Extensive Riparian Vegetation Monitoring, Model Transferability Testing

Agreement No. 93-098916 (1)

Andrew Cooke[1], Warren Devine[2]

[1] Research Scientist, Natural Resource Spatial Informatics Group (NRSIG), Precision Forestry Cooperative (PFC), University of Washington (UW)

[2] Data Management Specialist, Washington State Department of Natural Resources (WADNR)

Acknowledgements

Editors

Table of Contents

List of Figures

List of Tables

Purpose

The goal of this project is to test the transferability of several forest inventory models developed in the Mashel watershed under the "Extensive Riparian Vegetation Monitoring - Remote Sensing Pilot Study Agreement No. IAA 16-205". The models were tested using forest inventory plots that were established in the Olympic Experimental State Forest (OESF). This project examined the DBH, Basal Area, and Stand Density models. It is hypothesized that forest structures associated with different forest types impact the accuracy of forest inventory models, limiting their transferability from one forest type to another.

Background

Model Transferability

There is an increasing amount of research into the transferability of inventory models, but it is a relatively new area of study, and it is not yet known how effectively plot level inventory models can be applied across the different forest types present in Washington State. Scandinavian researchers have looked at transferring tree-level models between sites (Karjalainen, et al. 2019). Transferring basal area and stem density models from sites with plot data to sites without plot data has been tested in Northern Idaho (Fekety, Falkowski, et al. 2018). Researchers in British Columbia tested the transferability of Lorey's height, quadratic mean diameter, and volume models under various scenarios at sites on northern Vancouver Island (Tompalski, et al. 2019). The transferability of models across time at the same sites has been tested in Northern Idaho (Fekety, Falkowski and Hudak 2015). These examples present some of the ongoing work, and research in this area will likely become more common.

Extensive Riparian Vegetation Monitoring

From 2015 to 2017, the Precision Forestry Cooperative (PFC), at the University of Washington (UW) School of Environmental and Forest Sciences (SEFS), undertook the Extensive Riparian Vegetation Monitoring – Remote Sensing Pilot Study (Moskal, et al. 2017). The purpose of the Pilot Study was to investigate the effectiveness of using remote sensing methods as the basis for monitoring the status and trends of riparian forest stands on private lands in Washington State. This Pilot Study took place in the Mashel Watershed in the Cascade Mountains of Pierce County, and demonstrated the viability of remote sensing for monitoring certain riparian vegetation metrics. However, the applicability of the models developed in the Mashel watershed to other forest types in other parts of the State is unknown.

In 2018, PFC conducted a further scoping study (Cooke and Moskal 2018) to provide recommendations on where the next phase of the Extensive Riparian Vegetation Monitoring Project could take place. Part of this study was to identify areas of the State with available LIDAR and forest types that were substantially different from the Mashel watershed. Testing Mashel inventory models in these other forest types with different climatic regimes and geographic locations would help determine the number of models necessary for the monitoring of riparian forests statewide. One of the locations identified as a potential future study location was the coastal Olympic Peninsula.

Status and Trends Monitoring in the Olympic Experimental State Forest

Beginning in 2012, the Washington State Department of Natural Resources (WADNR) began a long term monitoring project in the Olympic Experimental State Forest (OESF) to see how WADNR's state lands Habitat Conservation Plan (HCP) and the OESF Forest Land Plan affect the habitat conditions in forests managed for timber production and habitat conservation. As part of this project, 62 Type 3 watersheds

were selected for monitoring in the OESF. Type 3 streams (aka Type F) are defined in WAC 222-16-031, but are, generally speaking, the smallest fish-bearing streams, and are relatively low gradient compared to headwater Type 4 streams (aka Type N) and at least two feet wide. Of these watersheds, 52 are on DNR-managed land, six are in the Olympic National Forest, and four are in the adjacent Olympic National Park (Figure 1) (Minkova, Teodora; Devine, Warren 2016). Sample reaches of at least 100 meter length were established for each watershed on the most-downstream Type 3 section of stream.
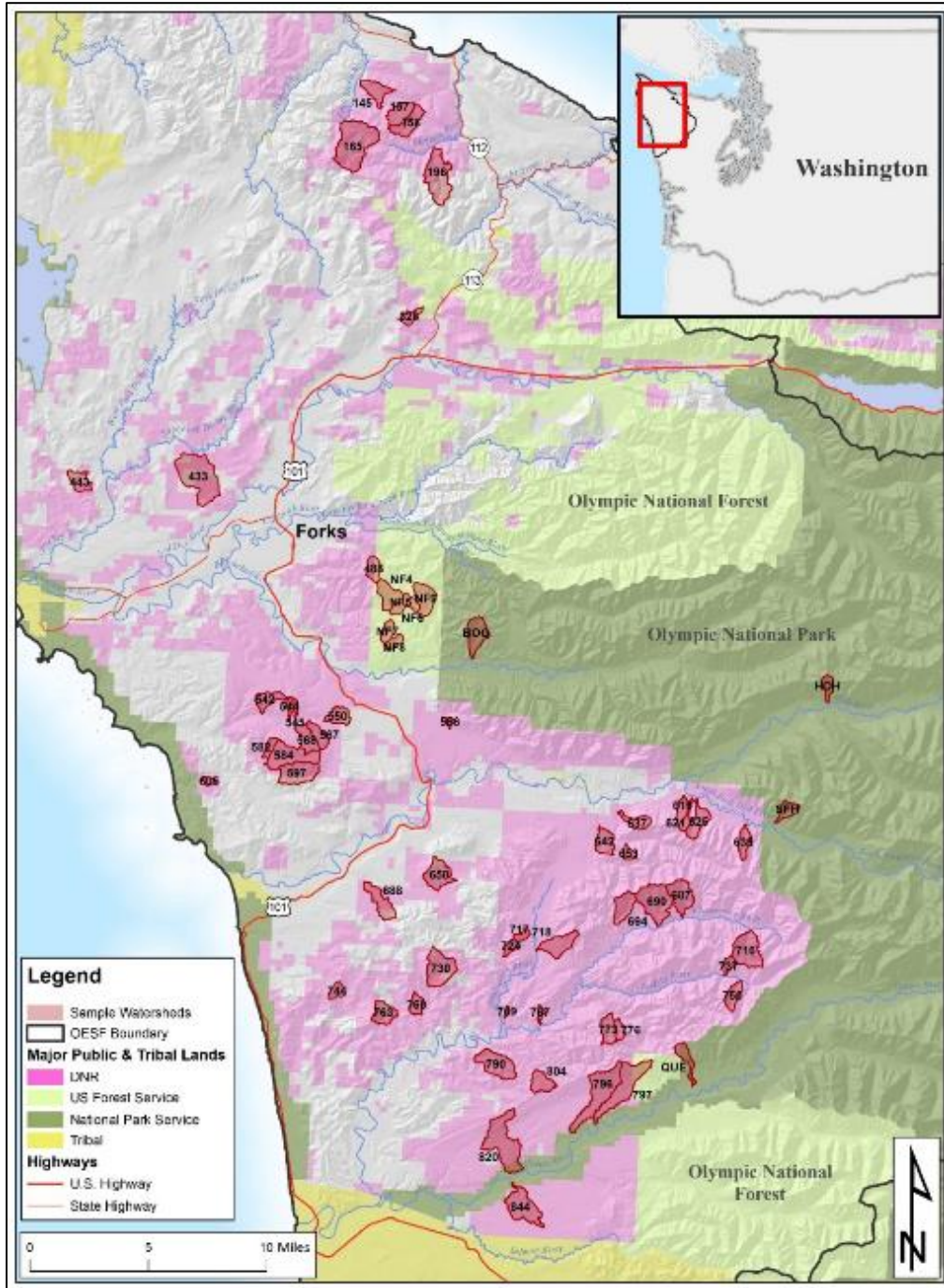
Figure 1: Map of the study area. Fifty monitored watersheds are located on DNR-managed land, six are located in Olympic National Forest, and four are located in Olympic National Park. Source: Warren Devine, WADNR.

Olympic Experimental State Forest Vegetation Plot Description

As part of the OESF monitoring project, Riparian Vegetation plots were established between 2013 and 2018, in all sample watersheds (Minkova, Teodora; Foster, Alex 2017).

Along each sample reach there are six stream transects (A – F) spaced at equal intervals of approximately 20 meters, perpendicular to the channel azimuth. Two fixed area plots were established at two randomly selected transects, one on each side of the stream (Figure 2). The plots were established by locating the plot centerline parallel to the selected transects and extending the transects away from the stream. The plot centerlines are 60 meters long (horizontal distance). The plot end lines were located by traversing 15 meters away from the plot centerline at each end of and perpendicular to the centerline.

Plots are 60 m by 30 m rectangles (horizontal distance), with an area of 0.18 ha (0.44 ac). Each plot was also subdivided along its 60-m centerline into three zones. Each zone was 20 m long and 30 m wide, with an area of 0.06 ha (0.15 ac). Zone 1 is nearest to the stream, Zone 3 is farthest from the stream.

All trees in the plots meeting the following criteria were measured (Minkova, Teodora; Foster, Alex 2017):

1. Live tree >12.5 cm diameter at breast height (DBH; 1.3 m above the forest floor).
2. Dead tree >12.5 cm DBH and ≥6.0 m in height.
3. Dead tree >12.5 cm DBH and < 6.0 m in height with an intact top.

Each qualifying tree was tagged and measured for diameter (DBH), species, and whether it was alive or dead. The zone in which each tree was located was also noted.

Sample Reach (A through F ≥ 100 m)

Channel cross-sections (A-F) spaced at 5 equal intervals along the sample reach; monumented with rebar above bankfull stage

Zone 3

Zone 2

Zone 1

Bankfull stage

E

F

D

Overstory plot, 60 by 30 meters

Center line/ travel corridor

C

Reference point (RP)

B

A

Hemispherical canopy photo point

Nested understory plot

Both spaced at 0, 10, 20, 40, and 60 meters from rebar

4m

6m

2m

Center line

Understory plot, enlarged to show detail

Figure 2: Vegetation Plot layout. Source: (Minkova, Teodora; Foster, Alex 2017)

LIDAR Availability

The availability of LIDAR data is a limiting factor on testing Mashel models. As part of the 2018 Scoping Project (Cooke and Moskal 2018), the locations of publically available LIDAR were examined. Figure 3 shows the locations of publically available LIDAR data sets on the Olympic Peninsula and the year of their acquisition.
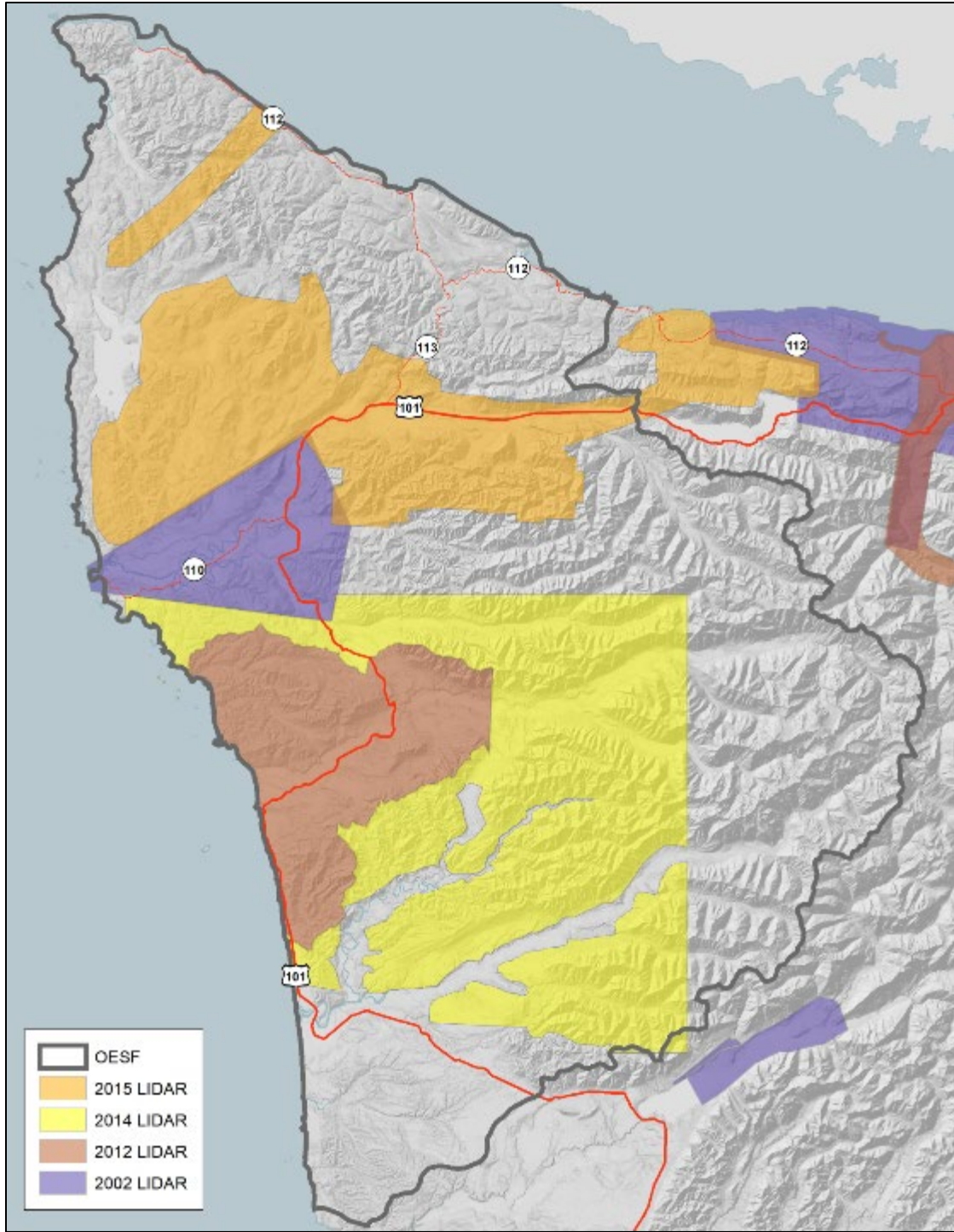
Figure 3: Locations of available LIDAR datasets within the OESF.

Methods

Overview

If forest inventory models are transferable from one forest type to another, it is expected that the models from one forest type will have similar accuracies when used in another forest type as they do when used at their original location. To test this theory, the following steps were used. They are described in further detail below.

1. Select inventory plots in the OESF with LIDAR coverage
2. Collect precise plot locations using survey grade GPS units. These locations are used to clip out the LIDAR points in each OESF plot from the full acquisition point clouds
3. Summarize OESF inventory data to plot level values
4. Process LIDAR points for each plot into plot level LIDAR metrics
5. Use Mashel models to predict OESF plot values, using data from steps 3 and 4. Examine model accuracies
6. Develop new OESF inventory models, using data from steps 3 and 4, and use these models to predict summarized Mashel plot values from the Pilot Study. Examine model accuracies.

In total 30 model-plot combinations were tested (Table 1). Further details on the tested models are provided below in the sections Application of Mashel Models and New OESF Model Development.

Table 1: Combinations of models and plots tested in this report.

Model Source	Plots	Models Tested		
Mashel	OESF Full	Existing Basal Area	Existing Diameter	Existing Density
		Refit Basal Area	Refit Diameter	Refit Density
		Adapted Basal Area	Adapted Diameter	Adapted Density
Mashel	OESF Zone 1	Existing Basal Area	Existing Diameter	Existing Density
		Refit Basal Area	Refit Diameter	Refit Density
		Adapted Basal Area	Adapted Diameter	Adapted Density
OESF Full	OESF Full	Basal Area	Diameter	Density
OESF Zone 1	OESF Zone 1	Basal Area	Diameter	Density
OESF Full	Mashel	Basal Area	Diameter	Density
OESF Zone 1	Mashel	Basal Area	Diameter	Density

Plot Selection

Plot selection was performed by Warren Devine of WADNR. Each of the 62 OESF sample watersheds has two plots, for a total of 124 available plots. Not all of those plots have LIDAR coverage, and those that do are covered by one of four different LIDAR acquisitions.

The two acquisitions that covered the most plots were selected. The 2012 and 2014 acquisitions within OESF cover a combined total of 66 plots. The 2012 LIDAR data covers 28 plots, and the 2014 data covers 38 plots. The target was to collect 40 plots over the one month of available fieldwork time. Metadata for the LIDAR acquisitions is available in vendor reports (Watershed Sciences 2012) (Woolpert 2014).

The majority of the plots were installed between 2013 and 2015, meaning the 2012 LIDAR was flown one to three years before the plots were measured, and the 2014 LIDAR was flown one year before to one year after the plot measurements. Minimizing the time between LIDAR and plot data acquisition increases the likelihood that both datasets measured the same forest conditions. The Mashel Pilot Study had a time difference of five years between LIDAR acquisition and field plot measurement.

Selected Models

The field crews for the OESF Riparian Vegetation Plots measured diameter (DBH), species, whether the trees were alive or dead, and the zone in which each tree was located. This limits the number of Mashel models that can be tested to the diameter (quadratic mean diameter for the plot), basal area, and density models.

GPS Data Collection Protocol

The GPS data collection protocol was developed by Warren Devine (WADNR) and Andrew Cooke (UW).

OESF plots were established using the plot centerline, which have fence posts, rebar, or flagging along the centerline as the only permanent markers. Plot corners may have flagging remaining from when the plots were measured, but are otherwise unmarked. All large trees (DBH >= 12.5 cm) within the plots were measured and tagged. The 12.5 cm diameter threshold was defined in the WADNR's vegetation monitoring protocol (Minkova, Teodora; Foster, Alex 2017).

Because the plot centerlines were the only parts of the plot with permanent markers, it was decided that the best method for locating the field plots was to collect GPS positions on the two endpoints of the plot centerlines (Figure 4).
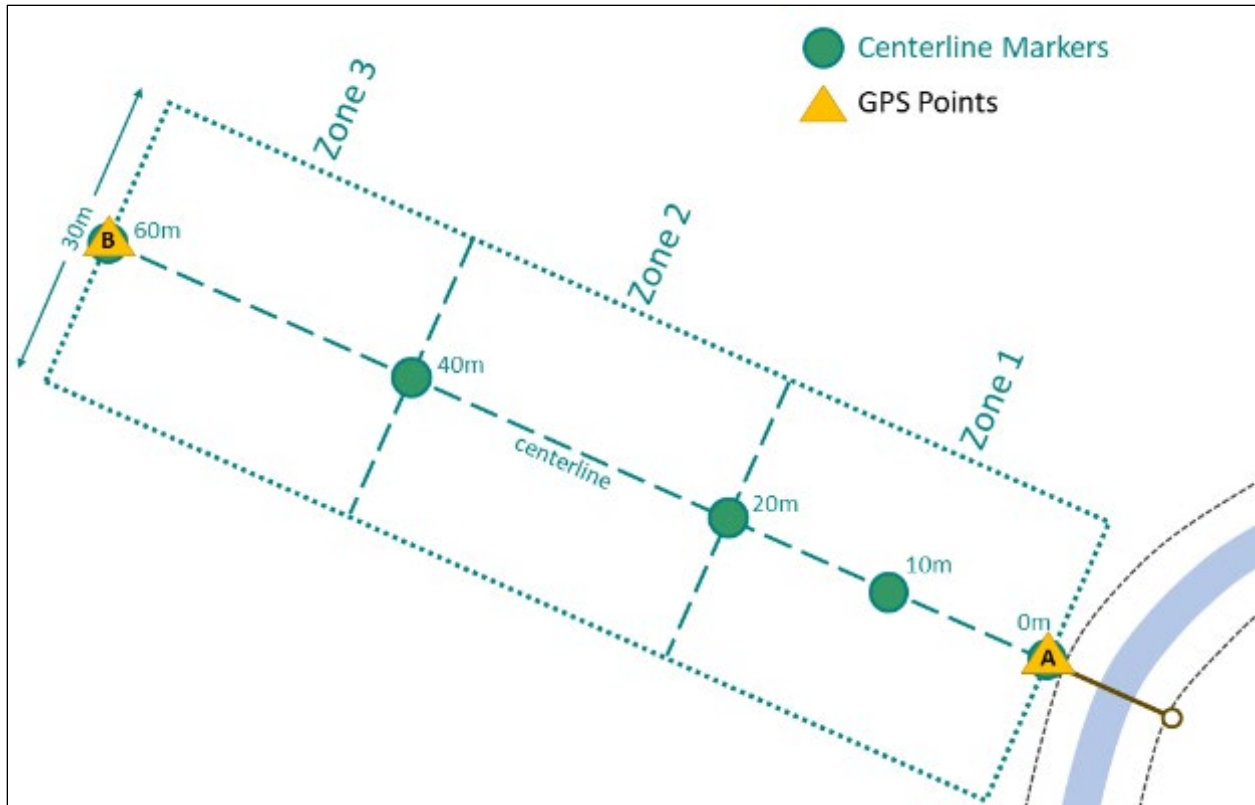
Figure 4: GPS position locations for each plot. GPS point A is nearest to the stream on the 0 meter centerline marker, and GPS point B is farthest from the stream on the 60 meter centerline marker.

GPS position data was collected by Anchal Rikhi (WADNR) during one month of fieldwork in April, 2019, using two Javad Triumph 2 GNSS (Global Navigation Satellite System) receivers. The receivers were configured to use United States' GPS satellites and Russia's GLONASS satellites.

Literature (Mcgaughey, et al. 2017) suggests that a minimum occupation time of 15 minutes should be used to collect GPS positions under dense forest canopy. The GPS units used for this project were set up to collect one position measurement per second, meaning that in ideal conditions, 15 minutes worth of data records 900 position measurements. Assuming that the GPS may not be able to calculate a position every second due to terrain, forest cover, and satellite geometry, the field protocol specified that the GPS should record 900 position measurements at each location even if that takes longer than 15 minutes.

Measuring two GPS points per plot, each with an occupation time of at least 15 minutes, provides enough information to build accurate plot polygons, while also being quick enough to collect enough plots in the available month of fieldwork time.

Data Processing
The GPS points were used to develop polygons for each plot by connecting the A and B GPS points for each plot with a line, and buffering these lines 15 meters on each side. The resulting plots are 30 meters wide and have varying lengths, depending on the positions of the GPS points.

Once plot polygons were created, they were subdivided along their centerlines every 20 meters from GPS point A. This resulted in three zone polygons for each plot with Zones 1 and 2 each having a length of 20 meters, and Zone 3 having a variable length, dependent on overall plot length (Figure 5).
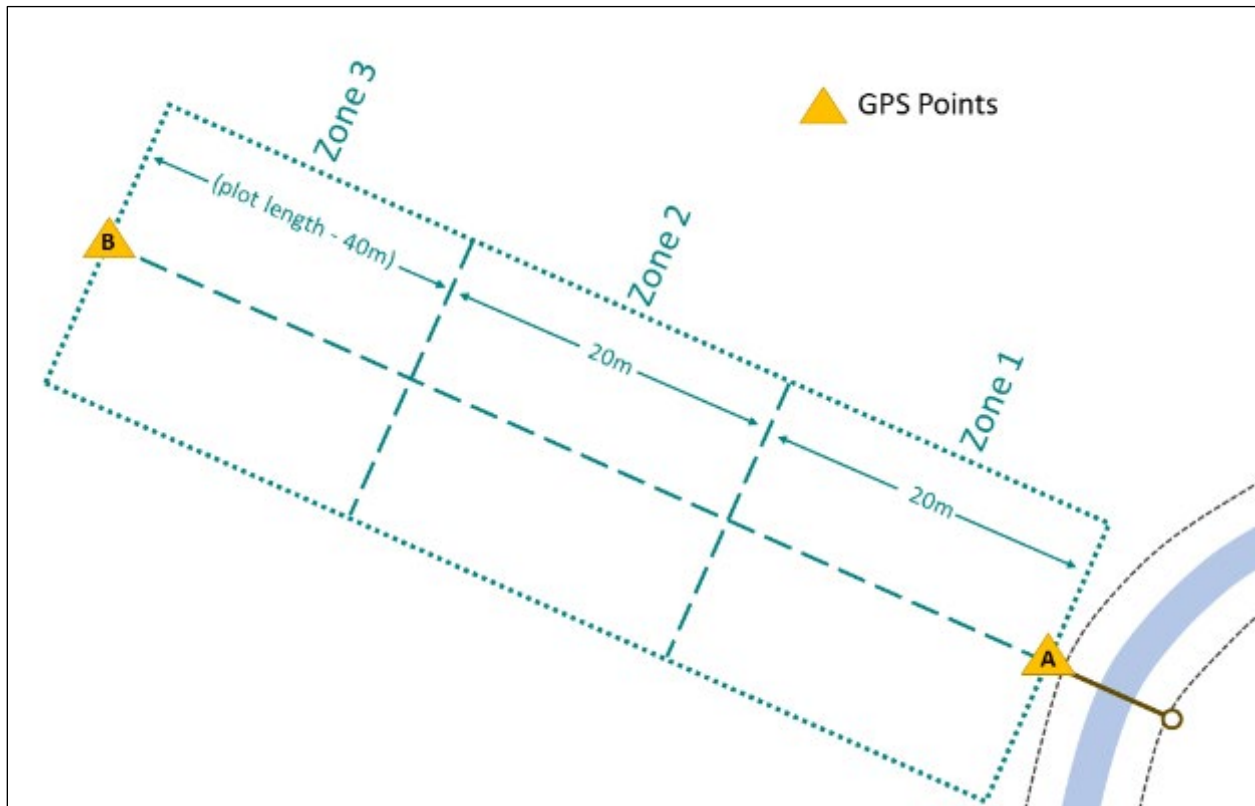
Figure 5: Length of plot zones.

One important factor considered in this study is plot size. Plots in the Mashel Pilot Study were 1/8th acre (42 foot radius) circles. The OESF plots are 0.44 acre rectangles. It was unclear whether or not the differing plot sizes would affect the accuracy of the models. Because the field data collected on OESF plots noted the location of each tree in the three plot zones, it was possible to test the Mashel models against not only the full plots, but also against individual or combined zones.

Zone 1 was chosen as the subzone against which to test the Mashel models because:

- GPS point A has the higher likelihood of the two GPS points of being accurately located
- any errors in centerline azimuth will exaggerate misalignment between the plot and LIDAR data as the distance from GPS point A increases
- using Zone 1 allows plot length differences from 60 meters to be ignored

The full plot and Zone 1 polygons were used to clip out LIDAR points for each plot from the full acquisitions using the lidR package (Roussel 2018) in R (R Core Team 2018).

The CloudMetrics executable, which is part of the Fusion LIDAR processing software (McGaughey 2019), was used to calculate plot summary metrics for each full plot and Zone 1 subplot. The same settings used in Mashel Pilot Study were used to calculate the plot metrics for the OESF plots, and metrics were

calculated using all LIDAR returns. The LIDAR points for all plots and subplots were elevation normalized to height above ground, using the vendor supplied digital elevation models.

OESF Plot Summaries

Tree data provided by Warren Devine (WADNR) was summarized to plot level diameter (quadratic mean diameter) in inches, basal area in square feet per acre, and density in trees per acre, for both full plots and Zone 1 subplots.

As with the Mashel Pilot Study, plot summaries were calculated using only live trees, both hardwoods and softwoods. The trees measured in the OESF plots had a diameter of 12.5 cm or larger, which is similar to the 5 inch (12.7cm) minimum diameter used in the Mashel plots. For this project, trees smaller than 5 inches were excluded from the plot summary metrics, for consistency with the Pilot Study. This removed 56 trees with a diameter >= 12.5 cm and < 5 inches.

Application of Mashel Models

The models from the Mashel Pilot Study were used to predict values for full and Zone 1 OESF plots in three different ways: an existing model, a refit model, and an adapted model, each described below. The existing models are the only ones that can be applied without collecting new plot data, but the refit and adapted models are included for comparison, to see if model performance improves if you have access to plot data.

The best performing density model in the Pilot Study used an independent variable that was not available for OESF plots, the stratification bin for each plot. The Mashel watershed was classified into bins using lidar-derived height and cover metrics, and each plot had the bin in which it was located as an attribute. A model without the bin variable was also developed as part of the Pilot Study, which performed slightly less well with an R^2 of 0.46 versus 0.49. The model without the bin variable was used for this project.

The existing Mashel models use the same independent variables, the same coefficient values, and the same terms, including interaction terms (if there are any), as the models from the Pilot Study. These models are the same for both the full plots and the Zone 1 plots. The model variables are described in Table 2.

The refit Mashel models use the same independent variables and the same interaction terms, but have new coefficient values, which are calculated using the OESF plots. These models cannot be applied to new areas without collecting new plot data. The models for the full plots and Zone 1 plots are different. The model variables are described in Table 2.

The adapted Mashel models use the same independent variables, but have new coefficients, which are calculated using the OESF plots, and have non-significant interaction terms removed. These models cannot be applied to new areas without collecting new plot data. The models for the full plots and Zone 1 plots are different. The model variables are described in Table 2.

Table 2: Mashel model variables

Model Variable	Fusion CloudMetrics Variable	Variable Explanation
Crr	Canopy.relief.ratio	(mean height - minimum height) / (maximum height - minimum height), all returns
Ek	Elev.kurtosis	measure of whether distribution of heights are heavy-tailed or light-tailed relative to a normal distribution, all returns. Heavy-tailed distributions have their tails return to zero more slowly than a normal distribution, and light-tailed distributions have their tails return to zero more quickly.
Emax	Elev.maximum	maximum height, all returns
Emm	Elev.MAD.median	Median of the absolute deviations from the overall median height, all returns
P20	Elev.P20	20th percentile of the heights, all returns
Pa2	Percentage.all.returns.above.6.56	percent cover; number of all returns above 2m / number of all returns
Pamd	Percentage.all.returns.above.mode	percent cover; number of all returns above mode height / number of all returns

The tables below provide coefficients for all versions of the Mashel models used in this project.

Table 3: Coefficients for Mashel basal area models

	sqrt(basal area)				
model	Existing	Refit		Adapted	
plot type	Full & Zone 1	Full Plot	Zone 1 Plot	Full Plot	Zone 1 Plot
intercept	4.63914 **	-1.97464	-5.72512	22.241836 *	-5.05004
Emm	-0.39478 **	0.1488	0.22004	-1.147746 *	0.18395 **
Crr	-5.63122	9.66773	9.29147	9.676813 **	8.07385 .
Pa2	0.11590 ***	0.11712 **	0.16163 ***	-0.163778	0.16068 ***
Emm * Crr	1.25605 ***	0.02237	-0.07578		
Emm * Pa2				0.015101 **	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 4: Coefficients for Mashel density models

sqrt(density)					
model	Existing	Refit		Adapted	
plot type	Full & Zone 1	Full Plot	Zone 1 Plot	Full Plot	Zone 1 Plot
intercept	-1.9701	22.7797	-2.0547	35.7279 ***	33.1449 ***
log(Emax)	2.1654	-7.2652	-2.8083	-8.7403 ***	-9.0353 ***
log(Ek)	16.4874 .	15.8119	36.8973	2.3486 *	0.9971
sqrt(Pa2)	0.30570	2.641	3.0966	2.0208 **	2.5695 ***
log(Emax) * log(Ek)	-5.8437 *	-1.5212	-6.4618		
log(Ek) * sqrt(Pa2)	1.5078 *	-0.6504	-0.4767		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 5: Coefficients for Mashel diameter models

sqrt(quadratic mean diameter)					
model	Existing	Refit		Adapted	
plot type	Full & Zone 1	Full Plot	Zone 1 Plot	Full Plot	Zone 1 Plot
intercept	0.605765	-2.22441	-6.15279 *	-0.305427	-5.08732 *
Emm	-0.025342	0.068856	0.30215 .	0.060846 ***	0.36403 **
log(P20)	0.77423 ***	1.317015 *	2.3123 **	0.782004 ***	2.02267 ***
sqrt(Pamd)	0.32355 **	0.311498	0.36136	0.035982	0.01803
Emm * log(P20)	0.022183 **	-0.001679	-0.06028		-0.07608 *
log(P20) * sqrt(Pamd)	-0.106525 **	-0.07906	-0.09045		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

New OESF Model Development

In addition to using the Mashel models, new basal area, diameter, and density models were built using OESF plot data and the OESF LIDAR using the Comprehensive Model Selection process described in the Mashel Pilot Study report (Moskal, et al. 2017). These new models, referred to from here on as the OESF Models, were developed to demonstrate the accuracy that could be expected if native models were built for the OESF using the same procedure as the Mashel Pilot Study. Models were built to estimate full plot and Zone 1 plot inventory. The tables below provide the model coefficients and descriptions of the model variables.

Full Plot OESF Models

Table 6, Table 7, and Table 8 provide the coefficients of the new OESF full plot models, basal area, density, and diameter respectively.

Table 6: Coefficients for new OESF full plot basal area model

sqrt(basal area)	
intercept	-23.6140 ***
log(P01)	-0.5074
log(Esms)	7.0502 ***
Pa2	0.1162 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1	

Table 7: Coefficients for new OESF full plot density model

log(density)	
intercept	2.4259
log(Es)	1.2835
P60	-0.0317 **
Pfmn	0.0872 *
log(Es) * Pfmn	-0.0317 *
P60 * Pfmn	0.0004 *
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1	

Table 8: Coefficients for new OESF full plot diameter model

log(quadratic mean diameter)	
intercept	-9.7792 *
Eskw	-14.4901 **
sqrt(P90)	1.1770 **
log(Pamnf)	2.5623 *
Eskw * sqrt(P90)	1.1220 *
Eskw * log(Pamnf)	3.7975 **
sqrt(P90) * log(Pamnf)	-0.2374 *
Eskw * sqrt(P90) * log(Pamnf)	-0.2993 *
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1	

Zone 1 OESF Plot Models

Table 9, Table 10, and Table 11 provide the coefficients of the new OESF Zone 1 plot models, basal area, density, and diameter respectively.

Table 9: Coefficients for new OESF Zone1 plot basal area model

sqrt(basal area)	
intercept	-11.8214 *
log(Ev)	2.2248 ***
sqrt(Ek)	3.7031 **
Pa2	0.0782 **
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1	

Table 10: Coefficients for new OESF Zone 1 plot density model

log(density)	
intercept	38.0285 ***
log(Eiq)	-13.6422 **
log(P40)	-6.8997 ***
Pa2f	-0.0080
log(Eiq) * log(P40)	2.9206 ***
log(Eiq) * Pa2f	0.0369 *
log(Eiq) * log(P40) * Pa2f	-0.0078 *
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1	

Table 11: Coefficients for new OESF Zone 1 plot diameter model

sqrt(quadratic mean diameter)	
intercept	-3.4907 ***
log(Eiq)	1.0782 ***
log(P30)	0.9049 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1	

Table 12 provides a description of all of the independent variables used in the new OESF models, both the full plot models and the Zone 1 plot models.

Table 12: Descriptions of the variables used in the new OESF models

Model Variable	Fusion CloudMetrics Variable	Variable Explanation
P01	Elev.P01	1st percentile of the heights, all returns
Eiq	Elev.IQ	Interquartile distance, all returns
Ek	Elev.kurtosis	measure of whether distribution of heights are heavy-tailed or light-tailed relative to a normal distribution, all returns. Heavy-tailed distributions have their tails return to zero more slowly than a normal distribution, and light-tailed distributions have their tails return to zero more quickly.
Es	Elev.stddev	standard deviation of the heights, all returns
Eskw	Elev.skewness	skewness of the heights, all returns
Esms	Elev.SQRT.mean.SQ	quadratic mean height, all returns
Ev	Elev.variance	variance of the heights, all returns
P30	Elev.P30	30th percentile of the heights, all returns
P40	Elev.P40	40th percentile of the heights, all returns
P60	Elev.P60	60th percentile of the heights, all returns
P90	Elev.P90	90th percentile of the heights, all returns
Pa2	Percentage.all.returns.above.6.56	percent cover; number of all returns above 2m / number of all returns
Pa2f	X.All.returns.above.6.56.....Total.first.returns....100	percent cover; number of all returns above 2m / number of first returns
Pamnf	X.All.returns.above.mean.....Total.first.returns....100	percent cover; number of all returns above the mean height of all returns / number of first returns
Pfmn	Percentage.first.returns.above.mean	percent cover; number of first returns above mean height of all returns / number of first returns

Applying OESF Models to Mashel Plots

The new OESF models were used to predict Mashel plot basal area, density, and diameter using Mashel LIDAR data, without refitting or adapting the models. This provides a second test of the transferability of inventory models from one location to another.

Results

Comparison of Plot Inventory Data

For reference, Table 13 provides summarized plot inventory values for the three different plot types. These are the values measured by the field crews. Diameters are similar for the three plot types, but the OESF plots have higher basal areas and higher stand densities.

Table 13: Summarized field-measured plot statistics.

Plot Type	Basal Area (ft^2/acre)					Density (trees/acre)					Quadratic Mean Diameter (inches)				
	mean	median	std. dev.	min	max	mean	median	std. dev.	min	max	mean	median	std. dev.	min	max
Mashel	175.0	152.6	120.9	8.9	582.1	162	144	95	16	520	14.4	13.9	5.1	6.1	28.2
OESF Full	235.7	232.0	74.8	67.8	490.8	192	167	86	52	405	16.1	15.7	5.0	8.3	32.8
OESF Zone 1	241.2	238.9	92.9	87.4	466.4	189	172	100	40	479	16.5	16.4	5.0	8.6	29.0

Data Collection

Position data was collected for 55 plots in the month of April, 2019, exceeding the targeted 40 plots. Of these 55 plots, 27 plots had either rebar stakes or fence posts permanently marking the centerline, while 28 plots were not permanently monumented. 53 of the 55 plots were located in the OESF, the final two plots were located in the Olympic National Forest.

Plot Sizes and Orientations

Plot centerline lengths and azimuths derived from the GPS points were compared to the intended 60 meter length and the azimuth values listed in the field sheet provided by WADNR. Field sheet azimuths were recorded by the technicians that originally installed and measured the plots. It is possible that the GPS values are a more accurate representation of the plots than the intended 60 meter length and the azimuth values listed in the field sheet. Plots were established, according to the DNR's monitoring protocol (Minkova, Teodora; Foster, Alex 2017), using a laser rangefinder or hypsometer and a compass. With understory vegetation, trees, and terrain, getting clear views along the full plot centerlines was likely difficult in many plots.

Plots should be 60 meters long (horizontal distance). Table 14 summarizes the length differences of the GPS-derived plots from the intended 60 meter length.

Table 14: GPS-derived plot length summary.

absolute value length difference from 60 meters			
	Average	Median	Max
plots with stakes / fence posts	1.6	1.4	4.9
no markers	3.8	3.2	11.4

Plots with stake or fencepost markers had a range of length differences from 60 meters of -4.2 to 4.9 meters. Plots without permanent markers had a range of length differences from 60 meters of -11.4 to 4.2 meters. Length errors are slightly larger in plots that are not permanently marked.

Table 15 summarizes the azimuth differences between the plot centerline azimuths listed in the field sheet and those derived from the plot GPS points. In general, it appears there is good alignment between the two sets of centerline azimuths.

Table 15: GPS-derived plot azimuth summary.

absolute value azimuth difference (degrees) from field sheet			
	Average	Median	Max
Plots with stakes / fence posts	4.0	2.4	23.2
no markers	7.5	8.3	16.6

Model Accuracies

The tables below provide two accuracy statistics, the coefficient of determination (r^2), and the root mean square error (RMSE). The coefficient of determination ranges from zero to one, with zero meaning that the model cannot predict any of the variance in the dependent variable, and one meaning that the model can predict all of the variance in the dependent variable. The root mean square error will be zero or greater and is in the units of the dependent variable (e.g. the RMSE of diameter is in inches). The larger the RMSE, the larger the differences between the predicted and actual values.

The tables below provide the r^2 and RMSE values for all models developed and tested for this project. While both r^2 and RMSE are used to describe the accuracy of the model, the interpretation in this report focuses on r^2 with RMSE provided for further context. Coefficient of determination values allow comparison of models to one another, which is not possible with RMSE. RMSE requires an understanding of what valid values are for each model in order to interpret results fairly, and valid ranges may be different for the same model in different forest types.

Table 16 provides the accuracies of the models as they performed in the Mashel Pilot Study, and is provided for comparison.

Table 16: Accuracies of Mashel models in the Pilot Study

	Mashel Pilot Study Results	
	r^2	RMSE
Basal Area	0.72	63.12
Density	0.46	68.96
Diameter	0.70	2.77

Using Mashel Models to Predict OESF Plots

Table 17 provides the results of using the Mashel models to predict full OESF plot basal area, density, and diameter. Table 18 provides the results of using the Mashel models to predict Zone 1 OESF plot values.

Table 17: Accuracies of the Mashel Models, and new OESF models, predicting full OESF plots

	Existing Mashel Models		Refit Mashel Models		Adapted Mashel Models		New OESF Models	
	r^2	RMSE	r^2	RMSE	r^2	RMSE	r^2	RMSE
Basal Area	0.32	59.50	0.36	57.90	0.44	54.16	0.61	44.87
Density	0.44	57.40	0.53	52.58	0.53	52.67	0.71	41.51
Diameter	0.70	2.28	0.77	2.02	0.76	2.05	0.86	1.55

Table 18: Accuracies of the Mashel Models, and new OESF models, predicting Zone 1 OESF plots

	Existing Mashel Models		Refit Mashel Models		Adapted Mashel Models		New OESF Models	
	r^2	RMSE	r^2	RMSE	r^2	RMSE	r^2	RMSE
Basal Area	0.24	78.16	0.32	73.70	0.32	73.71	0.37	57.20
Density	0.31	77.65	0.40	72.00	0.39	72.95	0.58	50.00
Diameter	0.54	3.08	0.68	2.55	0.68	2.57	0.59	2.67

The performance of all Mashel models decreased when applied to the OESF plots, except for the existing diameter model when used on the full OESF plots, which performed equally as well. The existing density model when used on the full OESF plots had only a slight decrease in performance. Refitting and adapting the models improved performance for all models, but still did not achieve the accuracy of the Pilot Study, except for the existing diameter model and the existing density model when used on the full OESF plots.

The new OESF models outperformed the three Mashel models in all cases, except for the new Zone 1 diameter model, which was outperformed by the refit and adapted Mashel diameter models.

The basal area model, which was the best performer of the models in the Pilot Study, saw the largest decrease in accuracy. Neither of the new OESF basal area models (full plots and Zone 1 plots) perform as well as the basal area model in the Pilot Study.

In all cases, the full plot models performed better than the Zone 1 plot models.

Using OESF Models to Predict Mashel Plots

Model accuracy was also checked in the reverse direction, using the new OESF models to predict the values for the plots in the Mashel Pilot Study. Table 19 provides the results of using the OESF full plot models to predict the Mashel plots. Table 20 provides the results of using the OESF Zone 1 plot models to predict the Mashel plots.

	OESF Full, Predicting Mashel	
	r^2	RMSE
Basal Area	0.68	68.06
Density	0.17	85.66
Diameter	0.58	3.29

Table 19: OESF full plot model accuracies predicting Mashel plots

Table 20: OESF Zone 1 plot model accuracies predicting Mashel plots

	OESF Zone 1, Predicting Mashel	
	r^2	RMSE
Basal Area	0.55	79.82
Density	0.01	94.67
Diameter	0.64	3.05

The OESF full plot basal area model performs nearly as well as the native Mashel basal area model. Both full plot and Zone 1 plot diameter models have some decrease in performance compared to the native Mashel models. Both density models had very poor performance, in particular, the Zone 1 density model.

Discussion

The data suggests that directly applying inventory models from the Mashel to the OESF or from the OESF to the Mashel should only be done with great care and caution. There will be noticeable decreases in accuracy in most circumstances.

For a model to be transferable to a new forest type, it must be useable without collecting new plot inventory data, and should achieve accuracies as good as models developed in that new forest type. For the purposes of this study, that means that in the OESF, the existing Mashel models should achieve accuracies as high as the new OESF models. Comparing existing Mashel model accuracies to new OESF model accuracies (Table 17 and Table 18) shows that this did not occur.

Additionally, for the Mashel plots, we would expect to see the new OESF models (Table 19 and Table 20) achieving the same accuracies as the Mashel models did in the pilot study (Table 16). Again, this did not occur, although the new OESF full plot basal area model and the new OESF Zone 1 diameter model were very close.

Possible reasons for this are discussed below. It will be up to decision makers to determine what levels of model accuracy are acceptable. If transferred models cannot achieve the required level of accuracy, new models will need to be developed.

Linear Regression Models and Selected Variables

The linear regression models from the Pilot study and developed for the OESF plots are dependent on the specific selected LIDAR variables. Any changes in the ranges or distributions of these variables between different locations will have an impact on the ability of the models to make accurate predictions. It should be expected that the ranges and distributions of LIDAR variables will be different for plots in forests of different heights and canopy densities, with these height and density differences dependent on species, climate, and management history.

The input variables, selected by the Comprehensive Model Selection process, are not the same between the Mashel models and the new OESF models. This suggests that there may be important differences between the two sets of plots, which make the selected metrics better at describing the forest structure in each model's respective location. The ranges and distributions of the selected LIDAR variables may not be the same from location to location.

For example, the new OESF Zone 1 density model has particularly poor performance in the Mashel plots (Table 20), with and r^2 of 0.01 and an RMSE of 94.67. The LIDAR variable Elev.IQ was selected as an independent variable in this model, and looking at the behavior of this variable across its range in the Mashel plots (Figure 6) shows that values below 20, approximately, result in massive overestimation of plot density values. However, the model's accuracy was improved (i.e., r^2 = 0.21 and RMSE = 67.0) when the plots with Elev.IQ values below 20 were removed.
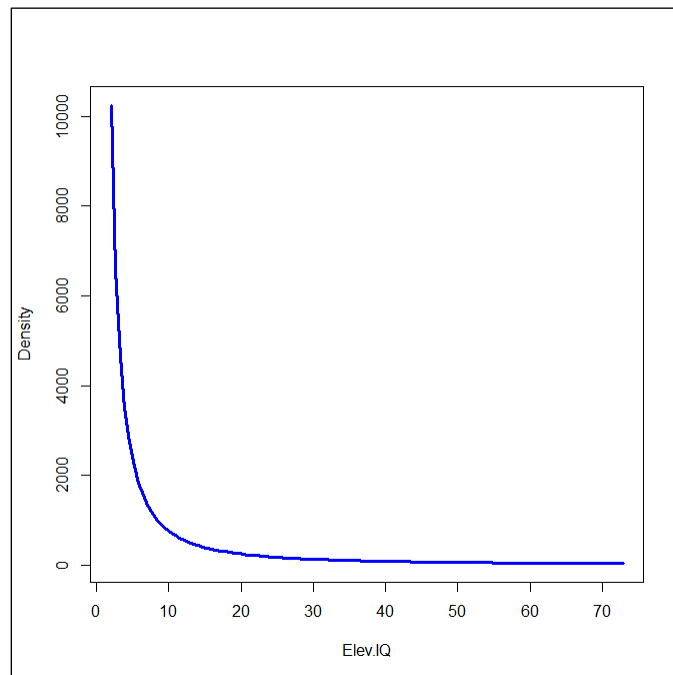

Figure 6: Predicted Mashel plot density using the OESF Zone 1 plot density model. The P40 and Pa2f variables were held constant, while varying the Elev.IQ variable across its range.

Plot Dissimilarity

Ideally, Mashel plots and OESF plots would be equally representative samples of the range of forest structures that exist in their study areas. If they are not, this disparity might decrease model performance when applied from location to location. While the majority of both study areas is managed for timber production, the different purposes of the Mashel and OESF projects resulted in different plot selection protocols, and different plot designs.

Mashel plot locations were selected to sample all forest conditions in the riparian buffers of the watershed, from young regen through old, complex stands. The Mashel watershed has a wide range of owner types with different management practices and the watershed has a complex management history and covers a wide elevation range and all stream types (fish-bearing and non-fish-bearing). Details on the selection of plot locations are provided in the Pilot Study Final Report (Moskal, et al. 2017).

While the OESF watersheds were selected to be a representative sample of the ecological diversity and management history of the OESF, they are all on Type 3 (small, fish-bearing) streams on WADNR or Olympic National Forest ownership. The OESF plots are also more similar in age and management history than the Mashel plots, with trees typically in the 40 to 50 year age range, replanted after clear cutting and disturbance (Figure 7).
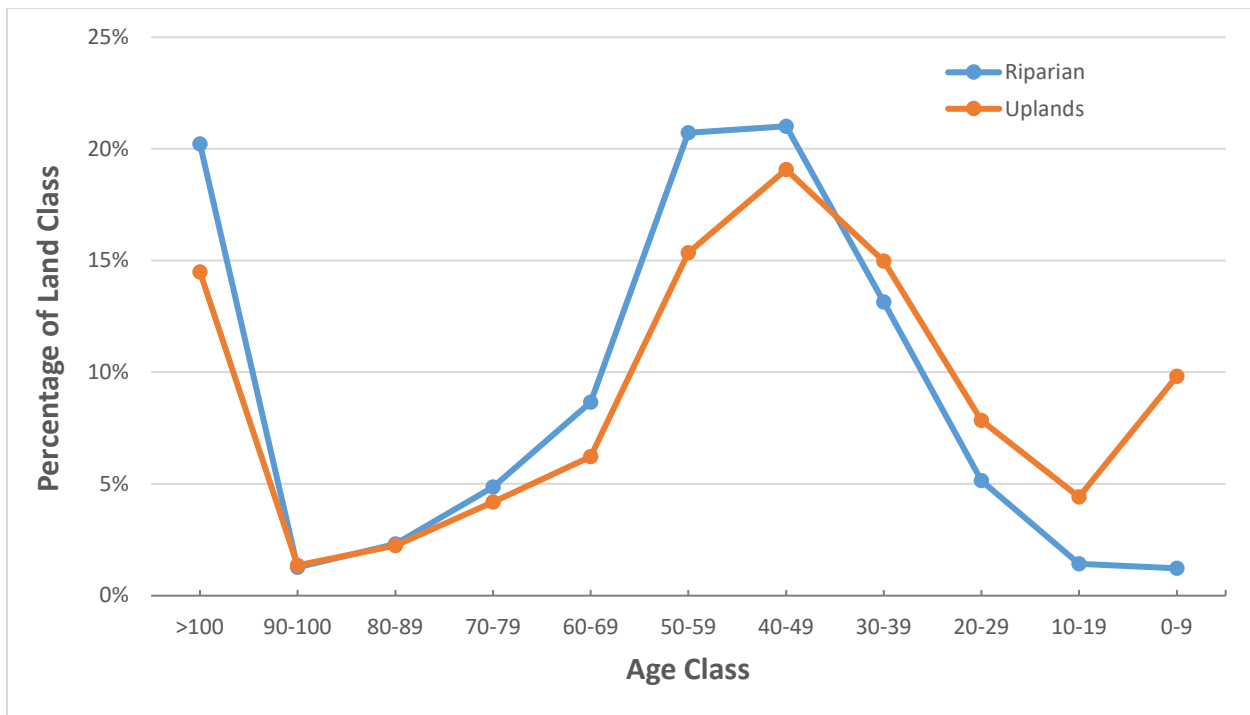

Figure 7: Age class distribution of the OESF by land class. Source: Teddy Minkova, WADNR.

Plot Size

It is interesting to note that Mashel models performed better when applied to full OESF plots than they did when applied to Zone 1 OESF plots. The full plot OESF basal area and density models also performed

better when applied to the Mashel plots than the Zone 1 OESF models. Zone 1 OESF plots are approximately the same size as the Mashel plots, while the full OESF plots are three times larger. It was hypothesized that using the Mashel models, which were developed using 1/8th acre plots, on 0.44 acre OESF plots, might reduce their accuracy, but the opposite occurred.

OESF plots were all located within a few meters of the stream. Because of this, the Zone 1 plots are likely capturing more riparian vegetation, in particular, an alder hardwood component in the first five to 10 meters from the stream, while the full OESF plots, due to their length, include more upland vegetation, with much less alder present. The Mashel plots on type N streams were within 75 feet of the streams, which is similar to the OESF Zone 1 plots. The Mashel plots on type F and S streams needed to be within 225 feet of the stream, which means they may be capturing more upland forest vegetation. It is possible that using all of the Mashel plots (all stream types) results in the Mashel plots capturing a wider range of forest conditions than is present in the OESF Zone 1 plots, and that the OESF full plots are more similar, in terms of forest conditions, even though they are much larger.

Each tree in a smaller plot has a larger impact on plot basal area, quadratic mean diameter, and density than a single tree in a larger plot does, meaning any measurement errors or inclusion/exclusion of trees in the smaller plots will matter more.

If the GPS positions are not located on the true OESF plot endpoints, due to data collection or data processing errors, there may be a misalignment between the measured plot trees and the clipped LIDAR point clouds for each plot. Additionally, if the plots are not actually 30 meters wide, the clipped LIDAR may be too wide or narrow compared to the plots. This would result in including LIDAR points that are not part of the plots and/or excluding LIDAR points that are part of the plots, which would have some effect on the LIDAR summary metrics (although the effect may be small). This may matter more on smaller-sized plots.

It was assumed in the data processing, that Zone 1 was 20 meters long along the centerline. If there is significant variability of the lengths of Zone 1, the clipped LIDAR may again be too long or too short, having some impact on the LIDAR summary metrics, possibly reducing the accuracy of the Zone 1 models.

LIDAR

Two separate LIDAR data sets were used to test the Mashel models and to build the new OESF models. If there are significant differences in the quality of the data in these two acquisitions, that may also have some impact on the results presented here. The acquisitions were flown by two different vendors, but used similar LIDAR sensors and data processing software and had equivalent data collection specifications. Further detailed study of the datasets would be required to know if there are any important differences. If future analysis is done over a large area, it is likely that multiple LIDAR acquisitions will be used.

Time Lag

In the Mashel Pilot Study, the time difference between when the LIDAR was flown and the field plots were measured was five years. For the OESF plots, there was a shorter range of time differences with some plots being measured one year before the LIDAR was flown, and the rest measured between one and three years after the LIDAR was flown. In the Mashel, five years of growth took place between the LIDAR measurements and the field measurements. This change is built into the models, with the models

predicting what forest conditions occurred five years after the LIDAR. This time relationship does not exist in the OESF plots, which likely has some impact on model accuracy.

Recommendations

Inventory models for different forest types will be necessary. How many is still unknown. Further work needs to be done to better understand the number of models that will be necessary for statewide monitoring. This will involve installing forest inventory plots in additional forest types, and also looking at the types of models being used. Eastern Washington is the area of the state that should be prioritized for new plot installation.

There are many factors that can influence the accuracy of the models, so it will be important to consider the quality of the LIDAR data, the timing between LIDAR acquisition and field data collection, the plot design, and the plot location selection, in particular, how well the plots represent the landscape of interest.

Testing Alternative Modeling Approaches

Because the linear regression models presented here are dependent on the selected LIDAR metrics and the values of those metrics, it could be valuable to examine whether or not there are certain LIDAR metrics that consistently appear in models in different forest types. If specific metrics are present in some of the best models across forest types, those metrics may be worth using even if it reduces model accuracy in some locations.

One of the difficulties of using linear regression models is identifying the appropriate number of non-correlated independent variables. LIDAR summary metrics are often subtle variations on the same types of metrics: heights, variability of the heights, shapes of the height distributions, etc. A way to solve these issues would be to build regression models using principal components rather than individual LIDAR metrics. Principal Components Analysis is a statistical method that builds uncorrelated principal components from the provided variables removing the need to select non-correlated independent variables. Many variables will contribute to each principal component making the model less easy to interpret, but in exchange, the models are simpler to build and less dependent on individual LIDAR metrics, which may increase their transferability.

There are modeling approaches other than linear regression that may be worth testing as well, such as regression trees, like Random Forest. This approach would again remove the need to identify the appropriate number of non-correlated independent variables.

Developing a Database of Plots and LIDAR Data

This project has now developed plot level inventory and LIDAR metrics for two forest types. If work continues on this project, and more plots are measured in new forest types, it would be a good idea to combine the existing and new data into a single database. This would create a platform that could be used to test many different modeling approaches and different combinations of plots from different forest types. Existing models could be tested against new plots, new models could be tested against all existing plots, and identifying areas where new plots are necessary would become more straightforward.

References

Cooke, Andrew, and L. Monika Moskal. 2018. *Scoping and Recommendations for Extensive Riparian Monitoring Implementation Pilot Project.* Olympia: Prepared for the Washington State Department of Natural Resources.

Fekety, Patrick A., Michael J. Falkowski, and Andrew T. Hudak. 2015. "Temporal transferability of LiDAR-based imputation of forest inventory attributes." *Canadian Journal of Forest Research* 45 (4): 422-435. doi:10.1139/cjfr-2014-0405.

Fekety, Patrick A., Michael J. Falkowski, Andrew T. Hudak, Theresa B. Jain, and Jeffrey S. Evans. 2018. "Transferability of Lidar-derived Basal Area and Stem Density Models within a Northern Idaho Ecoregion." *Canadian Journal of Remote Sensing* 44 (2): 131-143. doi:10.1080/07038992.2018.1461557.

Karjalainen, Tomi, Lauri Korhonen, Petteri Packalen, and Matti Maltamo. 2019. "The transferability of airborne laser scanning based tree-level models between different inventory areas." *Canadian Journal of Forest Research* 49 (3): 228-236. doi:10.1139/cjfr-2018-0128.

Mcgaughey, Robert J, Ahmed Kamal, Hans-Erik Andersen, and Stephen E Reutebuch. 2017. "Effect of Occupation Time on the Horizontal Accuracy of a Mapping-Grade GNSS Receiver under Dense Forest Canopy." *Photogrammetric Engineering & Remote Sensing* 83 (12): 861-868. doi:10.14358/PERS.83.12.861.

McGaughey, Robert J. 2019. *FUSION/LDV: Software for LIDAR Data Analysis and Visualization.* Seattle, WA: United States Department of Agriculture Forest Service Pacific Northwest Research Station. http://forsys.cfr.washington.edu/fusion.html.

Minkova, Teodora; Devine, Warren. 2016. *Status and Trends Monitoring of Riparian and Aquatic Habitat in the Olympic Experimental State Forest: Habitat Status Report and 2015 Project Progress Report.* Olympia: Washington Department of Natural Resources. https://www.dnr.wa.gov/publications/lm_oesf_status_report_final_20161122.pdf.

Minkova, Teodora; Foster, Alex. 2017. *Status and Trends Monitoring of Riparian and Aquatic Habitat in the Olympic Experimental State Forest: Monitoring Protocols.* Olympia: Washington Department of Natural Resources. https://www.dnr.wa.gov/publications/lm_oesf_2017_monitoring_protocol.pdf.

Moskal, L. Monika, Andrew G. Cooke, Travis Axe, and Jeffrey M. Comnick. 2017. *Extensive Riparian Vegetation Monitoring - Remote Sensing Pilot Study. Final Report.* Olympia: Prepared for the Washington State Department of Natural Resources.

R Core Team. 2018. "R: A Language and Environment for Statistical Computing." Vienna, Austria. https://www.R-project.org/.

Roussel, Jean-Romain. 2018. "lidR: R package for Airborne LiDAR Data Manipulation and Visualization for Forestry Applications." https://github.com/Jean-Romain/lidR.

Tompalski, Piotr, Joanne C. White, Nicholas C. Coops, and Michael A. Wulder. 2019. "Demonstrating the transferability of forest inventory attribute models derived using airborne laser scanning data." *Remote Sensing of Environment* 227: 110-124. doi:10.1016/j.rse.2019.04.006.

Watershed Sciences. 2012. "LiDAR Remote Sensing, Hoh River Watershed, Washington: Delivery 1, June 29, 2012." Corvallis, Oregon.

Woolpert. 2014. "USGS Olympic Peninsula WA QL1 Lidar." Dayton, Ohio.

Appendix A: Plot Measurement Tables

Table 21: GPS-derived plot lengths and azimuths and field sheet values.

Plot ID	GPS		Fieldsheet	
	Length (meters)	Azimuth	Length (meters)	Azimuth
542L	55.7	68.3	60	60
542R	58.2	268.8	60	260
544L	59.7	94.2	60	92
544R	57.4	244.1	60	240
545L	55.8	90.8	60	114
545R	56.7	291.3	60	276
550L	52.3	135.2	60	130
550R	56.1	313.9	60	310
566L	59.5	75.6	60	75
566R	59.8	237.7	60	240
567L	61.6	250.5	60	240
567R	59.6	95.3	60	90
568L	60.6	256.8	60	250
568R	59.8	73.6	60	70
582L	48.6*	146.6	60	130
582R	59.2	11.3	60	10
584L	61.7	159.6	60	N/A
584R	57.9	329.4	60	320
597L	55.4	175.7	60	170
597R	54.1	340.8	60	340
605L	52.3	235.1	60	230
605R	58.9	60.3	60	50
619R	59.5	120.6	60	N/A
621L	57.6	275.9	60	274
621R	58.7	90.7	60	80
625L	56.8	203.1	60	190
625R	57.9	80.5	60	70
637L	52.3	170.3	60	170
637R	53.1	20.4	60	10
642L	58.6	263.3	60	260
642R	64.5	94.1	60	100
653L	54.1	265.4	60	254
653R	52.8	90.0	60	76
658L	57.0	168.8	60	168

Plot ID	GPS		Fieldsheet	
	Length (meters)	Azimuth	Length (meters)	Azimuth
658R	58.6	320.2	60	316
688L	61.3	226.0	60	226
688R	59.7	92.4	60	90
694L	64.9	127.2	60	126
694R	63.3	325.6	60	324
724L	57.4	124.1	60	130
724R	60.3	353.6	60	356
744L	60.1	31.7	60	30
744R	58.7	202.4	60	200
760L	61.6	34.2	60	30
760R	59.7	197.7	60	200
763L	53.9	13.1	60	12
763R	64.2	214.7	60	200
773L	61.1	287.6	60	290
773R	58.6	139.8	60	140
790L	61.9	285.8	60	280
790R	61.1	106.1	60	110
804L	60.0	210.2	60	207
804R	59.2	27.8	60	24
NF8L	61.7	291.0	60	290
NF8R	60.8	112.1	60	110

* The WADNR technician collecting the GPS data noted that plot 582L was short

Table 22: Calculations of length and azimuth differences between GPS-derived plots and field sheet values.

Plot ID	Centerline Stakes (S) or Fence Posts (F)	Azimuth Difference GPS to Field Sheet	GPS Length Difference from 60 meters	Distance (meters) between GPS B Point and Field Sheet B Point
542L	none	8.3	-4.3	7.0
542R	none	8.8	-1.8	6.5
544L	S	2.2	-0.3	0.7
544R	S	4.1	-2.6	2.9
545L	F	-23.2	-4.2	26.4
545R	F	15.3	-3.3	13.1
550L	none	5.2	-7.7	8.1
550R	none	3.9	-3.9	4.0
566L	S	0.6	-0.5	2.2

Plot ID	Centerline Stakes (S) or Fence Posts (F)	Azimuth Difference GPS to Field Sheet	GPS Length Difference from 60 meters	Distance (meters) between GPS B Point and Field Sheet B Point
566R	S	-2.3	-0.2	5.2
567L	none	10.5	1.6	8.3
567R	none	5.3	-0.4	2.6
568L	S	6.8	0.6	4.2
568R	S	3.6	-0.2	0.9
582L	none	16.6	-11.4	17.2
582R	none	1.3	-0.8	1.7
584L	none	N/A	1.7	N/A
584R	none	9.4	-2.1	7.0
597L	none	5.7	-4.6	5.4
597R	none	0.8	-5.9	6.2
605L	none	5.1	-7.7	8.0
605R	none	10.3	-1.1	7.8
619R	none	N/A	-0.5	N/A
621L	none	1.9	-2.4	2.5
621R	none	10.7	-1.3	8.5
625L	none	13.1	-3.2	11.2
625R	none	10.5	-2.1	8.4
637L	none	0.3	-7.7	8.0
637R	none	10.4	-6.9	10.4
642L	F	3.3	-1.4	1.6
642R	F	-5.9	4.5	10.4
653L	none	11.4	-5.9	10.6
653R	none	14.0	-7.2	13.3
658L	S	0.8	-3.0	3.6
658R	S	4.2	-1.4	2.1
688L	S	0.0	1.3	3.2
688R	S	2.4	-0.3	0.5
694L	F	1.2	4.9	5.2
694R	F	1.6	3.3	3.4
724L	F	-5.9	-2.6	9.1
724R	F	-2.4	0.3	5.3
744L	S	1.7	0.1	1.1
744R	S	2.4	-1.3	1.4
760L	S	4.2	1.6	2.2
760R	S	-2.3	-0.3	5.3
763L	none	1.1	-6.1	6.2

Plot ID	Centerline Stakes (S) or Fence Posts (F)	Azimuth Difference GPS to Field Sheet	GPS Length Difference from 60 meters	Distance (meters) between GPS B Point and Field Sheet B Point
763R	none	14.7	4.2	13.6
773L	S	-2.4	1.1	5.3
773R	S	-0.2	-1.4	3.3
790L	F	5.8	1.9	3.8
790R	F	-3.9	1.1	7.1
804L	none	3.2	0.0	0.6
804R	none	3.8	-0.8	1.4
NF8L	S	1.0	1.7	2.5
NF8R	S	2.1	0.8	1.1