oak-pine forest type encroaches on the areas near the PEF and Demeritt Forests. There are inter-species size and structural differences and intra-species differences - especially in the spruce-fir forests where many industrial forestry companies manage land commercially.

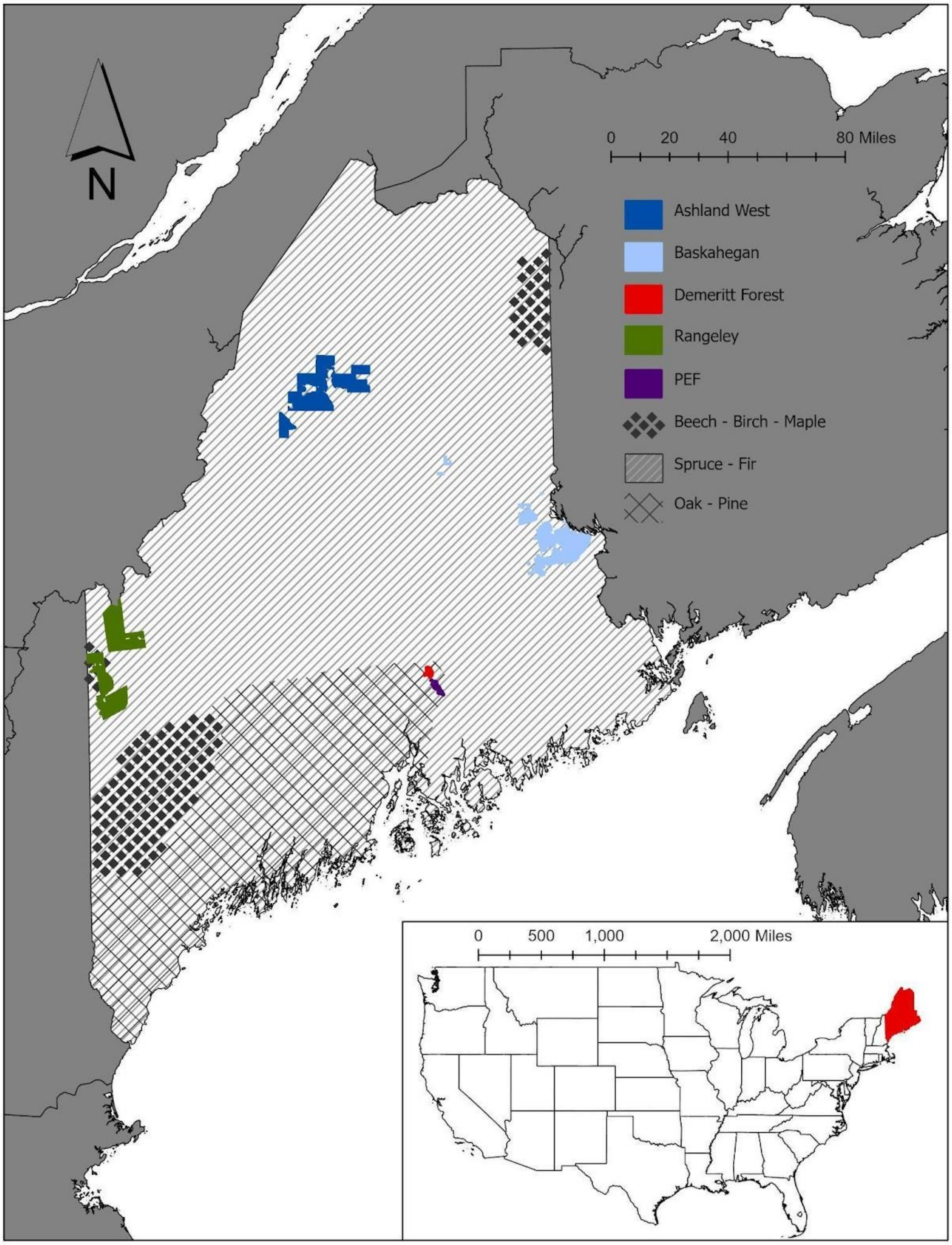


Figure 2.1. Our sample domain includes areas of interest owned by several forestry companies and university land holdings across the state of Maine. Maine comprises three distinct forest ecotypes (McCaskill et al., 2016). Spruce-fir dominates Northern Maine whereas the northern hardwoods and oak/pine forests are found in Southern Maine.

2.2.2 Data

2.2.2.1 In-Situ Data

Our partners at SILC and Baskahegan provided plot data in western, northern, and eastern Maine, while data was collected from the PEF and Demeritt in the summer of 2022 (Table 2.2). SILC supplied us with two datasets: one in Rangely and one in Ashland West. There are 384 circular FRPs in the Rangely site with a sampling intensity of 0.364%. The sample design used was simple random sampling (SRS). The Ashland West site had 1,387 VRPs using principal component analysis (PCA) for plot placement. A PCA for plot placement works by inputting ALS data and the number of plots needed or wanted for analysis. The PCA will use the ALS data to assess the structural variability of the area and will determine where plots should be placed to capture more of that variability. Sampling intensity could not be determined for this site because each VRP has a unique and unknown shape and area. Baskahegan used 183 FRPs using a systematic sample (SYS) design with a sampling intensity of around 0.012%.

We collected data in the PEF in 2022 to compare the different plot types, sample sizes, and sample designs in one location. A PCA for plot placement was used to determine FRP and VRP placement for 92 and 143 points. FRPs were 32.8 feet in radius, and sampling intensity was 0.304% and 0.472%, respectively. We also determined the plot radius factor for each tree to create variable radius data using a nine basal area factor (BAF) prism. An SRS design had been used previously in the PEF for a continuous forest inventory (CFI) in which there were 92 FRPs and each plot had a radius of 37.2 feet. We grew the CFI data using the USFS Forest Vegetation Simulator (FVS) Acadian Variant (Weiskittel & Kershaw, 2018) to 2022 and clipped the stem map to a 32.8-foot radius. A sample size of 143 random FRPs was generated by combining the existing 92 CFI FRPs with the 143 PCA FRPs and running a random number generator to extract 143 data points. The plot radius factor for all trees within the 92 and 143 FRPs in the SRS design was calculated to create variable radius data using a 9 BAF prism.

Because of the proximity of the Demeritt to the PEF, we were able to collect data there as well in 2022 to be used as validation. There were 53 plot locations with a sampling intensity of 0.359% in the Demeritt Forest. These data were collected using 1/10-acre plots. Calculations were done post-cruising to create 32.8-foot radius FRPs and to determine the plot radius factor for VRPs using a 9 BAF prism. Tree measurements collected in the PEF and Demeritt were species, diameter at breast height (DBH), tree height, crown class, azimuth, and distance. In each plot, only trees greater than four inches DBH were measured. In the PEF, tree heights were measured for one tree in each DBH class in each plot, while in the Demeritt, every tenth tree was measured for height in each plot. For all the VRP data, we could not determine sampling intensity because each plot has a different area. The differing plot sizes make it difficult to calculate sampling intensity (Keene & Barlow, 2019).

Table 2.2. Summary of plot types, sample size, sample intensity, and sample design of our areas of interest.

Study Area	Plot Type	Sample Size	Sample Intensity	Sample Design
PEF	FRP	143	0.472%	PCA
	FRP	92	0.304%	Random
PEF	VRP	143	N/A	PCA
	VRP	92	N/A	Random
Demeritt	FRP	53	0.359%	Random
Ashland West (SILC)	VRP	1,387	N/A	PCA
Rangely (SILC)	FRP	384	0.364%	Random
Baskahegan	FRP	183	0.009%	Grid

2.2.2.2 ALS Data

We used NASA's Goddard's LiDAR, Hyperspectral & Thermal Imager (G-LiHT) data for the Demeritt Forest and PEF (Table 2.3). G-LiHT combines ALS, imaging spectroscopy, and thermal measurements that can map the structure and composition of forests (Cook et al., 2013). G-LiHT LiDAR data have a high point density, and the sensor is flown during leaf-on conditions, which may allow for more accurate forest metric predictions (Næsset, 2005; Imangholiloo et al., 2020; Bouvier et al., 2015).

We used USGS's 3D Elevation Program (3DEP) data for all other study areas. This program aims to collect and make publicly available ALS datasets for the United States with high-resolution elevation data (USGS, n.d.). 3DEP datasets are available for Maine and classified as Quality Level 2 (QL2). The QL2 dataset has a slightly lower point density than QL0 or QL1 datasets (USGS, n.d.). Because the goal is to provide elevation data, ALS acquisition is during leaf-off events to ensure more ground returns. Both G-LiHT and 3DEP datasets were normalized by calculating height above the ground.

Table 2.3. Summary of ALS data and associated characteristics for each study area.

Study Area	ALS Data	Point Density (pls/m^2)	Temporal Resolution	Land Characteristics
PEF	G-LiHT	12-15	2021	Leaf-on
Demeritt	G-LiHT	12-15	2021	Leaf-on
Ashland West	3DEP	2-3	2017	Leaf-off
Rangely	3DEP	2-3	2018	Leaf-off
Baskahegan	3DEP	2-3	2017	Leaf-off

2.2.2.3 Covariate Data

In addition to the ALS data, we used Sentinel-2 multispectral data as covariates in our models (Table 2.4). Sentinel-2 data less than 30% cloud cover for 2021 were downloaded from USGS's Earth Explorer before the archive ended. These data were used to create a Normalized Difference Vegetation Index (NDVI), a Principal Component Analysis composite containing principal components 1 and 2, which explained 94.24 and 5.76% percent of the variance, respectively, and a four-band composite of the blue, green, red, and near-infrared bands at a 10-meter resolution to assist the model in vegetation recognition. A digital elevation model (DEM) at a 10-meter resolution was created using an ALS dataset corresponding to our study areas. We filtered for the last returns to create the DEM from which we calculated slope and aspect and added them as model covariates for all study areas except the PEF and Demeritt Forest. Slope and aspect were not used for these two study sites because of the flat topography of these areas.

Furthermore, we created a canopy cover dataset by filtering noise and points over 164 feet in the ALS dataset, and then a percentage of all returns over the mean height was calculated. This dataset was used to describe the canopy structure and was calculated at a 10-meter resolution. Lastly, we used various standard height and intensity metrics within the ALS dataset, as described in Table 2.5 below. Height metrics are calculated from the distance of each point from the ground (Hütt et al., 2022), whereas intensity metrics measure the laser pulse's return strength for each point (Kashani et al., 2015). These metrics were subset using the lidR package in R (Roussel et al., 2020; Roussel & Auty, 2023).

Table 2.4. Datasets and the respective products created from them.

Dataset	Products Created	Spatial Resolution (meters)
Sentinel-2	NDVI (exts1)	10
	PCA (exts2, exts3)	10
	4 Band Composite (B, G, R, NIR) (exts4, exts5, exts6, exts7)	10
*ALS	Canopy Cover (extsCV)	10
	Slope (ext_slope)	10
	Aspect (ext_aspect)	10
	Height Metrics	10
	Intensity Metrics	10

*G-LiHT or 3DEP were used to derive certain datasets depending on the study area. Slope and Aspect were not used in the models for the PEF or Demeritt because the topography for these areas is homogenous, so the elevation factors were not important for model performance.

Table 2.5 ALS metrics used in this study and a description.

Metric Name	Description
zmax	elevation maximum
zmean	elevation minimum
zsd	elevation standard deviation
zskew	elevation skewness
zkurt	elevation kurtosis
zentropy	elevation entropy of returns
zq5, zq10, …, zq90, zq95	elevation percentiles
zpcum1, zpcum2, …, zpcum8,	cumulative percentage of returns in each layer of the
zpcum9	elevation range
imax	intensity maximum
imean	intensity mean
isd	intensity standard deviation
ikurt	intensity kurtosis

2.2.3 Data Analysis

We used the area-based approach (ABA) for modeling (Figure 2.2). According to White et al. (2013), the ABA consists of two stages. In stage one, plot data are clipped to ALS data. From the clipped data, standard ALS metrics were subset that included mean height, height percentiles, and maximum intensity, among other metrics (Roussel et al., 2020; Roussel & Auty, 2023). Then predictive models were developed where in-situ data are the response variable, and predictors are the ALS metrics. We employed random forest for our predictive models, which used the clipped ALS metrics as the training data. From the random forest model, we could determine which predictor variables were most important or which variables the model used the most to determine the predictions. In stage two, wall-to-wall estimates were generated for the whole area of interest (Figure A.5). The resulting grid cell, or pixel size, is related to the size of the plot. We used 32.8-feet radius plots for FRP calibration, so the pixel size is 65.6 feet. The prediction equation created from the model in stage one is applied within the grid cells. Once applied, each pixel will contain a forest attribute prediction. This information can then be expanded to stand-level predictions (White et al., 2013; White et al., 2017).

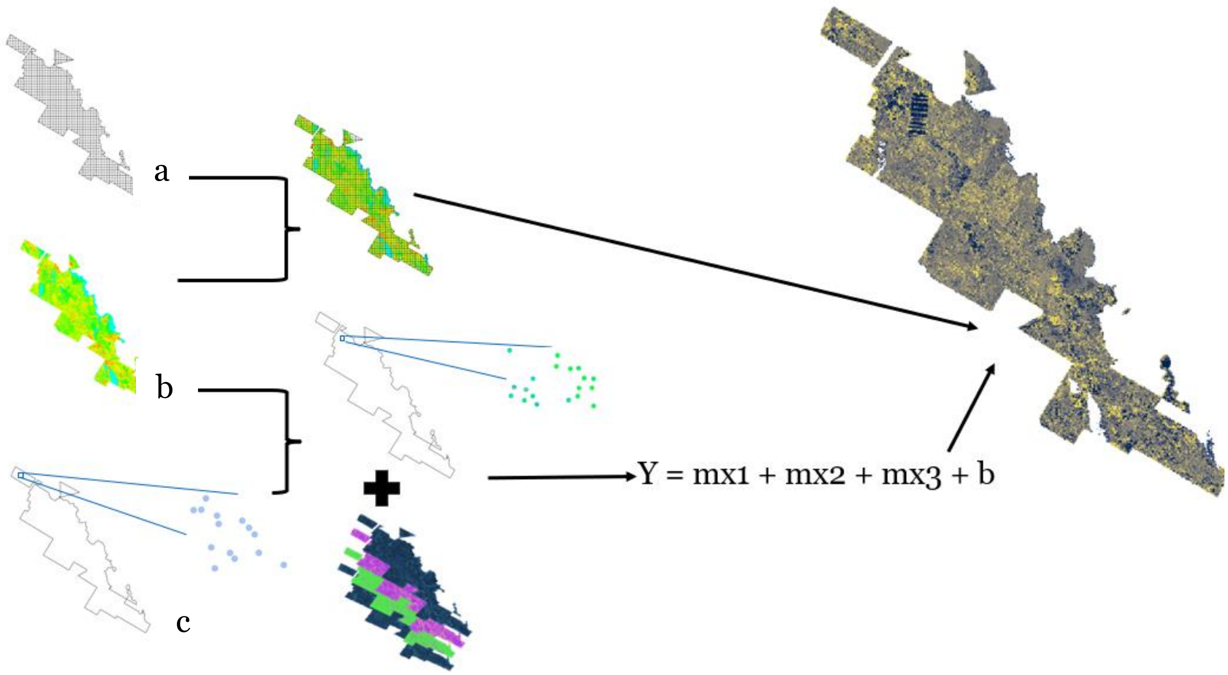


Figure 2.2. The ABA consists of several steps which are (a) tessellating the area of interest based on plot size, (b) gathering ALS data (c) gathering ground plot data, (d) generating wall-to-wall metrics, (e) clipping ALS data to plots, (f) compiling covariate data, (g) applying the predictive model, and finally (h) generating wall-to-wall maps.

We used random forests for the predictive model mentioned in stage one of the ABA. The random forests modeling approach allows a group of trees to “grow” by voting for the most popular decision (Breiman, 2001). Our clipped ALS data were used as training data - the model used the training data without replacement to grow the groups of trees. This is known as bagging (Breiman, 1996).

2.3. Results

The best results in terms of model performance consistently came from the percent softwood (PSW) predictions. The Baskahegan model had the most considerable R^2 value out of all the other models (Table A.1). This indicates that there is a stronger correlation between the predicted and observed percent softwood values than that of the other models (Figure A.1). The normalized root mean square error (NRMSE) metric of 0.3% shows that the model predictions are closer to the observed values on average than with the other models. The low coefficient of variation (CV) value indicates a higher precision in the predicted values, and the normalized mean bias (NMB) demonstrates that the error was underestimated but small. The PEF model using 143 FRPs with an SRS design was the lowest model performance in predicting PSW (Table 2.6). The R^2 shows that the PSW predictions and observed values do not fit the regression as well as with the Baskahegan model. The higher NRMSE indicates that the predictions are further away from the fitted line. The higher CV demonstrates a lower precision in the model's predictions, and the NMB shows that the error was overestimated and is substantial comparatively. Of all the PEF models, the best-performing one used 92 FRPs and a PCA for plot placement. None of the models performed terribly in predicting PSW and can be attributed to the addition of covariate data. The added imagery assisted the models by aiding the distinction between hardwoods and softwoods and can be surmised from the three most important variables from all the models. The important imagery variables include Sentinel-2 band 8 (near infrared) (exts7), NDVI (exts1), intensity kurtosis (ikurt), and standard deviation (isd) (Figure 2.3).

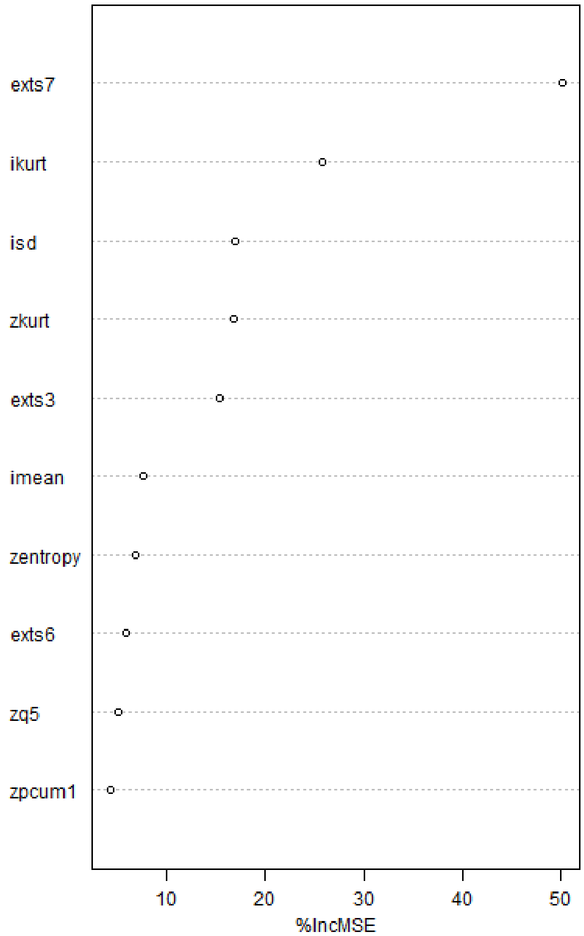


Figure 2.3. The ten most important percent softwood predictor variables for the PEF model using 92 FRPs and a PCA for plot placement.

Model performance across all study sites begins to decline when predicting volume. The Baskahegan model showed the best performance for this attribute. The evaluation metrics indicate that the predicted volume values are closer to the observed volume and are more precise than the worst-performing model - the PEF model using 143 FRPs with an SYS design (Figure A.2). Within the PEF, the best-performing model was the one that used 92 FRPs and an SYS design. The most important variables in predicting volume were consistently canopy cover (extsCV), elevation in the fifth percentile (zp5), and the cumulative percentage of returns in the lower 30% of the maximum elevation (zpcum3) (Figure 2.4). The decline in model performance continues with the prediction of BA. The Baskahegan model, once again, was the best-performing model compared to the PEF model using 143 FRPs and an SYS design. The evaluation statistics show that the predicted BA values are closer to the observed values and more precise (Figure A.3). Compared to the models in the PEF, the best evaluation statistics came from the one using 92 VRPs and PCA for plot placement. The topmost important variables in predicting BA were consistently zp5, exts7, and extsCV (Figure 2.4). Since volume and BA are somewhat related, overlapping importance variables are to be expected.

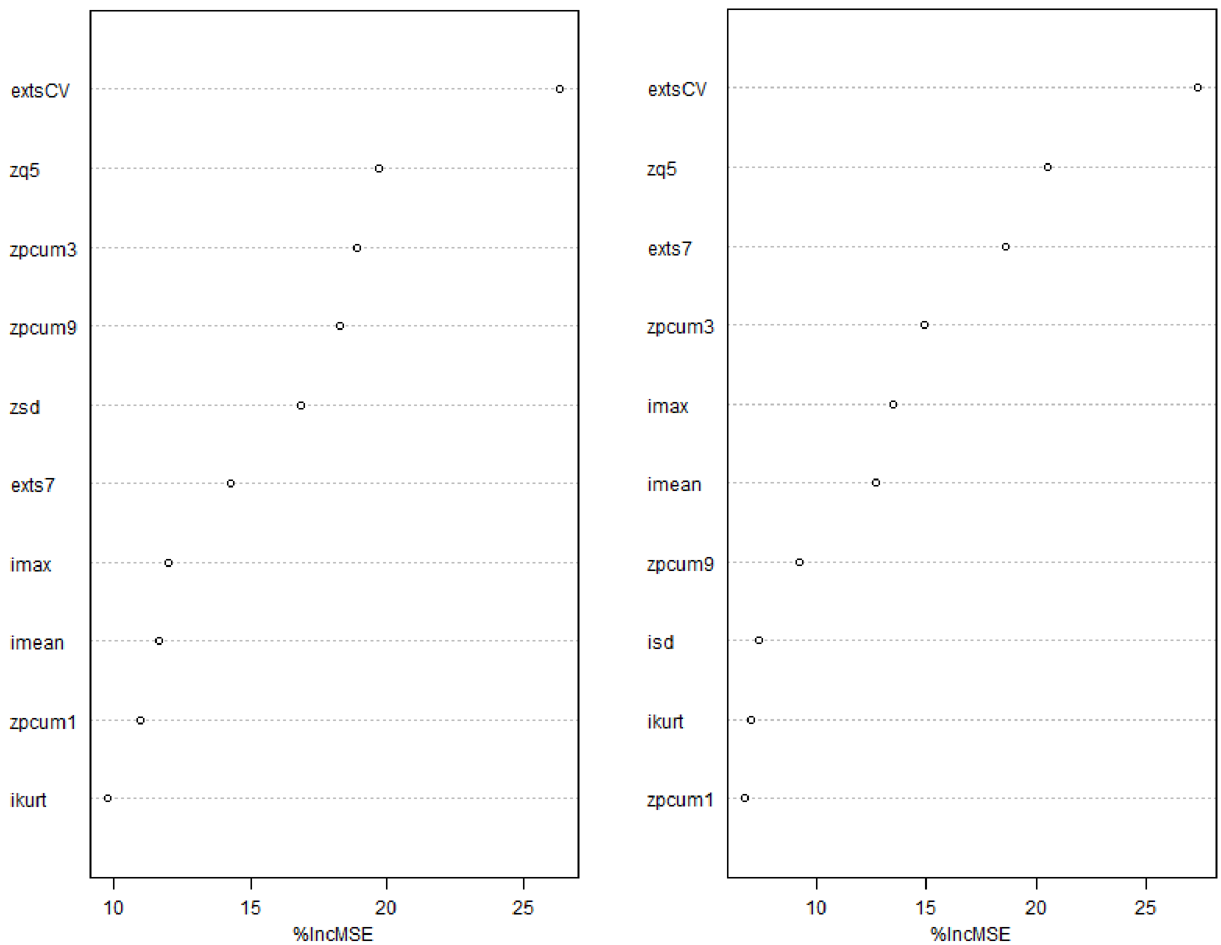


Figure 2.4. The ten most important volume (left) and basal area (right) predictor variables for the PEF model using 143 FRPs (left) and 92 FRPs (right) both with a SYS design.

Tree count predictions consistently underperformed across all sites, with the highest R2 coming from the PEF model using 92 FRPs and a PCA for plot placement (Figure A.4). The NRMSE shows that the range in predictions from this model are closer to the range of in-situ data compared to the other models. The NMB from this model indicates that the error was underestimated, but there was no significant error. This contrasts with the model using 143 VRPs and an SYS design in the PEF and the Ashland West model, which showed the worst performance in predicting tree count. The R^2 demonstrates that the predicted TPA values are not close to the observed TPA values, while the higher NRMSE indicates that the predictions lie further away from the fitted line. The high CV indicates that the predictions were less precise than other models. The most important variables that help predict tree count were zq5, the standard deviation of the elevation (zsd), and the standard deviation of the intensity (isd) (Figure 2.5). The most notable difference between the two models is the type of ALS data used for the analysis. G-LiHT data (leaf-on) was used for the PEF models, while 3DEP data (leaf-off) was used for the Ashland West model. The quality of ALS data may affect tree count predictions.

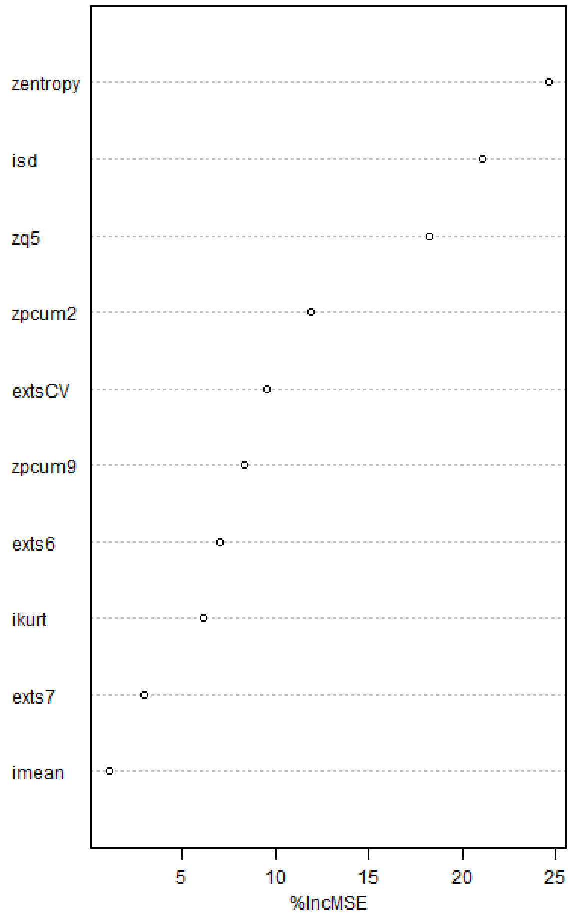


Figure 2.5. The ten most important tree count predictor variables for the PEF model using 92 FRPs.

Even with a large sample size, the Ashland West model (using VRPs) performed the poorest across all forest attribute predictions compared to the Rangeley and Baskahegan sites which used FRPs. The high NRMSE value of the Ashland West model demonstrates that the model was over-predicting, while the higher CV value means that the predictions could be more precise than the other two models. In the PEF, the models using FRPs slightly outperformed the models using VRPs with the same sample size across most of the calculated forest metrics. Comparing the models using a sample size of 92, the FRP types see a better model fit, usually better prediction values, usually more precise predictions, and smaller errors. Comparing the models using a sample size of 143, again, FRP types show better model performance. However, the model using 92 VRPs did perform better than the model using 143 FRPs for the PSW, volume, and BA predictions.

Sample sizes, hinted at in the previous paragraph, can also provide insight into model performance. Sample sizes range from large (n = 1,387) to small (n = 92). The Ashland West model with the largest sample size seemed to perform the worst compared to the Rangeley and Baskahegan models – even though the sample size was four to nine times larger than the other two sites. Comparing the four PEF models, the smaller sample size showed better model performance regardless of plot type. This may be because of how our random forest model worked. Our random forest model's default was bootstrapping, meaning all our observations were sampled with replacement (Breiman, 1996). Breiman (2001) explained that when the sample size is decreased, there are fewer correlations between trees. A lower correlation between trees means a potential improvement in prediction accuracy. However, a smaller sample size does not necessarily imply that more variability is captured in a land base.

Lastly, the models can be compared by sample designs. Half of the PEF models and the Ashland West model used a PCA for plot placement, whereas the other sites used either an SRS or an SYS (Table 2.2). Of the four PEF models, the two that performed slightly worse in predicting PSW and BA were the ones that used a PCA for plot placement. The R^2 for these two models is slightly worse, indicating that the predicted values were not as close to the observed values than the other models that used an SRS. The NRMSE, CV, and NMB values for the worst-performing PEF models also show that the prediction values are not as precise or accurate as the others. The PEF models that performed worse in predicting volume were the models that used 143 VRPs (PCA) and 92 VRPs (random).

Furthermore, the models that performed worse in predicting tree count were the ones that used 92 VRPs and 143 FRPs (PCA). Comparing across all models, the Baskahegan (PCA) model performed the best across all sites for PSW, volume, and BA predictions. The worst-performing model was the PEF (143 VRPs, PCA) in predicting PSW, volume, and BA and the Ashland West model in predicting tree count.

Table 2.6. R^2, normalized root mean square error (NRMSE), coefficient of variation (CV), and normalized mean bias (NMB) results for the PEF for each of the four forest metrics predicted as well as both plot types and sample sizes.

Plot Type	Sample Size (n)	Percent Softwood (%)				Total Volume (CDS/ac)				Basal Area (ft²/ac)				Tree Count (trees/ac)			
		R^2 (%)	NRMSE (%)	CV (%)	NMB (%)	R^2 (%)	NRMSE (%)	CV (%)	NMB (%)	R^2 (%)	NRMSE (%)	CV (%)	NMB (%)	R^2 (%)	NRMSE (%)	CV (%)	NMB (%)
FRP	92	72.7	47.1	31.0	-0.18	31.9	11.4	57.7	-1.74	38.4	21.6	45.4	-1.34	39.8	46.4	33.3	-1.27
FRP	143	59.1	41.5	38.6	0.89	35.8	36.0	37.9	-1.16	25.2	33.5	34.2	-0.94	26.3	51.0	49.7	-1.58
VRP	92	61.7	43.3	38.9	1.79	40.1	48.4	38.7	-0.23	61.3	22.1	52.8	-1.06	22.2	53.4	52.3	-0.61
VRP	143	58.5	40.6	39.8	0.89	30.7	37.2	39.3	-0.92	22.5	34.1	35.2	-1.02	27.7	50.5	49.6	-1.80
FRP	92	64.6	36.0	37.9	0.83	57.0	18.0	41.7	-0.54	40.7	21.3	33.0	0.94	32.5	34.1	32.6	-0.11
FRP	143	50.8	38.1	40.6	0.12	22.0	12.4	69.3	-2.23	22.0	43.3	57.7	-1.68	27.0	36.3	32.2	-0.42
VRP	92	63.3	42.8	39.6	1.07	33.3	34.2	39.3	-0.07	28.7	34.5	32.6	-0.13	21.4	52.0	46.1	-0.98
VRP	143	54.2	44.0	41.1	1.77	22.7	12.3	49.8	-0.06	35.1	12.3	77.7	-0.35	12.4	45.2	57.2	-0.65

*The first four rows of information reflect the PCA for plot placement while the last four rows of information reflect the simple random sampling design.

2.3.1. Validation

A model was created for the Demeritt forest, and four PEF FRP models were applied to this study site as a comparison. The created model data was prepared using G-LiHT data to separate the Demeritt into 64 PCA classes and assumed we needed 53 plots. Existing CFI plots were then used and categorized into each PCA class. Some classes did not need plots as these sites were devoid of trees, such as non-forested wetlands, so no plots were placed in those categories. Once plot placement had been determined, the stem map could be clipped to 32.8-foot radius plots. Plot metrics (i.e., PSW, volume, BA, and tree count) were calculated and fed into the model along with covariate data, which produced a raster dataset (Figure A.6). Results showed mediocre performance in predicting PSW – even with the addition of covariate data that was assumed to have aided PSW predictions in other models (Table 2.6). Although the statistics for the PSW predictions were poor, that was not the case for the other predictions. Volume and BA saw an improved model performance, either on par or better than other models. Tree count predictions also showed good model performance, comparatively. A visualization of how well the predicted values compared to the observed values can be seen below in Figure 2.6, and the results are shown in Table 2.7.

Table 2.7. R^2, NRMSE, CV, and NMB results for the Demeritt model for each of the four predicted forest metrics.

Plot Type	Sample Size (n)	Percent Softwood (%)				Total Volume (CDS/ac)				Basal Area (ft^2/ac)				Tree Count (trees/ac)			
		R^2 (%)	NRMSE (%)	CV (%)	NMB (%)	R^2 (%)	NRMSE (%)	CV (%)	NMB (%)	R^2 (%)	NRMSE (%)	CV (%)	NMB (%)	R^2 (%)	NRMSE (%)	CV (%)	NMB (%)
FRP	53	28.3	0.2	29.4	0.47	52.2	17.3	38.6	0.45	74.7	0.7	30.0	-0.09	41.4	57.4	37.3	-0.71

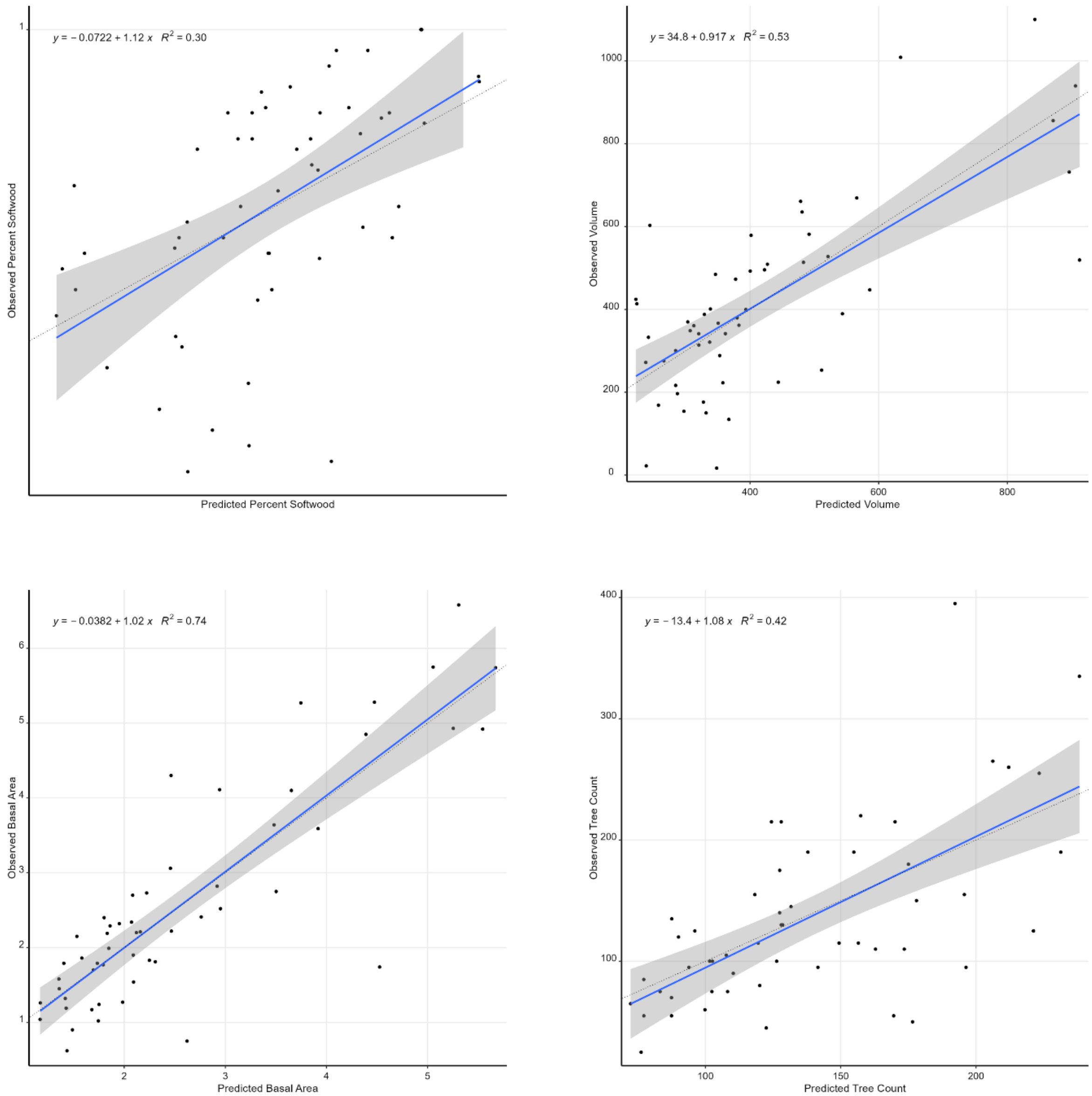


Figure 2.6. Plotted relationships between predicted (x-axis) and observed values (y-axis) for percent softwood, volume, basal area, and tree count predictions in the Demeritt Forest.

Looking at the top three most important variables for the PSW model (Figure 2.7), exts7, ikurt, and isd are the most used in the predictions. These are the same variables used in the PEF (92 FRPs) model to predict PSW (Figure 2.3). The top three variables for predicting volume were zq20, zq25, and zmean. For BA in the Demeritt, the top three variables were zq85, zq90, and zq45. Lastly, Demeritt's top three predictor variables for tree count were zpcum1, zpcum8, and zsd. These variables do not deviate from those used to predict the different forest metrics in the other models.

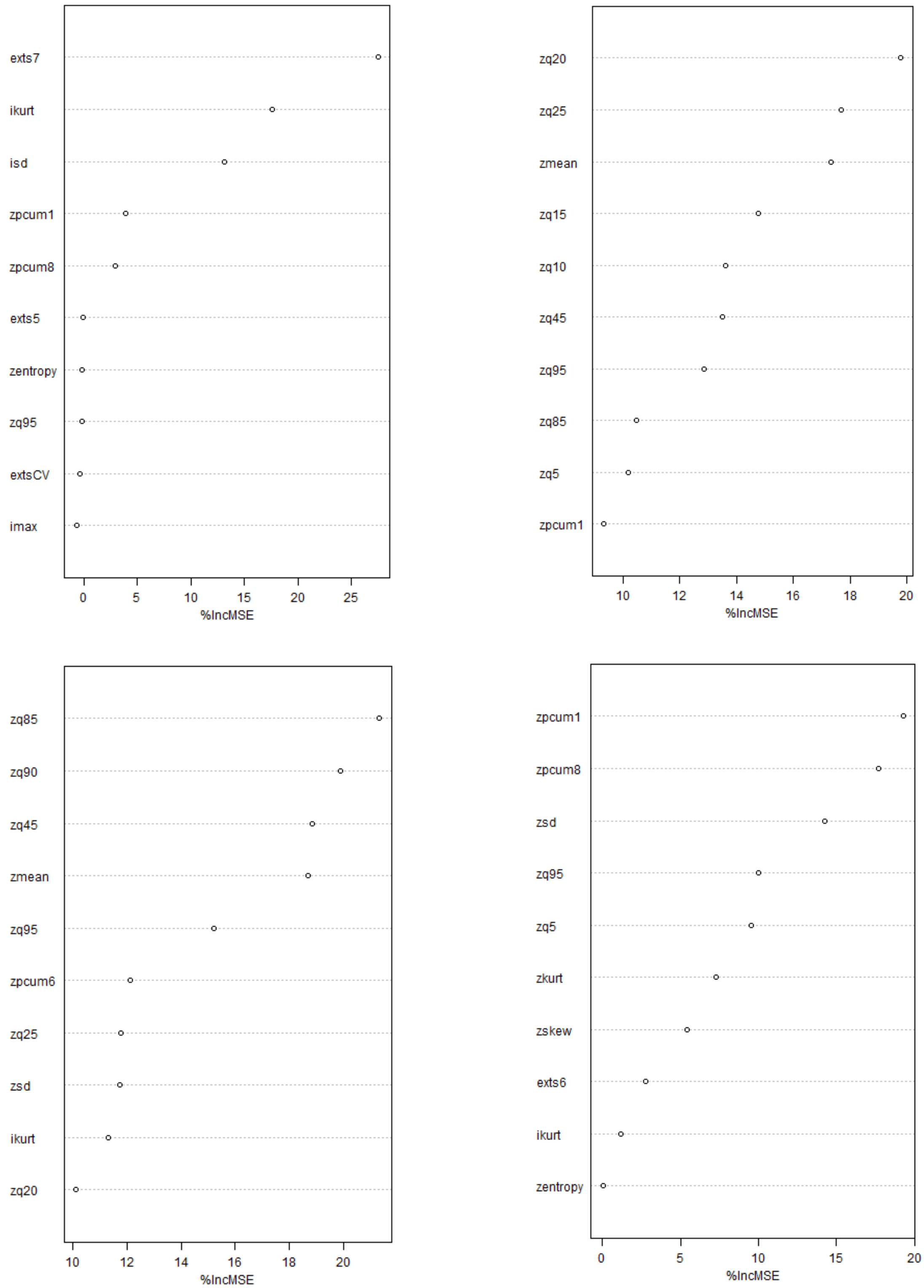


Figure 2.7. The ten most important predictor variables for the Demeritt model where percent softwood is upper left, volume is upper right, basal area is lower left, and tree count is lower right.

Another comparison that we can make is stand-level volume predictions from the model and stand-level volume estimates derived from the traditional forest inventory. The total predicted volume in cords per acre from the model was 3,818,469.77, and the total estimated volume from the traditional forest inventory was 2,512,658.50, with a difference of 1,305,811.27. This difference indicates that the EFI model is predicting more volume than the estimates from the traditional forest inventory. Table 2.8 shows a subsample of stands in the Demeritt Forest with corresponding EFI and traditional forest inventory volume predictions. The minimum and maximum values associated with each stand are shown with the EFI volume predictions. Stand B36 was predicted to have 52.30 CDS/ac, while the traditional forest inventory estimated 33.30 CDS/ac. Even though the two estimates are off by 19 CDS/ac, it is essential to understand that the EFI predictions are presented in ranges because the predictions within each raster pixel were aggregated to the stand boundaries. This means that some areas of the stand may have a lower predicted volume estimate and depend on the pixel's location.

In contrast, the traditional inventory estimate is derived from measurements in one to several plots per stand that are not necessarily spatially dependent. This is where EFIs come at an advantage because, with these predictions, foresters can have complete spatial knowledge of their land base instead of using a traditional forest inventory. In some instances, the estimated volume is higher than the predictions from the EFI. Stand E26, for example, was predicted to have 46.96 CDS/ac, whereas the estimate from the traditional forest inventory is 66.33 CDS/ac. While some stands have differing predictions and estimates of volume, others are more agreeable, like stand C49, where there is a difference of 0.24 CDS/ac. Given the comparison below, it is evident that predictions and estimates of forest metrics can vary no matter what type of inventory is used. Traditional forest inventories have been used for decades and can be reliable, but EFIs are time and cost-effective and provide expansive spatial knowledge of the land base.

Table 2.8. A subset of Demeritt **s**tands with corresponding EFI total volume estimates and the associated minimum and maximum values compared to the traditional forest inventory volume estimates all in cords per acre.

Stand	EFI CDS/ac	EFI Min CDS/ac	EFI Max CDS/ac	Estimated CDS/ac
B36	52.30	15.55	79.04	33.30
C48	33.94	13.47	71.57	43.08
C49	38.27	16.61	90.30	38.03
C50	44.86	21.27	91.71	34.55
E26	46.96	14.91	102.71	66.33
F7	36.07	13.69	90.67	29.02
I75	52.78	24.61	100.00	28.40

We can validate the Demeritt model by comparing means and standard deviations of the volume predictions from the EFI model and volume estimates from a traditional forest inventory (Figure 2.8). Most standard deviation error bars of predicted and estimated values for the stands listed in Table 2.8 above overlap, indicating that the values are not statistically significant from each other. In other words, our predicted EFI volume metrics are comparable to the traditional inventory volume estimates. However, in stands such as B36, the standard deviations of the predicted and estimated values do not overlap. This indicates that volume from the EFI model may have been overestimated in this stand. Since a sample of stands are statistically different, we can assume that the model may not work for all stands. Overall, these results indicate that the ABA may provide acceptable forest attribute predictions for forest management.

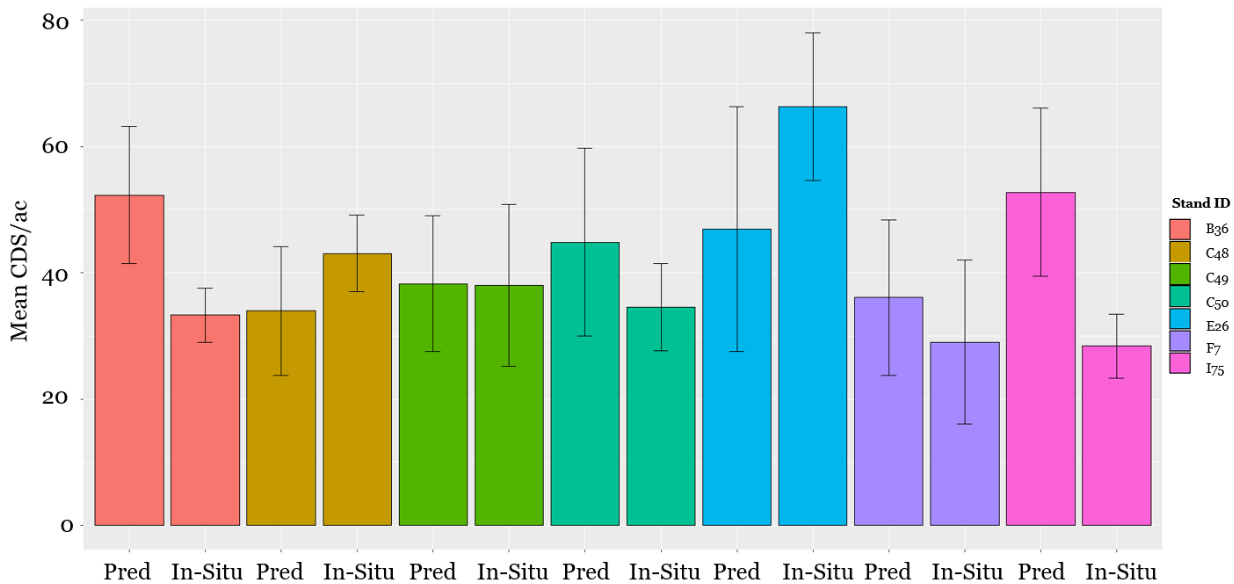


Figure 2.8. Means and standard deviations of volume (in cords per acre) in a subset of stands in the Demeritt for the predicted EFI and estimated (in-situ) values.

Comparing the Demeritt model to the other four PEF FRP models applied to the Demeritt, it becomes evident that there is little to no statistical difference between the five datasets for most forest attribute predictions (Figure 2.9). For the PSW prediction, the standard deviations of the Demeritt model and the first three applied models (143 and 92 FRPs using a PCA for plot placement and 143 FRPs using an SRS design) show overlap meaning that the models are working as expected. Additionally, there is much overlap in all model's standard deviations in the BA and TPA categories. However, when assessing volume, some models' standard deviations do not overlap others, demonstrating that these models may not be working as well as they should.

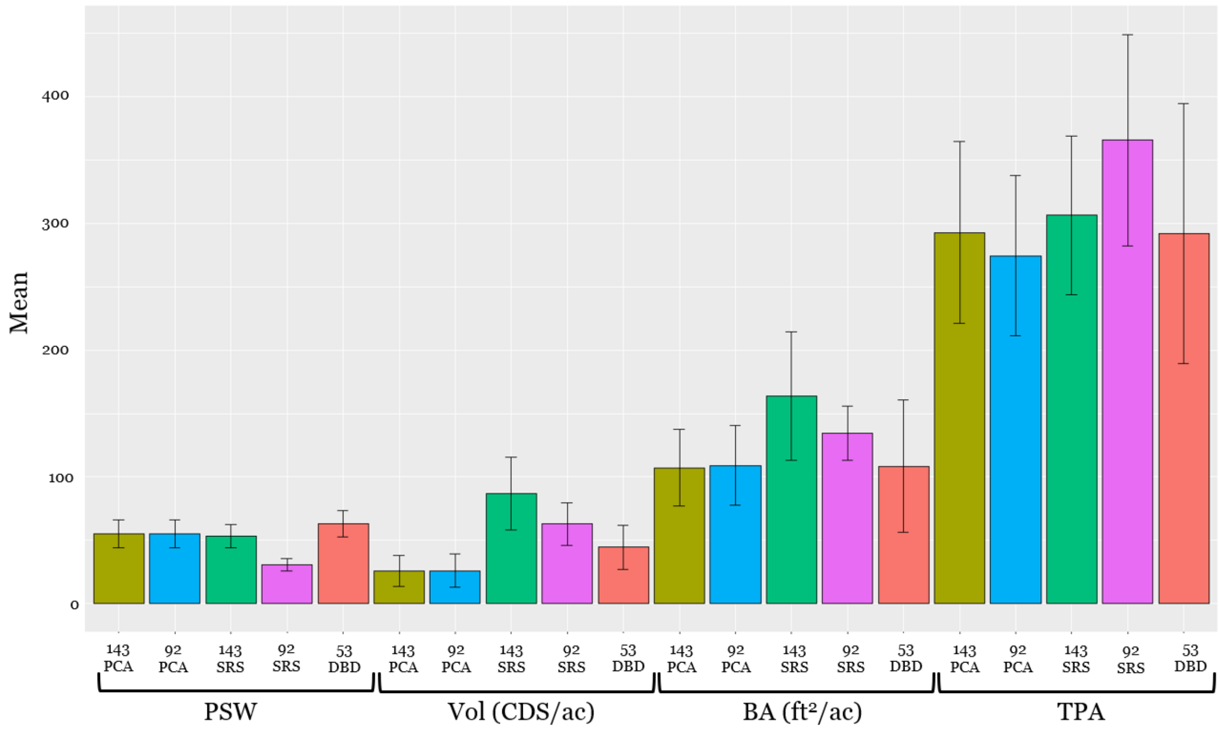


Figure 2.9. Means and standard deviations of the four PEF FRP models applied in the Demeritt and the Demeritt model for the percent softwood (PSW), volume (Vol), basal area (BA), and trees per acre (TPA) predictions.

2.4. Discussion

2.4.1 Analysis of Results

PSW predictions did well because of the covariate data used to assist the model in distinguishing between softwoods and hardwoods. The model performance declines after PSW, with tree count predictions having the worst evaluation statistics. The models may not be able to predict tree count, depending on the ALS data used, since some data were flown during leaf-on conditions and others were flown during leaf-off conditions. White et al. (2015) conducted a study comparing leaf-on to leaf-off ALS and found that models using leaf-on ALS data had slightly better model performance than models using leaf-off data (Table 2.9).

Table 2.9. RMSE and mean bias values of the White et al. (2015) study comparing leaf-on and leaf-off conditions compared to this study's results.

Source	**Forest Type**	**Target Forest Attributes**	**Results**
White et al. (2015)	Boreal	Lory's mean height* and crown closure	Leaf-on: RMSE = 3.738 mean bias = 0.098 Leaf-off: RMSE = 4.612 mean bias = 0.005
Chapter 2 data	Mixed	PSW*, volume, basal area, and tree count	Leaf-on: RMSE = 19.9 mean bias = 0.47 Leaf-off: RMSE = 0.17 mean bias = -0.004

*The results shown for White et al. (2015) are reflective of Lory's mean height for leaf-on and leaf-off data while this paper's results are reflective of the best performing models which were PEF using 92 VRPs (leaf-on) and Baskahegan (leaf-off).

Other studies using leaf-off ALS data or comparing leaf-on to leaf-off ALS data show mixed results (Table 2.10). Hawbaker et al. (2010) used leaf-off ALS data to predict various forest attributes, in which the evaluation statistics could have been better. Næsset (2005), contrary to the previous study mentioned, found that leaf-off ALS provided slightly better results. There was also a difference in the models when comparing plot types. Models that used FRPs tended to have better statistical outcomes than VRPs. A higher sample size intuits a better model performance, but results show the inverse. This may be due to the inner workings of random forest.

Table 2.10. R^2, RMSE, and mean bias values of studies that only investigated leaf-off conditions or compared leaf-on and leaf-off ALS conditions.

Source	**Forest Type**	**Target Forest Attributes**	**Results**
Hawbaker et al. (2010)	Mixed Hardwood	Tree density, basal area*, Lory's mean height, and volume	Leaf-off: R^2 = 0.13 mean bias = -13.0
Næsset (2005)	Mixed	Lory's mean height, basal area*, volume	Leaf-on: R^2 = 0.62 Leaf-off: R^2 = 0.66
Chapter 2 data	Mixed	PSW, volume, basal area*, and tree count	Leaf-on: RMSE = 19.9 mean bias = 0.47 Leaf-off: RMSE = 0.17 mean bias = -0.004

*The results shown for Hawbaker et al. (2010), Næsset (2005), and the PEF model using 92 VRPs and Baskahegan model are reflective of basal area for leaf-on and leaf-off data.

Since the models use many of the same variables to predict the different forest metrics, the models themselves may not be the issue. Further data collection is needed to compare the results from 2021 in the PEF and Demeritt to assess if observations are similar and provide comparable predictions. The structural variability may also impact the prediction outputs and evaluation statistics. The PEF and Demeritt Forest are both research forests with many different silvicultural applications creating areas with varying structures and species compositions contrary to the working industrial forests in Northern Maine. The increased variability of these two forests may prove difficult to capture without sacrificing time and cost efficiency and is potentially reflected in the results.

2.4.2 Future Work

More data should be collected to compare different sample sizes. For the purposes and timeline of this research, only two sample sizes and sampling intensities were compared in one study location. More work should be done to expand upon this by considering higher and lower sample sizes and sampling intensities. Future research should also focus on using different sample sizes and sampling intensities in areas beyond the PEF to ensure proper recommendations are being made for different forest types. Even though a smaller sample size shows the better model performance when using random forest, larger sample sizes, theoretically, capture more variability in the population, which should provide better results. This may be because a small sample size has less correlation and due to the inner workings of random forest. The user must determine what type of predictive model should be used for their research. Random forest is used in remote sensing because it yields reliable classifications (Breiman, 2001) but linear models have long been established and are well-understood (Means et al., 2000). Different sample designs should also be considered. We could only compare three different sample designs in this research, but there are others such as strip, cluster, or stratified random sampling.

Lastly, VRPs in the Ashland West site performed worse than the Baskahegan and Rangeley sites. These same trends were absent in the PEF. FRPs did, however, produce better evaluation statistics overall, while VRPs for some forest metrics performed slightly better or equally as well as FRPs – specifically when predicting volume and BA. The PEF, since it is a research forest with various silvicultural implementations, is a highly variable forest, most likely not representative of the industrial forests in northern Maine. This could be one explanation for the differing results between all four sites – why VRPs performed poorly in all models except for some forest metric predictions in the PEF. Specifically, the Baskahegan and Rangeley sites in the spruce-fir forest type (some parts of the Rangeley site in the beech-birch-maple forest type) had the best overall results, and even the Ashland West site in the spruce-fir forest type outperformed the PEF models in some predictions (Table 2.11). Maine's forests are intensively managed, especially the spruce-fir forests in northern Maine, where much of the industry is located.

Table 2.11. Summary of study areas (and validation study area) and associated forest ecoregions (McCaskill et al., 2016).

Study Area	**Forest Ecoregions**
PEF	Oak-Pine
Demeritt	Oak-Pine
Ashland West	Spruce-Fir
Rangely	Beech-Birch-Maple/Spruce-Fir
Baskahegan	Spruce-Fir

2.5. Conclusions

Although there is still much research to be done before making formal recommendations, there were trends in our findings that support our hypothesis. Our results show that a PCA for plot placement, FRPs, and a smaller sample size and sampling intensity provided slightly better results. However, the Baskahegan model, which used an SYS design to place plots, also showed good performance. Although smaller sample sizes had better results, that may not be the case if other predictive models are used. If using a random forest, smaller sample sizes reduce the number of correlated variables in each tree, resulting in improved evaluation statistics (Breiman, 2001). Results may be different if using a parametric model, such as linear regression. Employing these plot types and smaller sample sizes could allow the user to create more accurate and precise EFIs if utilizing random forest for the predictive model. This may be especially true in the spruce-fir forest types in northern Maine, where most of the forested land is managed for timber. The following steps should consist of attempting various combinations of these designs in other forest types to define what approach works best in southern Maine's oak-pine or beech-birch-maple forests. Other plot designs and the type of GPS used to collect plot location data, plot shape, and sample design should be further investigated. However, some plot types and sample designs may be cheaper and faster to execute than recommended here. Foresters should use a plot type and design that best fits their needs and is economically feasible for their company. If implemented in the future, foresters can effectively balance economically sustainable practices with ecologically sound interventions conserving ecosystem services that will continue to reward forest product companies and the public for future generations.

CHAPTER 3: USING BEST PRACTICES TO ESTIMATE ABOVEGROUND FOREST ABOVEGROUND BIOMASS WITH AIRBORNE LASER SCANNING IN MAINE

3.1Introduction

Trees play a substantial role in mitigating the effects of climate change by sequestering and storing carbon from the atmosphere (Pan et al., 2011). It is estimated that forests can store up to 70-90% of total global aboveground carbon depending on the forest type (i.e., boreal, temperate deciduous, rainforest, etcetera) (Cusack et al., 2014; Hao et al., 2019; Simard et al., 2020). Researchers and land managers need to understand how trees affect the concentration of anthropogenic carbon in the atmosphere, especially in areas that are intensively managed since young trees sequester more carbon as they tend to accumulate growth rapidly, whereas older trees, once the growth curve flattens out, store more carbon (Stephenson et al., 2014; Waring et al., 2020). Understanding how trees capture carbon in intensively managed forests and accurately estimating stored carbon is the first step in actively formulating a response to climate change (Pang et al., 2022).

Simple allometric equations using tree height and diameter at breast height (DBH) are used to estimate aboveground biomass from which carbon estimations are derived (Sun & Liu, 2020). This inventory-based approach to carbon estimation requires in-situ measurements of tree height, DBH, and potentially wood density for species-specific equations (Daba & Soromessa, 2019; Vorster et al., 2020). Although these inventory-based practices are the norm and accepted, biases and errors are inherent. For example, measuring the entire population at larger landscape scales is almost impossible, so forests are often sampled using plots. Statistics are applied to the sample of trees and expanded to the population, therefore, introducing error (Tomppo, 2004). Another error is that estimates derived from these inventories are not spatially explicit (Knoke et al., 2021) and are usually made at the stand or landscape scale (Ståhl, 1992). While stand or landscape-scale estimates of forest attributes may give foresters a better interpretation of forest characteristics like volume and basal area, spatially explicit estimates provide a finer detail that improves that understanding. Time and money constraints may also provide foresters with bias and error since repeat measurements may not occur for several years. Additionally, collecting in-situ measurements in intensively managed forests can negatively impact forest attribute estimates if the structural variability is not sufficiently captured (Didion et al., 2009). Remote sensing products may be able to help overcome these restraints.

Where there are missing data in a forest inventory, remote sensing products can fill in the knowledge gaps (Lister et al., 2020). These products can assist in estimating forest metrics such as BA, volume, biomass, and more (White et al., 2013; Esteban et al., 2019; Fraser & Congalton, 2021). Traditional forest inventories that use remote sensing products are known as enhanced forest inventories and still require in-situ measurement collection to produce accurate and precise results (Wulder, 1998). As more remote sensing products emerge and become publicly available, interest in their use in EFIs grows. Remote sensing products can be used to aid the generation of wall-to-wall forest attribute prediction maps for an area of interest. In-situ data are fed into a model with remote sensing products in which forest attribute predictions are made across areas of interest - even in areas with no ground plots (Waser et al., 2015).

While many remote sensing products, like satellite imagery, provide adequate detail pertaining to an area of interest, it lacks the ability to capture the 3D structure of the forest. Additionally, there is no limit on geographic extent; the only limitation is the size of the remote sensing dataset. With a high return interval of some satellite remote sensing products (i.e., a weekly return interval), foresters and researchers can enjoy a finer temporal resolution than a typical traditional forest inventory which may be completed every 5, 10, or 15-plus years.

Airborne laser scanning, a remote sensing product, can capture the structure of the forest. A 3D point cloud displaying the forest structure is created when a sensor attached to the underside of an aerial vehicle emits laser pulses at target objects. In this case, the target objects are trees. These laser pulses hit the trees and are reflected to the sensor, where the computer calculates the distance of each return. Information from 3D point clouds can be used for measurements like tree height and individual tree segmentation analyses (Weiser et al., 2022). Tree segmentation helps extract tree height information without measuring tree heights in-situ (Weinstein et al., 2021). Point clouds can also help develop stand-level predictions across a landscape using the area-based approach (ABA) to modeling (White et al., 2013). After collecting an inventory, missing information about forest characteristics can be predicted from ABA models that use calibrated ALS data (Brosofske et al., 2014).

The ABA has two distinct stages (White et al., 2013). ALS data are clipped to the plot locations in the first stage and fed through a predictive model. In the second stage, wall-to-wall estimates are made, resulting in a raster dataset for the area of interest. This has become a popular modeling approach for many users and is recommended by researchers (White et al., 2013; Næsset et al., 2002; Woods et al., 2011). Two main advantages of the ABA are 1. users can have complete spatial knowledge or forest attribute estimates (White et a., 2013), and it has been shown to predict forest metrics more accurately and precisely (Coomes et al., 2017).

The raster outputs from these models can be used in countless ways to support the sustainable management of a forest. The most obvious is creating stand-level predictions that can be used to assess forest growth and productivity for future harvest operations. Another way to use the outputs is to monitor change after harvesting. In this chapter, we applied the best practice recommendations from our previous study to the Penobscot Experimental Forest (PEF) on two different dates – 2017 and 2021 – to estimate aboveground biomass (AGB). We accomplished a change detection analysis using those model outputs and canopy height models (CHMs) as a reference to determine increases and decreases in biomass and canopy height in harvested areas.

3.2. Methodology

3.2.1 Study Site

The area of interest (AOI) chosen for this study was the Penobscot Experimental Forest (PEF) located in Bradley and Eddington, Maine (Figure 3.1). The PEF is about 3,800 acres managed by the U.S. Forest Service on a 99-year loan with the goal of long-term management research in the northern mixed forest type. It is located on the eastern side of the Penobscot River, with an offshoot stream on its perimeter that feeds into Chemo Pond. According to the NRCS Soil Survey, the PEF contains soils from very poorly drained (nearest Blackman Stream and Chemo Pond) to well-drained, allowing various vegetation to exist. It contains many peatlands and (forested and non-forested) wetlands. It is situated in Maine between the spruce-fir forests of the north and oak-pine forests of the south, creating a variety of habitats suitable for various wildlife species. Its proximity to the University of Maine, the many forest management projects, and its characteristics made this forest a convenient and valuable location for data collection.

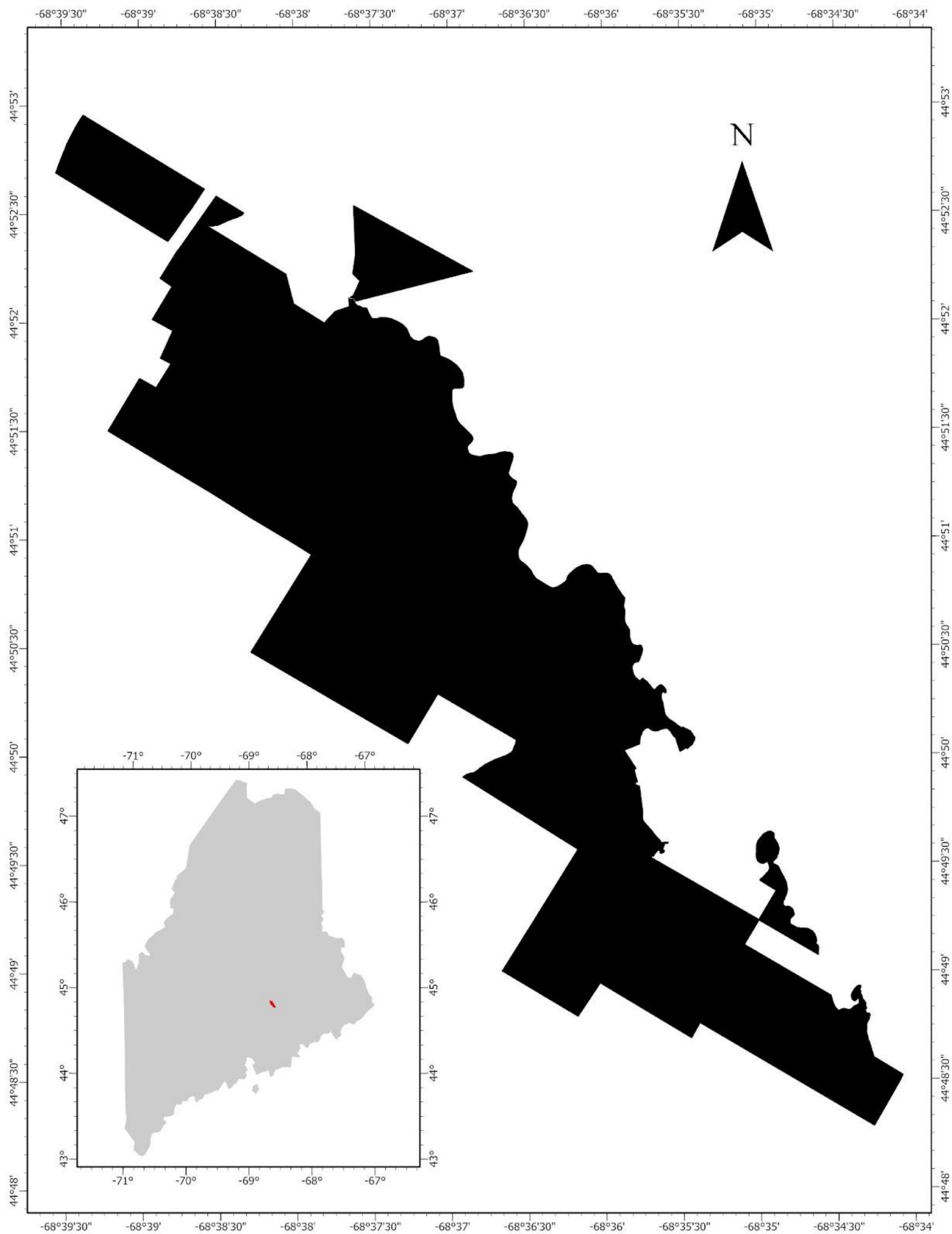


Figure 3.1. Location of the PEF within the state of Maine. Study site is red in the inset map denoting the location within the state of Maine.

3.2.2 Data

3.2.2.1 In-Situ Data

ALS data for the PEF was downloaded, normalized, and then run through a Principal Component Analysis (PCA) to determine the location of plots. 93 was the input for the PCA since this is how many CFI plots are within the PEF. One plot needed to be dropped because it was outside the ALS data, so our sample size was 92. In Chapter 1, we determined that using a smaller sample size and a PCA for plot placement in the PEF returned better results, so we chose this design for this application. Using the ALS data to determine variability, the PEF was split into 64 categories, and each "bin" required a certain number of plots or no plots at all. Areas deemed to have the most variability by the PCA necessitated more plots. An advantage to using a PCA for plot placement is that plots may be placed closer to areas that are more easily accessible such as roads, which makes data collection easier. The plot shape was circular, and the radius was 10 meters corresponding to the recommendations in White et al. (2013). The plot center location was collected using a survey-grade GPS to minimize error. Data collection in the PEF was completed between May and August in the summer of 2022.

3.2.2.2 ALS Data

NASA's Goddard's LiDAR, Hyperspectral & Thermal Imager (G-LiHT) data was used for the PEF in the model and for the PCA, as mentioned above. G-LiHT combines ALS, imaging spectroscopy, and thermal measurements that can map the structure and composition of forests (Cook et al., 2013). While these data come from an amalgamation of sensors, this project uses ALS data. G-LiHT, like other ALS data, creates 3D point clouds from which height measurements and structure can be determined – this is how the PCA determined the variability in forest structure in the PEF. G-LiHT has a high spatial resolution (12-15 pls/m2), and the sensor is flown during leaf-on conditions which may allow for more accurate forest metric predictions. The small footprint (10cm diameter) allows the user to characterize better forest structure and disturbances (Cook et al., 2013). G-LiHT data was collected in July/August of both 2017 and 2021, which is temporally close to the in-situ data collection timeframe.

3.2.2.3 Covariate Data

National Land Cover Database (NLCD) land cover data for 2019 and Sentinel- 2 data were used in addition to the ALS data in the models. 2019 NLCD land cover data was chosen because it is close to the ALS acquisition date and the in-situ collection date. Land cover data were used to categorize the biomass per forest type in the PEF since NLCD uses four forest type classifications that are classified as deciduous forest, evergreen forest, mixed forest, and woody wetlands (Dewitz, 2021). Sentinel-2 data with less than 30% cloud coverage was downloaded for 2021 from USGS's Earth Explorer before the archive ended. Sentinel-2-derived products included a normalized difference vegetation index (NDVI), a four-band composition using the red, green, blue, and infrared bands, and a principal component analysis (PCA) composite with principal components 1 and 2 that accounted for 94.24 and 5.76% of the variation, respectively. Another supplementary dataset used was a canopy cover raster created with the G-LiHT data. All noise and points over 164 feet were filtered in the ALS data. Then a percentage of all returns over the mean height was calculated at a spatial resolution of 10 meters. This canopy cover dataset was used to describe the canopy structure. Lastly, we used various standard height and intensity metrics within the ALS dataset, as described in Table 3.1 below. These metrics were subset using the lidR package in R (Roussel et al., 2020; Roussel & Auty, 2023).

Table 3.1 ALS metrics used in this study and a description.

Metric Name	Description
zmax	elevation maximum
zmean	elevation minimum
zsd	elevation standard deviation
zskew	elevation skewness
zkurt	elevation kurtosis
zentropy	elevation entropy of returns
zq5, zq10, ..., zq90, zq95	elevation percentiles
zpcum1, zpcum2, ..., zpcum8,	cumulative percentage of returns in each layer of the
zpcum9	elevation range
imax	intensity maximum
imean	intensity mean
isd	intensity standard deviation
ikurt	intensity kurtosis

3.2.3 Data Analysis

We used the component ratio method (CRM) to estimate aboveground biomass (Woodall et al., 2011). The CRM uses equations derived from Jenkins et al. (2003) which used wood-specific gravity to estimate the biomass of particular components of trees like branches and foliage. AGB data for the PEF was then applied to our model using the area-based approach (ABA) (Figure 3.2). Stage 1: ALS, in-situ, and covariate data were collected. Once these two datasets had been collected, wall-to-wall metrics were calculated for the PEF. Since we used 10-meter radius plots, the tessellation was done at 20 meters. Plot data were then clipped to the ALS and covariate data creating plot-level ALS/covariate data metrics that the model used as training data. These metrics contain forest metrics such as height percentiles, intensity values, NDVI values, and canopy cover values. A random forest model was used where the in-situ data were the response variables, and the ALS and covariate metrics were the predictors. Stage 2: Wall-to-wall estimates were generated for the PEF, resulting in a raster dataset containing those estimates. This information can then be expanded to stand-level predictions for easier use (White et al., 2013; White et al., 2017).

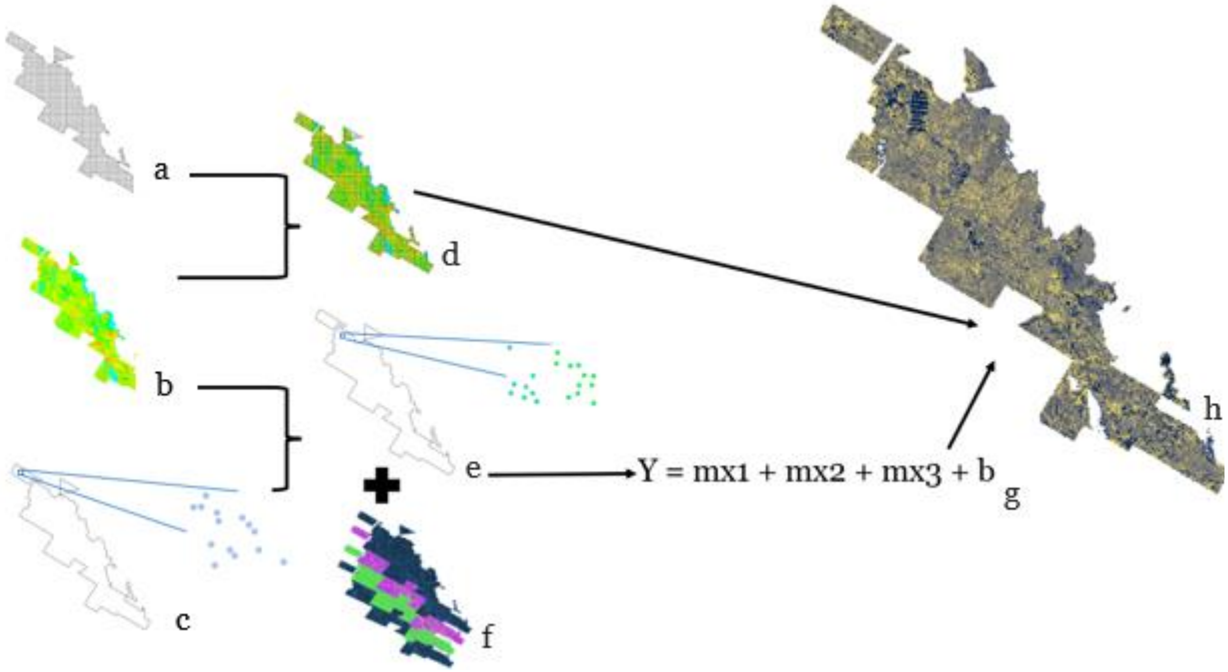


Figure 3.2. The ABA consists of several steps which are (a) tessellating the area of interest based on plot size, (b) gathering ALS data (c) gathering ground plot data, (d) generating wall-to-wall metrics, (e) clipping ALS data to plots, (f) compiling covariate data, (g) applying the predictive model, and finally (h) generating wall-to-wall maps.

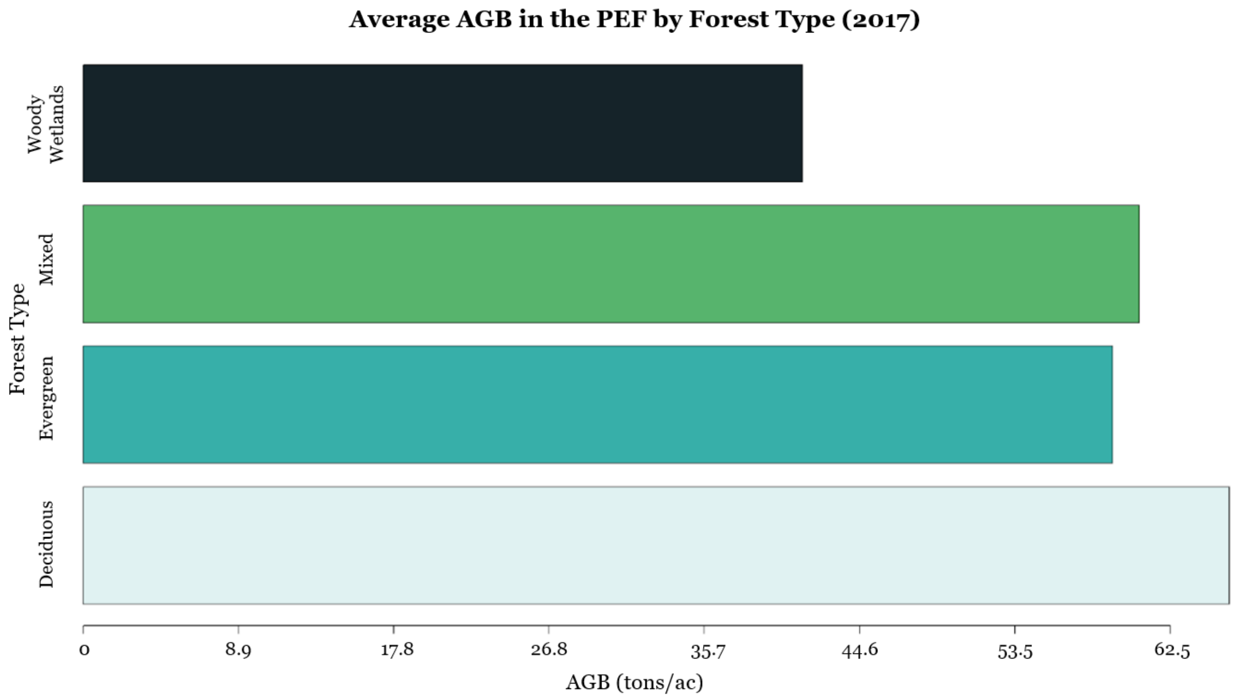
Once both models had been completed and raster products downloaded, we could begin a change detection analysis. Four different datasets were used for this analysis: the two raster outputs from the PEF models (AGB for 2017 and 2021) and two canopy height models (CHMs) for 2017 and 2021. CHMs were used because canopy height can be a crucial indicator of AGB (Tao et al., 2016). The outputs from the change detection analysis using the CHMs were used in conjunction with the results of the analysis using AGB to validate the changes in AGB and to show how the two data are interrelated. The analysis was done by subtracting the 2021 dates from the 2017 dates. Forest-type data were then extracted from an NLCD land cover dataset. This was done to understand how canopy height and AGB changed from 2017 to 2021 in the PEF.

3.3. Results

Comparing the R2 and root mean square error (NRMSE) of the two models, it is apparent that the 2021 model outperforms the other (Table 3.2). The R^2 for the 2021 model is slightly higher, indicating that the predicted biomass versus the observed biomass values are closer to each other than the predicted and observed values in the 2017 model. Similarly, the lower normalized root mean square error (NRMSE) metric of the 2021 model indicates that the predicted biomass value ranges are closer to the ranges of the observed values. Both models' normalized mean bias values (NMB) are negative, revealing that these predictions underestimate the actual value. However, the NMB metric of the 2021 model, being the most negative, shows that this model underestimates biomass more than the 2017 model. Lastly, the lower coefficient of variation (CV) metric of the 2021 model indicates that this model has smaller residuals comparatively.

Table 3.2. Summary of evaluation statistics for the two PEF models – the first using 2017 G-LiHT data and the second using 2021 G-LiHT data.

	Aboveground Biomass (tons/ac)			
Year	**R^2 (%)**	**NRMSE (%)**	**NMB (%)**	**CV (%)**
2017	50.12	62.8	-0.01	37.38
2021	54.99	59.5	-0.04	36.26



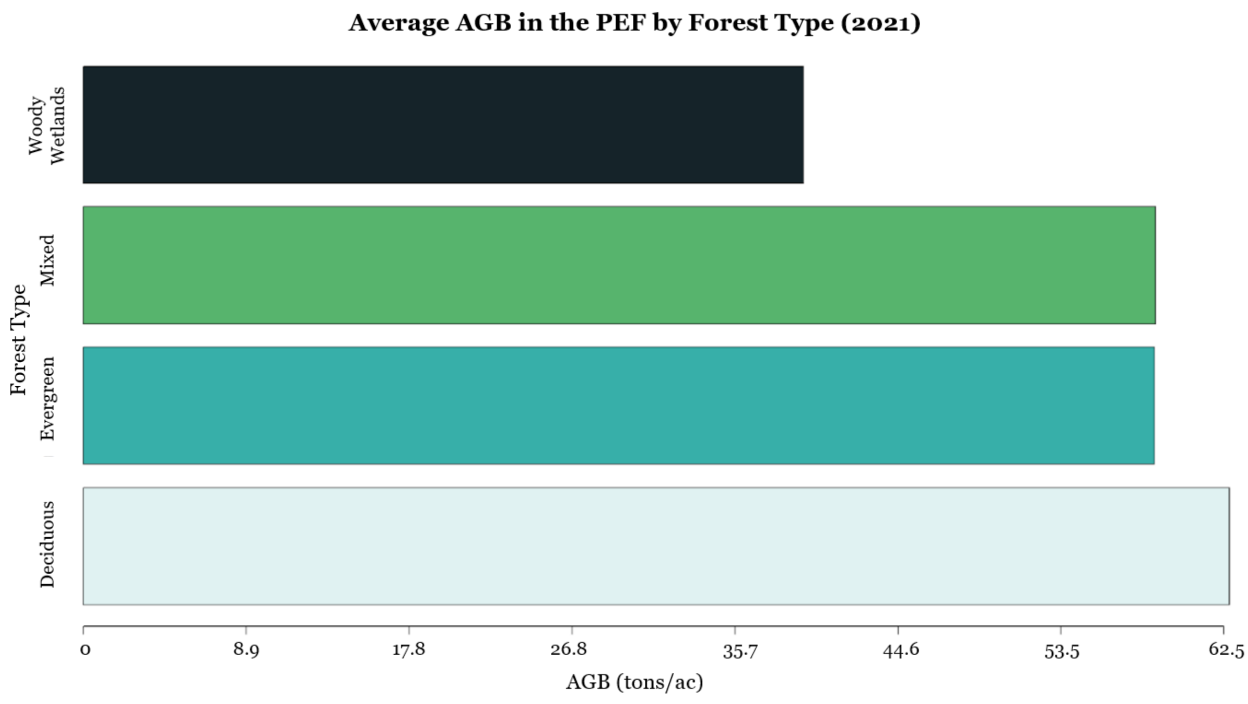


Figure 3.3. Chart showing the amout of aboveground biomass (AGB) by forest type in the PEF in 2017 (top) and 2021 (bottom).

Comparing the AGB predictions from the two models, some differences appear (Figure 3.3). From 2017 to 2021, AGB decreased in the deciduous forest type while it increased in the evergreen forest type. Over the five years, there was a change in canopy height and AGB (Figures

3.4 & 3.5). In 2017, there was a research experiment that required several strip clearcuts. We can see from those figures that the clearcut patches were not visible in the CHM and AGB images in 2017 but appear in 2021. This is also seen in the change detection images (outlined in green), where the strips appear white or light blue, indicating a drastic decrease in AGB and canopy height in those areas. Figure 3.6 shows what tree classification (deciduous, evergreen, etcetera) was removed from those clearcuts. It shows that most trees removed fall into the deciduous forest type followed by mixed forest type, woody wetlands, and evergreen forests close behind.

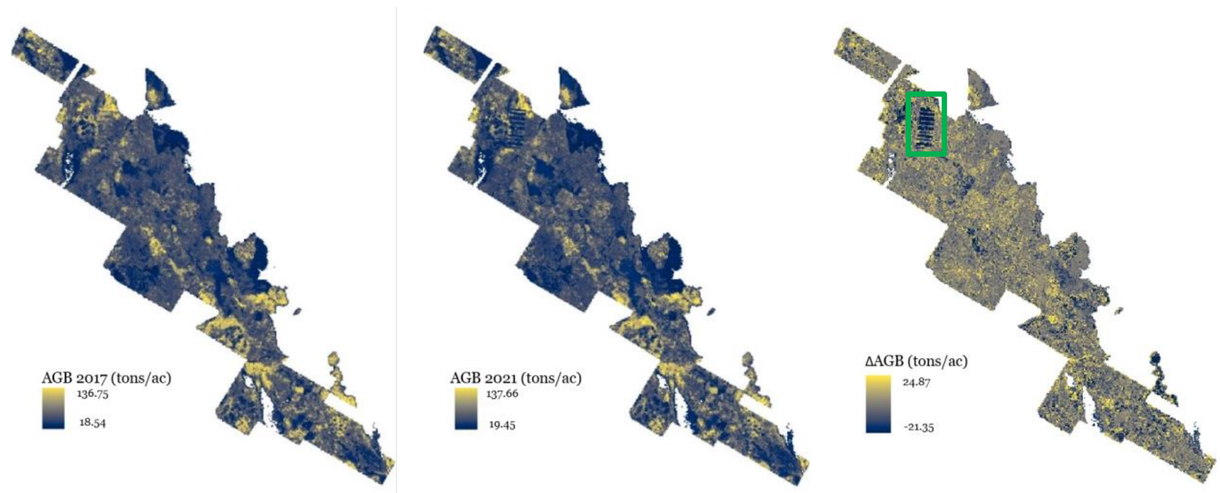


Figure 3.4. AGB in the PEF in 2017 (left), 2021 (middle), and the change in AGB (right) where yellow hues indicate higher biomass, and white or blue hues indicate lower biomass all shown in 20-meter pixels.

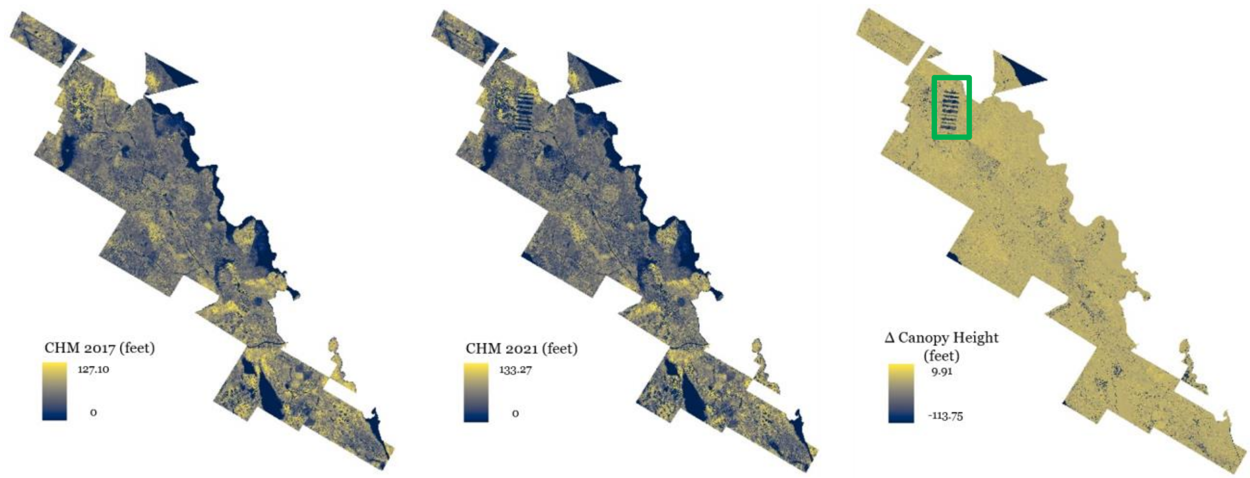


Figure 3.5. Canopy height in the PEF in 2017 (left), 2021 (middle), and the change in canopy height (right) where yellow hues indicate increased height, and white or blue hues indicate decreased height all shown in 1-meter pixels.

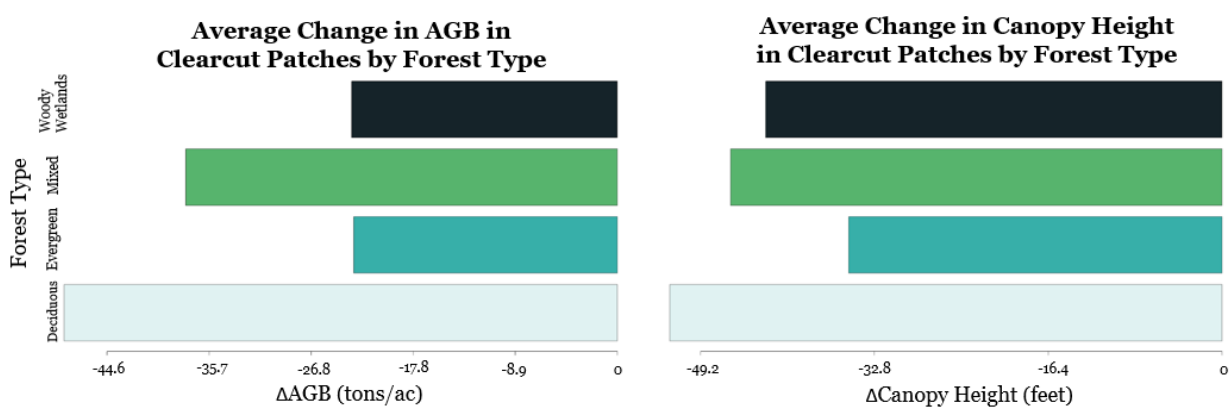


Figure 3.6. Chart showing the average change from 2017 to 2021 in AGB (left) and canopy height (right) in clear cuts in the PEF by forest type.

Other silvicultural removal practices in the PEF between 2017 and 2021 were expanding gap shelterwoods, overstory removals (OSRs), and shelterwoods with the addition of an OSR. The average change in AGB in expanding gap shelterwoods, we see that the most decrease occurred in the mixed forest type followed by evergreen and deciduous (no removals occurred in the woody wetlands category) (Figure B.1). In contrast, according to the change in canopy height, the most decrease in height occurred in the evergreen forest type followed by mixed, deciduous, and lastly the woody wetlands. Although the first two rankings were different, the trend remains.

The average change in AGB in OSRs shows the most biomass decrease in the deciduous forest type, followed by evergreen. Lastly, the mixed forest type (Figure B.2). Instead of biomass decreasing, there was an increase in the woody wetlands type. This contrasts with the average change in canopy height in OSRs. This model showed an increase in height in the mixed forest type, while the most decrease was in the deciduous forest type, followed by the evergreen and woody wetland types.

The last silvicultural treatment examined was the shelterwood with an additional OSR operation (Figure B.3). The average change in AGB showed an increase in woody wetlands. This is like the findings mentioned previously for the OSR treatment. The most decrease in biomass occurred in the deciduous forest type, followed by the mixed and evergreen forest types. There were no increases in average canopy height in these treatment areas from 2017 to 2021. The evergreen forest type saw the most decrease in height, followed by the mixed forest and woody wetlands. There was a decrease in canopy height in the deciduous forest type, but it was less relative to the other forest types.

3.4. Discussion

3.4.1 Data Analysis

This chapter explored the application of our best results from the previous chapter. This means that we used the smaller of the two sample sizes in the PEF from our previous research, which was 92 plots. Plot locations were determined using a PCA, and the plot type was a fixed radius. Our results using this design in the PEF in the previous study did not provide results on par with other models. As considered in the chapter 2 discussion, additional work should be done to collect more inventory data and investigate the effects of different sample sizes and sampled designs. Although this design did not provide better results than other models used in

Chapter 2, this did provide the best results for the PEF overall. The evaluation statistics for predicting AGB could be more robust. The R^2 values around 50% indicate that the predicted values are moderately close to the observed values.

Moreover, the NMB and CV values indicate that the model could be more precise and accurate. However, the PEF is a variable forest with many different silvicultural practices implemented throughout the years. Because of the heterogeneous state of the structure, species composition, and intense management of the PEF, the model may have difficulties predicting forest metrics since variability may not have been fully captured (Didion et al. 2009). If this model were to be implemented in a different forest, we might see different results. Foresters may need to implement a type of plot and sample design that best fits their needs.

The differences in these results were also seen in the graphs comparing the average change in AGB or canopy height to the different silvicultural treatments by forest type. There were slight differences in the two changes for the clearcut patches, and the overall trend remained. It was only when the other three silvicultural treatments were compared that the differences became more evident. This may have to do with the spatial resolution of all the datasets. AGB was predicted at 20-meter pixels, while the CHM was created at 1-meter pixels. Since we wanted to compare the average increases and decreases of AGB and canopy height in our analysis, we decided to use NLCD land cover data available at 30-meter pixels. The differences in pixel size across the datasets, especially when calculating the mean between AGB and forest type – and canopy height and forest type – may have resulted in errors that either over or underestimated the increases and decreases in AGB and canopy height in each silvicultural treatment.

3.4.2 Future Work

Like the discussion in Chapter 1, more research is needed to determine what plot types and sample designs will work best in different forest ecotypes. The calibration plot type used here may work slightly better for other study sites that happen to be in the spruce-fir forest type.

The PEF is located between the spruce-fir and oak-pine forest ecotypes, which have different tree species and more structural variability. This may influence model performance and must be explored further.

Pixel size for the AGB prediction raster data was determined by the size of our plots, which were 10-meter radius. That meant that our pixel size had to be 20 meters. If using NLCD data to determine the average change in AGB, future research should include a plot size of 15 meters so that the pixel size is the same as the NLCD datasets. Also, the CHM should be created with a coarser resolution to match the forest-type dataset's resolution. This may eliminate some of the errors shown in this chapter.

3.5. Conclusions

The results from this chapter relate to the concerns expressed in Chapter 2 since this was the application of our best practice recommendations. This was an example of how our partners and others may use their EFIs' results. Not only can this process be used to predict any range of forest metrics, but it can also be used to monitor changes in the forest. In this example, we estimated AGB in our study area and completed a change detection analysis; this was done in conjunction with a CHM change analysis since canopy height is related to forest biomass. We knew of past harvest activities in the PEF between 2017 and 2021. We were able to visualize those changes through the predictive outputs of the model and by assessing the average change in biomass and canopy height. This exemplifies the power of this modeling approach and its usefulness to foresters who may want to predict forest metrics more accurately and precisely, like volume or biomass. Not only can they improve forest metric predictions, but they can also monitor the removal of timber via a change detection analysis. This has the potential to assist foresters in assessing forests – how their past histories may affect production and how to manage them better to remain both economically viable and ecologically sound.

CHAPTER 4: CONCLUSION

We used the ABA to model forest attribute estimates, a recommended best practice for western Canadian forests by White et al. (2013) and shown to provide precise and accurate results in Norway's forests under operational conditions (Næsset, 2007). By assessing recent data in the PEF using two plot types, two sample sizes, and sampling intensities, and two sample designs, we were able to compare model performance better, especially to data collected in other sites using only one type of plot, one sample size and sampling intensity, and one sample design. We were also able to identify recommendations from this research and use them in the PEF to estimate AGB and demonstrate the applicability of these best practice suggestions in a forest monitoring capacity. The results from this research are intended for both foresters and researchers whose concerns lie with effectively integrating remote sensing products in their inventory analysis and how to formulate the inventory and plot design best to get acceptable results for their needs.

4.1 Key Findings

4.1.1 Best Practices Results

Our overall goal for this research was to develop models that compared fixed versus variable radius plots, sampling size and intensity, and sample design with ALS data to map EFI variables, while our objective for Chapter 2 was to quantify and compare the accuracy of those forest inventory designs at predicting percent softwood, volume, BA, and tree count. We found that a PCA for plot placement yielded the best model results when compared to plots that had been randomly placed. This works by using the calculated ALS metrics and inputting the desired number of plots. The model then assesses how much variation in the overall dataset is accounted for in the first two principal components and assigns plots to capture this variation in the placement of EFI calibration plots. Our second finding was that models using a smaller sample size outperformed those with a larger sample size. This may be because of the predictive model (random forest) that we used, where a smaller sample size means decreased correlation. Our third finding was that FRPs are a better plot type to calibrate ALS data with potentially depending on the site. In the spruce-fir forest types, FRPs performed best overall. However, we saw mixed results in the more mixed forest types, like in the PEF. FRPs still had high evaluation statistics, but VRPs performed equally, if not better, when predicting some forest metrics like volume and BA.

4.1.2 Application of Best Practices Results

Our objective for Chapter 3 was to use the best practice recommendations created from the first objective to map forest structures and determine the change in aboveground biomass from 2017 to 2021. We used a PCA for plot placement, a small sample size (n = 92), and FRPs for this analysis based on the results in Chapter 2. Once these recommendations from Chapter 2 had been applied at two dates to predict AGB, we then completed a change detection analysis which was also done with two CHMs for comparison. We extracted harvested areas and established what the increases and decreases in AGB and canopy height were by forest type. While the average changes in biomass and canopy height by forest type in the clearcut treatments were similar, there were disagreements in these changes by forest type in the expanding gap shelterwood, OSRs, and shelterwoods with OSR treatments. These disagreements may be due to the differences in forest cover type pixel size, the AGB and CHM data, and errors in predicting biomass. More research should be done looking at plot type and design in the creation of EFIs in Maine. This study laid the framework for applying the best practice recommendations in Chapter 2 and what analysis could look like for foresters who want to continue to manage the forests sustainably.

4.2 Future Research & Final Thoughts

Overall, this study offers valuable insight into how plot type and design can affect the outcomes of EFI models and their applicability. This research provides a crucial step in the creation of a best practices manual that will help guide foresters through the setup and success of an EFI for their area of interest. This will ensure that foresters in Maine can practice sustainable forest management that may have the potential to be more economically feasible and ecologically sound.

BIBLIOGRAPHY

3D Elevation Program | U.S. Geological Survey. (n.d.). Retrieved February 22, 2023, from https://www.usgs.gov/3d-elevation-program

Alberdi, I. (2021). Forest Resources Assessments: Mensuration, Inventory, and Planning. *Forests*, *12*(3), 296. https://doi.org/10.3390/f12030296

Bolton, D. K., White, J. C., Wulder, M. A., Coops, N. C., Hermosilla, T., & Yuan, X. (2018). Updating stand-level forest inventories using airborne laser scanning and Landsat time series data. *International Journal of Applied Earth Observation and Geoinformation*, *66*, 174–183. https://doi.org/10.1016/j.jag.2017.11.016

Bouvier, M., Durrieu, S., Fournier, R., & Renaud, J.-P. (2015). Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote Sensing of Environment*, *156*, 322–334. https://doi.org/10.1016/j.rse.2014.10.004

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140. https://doi.org/10.1007/BF00058655

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brosofske, K. D., Froese, R., Falkowski, M. J., & Banskota, A. (2014). A review of methods for mapping and prediction of inventory attributes for operational forest management. *Forest Science*, *60*(4), 733–756. https://doi.org/10.5849/forsci.12-134

Busby, P. E., & Canham, C. D. (2011). An exotic insect and pathogen disease complex reduces aboveground tree biomass in temperate forests of eastern North America. *Canadian Journal of Forest Research*, *41*(2), 401–411. https://doi.org/10.1139/X10-213

Chen, S., Liu, H., Feng, Z., Shen, C., & Chen, P. (2019). Applicability of personal laser scanning in forestry inventory. *PLOS ONE*, *14*(2), e0211392. https://doi.org/10.1371/journal.pone.0211392

Cheng, K., & Wang, J. (2019). Forest Type Classification Based on Integrated Spectral-Spatial-Temporal Features and Random Forest Algorithm—A Case Study in the Qinling Mountains. *Forests*, *10*(7), Article 7. https://doi.org/10.3390/f10070559

Cook, B. D., Corp, L. A., Nelson, R. F., Middleton, E. M., Morton, D. C., McCorkel, J. T., Masek, J. G., Ranson, K. J., Ly, V., & Montesano, P. M. (2013). NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. *Remote Sensing*, *5*(8), Article 8. https://doi.org/10.3390/rs5084045

Coomes, D. A., Dalponte, M., Jucker, T., Asner, G. P., Banin, L. F., Burslem, D. F. R. P., Lewis, S. L., Nilus, R., Phillips, O. L., Phua, M.-H., & Qie, L. (2017). Area-based vs tree-centric approaches to mapping forest carbon in Southeast Asian forests from airborne laser scanning data. *Remote Sensing of Environment*, *194*, 77–88. https://doi.org/10.1016/j.rse.2017.03.017

Cooperative Forestry Research Unit. (2019). *Cooperative forestry research unit: 2019 annual report*. AR2019_FINAL_web.pdf (umaine.edu)

Cusack, D. F., Axsen, J., Shwom, R., Hartzell-Nichols, L., White, S., & Mackey, K. R. (2014). An interdisciplinary assessment of climate engineering strategies. *Frontiers in Ecology and the Environment*, *12*(5), 280–287. https://doi.org/10.1890/130030

Daba, D. E., & Soromessa, T. (2019). *The accuracy of species-specific allometric equations for estimating aboveground biomass in tropical moist montane forests: Case study of Albizia grandibracteata and Trichilia dregeana | SpringerLink*. https://link.springer.com/article/10.1186/s13021-019-0134-8

Daily, G. C. (2013). Nature's Services: Societal Dependence on Natural Ecosystems (1997): In *Nature's Services: Societal Dependence on Natural Ecosystems (1997)* (pp. 454–464). Yale University Press. https://doi.org/10.12987/9780300188479-039

Deo, R. K., Froese, R. E., Falkowski, M. J., & Hudak, A. T. (2016). Optimizing Variable Radius Plot Size and LiDAR Resolution to Model Standing Volume in Conifer Forests. *Canadian Journal of Remote Sensing*, *42*(5), 428–442. https://doi.org/10.1080/07038992.2016.1220826

Dewitz, J. (2021). *National Land Cover Database (NLCD) 2019 Products* [Data set]. U.S. Geological Survey. https://doi.org/10.5066/P9KZCM54

Didion, M., Kupferschmid, A. D., Lexer, M. J., Rammer, W., Seidl, R., & Bugmann, H. (2009). Potentials and limitations of using large-scale forest inventory data for evaluating forest succession models. *Ecological Modelling*, *220*(2), 133–147. https://doi.org/10.1016/j.ecolmodel.2008.09.021

Fassnacht, F. E., Hartig, F., Latifi, H., Berger, C., Hernández, J., Corvalán, P., & Koch, B. (2014). Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sensing of Environment*, *154*, 102–114. https://doi.org/10.1016/j.rse.2014.07.028

Flatebo, G., Sulzer, C. R., & Steven K., P. (1999). *Biodiversity in the Forests of Maine: Guidelines for Land Management*.

Frank, B., Mauro, F., & Temesgen, H. (2020). *Remote Sensing | Model-Based Estimation of Forest Inventory Attributes Using Lidar: A Comparison of the Area-Based and Semi-Individual Tree Crown Approaches*. https://www.mdpi.com/2072-4292/12/16/2525

Ganz, S., Käber, Y., & Adler, P. (2019). Measuring Tree Height with Remote Sensing—A Comparison of Photogrammetric and LiDAR Data with Different Field Measurements. *Forests*, *10*(8), Article 8. https://doi.org/10.3390/f10080694

Goetz, S., & Dubayah, R. (2011). Advances in remote sensing technology and implications for measuring and monitoring forest carbon st. *Carbon Management*, 15.

Goodbody, T. R. H., Coops, N. C., & White, J. C. (2019). Digital Aerial Photogrammetry for Updating Area-Based Forest Inventories: A Review of Opportunities, Challenges, and Future Directions. *Current Forestry Reports*, *5*(2), 55–75. https://doi.org/10.1007/s40725-019-00087-2

Hao, H., Li, W., Zhao, X., Chang, Q., & Zhao, P. (2019). Estimating the Aboveground Carbon Density of Coniferous Forests by Combining Airborne LiDAR and Allometry Models at Plot Level. *Frontiers in Plant Science*, *10*. https://www.frontiersin.org/articles/10.3389/fpls.2019.00917

Hawbaker, T., Gobakken, T., Lesak, A., Trømborg, E., Contrucci, K., & Radeloff, V. (2010). Light Detection and Ranging-Based Measures of Mixed Hardwood Forest Structure. *Forest Science*, *56*, 313–326.

Hayashi, R., Weiskittel, A., & Sader, S. (2014). Assessing the Feasibility of Low-Density LiDAR for Stand Inventory Attribute Predictions in Complex and Managed Forests of Northern Maine, USA. *Forests*, *5*(2), Article 2. https://doi.org/10.3390/f5020363

Hoover, C. M., Bush, R., Palmer, M., & Treasure, E. (2020). Using Forest Inventory and Analysis Data to Support National Forest Management: Regional Case Studies. *Journal of Forestry*, *118*(3), 313–323. https://doi.org/10.1093/jofore/fvz073

Hütt, C., Bolten, A., Hüging, H., & Bareth, G. (2022). UAV LiDAR Metrics for Monitoring Crop Height, Biomass and Nitrogen Uptake: A Case Study on a Winter Wheat Field Trial. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*. https://doi.org/10.1007/s41064-022-00228-6

Imangholiloo, M., Saarinen, N., Holopainen, M., Yu, X., Hyyppä, J., & Vastaranta, M. (2020). Using Leaf-Off and Leaf-On Multispectral Airborne Laser Scanning Data to Characterize Seedling Stands. *Remote Sensing*, *12*(20), 3328. https://doi.org/10.3390/rs12203328

Iqbal, I. A., Musk, R. A., Osborn, J., Stone, C., & Lucieer, A. (2019). A comparison of area-based forest attributes derived from airborne laser scanner, small-format and medium-format digital aerial photography. *International Journal of Applied Earth Observation and Geoinformation*, *76*, 231–241. https://doi.org/10.1016/j.jag.2018.12.002

Jenkins, J. C., Chojnacky, D. C., Heath, L. S., & Birdsey, R. A. (2003). National scale biomass estimators for United States tree species. *Forest Science. 49: 12-35*. https://www.fs.usda.gov/research/treesearch/6996

Jolliffe, T., & Cadima, J. (2016). *Principal component analysis: A review and recent developments | Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. https://royalsocietypublishing.org/doi/full/10.1098/rsta.2015.0202

Kangas, A., Astrup, R., Breidenbach, J., Fridman, J., Gobakken, T., Korhonen, K. T., Maltamo, M., Nilsson, M., Nord-Larsen, T., Næsset, E., & Olsson, H. (2018). Remote sensing and forest inventories in Nordic countries – roadmap for the future. *Scandinavian Journal of Forest Research*, *33*(4), 397–412. https://doi.org/10.1080/02827581.2017.1416666

Kashani, A. G., Olsen, M. J., Parrish, C. E., & Wilson, N. (2015). A Review of LIDAR Radiometric Processing: From Ad Hoc Intensity Correction to Rigorous Radiometric Calibration. *Sensors*, *15*(11), Article 11. https://doi.org/10.3390/s151128099

Kershaw Jr, J. A., Ducey, M. J., Beers, T. W., & Husch, B. (2016). *Forest Mensuration*. John Wiley & Sons.

Knoke, T., Kindu, M., Schneider, T., & Gobakken, T. (2021). Inventory of Forest Attributes to Support the Integration of Non-provisioning Ecosystem Services and Biodiversity into Forest Planning—From Collecting Data to Providing Information. *Current Forestry Reports*, *7*(1), 38–58. https://doi.org/10.1007/s40725-021-00138-7

Köhl, M., & Magnussen, S. (2016). Sampling in Forest Inventories. In L. Pancel & M. Köhl (Eds.), *Tropical Forestry Handbook* (pp. 777–837). Springer. https://doi.org/10.1007/978-3-642-54601-3_72

Krug, T., & dos Santos, J. R. (2004). RESOURCE ASSESSMENT | Forest Change. In J. Burley (Ed.), *Encyclopedia of Forest Sciences* (pp. 989–997). Elsevier. https://doi.org/10.1016/B0-12-145160-7/00163-0

Liang, X., Kankare, V., Hyyppä, J., Wang, Y., Kukko, A., Haggrén, H., Yu, X., Kaartinen, H., Jaakkola, A., Guan, F., Holopainen, M., & Vastaranta, M. (2016). Terrestrial laser scanning in forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, *115*, 63–77. https://doi.org/10.1016/j.isprsjprs.2016.01.006

Lisańczuk, M., Mitelsztedt, K., Parkitna, K., Krok, G., Stereńczak, K., Wysocka-Fijorek, E., & Miścicki, S. (2020). Influence of sampling intensity on performance of two-phase forest inventory using airborne laser scanning. *Forest Ecosystems*, *7*(1), 65. https://doi.org/10.1186/s40663-020-00277-6

Ma, B., Hauer, R. J., Östberg, J., Koeser, A. K., Wei, H., & Xu, C. (2021). A global basis of urban tree inventories: What comes first the inventory or the program. *Urban Forestry & Urban Greening*, *60*, 127087. https://doi.org/10.1016/j.ufug.2021.127087

Maine Forest Service: Forest Policy & Management Division. (2020). *Maine forest action plan 2020*. https://www.maine.gov/dacf/mfs/about/action_plans/Maine%20Forest%20Action%20Plan%202020.pdf

Mauldin, T. E., Plantinga, A. J., & Alig, R. J. (1999). Determinants of Land Use in Maine with Projections to 2050. *Northern Journal of Applied Forestry*, *16*(2), 82–88. https://doi.org/10.1093/njaf/16.2.82

McCaskill, G. L., Albright, T., Barnett, C. J., Butler, B. J., Crocker, S. J., Kurtz, C. M., McWilliams, W. H., Miles, P. D., Morin, R. S., Nelson, M. D., Widmann, R. H., & Woodall, C. W. (2016). *Maine forests 2013* (NRS-RB-103; p. NRS-RB-103). U.S. Department of Agriculture, Forest Service, Northern Research Station. https://doi.org/10.2737/NRS-RB-103

Means, J., Acker, S., Fitt, B. J., Renslow, M., Emerson, L., & Hendrix, C. (2000). Predicting forest stand characteristics with airborne scanning lidar. *Photogrammetric Engineering and Remote Sensing*. https://www.semanticscholar.org/paper/Predicting-forest-stand-characteristics-with-lidar-Means-Acker/15fe6ec12e1b9538718bd4d1b51a81ad8a95b4cf

Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, *80*, 88–99. https://doi.org/10.1016/S0034-4257(01)00290-5

Næsset, E. (2005). Assessing sensor effects and effects of leaf-off and leaf-on canopy conditions on biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment*, *98*(2), 356–370. https://doi.org/10.1016/j.rse.2005.07.012

Næsset, E. (2007). *Airborne laser scanning as a method in operational forest inventory: Status of accuracy assessments accomplished in Scandinavia: Scandinavian Journal of Forest Research: Vol 22, No 5*. https://www.tandfonline.com/doi/abs/10.1080/02827580701672147

Nelson, R. (2014). How did we get here? An early history of forestry lidar 1. *Canadian Journal of Remote Sensing*, *39*, S6–S17. https://doi.org/10.5589/m13-011

Pan, Y., Birdsey, R., Fang, J., Houghton, R., Kauppi, P., Kurz, W., Phillips, O., Shvidenko, A., Lewis, S., Canadell, J., Ciais, P., Jackson, R., Pacala, S., McGuire, A., Piao, S., Rautiainen, A., Sitch, S., & Hayes, D. (2011). A Large and Persistent Carbon Sink in the World's Forests. *Science (New York, N.Y.)*, *333*, 988–993. https://doi.org/10.1126/science.1201609

Pang, Z., Zhang, G., Tan, S., Yang, Z., & Wu, X. (2022). Improving the Accuracy of Estimating Forest Carbon Density Using the Tree Species Classification Method. *Forests*, *13*(12), Article 12. https://doi.org/10.3390/f13122004

Roussel, J., Auty, D., (2023). *Airborne LiDAR Data Manipulation and Visualization for Forestry Applications*. R package version 4.0.3, https://cran.r-project.org/package=lidR.

Roussel, J., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R., Meador, A. S., 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.

Schumacher, J., Rattay, M., Kirchhöfer, M., Adler, P., & Kändler, G. (2019). Combination of Multi-Temporal Sentinel 2 Images and Aerial Image Based Canopy Height Models for Timber Volume Modelling. *Forests*, *10*(9), Article 9. https://doi.org/10.3390/f10090746

Scott, C. T., & Gove, J. H. (2002). *Encyclopedia of environmetrics*. Wiley.

Sheridan, R. D., Popescu, S. C., Gatziolis, D., Morgan, C. L. S., & Ku, N.-W. (2015). Modeling Forest Aboveground Biomass and Volume Using Airborne LiDAR Metrics and Forest Inventory and Analysis Data in the Pacific Northwest. *Remote Sensing*, *7*(1), Article 1. https://doi.org/10.3390/rs70100229

Simard, S. W., Roach, W. J., Defrenne, C. E., Pickles, B. J., Snyder, E. N., Robinson, A., & Lavkulich, L. M. (2020). Harvest Intensity Effects on Carbon Stocks and Biodiversity Are Dependent on Regional Climate in Douglas-Fir Forests of British Columbia. *Frontiers in Forests and Global Change*, *3*. https://www.frontiersin.org/articles/10.3389/ffgc.2020.00088

Soucy, A., De Urioste-Stone, S., Rahimzadeh-Bajgiran, P., Weiskittel, A., & McGreavy, B. (2020). Understanding characteristics of forest professionals and small woodlot owners for communicating climate change adaptation. *Trees, Forests and People*, *2*, 100036. https://doi.org/10.1016/j.tfp.2020.100036

Ståhl, G. (1992). A study on the quality of compartmentwise forest data acquired by subjective inventory methods. *Rapport - Sveriges Lantbruksuniversitet, Institutionen foer Biometri och Skogsindelning (Sweden)*. https://scholar.google.com/scholar_lookup?title=A+study+on+the+quality+of+compartmentwise+forest+data+acquired+by+subjective+inventory+methods&author=Staahl%2C+G.&publication_year=1992

Stephenson, N. L., Das, A. J., Condit, R., Russo, S. E., Baker, P. J., Beckman, N. G., Coomes, D. A., Lines, E. R., Morris, W. K., Rüger, N., Álvarez, E., Blundo, C., Bunyavejchewin, S., Chuyong, G., Davies, S. J., Duque, Á., Ewango, C. N., Flores, O., Franklin, J. F., … Zavala, M. A. (2014). Rate of tree carbon accumulation increases continuously with tree size. *Nature*, *507*(7490), Article 7490. https://doi.org/10.1038/nature12914

Sun, W., & Liu, X. (2019). Review on carbon storage estimation of forest ecosystem and applications in China. *Forest Ecosystems*, *7*(1), 4. https://doi.org/10.1186/s40663-019-0210-2

Tewari, V. P. (2016). Forest inventory, assessment, and monitoring, and long-term forest observational studies, with special reference to India. *Forest Science and Technology*, *12*(1), 24–32. https://doi.org/10.1080/21580103.2015.1018962

Tomppo, E. (2004). RESOURCE ASSESSMENT | Forest Resources. In J. Burley (Ed.), *Encyclopedia of Forest Sciences* (pp. 965–973). Elsevier. https://doi.org/10.1016/B0-12-145160-7/00156-3

USGS. (n.d.). *Landsat Missions | U.S. Geological Survey*. USGS. Retrieved from https://www.usgs.gov/landsat-missions

USGS. (n.d.). *Topographic Data Quality Levels (QLs) | U.S. Geological Survey*. USGS. Retrieved from https://www.usgs.gov/3d-elevation-program/topographic-data-quality-levels-qls

Vorster, A. G., Evangelista, P. H., Stovall, A. E. L., & Ex, S. (2020). Variability and uncertainty in forest biomass estimates from the tree to landscape scale: The role of allometric equations. *Carbon Balance and Management*, *15*(1), 8. https://doi.org/10.1186/s13021-020-00143-6

Waring, B., Neumann, M., Prentice, I. C., Adams, M., Smith, P., & Siegert, M. (2020). Forests and Decarbonization – Roles of Natural and Planted Forests. *Frontiers in Forests and Global Change*, *3*. https://www.frontiersin.org/articles/10.3389/ffgc.2020.00058

Waser, L. T., Fischer, C., Wang, Z., & Ginzler, C. (2015). Wall-to-Wall Forest Mapping Based on Digital Surface Models from Image-Based Point Clouds and a NFI Forest Definition. *Forests*, *6*(12), Article 12. https://doi.org/10.3390/f6124386

Weinstein, B. G., Marconi, S., Bohlman, S. A., Zare, A., Singh, A., Graves, S. J., & White, E. P. (2021). A remote sensing derived data set of 100 million individual tree crowns for the National Ecological Observatory Network. *ELife*, *10*, e62922. https://doi.org/10.7554/eLife.62922

Weiser, H., Schäfer, J., Winiwarter, L., Krašovec, N., Fassnacht, F. E., & Höfle, B. (2022). Individual tree point clouds and tree measurements from multi-platform laser scanning in German forests. *Earth System Science Data*, *14*(7), 2989–3012. https://doi.org/10.5194/essd-14-2989-2022

Weiskittel, A., & Kershaw, J. (2018). *The Acadian Variant of the Forest Vegetation Simulator: Continued development and evaluation.*

What are the acquisition schedules for the Landsat satellites? | U.S. Geological Survey. (n.d.). Retrieved April 4, 2023, from https://www.usgs.gov/faqs/what-are-acquisition-schedules-landsat-satellites

White, J., Arnett, J., Wulder, M., Tompalski, P., & Coops, N. (2015). Evaluating the impact of leaf-on and leaf-off airborne laser scanning data on the estimation of forest inventory attributes with the area-based approach. *Canadian Journal of Forest Research*, *45*, 1498–1513. https://doi.org/10.1139/cjfr-2015-0192

White, J. C., Coops, N. C., Wulder, M. A., Vastaranta, M., Hilker, T., & Tompalski, P. (2016). Remote Sensing Technologies for Enhancing Forest Inventories: A Review. *Canadian Journal of Remote Sensing*, *42*(5), 619–641. https://doi.org/10.1080/07038992.2016.1207484

White, J. C., Tompalski, P., Vastaranta, M., Wulder, M. A., Saarinen, N., Stepper, C., & Coops, N. C. (2017). *A model development and application guide for generating an enhanced forest inventory using airborne laser scanning data and an area-based approach*. 48.

White, J., Wulder, M., Varhola, A., Vastaranta, M., Coops, N., Cook, B., Pitt, D., & Woods, M. (2013). A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. In *Canadian Forest Service, Information Report: Vol. FI-X*. https://doi.org/10.5558/tfc2013-132

Woodall, C. W., Heath, L. S., Domke, G. M., & Nichols, M. C. (2011). *Methods and equations for estimating aboveground volume, biomass, and carbon for trees in the U.S. forest inventory, 2010* (NRS-GTR-88; p. NRS-GTR-88). U.S. Department of Agriculture, Forest Service, Northern Research Station. https://doi.org/10.2737/NRS-GTR-88

Woods, M., Pitt, D., Penner, M., Lim, K., Nesbitt, D., Etheridge, D., & Treitz, P. (2011). Operational implementation of a LiDAR inventory in Boreal Ontario. *The Forestry Chronicle*, *87*(04), 512–528. https://doi.org/10.5558/tfc2011-050

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

Yu, Y., Saatchi, S., Domke, G. M., Walters, B., Woodall, C., Ganguly, S., Li, S., Kalia, S., Park, T., Nemani, R., Hagen, S. C., & Melendy, L. (2022). Making the US national forest inventory spatially contiguous and temporally consistent. *Environmental Research Letters*, *17*(6), 065002. https://doi.org/10.1088/1748-9326/ac6b47

Zavyalov, V., Bingham, G., Wojcik, M., Johnson, H., & Struthers, M. (2009). Application of principal component analysis to lidar data filtering and analysis. *Proceedings of SPIE - The International Society for Optical Engineering*, *7479*. https://doi.org/10.1117/12.830126

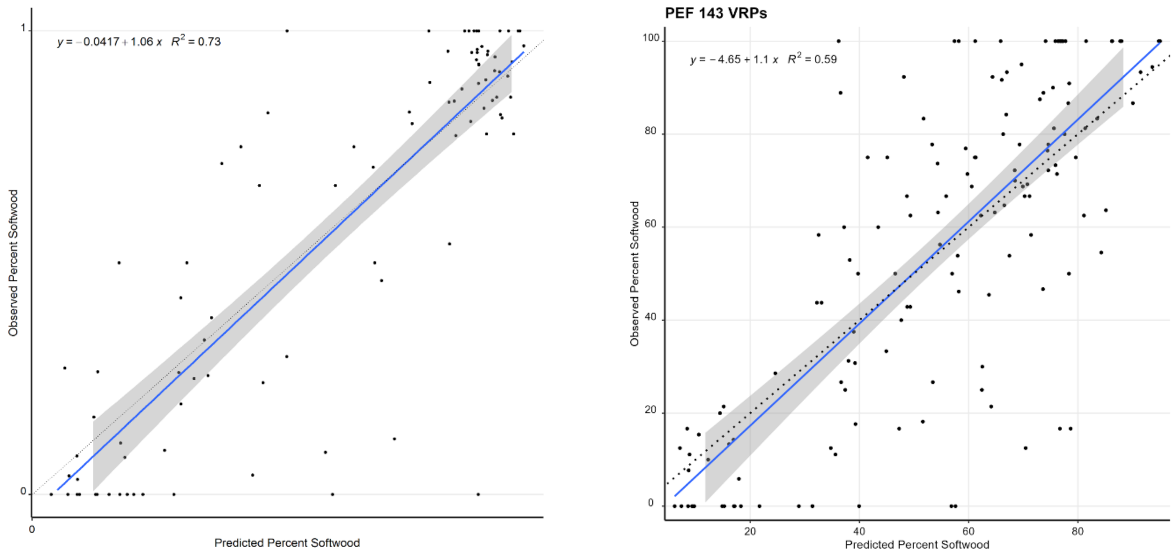
APPENDICES

APPENDIX A

MODEL COMPARISONS, SCATTER PLOTS, AND RASTER MODEL OUTPUTS USED AS SUPPLEMENTARY MATERIAL IN CHAPTER 1.

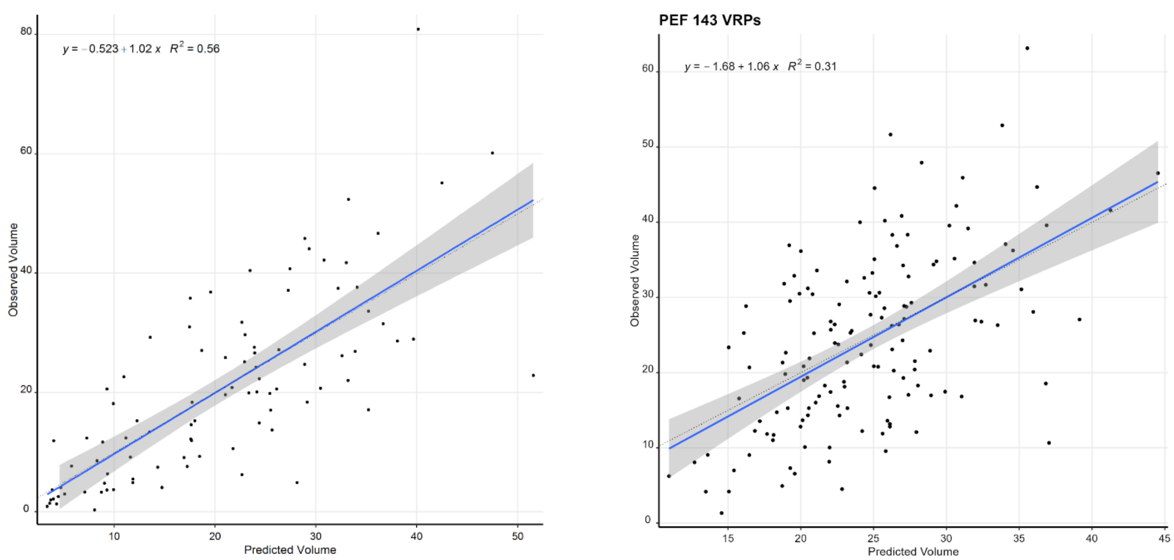
Table A.1. R^2, normalized root mean square error (NRMSE), coefficient of variation (CV), and normalized mean bias (NMB) metrics for model outputs predicting percent softwood, volume, basal area, and tree count in the Ashland West, Rangeley, and Baskahegan study areas.

AOI	Percent Softwood (%)				Total Volume (CDS/ac)				Basal Area (ft²/ac)				Tree Count (trees/ac)			
	R^2 (%)	NRMSE (%)	CV (%)	NMB (%)	R^2 (%)	NRMSE (%)	CV (%)	NMB (%)	R^2 (%)	NRMSE (%)	CV (%)	NMB (%)	R^2 (%)	NRMSE (%)	CV (%)	NMB (%)
Ashland West	58.9	43.3	44.4	0.35	43.6	59.2	55.1	-1.23	30.2	58.8	56.4	-1.31	15.8	86.2	83.9	-2.95
Rangeley	75.1	28.4	29.1	0.63	47.9	35.7	36.8	-0.65	37.7	33.5	34.0	-0.42	32.7	35.1	36.2	-0.46
Bask-ahegan	75.8	0.3	30.2	-0.08	59.2	47.2	47.8	-1.15	54.8	42.5	41.6	-1.22	22.4	90.1	45.6	-3.29



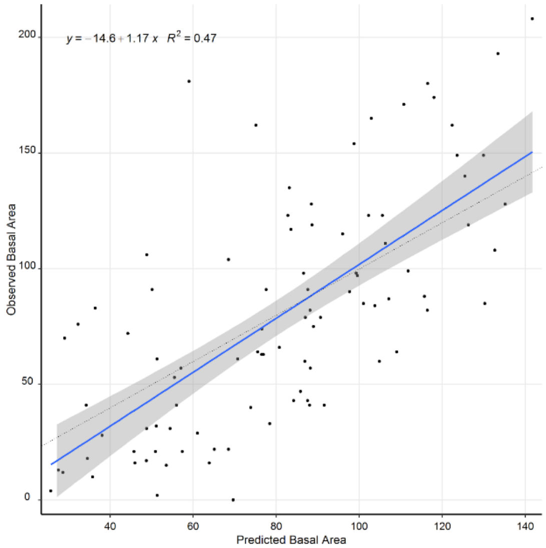
Baskahegan **PEF VRP 143**

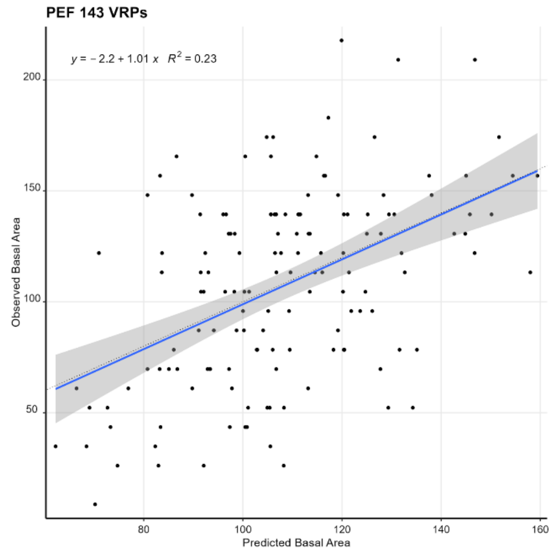
Figure A.1. Scatter plots of observed percent softwood as a function of predicted percent softwood in the Baskahegan (left) and PEF (143 VRPs, right) study sites.



Baskahegan **PEF VRP 143**

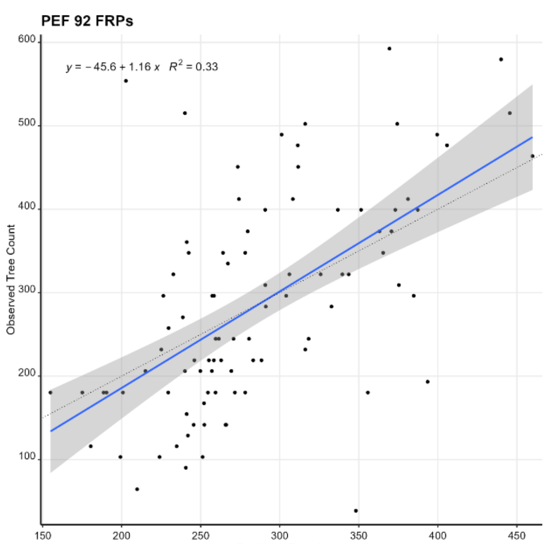
Figure A.2. Scatter plots of observed volume as a function of predicted volume in the Baskahegan (left) and PEF (143 VRPs, right) study sites.

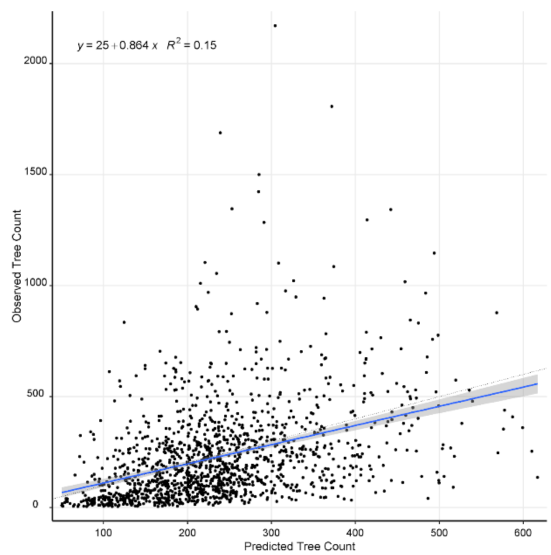




Baskahegan **PEF VRP 143**

Figure A.3. Scatter plots of observed basal area as a function of predicted basal area in the Baskahegan (left) and PEF (143 VRPs, right) study sites.





PEF FRP 92 **Ashland West**

Figure A.4. Scatter plots of observed tree count as a function of predicted tree count in the PEF (92 FRPs, left) and Ashland West (right) study sites.

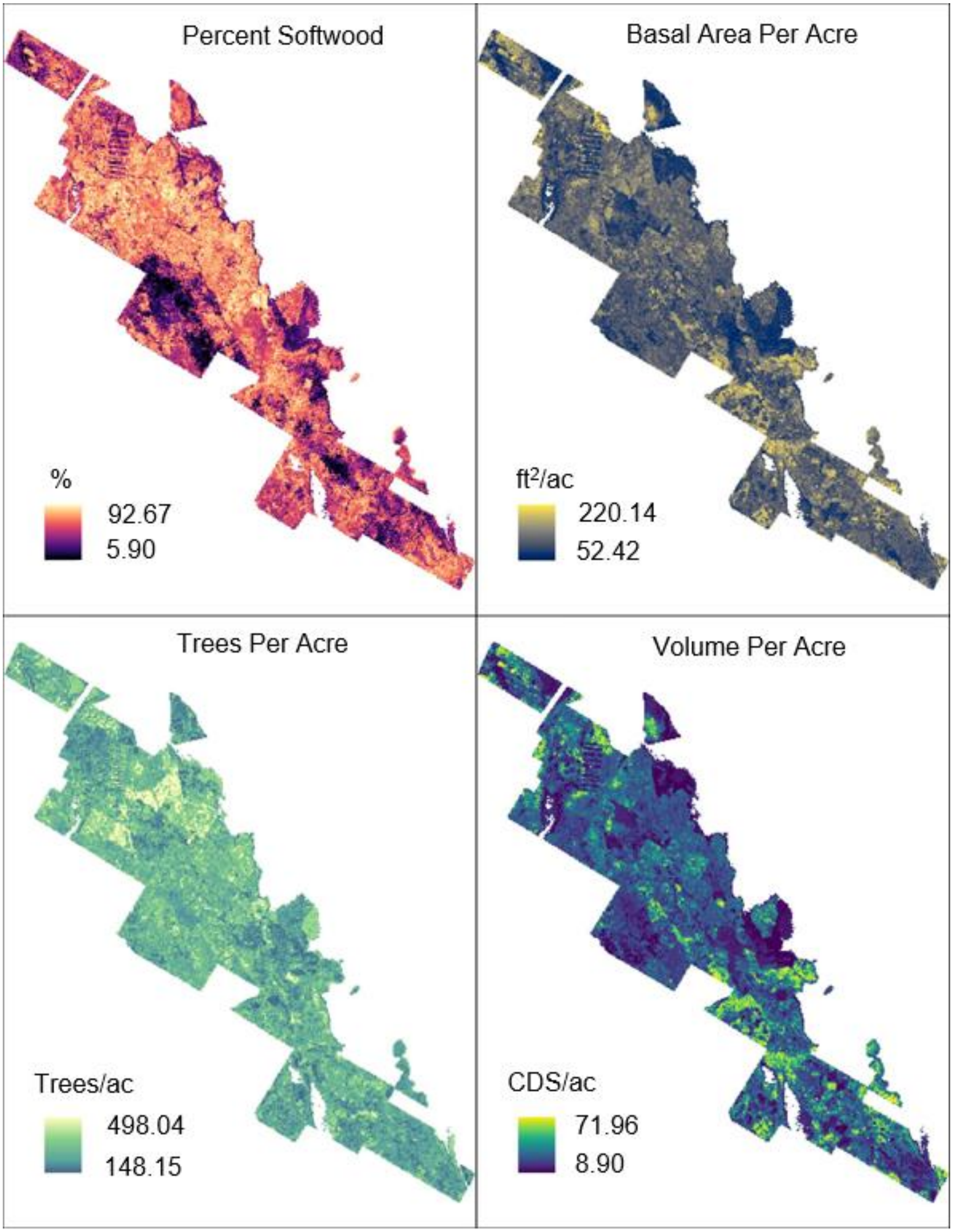


Figure A.5. Model outputs for the PEF using 92 FRPs predicting the percent softwood, basal area per hectare, trees per hectare, and cords per hectare. These predictions are at 65.6-foot resolution.

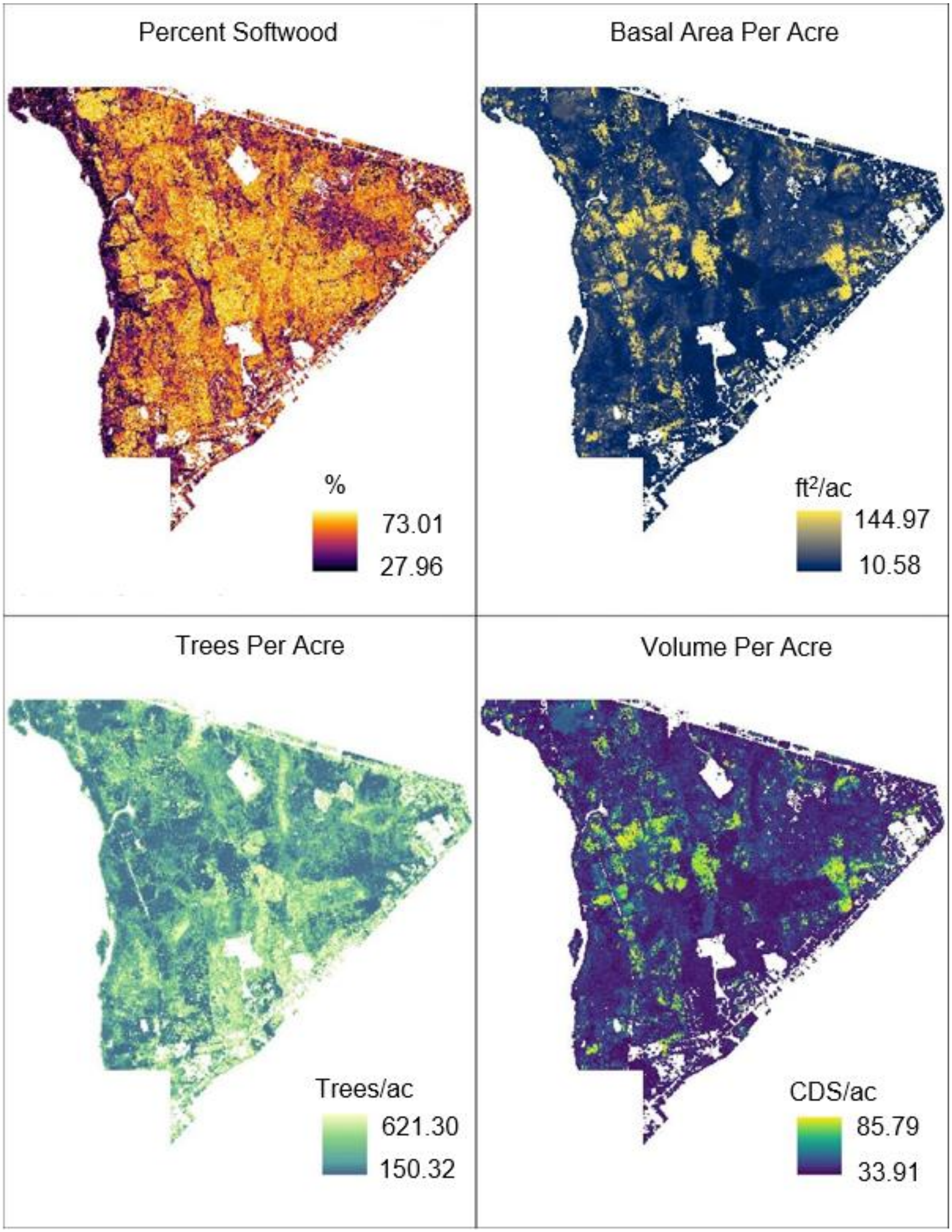


Figure A.6. Model outputs for the Demeritt predicting the percent softwood, basal area per hectare, trees per hectare, and cords per hectare. These predictions are at 65.6-foot resolution.

APPENDIX B

BAR PLOTS DESCRIBING ATTRIBUTES OF THE PEF BY FOREST TYPE USED IN CHAPTER 2.

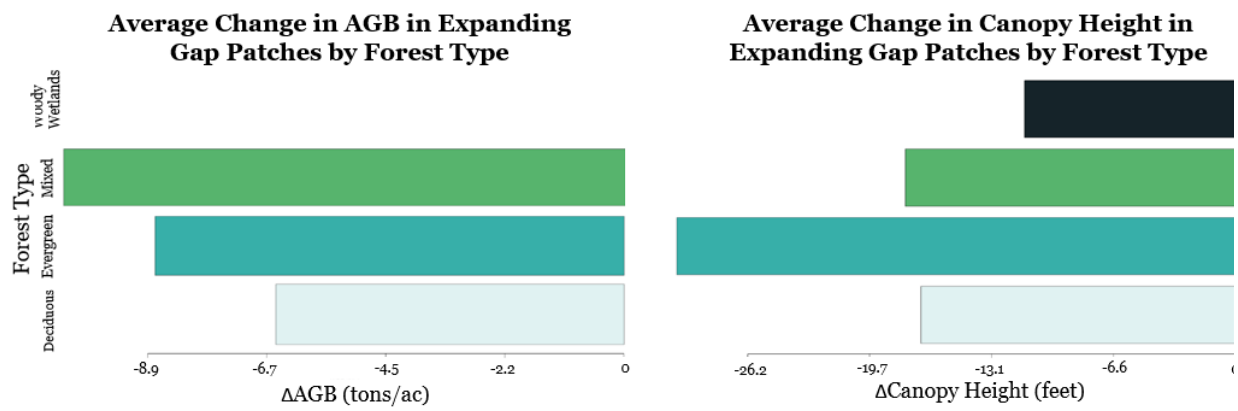


Figure B.1. Chart showing the average change from 2017 to 2021 in AGB (left) and canopy height (right) in expanding gap shelterwoods in the PEF by forest type.

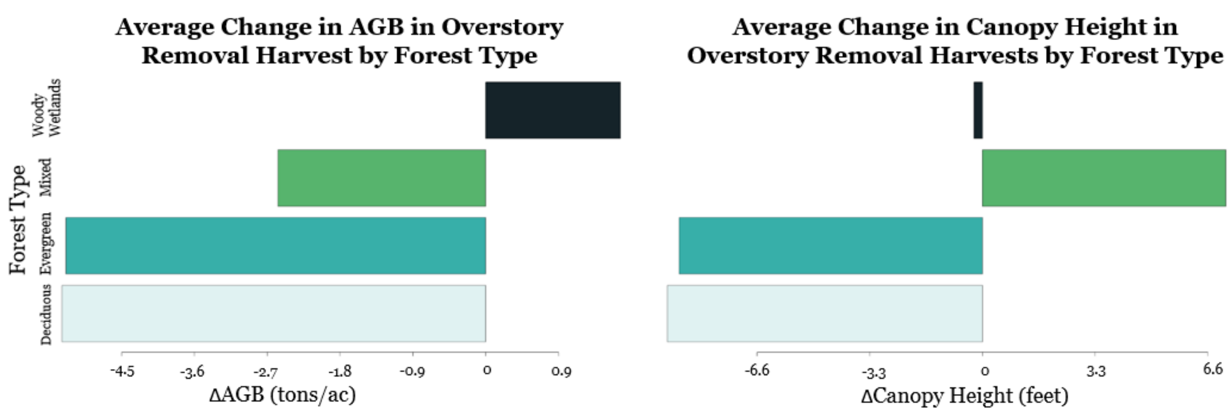


Figure B.2. Chart showing the average change from 2017 to 2021 in AGB (left) and canopy height (right) in overstory removals in the PEF by forest type.

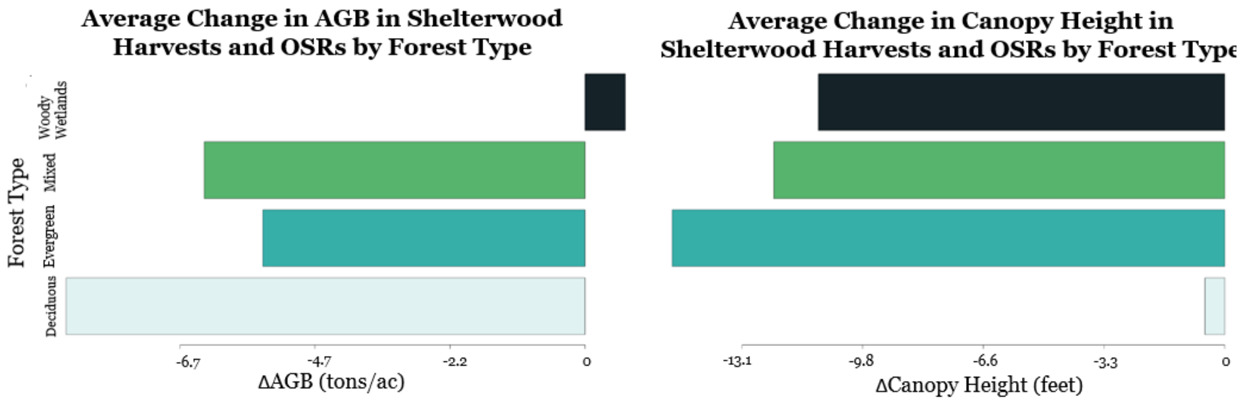


Figure B.3. Chart showing the average change from 2017 to 2021 in AGB (left) and canopy height (right) in shelterwoods with overstory removals in the PEF by forest type.

BIOGRAPHY OF THE AUTHOR

Stephanie Willsey was born in Cape Coral, Florida on July 28, 1994. She was raised in Elkhart, Indiana and graduated from Concord High School in 2013. She attended Indiana University and graduated in 2017 with a bachelor's degree in psychology. She enrolled at the University of Maine two years later as a Master of Forestry student and completed degree requirements in spring of 2021. In fall of 2021, she re-enrolled at the School of Forest Resources as a graduate research assistant. Stephanie is a candidate for the Master of Science degree in Forest Resources from the University of Maine in May 2023.