Ciceu, A., García-Duro, J., Seceleanu, I., & Badea, O. (2020). A generalized nonlinear mixed-effects height–diameter model for Norway spruce in mixed-uneven aged stands. Forest Ecology and Management, 477, 118507. https://doi.org/10.1016/j.foreco.2020.118507

A generalized nonlinear mixed-effects height–diameter model for Norway spruce in mixed-uneven aged stands

Albert Ciceu[a,b], Juan Garcia-Duro[a], Ioan Seceleanu[c], Ovidiu Badea[a,b]

[a] *Department of Forest Monitoring, "Marin Drăcea" Romanian National Institute for Research and Development in Forestry, 128 Eroilor Blvd., 077190 Voluntari, Ilfov, Romania*

[b] *Department of Forest Engineering, Forest Management Planning and Terrestrial Measurements, Faculty of Silviculture and Forest Engineering, "Transilvania" University, 1 Ludwig van Beethoven Str., 500123 Braşov, Romania*

[c] *The Academy of Agricultural and Forestry Sciences "Gheorghe Ionescu-Şişeşti", Bvd Mărăști 61, 011464 București, Romania*

Keywords:
Norway spruce (*Picea abies* (L.) H. Karst) Uneven aged stands
Mixed unmanaged spruce forest Nonlinear generalized mixed-effects height-diameter model
Random and fixed effects calibration

A B S T R A C T

Tree height measurements are laborious and require more time and effort compared to tree diameter mea-surements. That being the case, height-diameter (H-D) models are usually used to predict individual tree heights, which are necessary for estimating the tree volume and the site index, as well as for projecting the stand de-velopment over time. Using a permanent sampling network (400x400m) from Retezat National Park in Romania, twenty-one (H-D) functions were evaluated for their fit performance, sensitivity to outliers and prediction ability for Norway spruce in mixed uneven aged stands. A set of twenty-three stand variables, both spatial and non-spatial, were used to describe the stand structure, species inter-mingling and competition, in order to be used as stand predictors in a generalized H-D model. Nonlinear mixed effects model was used in modelling the H-D relationship of Norway spruce. We developed the first generalized height-diameter model in Romania using three stand predictors as measures of the stand vertical structure, density and competition. Random and fixed effects calibration techniques were compared, testing various sampling designs in order to improve the height prediction accuracy of the model on a new dataset. Measuring six trees around the median and the thickest tree gave the best result in calibrating both fixed and random effects. On average, the best calibration design in-creased the accuracy of the prediction by 50 cm compared to the fixed effects prediction. The use of the esti-mated coefficients and the calibration design will significantly decrease the amount of work done by forest management planners, while ensuring high accuracy and reducing costs.

1. Introduction

Tree height and diameter at breast height are two of the most important variables and commonly used measurements in forest inventory. Diameter at breast height and height models have been widely used in developing growth and yield models for volume and yield prediction. The height diameter (H-D) relationship allows forest managers to build H-D models for single stands in order to determine the volumes of single trees, the stand volume or the site index (Burkhart et al., 1972; Wykoff et al., 1982; Soares and Tomé, 2002).

Even though diameters are straightforward measurements, height is heavily affected by visual obstructions which results in higher measurement errors (Castaño-Santamaría et al., 2013). Therefore, the most common approach for estimating height for all trees in a stand is to measure a subsample of heights from trees that allow a good assessment of their height and build a simple (local) H-D model.

Local H-D models developed for single stands have a low applicability and they require great sampling effort as the H-D relationship changes over time and varies from stand to stand (Curtis, 1967). A more practical method with a wider applicability is to build a generalized H-D model that accounts for the stand conditions. Such models use additional stand predictors which help explain the local variation of the H-D relationship, instead of fitting simple H–D models that are specific to each stand. The stand predictors used are usually derived from common measurements e.g. quadratic mean diameter, basal area per hectare or dominant height (Temesgen and Gadow, 2004; Castedo-Dorado et al., 2006; Trincado et al., 2007; Özçelik et al., 2018).

Ordinary least squares (OLS) regression has served as the first tool in modelling H-D relationship (Kramer, 1964; Curtis, 1967; Huang et al., 2000). Lately, the popularity and use of mixed effects models has increased in forestry and H-D modelling practices (Mehtätalo, 2004; Trincado et al., 2007; Castaño-Santamaría et al., 2013; Bronisz and Mehtätalo, 2020). This method is more reliable because it takes into account the assumption of random and independent observations which is otherwise often violated, as well as the presence of autocorrelation which is also often ignored in OLS modelling.

In addition, mixed effects models provide both fixed (population specific) and random (plot specific) effects, allowing to develop stand specific H-D curves with a minimum sample size per plot (Lappi, 1991, 1997). Furthermore, if trees in a new plot are measured for height and diameter, the fixed and random parameters of mixed effects models can be easily calibrated for that particular stand (Temesgen et al., 2008; Mehtätalo et al., 2015).

A large number of models has been used for species in temperate forests. Norway spruce (*Picea abies* (L.) H. Karst) is one of the most economically and ecologically important evergreen coniferous native tree species found in Europe, including in Romania. In Romania, it usually forms a unique belt at the tree line of the Carpathian Mountains. At lower altitudes it coexists with beech (*Fagus sylvatica* L.) and fir (*Abies alba* Mill.) in mixed mountain forests (Bravo-Oviedo et al., 2014). Their compatible ecological requirements allow beech, Norway spruce and fir to create unique ecosystems with complex structures such as uneven aged forests.

Although the H-D relationship of Norway spruce has been studied across Europe (Mehtätalo, 2004; Sharma and Breidenbach, 2015; Schmidt et al., 2018), the main focus has been on pure even aged stands. In even aged stands H-D curves tend to shift over time and a distinct different layer can be identified at each inventory with an almost parallel curve to the x axis in mature stands. In uneven aged stands, however, with more or less the same tree number and diameter distribution, the H-D relationship remains constant over time. Trees in a given diameter class always have a similar position in the stand and the H-D curve of the stand follows an S-shape form (Pretzsch, 2009).

This study focuses on describing the H-D relationship of Norway spruce in uneven aged stands located in the Retezat National Park in Romania (South – Western Carpathians). It explores the application of H-D curves in unmanaged uneven aged structures which have hardly been studied, and it provides valuable information related to both forest conservation and sustainable forest management.

The main objectives of this study are to (i) evaluate the fit performance, sensitivity to outliers and prediction ability of twenty-one H-D models, (ii) test the impact of including spatial and non-spatial stand predictors in a generalized H-D model, (iii) compare the fixed effects and the random effects calibration methods for new stands, and (iv) determine the best calibration strategy for both calibration methods, by varying both the number of heights sampled and the diameter that heights are sampled from.

2. Materials and method

2.1. Study area

The Retezat National Park (RNP) is a 380,5 km^2 protected area located in the South – Western Carpathians in Romania. The park extends over the montane and alpine altitudinal belts. Forests cover more than 45 percent of the park area with approximately 4800 ha of virgin and quasi-virgin forests (Stelian, 2002). The most important tree species in the montane belt are Norway spruce, beech, fir, sycamore maple (*Acer pseudoplatanus* L.), birch (*Betula pendula* Roth) and swiss pine (*Pinus cembra* L.), but shrubby tree species like mugo pine (*Pinus mugo* Turra) and green alder (*Alnus alnobetula* (Ehrh.) K.Koch) can also be found.

2.2. Sampling design

A permanent sampling network (PSN) developed in 2015 (Fig. 1) to assess the main indicators of forest health status and the influence of climate change on the Retezat bio-geoclimatic ecosystem covers more than 2800 ha of unmanaged forests owned by the Romanian Academy.

The sampling strategy involves a grid of 400 × 400 m with a total number of 178 permanent sample plots (PSP), one for each 16 ha. Each PSP has two circular sub-sampling plots (SSP) with a distance of 60 m between them (30 m distance each from the PSP coordinates). The SSP surface varies depending on the maximum diameter at breast height (D): 200 m^2 (r = 7.98) if the maximum D is < 28 cm or 500 m^2 (r = 12.62) otherwise (Badea, 1999, 2008).

In order to estimate the forest characteristics with the desired accuracy the number and distance between PSPs were computed using prior information from the management plans. The standing volume (m^3/ha) variability 50%, the desired 95% confidence interval and a ±10% precision were introduced in the following formula (Cochran, 1963; Giurgiu, 1972):

$$n = \frac{Fu^2 s_{\%}^2}{F\Delta^2 + u^2 s_{\%}^2} \quad (1)$$

where n is the sample size, F is the total surface of the study area, u is the selected critical value of the desired confidence level, $s_{\%}$ is the variation coefficient of the volume and Δ is the desired level of precision.

From a total of 178 PSPs only 115 were reachable and covered by tree species, while 63 plots were found inaccessible due to steep slopes or due to them being at the edge of the tree line and therefore not covered by tree vegetation.

2.3. Data

For each SSP the following variables were recorded: species code, the circumference at breast height (mm) of all trees greater than 250 mm, the radius from the center of the SSP to all trees (m), the azimuth (°) and the height (m) of a random subsample of 10 trees. The circumference was measured with a tape band to the nearest mm, the height and radius were measured using VERTEX IV hypsometer (Häglofs, Sweden) to the nearest dm, the azimuth of trees was measured using a compass and the distance between every tree and the SSP center was measured using VERTEX IV and it was rounded to the closest cm. Other tree and stand variables were measured for each sample plot following the above-mentioned purposes of the PSN. All measurements were conducted in 2015 and no repeated measurements are involved. A total number of 6447 trees was measured of which 27 different species, with 3126 pairs of height-diameter measurements. A brief summary of relevant structural descriptors is given in Table 1.

Norway spruce represents 65% of the trees measured and is the main species in the study area. The presence of beech and fir, on the other hand, varies across the PSPs, being the dominated species in the mixed forests. We proceeded analyzing the H-D relationship for Norway spruce only. All trees that were dead, broken or with broken tops were removed and the dataset was split in a prediction (80%) and a calibration (20%) H-D dataset.

In order to apply different calibration designs, we selected nine SSPs in the calibration dataset, with at least ten trees sampled and with more than 90% of them measured for both height and diameter.

All three diameter distributions of the Norway spruce dataset (all sampled trees, prediction dataset and calibration dataset) (Fig. 2) display a reverse J-shape, which is considered an essential feature of uneven aged forests (Meyer, 1952).

2.4. Base model selection

The relationship between height and diameter has been described using both linear (Curtis, 1967; Fang and Bailey, 1998) and nonlinear models (Huang et al., 1992) with two and three parameters. Due to the relative ease of fitting nonlinear models and the nonlinear nature of the

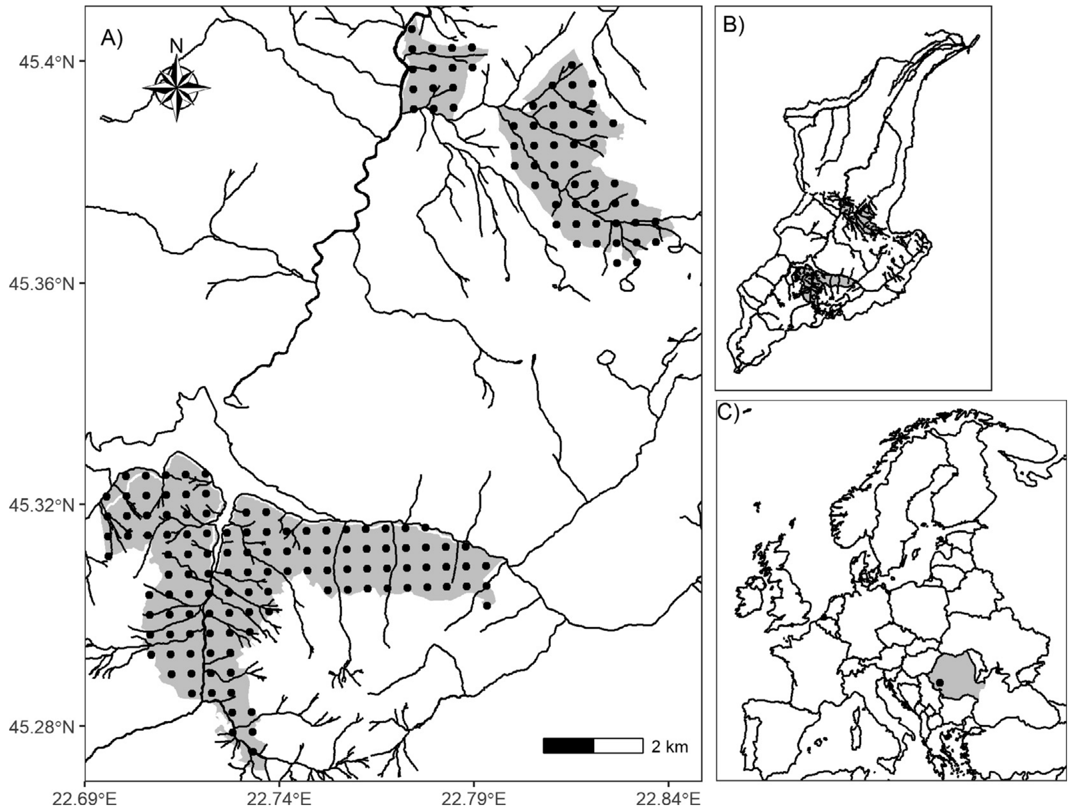

Fig. 1. Study area location where A) is the permanent sampling network (PSN), where the dots indicate the position of the PSP, B) the boundaries of Retezat National Park, C) Romania and the study area location in Europe.

Table 1
Summary of the main species descriptors of the RNP dataset: quadratic mean (Dq), maximum diameter at breast height (Max[D]), median height (Med[H]), maximum tree height (Max[H]), number of trees measured (N) and height-diameter pairs sampled (H-D pairs).

Species	D (cm)		H (m)		N	H-D pairs
	Dq	Max.	Med.	Max.		
Fir	39.9	127.6	21.5	48.4	584	391
Beech	37.4	107	21.3	43.1	923	609
Spruce	32.1	134.9	23.2	48.4	4259	1880
Other species	29.0	91.4	19.4	40.2	690	246

height-diameter relationship, 21 nonlinear models with two and three parameters were chosen from literature and tested for their fitness and predictive ability. Most of the 3-parameter models did not converge with all SSPs, although they were linearized in order to obtain starting values. The best ten models that managed to converge with all datasets are presented in the following table (Table 2) to summarize the model selection process.

The models tested have been widely used in modelling height – diameter data with high variability.

Model M2 (as indicated in the table) has been used for height prediction of uneven aged beech forests in northwestern Spain (Castaño-Santamaría et al., 2013). The latter, as well as M8 function are known to be the most flexible functions for modelling height-diameter relationships according to Yuancai and Parresol (2001). Models M3, M4 and M6 have been found to perform best in both mixed effects models and simple fixed-effects models using data from diverse ecological zones from tropical to boreal conditions (Mehtätalo et al., 2015). Model M9 has been tested for prediction for major tree species in complex stands of interior British Columbia (Temesgen and Gadow, 2004) and has given the best results for an interregional nonlinear height-diameter model for stone pine (Calama and Montero, 2004). Model M10 is widely used (Zeide, 1993; Mehtätalo, 2004; Lynch et al., 2005) as well as M1, M5 and M7 which are often used together on various datasets for comparison (Curtis, 1967; Huang et al., 1992; Soares and Tomé, 2002; Adame et al., 2008; Misik et al., 2016).

Each model was evaluated using 4 criteria: a) model fit performance for each SSP based on the root mean square error (*RMSE*) and the mean error (*ME*); b) model sensitivity to outliers for the entire prediction dataset using Prediction Sum-Of-Squares (*PRESS*) where small values indicate that the model is not overly sensitive to any single data point, and P-Square (P^2) - the equivalent to R-square (Allen, 1971); c) model prediction ability performance based on *RMSE* of 10 – fold cross – validation ($RMSE - CV_k$) using the complete prediction dataset; d) visual analysis of studentized residuals (*SR*) for each of the SSP.

The expressions of these statistics are summarized as follows:

$$RMSE = \sqrt{\frac{1}{n-p}\sum\nolimits_{j=1}^{n}(h_j - \hat{h}_j)^2} \tag{12}$$

$$ME = \frac{1}{n}\sum\nolimits_{j}^{n}(h_j - \hat{h}_j) \tag{13}$$

$$PRESS = \sum\nolimits_{j}^{n}(h_j - \hat{h}_{j,-j})^2 \tag{14}$$

$$P^2 = \frac{\sum_{j=1}^{n}(h_j - \hat{h}_{j,-j})^2}{\sum_{j}^{n}(h_j - \bar{h})^2} \tag{15}$$

$$RMSE - CV_k = \sqrt{\frac{1}{k}\sum\nolimits_{k=1}^{m}\frac{1}{n-p}\sum\nolimits_{j=1}^{n}(h_{kj} - \hat{h}_{kj})^2} \tag{16}$$

$$SR = \sum\nolimits_{j=1}^{n}\frac{d_j - \bar{d}}{sd} \tag{17}$$

whereh_j, $\hat{h}_j$ and $\bar{h}$ are the observed, the predicted and the average

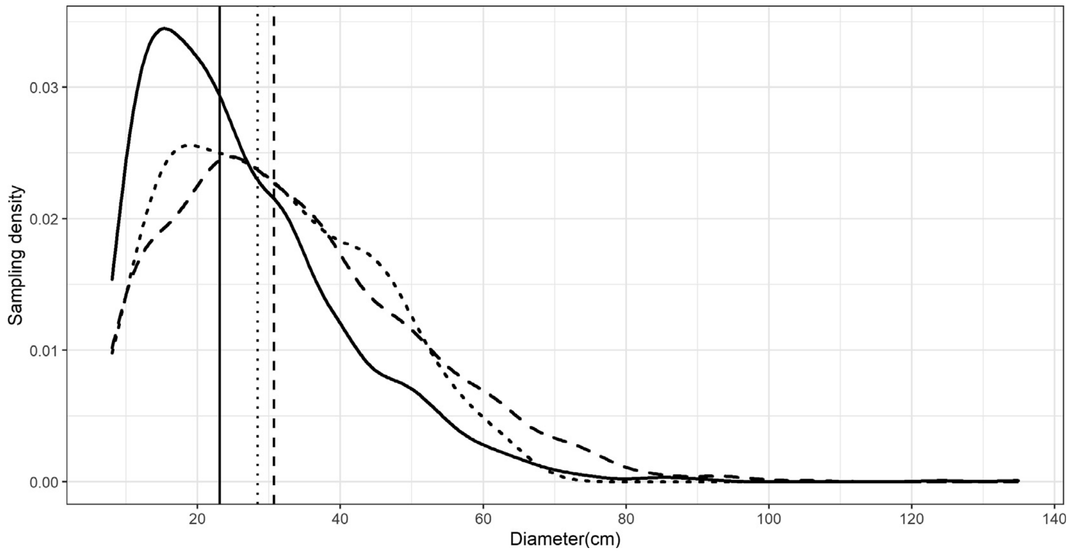

Fig. 2. The solid line is the diameter density distribution of all sampled spruce trees in the PSN, the dotted line represents the diameter density distribution of the -H-D prediction dataset and the dashed line is the diameter density distribution of the H-D calibration dataset. The vertical lines with the same line-types are the median diameter for each of the above-mentioned datasets respectively.

Table 2
H is the total tree height (m), D is the diameter at breast height (cm), b_1, b_2 and b_3 are the parameters of the model.

Model	Formulae	References	Eq.no.
M1	$H = 1.3 + \left(\frac{b_1 D}{b_2 + D}\right)$	Bates and Watts (1980)	(2)
M2	$H = 1.3 + b_1[1 - \exp(-b_2 D)]^3$	Von Bertalanffy (1957)	(3)
M3	$H = 1.3 + b_1\left(1 + \frac{1}{D}\right)^{-b_2}$	Curtis (1967)	(4)
M4	$H = 1.3 + b_1 \exp\left(\frac{b_2}{D}\right)$	Schumacher (1939)	(5)
M5	$H = 1.3 + b_1[1 - \exp(-b_2 D)]$	Meyer (1940)	(6)
M6	$H = 1.3 + \left[\frac{D}{(b_1 + b_2 D)}\right]^3$	Näslund (1936)	(7)
M7	$H = 1.3 + \frac{D^2}{b_1 + b_2 D + b_3 D^2}$	Prodan (1968)	(8)
M8	$H = 1.3 + b_1(1 - \exp(-b_2 D)^{b_3}$	Richards (1959)	(9)
M9	$H = 1.3 + exp\left[b_1 + \frac{b_2}{(D + 1)}\right]$	Wykoff et al. (1982)	(10)
M10	$H = 1.3 + b_1 \exp(-b_2 D^{-b_3})$	Zeide (1993)	(11)

values of the height of tree j; n is the total number of heights used to fit the model; p is the number of parameters used in the model; k is the subset (fold) left out from the training set and used to compute the error and m - the number of folds used; d_j, $\bar{d}$ are the diameter of the j tree and the mean diameter of each SSP respectively, while sd is the standard deviation of each SSP.

2.5. Stand variables

Generalized H-D models express the height as a function of tree diameter while using additional stand-level predictors to increase the explained variability. In order to fully grasp the H-D relationship variability, a set of twenty-three stand variables, both spatial and non-spatial, were used to describe the stand structure, species inter-mingling and competition (Pommerening, 2002; Maleki et al., 2020) for each PSP (Supplementary Table 1).

The stand structure was described using quadratic mean for diameter and height (Dq, Hq); the maximum diameter and height (Max[D], Max[H]); the range of diameters and heights (Range[D], Range[H]); the height of the thickest tree and the range of heights belonging to the thinnest and thickest trees (Max[DH], Range[DH]); the dominant height (Dom[H]), as well as the spatially explicit diameter dominance index (Dom[D] - Von Gadow and Hui, 2002) using four neighbors and the diameter variation (Dvar - Pretzsch, 2009). The horizontal structure of the plot was also estimated by two distance dependent indexes using one neighbor: the aggregation index (Agg index - Clark and Evans, 1954) and the Pielou index (Pielou, 1959).

Species diversity and inter-mingling were evaluated using the Shannon, Simpson index (Shannon, 1948; Simpson, 1949), the proportion of spruce basal area per hectare (spruce%) and the spatially explicit mingling factor (Ming - Füldner, 1995) using four neighbors.

The number of trees per hectare (N), the basal area per hectare (BA), the basal area of the largest trees (BAL - Temesgen and Gadow, 2004), the Reineke stand density index for even aged stands (SDI - Reineke, 1933) and uneven aged stands (SDI[uneven] - Shaw, 2000) and the nearest neighbors mean distance (NN mean) were used as competition and stand density variables.

The edge effect has an impact on spatial variables estimates as neighboring trees existing outside the plot border where no measurements are available lead to biased estimations. Edge correction methods are used to reduce (minus sampling) or increase in an artificial way the number of trees from a sample in order to reduce the bias of the estimators (Monserud and Ek, 1974; Radtke and Burkhart, 1998; Pommerening and Stoyan, 2006; Pretzsch, 2009). Nevertheless, no edge correction was made for the spatial variables computed because reducing the number of trees in small circular plots where the number of trees is already small leads to high bias values and they should only be applied to samples with sufficiently large number of trees (≥100) (Pommerening and Stoyan, 2006). Furthermore, edge correction methods which increase the number of trees such as the translation method (Illian et al., 2008) or the reflection method (Radtke and Burkhart, 1998) result in unrealistic periodicities, especially for circular sample plots (Windhager, 1997). Both sample reduction and sample increment methods were discarded after being initially assessed for RNP dataset.

2.6. Stand variable selection

The inclusion of stand variables can be done using various approaches; see Calama and Montero (2004). This study uses the two-stage approach (Ferguson and Leech, 1978) to include the stand predictors in the base model. In the first step all sampling plots were fitted by each model using the ordinary nonlinear least squares method (ONLS). In the second stage a redundancy analysis (RDA) was made to extract and summarize the variation of the model parameters explained by a stand predictor set. The stand variables that were statistically significant (> 0.001) in RDA and had a strong significant correlation (*p-value* < 0.001) with the models' parameters were included in the base model. In addition, each stand predictor was plotted against each of the model's parameters in order to assess the type of relationship between them.

2.7. Nonlinear generalized mixed effects model

The model that performed best was selected and used to develop a nonlinear generalized mixed effects model using additional stand predictors. The general form of a nonlinear mixed effects model (Pinheiro and Bates, 2006) is represented by:

$$y_i = f(\phi_i, x_i) + e_i \tag{18}$$

where y_i is the $n_i \times 1$ response vector of n_i observations of SSP i; x_i is the $n_i \times 1$ corresponding predictor vector of n_i observations of SSP i; ϕ is the parameter vector $r \times i$ of the nonlinear model for each SSP unit with r being the number of parameters; f is a nonlinear function of a SSP-specific parameter vector and predictor variables; and e_i is a $n_i \times 1$ vector, normally distributed, of the within SSPi error. The parameter vector can be defined as:

$$\phi_i = A_i\beta + B_i b_i,\ b_i\ N(0, \psi) \tag{19}$$

where $A_i$and B_i are design matrices for fixed and random-effects specific to SSP i and vector β and b_i are the fixed and random effects of SSP i with a variance–covariance matrix of ψ.

The model performance was tested using RMSE, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). During a preliminary analysis we observed that although our data has two levels of hierarchy (trees inside the SSP and SSP inside the PSP), a higher proportion of variance was explained when we used SSP as a single level of grouping than when applying the nested design of SSP inside of PSP. The model was tested for heteroscedasticity by applying the power-type variance function:

$$var(e_{ij}) = \sigma^2 {d_{ij}}^{2\delta} \tag{20}$$

where σ^2 is the scale, d_{ij} is the diameter j of SSP i and δ is the shape parameter of the power function.

2.8. Calibration of the nonlinear generalized fixed effects model

Mixed effects models allow the prediction of random parameters for a specific sample plot not included in the original dataset if supplementary observations are available (Lappi, 1991; Castaño-Santamaría et al., 2013; Mehtätalo et al., 2015). If no height measurements are available, the random parameters are set to 0 and the prediction of the fixed part is the standard generalized height–diameter curve.

When a set of heights from a new stand is available, the nonlinear generalized fixed effects model can be easily calibrated to that particular stand applying the following correction factor (Temesgen et al., 2008):

$$k_i^* = \frac{\sum_{j=1}^{n} [(\hat{h}_{ij} - 1.30)(h_{ij} - 1.30)]}{\sum_{j=1}^{n} (\hat{h}_{ij} - 1.30)^2} \tag{21}$$

where $k_i^*$is the correction factor of sample i, $\hat{h}_{ij}$ is the predicted height from the nonlinear generalized fixed effects model and h_{ij} is the observed height of tree j in sample i.

The predicted height based only on the fixed effects is then corrected as follows:

$$h_{ij} = k_i^* f(\phi_i, d_{ij}) + e_{ij} \tag{22}$$

where h_{ij} is the calibrated height of tree j, in the sample i after correction, the random-effects value of the nonlinear mixed effects function described above were set to 0 and d_{ij} is the diameter of tree j in sample i.

2.9. Calibration of the nonlinear generalized random effects model

A common option to predict random effects-parameters b_i is through the following approach (Chinchilli and Vonesh, 1997):

$$\hat{b}_i \approx \hat{D}\hat{Z}_i^T(\hat{R}_i + \hat{Z}_i\hat{D}\hat{Z}_i^T)^{-1}\hat{e}_i \tag{23}$$

where $\hat{D}$ is the variance–covariance matrix common to all plots for the among plot variability; $\hat{R}_i$ is the variance-covariance matrix for the within-plot variability; $\hat{e}_i$ is the residual error given by the difference between the observed height and the predicted height including only fixed effects in the model: $\hat{e}_{ij} = h_{ij} - f(\hat{\Phi}_i, d_{ij})$ where $\hat{\Phi}_i$ in this case equals $A_i\hat{\lambda}$ including only the fixed part of the estimated vector of the parameters, and $\hat{Z}_i$ is the evaluated at $\hat{\beta}$:

$$\begin{bmatrix} \frac{\partial f(d_{i1}, \hat{\Phi}_i)}{\partial \hat{\beta}_1} \cdots\cdots \frac{\partial f(d_{i1}, \hat{\Phi}_i)}{\partial \hat{\beta}_q} \\ \vdots \qquad\qquad \vdots \\ \frac{\partial f(d_{ik}, \hat{\Phi}_i)}{\partial \hat{\beta}_1} \cdots\cdots \frac{\partial f(d_{ik}, \hat{\Phi}_i)}{\partial \hat{\beta}_q} \end{bmatrix} \tag{24}$$

where $\hat{\Phi}_i$ is previously stated, $\hat{\beta}_1, \ldots, \hat{\beta}_q$ are the fixed parts of the mixed coefficients components of the vector for estimated fixed effects $\hat{\beta}$, and d_{ij} is the diameter of the j tree of plot i.

Once b_i is predicted, the value of the vector of heights, $\hat{h}_i$, for plot i, is defined by the expression:

$$\hat{h}_i = f(\hat{\Phi}_i, d_i) + e_i \tag{25}$$

$$\hat{\Phi}_i = A_i\hat{\beta} + U_i\hat{b}_i \tag{26}$$

2.10. Calibration design

The SSPs from the calibration dataset have more than 90% of the trees measured both for diameter and height which allowed us to sample heights from the entire range of the diameter distribution for each unit.

Different numbers of height sampling designs and sampling sizes were used to calibrate both the fixed and random effects parameters. We sampled the heights of three diameter categories: the thinnest, the thickest and the heights of the trees around the median tree diameter.

In the first sampling design -one-point sampling- we chose one, two and three trees from one diameter class over the curve: the thinnest, the median and the thickest tree. In the second sampling design -two-point sampling- we sampled one, two and three trees from two diameter classes (the thinnest, the median and the thickest). In the last sampling design, we used information from all the area over the curve -three-point sampling- and sampled from one to three trees from each diameter class (the thinnest, the median and the thickest).

Each SSP included in the calibration dataset was fitted with the base model (M9 – eq. no. 10) applying ONLS and using all heights measurements of the SSP. The prediction achieved was compared with the fit obtained by the calibration of the fixed and random effects using the sampling design described above.

The performance of each calibration design and sampling size was evaluated using RMSE statistic.

The nonlinear mixed effects analyses were performed using the nlme (Pinheiro et al., 2018) the graphs were built with ggplot2 package (Wickham, 2016) and the data manipulation was done with dplyr package (Wickham et al., 2019) of R statistical software (R Core Team, 2018).

Table 3
Simple fixed-effects H–D models fit performance, sensitivity to outliers and prediction ability.

Model	Fit performance		Sensitivity to outliers		Prediction ability
	RMSE	ME	P^2	PRESS	CV-RMSE
M1	2.949	−0.060	0.688	28,687	4.353
M2	2.737	0.104	**0.800**	25,216	4.145
M3	2.728	−0.081	0.780	23,684	3.997
M4	2.667	0.018	0.786	23,713	4.001
M5	2.664	0.013	0.719	24,907	4.081
M6	2.663	−0.011	0.767	23,811	4.017
M7	4.026	1.338	0.776	23,685	3.998
M8	2.668	−0.013	0.773	23,798	4.010
M9	**2.662**	**0.009**	0.778	**23,672**	**3.994**
M10	2.667	0.018	0.778	23,712	3.998

3. Results

3.1. Base model selection results

The difference between the assessed models in terms of fit performance, sensitivity to outliers and prediction ability is negligible (Table 3). Any of the 10 models except M7 can be utilized to build the generalized mixed effects model. Model 9 (M9) had the best fit, the lowest PRESS and also performed best on a new dataset. As the difference between the models performance is minor, we focused on the two parameters models in order to estimate the lowest possible number of parameters (model parsimony).

We proceeded with the analysis by comparing the standardized residuals of the models. No patterns were found in the residual plots and M9 was selected as the base model for developing the nonlinear generalized mixed effects model. The model is based on the linear relation proposed by Curtis (1967) where b_1 and b_2 are the rate and shape parameters and the value of $+\ 1$ is added to the diameter to avoid meaningless estimates when values of breast height diameter are near zero.

3.2. Variable selection

Several variables describing stand structure, stand species diversity and inter-mingling as well as stand competition exhibited a strong correlation with the M9 fixed effects parameters (Fig. 3).

The stand structure variables, in particular those measuring vertical differentiation (e.g Range[H]), showed the highest correlation with the model parameters. The distance dependent stand variables which measure the horizontal (Pielou index, Agg index) tree distribution underline a low or inexistent relation with the model parameters.

Among the stand variables measuring tree species diversity and inter-mingling (Simpson, Shannon, Ming, spruce%), only species mingling (Ming) and the proportion of Norway spruce in the species mixture (spruce%) showed a statistically significant correlation. The nearest neighbor mean (NN mean) and species mingling (Ming) are the only distance dependent variables which revealed a statistically significant relation with the model's parameters. The relationship between the variables with a significant correlation and the model's parameters was linear and no transformation was made. No statistically significant correlation was found for the stand density and competition variables BA (basal area), SDI (Reineke stand density index) for even and uneven aged stands or the BAL (basal area of the largest trees). The number of trees per hectare (N) and the NN mean (nearest neighbors mean distance) were the only two density-competition variables which explained the height variability at the plot level.

3.3. Nonlinear generalized mixed effects model

One of this paper's objectives is to develop a height diameter model to be used for height predictions in future inventories carried out by the national park administration and forest management planners. The development process was iterative, including height related variables one at a time and comparing the model's performance with the likelihood testing for a significant influence (*p-value* < 0.0001). The first generalized model we developed (Generalized 1) used the maximum height (Max[H]) and the range of heights (Range[H]) as the stand variables that had the strongest relationship with the model parameters. For field-work practical reasons, we chose the height of the thickest tree (Max[DH]) and the range of height (Range[DH]) computed as the difference between the height of the thinnest (Min[DH]) and the thickest tree (Max[DH]) as stand variables (Generalized 2). The correlation between Min[DH], Max[DH] and the real minimum and maximum heights measured is strong, over 0.96, (Fig. 4) and no significant differences were found in the model performance when the two sets of heights were used in the model (Table 4).

Other variables were tested such as Ming, spruce%, N and Dq. Although each variable slightly improved the Generalized 2 model's performance, only the number of trees per hectare had significant influence ($p < 0.0001$) and thus was added to the model (Generalized 3).

The residual plot (Fig. 5) of the nonlinear generalized mixed effects model Generalized 3 shows the presence of heteroscedasticity and confirms that variance weighting is necessary. The likelihood testing indicated that the variance function used has significant influence on the model fit ($p - value < 0.0001$) and therefore resulted in the final nonlinear generalized mixed effect model (Generalized 4).

The final nonlinear generalized mixed effects model developed has the following form:

$$H_{ij} = 1.3 + exp\left(\beta_1^{(1)} + \beta_2^{(1)}Max(DH)_i + b_i^{(1)} + \left(\frac{\beta_1^{(2)} + \beta_2^{(2)}Range(DH)_i + \beta_3^{(2)}N_i + b_i^{(2)}}{D_{ij} + 1}\right)\right) + e_{ij} \tag{27}$$

where H_{ij} is the height of tree j in sample SSPi, D_{ij} is the diameter at breast height of tree j in SSP i, $Max(DH)_i$ is the height of the maximum diameter at breast height of SSPi, $Range(DH)_i$ is the difference between the heights of trees with the maximum and the minimum diameters of SSP i, N_i is the number of trees per hectare estimation of SSP i, β_1 and β_2 are the fixed-effects parameters, and b_i represents the estimated random-effects.

The parameters of the four nonlinear generalized mixed-effects models are presented in the following table:

3.4. Calibration of fixed and random effects results

The calibration design, the sampling design and the number of heights were found to influence the calibration prediction accuracy (Table 5).

Both the calibration of random and fixed effects showed a lower RMSE as the number of trees increases. The fixed effects calibration is particularly sensitive to height variability and performed worse than the fixed effects prediction when a low number of trees was sampled among the thinnest trees. The same result was found for the random effect calibration prediction. The calibration of the fixed effects improved when the heights sampled were close to the median diameter and the thickest trees. The random effects calibration is less sensitive to height variability and it performed better than the fixed effects calibration in the one-point sampling calibration design.

For both calibration techniques the two-point sampling design where six heights around the median and thickest trees were sampled

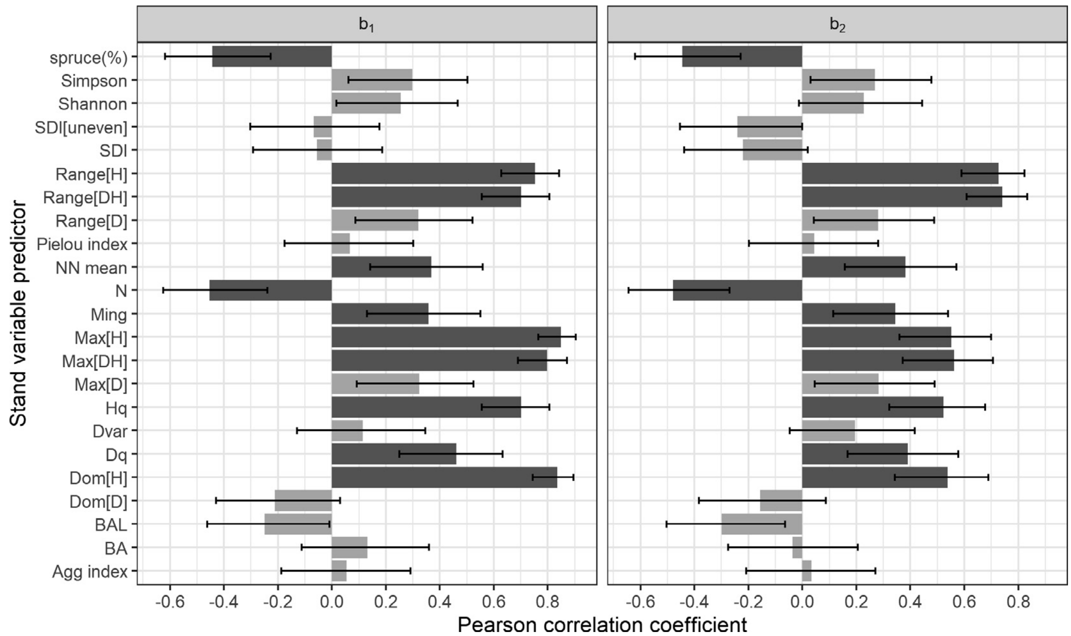

Fig. 3. Correlation barplot between the stand variables and the model parameters b_1 and b_2. The light gray bars indicate a low correlation with the model parameters while the dark gray ones show a significant, moderate and strong correlation. The whiskers represent the 0.95 confidence interval.

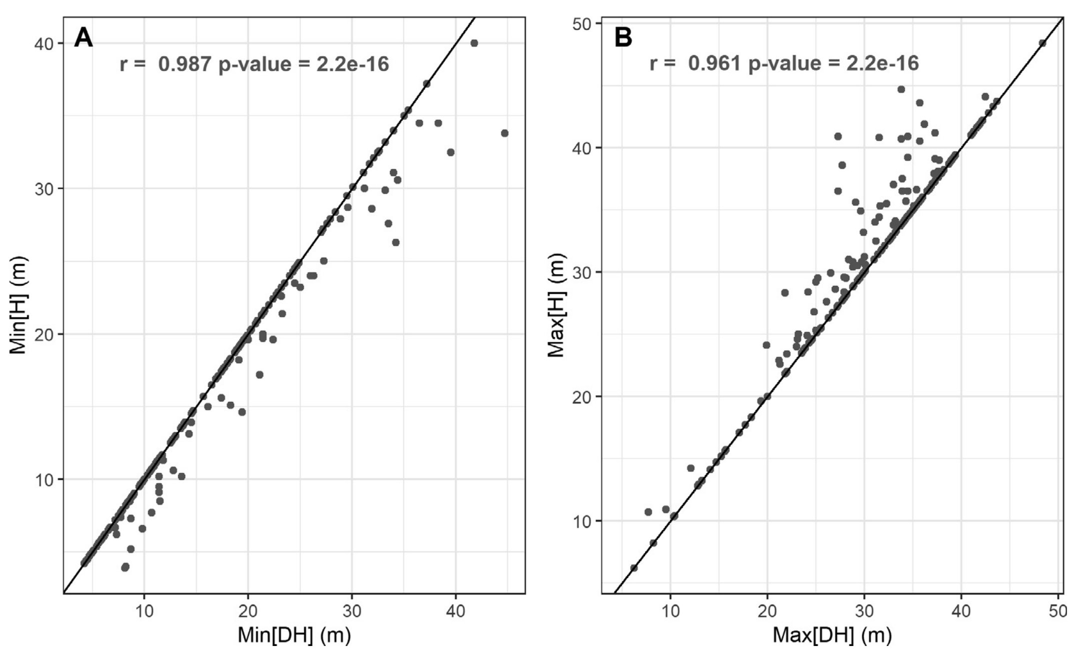

Fig. 4. The Pearson correlation (r) between the Min[DH] and Min[H] for each SSP in scatterplot A and the correlation between Max[DH] and Max[H] in scatterplot B. The points overlaying the line show perfect match between variables in a given SSP. All the other points indicate SSP height discrepancies.

resulted in the lowest RMSE increasing the prediction accuracy by up to 50 cm compared with the fixed effects prediction and proved to be the best calibration design.

Fig. 6 presents an example of the best calibration and sampling design of the calibrated fixed effects and random effects for one unit of the calibration dataset. Conjunctively it is also shown the fixed effect prediction of the generalized model and the ONLS prediction of the simple model (M9 – eq. no. 10).

4. Discussion and conclusion

4.1. Base model selection

In this study we developed the first generalized height diameter model for the Romanian forests. We used a nonlinear mixed effects regression technique (Pinheiro and Bates, 2006) to model the height-diameter relationship of Norway spruce in uneven aged stands, as it has been successfully applied in other temperate forests in Europe (Mehtätalo, 2004) .

A number of 21 models with two and three parameters were fitted separately for each SSP and compared for their fit performance, sensitivity to outliers and prediction ability. Ten of the models that managed

Table 4
The fixed effects parameters and their standard deviation (in parenthesis), the variance–covariance structure of the random effects and model performance and significance of the four nonlinear generalized mixed-effects models.

Type	Parameter	Generalized 1	Generalized 2	Generalized 3	Generalized 4
Fixed effects	$\beta_1^{(1)}$	2.9315 (0.0408)	2.9619 (0.0428)	2.9282 (0.0424)	2.9325 (0.0432)
	$\beta_2^{(1)}$	0.0236 (0.0011)	0.0237 (0.0012)	0.0249 (0.0012)	0.0248 (0.0012)
	$\beta_1^{(2)}$	−14.6695 (0.6717)	−13.8802 (0.6204)	−16.1872 (0.7321)	−16.2470 (0.7326)
	$\beta_2^{(2)}$	−0.1590 (0.0242)	−0.2471 0.0266	−0.2439 0.0249)	−0.2471 (0.0249)
	$\beta_3^{(2)}$			0.0031 (0.0005)	0.0031 (0.0005)
Random effects	$StdDev(b_i^{(1)})$	0.0822	0.0866	0.0850	0.0924
	$StdDev(b_i^{(2)})$	3.4314	2.6552	2.0336	2.3545
	$corr(b_i^{(1)}, b_i^{(2)})$	−0.7900	−0.6030	−0.6060	−0.6760
Variance function	σ^2				1.4042^2
	δ				0.1959
Model performance	AIC	8338	8358	8336	8315
	BIC	8380	8400	8384	8368
Significance testing	F-test ($p - value$)	> 0.0001	< 0.0001	< 0.0001	

to converge all SSPs were presented in the study. Non-convergence is largely caused by high variance associated with the uneven aged structure and the randomness associated with small samples. Non-convergence issues are common to other authors (Soares and Tomé, 2002; Mehtätalo et al., 2015) when applying nonlinear regression models.

The difference in the predictive performance of the ten tested base models was found to be negligible; most of the base models used showed similar statistics. Previous studies (Curtis, 1967; Huang et al., 1992) have shown that although most models found in literature give similar results within the observed range of the data, their predictive ability differs when they are used on a new dataset.

The Wykoff et al. (1982) model, M9, was chosen to build the nonlinear generalized mixed effects model as it provides the best fit, the lowest PRESS and the best predictive ability on a new dataset estimating only 2 parameters. This function was previously used for the development of the stand model PROGNOSIS (Wykoff et al., 1982), for height-diameter modelling of tree species in Southwestern Oregon and Inland Northwest (Larsen and Hann, 1987; Moore et al., 1996) and also for creating an interregional nonlinear height-diameter model for stone pine (Calama and Montero, 2004).

4.2. Stand variable selection

In order to build a generalized model and increase the applicability of the M9 model, twenty-three spatial and non-spatial plot-specific predictors describing the structure, the stand density, species diversity and inter-mingling, as well as the competition at the stand level, were assessed.

In general, structural descriptors provided a better correlation with the M9 parameters. However, all of these predictors provide more than one type of information because they are determined by a high number of ecological processes and thus they are often similar; Pommerening (2006), for example, points out that some competition and structure indices are alike.

Indeed, most of these indices are inter-correlated and even when

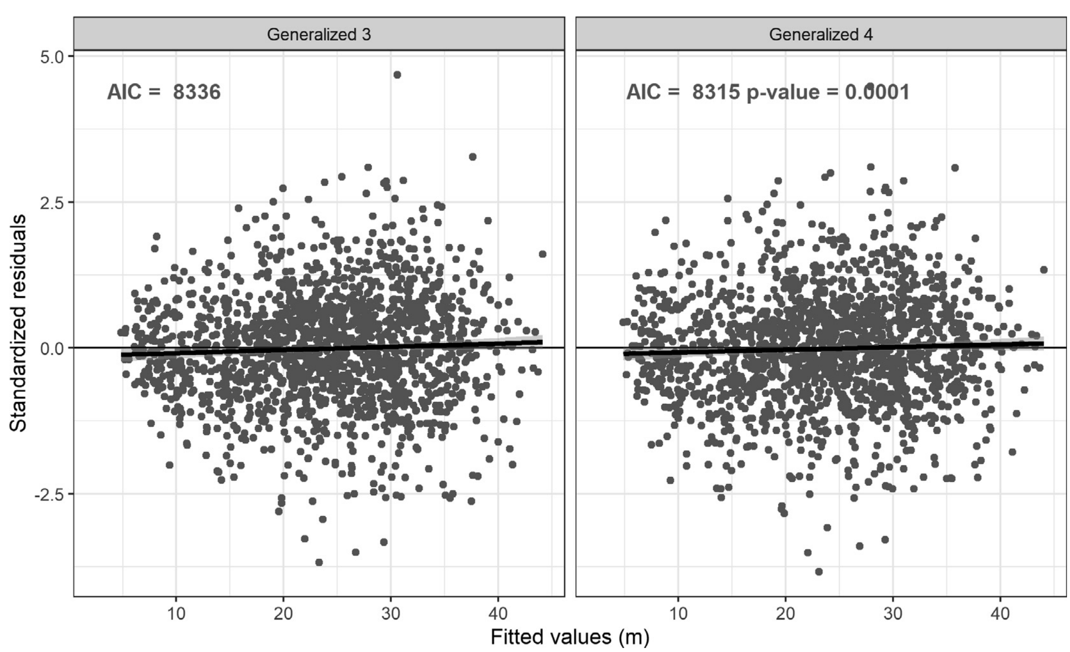

Fig. 5. Residues scatterplot of the Generalized 3 model, without variance weighting, (left) and residues scatterplot of the Generalized 4 model, with variance weighting (right).

Table 5
Calibration results.

Calibration design	Sampling design	No. trees	Fixed effects calibration		Random effects calibration	
			RMSE	sd	RMSE	Sd
Fixed effects	–	–	3.292	1.356	3.292	1.356
Base Model - ONLS	–	–	2.655	0.748	2.655	0.748
		1	4.710	1.893	3.311	1.263
	Thinnest	2	3.502	0.922	3.228	1.185
		3	3.303	0.991	3.151	1.230
		1	3.883	1.492	3.052	0.990
One-point sampling	Median	2	3.051	0.599	2.957	0.957
		3	2.888	0.808	2.887	0.968
		1	3.469	1.275	3.318	1.309
	Thickest	2	3.053	0.835	3.153	1.035
		3	2.846	0.769	2.997	0.910
		2	3.578	1.025	3.083	0.952
	Thinnest-Median	4	3.036	0.699	2.964	0.947
		6	2.928	0.792	2.895	0.975
		2	3.488	1.135	3.339	1.234
Two-point sampling	Thinnest-Thickest	4	3.086	0.855	3.111	0.996
		6	2.881	0.856	2.958	0.943
		2	2.985	0.848	3.117	1.040
	Median-Thickest	4	2.871	0.771	2.967	0.900
		6	2.744	0.781	2.821	0.856
		3	3.578	1.025	3.083	0.952
There-point sampling	Thinnest-Median-Thickest	6	3.036	0.699	2.964	0.947
		9	2.928	0.792	2.895	0.975

they had a significant correlation with the M9 model parameters individually, they did not improve the model's overall performance.

The diversity, the mingling factor and most competition variables were unsuccessful in improving the model. Similar findings were described by Huang et al. (2009), who indicated that species composition did not improve the precision of height prediction in their studies.

One likely reason is that complex stand structures such as mixed uneven aged forests contribute to niche differentiation. Indeed, Norway spruce, beech and fir tend to coexist (Bravo-Oviedo et al., 2014) and yet niche complementarity and facilitation processes diminish the competitive interferences quantified by the distance dependent predictors in mixed uneven aged stands (Dănescu et al., 2016); hence the high productivity of mixed uneven aged stands compared with even aged stands (Pretzsch et al., 2010).

In addition, most of the descriptors have certain limitations that can undermine their prediction capacity. For instance, nearest-neighbor based methods are largely affected by the interdependence between tree-distances measured and by the ecological processes in force at that scale (Pommerening, 2002).

No competition variables were used in the generalized model other than the number of trees per hectare. Although this variable had little effect on the model, it did manage to improve the model's fit. The inclusion of competition predictors in the H-D model is coherent with the fact that competition for light promotes higher vertical structure development. Therefore, even when the coefficient estimate is low, its presence is fully justified being commonly used as a stand predictor (Soares and Tomé, 2002; Calama and Montero, 2004; Sharma and Parton, 2007). The basal area per hectare is another competition variable that usually explains H-D variability and therefore it is likely to be used (Bronisz and Mehtätalo, 2020; Sharma and Parton, 2007). Nevertheless, in our study it did not show any significant correlation with the model parameters, most likely because of the uneven aged structure. However, the basal area of the largest trees underlined a significant correlation with the asymptotic part of the model M9 and it has been used by Temesgen and Gadow (2004) in similar complex structures as the one in our study.

The stand structure variables, in particular those measuring vertical differentiation (e.g Range[H]), tend to show the highest correlation with the model parameters. The quadratic mean diameter and the dominant height are the stand structure predictors most frequently used in H-D modelling as they are easy to measure or compute from available field data.

Although the dominant height, the maximum height and the height range showed a higher correlation with the stand parameters, their determination on the field is difficult. The dominant height requires a large number of height measurements, whereas the highest and the smallest trees from a stand are hard to establish. Furthermore, the determination of the dominant height in uneven aged stands is difficult due to the multiple development phases present, even in small stands (Barbir et al., 2010). This is the reason why the yield class for uneven aged stands are not computed using the dominant height as it is for even aged stands but using the height of a reference diameter.

Our model uses the height of the thickest tree, the range between the height of the thinnest and the tallest tree, and the number of trees per hectare Similarly to our study Huang et al. (2009) also found top height to be the most significant contributor among different stand-level variables. Although the variables underlined a strong correlation with both M9 model coefficients, each one of them was eventually used to modify only the parameters with which it showed the strongest correlation. By taking into account the model's parameters' mathematical meaning, the developed generalized H-D model will result in high growth rates with higher heights of the thickest trees and lower asymptotes, with reduced number of trees per hectare and higher range between the height of the thickest and thinnest tree.

4.3. Calibration design of fixed and random effects

The main purpose of the generalized H-D model is to be used as a tool for predicting tree heights for a new stand. Without any additional information, a fixed effects prediction usually provides high accuracy. However, where new heights are available, fixed effects predictions can usually be improved by calibrating the random effects for that particular stand (Trincado et al., 2007; Mehtätalo et al., 2015; Fu et al., 2017). In this study, we calibrated both the fixed and the random effects of the developed model using three calibration designs.

Although there is no accepted rule on what the number of heights or the calibration design used to calibrate the local curve should be, different studies have already addressed this issue (Calama and Montero, 2004; Temesgen et al., 2008). In our study, we varied both the number of trees sampled and their diameter in order to find the combination of both which gave the most accurate calibration. For both calibration techniques increasing the number of trees provided higher accuracy. However, the calibration of fixed effects performed weakest with a low number of trees compared with the random effects' calibration. Similar findings have been reported by Temesgen et al. (2008) and Özçelik et al. (2018) who found that the calibrated mixed-effects model performed better than the fixed effects calibration. However, Temesgen et al. (2008) state that increasing the number of trees reduces the difference between the two calibration techniques. This was the case for our study too. In fact, in some cases the fixed effects calibration outperformed the random effects' calibration when the number of sampled trees increased.

Another key point is where to take height measurements from or what trees should be selected. We excluded the random sampling as previous studies have shown that the accuracy of the calibration depends on both the number of the trees sampled and their diameter class (Calama and Montero, 2004). When we sampled from two different diameter classes (two-three-point sampling) the accuracy increased in comparison with the samples from a single diameter class, as was also

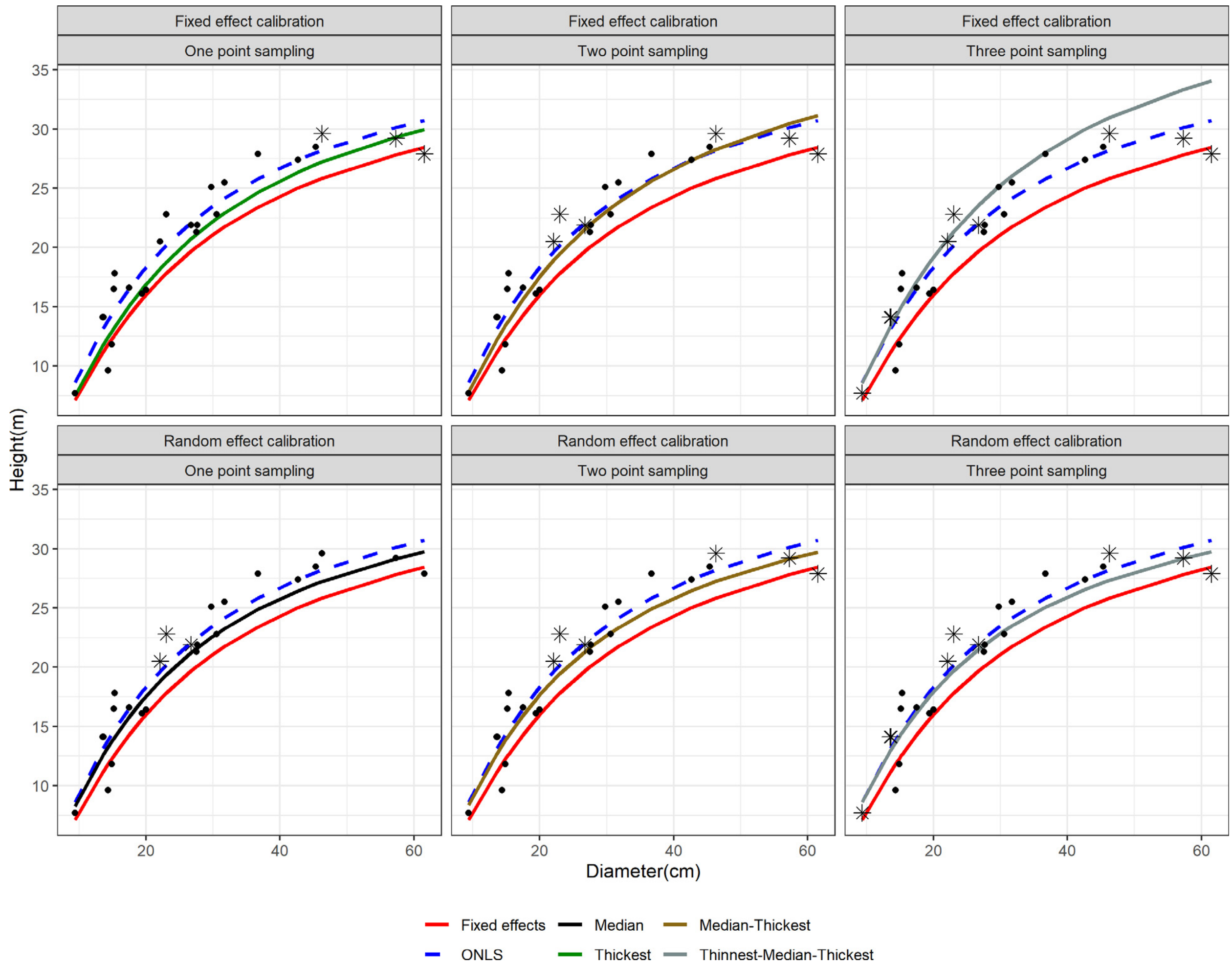

Fig. 6. Example of the best fixed and random effects calibration under different sampling designs compared to the fixed effect and ONLS prediction with the simple model (M9). The red solid line is the fixed effects prediction, the dashed blue line is the ONLS prediction and the colorful solid lines are the fixed and random effects calibrations using different calibration designs. The black dots are the heights of the sample and the star shape dots are the selected heights used for calibration. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

found by Castedo-Dorado et al. (2006) and Bronisz and Mehtätalo (2020).

The diameter classes selected are also important. Sampling from the median and thickest trees gave the best results for both calibration techniques. The same results were found by Crecente-Campo et al. (2014) when they sampled the medium size trees in mixed uneven aged stands and by Temesgen et al. (2008) when they tested different calibration techniques. A number of six trees sampled around the median and the thickest trees proved to be the best sampling design for both the fixed and the random effects calibration.

Mixed effects regression techniques performed very well in explaining the variation of the H-D relationship of Norway spruce in mixed uneven aged stands. The vertical structure, the stand density and the competition existing in an uneven forest, which all influence height variability, were explained by the stand predictors used in the model. The model has a high practical applicability as it is easy to determine both the stand level predictors and to calibrate the fixed and random effects. In future forest management planning tasks, the use of this model will not only lead to high accuracy and a more convenient field work, but it will also result in reduced costs.

CRediT authorship contribution statement

Albert Ciceu: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft. **Juan Garcia-Duro:** Methodology, Validation, Writing - review & editing. **Ioan Seceleanu:** Methodology, Writing - review & editing. **Ovidiu Badea:** Supervision, Resources, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The dataset used in this study was provided by a Core GENERESERV project supported by the Romanian Ministry of Research and Innovation within the National “Nucleu” Program – project code PN 09460117 aimed at long-term monitoring of the state of forest ecosystems. We would like to thank our colleagues from INCDS “Marin Drăcea” for their help with the sample collection and to Ms. Nataliya Nikolova for revising the English of the manuscript.

Appendix A. Supplementary material

References

Adame, P., del Río, M., Cañellas, I., 2008. A mixed nonlinear height-diameter model for pyrenean oak (Quercus pyrenaica Willd.). For. Ecol. Manage. https://doi.org/10.1016/j.foreco.2008.04.006.
Allen, D.M., 1971. Mean square error of prediction as a criterion for selecting variables. Technometrics. https://doi.org/10.1080/00401706.1971.10488811.
Badea, O., 2008. Manual on methodology for long term monitoring of forest ecosystem status under air pollution and climate change influences. Editura Silvica.
Badea, O., 1999. The forest condition monitoring in Romania (1990–1996). Imprimerie ONF, Fontainebleau, pp. 34–37.
Barbir, F.C., Roibu, C.C., Flutur, G., 2010. Structural research in the natural beech forest, situated at the eastern limit (Humosu Old Growth Beech Forest, Ia. i county, Romania). AES Bioflux 2 (2), 203–214.
Bates, D.M., Watts, D.G., 1980. Relative Curvature Measures of Nonlinearity. J. R. Stat. Soc. Ser. B. https://doi.org/10.1111/j.2517-6161.1980.tb01094.x.
Bravo-Oviedo, A., Pretzsch, H., Ammer, C., Andenmatten, E., Barbati, A., Barreiro, S., Brang, P., Bravo, F., Coll, L., Corona, P., Den Ouden, J., Ducey, M.J., Forrester, D.I., Giergiczny, M., Jacobsen, J.B., Lesinski, J., Löf, M., Mason, B., Matovic, B., Metslaid, M., Morneau, F., Motiejunaite, J., O'Reilly, C., Pach, M., Ponette, Q., Del Rio, M., Short, I., Skovsgaard, J.P., Soliño, M., Spathelf, P., Sterba, H., Stojanovic, D., Strelcova, K., Svoboda, M., Verheyen, K., Von Lüpke, N., Zlatanov, T., 2014. European mixed forests: Definition and research perspectives. For. Syst. https://doi.org/10.5424/fs/2014233-06256.
Bronisz, K., Mehtätalo, L., 2020. Mixed-effects generalized height–diameter model for young silver birch stands on post-agricultural lands. For. Ecol. Manage. https://doi.org/10.1016/j.foreco.2020.117901.
Burkhart, H.E., Parker, R.C., Strub, M.R., Oderwald, R.G., 1972. Yields of old-field loblolly pine plantations.
Calama, R., Montero, G., 2004. Interregional nonlinear height-diameter model with random coefficients for stone pine in Spain. Can. J. For. Res. https://doi.org/10.1139/x03-199.
Castaño-Santamaría, J., Crecente-Campo, F., Fernández-Martínez, J.L., Barrio-Anta, M., Obeso, J.R., 2013. Tree height prediction approaches for uneven-aged beech forests in northwestern Spain. For. Ecol. Manage. https://doi.org/10.1016/j.foreco.2013.07.014.
Castedo-Dorado, F., Diéguez-Aranda, U., Barrio Anta, M., Sánchez Rodríguez, M., von Gadow, K., 2006. A generalized height-diameter model including random components for radiata pine plantations in northwestern Spain. For. Ecol. Manage. https://doi.org/10.1016/j.foreco.2006.04.028.
Chinchilli, V.M., Vonesh, E.F., 1997. Linear and nonlinear models for the analysis of repeated measurements. M. Dekker.
Clark, P.J., Evans, F.C., 1954. Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations. Ecology. https://doi.org/10.2307/1931034.
Cochran 1963, W.G., n.d. Sampling techniques 2nd Edition, 1963 New York.
Crecente-Campo, F., Corral-Rivas, J.J., Vargas-Larreta, B., Wehenkel, C., 2014. Can random components explain differences in the height-diameter relationship in mixed uneven-aged stands? Ann. For. Sci. https://doi.org/10.1007/s13595-013-0332-6.
Curtis, R., 1967. Height-Diameter and Height-Diameter-Age Equations For Second-Growth Douglas-Fir. For. Sci. https://doi.org/10.1093/forestscience/13.4.365.
Dănescu, A., Albrecht, A.T., Bauhus, J., 2016. Structural diversity promotes productivity of mixed, uneven-aged forests in southwestern Germany. Oecologia. https://doi.org/10.1007/s00442-016-3623-4.
Fang, Z., Bailey, R.L., 1998. Height-diameter models for tropical forests on Hainan Island in southern China. For. Ecol. Manage. https://doi.org/10.1016/S0378-1127(98)00297-7.
Ferguson, I., Leech, J., 1978. Generalized Least Squares Estimation of Yield Functions. For. Sci. https://doi.org/10.1093/forestscience/24.1.27.
Fu, L., Zhang, H., Sharma, R.P., Pang, L., Wang, G., 2017. A generalized nonlinear mixed-effects height to crown base model for Mongolian oak in northeast China. For. Ecol. Manage. https://doi.org/10.1016/j.foreco.2016.09.012.
Füldner, K., 1995. Strukturbeschreibung von Buchen-Edellaubholz-Mischwäldern. [Describing forest structures in mixed beech-ash-maple-sycamore stands.]. Univ. Göttingen.
Giurgiu, V., 1972. Methods of mathematical statistics applied in forestry. Ceres Publishing House, Bucharest.
Huang, S., Price, D., Titus, S.J., 2000. Development of ecoregion-based height-diameter models for white spruce in boreal forests. For. Ecol. Manage. https://doi.org/10.1016/S0378-1127(99)00151-6.
Huang, S., Wiens, D.P., Yang, Y., Meng, S.X., Vanderschaaf, C.L., 2009. Assessing the impacts of species composition, top height and density on individual tree height prediction of quaking aspen in boreal mixedwoods. For. Ecol. Manage. https://doi.org/10.1016/j.foreco.2009.06.017.
Huang, S.M., Titus, S.J., Wiens, D.P., 1992. Comparison of nonlinear height-diameter functions for major Alberta tree species. Can. J. For. Res. https://doi.org/10.1139/x92-172.
Illian, J., Penttinen, A., Stoyan, H., Stoyan, D., 2008. Statistical Analysis and Modelling of Spatial Point Patterns. Statistical Anal. Model. Spatial Point Patterns. https://doi.org/10.1002/9780470725160.
Kramer, H., 1964. Die Genauigkeit der Massenermittlung nach dem "'Reihenverfahren"'—zu dem gleichlautenden Beitrag von Oberforstmeister von Laer [Accuracy of volume assessment using a sequential approach—response to Laer 1964]. Forst und Holzwirt 19, 140–141.
Lappi, J., 1997. A longitudinal analysis of height/diameter curves. For. Sci. 43, 555–570.
Lappi, J., 1991. Calibration of Height and Volume Equations with Random Parameters. For. Sci.
Larsen, D.R., Hann, D.W., 1987. Height-diameter equations for seventeen tree species in southwest Oregon. For. Res. Lab.
Lynch, T.B., Holley, A.G., Stevenson, D.J., 2005. A random-parameter height-dbh model for cherrybark oak. South. J. Appl. For. https://doi.org/10.1093/sjaf/29.1.22.
Maleki, K., Zeller, L., Pretzsch, H., 2020. Oak often needs to be promoted in mixed beech-oak stands – the structural processes behind competition and silvicultural management in mixed stands of European beech and sessile oak. IForest. https://doi.org/10.3832/ifor3172-013.
Mehtätalo, L., 2004. A longitudinal height-diameter model for Norway spruce in Finland. Can. J. For. Res. https://doi.org/10.1139/x03-207.
Mehtätalo, L., de-Miguel, S., Gregoire, T.G., 2015. Modeling height-diameter curves for prediction. Can. J. For. Res. 2015. https://doi.org/10.1139/cjfr-2015-0054.
Meyer, H.A., 1952. Structure, growth, and drain in balanced uneven-aged forests. J. For. 50, 85–92.
Meyer, H.A., 1940. A Mathematical Expression for Height Curves. J. For. https://doi.org/10.1093/jof/38.5.415.
Misik, T., Antal, K., Kárász, I., Tóthmérész, B., 2016. Nonlinear height-diameter models for three woody, understory species in a temperate oak forest in Hungary. Can. J. For. Res. https://doi.org/10.1139/cjfr-2015-0511.
Monserud, R.A., Ek, A.R., 1974. Plot edge bias in forest stand growth simulation models. Can. J. For. Res. 4, 419–423.
Moore, J.A., Zhang, L., Stuck, D., 1996. Height-diameter equations for ten tree species in the inland Northwest. West. J. Appl. For. https://doi.org/10.1093/wjaf/11.4.132.
Näslund, M., 1936. Skogsförsöksanstaltens gallringsförsök i tallskog.
Özçelik, R., Cao, Q.V., Trincado, G., Göçer, N., 2018. Predicting tree height from tree diameter and dominant height using mixed-effects and quantile regression models for two species in Turkey. For. Ecol. Manage. https://doi.org/10.1016/j.foreco.2018.03.051.
Pielou, E.C., 1959. The Use of Point-to-Plant Distances in the Study of the Pattern of Plant Populations. J. Ecol. https://doi.org/10.2307/2257293.
Pinheiro, J., Bates, D., 2006. Mixed-effects models in S and S-PLUS. Springer Science & Business Media.
Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Authors, E., Heisterkamp, S., Van Willigen, B., 2018. Package "nlme": Linear and nonlinear mixed effects models. Version.
Pommerening, A., 2006. Evaluating structural indices by reversing forest structural analysis. Forest Ecol. Manage. https://doi.org/10.1016/j.foreco.2005.12.039.
Pommerening, A., 2002. Approaches to quantifying forest structures. Forestry. https://doi.org/10.1093/forestry/75.3.305.
Pommerening, A., Stoyan, D., 2006. Edge-correction needs in estimating indices of spatial forest structure. Can. J. For. Res. https://doi.org/10.1139/X06-060.
Pretzsch, H., 2009. Forest dynamics, growth, and yield. In: Forest Dynamics, Growth and Yield. Springer, pp. 1–39.
Pretzsch, H., Block, J., Dieler, J., Dong, P.H., Kohnle, U., Nagel, J., Spellmann, H., Zingg, A., 2010. Comparison between the productivity of pure and mixed stands of Norway spruce and European beech along an ecological gradient. Ann. For. Sci. https://doi.org/10.1051/forest/2010037.
Prodan, M., 1968. Growth Functions, in: Forest Biometrics. https://doi.org/10.1016/b978-0-08-012441-4.50027-4.
R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Found. Stat. Comput. Vienna, Austria.
Radtke, P.J., Burkhart, H.E., 1998. A comparison of methods for edge-bias compensation. Can. J. For. Res. https://doi.org/10.1139/x98-062.
Reineke, L.H., 1933. Perfecting a stand-density index for even-aged forests.
Richards, F.J., 1959. A flexible growth function for empirical use. J. Exp. Bot. https://doi.org/10.1093/jxb/10.2.290.
Schmidt, M., Breidenbach, J., Astrup, R., 2018. Longitudinal height-diameter curves for Norway spruce, Scots pine and silver birch in Norway based on shape constraint additive regression models. For. Ecosyst. https://doi.org/10.1186/s40663-017-0125-8.
Schumacher, F.X., 1939. A new growth curve and its application to timber studies. J. For.
Shannon, C.E., 1948. A Mathematical Theory of Communication. Bell Syst. Tech. J. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.
Sharma, M., Parton, J., 2007. Height-diameter equations for boreal tree species in Ontario using a mixed-effects modeling approach. For. Ecol. Manage. https://doi.org/10.1016/j.foreco.2007.05.006.
Sharma, R.P., Breidenbach, J., 2015. Modeling height-diameter relationships for Norway spruce, Scots pine, and downy birch using Norwegian national forest inventory data. Forest Sci. Technol. https://doi.org/10.1080/21580103.2014.957354.
Shaw, J.D., 2000. Application of Stand Density Index to Irregularly Structured Stands. West. J. Appl. For. https://doi.org/10.1093/wjaf/15.1.40.
Simpson, E.H., 1949. Measurement of diversity [16]. Nature. https://doi.org/10.1038/163688a0.
Soares, P., Tomé, M., 2002. Height-diameter equation for first rotation eucalypt plantations in Portugal. For. Ecol. Manage. https://doi.org/10.1016/S0378-1127(01)00674-0.
Stelian, R., 2002. Inventar preliminar al pădurilor virgine şi cvasivirgine din teritoriul arondat şi învecinat Parcului Naţional Retezat (Preliminary inventory of virgin and quasi-virgin forests in the adjacent and neighboring area of Retezat National Park). APNR.
Temesgen, H., Gadow, K.V., 2004. Generalized height-dimater models - An application for major tree species in complex stands of interior British Columbia. Eur. J. For. Res. https://doi.org/10.1007/s10342-004-0020-z.
Temesgen, H., Monleon, V.J., Hann, D.W., 2008. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. Can. J. For. Res. https://doi.org/10.1139/X07-104.

Trincado, G., VanderSchaaf, C.L., Burkhart, H.E., 2007. Regional mixed-effects height-diameter models for loblolly pine (Pinus taeda L.) plantations. Eur. J. For. Res. https://doi.org/10.1007/s10342-006-0141-7.

Von Bertalanffy, L., 1957. Quantitative laws in metabolism and growth. Q. Rev. Biol. https://doi.org/10.1086/401873.

Von Gadow, K., Hui, G.Y., 2002. Characterizing forest spatial structure and diversity. Proc. SUFOR Int. Work. Sustain. For. Temp. Reg.

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis.

Wickham, H., François, R., Henry, L., Müller, K., 2019. Dplyr: A Grammar of Data Manipulation. R package version. Media. https://doi.org/10.1007/978-0-387-98141-3.

Windhager, M., 1997. Die Berechnung des Ek und Monserud (1974) Konkurrenzindex für Randbäume nach unterschiedlichen Berechnungsmethoden.[Calculation of the Ek and Monserud (1974) competition index for edge trees using alternative computing methods.]. Dtsch. Verband Forstl. Versuchsanstalten, Sekt. Ertragskunde, Proc. Jahrestagung 74–86.

Wykoff, W.R., Crookston, N.L., Stage, A.R., Service, F., Agriculture, U.S.D. of, 1982. User's Guide to the Stand Prognosis Model, General Technical Report (GTR). https://doi.org/https://doi.org/10.2737/INT-GTR-133.

Yuancai, L., Parresol, B.R., 2001. Remarks on Height-Diameter Modeling. South. Res. Station. For. Serv.

Zeide, B., 1993. Analysis of Growth Equations. For. Sci. https://doi.org/10.1093/forestscience/39.3.594.