forests

Article

A Flexible Height–Diameter Model for Tree Height Imputation on Forest Inventory Sample Plots Using Repeated Measures from the Past

Christoph Gollob *, Tim Ritter [ID], Sonja Vospernik, Clemens Wassermann and Arne Nothdurft

Department of Forest- and Soil Science, Institute of Forest Growth, University of Natural Resources and Life Sciences, Vienna (BOKU), Vienna 1180, Austria; tim.ritter@boku.ac.at (T.R.); sonja.vospernik@boku.ac.at (S.V.); clemens.wassermann@boku.ac.at (C.W.); arne.nothdurft@boku.ac.at (A.N.)
* Correspondence: christoph.gollob@boku.ac.at; Tel.: +43-147-6549-1418

check for updates

Abstract: In this study, height–diameter relations were modeled using two different mixed model types for imputation of missing heights from longitudinal data. Model Type A had a hierarchical structure of sample plot-specific and measurement occasion-specific random effects. In Model Type B, a possible temporal variance was modeled by a sample plot-specific linear time trend. Furthermore, various calibration strategies of random effects were performed on past and current data, and a combination of both. The performance of the mixed models was compared on independent data using bias and root mean square error (RMSE). The results showed that Model Type A achieved the highest precision (lowest RMSE), if sample plot-specific random effects were predicted from old data and measurement occasion-specific ones were predicted from new data. In comparison, however, Model Type B had a higher RMSE, and lower bias. Model performance was almost unaffected from the usage of past or current data for the prediction of random effects. Results revealed that a certain calibration strategy should be simultaneously applied to all random effects from the same hierarchy level. Otherwise, predictions would become imprecise and a serious bias may result. In comparison with traditional uniform height curves, the novel mixed model approach performed slightly better.

Keywords: height–diameter models; mixed models; random effects; calibration strategies

1. Introduction

The estimation of growing stock timber volume can be regarded as a major purpose of traditional forest inventories [1]. In traditional forest inventories with multiple sample plots, the total growing stock of a forest landscape is usually estimated by the Horwitz–Thompson estimator applied to single tree stem volumes together with the trees' inclusion probabilities. The direct measurement of tree volume is, however, not feasible in practice and single tree volume is usually estimated by allometric relationships or regression functions dependent on easy-to-measure auxiliary variables, such as tree height (h) and diameter at breast height (d); the reader is referred for further details to standard textbooks on forest mensuration, such as Kershaw et al. [2]. Compared with the diameter measurement, the measurement of tree height is time-consuming and has larger relative measurement errors. Tree height is therefore often measured only for a subsample per sample plot and the heights of the other trees are imputed. Tree heights may be estimated using functions solely based on d. This type of function is termed "local" [3] or "simple" [4] height–diameter curve. These equations do not include additional variables that may influence the height–diameter relation in different stands. The $h - d$ relationship, however, can vary with environmental conditions and stand structure [5,6] and $h - d$ curves have to be updated for a new survey. In addition, retaining the same sample trees for $h - d$ curve

fits on consecutive measurement occasions proved not to be optimal and can produce a serious bias [7]. A "generalized" or "regional" height diameter function estimates the specific relationship between individual tree heights and diameter using stand variables (e.g., [4,8,9]). This results in individual height–diameter relationships for each stand [5]. Variables commonly included in generalized height–diameter relationships are stand variables, such as stem density, basal area per hectare and quadratic mean diameter. In generalized height–diameter curves, tree-level attributes are often used in addition, e.g., the basal-area-of-larger-trees (BAL). The benefit of using the above-mentioned characteristics is that further valuable information is provided, but extra measurements besides d would not be required, as the latter attributes can be easily derived from the existing distribution of d-measurements per sample plot. Other approaches have been proposed using meteorological measures aggregated per growing season as additional fixed effects, resulting in a climate-sensitive parameterization of $h - d$ curves and providing predictions of site-productivity changes under a possible climate change [10]. The inclusion of additional variables generally improves the goodness of fit and extends the generality of height–diameter models [11]. However, these models still assume that sample plots with similar values of stand level predictors also have similar height–diameter curves [4].

Mixed model approaches allow sample plot specific predictions, and have been widely used for height–diameter modeling since the pioneer work by Lappi [12,13], who thereby also introduced mixed models to forestry application in general. Mixed model approaches enable the incorporation of hierarchically structured random effects to account for possible correlations of the observed heights among trees on the same sample plot [14]. Whereas the fixed model part alone provides unbiased mean curve predictions, even under unbalanced sampling designs, the random effect parameters together with the fixed effects represent more precisely the response specific to each grouping unit [15,16]. The prediction of the random effects is performed via empirical Bayesian calibration methods, providing an efficient $h - d$ curve adjustment by using only sparse data from subsamples. The calibration through a best linear unbiased predictor (BLUP) can even be conducted by means of new observations in terms of independent data, which have not been used for the model fit. The BLUP would then only require prior estimates of the variances and covariances of the random effects [15,17,18]. Thus, mixed models with BLUPs can be seen as the new standard approach in modeling height–diameter relationships.

In Central Europe, the traditional approach to impute missing heights in forest inventories using angle count sampling is "Einheitshöhenkurven" [19–26]. In the present paper, this traditional approach is henceforward termed as "uniform height curve approach". The uniform height curve (uhc) approach can be classified as a generalized height-curve approach [8]. When transformed by linearization, the rationale of uniform height curves is based on the assumption that the intercept parameter increases with age as well as site quality [6,25], whereas the slope coefficient varies only randomly. Uniform height curves therefore assume that for a given region it is sufficient to estimate the species-specific slope coefficient only once. In contrast, the intercept needs to be calibrated for each sample plot by measuring the height of one tree per species. When the uniform height curve is applied in forest inventory practice, height is usually measured on the median basal area tree per each species on a sample plot. This is because the median basal area tree can be easily determined in the field when angle count sampling is applied. For Austria, uniform height curves were developed for all species by Knieling [27] using data from the Austrian National Forest Inventory.

The Institute of Forest Growth of the University of Natural Resources and Life Sciences, Vienna (BOKU) maintains a forest inventory based on approximately 500 Bitterlich relascope sampling plots. Tree height has been measured of every sample tree from 1989 to 2014. To save labor costs associated with the fieldwork, tree height will be measured henceforth only for the median basal area tree per tree species.

The aim of this study was to develop a novel method for height prediction using not only height information from the median basal area tree per species, but also considering the complete height information from prior inventories in the past. For this purpose, a mixed model was fitted to the

$h - d$ data from the past inventories. It was further examined whether random effects on the sample plot level remain stationary or whether a time trend can be detected, which could be modeled by additional fixed effects. Furthermore, different strategies for the calibration were compared. In the different calibration strategies, only a specific subset of the random effects as predicted by means of new data and the other random effects were adopted from the past. The performance of the different candidate models was examined by means of independent data, which was not used for the model fitting. Finally, a comparison with the traditional uniform height-curve approach was conducted.

2. Materials and Methods

The BOKU Institute of Forest Growth maintains a permanently repeated forest inventory in the forest district Ofenbach located in the Austrian pre Alps, in the federal state of Lower Austria nearby the village Forchtenstein. The forest inventory was initiated by Schodterer [28] in 1989 and comprises in total 554 sample plot locations. Since then, approximately one fifth of the total sample size has been re-measured in each year, meaning that a complete follow-up of the entire sample has been conducted during a five-year interval. From 1989 to 2016, sampling plots were measured on five to six occasions. The most recent inventory cycle with the sixth measurement occasion includes so far only recordings for three years (2014, 2015, and 2016) meaning that data from 2017 and 2018 are missing before the follow-up is completed. Therefore, the sixth measurement period was discarded from our data analysis. In addition, inconsistencies with respect to the tree numbering between the first and the second measurement existed and could not be solved properly, because old data collection forms are lacking. Thus, data from the first measurement occasion were also discarded from the analysis.

Sample plots are systematically aligned in a regular grid having a mesh width of 141.4 m × 141.4 m. At each sample point, Bitterlich relascope sampling [29–31] was conducted using a basal area factor of 4 m^2 per ha. Trees were only recorded if their diameter was greater than or equal to a lower threshold of 5 cm.

Site, stand and tree variables measured at each sample point are shown in Table 1, although it should be noted that not all of the listed features were addressed on each measurement occasion during the entire period from 1989 to 2016. Furthermore, the following variables were additionally calculated additional by means of the forest to the inventory sample data: basal area, and stem number per hectare, stand density index (*SDI*), quadratic mean diameter (*QMD*) and *d*/*QMD* ratio.

Table 1. Variables assessed in the inventory.

Spatial and site variables	Growth district, sample plot coordinates, elevation, aspect, slope, soil type, hydrology
Stand variables	Stand age, vegetation type, mixture type, presence and type of regeneration, type of removal
Tree variables	Distance and angle from sample plot center, tree species, dbh, height, height to crown base, crown transparency, tree social class, stem quality, stem damage, broken top

The data were split into a model dataset used for model fitting and an extra dataset for evaluation purpose. The model dataset comprised 1491 sample plots measured in three periods (1994–1998: 468 plots; 1999–2003: 510 plots; and 2004–2008: 513 plots). The number of sample points varied between measurement occasions due to clear cuts in some forest stands. Data mining was necessary and a few outliers and missing values of the tree and stand variables were corrected. For a simple and clear overview, Tables 2 and 3 show summary statistics (number of observations, mean, min, and max) of the observed heights and the covariates used in the final models for both datasets.

Table 2. Summary statistics of the model data. *d* = diameter at breast height; *h* = height; *dm* = diameter of the median basal area tree per species; *hm* = height of the median basal area tree per species.

# of sample plots		1491	
# of height measurements		12,149	
# of height measurements/sample plot		8.15	
	Mean	**Min**	**Max**
d (cm)	31.03	5.7	99.9
h (m)	24.24	4	45
dm (cm)	31.95	5.7	79.1
hm (m)	24.81	5	43

Table 3. Summary statistics of the evaluation data. *d* = diameter at breast height; *h* = height; *dm* = diameter of the median basal area tree per species; *hm* = height of the median basal area tree per species.

|			
# of sample plots		512	
# of height measurements		4364	
# of height measurements/sample plot		8.52	
	Mean	**Min**	**Max**
d (cm)	32.41	5.5	99
h (m)	25.05	4	46
dm (cm)	33.33	5.2	86.6
hm (m)	25.78	4	46

Tables 4 and 5 contain summary statistics for tree species groups, which share a similar growth pattern in the survey area and which possess statistically not significant differences with respect to their $h - d$ ratios. The variables include the main regressor-variable *d*, the diameter of median basal area tree per species (*dm*), the height of median basal area tree per species (*hm*), and the tree height (*h*) as response variables. Figure 1 shows a smoothed scatterplot of the observed height–diameter pairs of the model dataset. Scatterplots of the different species groups can be found in Appendix A.

Table 4. Summary statistics of the model data with respect to species groups. *d* = diameter at breast height; *h* = height; *dm* = diameter of the median basal area tree; *hm* = height of the median basal area tree.

		Mean $\pm$ sd (Min–Max)			
Species Group	*n*	*d* (cm)	*h* (m)	*dm* (cm)	*hm* (m)
Spruce/Fir	5643	31.29 ± 12.84 (5.8–99.9)	24.58 ± 8.23 (5.0–44.0)	31.97 ± 11.46 (5.8–78.6)	25.01 ± 7.91 (5.0–43.0)
Larch/Scots pine	1629	34.07 ± 11.35 (6.0–79.1)	25.22 ± 6.82 (5.0–40.0)	34.67 ± 10.46 (6.9–79.1)	25.41 ± 6.75 (6.0–40.0)
Douglas fir	61	38.95 ± 19.44 (11.5–80.3)	27.05 ± 10.29 (13.0–45.0)	39.00 ± 19.06 (11.5–73.1)	27.46 ± 9.36 (15.0–42.0)
Beech/other broadleafs	4434	29.86 ± 14.87 (5.7–77.8)	23.68 ± 7.94 (4.0–42.0)	31.22 ± 13.41 (5.7–72.0)	24.58 ± 7.63 (5.0–42.0)
Oak	91	28.25 ± 10.80 (6.1–69.6)	20.26 ± 4.44 (6.0–29.0)	29.05 ± 10.75 (6.1–69.6)	20.15 ± 4.33 (6.0–29.0)
Hornbeam/Elm/Alder/Birch	291	26.07 ± 12.90 (6.0–70.5)	21.45 ± 6.67 (7.0–36.0)	26.73 ± 12.34 (6.0–70.5)	21.83 ± 6.51 (7.0–33.0)

Table 5. Summary statistics of the evaluation data with respect to species groups. *d* = diameter at breast height; *h* = height; *dm* = diameter of the median basal area tree; *hm* = height of the median basal area tree.

		Mean ± sd (Min-Max)			
Species Group	*n*	*d* **(cm)**	*h* **(m)**	*dm* **(cm)**	*hm* **(m)**
Spruce/Fir	1742	33.86 ± 13.57 (5.3–99.0)	25.99 ± 8.22 (4.0–45.0)	34.37 ± 12.06 (6.5–80.1)	26.62 ± 7.81 (4.0–43.0)
Larch/Scots pine	545	36.17 ± 11.81 (5.2–86.6)	26.73 ± 6.80 (4.0–40.0)	36.97 ± 11.05 (5.2–86.6)	26.93 ± 6.72 (4.0–40.0)
Douglas fir	26	39.70 ± 22.06 (11.5–85.3)	29.20 ± 10.06 (17.0–46.0)	41.21 ± 22.33 (21.3–79.9)	29.79 ± 9.40 (21.0–44.0)
Beech/other broadleafs	1921	30.22 ± 16.34 (5.3–86.2)	23.77 ± 8.32 (5.0–46.0)	31.55 ± 14.95 (5.3–86.2)	24.77 ± 7.90 (5.0–46.0)
Oak	35	30.10 ± 12.05 (9.8–72.0)	22.21 ± 5.28 (8.0–33.0)	31.48 ± 12.47 (9.8–72.0)	22.35 ± 5.10 (8.0–33.0)
Hornbeam/Elm/Alder/Birch	95	27.49 ± 12.90 (5.3–62.2)	23.98 ± 6.75 (8.0–38.0)	27.95 ± 11.77 (5.3–62.2)	24.48 ± 6.71 (8.0–35.0)
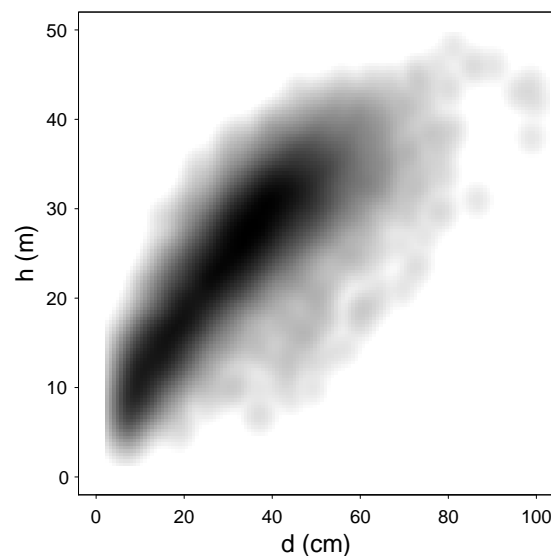
Figure 1. Smoothed scatterplot: height–diameter relationship of the model data.

2.1. Basic Model and Model Variants

Many height–diameter equations have been proposed in several studies. A proper height–diameter model possesses some important characteristics (e.g., [11]). The $h - d$ curve is monotonically increasing and has a functional inflection point as well as an asymptote. Parameter-parsimonious models that meet these requirements are the "Petterson function" (e.g., [32]) and the "Näslund function" (e.g., [33]). The Petterson function is commonly applied in German speaking countries of Central Europe and the Näslund function is applied in Nordic countries.

According to findings in [34–37] and recent suggestions by Mehtätalo et al. [4], the Näslund function was chosen in our work:

$$h_{ijk} = 1.3 + \left(\frac{d_{ijk}}{\alpha + \beta \times d_{ijk}} \right)^{\gamma} \tag{1}$$

where h_{ijk} and d_{ijk} are the total height (in m) and diameter at breast height (in cm) on the *k*-th measurement occasion of tree *j* at sample plot *i*, respectively. Parameters α, β and γ are fixed effects.

As fitting a hierarchical nonlinear mixed model often becomes challenging [18,38], the base function in Equation (1) was linearized. For parameter γ, a grid search was performed with regard to the minimum of bias and RMSE. After setting parameter γ fixed, linear mixed model theory was applied to the transform

$$y_{ijk} = \frac{d_{ijk}}{\left(h_{ijk} - 1.3\right)^{(1/\gamma)}} = \alpha + \beta \times d_{ijk} \tag{2}$$

Hereof, y_{ijk} is the transformed response.

The process of model construction was strictly based on guidelines outlined in the standard textbook of Pinheiro and Bates [39]: first, the structure of the random effects has to be defined and tested via LRTs (likelihood ratio test), and, afterwards, the fixed effects were selected guided by Wald-tests [40]. In mixed models, random effects can affect both the absolute term (intercept) and the slope. A first simple model that integrates the random effects into the basic linear function (Equation (2)) is:

$$\text{A1}: \ y_{ijk} = \beta_0 + \beta_1 \times d_{ijk} + a_i + \epsilon_{ijk} \tag{3}$$

where β_0 is the fixed intercept and β_1 is the fixed slope coefficient for d_{ijk}, and d is the diameter of tree j on sample plot i at measurement occasion k. Both terms provide an estimate for the conditional expectation of the population mean of the response variable. a_i is a random effect on the sample plot level, i.e., the random deviation on the sample plot level around the intercept β_0. ϵ_{ijk} is the residual error. That is, A1 considers the differences in the intercept between the individual sample plots. For the next step of the model construction, it is assumed that not only the intercept, but also the slope of the height curve can vary among the sample plots:

$$\text{A2}: \ y_{ijk} = \beta_0 + \left(\beta_1 + b_i^{(1)}\right) \times d_{ijk} + a_i + \epsilon_{ijk} \tag{4}$$

with $b_i^{(1)}$ being the random deviation of the coefficient for d_{ijk}. Since data comprised several measurement periods, the model should also consider a possible temporal variation of the parameters. Thus, an additional random effect is introduced, which considers the grouping with respect to the measurement occasions:

$$\text{A3}: \ y_{ijk} = \beta_0 + \left(\beta_1 + b_i^{(1)} + b_{ik}^{(1)}\right) \times d_{ijk} + a_i + a_{ik} + \epsilon_{ijk} \tag{5}$$

where a_{ik} and $b_{ik}^{(1)}$ are additional random effects for the measurement occasion k.

The R [41] package "nlme" [39] was used for fitting the models. However, nlme could not be applied for the calibration of Model A3 by means of new independent data. This is because nlme does not provide a straightforward implementation of the random effect prediction when new units for a grouping level occur. This case simply happened, as a future measurement occasion always represents a new unit k for a_{ik} and b_{ik}. Thus, a little extra algebra was necessary for the evaluation of the BLUPs. As an alternative solution to the extra programming effort associated with the calibration of Model A3, the $h - d$-model was further reformulated. Whereas it is assumed for Model A3 that the random effects a_{ik} and b_{ik} are independently distributed among the different measurement occasions, the reformulations assume that the model parameters follow a linear temporal trend. The latter assumption then no longer requires random effects on the measurement-occasion level. This approach was implemented by adding the measurement occasion m as an extra fixed effect to Model A1 (Equation (3)) and Model A2 (Equation (4)), meaning that a population-wide time trend is modeled for the parameters. The models are further enhanced by extra sample plot level random effects, which allow the temporal trend to vary among the sample plots. With these prerequisites, the models were specified as Model Type B:

$$\text{B1}: \ y_{ijk} = \beta_0 + \beta_1 \times d_{ijk} + \left(\beta_2 + b_i^{(2)}\right) \times m + a_i + \epsilon_{ijk} \tag{7}$$

$$\text{B2}: \ y_{ijk} = \beta_0 + \left(\beta_1 + b_i^{(1)}\right) \times d_{ijk} + \left(\beta_2 + b_i^{(2)}\right) \times m + a_i + \epsilon_{ijk} \tag{8}$$

According to the generalized $h - d$ curve approach, further regressor covariates were tested as fixed effects. The random effects at each level were plotted against all covariates and carefully examined for any nonlinear effect. In consequence, only the dm was transformed logarithmically. Diagnostic graphs for the other covariates did not show any relevant trends. The candidate variables were d, m (measurement occasion), logarithmic dm (diameter median basal area tree per species), hm (height median basal area tree per species), dummy categories for the different tree species groups, interaction effects between tree species groups and d, dummies for the mixture, SDI (Stand Density Index) [42], QMD (quadratic mean diameter), d/QMD ratio, basal area per hectare and stem number per hectare, elevation, slope and aspect. Based on Wald tests, the d, the measurement occasion, the logarithmic dm, the hm, the fixed offset effects for dummy-coded tree species groups, the interaction between tree species groups and d, the QMD and the d/QMD ratio proved to be as significant. However, the QMD was strongly correlated with the dm and did not lead to any improvement when the model was used for predictions. The latter also applied to the QMD/d ratio, which did not show any biological logics in connection with $h - d$ curves. Furthermore, the models were kept as simple as possible to produce reliable prediction results; and because of that the two variables were not included in the full models. The respective variables were considered with fixed effects in two full models, which differed with respect to their specification of the variance among the measurement occasions. Model B3 is analogous to Model B2 and also includes a sample plot level random time trend for the intercept parameter as well as for the slope dependent on d.

$$\begin{aligned} \text{B3}: \ y_{ijk} = \beta_0 + &\left(\beta_1 + b_i^{(1)}\right) \times d_{ijk} + \left(\beta_2 + b_i^{(2)}\right) \times m + \beta_3 \times \log(dm_{ik}) + \beta_4 \times hm_{ik} \\ &+ \beta_5 \times I\{spec = Douglas\ Fir\} + \cdots + \beta_9 \times I\{spec = Larch/Scots\ pine\} \\ &+ \beta_{10} \times d_{ijk} \times I\{spec = Douglas\ Fir\} \\ &+ \cdots + \beta_{14} \times d_{ijk} \times I\{spec = Larch/Scots\ pine\} + a_i + \epsilon_{ijk} \end{aligned} \tag{9}$$

Model A4 is analogous to A3 and incorporates the measurement occasion itself as a random effect.

$$\begin{aligned} \text{A4}: \ y_{ijk} = \beta_0 + &\left(\beta_1 + b_i^{(1)} + b_{ik}^{(1)}\right) \times d_{ijk} + \beta_3 \times \log(dm_{ik}) + \beta_4 \times hm_{ik} \\ &+ \beta_5 \times I\{spec = Douglas\ fir\} + \cdots + \beta_9 \times I\{spec = Larch/Scots\ pine\} \\ &+ \beta_{10} \times d_{ijk} \times I\{spec = Douglas\ fir\} \\ &+ \cdots + \beta_{14} \times d_{ijk} \times I\{spec = Larch/Scots\ pine\} + a_i + a_{ik} + \epsilon_{ijk} \end{aligned} \tag{10}$$

y_{ijk} is the transformed tree height and $\beta_0, \ldots, \beta_{14}$ are fixed effects. d_{ijk} is the diameter at breast height, m is the measurement occasion, $\log(dm_{ik})$ is the logarithmic diameter of the median basal area tree per tree species, hm_{ik} is the height of the median basal area tree per tree species and $I\{spec = \cdots\}$ is the dummy categorical variable for the respective tree species. The order of tree species groups in Equations (9) and (10) is: Douglas fir, Oak, Spruce/Fir/other conif., Hornbeam/Elm/Alder/Birch and Larch/Scots pine. a_i is a random intercept effect on the sample plot level. In Model B3, $b_i^{(1)}$ and $b_i^{(2)}$ are sample plot level random effects; the former acts on the slope dependent on d_{ijk} and the latter reflects the sample plot specific temporal trend of the model intercept. In Model A4, a_{ik} is a random intercept-effect and $b_{ik}^{(1)}$ is a random slope-effect on the measurement-occasion level. ϵ_{ijk} is the residual error.

2.2. Matrix Formulation

The linear mixed model can be written in matrix notation [15] as

$$y = X\beta + Zb + \epsilon \tag{11}$$

where y is a column vector of the response values having length n (number of observations), β is a vector of the total p fixed effects, b is a vector of the total q Gaussian random effects, and ϵ is an independent and identically distributed residual Gaussian error. The model matrix X has dimension $n \times p$ and contains the values of the regressor covariables and columns of one-values for fixed effect intercept parameters as well as for dummy-categorical regressor variables. Matrix Z is a $n \times q$ design matrix containing the covariate values associated with the respective random slope-effects and one-values for the random intercept-effects. The summand $X\beta$ is the fixed part of the model and $Zb + \epsilon$ is the random part. It is assumed that the vector of random effects b is normally distributed with $b \sim N(0, D)$ with D denoting the $q \times q$ variance–covariance matrix of the random effects. It is assumed that ϵ is normally distributed with $\epsilon \sim N(0, R)$ with R being the corresponding $n \times n$ variance matrix containing $var(\epsilon)$ on the diagonal. Because residual errors of different trees are assumed to be uncorrelated, the off-diagonal elements of R have zero values.

2.3. Estimation of Fixed Effects and Model Calibration via BLUP

The fixed effects and the variances and covariances of the random effects were estimated using restricted maximum likelihood methods (REML) [38,39]. Hypothesis testing for the fixed effects is performed using Wald [40] tests sequentially for each of the parameters. The null hypothesis is $H_0 : \beta = 0$ and the alternative is $H_1 : \beta \neq 0$.

The linear mixed model approach allows for model calibration by means of prior information [15,17]. Different kinds of information of the predictors in the fixed part combined with additional local height measurements can be used to improve the prediction accuracy. The random effects in b were predicted via BLUPs:

$$\hat{b} = \hat{D} Z' \left(Z \hat{D} Z' + \hat{R} \right)^{-1} (y - X\hat{\beta}) \tag{12}$$

$\hat{D}$ contains the REML-estimates of the covariances, $\hat{R}$ the estimated residual variance and $\hat{\beta}$ the estimates of the fixed effects. The BLUP can be either applied to response observations, which have been used for the model fitting or to new observations from independent data. In the latter case, y is substituted by y^* denoting any prior information from existing $h - d$ measurements. An example of matrix algebra can be found in Appendix B.

The calibration of the mixed models was done by means of old and new data. The old data represent the model dataset and the evaluation dataset was formed by new independent measurements, which were not used for model fitting. For a comparison under fair conditions, the calibration by means of new data and the adjustment of the uhc were performed exclusively based on tree height information from the median basal area trees per species. For Model B3 in total eight calibration strategies exist, among them two extremes: one calibration variant in which all random effects were adopted from predictions by means of old model data and the other extreme for which all random effects were "re-predicted" by means of the new independent evaluation data. For Model A4, which has only sample plot level random effects, in total four calibration strategies exist. Likewise, in one extreme variant the random effects predictions were adopted from the model fitting data and in the other extreme they were entirely predicted by means of the calibration data. It is noted that random effects at the measurement occasion level (Model B) were always re-predicted by means of the new calibration data.

2.4. The Traditional Uniform Height Curve Method

The principle of the uniform height curve approach is explained by Knieling [27] who proposed a height curve type according to Pollanschütz [43]:

$$h = e^{a + \frac{b}{d}} + 1.3 \tag{13}$$

where h is the tree height and a and b are regression coefficients. In the study of Knieling [27], the permanent samples of the Austrian Forest Inventory 1981/1985 formed the basis for the calculation of the uhc. The coefficients a and b were separately estimated for each forest inventory sample plot and per each tree species. As expected, the sample plot-wise estimates of parameter a had a relatively large variance and estimates of b were almost constant. The rationale is that parameter a represents the asymptote of the $h - d$-curve that generally increases with stand age and site quality [6,25].

For the practical application, this means that b can be kept fixed depending on tree species, and parameter a can be estimated for each sample plot by the transform

$$a = \log(hm - 1.3) - \frac{b}{dm} \tag{14}$$

using dm and hm.

2.5. Model Performance and Evaluation

The goodness of model fits was assessed by means of the AIC (Akaike information criterion)

$$\text{AIC} = -2 \times \log(L) + 2 \times k \tag{15}$$

in which the negative logarithmic likelihood (L) is penalized by the number of k parameters. As inference was based on a restricted likelihood, only model candidates could be compared having the same structure in the fixed model part. Thus, models that differ in their specification of random effects were compared using a likelihood-ratio test (LRT)

$$\text{LR} = 2 \times [\log(L_F) - \log(L_R)] \tag{16}$$

with L_F being the likelihood of the full model and L_R the likelihood of a restricted model. Compared to the full model, a restricted model has at least one random effects set to zero. The LR test statistic is χ^2 distributed with degrees of freedom corresponding with the number of random effects set to zero. As stated above, the selection of fixed effects among the candidates of the regressor covariates was conducted using Wald tests. In addition, the coefficient of determination was calculated based on the model dataset twice: (i) by means of mean curve predictions (R_m^2); and (ii) for prognoses with random effect calibrations (R_c^2).

$$R^2 = \frac{var\left(\widehat{h_{ijk}}\right)}{\left(var\left(h_{ijk}\right) + var(\epsilon)\right)} \tag{17}$$

By Equation (2), the predicted tree heights of the different model variants were calculated. The evaluation of the mixed models was done by comparing the observed (h_{ijk}) and predicted ($\hat{h}_{ijk}$) heights from the evaluation dataset, with n trees, through a statistical analysis of the residuals. The performance of the models was examined by means of the precision and accuracy of the predictions for independent data from the 5th measurement occasion that was not used for the model fitting. Therefore, the RMSE

$$\text{RMSE} = \sqrt{\frac{\sum\left(h_{ijk} - \widehat{h_{ijk}}\right)^2}{n}} \tag{18}$$

as well as the bias

$$\text{BIAS} = \frac{\sum\left(h_{ijk} - \widehat{h_{ijk}}\right)}{n} \tag{19}$$

were evaluated. In addition, it was assessed, using diagnostic plots for the residuals against predictions and for the predictions against observations, whether the models behaved well or whether a bias or some signs of heteroscedasticity occurred.

3. Results

The grid search revealed that an optimum value of 3 existed for the parameter γ of the Näslund function (Equation (1)), and, accordingly, γ was kept fixed at that optimum. LRTs showed that Model A3 (Equation (5)) was superior to A2 (Equation (4)) and also to A1 (Equation (3)) with respect to the specification of the random effects structure (Table 6). This means that it was appropriate to place random effects on the intercept as well as on the slope that acted on both levels, the sample plot level and the measurement-occasion level. The full Model A4 (Equation (10)) for the A-model-type, with random effects on the measurement-occasion level, additionally includes significant fixed effects for d, logarithmic dm, hm and for dummy-coded tree species groups. When the random parameters on the measurement-occasion level (Model A3 of the A-type) were substituted by a temporal trend in the models belonging to the B-type, Model B2 (Equation (8)) was superior to B1 (Equation (7)), meaning that the inclusion of random effects for both regressor covariates, d and measurement occasion, was advantageous. The full Model B3 (Equation (9)) for the B-model-type includes significant fixed effects for d, measurement occasion, logarithmic dm, hm and dummy-coded for tree species groups.

Table 6. Model types. $\beta_0 \ldots \beta_{10}$ are fixed effects; a_i, $b_i^{(1)}$, $b_i^{(2)}$, a_{ik}, $b_{ik}^{(1)}$ are random effects; AIC is the Akaike information criterion; LRT is the likelihood ratio test with the χ^2 distributed LR (likelihood ratio) and the p-value; R_m^2 is the marginal coefficient of determination; and R_c^2 is the conditional coefficient of determination. Note that, in contrast to the covariance, the variance is reported as standard deviation squared.

	A1	A2	A3	B1	B2	B3	A4
β_0	1.629	1.578	1.506	1.881	1.866	1.111	0.989
β_1	0.295	0.303	0.307	0.297	0.308	0.313	0.312
β_2				−0.102	−0.124	−0.043	
β_3						0.773	0.824
β_4						−0.095	−0.103
β_5						0.542	0.499
β_6						−0.313	−0.299
β_7						0.262	0.268
β_8						−0.251	−0.246
β_9						0.145	0.155
β_{10}						−0.020	−0.019
β_{11}						0.015	0.014
β_{12}						−0.009	−0.010
β_{13}						0.012	0.012
β_{14}						−0.006	−0.006
$var(a_i)$	0.618^2	0.598^2	0.606^2	0.640^2	0.537^2	0.439^2	0.533^2
$var\left(b_i^{(1)}\right)$		0.029^2	0.031^2		0.032^2	0.025^2	0.023^2
$var\left(b_i^{(2)}\right)$				0.080^2	0.123^2	0.070^2	
$cov\left(a_i, b_i^{(1)}\right)$		−0.0101	−0.0107		−0.0038	−0.0085	−0.0111
$cov\left(a_i, b_i^{(2)}\right)$				−0.0095	−0.0003	0.0087	
$cov\left(b_i^{(1)}, b_i^{(2)}\right)$					−0.0027	−0.0010	
$var(a_{ik})$			0.221^2				0.108^2
$var\left(b_{ik}^{(1)}\right)$			0.005^2				0.002^2
$cov\left(a_{ik}, b_{ik}^{(1)}\right)$			−0.0007				−0.0002
AIC	16,465.72	13,092.16	12,492.45	16,014.60	11,914.33	8419.16	8555.86
LRT		A1 vs. A2	A2 vs. A3		B1 vs. B2		
LR		3377.55	605.72		4106.27		
p		<0.0001	<0.0001		<0.0001		
R_m^2	0.969	0.967	0.966	0.969	0.965	0.986	0.987
R_c^2	0.987	0.992	0.993	0.988	0.993	0.994	0.994

The performance of both full Models B3 and A4 was compared under different calibration strategies using prior observations from the independent evaluation data. In addition, a comparison was made with the traditional uhc approach (Table 7). The calibration of the mixed models was done by means of old and new data. Note that new data calibration and the localization of the uhc were performed exclusively based on median basal area trees per species to get a fair comparison. Systematically, we started with the prediction of all random effects from new data. For Model B3, the lowest RMSE of 2.095 and the lowest bias of −0.059 were achieved with Calibration Strategy 8 when the random effects were predicted by means of the model data and were assumed to be also valid for predictions of the new independent data. In this calibration strategy, a calibration via BLUPs using new data was not necessary and the prior predictions of random effects by means of the model data were likewise assumed to be valid for future measurement occasions. Thus, in this particular calibration variant, the random effects were entirely predicted using old data, and these predictions were henceforth plugged into Equation (11) to derive response prognoses for new covariate data in X representing the evaluation data. However, Calibration Strategy 1 in which the complete set of random effects was predicted by means of calibrations to the new data performed only slightly worse and had a RMSE of 2.148 and a bias of 0.071.

Table 7. RMSE and bias of different calibration strategies compared with the uhc. a_i, $b_i^{(1)}$, $b_i^{(2)}$, a_{ik}, $b_{ik}^{(1)}$ are random effects; B3 and A4 are full models of different model types; and uhc is the uniform height curve.

Model	Calibration Strategy	Predicted Random Effects Based on		RMSE	Bias
		Old Data	New Data		
B3	1		a_i, $b_i^{(1)}$, $b_i^{(2)}$	2.148	0.071
	2	a_i	$b_i^{(1)}$, $b_i^{(2)}$	3.536	0.155
	3	$b_i^{(1)}$	a_i, $b_i^{(2)}$	12.384	−4.675
	4	$b_i^{(2)}$	a_i, $b_i^{(1)}$	2.635	−0.059
	5	a_i, $b_i^{(1)}$	$b_i^{(2)}$	5.032	−1.974
	6	a_i, $b_i^{(2)}$	$b_i^{(1)}$	5.047	0.007
	7	$b_i^{(1)}$, $b_i^{(2)}$	a_i	8.988	−3.743
	8	a_i, $b_i^{(1)}$, $b_i^{(2)}$		2.095	−0.059
A4	1		a_i, a_{ik}, $b_i^{(1)}$, $b_{ik}^{(1)}$	2.178	0.214
	2	a_i	$b_i^{(1)}$, a_{ik}, $b_{ik}^{(1)}$	4.571	0.325
	3	$b_i^{(1)}$	a_i, a_{ik}, $b_{ik}^{(1)}$	10.714	−3.932
	4	a_i, $b_i^{(1)}$	a_{ik}, $b_{ik}^{(1)}$	2.081	0.332
uhc	1			2.167	−0.116

Calibration Strategies 3 and 7, in which $b_i^{(1)}$ or $b_i^{(1)}$ and $b_i^{(2)}$ were adopted from calibrations to the old data, had a poor performance; in the former case, RMSE was 12.384 and bias was −4.675, and, in the latter case, RMSE and bias were 8.988 and −3.743, respectively. Calibration Strategies 2 and 4 with only a_i or $b_i^{(2)}$ predicted using data from the past had medium performance. Strategies 5 and 6 with $b_i^{(2)}$ or $b_i^{(1)}$ predicted using new data had a relatively high RMSE of 5.032 and 5.047, respectively. The former also had a high bias with −1.974. The latter has a small bias with 0.007.

Model A4 had the smallest RMSE of 2.081, when the sample plot level random effects were adopted from calibrations to the old data and when the random effects on the measurement-occasion level were predicted by means of the new data (Calibration Strategy 4); however, bias was 0.332. The smallest bias of 0.214 was achieved with the A-model type, when the complete set of random effects was predicted by means of the new data (Calibration Strategy 1). However, RMSE was slightly

higher with 2.178 compared to the Calibration Strategy 4, which used sample plot level random effects from calibrations to the past data. The procedure, which used only $b_i^{(1)}$ from past calibrations, had the worst performance among all calibration strategies with Model A4, and using only a_i from calibrations by means of the old data yielded a medium prediction performance. Please note that, with Model A4, the random effects a_{ik} and $b_{ik}^{(1)}$, which consider the measuring occasion, always need to be estimated from new data. Therefore, this approach had fewer variants.

In summary, Model B3 yielded smaller biases than Model A4, when the random effects were entirely predicted either by means of the old data or completely by the new data. In terms of RMSE, Model A4 had the best performance, especially when sample plot level random effects were predicted by means of the old data and when the random effects on the measurement-occasion level were predicted by means of the new data. However, the absolute bias of the best performing calibration variant with Model A4 was higher than with the best performing Model B3 variant.

The traditional uhc approach had a RMSE of 2.167 and a bias of −0.116. Thus, in terms of RMSE and bias, uhc performed only slightly worse than the Model B3 Calibration Strategy 8 with random effects predicted by means of the old data. In terms of RMSE, the uhc approach was slightly inferior compared with Model A4 and random effects for the measurement occasion predicted by means of the new data (Calibration Strategy 4). In terms of bias, uhc was superior.

Graphical diagnostic checks were conducted for Model B3 in the Calibration Strategy 8, Model A4 in Calibration Strategy 4 and the uhc approach. Figures 2 and 3 contain all predicted heights for the evaluation dataset. The smoothed scatterplot of predicted versus observed heights (Figure 2) of the three models showed that the tuples were independently and closely distributed around the reference line having zero intercept and slope one. In addition, the smoothed scatterplots of the residuals versus predictions (Figure 3) showed for both mixed models and also the uhc, that residuals behaved well, and evidence was found neither for a trend nor for heteroscedasticity. The partially accurate estimates (residue = 0) of the uhc can be explained by the localization based on the height and diameter of the median basal area tree per species. Due to the high density in the residue = 0, deviations from it have a much lower density, which results in a lighter gray value compared to the mixed models. Although the mixed model also contains the diameter of the median basal area tree per species as a covariate, this is in addition to some other covariables, which usually leave a residual error. In the mixed model approach, the height curve is not forced to exactly pass the tuple consisting of the diameter and height of the median tree, as is the case with the traditional uhc method. The graphical diagnostic plots of Figures 2 and 3 can be seen for the different tree species in Appendix C. Figure 4 contains the Model B3 and Model A4 mean curves for each tree species group.
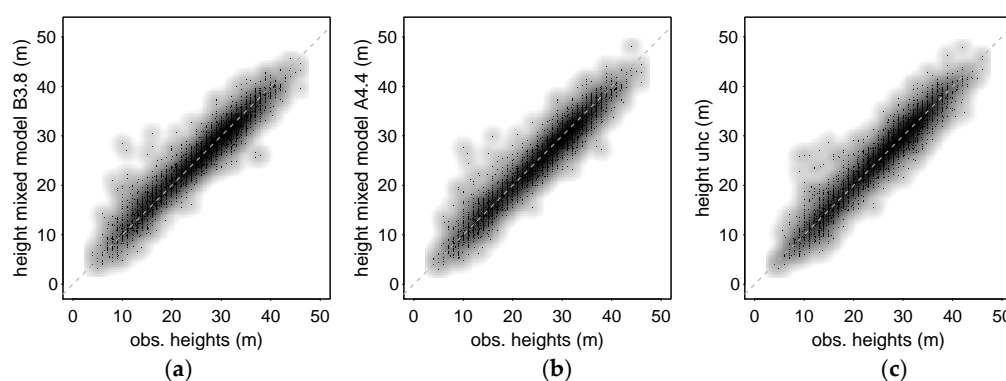
Figure 2. Predicted heights against observed heights. Smoothed densities of the corresponding scatterplots derived through two-dimensional kernel density estimates. The light-gray dashed line is the reference line with zero intercept and slope one. (**a**) Model B3, Calibration Strategy 8; (**b**) Model A4, Calibration Strategy 4; and (**c**) uhc model.
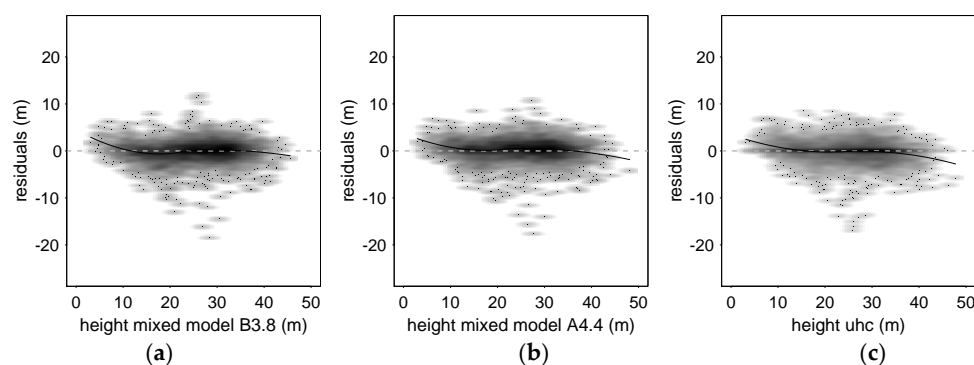
Figure 3. Residuals against predicted heights. Smoothed densities of the corresponding scatterplots derived through two-dimensional kernel density estimates. The light-gray dashed line is zero-reference; the black solid line is polynomial regression fit. (**a**) Model B3, Calibration Strategy 8; (**b**) Model A4, Calibration Strategy 4; and (**c**) uhc model.
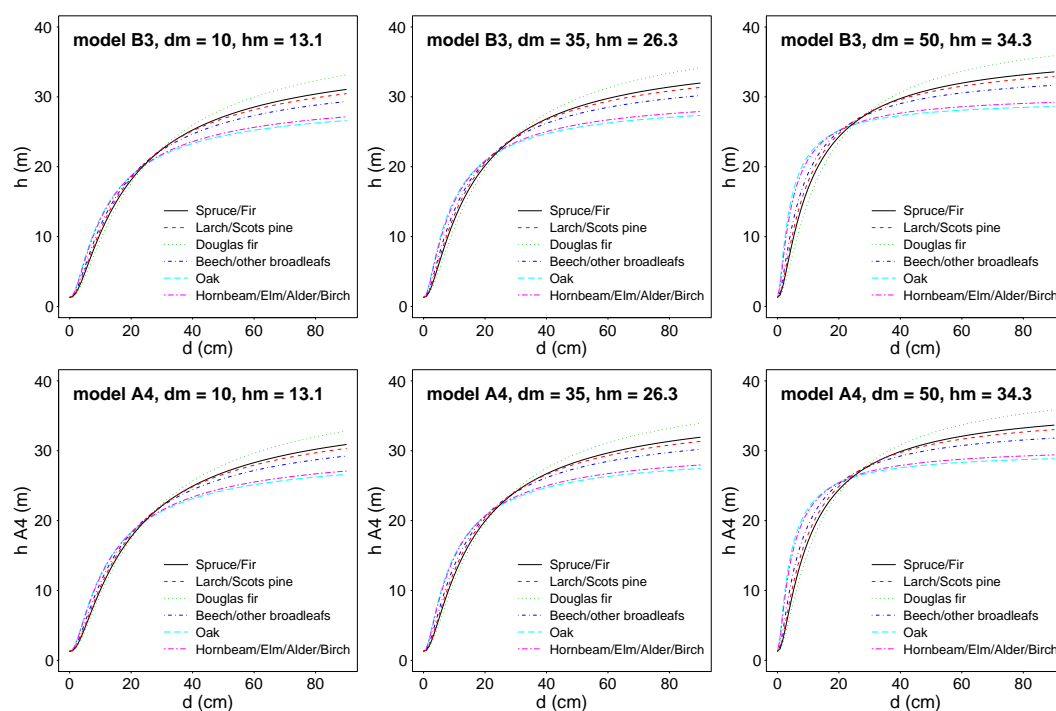
Figure 4. Mean curves separately for each tree species group (Model B3 and Model A4).

4. Discussion

4.1. Model Formulation and Calibration Strategy

The present paper demonstrates two different approaches to account for possible temporal changes of the sample plot level variance of $h - d$-curves. The first approach uses additional random effects at the level of the measurement occasion (Model A4), and the second includes a fixed time trend which can vary randomly among the sample plots (Model B3). In the context of the Näslund function (Equation (1)), Schmidt [44] wrote that it provides numerical advantages when specifying complex additive and mixed regression models as linear models. Using the linearized form of Näslund's function may introduce a (small) transformation bias into the model. However, the alternative would be to use a nonlinear mixed model and fit a true nonlinear curve type to avoid a transformation bias (e.g., [4,8,45]). The application of a true nonlinear curve type would change statistical inference

completely and would require imprecise first order Taylor expansions for the model fitting as well as for the calibration (random effect prediction). Thus, model fitting as well as calibration would easily become instable and may produce illogical results [18]. The rationale of choosing the Näslund function was that, besides its well-known proper statistical behavior, this curve type can be linearized and produces highly stable results even when applied with a large set of possible covariates. To linearize the Näslund function, the parameter γ was determined via a grid search. If optimized with respect to a small bias, Model B3, Strategy 8 would yield $\gamma = 2.6$. The smallest RMSE results with an exponent of 3. Model A4 Strategy 4 achieves the lowest bias with $\gamma = 1.5$ and the smallest RMSE with $\gamma = 3.2$. Kramer et al. [46] and Schmidt [47] suggested fixing gamma at a value of 3 to linearize the function. Based on the suggestions found in the literature [46,47] and with the prerequisite of a small RMSE, even under a negligible bias, the parameter was fixed at a value of 3. For comparison with nonlinear models, Temesgen et al. [8] obtained an RMSE between 1.4 and 4.6 and a bias between 0 and 0.1. Compared to our results (Table 7), this means that with linear mixed models a similar RMSE (especially for Model B3 Calibration Strategy 8 and Model A4 Calibration Strategy 4) but a slightly higher bias (especially for Model A4 Calibration Strategy 4) can be achieved. The latter probably results from the linearization, but still has a negligible small amount. However, as seen in Table 7, model performance depends much more on the calibration strategy.

The different calibration strategies were evaluated using straightforward matrix algebra, shown in Equation (12). Hence, under certain circumstances, it was feasible to use random effect predictions that stem from calibrations by means of different data sources. However, such a mixing of predicted random effects only worked well if the random effects from a certain model hierarchy were thoroughly calibrated for a single data source. Contrarily, it was not feasible to mix random effects predictions from different calibrations on a specific model hierarchy, as this resulted in a relatively high RMSE and bias. The rationale is that covariances between the random effects were not properly addressed in the latter case.

This means that in the case of Model B3 (time trend), either all random effects (Calibration Strategy 1) must be recalibrated by means of new data, or all must be calibrated based on old data (Calibration Strategy 8), to achieve a good model performance. With Model A4 (with random effects on measurement occasion), either all random effects (Calibration Strategy 1) or only those at the level of measurement occasion (Calibration Strategy 4) should be recalibrated by means of new data. In the latter case, sample plot level random effects would stem from old data. Furthermore, it turned out that a recalibration based on current observations is actually not necessary with Model B3, but it works surprisingly well, despite the very small number of observations that only contained d and h of the median basal area tree per species of each sample plot. The latter also applies to Model A4, where a recalibration or partial recalibration is always necessary, i.e., either the complete set of random effects was predicted via BLUPs or only a subset was predicted. Based upon these findings, the Models B3 Calibration Strategy 8 and A4 Calibration Strategy 4 can be applied as appropriate models for height imputations in forest inventory practice. Both models showed similar performance with respect to RMSE, whereas Model A4 Calibration Strategy 4 has a slightly higher bias. A benefit of Model B3 Calibration Strategy 8 is that predictions can be derived using standard routines that are implemented in the R package "nlme" [39]; this is because random effects are predicted based on old data. In particular, as the Model A4 Calibration Strategy 4 4 is calibrated based on old and new data, a little extra programming was necessary.

To account for the temporal variance of the model parameters, the measurement occasion was considered in Model B3 as a population-wide fixed effect as well as by using sample plot level random effects. It is crucial to note that with such a formulation, the temporal trend is linearly extrapolated from the model data into the future. To overcome possible limitations associated with the latter simplification, further trials were made exploring the fixed interaction effects of the measurement occasion and the mean stand height hm together with additional sample plot specific random effects. The advantage of this approach would be that hm could account for other unknown effects on tree

height development that could possibly result from different environmental or climatic conditions. However, the accuracy and precision of the latter model construction was not satisfactory.

In addition to the random effects on intercept (Model B3 and A4), d (Model B3 and A4) and measurement occasion (Model B3), additional random effects were also tested for dm and hm, but did result in an enhanced prediction performance. In some existing mixed $h - d$ models (e.g., [15]), additional tree-level random effects are applied that account for possible correlations due to multiple repeated measurements of a single tree. Mehtätalo [17] also tested the implementation of random effects at the tree level, but this approach was discarded because estimation was too computer expensive. Nevertheless, neglecting possible tree-level random effects should not significantly affect the prediction performance of the models, as long as the number of repeated measurements per tree is small compared to the number of trees in modeling data. The experiences made in the present study were in conformance with those presented in Mehtätalo [17], whereby no random effect was placed at the tree level.

4.2. Covariates

The final mixed generalized height–diameter models included fixed effects for dummy categorical tree-species groups, for interaction effects between tree species groups and d and for the diameter and height of the median basal area tree per species. Knieling [27] also showed, in estimating the coefficients for the uhc, that tree-specific differences should be taken into account. Therefore, it was also tested whether the slope of the $h - d$ curve varies among the different tree species groups [27]. However, hypothesis tests revealed that a significant tree species effect only exists for the model intercept.

Following Mehtätalo [17], the diameter and height of the median basal area tree were used to describe mean tree size. Both measures can be easily derived from angle-count sample plots [26,43]. Alternatively, Schmidt [36] and Calama et al. [48] integrated age rather than average tree size as a predictor in the model. Calama et al. [48] also showed that the inclusion of age in the model in addition to mean tree size did not lead to an improvement. The inclusion of age has rather a higher relevance in conjunction with growth models, but is actually necessary for the modeling of $h - d$ relationships in a forest inventory context. In addition, the application of stand age is generally restricted to even-aged stands due to the limited explanatory power of age in uneven-aged stands (e.g., [49] and references therein).

As part of the modeling process, the site variables elevation, slope and aspect were also tested as fixed effects. None of these parameters could describe a significant correlation with tree height. Köhler [50] investigated the relationships between site factors and height growth of Norway spruce. In contrast to the findings of our work, it was found that the factors sunshine class (including terrain and slope), elevation and relative humidity proved to be relevant regressor covariates for modeling stand height. Similarly, in [51], it has been shown for the subalpine zone that stand height is decreased by 13% per an increase in elevation of 70 m. However, site elevation in the present study region of the BOKU training forest has only a small range from 350 m to 725 m, overall representing lower mountainous sites. Accordingly, on such sub-montane sites in Austria, only little differences can be expected with respect to height growth [52].

The tested SDI (stand density index) showed no significant influence on the shape of the height curve. Inclusion of SDI in the models also resulted in a worse RMSE and bias and was therefore not considered as a fixed effect in the full models. Some authors (e.g., [53]) demonstrated that stand density is a highly relevant factor influencing the $h - d$ relationship in a forest stand. However, in other studies (e.g., [16]), less influence is attributed to stand density. In this study, both the diameter and the height of the median basal area tree per species are included, which represent the average $h - d$ ratio of the stand. The $h - d$ ratio is highly correlated with stand density in that higher ratios generally occur in denser stands (e.g., [54,55]). The rationale behind this phenomenon is a well-known individual resource reallocation strategy to maintain survival. If high competition is present, trees react with reduced diameter growth rates, whereas height growth remains nearly unchanged. This may also

explain why additional measures for tree competition or density (basal area and stem number) likewise proved to be not significant in the model, as long as the $h - d$ ratio is included. Due to the included $h - d$ ratio, the QMD and the d/QMD ratio also show no improved height predictions.

Based on the assumption that trees in mixed stands grow differently than in pure stands, a dummy variable for the mixture-type was tested during the model construction process. Sharma et al. [35] showed that the $h - d$ relationship in mixed stands has a higher variance than in pure stands. This is because mixed stands often form multiple layers. However, it was found in the present study that species mixture does not affect tree height.

5. Conclusions

When applied to new independent data, the novel mixed model $h - d$-curve approach provides more precise height predictions than the traditional uhc method. However, the relatively simple uhc method achieved surprisingly good results and has only a slightly higher bias and RMSE. Consideration of repeated measurements in mixed models can be implemented either via additional random effects for the measurement occasion (Model A4) or via a co-modeling of the temporal trend (Model B3) of the relevant model parameters.

Results gave evidence that a recalibration of a mixed $h - d$ model is not strictly necessary, when a temporal trend model (Model B3) is applied for predictions at new measurement occasions. For Model A4 and B3, it is appropriate to use random effect predictions obtained from calibrations by means of measurements in the past. However, this procedure must be consequently applied for all random effects on a specific model hierarchy. Otherwise, predictions would become imprecise and a serious bias may result.

As fixed covariables, the diameter at breast height, the logarithmic diameter, and the height of the median basal area tree per species, as well as the measuring period are used in the created model. Regarding the structure of the random effects, the best results could be achieved with sample plot specific effects on the intercept (Model B3 and A4), diameter at breast height (Model B3 and A4) and measurement period (Model B3).
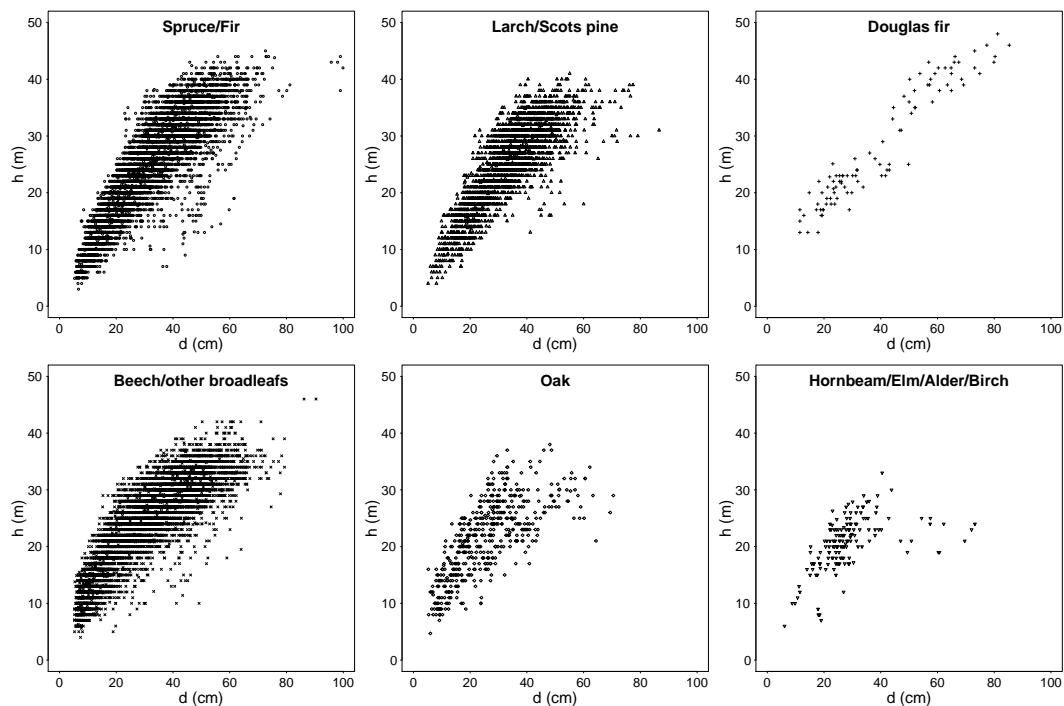
Author Contributions: A.N. and C.G. designed the study; C.G. and C.W. collected the field data of the current inventory; C.W. and C.G. processed the data; C.G. and A.N. performed the model fitting; S.V. and T.R. supported data analysis; and C.G., A.N., S.V., T.R. and C.W. wrote the paper.

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

Appendix A

Figure A1. Smoothed scatterplots of the observed height–diameter pairs for the different tree species of the model dataset.

Appendix B

Matrix algebra for example calibrations of Model B3 on a sample plot in model data via BLUP (Equation (12)):

Table A1. Model data sample plot 12. y is the transformed tree height; m —*Species group* are the covariates of the model.

y	m	d	hm	$log(dm)$	*Species Group*
7.91	2	21.0	16	3.13	Spruce/Fir
9.31	2	22.8	16	3.13	Spruce/Fir
8.27	2	20.7	16	3.13	Spruce/Fir
10.85	2	29.3	16	3.13	Spruce/Fir
8.51	3	21.3	19	3.25	Spruce/Fir
9.77	4	26.4	21	3.27	Spruce/Fir
9.03	4	24.8	21	3.27	Spruce/Fir
9.44	4	24.6	21	3.27	Spruce/Fir
11.26	2	30.4	16	3.13	Spruce/Fir
12.40	3	35.1	19	3.25	Spruce/Fir
8.60	3	23.6	19	3.25	Spruce/Fir
12.08	3	34.2	19	3.25	Spruce/Fir
8.81	3	23.8	19	3.25	Spruce/Fir
12.45	4	39.4	21	3.27	Spruce/Fir
13.30	4	38.2	21	3.27	Spruce/Fir
8.69	4	24.6	21	3.27	Spruce/Fir
11.63	4	33.4	21	3.27	Spruce/Fir
9.86	3	25.7	19	3.25	Spruce/Fir
14.37	3	45.0	32	3.81	Beech/other broadleafs
13.59	2	41.6	30	3.73	Beech/other broadleafs
15.55	4	48.7	32	3.89	Beech/other broadleafs

$$\hat{\beta}_{15\times1} = \begin{pmatrix} 1.111 \\ 0.313 \\ 0.542 \\ -0.313 \\ 0.262 \\ -0.251 \\ 0.145 \\ 0.773 \\ -0.095 \\ -0.043 \\ -0.020 \\ 0.015 \\ -0.009 \\ 0.012 \\ -0.006 \end{pmatrix} \quad X_{o\ 21\times15} = \begin{pmatrix} 1 & 21.0 & 0 & 0 & 1 & 0 & 0 & 3.13 & 16 & 2 & 0 & 0 & 21.0 & 0 & 0 \\ 1 & 22.8 & 0 & 0 & 1 & 0 & 0 & 3.13 & 16 & 2 & 0 & 0 & 22.8 & 0 & 0 \\ 1 & 20.7 & 0 & 0 & 1 & 0 & 0 & 3.13 & 16 & 2 & 0 & 0 & 20.7 & 0 & 0 \\ 1 & 29.3 & 0 & 0 & 1 & 0 & 0 & 3.13 & 16 & 2 & 0 & 0 & 29.3 & 0 & 0 \\ 1 & 21.3 & 0 & 0 & 1 & 0 & 0 & 3.25 & 19 & 3 & 0 & 0 & 21.3 & 0 & 0 \\ 1 & 26.4 & 0 & 0 & 1 & 0 & 0 & 3.27 & 21 & 4 & 0 & 0 & 26.4 & 0 & 0 \\ 1 & 24.8 & 0 & 0 & 1 & 0 & 0 & 3.27 & 21 & 4 & 0 & 0 & 24.8 & 0 & 0 \\ 1 & 24.6 & 0 & 0 & 1 & 0 & 0 & 3.27 & 21 & 4 & 0 & 0 & 24.6 & 0 & 0 \\ 1 & 30.4 & 0 & 0 & 1 & 0 & 0 & 3.13 & 16 & 2 & 0 & 0 & 30.4 & 0 & 0 \\ 1 & 35.1 & 0 & 0 & 1 & 0 & 0 & 3.25 & 19 & 3 & 0 & 0 & 35.1 & 0 & 0 \\ 1 & 23.6 & 0 & 0 & 1 & 0 & 0 & 3.25 & 19 & 3 & 0 & 0 & 23.6 & 0 & 0 \\ 1 & 34.2 & 0 & 0 & 1 & 0 & 0 & 3.25 & 19 & 3 & 0 & 0 & 34.2 & 0 & 0 \\ 1 & 23.8 & 0 & 0 & 1 & 0 & 0 & 3.25 & 19 & 3 & 0 & 0 & 23.8 & 0 & 0 \\ 1 & 39.4 & 0 & 0 & 1 & 0 & 0 & 3.27 & 21 & 4 & 0 & 0 & 39.4 & 0 & 0 \\ 1 & 28.2 & 0 & 0 & 1 & 0 & 0 & 3.27 & 21 & 4 & 0 & 0 & 38.2 & 0 & 0 \\ 1 & 24.6 & 0 & 0 & 1 & 0 & 0 & 3.27 & 21 & 4 & 0 & 0 & 24.6 & 0 & 0 \\ 1 & 33.4 & 0 & 0 & 1 & 0 & 0 & 3.27 & 21 & 4 & 0 & 0 & 33.4 & 0 & 0 \\ 1 & 25.7 & 0 & 0 & 1 & 0 & 0 & 3.25 & 19 & 3 & 0 & 0 & 25.7 & 0 & 0 \\ 1 & 45.0 & 0 & 0 & 0 & 0 & 0 & 3.81 & 32 & 3 & 0 & 0 & 0 & 0 & 0 \\ 1 & 41.6 & 0 & 0 & 0 & 0 & 0 & 3.73 & 30 & 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 48.7 & 0 & 0 & 0 & 0 & 0 & 3.89 & 32 & 4 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$y_{o\ 21\times1} = \begin{pmatrix} 7.91 \\ 9.31 \\ 8.27 \\ 10.85 \\ 8.51 \\ 9.77 \\ 9.03 \\ 9.44 \\ 11.26 \\ 12.40 \\ 8.60 \\ 12.08 \\ 8.81 \\ 12.45 \\ 13.30 \\ 8.69 \\ 11.63 \\ 9.86 \\ 14.37 \\ 13.59 \\ 15.55 \end{pmatrix} \quad \hat{D}_{3\times3} = \begin{pmatrix} 0.1927 & -0.0085 & 0.0087 \\ -0.0085 & 0.0006 & -0.0010 \\ 0.0087 & -0.0010 & 0.0049 \end{pmatrix}$$

$$
Z_{o\ 21\times3} = \begin{pmatrix} 1 & 21.0 & 2 \\ 1 & 22.8 & 2 \\ 1 & 20.7 & 2 \\ 1 & 29.3 & 2 \\ 1 & 21.3 & 3 \\ 1 & 26.4 & 4 \\ 1 & 24.8 & 4 \\ 1 & 24.6 & 4 \\ 1 & 30.4 & 2 \\ 1 & 35.1 & 3 \\ 1 & 23.6 & 3 \\ 1 & 34.2 & 3 \\ 1 & 23.8 & 3 \\ 1 & 39.4 & 4 \\ 1 & 38.2 & 4 \\ 1 & 24.6 & 4 \\ 1 & 33.4 & 4 \\ 1 & 25.7 & 3 \\ 1 & 45.0 & 3 \\ 1 & 41.6 & 2 \\ 1 & 48.7 & 4 \end{pmatrix}
\qquad
\hat{R}_{21\times21} = \begin{pmatrix} 0.091 & 0 & \cdots & 0 \\ 0 & 0.091 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0.091 \end{pmatrix}
$$

$$
\hat{b}_{3\times1} = \hat{D}_{3\times3}\, Z'_{3\times21}\, \left(Z_{21\times3}\, \hat{D}_{3\times3}\, Z'_{3\times21} + \hat{R}_{21\times21} \right)^{-1} \left(y_{0\ 21\times1} - X_{0\ 21\times15}\, \hat{\beta}_{15\times1} \right)
$$

$$
\hat{b}_{3\times1} = \begin{pmatrix} 0.200 \\ -0.017 \\ 0.019 \end{pmatrix}
$$

$$
\hat{y}_{21\times1} = X_{n\ 21\times15}\, \hat{\beta}_{15\times1} + Z_{n\ 21\times3}\, \hat{b}_{3\times1} + \epsilon
\qquad
y_{21\times1} = \begin{pmatrix} 8.43 \\ 8.94 \\ 8.34 \\ 10.80 \\ 8.29 \\ 9.56 \\ 9.10 \\ 9.04 \\ 11.11 \\ 12.24 \\ 8.95 \\ 11.98 \\ 9.01 \\ 13.28 \\ 12.93 \\ 9.05 \\ 11.56 \\ 9.55 \\ 14.43 \\ 13.58 \\ 15.56 \end{pmatrix}
$$

$$d_{21 \times 1} = \begin{pmatrix} 21.0 \\ 22.8 \\ 20.7 \\ 29.3 \\ 21.3 \\ 26.4 \\ 24.8 \\ 24.6 \\ 30.4 \\ 35.1 \\ 23.6 \\ 34.2 \\ 23.8 \\ 39.4 \\ 38.2 \\ 24.6 \\ 33.4 \\ 25.7 \\ 45.0 \\ 41.6 \\ 48.7 \end{pmatrix} \qquad h_{ijk} = 1.3 + \left(\frac{d_{21 \times 1}}{y_{21 \times 1}} \right)^3 \qquad h_{ijk} = \begin{pmatrix} 16.8 \\ 17.9 \\ 16.6 \\ 21.3 \\ 18.2 \\ 22.4 \\ 21.5 \\ 21.4 \\ 21.8 \\ 24.9 \\ 19.6 \\ 24.5 \\ 19.7 \\ 27.4 \\ 27.1 \\ 21.4 \\ 25.4 \\ 20.8 \\ 31.6 \\ 30.0 \\ 32.0 \end{pmatrix}$$

Appendix C
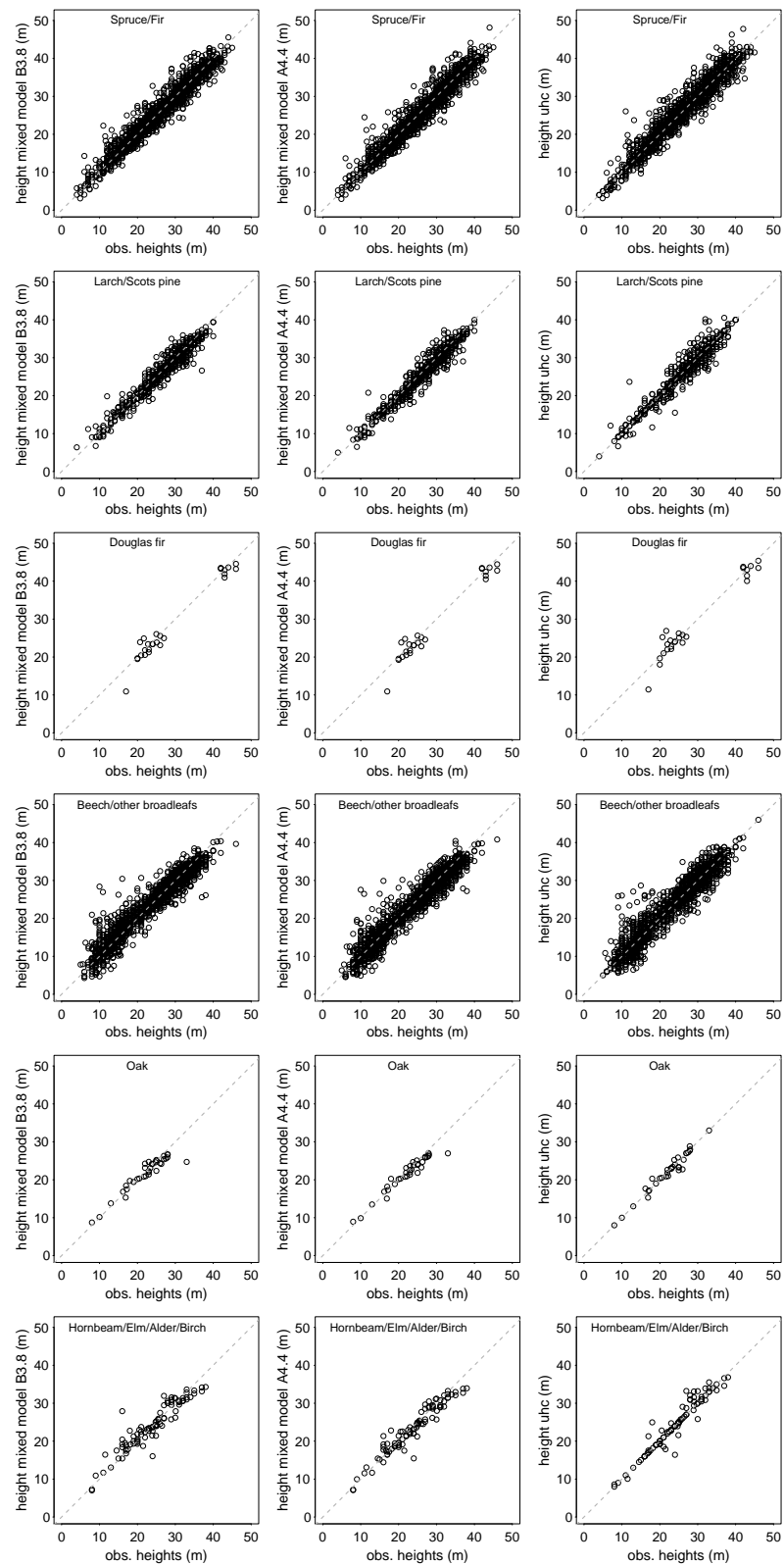
Figure A2. Predicted heights against observed heights.
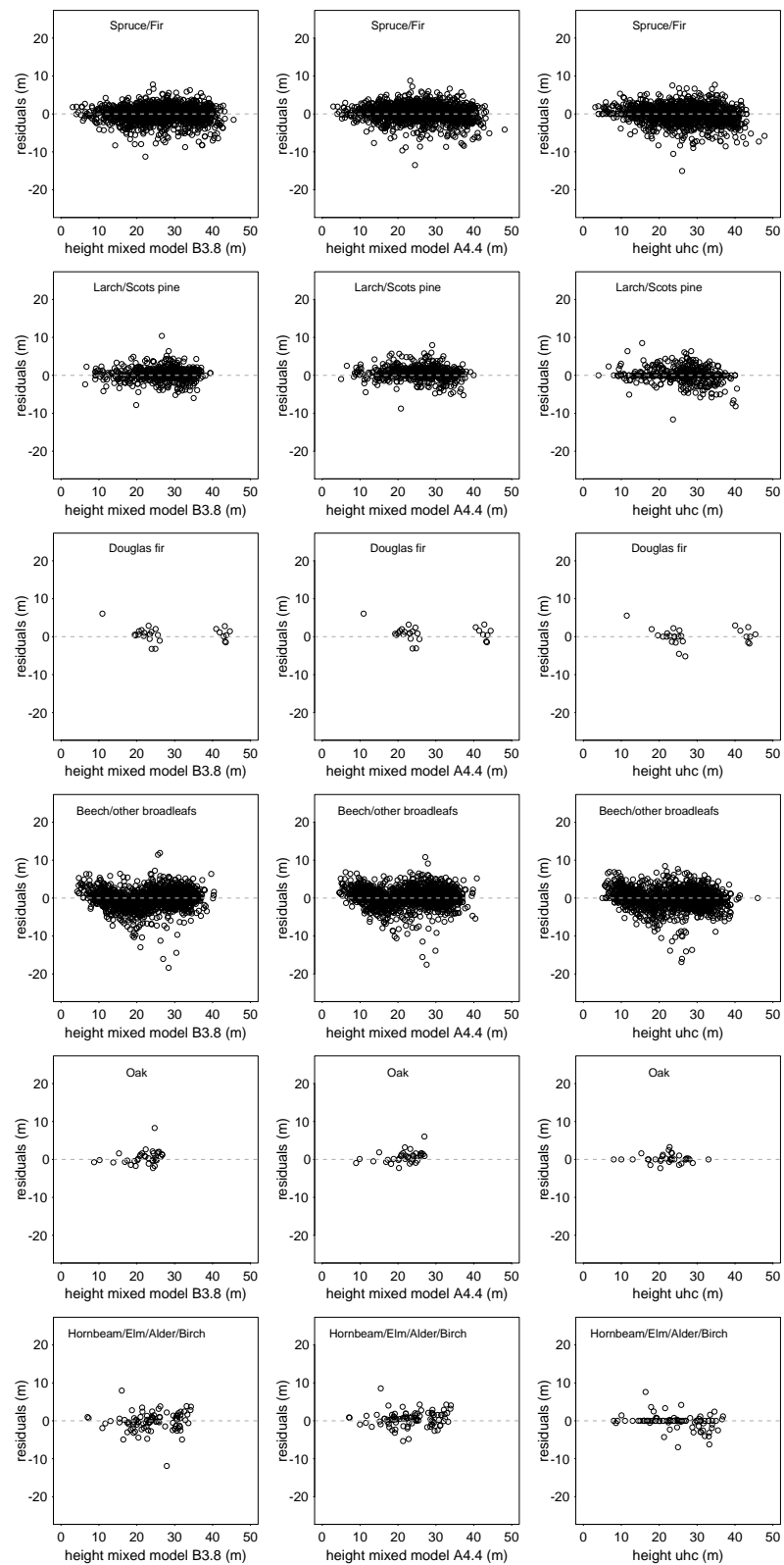
Figure A3. Residuals against predicted heights.

References

1. Köhl, M.; Magnussen, S.; Marchetti, M. *Sampling Methods, Remote Sensing and GIS Multiresource Forest Inventory*; Tropical Forestry; Springer: Berlin/Heidelberg, Germany, 2006; ISBN 978-3-540-32571-0.

2. Kershaw, J.A.; Ducey, M.J.; Beers, T.W.; Husch, B. *Forest Mensuration*; John Wiley & Sons, Ltd.: Chichester, UK, 2016; ISBN 9781118902028.

3. Soares, P.; Tomé, M. Height-diameter equation for first rotation eucalypt plantations in Portugal. *For. Ecol. Manag.* **2002**, *166*, 99–109. [CrossRef]

4. Mehtätalo, L.; de-Miguel, S.; Gregoire, T.G. Modeling height-diameter curves for prediction. *Can. J. For. Res.* **2015**, *45*, 826–837. [CrossRef]

5. Curtis, R.O. Height-Diameter and height-diameter-age equations for second-growth douglas-fir. *For. Sci.* **1967**, *13*, 365–375. [CrossRef]

6. Pretzsch, H. Forest Dynamics, Growth, and Yield. In *Forest Dynamics, Growth and Yield*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 1–39.

7. Arabatzis, A.A.; Burkhart, H.E. An evaluation of sampling methods and model forms for estimating height-diameter relationships in loblolly pine plantations. *For. Sci.* **1992**, *38*, 192–198.

8. Temesgen, H.; Gadow, K.V. Generalized height–diameter models—An application for major tree species in complex stands of interior British Columbia. *Eur. J. For. Res.* **2004**, *123*, 45–51. [CrossRef]

9. Larsen, D.R.; Hann, D.W.; Forest Research Laboratory of Oregon State University. *Height-Diameter Equations for Seventeen Tree Species in Southwest Oregon*; Forest Research Laboratory of Oregon State University. Height-Diameter Equations for Seventeen Tree Species in Southwest Oregon; Forest Research Laboratory, College of Forestry, Oregon State University: Corvallis, OR, USA, 1987; p. 49.

10. Albert, M.; Schmidt, M. Climate-sensitive modelling of site-productivity relationships for Norway spruce (*Picea abies* (L.) Karst.) and common beech (*Fagus sylvatica* L.). *For. Ecol. Manag.* **2010**, *259*, 739–749. [CrossRef]

11. Burkhart, H.E.; Tomé, M. *Modeling Forest Trees and Stands*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 53, ISBN 9788578110796.

12. Lappi, J.; Bailey, R.L. A height prediction model with random stand and tree parameters: An alternative to traditional site index methods. *For. Sci.* **1988**, *34*, 907–927.

13. Lappi, J. Calibration of height and volume equations with random parameters. *For. Sci.* **1991**, *37*, 781–801.

14. Saud, P.; Lynch, T.B.; Anup, K.C.; Guldin, J.M. Using quadratic mean diameter and relative spacing index to enhance height–diameter and crown ratio models fitted to longitudinal data. *Forestry* **2016**, *89*, 215–229. [CrossRef]

15. Lappi, J. A longitudinal analysis of height/diameter curves. *For. Sci.* **1997**, *43*, 555–570.

16. Castedo Dorado, F.; Diéguez-Aranda, U.; Barrio Anta, M.; Sánchez Rodríguez, M.; von Gadow, K. A generalized height–diameter model including random components for radiata pine plantations in northwestern Spain. *For. Ecol. Manag.* **2006**, *229*, 202–213. [CrossRef]

17. Mehtätalo, L. A longitudinal height–diameter model for Norway spruce in Finland. *Can. J. For. Res.* **2004**, *34*, 131–140. [CrossRef]

18. Nothdurft, A.; Kublin, E.; Lappi, J. A non-linear hierarchical mixed model to describe tree height growth. *Eur. J. For. Res.* **2006**, *125*, 281–289. [CrossRef]

19. Lang, A. Bestandeskurven der Würtembergischen Forsteinrichtungsanstalt. *Allg. Forst- und Jagdzeitung* **1938**, *114*, 168–176. (In German)

20. Hohenadl, W. Einführung in die Bestandesberechnung mit Hilfe von zwei Mittelstämmen. *Forstwissenschaftliches Cent.* **1939**, *61*, 261–280. (In German) [CrossRef]

21. Kramer, H. Die Genauigkeit der Massenermittlung nach dem "Reihenverfahren"- zu dem gleichlautenden Beitrag von Oberforstmeister von Laer. *Forst und Holzwirt* **1964**, 140–141. (In German)

22. Von Laer, W. Die Genauigkeit der Massenermittlung nach dem "Reihenverfahren". *Forst und Holzwirt* **1964**, 139–140. (In German)

23. Hui, G.Y.; Gadow, K. von Zur Entwicklung von Einhietshöhenkurven am Beispiel der Baumart Cunninghamia lanceolata. *Allg. Forst- und Jagdzeitung* **1993**, *12*, 218–220.

24. Marschall, J. Einheitshöhenkurven zur Vereinfachung der Probeflächenaufnhame. *IUDEO-Symposium* **1974**, 18–30. (In German)

25. Sterba, H.; Marschall, J.; da Silva, A. Einheitshöhenkurven aus und für Stichprobeninventuren. *Allg. Forstzeitg.* **1976**, *87*, 349–350. (In German)
26. Eckmüllner, O. Einheitshöhenkurven und Alters-Höhen-Durchmesser-Funktionen für Fichte und Buche im Lehrforst. *Univ. Natl. Res. Appl. Life Sci.* **1985**, 66. (In German)
27. Knieling, A. Methodische Beiträge zur Auswertung der Österreichischen Forstinventur nach 1980. Ph.D. Thesis, University of National Research and Life Sciences, Universität für Bodenkultur, Wien, Austria, 1994. (In German)
28. Schodterer, H. Einrichtung eines Permanenten Stichprobennetzes im Lehrforst. Dipl.-Arb Thesis, University of für Bodenkultur, Wien, Austria, 1987. (In German)
29. Bitterlich, W. Die Winkelzählprobe. *Allg. forst- und Holzwirtschaftliche Zeitung* **1948**, *59*, 4–5. (In German) [CrossRef]
30. Bitterlich, W. Die Winkelzählprobe. *Forstwissenschaftliches Cent.* **1952**, *71*, 215–225. (In German) [CrossRef]
31. Bitterlich, W. *The Relascope Idea. Relative Measurements in Forestry*; Commonwealth Agricultural Bureau: Wallingford, UK, 1984; ISBN 0851985394.
32. Kramer, H. *Leitfaden zur Waldmesslehre*, 4th ed.; Sauerländer: Frankfurt am Main, Germany, 1995; (In German). ISBN 9783793908302.
33. Näslund, M. *Skogsförsöksanstaltens gallringsförsök i tallskog*; Medd. Statens Skogsförsöksanstal: Stockholm, Sweden, 1936.
34. Sharma, R.P.; Vacek, Z.; Vacek, S. Nonlinear mixed effect height-diameter model for mixed species forests in the central part of the Czech Republic. *J. For. Sci.* **2016**, *62*, 470–484. [CrossRef]
35. Sharma, R.P.; Breidenbach, J. Modeling height-diameter relationships for Norway spruce, Scots pine, and downy birch using Norwegian national forest inventory data. *For. Sci. Technol.* **2015**, *11*, 44–53. [CrossRef]
36. Schmidt, M. Ein standortsensitives, longitudinales Höhen-Durchmesser-Modell als eine Lösung für das Standort-Leistungs-Problem in Deutschland. *Tagungsband der DVFFA Sektion Ertragskunde Körbecke am Möhnesee* **2010**, *17*, 131–152. (In German)
37. Kangas, A.; Maltamo, M. Anticipating the variance of predicted stand volume and timber assortments with respect to stand characteristics and field measurements. *Silva Fenn.* **2002**, *36*. [CrossRef]
38. Models, M. *Mixed-Effects Models in S and S-PLUS:Mixed-Effects Models in S and S-PLUS*; Springer: Berlin/Heidelberg, Germany, 2001; Volume 43, ISBN 0-387-98957-9.
39. Pinhero, J.; Bates, D.; Debroy, S.; Sarkar, D. Linear and Nonlinear Mixed Effects. R Packag. version 3.1-122 2015. Available online: https://CRAN.R-project.org/package=nlme (accessed on 2 May 2018).
40. Wald, A. Sequential Tests of Statistical Hypotheses. *Ann. Math. Stat.* **1945**, *16*, 117–186. [CrossRef]
41. Team, R.C. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria, 2016.
42. Reineke, L.H. Perfecting a stand-density index for evenage forests. *J. Agric. Res.* **1933**, *46*, 627–638.
43. Pollanschütz, J. Auswertung von Waldinventuren. *Forst- und Holzwirtschaft* **1973**, 355–367.
44. Schmidt, M. Ein longitudinales Höhen-Durchmesser Modell für Fichte in Nordwestdeutschland. In *Sektion Ertragskunde Jahrestagung 2008*; Nagel, J., Ed.; Deutscher Verband Forstlicher Forschungsanstalten: Göttingen, Germany, 2008; pp. 22–31.
45. Temesgen, H.; Hann, D.W.; Monleon, V.J. Regional height diameter equations for major tree species of southwest Oregon. *West. J. Appl. For.* **2007**, *22*, 213–219.
46. Kramer, H.; Akça, A. *Leitfaden zur Waldmesslehre*, 3rd ed.; Sauerländer, J.D., Ed.; Bad Orb: Aarau, Switzerland, 1995; ISBN 9783793908807.
47. Schmidt, A. Der rechnerische Ausgleich von Bestandeshöhenkurven. *Forstwissenschaftliches Cent.* **1967**, *86*, 370–382. [CrossRef]
48. Calama, R.; Montero, G. Interregional nonlinear height–diameter model with random coefficients for stone pine in Spain. *Can. J. For. Res.* **2004**, *34*, 150–163. [CrossRef]
49. Weiskittel, A.R.; Hann, D.W.; Kershaw, J.A.; Vanclay, J.K. *Forest Growth and Yield Modeling*; John Wiley & Sons, Ltd.: Chichester, UK, 2011; ISBN 9780470665008.
50. Köhler, H. Untersuchungen über die Beziehungen zwischen Standortfaktoren, Ernährungszustand und Wachstum von Fichten der Herkunft Westerhof auf verschiedenen Standorten Niedersachsens. Ph.D. Thesis, Forstlichen Fakultät, Georg-August-Universität Göttingen, Goettingen, Germany, 1984.
51. Forstliche Bundesversuchsanstalt (BFW). *Waldbau in der Subalpinen Stufe*; BFW: Vienna, Austria, 1997.

52. Nachtmann, G. Height increment models for individual trees in Austria depending on site and competition. *Austrian J. For. Sci. Cent. für das Gesamte Forstwes.* **2006**, *123*, 199–222.

53. Crecente-Campo, F.; Tomé, M.; Soares, P.; Diéguez-Aranda, U. A generalized nonlinear mixed-effects height–diameter model for *Eucalyptus globulus* L. in northwestern Spain. *For. Ecol. Manag.* **2010**, *259*, 943–952. [CrossRef]

54. Röhle, H. *Zum Wachstum der Fichte auf Hochleistungsstandorten in Südbayern*; Mitteilungen aus der Staatsforstverwaltung Bayerns; Bayerisches Staatsministerium für Ernährung, Landwirtschaft und Forsten, Referat Forstliche Aus- und Fortbildung: Munich, Germany, 1995; Volume 48.

55. Mäkinen, H.; Hynynen, J.; Isomäki, A. Intensive management of Scots pine stands in southern Finland: First empirical results and simulated further development. *For. Ecol. Manag.* **2005**, *215*, 37–50. [CrossRef]