Article

Predicting Volume and Biomass Change from Multi-Temporal Lidar Sampling and Remeasured Field Inventory Data in Panther Creek Watershed, Oregon, USA

Krishna P. Poudel [1], James W. Flewelling [2] and Hailemariam Temesgen [1,*]

[1] Department of Forest Engineering, Resources, and Management, College of Forestry, Oregon State University, 280 Peavy Hall, Corvallis, OR 97331, USA; Krishna.Poudel@oregonstate.edu

[2] 9320 40th Ave. NE, Seattle, WA 98115, USA; jim.flew@yahoo.com

* Correspondence: hailemariam.temesgen@oregonstate.edu; Tel.: +1-541-737-8549

Abstract: Using lidar for large-scale forest management can improve operational and management decisions. Using multi-temporal lidar sampling and remeasured field inventory data collected from 78 plots in the Panther Creek Watershed, Oregon, USA, we evaluated the performance of different fixed and mixed models in estimating change in aboveground biomass (ΔAGB) and cubic volume including top and stump (ΔCVTS) over a five-year period. Actual values of CVTS and AGB were obtained using newly fitted volume and biomass equations or the equations used by the Pacific Northwest unit of the Forest Inventory and Analysis program. Estimates of change based on fixed and mixed-effect linear models were more accurate than change estimates based on differences in LIDAR-based estimates. This may have been due to the compounding of errors in LIDAR-based estimates over the two time periods. Models used to predict volume and biomass at a given time were, however, more precise than the models used to predict change. Models used to estimate ΔCVTS were not as accurate as the models employed to estimate ΔAGB. Final models had cross-validation root mean squared errors as low as 40.90% for ΔAGB and 54.36% for ΔCVTS.

Keywords: LiDAR; Pacific Northwest; aboveground biomass; cubic volume; change estimation

1. Introduction

Using lidar for large-scale forest management can improve operational management decisions. Wall-to-wall lidar is increasingly available for large forested areas in the western United States. Acquisition is often funded by public agencies, though private companies are also funding some acquisition campaigns. The State of Oregon has recently been acquiring lidar data at a rate of over 100,000 ha per year (Oregon Department of Geology and Mineral Industries, 2009). Current campaigns have lidar densities averaging eight pulses per m^2. The expense of acquiring the data can often be justified without considering the use of the data for intensive forest inventory. However, public agencies including the U.S. Bureau of Land Management (BLM) are using the lidar data for stand delineation and the estimation of per hectare attributes. The later are typically based on analysis techniques similar to those described by Næsset [1].

Data generated by aerial lidar has been used as an information base for mapping (hydrography, topography), civil engineering (roads, urban planning), and natural resource management. The cost of lidar acquisition has decreased in recent years, and it has become affordable for most land management agencies and forest companies [2]. The high-precision ground and vegetation information that it provides generates significant benefits in terms of savings. BLM and the USDA Forest service have already used lidar over large areas and have benefited from such an investment.

There is a common interest in exploiting the lidar data and fused imagery to obtain detailed forest inventories. Of particular interest are inventories with better information on species distribution and the mapping of dominant and co-dominant trees. Additionally, inferences related to habitat, fire risk, down woody debris, and numerous other landscape features are of interest. To examine and evaluate suitable methods, the BLM, together with other governmental agencies and private parties, have a cooperative research effort focusing on the Panther Creek watershed. The watershed has been the target of multiple efforts to collect remotely sensed data, particularly airborne lidar data. In support of the inventory effort, stem-mapped forest plots have been repeatedly measured. The U.S. Environmental Protection Agency (EPA) is conducting an intensive soil survey of the area. Other data gathering efforts are ongoing, including the collection of terrestrial lidar data and meteorological data [3].

The use of remote sensing technologies in describing forest attributes over time has been attractive among forest managers and researchers due to its advantage of covering large spatial domains compared to the traditional field-based approach. However, the accuracy of such estimates may not be the same as the ones obtained from the traditional field-based measurements. Therefore, the primary focus of research on the application of lidar technology in forestry has been on the identification of one or more lidar derived variables that explain the highest variation in the forest attribute of interest such as dominant height, basal area, tree density, volume, and biomass [4].

A variety of model forms have been used to predict forest attributes using lidar derived variables. Tonolli et al. [5] used multiple linear regression to estimate tree volume using forest inventory and lidar data. Goerndt et al. [6] used small area estimation methods to predict selected stand attributes. Chen [7] used power models to estimate aboveground tree woody biomass using airborne lidar data. Temesgen et al. [4] used exponential models to estimate biomass increment in south-central Alaska. Each of these model forms has both advantages and disadvantages. For example, the exponential models ensure that the predicted biomass and volume are non-negative. However, these models are not flexible in estimating change in volume and biomass because the change is not always greater than or equal to zero, especially if the change is estimated over a longer time period. The purpose of using remotely sensed data is to obtain information at a large scale over time. Thus, one would expect an increase in aboveground biomass (AGB) and cubic volume including top and stump (CVTS) in some areas, but loss in others due to factors such as disturbance regimes and harvest.

In addition to different model forms, different fitting techniques such as fixed effects and mixed effects models have been used. The relationship of volume and biomass with lidar metrics differs by stands [8]. Therefore, the mixed effects models are advantageous over the fixed effects models because they address the hierarchical nature of the data by incorporating plot or stand level variation in the model; however, their application is limited in the absence of a subsample of the volume and biomass estimates for the plots or stands not included in the modeling data [4].

Modeling change in volume and biomass is important to understand forest productivity, assess the impact of fire and other disturbance regimes, and forecast environmental and economic potential [9]. The lidar-based estimation improves our abilities in monitoring, reporting, and verifying the status and change of important forest attributes such as volume and biomass over time. The traditional field-based estimates become outdated very quickly because of the dynamic nature of the forest environment and thus the remotely sensed data can supplement and or substitute the ground-based measurements [10].

Recently, remote sensing techniques have been used to predict change in AGB over a given time interval. Hudak et al. [11] used the Random Forest machine learning algorithm to quantify AGB change and carbon pools and fluxes from mean canopy heights derived from repeated lidar surveys. In comparing different approaches to model change in the biomass of a Norwegian mountain forest area, Bollandsås et al. [12] found the methods that modeled change in AGB directly from the change in different lidar variables to perform better than the methods that estimated change in AGB as the difference between predicted biomass at two measurement occasions. These approaches were also used

by Næsset et al. [13] to estimate change in forest biomass over an 11-year period. Temesgen et al. [4] used similar approaches to model change in AGB in south-central Alaska and obtained results that were consistent with these findings.

Naesset and Gobakken [14] estimated volume growth as the difference between predictions of volume in two occasions obtained from successive lidar metrics. Yu et al. [15] found the differences in digital surface models to be the best predictor of volume growth. Nakajima et al. [16] used the lidar metrics to predict the crown surface area, which was later successfully used to model the volume growth of sugi (*Cryptomeria japonica*) in Japan. Recently, Nakajima [17] used crown metrics derived from lidar data to estimate growth in diameter at breast height, diameter 4 m above the ground, total tree height, and volume. In their study, the cross validation standard deviation of the error in volume growth was 43.3% of the ground truth mean volume. All of these studies suggested that there is potential for the use of lidar metrics in modeling volume and biomass and their change over time.

The objectives of this study were to: (1) identify and examine selected methods for estimating volume and biomass increment using repeatedly measured lidar and ground data; (2) examine the performance of alternative techniques for estimating volume and biomass increment in the Panther Creek watershed; and (3) examine the potential use of mixed effect models to predict volume/biomass increment. This study is part of a larger project aimed at examining the potential use of lidar to inventory and monitor the Pacific Northwest forests.

2. Materials and Methods

2.1. Study Area

The study was conducted at the Panther Creek watershed, Oregon (Figure 1). The Panther Creek study area is an approximately 2580-hectare forested watershed on the east side of the coastal mountain range of Oregon, USA. It is located 57 km southeast of Portland, Oregon, USA at 45°18′ N, 123°21′ W [3]. The elevation ranges from 100 m to 700 m. Annual precipitation is about 150 cm. The forests are mainly planted or natural stands of Douglas fir, with significant amounts of western hemlock, western red cedar, grand fir, red alder, bigleaf maple, and several other species. Tree heights are up to 60 m [3]. Management intensity throughout the watershed has been variable, with varying planting densities, and both thinned and unthinned regimes. The ecoregion classification is "Cascade mixed forest" [18].

2.2. Data

2.2.1. Ground Data

Ground data for this study came from 78 field plots of an approximately 16.0 m radius (about 0.08 ha). Individual tree measurements such as species, diameter at breast height (DBH, 1.3 m above ground), total tree height, and height-to-live-crown were recorded for all trees with 0.5 cm and larger DBH.

One of the critical steps in using remotely sensed data for estimating forest attributes through regression is obtaining the observed values of the dependent variables—volume and biomass in our case. These estimates are usually obtained through volume and biomass equations commonly used in the region. Poudel and Temesgen [19] recently developed methods for estimating aboveground biomass equations for five Pacific Northwest tree species, namely Douglas-fir, grand fir, western hemlock, red alder, and lodgepole pine. Data were also available to fit cubic volume equations for those tree species. Poudel and Temesgen [19] also found that the locally fitted equations were more accurate than the regional and national scale equations. Therefore, when the data were available, we used the locally fitted volume and biomass equations to obtain the individual tree observed volume and biomass. Local volume equations were functions of DBH and height, while the biomass equations were simple DBH-based logarithmic equations. The volume and biomass equations used by the Pacific

Northwest unit of the Forest Inventory and Analysis (FIA PNW) program of the USDA Forest Service were used to obtain the observed volume and biomass equations for the species for which the local data were unavailable. Per plot volume and biomass was obtained by summing individual tree volume and biomass and per hectare estimates were obtained by simple scaling.
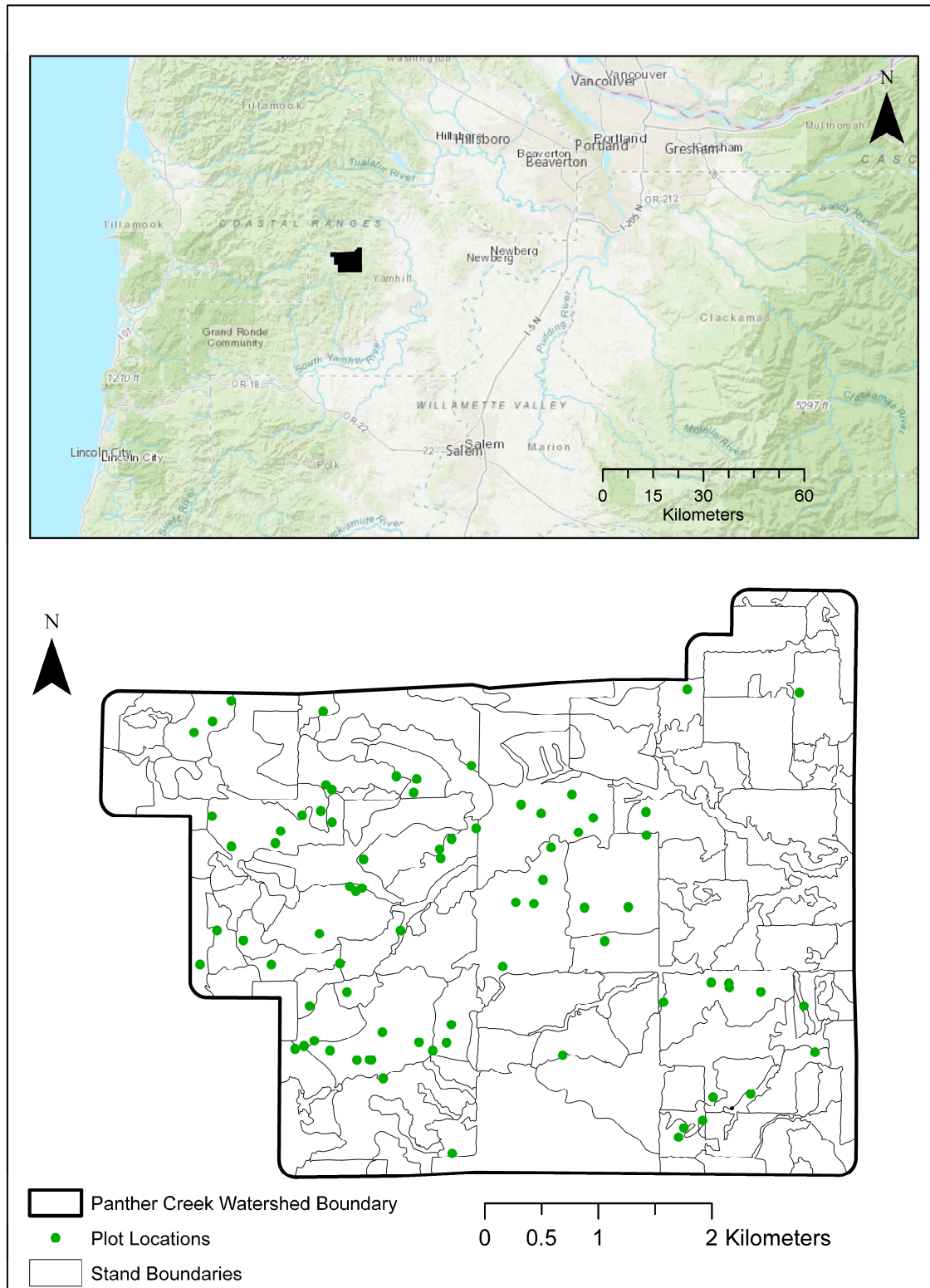

Figure 1. Map of the study area—the panther creek watershed.

2.2.2. Lidar Data

We used two leaf-off airborne laser scanning survey datasets acquired over the Panther Creek Watershed in September 2007 and September 2012 for this study. Both datasets were collected using small-footprint, discrete-return lidar systems that recorded a target density of 4 pulses/m^2 returns. The data were processed to create plot-level lidar-derived metrics using Fusion 2.0 [20]. Over 20 lidar-derived variables were evaluated as potential predictors. We selected a number of lidar variables to examine their relationships with volume and biomass and their increment. The selected predictor variables include lidar canopy height (LCH) estimated as the difference of first and last return, variance of lidar canopy height (LCHV), percent canopy cover (PCC), differences between successive percent canopy cover values (ΔPCC), and differences between successive values in the 60th percentile of all-returns lidar canopy heights.

LCH4007 and LCH4012: 40th percentiles of all-returns lidar canopy height above 1 m in 2007 and 2012, respectively.
LCH6007 and LCH6012: 60th percentiles of all-returns lidar canopy height above 1 m in 2007 and 2012, respectively.
LCHV07 and LCHV12: variance of all-returns lidar data in 2007 and 2012.
PCC07 and PCC12: percent of all-returns lidar heights, within each circular plot, above 1 m for the 2007 and 2012 data.
ΔLCH40, ΔLCH60 and ΔPCC: differences in the 40th percentiles, 60th percentiles, and percent of all-returns lidar heights above 1 m between 2007 and 2012, respectively.

Both ground and lidar data were obtained for years 2007 and 2012. Summary statistics for the selected variables are given in Table 1.

Table 1. Minimum (Min), average (Mean), maximum (Max), and standard deviation (SD) of the selected variables obtained from ground measurement or lidar. A total of 3479 and 3268 trees were observed in the ground measurements in the year 2007 and 2012, respectively.

Variable	2007				2012			
	Min	Mean	Max	SD	Min	Mean	Max	SD
DBH	0.50	26.18	162.00	18.75	0.90	28.36	165.40	18.92
HT	1.40	21.28	63.40	12.05	0.20	22.91	63.10	12.08
VPH	3.51	575.49	1744.22	397.23	21.00	622.60	1836.80	401.48
AGBPH	8.97	302.06	828.81	175.17	37.97	321.02	851.86	174.50
LCH40	1.03	16.79	37.16	9.37	1.52	19.76	39.56	8.86
LCH60	1.93	22.22	41.80	10.22	2.31	24.78	43.10	9.54
LCHV	0.72	111.30	380.33	86.68	5.18	114.24	381.78	83.22
PCC	1.09	80.02	96.00	20.37	29.94	85.36	96.42	12.59
Growth								
VPH					−174.91	47.12	136.93	41.14
AGBPH					−79.04	18.96	43.51	20.12
LCH40					−10.91	2.97	20.74	3.53
LCH60					−4.46	2.56	6.03	1.55
PCC					−21.30	5.34	50.81	11.17

DBH = diameter at breast height (1.3 m), (cm); HT = total tree height (m); VPH = volume per hectare, (m^3 ha^{-1}); AGBPH = aboveground biomass per hectare, (Mg ha^{-1}); LCH40 = 40th percentile of all-returns lidar canopy height; LCH60 = 60th percentile of all-returns lidar canopy height; LCHV = variance of all-returns lidar data; PCC = percent canopy cover. DBH, HT, VPH, and AGBPH were based on ground measurement and other variables were based on lidar.

2.3. Statistical Analysis

We adopted the modeling approach outlined by Temesgen et al. [4]. However, the model forms and the explanatory variables were different than those used in Temesgen et al. [4]. These approaches are briefly discussed below. Similar to Temesgen et al. [4], the first three approaches are based on fixed effect models and the remaining three approaches are based on mixed effects models.

Approach 1 (A1): Estimate change in volume and biomass using differences in predicted volume and biomass in 2007 and 2012:

$$Y_{07_i} = \beta_0 + \beta_1 LCHV07 + \beta_2 PCC07 + \epsilon_i \tag{1}$$

$$Y_{12_i} = \beta_0 + \beta_1 LCHV12 + \beta_2 PCC12 + \epsilon_i \tag{2}$$

$$\Delta \hat{Y}_i = \hat{Y}_{12_i} - \hat{Y}_{07_i} \tag{3}$$

where, Y_{07_i} and Y_{12_i} are the volume ($m^3 \, h^{-1}$) or aboveground biomass ($Mg \, h^{-1}$) based on the ith plot in 2007 and 2012, respectively; $\hat{Y}_{07_i}$ and $\hat{Y}_{12_i}$ are their predicted values, respectively; β_0, β_1, and β_2 are regression coefficients to be estimated from the data; $\Delta \hat{Y}_i$ is the change in volume or biomass; ϵ_i is the model error; and all other variables are as defined previously.

Approach 2 (A2): Model change in volume and biomass using the difference in difference in lidar metrics:

$$\Delta \hat{Y}_i = \beta_0 + \beta_1 \Delta LCH_P + \beta_2 \Delta PCC + \epsilon_i \tag{4}$$

where, ΔLCH_P is $\Delta LCH40$ for the ΔAGB model and $\Delta LCH60$ for the $\Delta CVTS$ model and all other variables are as defined previously.

Approach 3 (A3): Calibrate the predicted change in volume and biomass in A1 with a simple linear regression. The calibration coefficients are obtained by fitting a simple linear regression in the form of Equation (5).

$$\Delta \hat{Y}_i = \beta_0 + \beta_1 (\hat{Y}_{12_i} - \hat{Y}_{07_i}) + \epsilon_i \tag{5}$$

Approaches 4–6 are based on mixed effects models. We grouped the 78 plots in the study into five plot types based on basal area per hectare in the given plot—less than 20 $m^2 \, ha^{-1}$ (Type I), 20–40 $m^2 \, ha^{-1}$ (Type II), 40–60 $m^2 \, ha^{-1}$ (Type III), 60–80 $m^2 \, ha^{-1}$ (Type IV), and greater than 80 $m^2 \, ha^{-1}$ (Type V). For this study, we had observed basal area per hectare information for all the plots. However, in practice, that information may not be available and it might be necessary to use the lidar predicted basal area per hectare. Therefore, we fitted a simple linear regression that related basal area per hectare in 2007 to $LCH4007$ and $LCHV07$ ($BAPH_{07} = 10.74954 + 1.28667 \times LCH40_{07} + 0.11982 \times LCHV_{07}$, Adjusted-$R^2 = 0.60$). We then used these plot types as the random effects in the mixed effects models to predict change in volume and biomass per hectare. Thus, we are assuming that the change in volume and aboveground biomass is similar in stands that had a similar initial basal area per hectare compared to the stands with different initial basal areas per hectare.

Approach 4 (A4): Estimate change in volume and biomass using differences in volume and biomass in 2007 and 2012 obtained using the mixed effects model. In other words, A4 is the same as A1, except that Y_i's are predicted using the mixed effects models.

$$Y_{07_{im}} = \beta_0 + bi + \beta_1 LCHV07 + \beta_2 PCC07 + \epsilon_i \tag{6}$$

$$Y_{12_{im}} = \beta_0 + bi + \beta_1 LCHV12 + \beta_2 PCC12 + \epsilon_i \tag{7}$$

$$\Delta \hat{Y}_i = \hat{Y}_{12_{im}} - \hat{Y}_{07_{im}} \tag{8}$$

where, bi is the random plot type effect and $bi \sim N(0, \sigma_b^2)$ and is independent of $\epsilon_i \sim N(0, \sigma_e^2)$. Variance of bi represents the variability in volume and biomass due to plot type or estimated initial basal area classification. All other variables are the same as defined previously.

Approach 5 (A5): Model change in volume and biomass using the difference in lidar metrics and random plot type effect.

$$\Delta \hat{Y}_i = \beta_0 + bi + \beta_1 \Delta LCH_P + \beta_2 \Delta PCC + \epsilon_i \tag{9}$$

This approach is the same as approach 2 but a random intercept term that varies by plot type is added in the model. Once again, ΔLCH_P is $\Delta LCH40$ for the ΔAGB model and $\Delta LCH60$ for the $\Delta CVTS$ model.

Approach 6 (A6): Calibrate the predicted change in volume and biomass in A4 with the coefficients of simple linear regression:

$$\Delta \hat{Y}_i = \beta_0 + \beta_1 \left(\hat{Y}_{12_{im}} - \hat{Y}_{07_{im}} \right) + \epsilon_i \tag{10}$$

where all variables are the same as defined previously.

All statistical analyses were performed using statistical software R [21]. The mixed effects models were performed with the lme function in R-package nlme [22]. The mixed effects models incorporate the variability between stand types when modeling AGB and CVTS or their change.

These approaches were evaluated based on bias (mean difference in observed and predicted values), bias percent, root mean squared errors (RMSE), and RMSE percent that they produced in estimating the change in volume and biomass per hectare.

$$\text{Bias} = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)}{n} \tag{11}$$

$$\text{Bias \%} = \frac{\text{Bias}}{\overline{Y}} \tag{12}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} \tag{13}$$

$$\text{RMSE \%} = \frac{\text{RMSE}}{\overline{Y}} \tag{14}$$

where, n is the number of plots; y_i and $\hat{y}_i$ are the observed and predicted values of volume or biomass per hectare, respectively; and $\overline{Y}$ is the mean volume or biomass per hectare based on n plots. Leave one (plot) out cross validation was carried out to validate the performance of each approach.

3. Results and Discussion

Except for five plots for AGB and three plots for CVTS, all other plots exhibited positive change, i.e., there was an increase in both volume and biomass in 73 out of 78 plots between the year 2007 and 2012 (Figure 2). Interestingly, there were two plots in which the CVTS per hectare was increased but AGB was decreased from year 2007 to 2012. This could be because the AGB equations were a function of DBH only and the CVTS was a function of DBH and height. These five plots were tested for possible outliers using the median absolute deviation (MAD) statistics calculated as follows (Leys et al. [23]):

$$\text{MAD} = bM_i \left(\left| x_i - M_j(x_j) \right| \right) \tag{15}$$

where x_j is the set of n original observations, M is the median function, and b = 1.4826 is a constant associated with the normality assumption of the data without regarding the abnormality due to outliers. Once MAD is calculated, any observation outside Median $\pm 3 \times$ MAD was considered an outlier and removed from the final dataset used for modeling change in AGB and CVTS. This resulted in 74 plots available for modeling AGB and 76 plots for CVTS change.
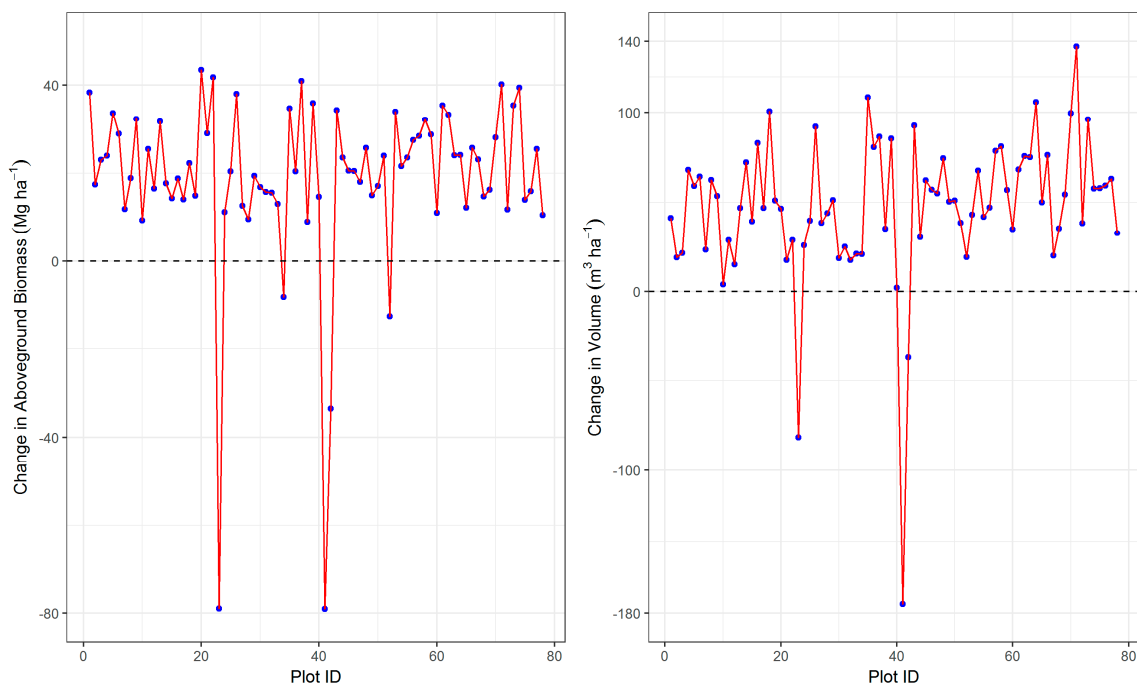
Figure 2. Change in biomass (Mg ha^{-1}) and volume (m^3 ha^{-1}) in different plots from year 2007 to 2012 based on repeated ground measurements.

Parameter estimates and their standard errors for the models used to predict AGB and its change from 2007 to 2012 are presented in Table 2. The evaluation statistics—bias, bias percent, RMSE, and RMSE percent obtained from the leave-one-out cross validation [24,25], are given in Table 3. When the fixed effect modeling approach is used, Approach 2 (A2) in which the change in AGB is directly modeled using the difference in lidar metrics obtained in 2007 and 2012 produced the smallest absolute bias. It also produced the smallest root mean squared error (9.30 Mg ha^{-1}, 40.90%) compared to other approaches based on fixed effects models. In approach 1 (A1), change in AGB is obtained as the difference in predicted biomass in 2007 and 2012 based on lidar variables obtained in respective years. This approach produced a comparable absolute bias percent to Approach 2, but it had the highest root mean squared error (31.30 Mg ha^{-1}, 137.64%). Approach 3, in which the estimated change in AGB in Approach 1 was calibrated using the coefficients of a simple linear regression, performed similarly to Approach 2 (42.90% vs. 40.90% RMSE).

Table 2. Parameter estimates and their standard errors of the models used to predict aboveground biomass and its change using lidar variables. Model 1.0 and Model 1.1 predicted aboveground biomass per hectare in 2007 and 2012, respectively, using the fixed effects model. Model 4.0 and 4.1 did the same based on the mixed effects model. All other models predicted the change in aboveground biomass per hectare, directly. Observed values of changes were obtained based on measurements of 78 field plots.

Approach	Model	Parameter (Standard Error)			R^2
		β_0	β_1	β_2	
A1	1.0	−112.7205 (49.0673)	1.4335 (0.1602)	3.2519 (0.5995)	0.64
A1	1.1	−206.7080 (84.4578)	1.5970 (0.1660)	4.1412 (0.9876)	0.64
A2	2	22.6334 (1.5774)	−0.6031 (0.3416)	0.3535 (0.0990)	0.17
A3	3	20.4517 (1.3812)	0.1008 (0.0366)	-	0.10
A4	4.0	−71.6786 (68.8690)	1.2609 (0.1942)	3.1908 (0.7429)	0.69
A4	4.1	−145.9440 (93.9257)	1.4984 (0.1838)	3.5907 (1.0721)	0.66
A5	5	23.0565 (1.8158)	−0.6644 (0.3471)	0.3434 (0.1044)	0.19
A6	6	21.6315 (1.8878)	0.0769 (0.0444)	-	0.10

Table 3. Bias, bias percent, RMSE, and RMSE percent obtained from leave-one-out cross validation in estimating change in aboveground biomass per hectare using different approaches.

Approach	Bias (Mg ha^{-1})	Bias Percent	RMSE (Mg ha^{-1})	RMSE Percent
A1	−0.21	−0.92	31.30	137.64
A2	−0.04	−0.18	9.30	40.90
A3	0.07	0.31	9.61	42.26
A4	−0.26	−1.13	28.50	125.30
A5	−0.02	−0.07	9.42	41.42
A6	0.06	0.25	9.88	43.45

Approaches 4–6 were based on modeling AGB and its change using mixed effects models. These approaches allowed us to evaluate the effects of initial basal area per hectare in change in AGB. Approach 4 produced an absolute bias percent that was similar to its fixed effects counterpart, Approach 1 (−0.21% vs. −0.26%). However, it reduced the root mean squared error by 12.34% compared to Approach 1. Approach 5, however, had a slightly higher RMSE than Approach 2 (41.42% vs. 40.90%), its fixed effect counterpart implying that if the change in AGB is estimated directly using the difference in lidar metrics, there is no difference in biomass change due to initial basal area per hectare values. In fact, the standard deviation of the random effect parameter was 0.00038 $\left(\sigma_b^2 = 1.42 \times 10^{-7}\right)$ and the fixed effect parameter estimates were also very similar. Approach 6 is the calibration of Approach 4 and it reduced the RMSE from 125.30% to 43.45% (Table 3). Instead of using a simple ratio obtained from OLS through origin, we used both slopes and intercept of the OLS because Poudel and Temesgen [26] found that calibration using both the slope and intercept of an OLS model further improved the calibration of regional volume and biomass equations compared to the OLS through origin. Additionally, having both coefficients in the calibration equation produces unbiased estimates.

We found that biomass at one point in time can be better predicted using lidar derived variables than the prediction of growth or change. For example, using the fixed effects models, the RMSE for the observed data was 34.71% and 32.83% for 2007 and 2012, respectively, but the RMSE in estimating change was 39.50%. Additionally, when the mixed effects models are used, the RMSE values for 2007 and 2012 were, respectively, 32.47% and 31.64%, but the RMSE in estimating change was 39.01%.

We performed the same analysis for change in cubic volume per hectare. Parameter estimates and their standard errors for the models used to predict CVTS and its change from 2007 to 2012 are presented in Table 4. The evaluation statistics—bias, bias percent, RMSE, and RMSE percent obtained from the leave-one-out cross validation, are given in Table 5. Average bias ranged from −0.60 m^3 ha^{-1} to 0.06 m^3 ha^{-1} (Table 5). Similar to estimating change in AGB, Approach 2 produced the smallest absolute bias percent among the first three approaches based on fixed effects models. It also produced the smallest root mean square error (28.12 m^3 ha^{-1} or 54.36%). Approach 3, which calibrates Approach 1 with OLS coefficients, reduced RMSE by 82.35%, but this RMSE was still higher than those obtained from Approach 2. Therefore, Approach 2 was the best fixed effects model-based approach for estimating change in CVTS using lidar derived variables.

Change in CVTS was also modeled using mixed effects models. Approaches 4, 5, and 6 reduced absolute percent bias by 1.05%, 0.13%, and 0.23% compared to their respective fixed effects counterparts—Approaches 1, 2, and 3. Additionally, approaches 4 and 6 reduced RMSE by 39.82% and 1.54%, but Approach 5 produced an RMSE value that was 0.61% higher than its fixed effect counterpart—Approach 2. Among the mixed effects model-based approaches, Approach 5 was the best in terms of RMSE, but Approach 4 had a slightly smaller absolute bias (Table 5). Our results are consistent with the findings of Bollandsas et al. [12], Naesset et al. [13], and Temesgen et al. [4].

Table 4. Parameter estimates and their standard errors of the models used to predict volume per hectare and its change using lidar variables. Model 1.0 and Model 1.1 predicted volume per hectare in 2007 and 2012, respectively, using the fixed effects model. Model 4.0 and 4.1 did the same based on the mixed effects model. All other models predicted the change in volume per hectare, directly.

Approach	Model	Parameter (Standard Error)			R^2
		β_0	β_1	β_2	
A1	1.0	−269.1384 (102.4261)	3.5070 (0.3069)	5.7652 (1.2567)	0.710
A1	1.1	−524.0004 (182.6992)	3.6709 (0.3307)	8.6432 (2.1417)	0.684
A2	2	33.1701 (6.8663)	8.3630 (2.4188)	−0.6203 (0.2999)	0.153
A3	3	50.3331 (4.3807)	0.0271 (0.0549)	-	0.003
A4	4.0	−131.4910 (149.1598)	3.0154 (0.4063)	5.1304 (1.5635)	0.756
A4	4.1	−148.2289 (233.4982)	2.9270 (0.4686)	5.4781 (2.4128)	0.735
A5	5	27.7979 (8.3186)	8.1724 (2.4947)	−0.3344 (0.3395)	0.221
A6	6	39.5584 (7.5935)	0.1225 (0.0854)	-	0.151

Table 5. Bias, bias percent, RMSE, and RMSE percent obtained from leave-one-out cross validation in estimating change in volume per hectare using different approaches.

| Approach	Bias (m^3 ha^{-1})	Bias Percent	RMSE (m^3 ha^{-1})	RMSE Percent
A1	−0.60	−1.16	72.16	139.47
A2	−0.13	−0.26	28.12	54.36
A3	−0.33	−0.63	29.55	57.12
A4	0.06	0.11	51.56	99.65
A5	−0.07	−0.13	28.44	54.97
A6	−0.21	−0.40	28.75	55.58

4. Summary and Conclusions

Estimating the status of forest resources and their change over time is critical for sustainable forest management. Volume and biomass estimation requires destructively sampled data [27]. Thus, the application of sampling methods and regression modeling is imperative. Use of remote sensing technology to monitor forest resources over time has gained substantial momentum in recent years due to its ability to cover a larger spatial scale than the traditional ground-based measurements. However, the accuracy of such methods is still questionable in comparison to the ground-based measurements, making the practitioner hesitant to use models developed based on remotely sensed data.

One of the advantages of using lidar metrics in estimating forest attributes is that they can provide timely estimates and help monitor change over a given time period, possibly under changing climate. However, most of the modeling efforts have focused on developing models for estimating inventory at a point in time. In this study, we compared different regression techniques to estimate the change in volume and biomass over a five-year period using multi-temporal lidar and ground based remeasurement data from a 2580 hectare watershed in the Western US. Among the approaches compared, we found that modeling change in volume and biomass (ΔCVTS and ΔAGB) directly as the function of the change in lidar metrics (ΔLidarMetrics) was superior to the approach in which AGB and CVTS in time 1 and time 2 were modeled as the functions of lidar metrics observed in respective time and the changes in CVTS and AGB were estimated as the difference in those predicted values at two points in time. Therefore, we suggest that if the objective is to estimate change over a given period of time, the objective function of the regression models should be the error in change itself. Final models suggested in this study had cross-validation root mean squared errors of 40.90% and 41.42% for ΔAGB and 54.36% and 54.97% for ΔCVTS. The residual analysis of the selected models did not show a severe problem with the assumption of homogenous variance (Figures 3 and 4). Models to predict volume and biomass based on lidar metrics at a given time were, however, more precise than

the models to predict change, as expected. Additionally, the accuracy of models for estimating change in CVTS was not as good as the models used to estimate change in AGB (Table 3 vs. Table 5).
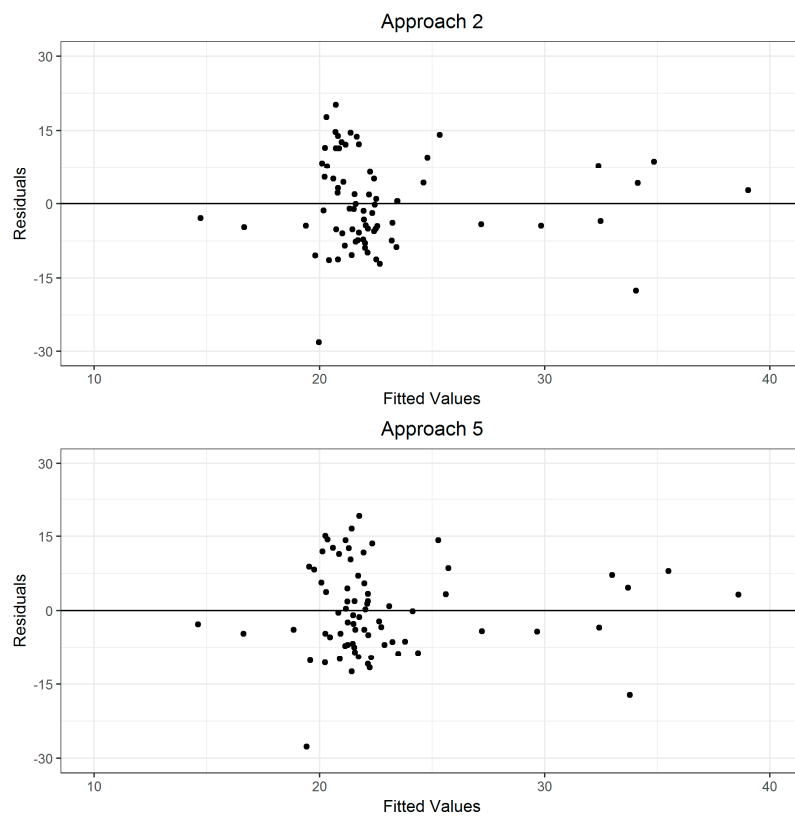

Figure 3. Plots of fitted values vs. residuals obtained in modeling change in biomass (Mg ha^{-1}) using change in lidar metrics using fixed (Approach 2) and mixed (Approach 5) effects models.
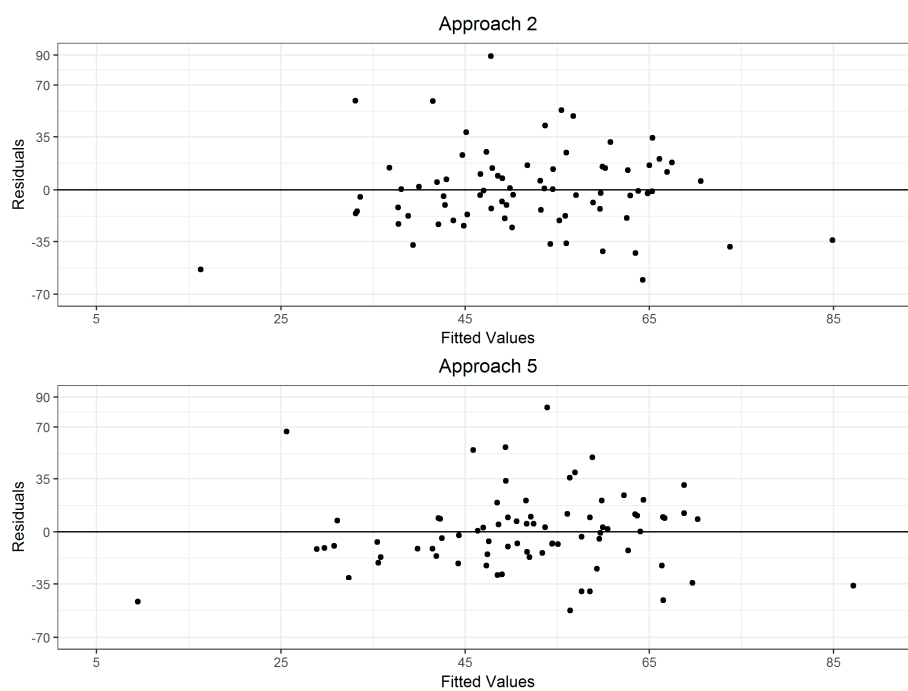
Figure 4. Plots of fitted values vs. residuals obtained in modeling change in volume (m^3 ha^{-1}) using change in lidar metrics using fixed (Approach 2) and mixed (Approach 5) effects models.

Mixed effects models have been used in the past to model ΔCVTS and ΔAGB to account for the variability due to the eco-region (e.g., [4]). Generally, plots or stands are used as the random component of the mixed models. Such models, however, have a limitation in their application because they require samples of ground measurements in order to make use of the random effect parameters. We used lidar predicted initial basal area per hectare in each plot to group them into to different "stand types". We did not find any improvement by using the mixed effects models when the change is modeled directly by the change in lidar metrics. In fact, the cross validation RMSEs were 0.52% and 0.61% higher when mixed models instead of fixed effects models were used to model ΔAGB and ΔCVTS using ΔLidarMetrics (Approach 2 and 5, Table 3). The prediction intervals obtained from leave-one-out cross validation are shown in Figure 5.
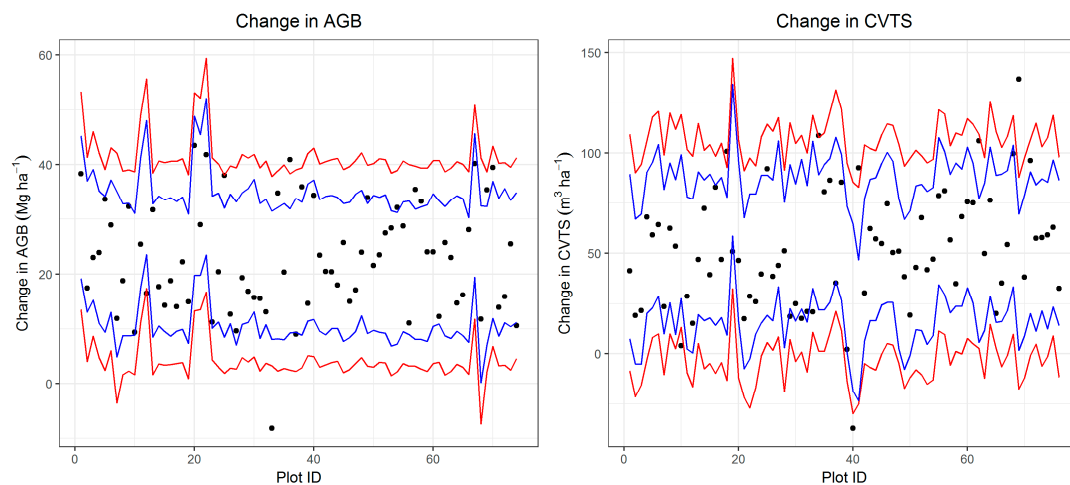
Figure 5. Prediction intervals obtained from Approach 2 (red) and Approach 5 (blue) in estimating change in aboveground biomass and cubic volume. Black points are the observed values of change in AGB and CVTS from 2007 to 2012.

The successful use of remote sensing methods such as lidar is predicated on the accuracy of the equations used to obtain "observed" values of volume and biomass for the modeling datasets. This further establishes the need for developing more accurate ground-based allometric equations covering a larger spatial scale. In this study, we used the volume and biomass equations developed using data from within the region. We believe that the test of accuracy of lidar in predicting change should be based on the actual volume and biomass obtained from the same stand. Research on error propagation by using previously published equations versus the equations developed using samples collected from the same stand or forest for which the lidar metrics are derived would validate this postulation. Additionally, given the smaller sample size in this study, we did not test more complex model forms that may be necessary to better predict the change in volume and biomass using lidar metrics.

Author Contributions: Krishna P. Poudel wrote the manuscript and R codes, and conducted the analyses. James W. Flewelling conceived and designed the study and data collection procedures, provided significant input throughout the manuscript preparation, and supervised the study. Hailemariam Temesgen formulated the idea and contributed input into the manuscript through many reviews. All authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Næsset, E. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* **2002**, *80*, 88–99. [CrossRef]
2. White, J.; Wulder, M.; Varhola, A.; Vastaranta, M.; Coops, N.; Cook, B.; Pitt, D.; Woods, M. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. In *Information Report FI-X-010*; Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre: Victoria, BC, Canada, 2013; p. 50.
3. Flewelling, J.W.; McFadden, G. LiDAR data and cooperative research at Panther Creek, Oregon. In Proceedings of the SilviLaser, Hobart, Austria, 16–20 October 2011.
4. Temesgen, H.; Strunk, J.; Anderson, H.-E.; Flewelling, J. Evaluating different models to predict biomass increment from multi-temporal lidar sampling and remeasured field inventory data in South-central Alaska. *Math. Comput. For. Nat. Res. Sci.* **2015**, *7*, 66–80.
5. Tonolli, S.; Dalponte, M.; Vescovo, L.; Rodeghiero, M.; Bruzzone, L.; Damiano, G. Mapping and modeling forest tree volume using forest inventory and airborne laser scanning. *Eur. J. For. Res.* **2010**, *130*, 1764–1772. [CrossRef]
6. Goerndt, M.E.; Monleon, V.J.; Temesgen, H. A comparison of small-area estimation techniques to estimate selected stand attributes using LiDAR-derived auxiliary variables. *Can. J. For. Res.* **2011**, *41*, 1189–1201. [CrossRef]
7. Chen, Q. Modeling aboveground tree woody biomass using national-scale allometric methods and airborne lidar. *ISPRS J. Photogramm. Remote Sens.* **2015**, *106*, 95–106. [CrossRef]
8. Næsset, E.; Gobakken, T.; Holmgren, J.; Hyyppä, H.; Hyyppä, J.; Maltamo, M. Laser scanning of forest resources: The Nordic experience. *Scand. J. For. Res.* **2004**, *19*, 482–499. [CrossRef]
9. Wulder, M.A.; White, J.C.; Fournier, R.A.; Luther, J.E.; Magnussen, S. Spatially explicit large area biomass estimation: Three approaches using forest inventory and remotely sensed imagery in a GIS. *Sensors* **2008**, *8*, 529–560. [CrossRef] [PubMed]
10. Temesgen, H.; Affleck, D.; Poudel, K.; Gray, A.; Sessions, J. A review of the challenges and opportunities in estimating above ground forest biomass using tree-level models. *Scand. J. For. Res.* **2015**, *30*, 326–335. [CrossRef]
11. Hudak, A.T.; Strand, E.K.; Vierling, L.A.; Byrne, J.C.; Eitel, J.U.H.; Martinuzzi, S.; Falkowski, M.J. Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys. *Remote Sens. Environ.* **2012**, *123*, 25–40. [CrossRef]
12. Bollandsås, O.M.; Gregoire, T.; Næsset, E.; Øyen, B.H. Detection of biomass change in a Norwegian mountain forest area using small footprint airborne laser scanner data. *Stat. Methods Appl.* **2013**, *22*, 113–129. [CrossRef]
13. Næsset, E.; Bollandsås, O.M.; Gobakken, T.; Gregoire, T.; Ståhl, G. Model-assisted estimation of change in forest biomass over an 11 year period in a sample survey supported by airborne LiDAR: A case study with post-stratification to provide "activity data". *Remote Sens. Environ.* **2013**, *128*, 299–314. [CrossRef]
14. Næsset, E.; Gobakken, T. Estimating forest growth using canopy metrics derived from airborne laser scanner data. *Remote Sens. Environ.* **2005**, *96*, 453–465. [CrossRef]
15. Yu, X.; Hyyppä, J.; Kaartinen, H.; Maltamo, M.; Hyyppä, H. Obtaining plotwise mean height and volume growth in boreal forests using multi-temporal laser surveys and various change detection techniques. *Int. J Remote Sens.* **2008**, *29*, 1367–1386. [CrossRef]
16. Nakajima, T.; Hirata, Y.; Hiroshima, T.; Furuya, N.; Tatsuhara, S.; Tsuyuki, S.; Shiraishi, N. A growth prediction system for local stand volume derived from lidar data. *GISci. Remote Sens.* **2011**, *48*, 394–415. [CrossRef]
17. Nakajima, T. Estimating tree growth using crown metrics derived from lidar data. *J. Indian Soc. Remote Sens.* **2016**, *44*, 217–223. [CrossRef]
18. Bailey, R.G. *Description of the Ecoregions of the United States*, 2nd ed.; USDA Forest Service: Washington, DC, USA, 1995; p. 108.
19. Poudel, K.P.; Temesgen, H. Methods for estimating aboveground biomass and its components for Douglas-fir and lodgepole pine trees. *Can. J. For. Res.* **2016**, *46*, 77–87. [CrossRef]
20. McGaughey, R.J. *FUSION/LDV: Software for LIDAR Data Analysis and Visualization*; USDA Forest Service, Pacific Northwest Research Station, University of Washington: Seattle, WA, USA, 2004.

21. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2017. Available online: https://www.R-project.org/ (accessed on 12 December 2017).

22. Pinheiro, J.; Bates, D.; DebRoy, S.; Sarkar, D.; R Core Team. nlme: Linear and Nonlinear Mixed Effects Models. R Package Version 3.4.1. 2017. Available online: https://CRAN.R-project.org/package=nlme (accessed on 12 December 2017).

23. Leys, C.; Christophe, L.; Klein, O.; Bernard, P.; Licata, L. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *J. Exp. Soc. Psychol.* **2013**, *49*, 764–766. [CrossRef]

24. Efron, B.; Gong, G. A leisurely look at the bootstrap, the jackknife, and cross-validation. *Am. Stat.* **1983**, *37*, 36–48.

25. Kearns, M.; Ron, D. Algorithmic stability and sanity-check bounds for leave-one-out cross-validation. *Neural Comput.* **1997**, *11*, 1427–1453. [CrossRef]

26. Poudel, K.P.; Temesgen, H. Calibration of volume and component biomass equations for Douglas-fir and lodgepole pine in Western Oregon forests. *For. Chron.* **2016**, *92*, 172–182. [CrossRef]

27. Temesgen, H.; Monleon, V.J.; Weiskittel, A.R.; Wilson, D.S. Sampling strategies for efficient estimation of tree foliage biomass. *For. Sci.* **2011**, *57*, 153–163.