Article

Tree Height Increment Models for National Forest Inventory Data in the Pacific Northwest, USA

Hyeyoung Woo [1], Bianca N. I. Eskelson [1,*] and Vicente J. Monleon [2]

[1] Department of Forest Resources Management, Faculty of Forestry, The University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada; hyeyoung.woo@alumni.ubc.ca

[2] USDA Forest Service, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA; vicente.monleon@usda.gov

* Correspondence: bianca.eskelson@ubc.ca; Tel.: +1-604-827-0629

Abstract: The United States national inventory program measures a subset of tree heights in each plot in the Pacific Northwest. Unmeasured tree heights are predicted by adding the difference between modeled tree heights at two measurements to the height observed at the first measurement. This study compared different approaches for directly modeling 10-year height increment of red alder (RA) and ponderosa pine (PP) in Washington and Oregon using national inventory data from 2001–2015. In addition to the current approach, five models were implemented: nonlinear exponential, log-transformed linear, gamma, quasi-Poisson, and zero-inflated Poisson models using both tree-level (e.g., height, diameter at breast height, and compacted crown ratio) and plot-level (e.g., basal area, elevation, and slope) measurements as predictor variables. To account for negative height increment observations in the modeling process, a constant was added to shift all response values to greater than zero (log-transformed linear and gamma models), the negative increment was set to zero (quasi-Poisson and zero-inflated Poisson models), or a nonlinear model, which allows negative observations, was used. Random plot effects were included to account for the hierarchical data structure of the inventory data. Predictive model performance was examined through cross-validation. Among the implemented models, the gamma model performed best for both species, showing the smallest root mean square error (RSME) of 2.61 and 1.33 m for RA and PP, respectively (current method: RA—3.33 m, PP—1.40 m). Among the models that did not add the constant to the response, the quasi-Poisson model exhibited the smallest RMSE of 2.74 and 1.38 m for RA and PP, respectively. Our study showed that the prediction of tree height increment in Oregon and Washington can be improved by accounting for the negative and zero height increment values that are present in inventory data, and by including random plot effects in the models.

Keywords: Forest Inventory and Analysis (FIA); tree-level increment; random plot effect; negative increment; gamma model; quasi-Poisson model; zero-inflated Poisson model

1. Introduction

Forest growth models are used to quantitatively generalize forest stand development [1]. A variety of different data sources have been used to develop growth models: permanent plots measured over a full rotation, interval plots measured over one or more growth periods, and temporary plots [2]. Permanent and interval plots may come from inventory or monitoring programs or growth and yield studies that assess silvicultural treatment effects (e.g., thinning, fertilization). Köhl et al. [3] compared permanent inventory and growth and yield study plots for the purpose of growth model development and identified the fact that inventory plots are representative of the total population as major strength for using permanent inventory plots for model development. Inventory plot data have been used to develop and calibrate equations for growth models [4–7]. Growth model

development is challenging to validate if the data used for modeling are not from a well-controlled growth and yield experiment [1], for which detailed information about site conditions and management history are available. Yet, these experimental data sets are not representative of the total population [3] across a large geographical area and gradients of environmental conditions.

Tree-level growth models typically contain equations for diameter increment, height increment, crown recession, and mortality [1]. Tree height and tree volume are closely related [8], and therefore height increment equations are frequently used to predict future tree height in stand development simulations [9–11] or to update previously measured tree height in inventories with the ultimate goal to predict and estimate tree volume. Two approaches are commonly used to model height increment [1]: (1) a potential multiplied by modifier approach, where the potential height increment is adjusted by a modifier that accounts for size, site, and competition [12,13] or (2) a linear or nonlinear equation that models realized height increment [14,15]. The use of random effects to account for and take advantage of hierarchical data structure has emerged in the late 1990s [1]. Mixed-effects models allow incorporating locality effects [16,17] and the hierarchical data structure of plot-based forest inventory data into height increment models [18,19].

The Forest Inventory and Analysis (FIA) Program of the United States Forest Service monitors forest resources with a nationwide systematic sample of all forested lands [20]. Due to budgetary constraints, the data collection protocol in the Pacific Coast states was changed in 2015 to only sample a subset of trees for height measurement within each plot [21], which leads to the need of predicting the height of the remaining trees. Currently, the Pacific Northwest (PNW) research station employs a multi-step approach for tree height estimation at a particular time: (1) obtain diameter at both times either by measurement or modeling; (2) predict the height at both times from the diameter using height-diameter equations [22]; (3) estimate the height increment as the difference in the predicted heights at time 1 and time 2; (4) add the estimated height increment to the height at time 1 to estimate the height at time 2. This approach ensures the height at time 2 is always greater than that at time 1. The same method is used in other FIA work units (e.g., Northern and Southern research stations), based on regional calibration of diameter increment models [6,23]. Regional height–diameter equations are often built using plantation-grown trees [24–26] that do not represent the total population, and height increment is indirectly obtained as the difference between two heights estimated from these regional height-diameter equations [27]. This approach requires multiple modeling steps that are each associated with model uncertainties [28], and demand future measurements or predictions of tree diameters to obtain height increment [25]. Moreover, this approach does not incorporate plot- or stand-level variation. Direct modeling of tree height increment and a mixed effect modeling approach reduce the steps for future height increment prediction and localize prediction, respectively [27].

Even though height increment models developed from inventory and monitoring data can account for regional and species-specific characteristics by incorporating random effects in the model [18], they may still encounter difficulties in application pertaining to the availability of predictive variables and measurements. Age, one of the tree attributes that is frequently used in height increment modeling, is hard to measure in natural stands [29–31] and is not necessarily available from inventory and monitoring data. Forest inventory field measurements often include many unusual data points of tree heights (e.g., negative or zero height increments) due to mensuration errors [32] or any external events such as disease, fire, weather, or mechanical damage [33,34]. However, dropping these unusual measurements from the dataset used for modeling may result in the equations not being representative of the whole population of trees, which is one of the major strengths of using inventory data for growth modeling [3]. Dropping negative or zero height increments would tend to bias estimates of height increment upwards. Therefore, models which allow height increment predictions to all tree records available in the inventory and monitoring database—including trees with atypical height increment observations—are needed. Height increment is defined as the rate of change in height over a specified period of time, which is not the same as height growth, the increase in height of each individual tree through time [1]. Therefore, height growth is always positive, while height increment could also be negative or zero [35].

The purpose of this study was to develop a tree height increment model that is compatible with continuous regional and national forest inventory and monitoring programs and does not use stand age as predictor variable. The model should be applicable to the inventory data from a broad population of trees with different growth conditions and damage records. For this purpose, we utilized a large sample of FIA inventory plots with 10-year remeasurement interval for ponderosa pine and red alder in forests of Oregon and Washington, USA. The main objectives were to (1) predict realized tree height increment in a single step, thus simplifying the current multi-step approach used; (2) include observations of negative or zero height increment in the modeling process; and (3) achieve more accurate prediction than the current procedure which applies regional height-diameter equations to obtain local height increment predictions. To meet these objectives, we examined six different models that (1) use predictor variables that are commonly measured in regional and national inventories; (2) account for unusual observations (i.e., negative or zero height increment) in the model form; and (3) incorporate random plot effects to improve prediction accuracy by localizing the predictions. We compared model performance through cross-validation.

2. Materials and Methods

2.1. Data

Forest Inventory and Analysis (FIA) data from Oregon (OR) and Washington (WA) measured between 2001 and 2015 for two tree species—ponderosa pine (PP, *Pinus ponderosa* Dougl. ex P. Laws. & C. Laws.) and red alder (RA, *Alnus rubra* Bong.)—were used in this analysis. FIA data contain a spatially balanced sample of inventory plots every 24 km^2 [20]. All live trees over 2.54 cm of diameter at breast height (DBH) were measured within accessible forested plots (i.e., at least 10 percent of potential canopy cover, excluding agricultural and urban areas). We retained plots with five or more trees for modeling and validation purposes. This resulted in 2172 RA trees from 193 plots and 6977 PP trees from 573 plots. Each tree had repeated observations (Time 1 and Time 2) with a 10 year remeasurement interval, except for a few remeasurements with a 9 year interval (i.e., 46 for PP and 26 for RA).

For each tree, the field crews noted if the top was broken and missing, and then (1) recorded the actual height of the tree, and (2) if the top was broken, estimated the height that the tree would have had if the crown was intact [36] (p. 92). Thus, trees with a broken top have a different actual height and field-estimated total height, and were excluded from the analysis. A previously broken top was considered recovered and, therefore, unbroken, if the diameter of the new leader was 1/3 the diameter at the point of breakage. When measuring heights, forked trees were treated in the same way as unforked trees [36]. In our analysis, only the trees with an unbroken top were included in the data. However, the analysis included trees with previously broken and recovered tops, even though the recovered height may be less than that before breakage.

The response variable—ΔHT, 10 year height increment in meters—was computed as the difference in heights measured at Time 2 and Time 1 for trees with unbroken tops, as defined in the previous paragraph (i.e., $\Delta HT = HT_2 - HT_1$). The height increment ranged from −6.70 to 15.24 m for PP, and from −7.62 to 17.68 m for RA (Figure 1). Most height increment equations allow positive values only, therefore it is impossible to model negative or zero values of ΔHT with those equations. In our analysis, however, the negative or zero ΔHT accounted for noticeable percentages of the tree data (e.g., 13.21% for PP and 15.75% for RA). We tried to exclude negative observations caused by physical damage or measurement error based on information collected by field crews as damage codes and optional notes in the tally sheet. The attempt of excluding negative height increment values this way, however, only accounted for a small portion of all negative values. When we filtered the data by damage codes that addressed damage to tree tops (e.g., broken or suppressed due to weather or any mechanical injury), 230 out of 922 and 27 out of 342 negative height increment observations were removed for PP and RA, respectively. When we examined the optional field crew notes including the words referencing broken tops or measurement error, additional 42 and 40 negative height increment values for PP and RA were removed, respectively. However, both filtering methods

eliminated observations with positive growth as well (e.g., 1196 using the damage code and 822 using the optional field crew notes out of all 7885 positive height increment observations). Because our approach of selectively removing observations based on the available information could not successively withdraw all the observed negative values, we decided to include all data as is and deal with the negative values during the modeling stage.
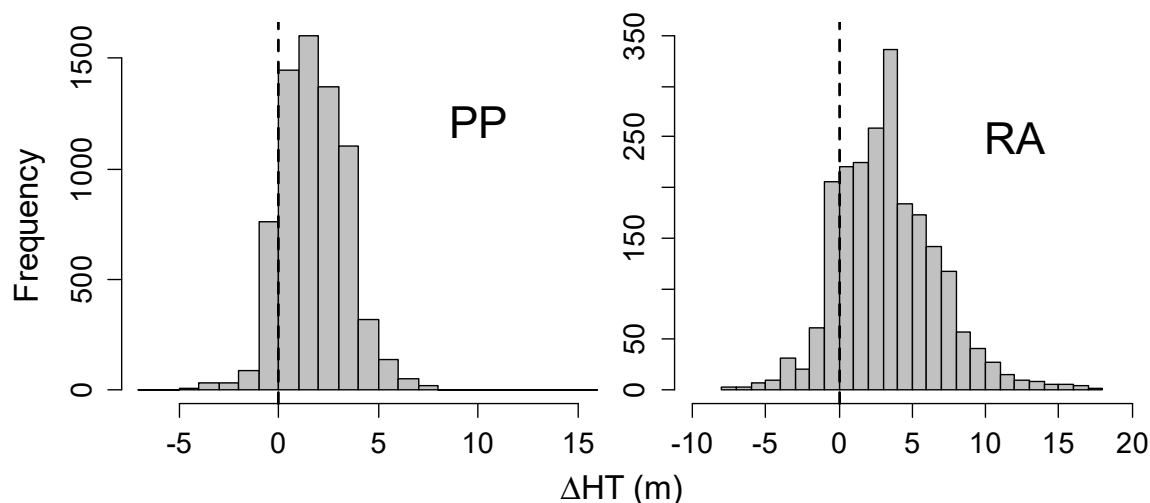

Figure 1. Histograms of height increment (ΔHT) for 10 years of ponderosa pine (PP) and red alder (RA) from the Forest Inventory and Analysis (FIA) data. Dashed vertical lines are marked at zero separating negative from positive height increment.

Auxiliary tree-level variables such as height, DBH, compacted crown ratio, and crown class were selected as potential predictor variables based on existing literature [37–39]. Plot-level variables such as site productivity, basal area of all species, elevation, slope, and aspect were also considered to be important predictors [5,15,37]. Full description of the FIA-measured variables can be found in [21]. All tree-level predictor variables and basal area per ha were from Time 1 measurements. Summary statistics of the response and the auxiliary variables obtained from the FIA database are displayed in Table 1.

Table 1. List and summary statistics of tree-level and plot-level variables used for modeling PP and RA tree height increment in Washington (WA) and Oregon (OR).

	PP (n = 6977)			RA (n = 2172)		
	Minimum	Mean (SD)	Maximum	Minimum	Mean (SD)	Maximum
Tree-level						
ΔHT (m)	−6.71	1.81 (1.60)	15.24	−7.62	3.33 (3.36)	17.68
DBH (cm)						
Time 1	2.54	33.02 (22.05)	149.61	2.54	22.97 (9.74)	79.76
Time 2	2.54	35.91 (21.90)	147.83	3.56	26.76 (9.64)	82.55
Height (m)						
Time 1	1.52	16.54 (9.46)	56.69	3.56	18.33 (5.86)	43.28
Time 2	2.13	18.35 (9.16)	57.91	3.66	21.67 (5.39)	46.03
Compacted crown ratio (%) [1]	1.00	54.01 (18.38)	99.00	5.00	37.14 (16.48)	95.00
	Factor with 2 levels (dominant/intermediate)					
Crown class code	Number trees per class: 5889/1088			Number trees per class: 1795/377		
Plot-Level						
Basal area (m² ha⁻¹)	2.14	22.98 (10.82)	63.50	4.12	34.65 (15.94)	112.97

Elevation (m)	1300	4386 (1108.33)	6700	100	932.9 (691.55)	3100
Slope (percent)	0.00	18.63 (16.85)	110.00	0.00	31.42 (25.42)	100.00
Aspect (cosine value of degree)	−1.00	0.18 (0.76)	1.00	−1.00	0.23 (0.72)	1.00
Site productivity [2]	Factor with 3 levels (0–3.5, 3.5–8.4, >8.4 m³ ha⁻¹ year⁻¹)					
	Number trees per class: 107/4223/2647			Number trees per class: 1719/431/22		

[1] Compacted crown ratio: percent of tree bole with live foliage. [2] Site productivity: capacity to grow crops of industrial wood (potential growth based on the culmination of mean annual increment of fully stocked natural stands). The FIA database classifies site productivity into seven categories based on the potential volume growth [21], and the categories were reclassified in this analysis into three factors.

2.2. Models

The current process for estimating tree height increment in the FIA program of the Pacific Northwest (PNW) research station requires DBH measurements at both Time 1 and Time 2, because the tree heights are estimated at both times using the Chapman–Richards height-diameter equation [40] and the difference between the estimated tree heights is added to the height at Time 1. The current method was considered as our baseline model estimating the height increment ($\Delta \widehat{HT_{base}}$) over time (t) as follows;

$$\widehat{HT}_t = 1.37 + \alpha\left(1 - e^{-\beta*DBH_t}\right)^{\gamma}, t = 1,2, \tag{1}$$

$$\Delta\widehat{HT_{base}} = \widehat{HT}_2 - \widehat{HT}_1, \tag{2}$$

where α = 28.26, β = 0.0590, and γ = 1.3040 for red alder (RA), and α = 58.81, β = 0.0123, and γ = 1.1766 for ponderosa pine (PP), respectively [22]. In this approach, the predicted height increment is always positive (Table 2).

The models that we considered as alternatives to this baseline approach needed to accommodate the negative height increment observations in the dataset. Three strategies for dealing with negative response values were taken into consideration: (i) adding a constant to the height increment values to shift the entire distribution of the response to be positive; (ii) setting all negative values to zero; and (iii) using a nonlinear model, which allows negative height increment observations but only predicts positive values. These three strategies are not uncommon in regression analyses that involve non-negative variable transformation or mixed distributions [41,42]. The first strategy requires prior knowledge about the distribution of the response variable, and the second resembles the replacement of missing data strategy [43]. Two models for each of the three options of treating the observed negative height increment values were explored (Table 2). In order to have positive height increment for all observations, we chose the rounded absolute of the maximum negative value +1 as the constant c (i.e., c = 8 m for PP and c = 9 m for RA) for the log-transformed linear and gamma models. The negative increment observations were also dealt with using a nonlinear exponential model (i.e., $\Delta\widehat{HT} = \exp(\mathbf{X\beta} + \mathbf{Zu})$), where $\mathbf{X}$ and $\mathbf{Z}$ are the fixed and random-effects matrices and $\boldsymbol{\beta}$ and $\mathbf{u}$ are the vectors parameters to be estimated. This resulted in a total of six height increment models that were examined in this study (Table 2).

Depending on the model type, different procedures (Table 2) in SAS 9.4 (SAS Institute, Cary, NC, USA) were chosen to fit the models. Except for the baseline model, the models included fixed effects of the tree- and plot-level variables (Table 1) as well as a random plot effect. After fitting the models by species using the full data (i.e., n = 6997 for PP and n = 2172 for RA), variables that were not significant at α = 0.05 were excluded from the model. Significance of the variables was tested using likelihood ratio tests.

Table 2. Details of models and SAS procedures used to fit models. In each model, the negative height increment values were incorporated 1) as is, 2) as positive by adding constant, or 3) as zero. The resulting predicted height increment could be either only positive, or both positive and negative.

Model	Plot-Level Random Effect	Negative Height Increment Strategy	Predicted Height Increment	SAS Procedure	References
(1) Baseline	No	Data used as is	Positive	GLM	[22]
(2) Nonlinear exponential	Yes	Data used as is	Positive	NLMIXED	[44]
(3) Log-transformed linear	Yes	Constant added	Positive/negative	GLM	[45]
(4) Gamma	Yes	Constant added	Positive/negative	GLIMMIX	[46]
(5) Quasi-Poisson	Yes	Replaced by zero	Positive	GLIMMIX	[47]
(6) Zero-inflated Poisson	Yes	Replaced by zero	Positive	NLMIXED	[47]

2.3. Evaluation of Predictive Performance of Models

We performed leave-one-out cross-validation for model evaluation [48], similar to Temesgen et al. [18]. To account for the random plot effect during the process, we conducted the cross-validation as described in Figure 2. The cross-validation process produced one validation dataset accommodating observations from every plot during steps 1–6 (Figure 2) for each species, and the height increments of the validation dataset were predicted by the six models (Table 2). With 200 iterations of random subsampling and variation in the number of subsamples (m = 1, 2, 3, 4), 800 validation datasets by species were created and used for model performance evaluation.
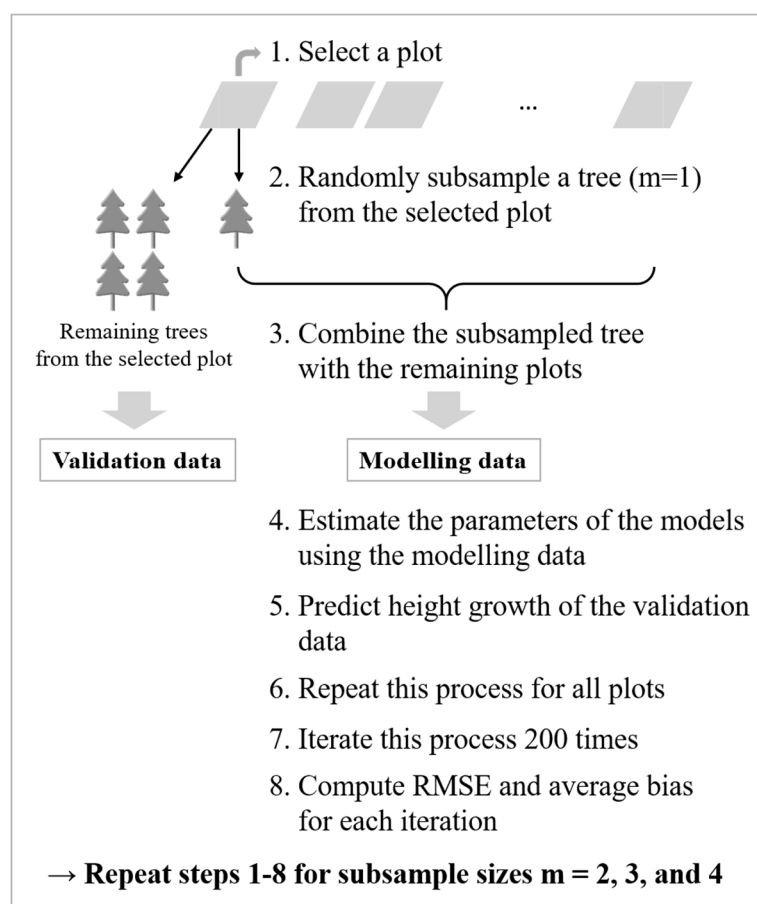

1. Select a plot

2. Randomly subsample a tree (m=1) from the selected plot

Remaining trees from the selected plot

3. Combine the subsampled tree with the remaining plots

Validation data

Modelling data

4. Estimate the parameters of the models using the modelling data

5. Predict height growth of the validation data

6. Repeat this process for all plots

7. Iterate this process 200 times

8. Compute RMSE and average bias for each iteration

→ **Repeat steps 1-8 for subsample sizes m = 2, 3, and 4**

Figure 2. Schematic description of the cross-validation process used in this study.

With the predicted height increment for the validation dataset, we calculated the bias (average difference) and root mean square error (RMSE) at iteration j (j = 1, …, 200) of model k (k = 1, … , 6) as follows:

$$Bias_k = \frac{1}{200} \sum_{j=1}^{200} \left[\sum_{i=1}^{n-lm} \left(y_{ijk} - \hat{y}_{ijk} \right) \Big/ (n - lm) \right], \tag{3}$$

$$RMSE_{jk} = \sqrt{ \sum_{i=1}^{n-lm} \left(y_{ijk} - \hat{y}_{ijk} \right)^2 / (n - lm)}, \tag{4}$$

where y_{ijk} and $\hat{y}_{ijk}$ are the observed and predicted height increment of the ith tree for model k at iteration j, respectively. The number of plots (l) in the data were 193 and 573 for RA and PP, respectively. Consequently, each validation dataset with the subsample size m had $n - lm$ observations. For the log-transformed linear and the gamma models, y_{ijk} was corrected with the constant added, while the other models used the actual height increment including negative values. Log-bias in the log-transformed linear model was corrected following Duan's smearing estimator [49]. All cross-validation and following analysis were conducted using the SAS statistical software 9.4 (SAS Institute, Cary, NC, USA).

3. Results

There were 922 observations showing negative or zero height increment for PP and 342 for RA, accounting for approximately 13.21% and 15.75% of the PP and RA data sets, respectively. The observed 10-year mean height increment when negative or zero values were included was 3.3 m for RA, and 1.8 m for PP. Excluding the negative or zero values, the observed 10-year mean height increment was 4.2 m for RA, and 2.2 m for PP. When the negative values were replaced with zero, the mean increment was 3.5 and 1.9 m for RA and PP, respectively.

3.1. Model Fitting

Among the eight tree-level and plot-level predictor variables, DBH at Time 1, height at Time 1, crown class, site productivity, and slope were found to be significant in predicting height increment of RA in log-linear, gamma, and quasi-Poisson models (Table 3). Height increment of PP had the same predictor variables, except slope was replaced by elevation (Table 3). Compacted crown ratio also exhibited a strong relationship (e.g., $p < 0.0001$) with the height increment of PP. In the zero-inflated Poisson model, the production of negative or zero height increment (i.e., zero model) was associated with DBH and height for RA, and DBH, height, compacted crown ratio, and elevation for PP (Table 3). Basal area per ha was an additional predictor variable for RA height increment. In the zero-inflated models, the probability of having negative or zero tree increment was larger for taller trees in both species. DBH of RA, and compacted crown ratio and elevation of PP were positively related to height increment, while they were negatively associated in the zero model. The plot-level random effect was significant for all models (i.e., $p < 0.0001$).

Table 3. Coefficients of suggested models using all data (n = 6997 for PP and n = 2172 for RA).

	Model Coefficients					
	Nonlinear	Log-Trans-Formed Linear	Gamma	Quasi-Poisson	Zero-Inflated Poisson	
					Fixed Effects	Zero Model
RA						
Intercept	3.4546	3.9072	3.8864	3.0068	3.6571	−4.5326
DBH (cm)	0.0420	0.0140	0.0136	0.0451		−0.0883
Height (m)	−0.0238	−0.0111	−0.0101	−0.0296	−0.0154	0.0528
Crown class (dominant)		0.1853	0.1726	0.5621		
Site productivity (medium)		0.1158	0.1175	0.3593		
(high)		−0.0305	−0.0267	−0.1018		
Basal area (m³)					−0.0028	
Slope (degrees)	−0.0030	−0.0016	−0.0016	−0.0053		0.0079
Variance of random plot effect	62.2362	0.0240	0.0212	0.2975		0.1930
PP						
Intercept	2.1331	3.3161	3.3201	1.2939	1.4919	−3.6185
DBH (cm)	0.0098 **	0.0029	0.0029	0.0100		0.0336 **
Height (m)	−0.0090	−0.0032	−0.0030	−0.0110	−0.0017	0.0217
Compacted crown ratio (percent)	0.0085	0.0014	0.0014	0.0066	0.0084	−0.0192
Crown class code (dominant)		0.1153	0.1116	0.5409		
Site productivity (medium)		0.1382	0.1411	0.6489		
(high)		0.0831	0.0789	0.3518		
Elevation (m)	−0.0001	−0.00001 **	−0.00001 **	−0.00004 *		0.0001 *
Variance of random plot effect	17.4055	0.0084	0.0075	0.1948		0.1587

All variables are strongly significant (i.e., $p < 0.0001$) unless marked otherwise: ** $p < 0.001$; * $p < 0.05$
DBH, height, and crown class are Time 1 measurements.

3.2. Comparison of Model Performance

The baseline model substantially overestimated the height increment for both species (e.g., 1.4 m for RA and 0.5 m for PP on average) (Table 4). The magnitude of mean bias relative to the observed mean height increment was larger for RA (42.81%) than for PP (27.62%) in the baseline model. The mean bias was lowest for the log-transformed linear model for both species across all subsample sizes m, except for the gamma model with $m = 1$ and $m = 2$. The average mean bias across subsample sizes was the largest for the quasi-Poisson model (0.20 m for RA and 0.08 m for PP). All suggested models reduced the prediction bias at least by 83% compared to the baseline model (Table 4).

Table 4. Mean bias and root mean square error (RMSE; standard deviation in parentheses) of predicted height increment for the six models for both species, RA and PP, across subsample sizes *m* = 1, 2, 3, 4.

	Mean Bias (m)				RMSE (std. dev) (m)			
	m = 1	*m* = 2	*m* = 3	*m* = 4	*m* = 1	*m* = 2	*m* = 3	*m* = 4
RA								
Baseline	1.41	1.42	1.43	1.44	3.31 (0.34)	3.32 (0.40)	3.33 (0.44)	3.35 (0.52)
Nonlinear	0.07	0.04	0.03	0.03	2.86 (0.52)	2.74 (0.50)	2.68 (0.50)	2.65 (0.49)
Log-trans-formed linear	0.00	−0.01	−0.01	0.00	3.43 (1.08)	2.86 (0.64)	2.70 (0.59)	2.60 (0.52)
Gamma	−0.13	−0.16	−0.17	−0.18	2.72 (0.47)	2.62 (0.46)	2.56 (0.48)	2.52 (0.48)
Quasi-Poisson	0.23	0.18	0.19	0.18	3.02 (1.04)	2.74 (0.62)	2.64 (0.55)	2.57 (0.50)
Zero-inflated Poisson	0.13	0.12	0.11	0.10	2.90 (0.61)	2.78 (0.58)	2.72 (0.57)	2.68 (0.55)
PP								
Baseline	0.49	0.50	0.50	0.51	1.41 (0.13)	1.41 (0.16)	1.40 (0.19)	1.39 (0.20)
Nonlinear	0.05	0.04	0.04	0.03	1.43 (0.17)	1.39 (0.17)	1.36 (0.18)	1.34 (0.19)
Log-trans-formed linear	−0.00	−0.04	−0.05	0.02	1.78 (0.45)	1.52 (0.27)	1.42 (0.20)	1.38 (0.20)
Gamma	0.00	0.00	0.07	0.07	1.38 (0.14)	1.34 (0.16)	1.31 (0.17)	1.29 (0.18)
Quasi-Poisson	0.08	0.08	0.08	0.07	1.45 (0.33)	1.39 (0.24)	1.35 (0.20)	1.32 (0.19)
Zero-inflated Poisson	0.06	0.07	0.06	0.06	1.45 (0.25)	1.40 (0.21)	1.37 (0.19)	1.35 (0.19)

For RA, we obtained lower mean RMSE values than for the baseline model for all suggested models (Figure 3). The magnitude of RMSE reduction was different by model and subsample size. There was a clear tendency for RMSE values to decrease for all models with random plot effects as the number of subsamples included in the modeling set increased (Table 4; Figure 3). When *m* = 1, RMSE was lowest for the gamma model followed by the zero-inflated Poisson, quasi-Poisson, and then the log-transformed linear model. The gamma model consistently showed the lowest RMSE value across all models, while the log-transformed linear model exhibited a rapid decrease of RMSE as *m* increased (Figure 3). When *m* = 4, the quasi-Poisson and log-transformed linear models outperformed the zero-inflated Poisson model. The largest difference from the mean RMSE of the baseline model (3.3 m) for RA occurred in the gamma model at *m* = 4 (2.5 m) with 24.32% reduction in RMSE (Table 4).
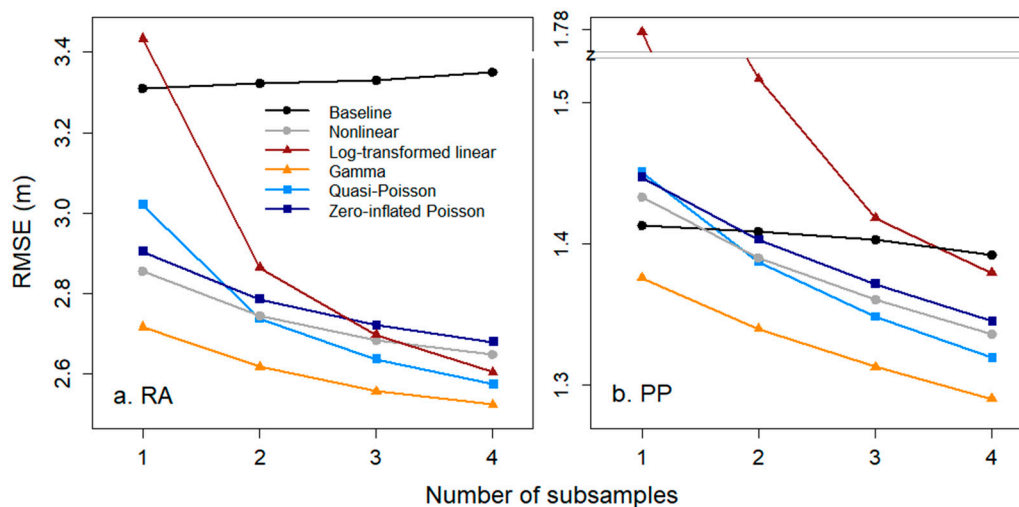

Figure 3. Mean RMSE of six models across subsample size *m* = 1, 2, 3, 4 plotted within the range of obtained values. RMSE of the baseline model was calculated using the same set of validation data used for the other models.

　　　For PP, on the other hand, the baseline model had lower RMSE at *m* = 1 than the other models except for the gamma model. There is a similar pattern of decreasing RMSE with increasing subsample size *m* as was observed for RA (Figure 3). Nonetheless, the log-transformed linear model performed particularly poorly with small *m* values (e.g., *m* = 1, 2) and only achieved a smaller RMSE than the baseline model for *m* = 4 (Figure 3). This was due to extreme observations (i.e., height increment of −6.7 and −6.4 m) that were included in the modeling or validation dataset, but the impact of such unusual observation drastically subsided as *m* reached 4. All the other models managed to outperform the baseline model with subsample sizes of *m* ≥ 2. The gamma model performed best in terms of the smallest RMSE for PP as well as for RA (Table 4), and the amount of reduction in RMSE was about 8% of the currently used baseline model (mean RMSE of the baseline model =1.4 m).

　　　There was large variation in RMSE across the 200 simulation sets for the log-transformed linear and quasi-Poisson models especially for *m* = 1 for both species, while the gamma model showed less variability in RMSE regardless of *m* (Figure 4). The large variation of RMSE subsided as the modeling dataset included more subsamples *m* per plot (i.e., *m* = 3, 4). For PP, the variability in RMSE was substantial for both *m* = 1 and *m* = 2, but leveled-off for *m* ≥ 3. The variability in RMSE of the baseline model slightly increased as the number of trees in the validation dataset decreased along with increases in *m*.
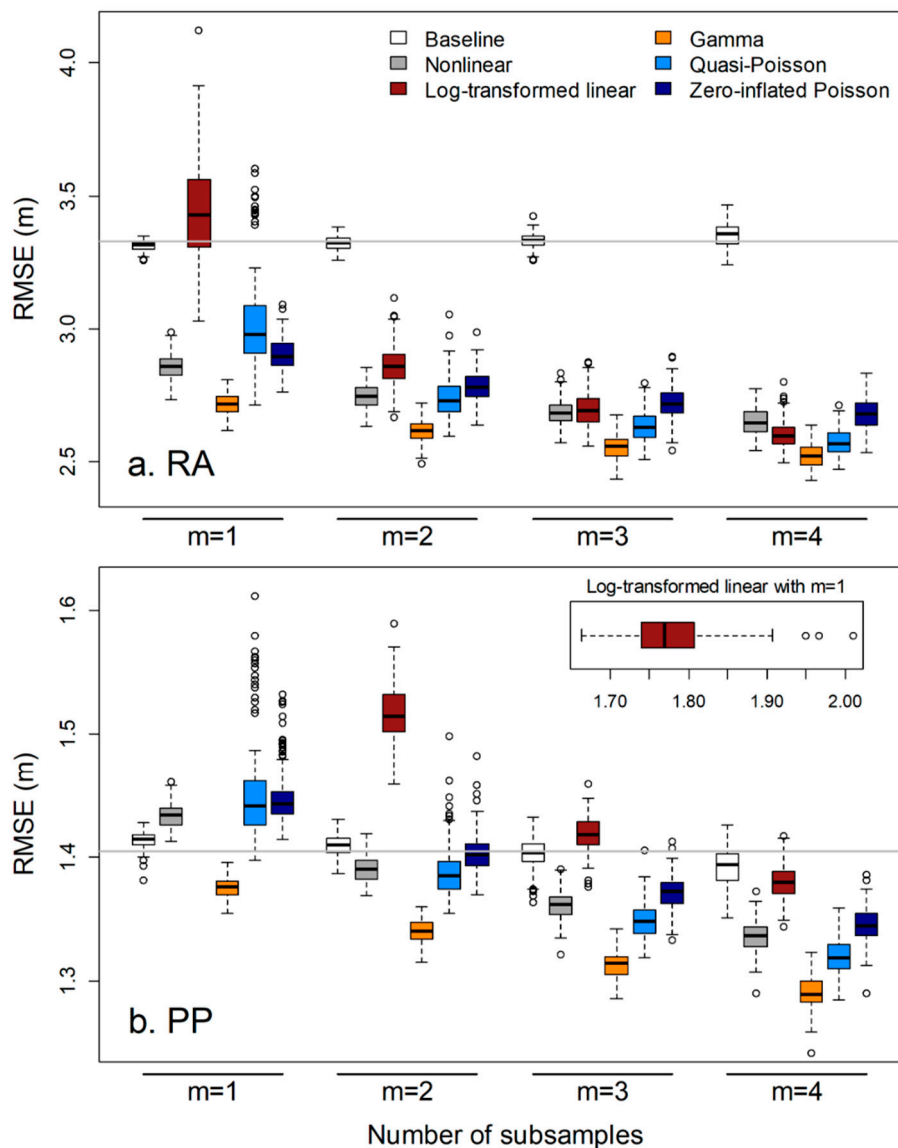

Figure 4. Boxplots of RMSE by each model for (**a**) RA and (**b**) PP based on 200 iterations of validation process. The gray horizontal line is the mean RMSE of the baseline model across all subsample sizes.

4. Discussion

The negative height increment values may be due to measurement errors. When height increments are recorded based on remeasurements on the same tree, the measurement errors can be substantial, especially for slow-growing trees and dense stands with steep terrain [27]. Observed negative or zero height increment values may further be due to poor tree shape or physical and mechanical damages such as topped and partially recovered trees [36]. The proportions of negative height increment values (i.e., 13.21% for PP and 15.75% for RA) were too large to be ignored and to be dropped in the modeling process without substantially biasing the resulting estimators. The mean increment values for both species also differed substantially when calculated with and without the negative or zero values, which further indicated that excluding the negative or zero values from the modeling process may impact modeling results as truncating the data from below will result in overpredictions. This will produce models that are not representative of the whole population of trees [3]. Thus, we decided to explicitly include all the negative or zero observations in the modeling process.

4.1. Interpretation of Model Fit and Variables

The general relationship between the height increment and the tree- and plot-level variables was consistent among the species and models. Shorter, fatter, and dominant trees tended to grow more in less steep and lower elevation conditions. The tendencies mostly agree with previous literature characterizing tree height increment in relation with individual tree attributes and competition measures [50–52]. Initial height and DBH were among the most important covariates for predicting height increment in previous research [53,54]. Similar to our findings plot or stand-level variables, such as basal area per hectare and site factors (e.g., slope, aspect, elevation, and forest type) have also been found to be highly associated with tree height increment [5,15,37,55].

During data exploration, we found that negative values tended to have higher prevalence in tall trees than in short trees. These tendencies appeared in the estimated coefficients as well. This may be explained by the fact that when measuring standing trees in the field, the trigonometry-based instruments such as hypsometers are more sensitive to errors as the top height is higher, thus there may be more chance of obtaining erroneous height measurement for taller trees [56,57]. Another explanation could be that taller trees have slower height growth than shorter trees [58]. The opposite trends of variables in the fixed effects and in the zero models for the zero-inflated Poisson model suggest that the trees with negative or zero growth may possess a different data-generating process than the positive increment values. Thus, zero-inflated or hurdle models [59], which allow modeling the negative or zero increment observations separately from the positive increment observations may be useful in modeling tree height increment when the data include negative or zero increment [60].

The significant plot-level random effects across all models suggest that the inclusion of a random plot effect contributed substantially to the reduction of random variance [18]. This was also confirmed by the reduction in RMSE as subsample size *m* increased, because a larger number of subsamples *m* included in the modeling dataset improved the estimation of the random plot effect. Previously shown gains in precision of height prediction by inclusion of random effects [18] support our decision of incorporating the plot-level random effect to account for the hierarchical data structure of the inventory data, thus allowing localized height increment predictions. As subsamples of height are typically measured on regional and national inventory and monitoring plots, random effects can be estimated from existing databases during future model applications when height increment is predicted for trees without height measurements.

4.2. Comparison of Model Performance

Differences in the relative mean bias between RA and PP (e.g., 42.81% vs. 27.62%) imply that the current height diameter equation [22] may be less accurate for RA than for PP. It is generally considered that height measurements tend to be challenging for irregular or round-crowned trees, hence height increment models are more difficult to develop for hardwood species than for conifers

[1]. Average absolute gains in RMSE were smaller for PP (0.04–0.10m) than for RA (0.60–0.83m) for the gamma model that performed best among any of the suggested models. At least for RA, this may be a considerable improvement over the baseline model, as RA trees are much shorter than PP trees (e.g., Hanus et al. [61]; average height of 25.6 m for PP vs. 9.2 m for RA). Differences in performance among the suggested models themselves were quite small, especially for larger subsample sizes ($m \geq$ 3; approximately 0.15 m).

Based on the previously mentioned small differences in performance among the suggested models, it may not matter too much which model form is chosen as long as the model accounts for negative or zero observations in some way and is not too sensitive to extreme observations in the data set. Especially, the log-transformed linear model was shown to be sensitive to extreme values when small subsample sizes were used. Changyong et al. [62] demonstrated that log transformation can result in more variable and more skewed data through simulation. Another model performance characteristic to consider when selecting the model may be the variability of the model across subsample sizes m. The gamma model, which had the lowest RMSE for all subsample sizes, demonstrated very low variability in performance across subsample sizes, which provides another advantage of this modeling approach. Similarly, the nonlinear model showed little variability across the subsample sizes. On the contrary, the zero-inflated and quasi-Poisson models exhibited high variability, especially for smaller subsample sizes. This may suggest sensitivity of these models to extreme values when only small subsample sizes are available [63], and makes these two models less desirable to use for height increment predictions.

Zero-inflated and quasi-Poisson models adopted the second strategy of negative value treatment (i.e., set negative values to zero), inevitably showed larger mean bias, because the response values were modified during the modeling process (Table 2). However, the differences in bias across the models were minor, and all the suggested models managed to reduce the prediction bias at least by 83% compared to the baseline model. Although this result is possibly also due to the fact that the suggested models were built from the same dataset used for the cross-validation while the baseline model was not, the improvement in the accuracy of prediction is noteworthy. Acquisition of a new dataset to validate the model will confirm the true performance of our suggested models contrasted with the current method of height increment prediction.

The quasi-Poisson model exhibited the largest mean bias but smaller RMSE, showing typical variance-bias trade-off in model prediction [64]. It is noteworthy that for the shifted models with the constant-added response (e.g., log-transformed linear model and gamma model), the constants should be determined before the modeling process. We chose the constants from the observed maximum absolute value of the negative height increment, c = 8 m for PP and c = 9 m for RA, but these values may vary with different datasets. Considering the need of prior information to find the constant to be added as a drawback of the gamma model, we recommend the use of the quasi-Poisson model for future height increment predictions based on its second-lowest RMSE value across the two species examined in this study.

The zero-inflated Poisson model provides the ability to explicitly model negative and zero height increment values separately from the positive values [65]. Therefore, this model may be able to capture the mechanism that creates the negative or zero height increment values, provided any of the available predictor variables have the explanatory power to do so. We also implemented a zero-inflated negative binomial model during the model fitting and validation process. A zero-inflated negative binomial model would similarly to the zero-inflated Poisson model allow capturing the mechanism that creates the negative or zero values [65]. However, the results were not reported in this paper because the differences in performance between the zero-inflated negative binomial and the zero-inflated Poisson were minimal. However, the zero-inflated negative binomial model required excessive time for model fitting with low convergence rate compared to the zero-inflated Poisson model.

The main improvement for all models appeared to be for larger subsample sizes ($m \geq 3$). This may suggest that the inclusion of the random effect, which improves with larger subsample sizes [18], may have provided the most improvement over the baseline model, which does not include a

random effect and is therefore unable to localize height increment predictions. The inclusion of a random effect to allow localized predictions may be the most important strategy in improving upon the currently used baseline model. It may be possible to include random plot effects in the currently used height prediction models [22] that are used in the baseline model. This could lead to similar results for height increment predictions as were achieved by the gamma and the other models that allowed localized height increment predictions.

4.3. Height Increment vs. Height Growth Models

From an inventory and monitoring point of view, it is necessary to predict height increment for trees for which heights were not measured. The main interest for inventory and monitoring applications is therefore to estimate changes in measured height in the population of all trees, under the same measurement conditions as those of the measured trees. Therefore, the focus is strictly on height increment prediction, where increment is defined as the rate of change in height in a specified period of time [1]. On the other hand, height growth refers to the actual increase in height of each individual tree through time [1]. Height growth can be considered a latent process [66], which contributes to the observed differences in height increment among other possible factors (e.g., diameter increment, crown recession, and mortality). Using height increment models to infer height growth assumes that a population-level change is the same as growth at the individual tree level. This is especially a big assumption when negative or zero height increment values are included in the modeling process as suggested in this study for predicting height increment in the framework of inventory and monitoring databases.

5. Conclusions

Height increment models are major components of growth and yield models and are also needed to update inventory and monitoring databases when tree heights are only subsampled. Inventory data contain many negative or zero height increment observations that cannot be explained by damage records. Therefore, these negative or zero height increment values are part of the total population of trees, as measured using the inventory field protocols. Yet, when developing height increment equations from inventory and monitoring databases, negative or zero height increment observations are often dropped from the modeling data set, which results in models that may not be fully representative of the total population of tree measurements. Therefore, negative or zero height increment observations should not be ignored in model development.

In this study, we demonstrated that the currently used baseline method of the FIA inventory in the Pacific Northwest region—predicting height increment as difference in modeled heights—can be improved upon by using random effects models that explicitly account for negative or zero height increment observations in the modeling process and include random plot effects to allow localized predictions. We employed a nonlinear model, two models that added constants to the negative observations to shift data into the positive range, and two models that set all negative values to zero.

Models for red alder (RA) improved the baseline model substantially, while the reduction in RMSE was not that large for ponderosa pine (PP) models. The gamma model performed best for both species across all subsample sizes ($m = 1, 2, 3, 4$), followed by the quasi-Poisson and non-linear models for larger subsample sizes ($m \geq 3$). Gains in model performance over the baseline model were more pronounced for larger subsample sizes.

The somewhat moderate observed gains in model performance, especially for PP, may not justify the need to fit more complicated models. As observed differences between the baseline model and models that account for negative or zero observations and include a random effect were not that large, we suggest that models are used that are most convenient for a specific application. The inclusion of random effects to localize predictions are suggested for any model and could improve the currently used baseline model as well.

Author Contributions Conceptualization, B.N.I.E. and V.J.M.; Methodology, H.W., B.N.I.E., and V.J.M.; Software, H.W.; Formal analysis and investigation, H.W.; Resources, B.N.I.E. and V.J.M.; Data curation, V.J.M.

and H.W.; Writing—original draft preparation, H.W and B.N.I.E.; Writing—review and editing, B.N.I.E. and V.J.M.; Visualization, H.W.; Supervision, B.N.I.E.; Project administration and funding acquisition, B.N.I.E. and V.J.M. All authors have read and agreed to the published version of the manuscript.

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

References

1. Weiskittel, A.R.; Kershaw, J.A.; Vanclay, J.K.; Hann, D.W. *Forest Growth and Yield Modeling*; Wiley-Blackwell: Chichester, UK, 2011.
2. Burkhart, H.E.; Tomé, M. Modeling Forest Stand Development. In *Modeling Forest Trees and Stands*; Springer: Dordrecht, The Netherlands, 2012; pp. 233–244, ISBN 978-90-481-3170-9.
3. Köhl, M.; Scott, C.T.; Zingg, A. Evaluation of permanent sample surveys for growth and yield studies: A Swiss example. *For. Ecol. Manag.* **1995**, *71*, 187–194.
4. Walters, D.K.; Ek, A.R. Whole stand yield and density equations for fourteen forest types in Minnesota. *North. J. Appl. For.* **1993**, *10*, 75–85.
5. Lacerte, V.; Larocque, G.R.; Woods, M.; Parton, W.J.; Penner, M. Calibration of the forest vegetation simulator (FVS) model for the main forest species of Ontario, Canada. *Ecol. Modell.* **2006**, *199*, 336–349.
6. Westfall, J.A. Predicting past and future diameter growth for trees in the northeastern United States. *Can. J. For. Res.* **2006**, *36*, 1551–1562.
7. Russell, M.B.; Weiskittel, A.R.; Kershaw, J.A., Jr. Benchmarking and calibration of Forest Vegetation Simulator individual tree attribute predictions across the northeastern United States. *North. J. Appl. For.* **2013**, *30*, 75–84.
8. Zianis, D.; Muukkonen, P.; Mäkipää, R.; Mencuccini, M. *Biomass and Stem Volume Equations for Tree Species in Europe*; Finnish Society of Forest Science, Finnish Forest Research Institute: Helsinki, Finland, 2005; ISBN 9514019830.
9. Stage, A. *Prognosis Model for Stand Development*; U.S. Dept. of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station: Ogden, UT, USA, 1973.
10. Pretzsch, H.; Biber, P.; Ďurský, J. The single tree-based stand simulator SILVA: Construction, application and evaluation. *For. Ecol. Manag.* **2002**, *162*, 3–21.
11. Dixon, G.E. *Essential FVS: A User's Guide to the Forest Vegetation Simulator*; Internal Rep; US Department of Agriculture, Forest Service, Forest Management Service Center: Fort Collins, CO, 2002; 226 p.
12. Hann, D.W.; Hanus, M.L. *Enhanced Height-Growth-Rate Equations for Undamaged and Damaged Trees in Southwest Oregon*; Forestry Communications Group, Oregon State University: Corvallis, OR, USA, 2002.
13. Crookston, N.L.; Dixon, G.E. The forest vegetation simulator: A review of its structure, content, and applications. *Comput. Electron. Agric.* **2005**, *49*, 60–80.
14. Crecente-Campo, F.; Soares, P.; Tomé, M.; Diéguez-Aranda, U. Modelling annual individual-tree growth and mortality of Scots pine with data obtained at irregular measurement intervals and containing missing observations. *For. Ecol. Manag.* **2010**, *260*, 1965–1974.
15. Boisvenue, C.; Temesgen, H.; Marshall, P. Selecting a small tree height growth model for mixed-species stands in the southern interior of British Columbia, Canada. *For. Ecol. Manag.* **2004**, *202*, 301–312.
16. Fang, Z.; Bailey, R.L. Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. *For. Sci.* **2001**, *47*, 287–300.
17. Calegario, N.; Daniels, R.F.; Maestri, R.; Neiva, R. Modeling dominant height growth based on nonlinear mixed-effects model: A clonal Eucalyptus plantation case study. *For. Ecol. Manag.* **2005**, *204*, 11–21.
18. Temesgen, H.; Monleon, V.J.; Hann, D.W. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. *Can. J. For. Res.* **2008**, *38*, 553–565.
19. Wang, Y.; LeMay, V.M.; Baker, T.G. Modelling and prediction of dominant height and site index of Eucalyptus globulus plantations using a nonlinear mixed-effects model approach. *Can. J. For. Res.* **2007**, *37*, 1390–1403.
20. Bechtold, W.A.; Patterson, P.L. *The Enhanced Forest Inventory and Analysis Program-National Sampling Design and Estimation Procedures*; Gen. Tech. Rep. SRS-80; U.S. Department of Agriculture, Forest Service, Southern

Research Station: Asheville, NC, USA, 2005.

21. Burrill, E.; Wilson, A.; Turner, J.; Pugh, S.; Menlove, J.; Christiansen, G.; Conkling, B.; David, W. *The Forest Inventory and Analysis Database: Database Description and User Guide Version 8.0 for Phase 2*; US Department of Agriculture Forest Service: Washington, DC, USA, 2018.

22. Barrett, T.M. Optimizing efficiency of height modeling for extensive forest inventories. *Can. J. For. Res.* **2006**, *36*, 2259–2269.

23. McRoberts, R.E.; Woodall, C.W.; Lessard, V.C.; Holdaway, M.R. Diameter growth models for inventory applications. In Proceedings of the Second Forest Vegetation Simulator Conference, Fort Collins, CO, USA, 12–14 February 2002; pp. 158–163.

24. Weiskittel, A.R.; Garber, S.M.; Johnson, G.P.; Maguire, D.A.; Monserud, R.A. Annualized diameter and height growth equations for Pacific Northwest plantation-grown Douglas-fir, western hemlock, and red alder. *For. Ecol. Manag.* **2007**, *250*, 266–278.

25. Trincado, G.; VanderSchaaf, C.L.; Burkhart, H.E. Regional mixed-effects height-diameter models for loblolly pine (Pinus taeda L.) plantations. *Eur. J. For. Res.* **2007**, *126*, 253–262.

26. Gonda, H.E. Height-diameter and volume equations, growth intercept and needle length site quality indicators, and yield equations for young ponderosa pine plantations in Neuquén, Patagonia, Argentina. Ph.D. Thesis, Oregon State University, Corvallis, OR, USA, 1998.

27. Hasenauer, H.; Monserud, R.A. Biased predictions for tree height increment models developed from smoothed 'data.' *Ecol. Modell.* **1997**, *98*, 13–22.

28. Kangas, A.S. Methods for assessing uncertainty of growth and yield predictions. *Can. J. For. Res.* **1999**, *29*, 1357–1364.

29. Leak, W.B. *Relationships of Tree Age to Diameter in Old-Growth Northern Hardwoods and Spruce-Fir*; US Department of Agriculture, Forest Service, Northeastern Forest Experiment Station: Fort Collins, CO, USA, 1985.

30. António, N.; Tomé, M.; Tomé, J.; Soares, P.; Fontes, L. Effect of tree, stand, and site variables on the allometry of Eucalyptus globulus tree biomass . *Can. J. For. Res.* **2007**, *37*, 895–906.

31. Shaiek, O.; Loustau, D.; Trichet, P.; Meredieu, C.; Bachtobji, B.; Garchi, S.; Aouni, M.H.E.L. Generalized biomass equations for the main aboveground biomass components of maritime pine across contrasting environments. *Ann. For. Sci.* **2011**, *68*, 443–452.

32. Garcia, O. A stochastic differential equation model for the height growth of forest stands. *Biometrics* **1983**, *39*, 1059–1072.

33. DeBell, D.S.; Clendenen, G.W.; Harrington, C.A.; Zasada, J.C. Tree growth and stand development in short-rotation Populus plantings: 7-year results for two clones at three spacings. *Biomass Bioenergy* **1996**, *11*, 253–269.

34. Spiecker, H. Overview of Recent Growth Trends in European Forests. In *Forest Growth Responses to the Pollution Climate of the 21st Century*; Sheppard, L.J., Cape, J.N., Eds.; Springer: Dordrecht, The Netherlands, 1999; pp. 33–46, ISBN 978-94-017-1578-2.

35. Huang, S.; Titus, S.J. An individual tree height increment model for mixed white spruce–aspen stands in Alberta, Canada. *For. Ecol. Manag.* **1999**, *123*, 41–53.

36. Woudenberg, S.W.; Conkling, B.L.; O'Connell, B.M.; LaPoint, E.B.; Turner, J.A.; Waddell, K.L. *The Forest Inventory and Analysis Database: Database Description and Users Manual Version 4.0 for Phase 2*; Gen. Tech. Rep. RMRS-GTR-245; U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station: Fort Collins, CO, USA, 2010.

37. Salas, C.; Stage, A.R.; Robinson, A.P. Modeling Effects of Overstory Density and Competing Vegetation on Tree Height Growth. *Soc. Am. For.* **2008**, *54*, 107–122.

38. Vospernik, S.; Monserud, R.A.; Sterba, H. Do individual-tree growth models correctly represent height:diameter ratios of Norway spruce and Scots pine? *For. Ecol. Manag.* **2010**, *260*, 1735–1753.

39. Nunifu, T.K. Compatible diameter and height increment models for lodgepole pine, trembling aspen, and white spruce. *Can. J. For. Res.* **2009**, *39*, 180–192.

40. Richards, F.J. A flexible growth function for empirical use. *J. Exp. Bot.* **1959**, *10*, 290–301.

41. Atkinson, A.C. *Plots, Transformations and Regression; An Introduction to Graphical Methods of Diagnostic Regression Analysis*; Clarendon Press: Oxford, UK, 1985; ISBN No. 519.536 A875.

42. Hyndman, R.J.; Grunwald, G.K. Applications: Generalized Additive Modelling of Mixed Distribution Markov Models with Application to Melbourne's Rainfall. *Aust. N. Z. J. Stat.* **2000**, *42*, 145–158.

43. Hosmer, D.W., Jr.; Lemeshow, S.; Sturdivant, R.X. *Applied Logistic Regression*; John Wiley & Sons: Hoboken, NJ, USA, 2013.

44. Davidian, M. *Nonlinear Models for Repeated Measurement Data*; Routledge: Abingdon-on-Thames, UK, 2017; ISBN 0203745507.

45. Zeger, S.; Liang, L. Longitudinal data analysis using generalized linear models. *Biometrika* **1986**, *73*, 13–22.

46. McCullagh, P. *Generalized Linear Models*; CRC press: Boca Raton, FL, USA, 2019.

47. Bolker, B.M.; Brooks, M.E.; Clark, C.J.; Geange, S.W.; Poulsen, J.R.; Stevens, M.H.H.; White, J.-S.S. Generalized linear mixed models: A practical guide for ecology and evolution. *Trends Ecol. Evol.* **2009**, *24*, 127–135.

48. Stone, M. Cross-Validatory Choice and Assessment of Statistical Predictions. *J. R. Stat. Soc. Ser. B* **1974**, *36*, 111–133.

49. Duan, N. Smearing estimate: A nonparametric retransformation method. *J. Am. Stat. Assoc.* **1983**, *78*, 605–610.

50. Thomas, S.C. Asymptotic height as a predictor of growth and allometric characteristics in Malaysian rain forest trees. *Am. J. Bot.* **1996**, *83*, 556–566.

51. LeBlanc, D.C. Relationships between breast-height and whole-stem growth indices for red spruce on Whiteface Mountain, New York. *Can. J. For. Res.* **1990**, *20*, 1399–1407.

52. Kaufmann, M.R.; Ryan, M.G. Physiographic, stand, and environmental effects on individual tree growth and growth efficiency in subalpine forests. *Tree Physiol.* **1986**, *2*, 47–59.

53. Wang, J.R.; Kimmins, J.P. Height growth and competitive relationship between paper birch and Douglas-fir in coast and interior of British Columbia. *For. Ecol. Manag.* **2002**, *165*, 285–293.

54. Chen, H.; Klinka, K.; Kabzems, R. Height growth and site index models for trembling aspen (Populus tremuloides Michx.) in northern British Columbia. *For. Ecol. Manag.* **1998**, *102*, 157–165.

55. Uzoh, F.C.C.; Oliver, W.W. Individual tree height increment model for managed even-aged stands of ponderosa pine throughout the western United States using linear mixed effects models. *For. Ecol. Manag.* **2006**, *221*, 147–154.

56. Williams, M.S.; Bechtold, W.A.; LaBau, V.J. Five instruments for measuring tree height: An evaluation. *South. J. Appl. For.* **1994**, *18*, 76–82.

57. Larjavaara, M.; Muller-Landau, H.C. Measuring tree height: A quantitative comparison of two common field methods in a moist tropical forest. *Methods Ecol. Evol.* **2013**, *4*, 793–801.

58. Ryan, M.G.; Yoder, B.J. Hydraulic Limits to Tree Height and Tree Growth. *Bioscience* **1997**, *47*, 235–242.

59. Mullahy, J. Specification and testing of some modified count data models. *J. Econom.* **1986**, *33*, 341–365.

60. Zorn, C.J.W. An analytic and empirical examination of zero-inflated and hurdle Poisson specifications. *Sociol. Methods Res.* **1998**, *26*, 368–400.

61. Hanus, M.L.; Hann, D.W.; Marshall, D.D. *Predicting Height to Crown Base for Undamaged and Damaged Trees in Southwest Oregon*; Oregon State University, Forest Research Laboratory: Corvallis, OR, USA, 2000.

62. Changyong, F.; Hongyue, W.; Naiji, L.U.; Tian, C.; Hua, H.E.; Ying, L.U. Log-transformation and its implications for data analysis. *Shanghai Arch. Psychiatry* **2014**, *26*, 105.

63. Agarwal, D.K.; Gelfand, A.E.; Citron-Pousty, S. Zero-inflated models with application to spatial count data. *Environ. Ecol. Stat.* **2002**, *9*, 341–355.

64. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 26.

65. Hall, D.B. Zero-inflated Poisson and binomial regression with random effects: A case study. *Biometrics* **2000**, *56*, 1030–1039.

66. Kempes, C.P.; West, G.B.; Crowell, K.; Girvan, M. Predicting Maximum Tree Heights and Other Traits from Allometric Scaling and Resource Limitations. *PLoS ONE* **2011**, *6*, e20551.