

Comparison of parametric and nonparametric methods for modeling height-diameter relationships

Zdenek Adamec,
Karel Drápela

This paper focuses on the problem of regionalization of the height-diameter model at the stand level. To this purpose, we selected two different modeling techniques. As a parametric method, we chose a linear mixed effects model (LME) with calibrated conditional prediction, whose calibration was carried out on randomly selected trees either close to mean diameter or within three diameter intervals throughout the diameter range. As a nonparametric method, the technique of classification and regression trees (CART) was chosen. These two methods were also compared with the local model created by ordinary least squares regression. The results show that LME with calibrated conditional prediction based on measurements of height at three diameter intervals provided results very close to the local model, especially when six to nine trees are measured. We recommend this technique for the regionalization of the global model. The CART method provided worse results than LME, with the exception of parameters of the residual distribution. Nevertheless, the latter approach is very user-friendly, as the regression tree creation and especially its interpretation are relatively simple, and could be recommended when larger deviations are allowed.

Keywords: Calibration, Classification and Regression Trees, Hierarchical Structure, Linear Mixed Effects Model, Spatial Heterogeneity

Introduction

Modeling the relationship between the tree height and its diameter at breast height is one of the most common and oldest paradigms in forestry, especially in forest inventory (Schmidt et al. 2011). Various linear and nonlinear models have been used to study such relationship (Zhang et al. 2005). Estimates are usually obtained by the ordinary least squares (OLS) method, though the estimation of model parameters is often hampered by the failure in satisfying the assumptions of the method. Indeed, the independence between observations is very often violated due to the spatial or temporal dependencies between data, leading to biased estimates of standard errors of parameters and their confidence intervals, as well as to misleading results of significance tests (Fox et al. 2001). One of the main causes of the dependence of data in forestry models is their hierarchi-

cal structure. Data collected from the same inventory units are more similar to each other than data from different inventory units (Fox et al. 2001). To face this problem, we must consider that spatial variability is an important feature of the ecosystem, rather than a mere statistical nuisance (Legendre 1993), and its accurate assessment allows to better understand the complexity and structure of an ecosystem (Zhang et al. 2008). For this reason, models must aim to explain the spatial variability of data as much as possible. In previous studies, researchers often disregarded the problem of spatial and temporal non-stationarity, using global or local models instead. The real spatial structure is commonly neglected by global models, as they assume temporal and spatial homogeneity of the data, and this results in a high portion of unexplained variation. On the contrary, local models cannot be generalized

to populations (e.g., large areas) – as fitting many individual models requires a large number of measurements – and are unsuitable to model the variability among different plots. Several authors reported different ways to overcome these drawbacks (Calama & Montero 2005). Ferguson & Leech (1978) used two-stage models, Huang et al. (2000) applied models for each geographical or ecologically defined area, while Zhang et al. (2004) used various geostatistical methods. Among nonparametric methods used on spatially correlated data, kriging has become very popular in the last decades; several examples are reported in Kangas & Haara (2012).

The use of models with mixed effects is a successful approach to overcome the problem of spatial variability in the data. Mixed models include a fixed component, which covers the entire analyzed dataset (e.g., population), and a random component that involves various levels of hierarchy in the dataset (Calama & Montero 2005). Mixed effects model allows the simultaneous estimation of parameters of the fixed and random components, thus combining the “average” characteristics of the analyzed population (through fixed parameters of the global model) with the spatial or temporal variability of data (through the random component parameters – Adamec et al. 2008). The fixed component of the model therefore explains the influence of different variables as in the classical OLS regression (Yang & Huang 2011). In contrast, the random component takes into

Department of Forest Management and Applied Geoinformatics, Faculty of Forestry and Wood Technology, Mendel University in Brno, Brno, 613 00 (Czech Republic)

@ Zdenek Adamec (zdenek.adamec@mendelu.cz)

Received: Nov 24, 2015 - Accepted: Jul 07, 2016

Citation: Adamec Z, Drápela K (2016). Comparison of parametric and nonparametric methods for modeling height-diameter relationships. *iForest* 10: 1-8. - doi: [10.3832/ifor1928-009](https://doi.org/10.3832/ifor1928-009) [online 2016-10-19]

Communicated by: Piermaria Corona

account the heterogeneity and randomness in the data caused by known and unknown factors (Vonesh & Chiinchilli 1997). Mixed models thus provide a better parameter estimates than OLS regressions (Budhathoki et al. 2008), providing a flexible tool to analyze a given area.

The main advantage of using a mixed model is the possibility to calibrate an already existing model (Mehtätalo 2004) by estimating its parameters on data not used in the construction of the original model, as in the case of a new experimental plot. Two calibration methods are available to this purpose (Sharma & Parton 2007):

- Random parameters have a value of 0. There are no measured values of the dependent variable, so the calibration is done with only fixed effects (often referred to as “fixed effects marginal prediction”), which is related to the entire population.
- Random parameters are estimated separately for each sampling unit. It includes at least one (or better several) measured values of the dependent variable. It is a calibration with fixed and random effects (also referred to as “calibrated conditional prediction”), which applies to individual areas.

These calibration techniques have a wide range of applications, though so far they have been most frequently used in modeling growth (Miguel et al. 2013), increment (Calama & Montero 2005, Budhathoki et al. 2008, Aakala et al. 2013) and height-diameter functions (Trincado et al. 2007, Adamec et al. 2008, Schmidt et al. 2011, Castaño-Santamaría et al. 2013, Lu & Zhang 2012).

When trees in the experimental area are not enough for applying the calibrated conditional prediction, LME with fixed effects marginal predictions has to be used for tree height modeling, instead of the LME model with calibrated conditional predictions. In such cases, a suitable alternative could be the method of classification and regression trees (CART), which is a non-

parametric technique that effectively use local data (Sironen et al. 2003), does not require a normal distribution of residuals (Fan et al. 2006), and is robust to outliers (De'ath & Fabricius 2000). According to Rejwan et al. (1999), this method is suitable for modeling with hierarchical data. This approach is based on the splitting of binary data into homogeneous groups (with respect to the dependent variable) by using independent variables (Dobbertin & Biging 1998). Using qualitative dependent variables, this method produces a classification tree, while a regression tree is obtained using quantitative dependent variables (Moisen & Frescino 2002). The main advantage of this approach is that no additional height measurements are required for modeling tree height at the local level when a CART model based on other data sets is available for the area of interest. In this case, only the stand variables required for the assembled classification or the regression tree have to be identified.

Although far less used in forestry than the LME, the CART method has been successfully applied for modeling tree mortality (Dobbertin & Biging 1998) and for the regionalization of global models of the volume of tally trees (Räty & Kangas 2008) or of floodplain forest occurrences (Anderson et al. 2010). Therefore, an accurate evaluation its advantages and disadvantages in comparison with the widely used LME method is required.

The aim of this paper is to compare the use of the LME (parametric) and the CART (nonparametric) models to regionalize the global model of height-diameter relationship.

Material and methods

Study area

The study site is located in the Training Forest Enterprise Masaryk Forest Křtiny (lat. 49° 17' 43" N, long. 16° 45' 01" E), southern Moravia (Czech Republic), north of Brno. Forty-six circular sample plots were

selected in 23 different Norway spruce (*Picea abies* Karst.) stands. The sample plot area was 250 m² (stands younger than 40 years), 700 m² (stands aged between 41-80 years) and 1200 m² (stands older than 80 years). For each tree in the sample plots, its height and diameter at breast height were measured, and the basal area of each tree was calculated. Volume of each tree was calculated using volume equations by Petráš & Pajtík (1991). We measured a total of 1,590 trees. For each of the 23 forest stands considered, the mean diameter, mean height, stand basal area, mean tree volume, number of trees per ha, site index (mean height in 100 years) and the stand age were obtained by averaging the sampling plot values. Tab. 1 summarizes the basic characteristics of tree and stand variables over the study area.

Statistical analysis

All the analyses were carried out using the software packages R (R Core Team 2015) and STATISTICA® v. 12.0 (StatSoft Inc., Tulsa, OK, USA). The performances of the parametric LME and the nonparametric CART methods in modeling the height-diameter relationship were compared. We fitted the global model of the height-diameter function for the entire surveyed area and local models for each forest stand. To fit a model with mixed effects, an appropriate height-diameter (H-D) function has to be selected. We have evaluated five common H-D functions: (i) the Michailov function (Michailov 1943 – eqn. 1):

$$\hat{h} = 1.3 + ae^{\frac{b}{dbh}}$$

(ii) the Petterson function (Petterson 1955 – eqn. 2):

$$\hat{h} = 1.3 + \frac{1}{\left(a + \frac{b}{dbh}\right)^3}$$

(iii) the Näslund function (Näslund 1947 – eqn. 3):

$$\hat{h} = 1.3 + \frac{dbh^2}{(a + b \cdot dbh)^2}$$

(iv) the Levakovic function (Levakovic 1935 – eqn. 4):

$$\hat{h} = 1.3 + a \left(\frac{dbh}{1 + dbh} \right)^b$$

(v) the Meyer function (Meyer 1940 – eqn. 5):

$$\hat{h} = 1.3 + a(1 - e^{-b \cdot dbh})$$

where \hat{h} is the fitted height of a tree, dbh is the diameter at breast height, and a , b , c are the parameters to be estimated.

To select the best function for the global model, the following criteria were considered: (i) mean of residuals; (ii) residual standard deviation; (iii) standard error of residuals; (iv) Akaike's information criterion (AIC – Akaike 1973).

Tab. 1 - Basic characteristics of tree and stand variables. (DBH): diameter at breast height of a tree; (H): height; (BA): basal area; (V): volume; (DBH_m): mean diameter; (H_m): mean height; (BA_s): stand basal area; (V_m): mean tree volume; (N): number of trees; (SI): site index; (A): stand age; (95% CI): 95% confidence interval of mean value; (min): minimal value; (Max): maximal value; (STD): standard deviation.

| Level | Variable | Mean ± 95% CI | Min | Max | STD |
|-------|--|---------------|------|------|------|
| Tree | DBH (cm) | 32.2 ± 0.5 | 9.0 | 75.0 | 10.9 |
| | H (m) | 28.4 ± 0.3 | 10.0 | 41.5 | 6.4 |
| | BA (cm ²) | 911.0 ± 29.7 | 63.6 | 4417 | 603 |
| | V (m ³) | 1.17 ± 0.04 | 0.01 | 6.55 | 0.91 |
| Stand | DBH _m (cm) | 33.1 ± 0.4 | 18.6 | 47.4 | 8.3 |
| | H _m (m) | 28.8 ± 2.8 | 14.6 | 36.5 | 6.4 |
| | BA _s (m ² ha ⁻¹) | 35.7 ± 0.3 | 22.2 | 46.3 | 6.34 |
| | V _m (m ³) | 1.17 ± 0.03 | 0.16 | 2.56 | 0.66 |
| | N (n ha ⁻¹) | 498.0 ± 13 | 221 | 1240 | 272 |
| | SI (m) | 33.6 ± 0.1 | 32 | 38 | 1.7 |
| | A (years) | 81.6 ± 1.6 | 30 | 136 | 33.0 |

Linear mixed effects model

We built the actual mixed model as a two-level model, according to the methodology proposed by Weiss (2007). The first-level model included only tree variables, which vary within the forest stand in relation to height and diameter. The second-level model also included stand variables. In terms of general matrix expression, the linear mixed model can be described as follows (Gregoire et al. 1995 – eqn. 6):

$$y = X\beta + Zb + \varepsilon$$

where y is the dimensional vector of the n observations of the response variable taken from n_v trees within n_u plots, X is the design matrix of dimension $n \times m$ including the explanatory variables associated with fixed components, β is the m -dimensional vector of fixed components (estimated fixed parameters of the model), Z is the design matrix of dimension $n \times q$ for the random components, b is the q -dimensional vector of random components for the levels tree and plot, and ε is the n -dimensional vector of residual components.

Making the first-level model consists of three steps: (1) the global model; (2) the model with random effect of intercept only; and (3) the model with random effect of intercept and regression parameter. Through these steps, we tested the significance of the structure in the data, the importance of the predictor, and the random effect of the model. The significance of each model was tested using the likelihood ratio test and the AIC. After the first-level model was obtained, we fitted the second-level model by testing the contributions of the following stand variables: number of trees per ha, age of the forest stand, stand basal area, mean diameter, and the site index. In some cases, the logarithmic transformation was applied to stand variables because of the non-linear relationships observed between stand variables and the model parameter estimates.

The inclusion of the stand variables in the second-level model was done using the following procedure: the stand variable showing the highest correlation with the parameter estimates was included in the model; after the inclusion, the relation of the remaining stand variables on parameter estimation was tested again. We repeated the above procedure until the parameter estimates showed a statistically significant relationship (correlation coefficient) with some of the tested stand variables.

The mixed model obtained was calibrated using the calibrated conditional prediction in order to extend its predictability to the other plots. This step requires at least one dependent variable value to be measured in a given plot (in this case, the tree height) in order to calculate the random component of the parameter estimates. We can write the calibrated conditional prediction using the equation by Calama & Montero

(2005 – eqn. 7):

$$E(y|b) = X\beta + Zb$$

where the meaning of the symbols is the same as in eqn. 6.

For the calculation of the random components of parameter estimation, we used the best linear unbiased predictor (BLUP) proposed by Robinson (1991). The estimation of random components of the model parameters can be expressed as follows (Calama & Montero 2005 – eqn. 8):

$$\hat{b} = \hat{D}\hat{Z}(\hat{R} + \hat{Z}\hat{D}\hat{Z}^T)^{-1}\hat{e}$$

where \hat{b} is the vector of BLUP for the random components, \hat{D} is the covariance matrix of the random effects, \hat{Z} is the design matrix for the random components, \hat{R} is the estimated matrix for the residual variance, \hat{e} is the vector whose dimension equals the number of observations and whose values are the residuals of the fixed effects marginal model.

The calibration process was carried out using two different procedures of tree selection. In the first procedure, we measured the height of 1 to 5 randomly selected trees that were in the range of ± 2 cm of the mean diameter. In the second case, three diameter (d) intervals of 4 cm were set as follows: (i) d_{\min} to $d_{\min} + 4$ cm; (ii) $d - 2$ cm to $d + 2$ cm; (iii) $d_{\max} - 4$ cm to d_{\max} . In this case, 1, 2 or 3 trees were randomly selected within each interval, obtaining 3, 6 or 9 trees, respectively, to be measured for height.

The calibration was performed on six forest stands that differed significantly in age and other stand characteristics. To assess the quality of the mixed model as compared with local models, we used the following criteria: (i) coefficient of determination (R^2); (ii) root mean square error (RMSE); (iii) mean of residuals; (iv) residual standard deviation; (v) Akaike's information criterion (AIC); (vi) mean deviation (Δ_i) of fitted values obtained by the LME model from the fitted values obtained by the OLS model, calculated as follows (eqn. 9):

$$\Delta_i = \frac{\sum_{i=1}^n |\hat{y}_{iLME} - \hat{y}_{iOLS}|}{n}$$

where n is the sample size, \hat{y}_{iLME} and \hat{y}_{iOLS} are the fitted values obtained for the i -th tree ($i = 1, 2, 3, \dots, n$) by the LME and the OLS models, respectively.

Classification and regression trees

Because tree height is a continuous variable, the regression tree method was selected, whose results are homogeneous groups of individuals sharing the same predicted height (μ_h). Individual groups are explained using several different predictors, which are independent variables easy to be measured in the stand and/or to be computed from the measured variables. In this case, we used diameter at breast

height, mean diameter of tree species, age of forest stand as continuous variables, as well as the site index as a categorical variable. We assessed the contribution of each variable to the regression tree grouping using the analysis of significance proposed by Breiman et al. (1984). The degree of significance take values from 1 to 100; the larger the value, the greater the significance of given variable in the tree formation.

The regression tree model was tested on the same data used for the mixed model described above, but the original dataset was divided into a training (860 records) and a validation (730 records) subsets. The training subset was used to compile the regression tree model, while the validation subset to verify the accuracy of the model.

In the process of binary data splitting of the tree nodes, the algorithm tries to reduce the residual sum of squares. To prevent model overfitting and ensure a good estimate of μ_h , we set the minimum number of values in one node to 50, i.e., when one of the two groups obtained by splitting has fewer than 50 values, the process is terminated.

Ten-fold cross-validation pruning technique was chosen to reduce the number of branching of the regression trees achieved. After cross-validation, the regression tree models were tested on the validation subset. The quality of the resulting trees was determined using both subsets based on the following criteria: (i) coefficient of determination; (ii) mean of residuals; (iii) standard error of residuals; (iv) residual standard deviation.

Comparison of the results of the mixed model and regression tree

We compared the results of the mixed model and regression tree using data from six forest stands for which the mixed model calibration was carried out. In addition, the results of both methods were compared with those obtained with a local linear model of the height-diameter function fitted using the classical ordinary least squares (OLS) method. As comparison criteria, we have used the coefficient of determination (R^2), mean of residuals (\bar{e}), the standard deviation of residuals ($\sigma_{\bar{e}}$) and the mean deviation of fitted values obtained by the mixed model or the regression tree from the fitted values of the local model (Δ_i). For the comparison of the LME model with CART and OLS methods, we used only the mixed model with calibrated conditional prediction carried out with heights measured in three diameter intervals, as we have identified this type of calibration to show the best goodness-of-fit for the linear mixed effects model.

Results

Linear mixed effects model

Tab. 2 shows the results of the comparison of functions used in the global height-

Tab. 2 - Comparison of global height-diameter models. ($\bar{\epsilon}_i$): mean of residuals; (σ_{ϵ_i}): residual standard deviation; (SE_{ϵ_i}): standard error of residuals; (AIC): Akaike information criterion.

| Function | $\bar{\epsilon}_i$ (m) | σ_{ϵ_i} (m) | SE_{ϵ_i} (m) | AIC |
|-------------------|------------------------|---------------------------|-----------------------|--------|
| Michailov (eqn.1) | 0.0103 | 2.8101 | 0.0705 | 7802.8 |
| Petterson (eqn.4) | -0.0006 | 2.8023 | 0.0703 | 7794.0 |
| Näslund (eqn.3) | -0.0071 | 2.8081 | 0.0704 | 7800.6 |
| Levakovic (eqn.2) | 0.0085 | 2.8069 | 0.0704 | 7799.3 |
| Meyer (eqn.5) | -0.0313 | 2.8244 | 0.0708 | 7819.2 |

Tab. 3 - Comparison of first-level models. (a, b): parameters of the model; (σu_{ai}): standard deviation of the random intercept; (σu_{bi}): standard deviation of the random regression parameter; (ϵ_{ik}): residual standard deviation; (AIC): Akaike's information criterion; (p): p-value after the likelihood ratio test.

| Model | a | b | σu_{ai} | σu_{bi} | ϵ_{ik} | AIC | p |
|------------------------------|--------|--------|-----------------|-----------------|-----------------|--------|---------|
| Global | 0.2648 | 2.0848 | - | - | 0.0149 | -8861 | - |
| Random effect of a only | 0.2952 | 1.2624 | 0.0192 | - | 0.0076 | -10863 | < 0.001 |
| Random effect of a and b | 0.2944 | 1.2919 | 0.0194 | 0.1726 | 0.0001 | -10881 | < 0.001 |

diameter models. Based on all the criteria chosen, the Petterson's height-diameter function provided the best results. However, similar model performances were observed using all the functions tested, indicating their potential applicability in the prediction of tree height.

The resulting global linear model using the Petterson H-D function can be represented by the following formula (eqn. 10):

$$\frac{1}{\sqrt[3]{\hat{h}-1.3}} = 0.2648 + 2.0848 \cdot \frac{1}{dbh}$$

where \hat{h} is the fitted height of the tree and

dbh is its diameter at breast height. The whole model and its parameters were statistically significant ($p < 0.001$), with a high coefficient of determination ($R^2 = 0.781$).

All subsequent stages of mixed model (global model, the model with random effect of intercept only and model with random effect of intercept and regression parameter) were compared using the likelihood ratio test. Tab. 3 shows the results of all first-level sub-models. In all cases, we observed statistically significant differences among the different models. The final first-level mixed model therefore

included the random part of the parameter estimation for both the intercept and the regression parameter.

When compiling the second-level model, we tested the contribution of individual stand variables. The obtained results suggest that the estimates of the parameter a are mainly dependent on the logarithm of the forest stand age ($R = 0.944$), while b estimates are dependent on the age of the forest stand ($R = 0.457$). Therefore, both forest stand age and its logarithmic transformation were included into the second-level model. Subsequently, the parameter estimates of the model including the stand age were tested for relationship with the remaining stand variables and no statistically significant correlations were found. Thus, the final mixed, second-level model included as predictors the stand age in parameter b and the logarithm of stand age in parameter a . We can express the final model with following equations (eqn. 11):

$$\frac{1}{\sqrt[3]{\hat{h}-1.3}} = (0.4572 - 0.0381 \ln T + u_{ai}) + (1.1109 + 0.0025 T + u_{bi}) \cdot \frac{1}{dbh} + \epsilon_{ik}$$

where (eqn. 12):

$$u_i = \begin{bmatrix} u_{ai} \\ u_{bi} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.0000303 & -0.0000001 \\ -0.0000001 & 0.0309408 \end{bmatrix} \right)$$

and (eqn. 13):

$$\epsilon_{ik} \sim N(0, 0.00563)$$

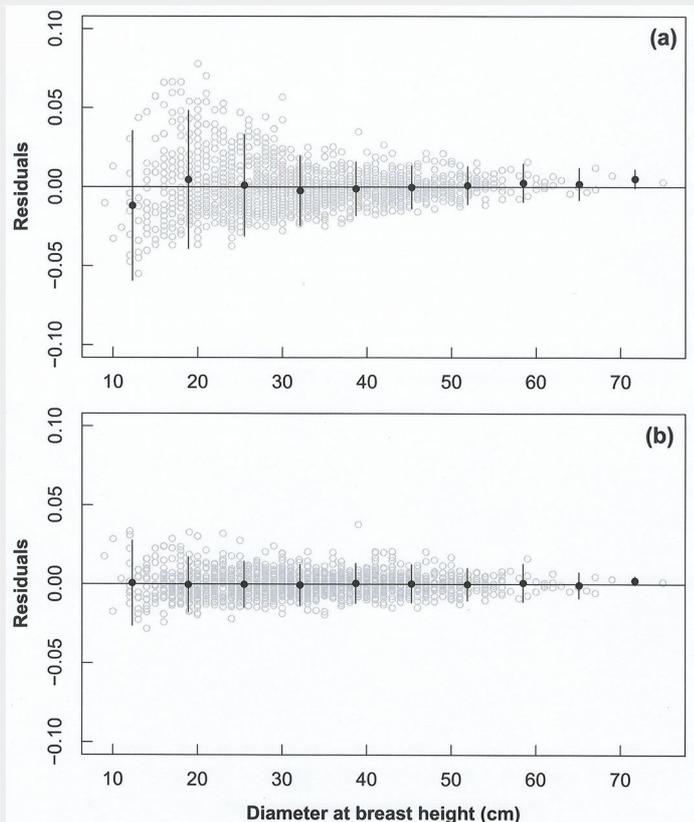
in which \hat{h} is the fitted height of a tree, dbh is the diameter at breast height, T is the age of the forest stand, u_{ai} is the random part of the intercept, u_{bi} is the random part of the regression parameter.

The second-level mixed-effects model showed a better fit on data as compared with the first-order model (likelihood ratio test: $p < 0.001$, $AIC = -10927.7$). This was confirmed by its lower dispersion of residuals, as depicted in Fig. 1.

The calibrated conditional prediction of tree height based on a random sample of 1 to 5 trees within ± 2 cm around the mean diameter provided biologically inconsistent results. Indeed, an inverse relationship was found, thereby tree height was predicted to decrease when diameter increases. Such inconsistency could not be eliminated even by using a larger number of measurements (max 5). Due to the poor quality of this calibration, a low coefficient of determination was obtained, as well as a high average deviation of fitted values of the mixed model and fitted values of the local model.

The second calibration method (tree height measured in three diameter intervals – see above) yielded significantly better results. Tab. 4 summarizes the goodness-of-fit statistics obtained using the calibrated second-level mixed effects model

Fig. 1 - Comparison of residuals (grey circles) of the global model (a) and the second-level mixed effects model (b) for the prediction of tree height. Black dots represent the mean of residuals in each diameter class, while the black lines represent its confidence intervals.



Tab. 4 - Goodness-of-fit of the calibrated second-level mixed effects model using 3 to 9 tree height measurements. (N): number of measured trees; (Δ_i): average deviation of fitted values obtained by the LME model and those obtained by the OLS model; (R^2): coefficient of determination; (AIC): Akaike's information criterion; (RMSE): root mean square error; ($\bar{\epsilon}_i$): mean of residuals; (σ_{ei}): residual standard deviation; (95% CI): 95% confidence interval of the mean.

| Statistics | N | Forest stand no. | | | | | | Mean ($\pm 95\%$ CI) |
|--------------------|-------|------------------|--------|--------|---------|---------|---------|--------------------------|
| | | 3 | 5 | 11 | 14 | 19 | 22 | |
| Δ_i | 3 | 0.549 | 0.752 | 1.144 | 1.032 | 0.694 | 1.315 | 0.914 \pm 0.310 |
| | 6 | 0.542 | 0.339 | 0.665 | 0.764 | 0.590 | 1.115 | 0.669 \pm 0.295 |
| | 9 | 0.400 | 0.239 | 0.632 | 0.642 | 0.299 | 0.493 | 0.451 \pm 0.177 |
| R^2 | 3 | 0.763 | 0.702 | 0.730 | 0.200 | 0.534 | 0.390 | 0.553 |
| | 6 | 0.777 | 0.817 | 0.812 | 0.265 | 0.569 | 0.433 | 0.612 |
| | 9 | 0.809 | 0.835 | 0.818 | 0.307 | 0.623 | 0.512 | 0.651 |
| | local | 0.848 | 0.854 | 0.860 | 0.368 | 0.644 | 0.534 | 0.685 |
| AIC | 3 | 18.421 | 42.177 | 98.792 | 120.580 | 103.958 | 131.355 | 85.880 |
| | 6 | 17.199 | 11.372 | 76.200 | 110.891 | 99.130 | 127.379 | 73.695 |
| | 9 | 10.095 | 6.847 | 74.243 | 106.892 | 87.499 | 119.566 | 67.524 |
| | local | -1.140 | -6.145 | 54.581 | 100.960 | 82.563 | 117.055 | 57.979 |
| RMSE | 3 | 1.191 | 1.333 | 1.901 | 2.458 | 1.785 | 3.186 | 1.976 |
| | 6 | 1.166 | 1.060 | 1.615 | 2.256 | 1.728 | 3.074 | 1.817 |
| | 9 | 1.084 | 1.012 | 1.591 | 2.190 | 1.617 | 2.858 | 1.725 |
| | local | 0.971 | 0.952 | 1.401 | 2.093 | 1.573 | 2.794 | 1.631 |
| $\bar{\epsilon}_i$ | 3 | 0.921 | 1.022 | 1.530 | 1.889 | 1.417 | 2.461 | 1.540 |
| | 6 | 0.893 | 0.789 | 1.288 | 1.808 | 1.394 | 2.376 | 1.425 |
| | 9 | 0.825 | 0.952 | 1.261 | 1.757 | 1.288 | 2.200 | 1.380 |
| | local | 0.768 | 0.698 | 1.074 | 1.636 | 1.251 | 2.176 | 1.267 |
| σ_{ei} | 3 | 7.003 | 9.597 | 6.695 | 11.070 | 15.330 | 14.798 | 10.749 |
| | 6 | 7.500 | 6.420 | 8.304 | 10.867 | 17.376 | 17.215 | 11.280 |
| | 9 | 5.565 | 6.510 | 8.097 | 9.912 | 13.852 | 15.041 | 9.830 |
| | local | 5.794 | 6.611 | 9.888 | 14.152 | 12.311 | 17.593 | 11.058 |

on the whole data set (training + validation) for the six forest stands analyzed. The results clearly showed that increasing the numbers of height measurements (1, 2, or 3 trees per interval, *i.e.*, 3, 6, and 9 trees), the accuracy of model predictions was largely improved. Indeed, a higher number of trees minimizes the impact of selecting dominant or suppressed trees on the position of the model curve. The results also show that the calibration using only three measured heights can provide confidence intervals of Δ_i higher than 1 m, which is obviously unacceptable from a practical point of view. On the contrary, Δ_i becomes significantly smaller when the model is run based on six or nine height measurements depending on the desired degree of accuracy, indicating the reliability of its predictions of tree height. We achieved the best results when the calibration was made using height measurements of nine trees. Subsequently, we used this model for comparison with the regression tree and the local model.

Classification and regression trees

In the CART model, all the selected variables were significant in the resulting regression tree, with a degree of significance of 100 for dbh, 97 for the mean diameter, 94 for stand age, and 57 for the site index. The resulting tree consists of 19 splitting nodes and 20 homogeneous groups (final nodes). Tab. 5 shows the actual values of the splitting criteria and the resulting values of regression tree, while Tab. 6 shows the goodness-of-fit statistics obtained for the resulting tree using the two data subsets (training and validation). The results

correctly indicated that tree height increases with increasing stand age and dbh, supporting the use of CART models in different stages of the stand development. However, the achieved regression tree showed that trees from stands with lower mean diameter are taller than trees from stands with higher mean diameter, which seems biologically counter-intuitive. We therefore explored the characteristics of forest stands from which these trees originated, finding highly heterogeneous density values across forest stands. Indeed, competition for sunlight in high-density stands results in tall and slender trees, while more resources are allocated by trees to radial growth in low density-stands due to the lower crown competition. In our opinion, this may explain the observed anomaly.

Further, based on the results from the CART model, trees from stands with lower site index were taller than those from stands with higher site index, which also seems illogical. A deeper analysis of the stands with lower site index revealed that the selected trees in these stands were mostly dominant and growing in better conditions than trees with mean diameter and height. Indeed, the site index in this study was not determined according to the height of dominant trees but to that of trees with mean height. The regression tree therefore correctly ranked these individuals into one homogeneous group based on more predictors than only site index (especially according to diameter at breast height). All the independent variables used in the CART model are thus biologically justified.

Comparison of OLS with LME and CART

The goodness-of-fit statistics obtained with each regression technique are compared in Tab. 7. The CART technique provided the lowest mean and standard deviation of residuals, but also a high average deviation of the fitted values from those predicted by the local model (in some cases close to 1 m). Although such deviation could still be considered acceptable for some applications (*e.g.*, the use in volume tables for determining the standing volume), we consider CART modeling significantly biased as compared with the OLS method. Similar statistics were obtained by using the mixed model calibrated with only 3 tree heights in 3 diameter intervals. In both cases, small deviations could be observed in individual stands, but the average deviation of the predictions obtained by the two methods from those of the OLS model was high (Tab. 7). On the other hand, once a regression tree has already been established, its application to other stands does not require any additional height measurement. In our case, to estimate the height of a specific tree we need to know its diameter at breast height (to be measured) and the stand variables, such as mean diameter, stand age, and site index, which can be easily drawn from forest management plans.

The LME technique provided good parameter estimates of the residual distribution, though their values were higher than those of the CART model. However, its predictions were very close to those achieved with the local model using the OLS method (Tab. 7), and in some cases even better. The average deviation of predictions by

Tab. 5 - Results of the regression tree construction. (LB): left branch; (RB): right branch; (N): number of trees in node; (μ_h): mean value of the height distribution; (SP): splitting variable; (dbh): diameter at breast height of a tree; (dbh_m): mean diameter; (A): stand age; (SI): site index; (f): final nodes.

| Node | LB of the node | RB of the node | N | μ_h (m) | SP | Value of SP |
|-----------------|----------------|----------------|-----|-------------|------------------|-------------|
| 1 | 2 | 3 | 860 | 27.97 | dbh _m | 30.1 |
| 2 | 4 | 5 | 380 | 22.09 | A | 48 |
| 4 | 6 | 7 | 197 | 19.09 | dbh | 20.5 |
| 6 | 8 | 9 | 93 | 17.04 | dbh | 12.5 |
| 8 ^f | - | - | 11 | 12.82 | - | - |
| 9 | 10 | 11 | 82 | 17.60 | dbh | 16.5 |
| 10 ^f | - | - | 26 | 16.29 | - | - |
| 11 | 12 | 13 | 56 | 18.21 | A | 39 |
| 12 ^f | - | - | 41 | 17.54 | - | - |
| 13 ^f | - | - | 15 | 20.07 | - | - |
| 7 | 14 | 15 | 104 | 20.92 | A | 39 |
| 14 ^f | - | - | 47 | 19.16 | - | - |
| 15 ^f | - | - | 57 | 22.37 | - | - |
| 5 | 18 | 19 | 183 | 25.33 | dbh | 23.5 |
| 18 | 20 | 21 | 78 | 22.54 | dbh | 16.5 |
| 20 ^f | - | - | 16 | 19.63 | - | - |
| 21 ^f | - | - | 62 | 23.29 | - | - |
| 19 | 24 | 25 | 105 | 27.40 | SI | 32 |
| 24 ^f | - | - | 46 | 29.01 | - | - |
| 25 | 26 | 27 | 59 | 26.14 | dbh | 30.5 |
| 26 ^f | - | - | 42 | 25.62 | - | - |
| 27 ^f | - | - | 17 | 27.44 | - | - |
| 3 | 28 | 29 | 480 | 32.62 | dbh | 32.5 |
| 28 | 30 | 31 | 126 | 28.99 | dbh | 23.5 |
| 30 ^f | - | - | 16 | 24.56 | - | - |
| 31 | 32 | 33 | 110 | 29.64 | A | 75 |
| 32 ^f | - | - | 23 | 27.85 | - | - |
| 33 | 34 | 35 | 87 | 30.11 | dbh _m | 39.3 |
| 34 ^f | - | - | 18 | 31.81 | - | - |
| 35 ^f | - | - | 69 | 29.67 | - | - |
| 29 | 40 | 41 | 354 | 33.91 | dbh | 46.5 |
| 40 | 42 | 43 | 269 | 33.18 | A | 75 |
| 42 ^f | - | - | 25 | 29.70 | - | - |
| 43 | 44 | 45 | 244 | 33.53 | dbh _m | 39.3 |
| 44 ^f | - | - | 45 | 35.52 | - | - |
| 45 | 46 | 47 | 199 | 33.08 | dbh | 40.5 |
| 46 ^f | - | - | 111 | 32.51 | - | - |
| 47 ^f | - | - | 88 | 33.80 | - | - |
| 41 ^f | - | - | 85 | 36.22 | - | - |

Tab. 7 - Comparison of goodness-of-fit statistics for LME (mixed model), CART (regression tree model) and OLS (local model) techniques. (R^2): coefficient of determination; ($\bar{\epsilon}_i$): mean of residuals; (σ_{ϵ_i}): residual standard deviation; (Δ_i): average deviation of the fitted values obtained using the LME model and the fitted values obtained using the OLS model.

| Model | Stats | Forest stand no. | | | | | | Mean |
|-------|-----------------------|------------------|-------|-------|--------|--------|--------|--------|
| | | 3 | 5 | 11 | 14 | 19 | 22 | |
| LME | R^2 | 0.809 | 0.835 | 0.818 | 0.307 | 0.623 | 0.512 | 0.651 |
| | $\bar{\epsilon}_i$ | 0.825 | 0.952 | 1.261 | 1.757 | 1.288 | 2.200 | 1.380 |
| | σ_{ϵ_i} | 5.565 | 6.510 | 8.097 | 9.912 | 13.852 | 15.041 | 9.830 |
| | Δ_i | 0.400 | 0.239 | 0.632 | 0.642 | 0.299 | 0.493 | 0.451 |
| CART | R^2 | 0.783 | 0.727 | 0.767 | 0.202 | 0.570 | 0.465 | 0.586 |
| | $\bar{\epsilon}_i$ | 0.089 | 0.094 | 0.148 | -0.530 | -0.021 | 0.144 | -0.013 |
| | σ_{ϵ_i} | 1.146 | 1.293 | 1.788 | 2.272 | 1.718 | 2.963 | 1.863 |
| | Δ_i | 3.546 | 2.303 | 2.220 | 1.258 | 1.179 | 1.216 | 1.954 |
| OLS | R^2 | 0.848 | 0.854 | 0.860 | 0.368 | 0.644 | 0.534 | 0.685 |
| | $\bar{\epsilon}_i$ | 0.768 | 0.698 | 1.074 | 1.636 | 1.251 | 2.176 | 1.267 |
| | σ_{ϵ_i} | 5.794 | 6.611 | 9.888 | 14.152 | 12.311 | 17.593 | 11.058 |

Tab. 6 - Goodness-of-fit statistics obtained using the CART model. (R^2): coefficient of determination; ($\bar{\epsilon}_i$): mean of residuals; (SE_{ϵ_i}): standard error of residuals; (σ_{ϵ_i}): residual standard deviation.

| Statistics | Training data | Validation data |
|-----------------------|---------------|-----------------|
| R^2 | 0.940 | 0.804 |
| $\bar{\epsilon}_i$ | 0.000 | -0.585 |
| SE_{ϵ_i} | 0.054 | 0.102 |
| σ_{ϵ_i} | 1.590 | 2.746 |

both the local model and the mixed model (based on 9 height measurements) were small both in the individual forest stands and overall. We concluded that the LME technique is superior to the CART technique for the prediction of height based on dbh, and that the mixed model can replace the local model without significant losses of goodness-of-fit of the height curve model. Moreover, the mixed model also guarantees a consisting reduction in the variability of residuals.

Discussion

The results of this study revealed that tree height can be accurately predicted based on diameter at the breast height by including the stand age and its logarithmic transformation as random factors in the mixed model considered. Moreover, the mixed model removed the heteroscedasticity occurring in the global model (Fig. 1), as previously reported by Drápela (2011). However, the inclusion of stand age as a stand variable makes the use of the above model suitable only to even-aged forests. Mixed models for height prediction were previously applied to even-aged forests, but using different stand variables, such as the dominant height and dominant diameter (Crecente-Campo et al. 2010), the mean diameter (Schmidt et al. 2011), or a combination of stand variables (Sharma & Parton 2007, Adame et al. 2008). Other authors have focused on uneven-aged forest using the mean/dominant diameter or the dominant height as predictors (Castaño-Santamaría et al. 2013), as the use of stand age is unfeasible. Eerikäinen (2003) stated that the inclusion of the dominant height in models may reflect the site index as an indicator of the production potential of a site. All the above evidences support the hypothesis that the inclusion of stand variables into the model improves its final accuracy (Soares & Tomé 2002).

The applicability of the proposed mixed model in practice has been assessed by comparing the effect of different ways of selecting trees for height measurement to be used in the calibrated conditional prediction. Using the first method, 1 to 5 trees within ± 2 cm from the mean diameter were selected, while the second was based on 3, 6, or 9 trees randomly selected within 4 cm intervals from the thinnest, medium

and thickest trees. Other papers also included two ways of tree selection, namely random selection of different number of trees and a selection focused only on a portion of the diameter range. For example, random selection of trees has been applied by Adame et al. (2008) and Kangas & Haara (2012), who selected 1 to 4 trees, while Trincado et al. (2007) chose only 1 tree. Sharma & Parton (2007) and Schmidt et al. (2011) carried out a selection of 1 to 3 trees. Crecente-Campo et al. (2010) and Castaño-Santamaría et al. (2013) randomly selected 1 to 10 trees in each of three ranges of minimum, mean and largest diameters, as well as 3, 6, or 9 trees in several ranges of the smallest, largest and mean diameter.

Calibration results showed that random selection of trees close to the mean diameter is not feasible, as it provides predictions too biased compared to local models or not biologically justified (height decreasing as the diameter increases). Crecente-Campo et al. (2010) reported similar results, with very high mean and standard error of residuals of the resulting model using 1 to 10 trees of mean diameter. Castaño-Santamaría et al. (2013) using trees of mean diameter did not find such high means of residuals, but recommended the use of trees with smaller diameter to achieve better results in the calibrated conditional prediction.

The second method, based on the selection of trees from more diameter intervals, proved to yield better results in this study. Similarly, Crecente-Campo et al. (2010) and Castaño-Santamaría et al. (2013) reported that this approach improves the predictive power of the model. The results of this study confirm that the goodness-of-fit of the model increases by increasing the number of measured trees, in accordance with Trincado et al. (2007), Adame et al. (2008) and Kangas & Haara (2012).

Sharma & Parton (2007) and Schmidt et al. (2011) compared both LME model with fixed effects marginal prediction and calibrated conditional prediction, concluding that the latter provides better results even when only one tree is used for model calibration.

Van Laar & Akça (2007) recommended a minimum number of 20-25 tree height measurements for determining a local model of the height curve, to be used for example for calculation of the standing volume using the volume tables. We found that the reduction in number (up to 6 or 9) of the required measured heights through the calibrated model did not invalidate the model predictability, but at the same time ensures a reduction in costs and time of data collection. According to Crecente-Campo et al. (2010), this reduction in the number of required measurements makes the application of mixed models highly effective and useful.

The results obtained using the CART model indicated that tree height depends on

the diameter at breast height, mean diameter, stand age, and site index. Strongly nonlinear relationships exist between the height and the above variables. According to Gómez et al. (2012), the CART method should be preferred to parametric models as it is more suitable to analyze nonlinear relationships as well as data with hierarchical structure (De'ath & Fabricius 2000, Shifley et al. 2006).

Indeed, the strong dependence among independent variables (e.g., dbh and age) could bring about problems of multicollinearity when parametric methods are used. Contrastingly, CART models have been reported to resist to multicollinearity (Shifley et al. 2006), as CART considers the high-hierarchy variables to be more important than low-hierarchy variables, and the independent variables to have greater influence on the results than their mutual correlations. Consequently, there are no biased estimates of parameters (Aertsen et al. 2010). Further, the easy description of interactions between variables make this method a useful tool for the interpretation of the results, particularly when the number of independent variables increases and other model types are very difficult to interpret (De'ath & Fabricius 2000).

Our results showed that tree height can be quite easily predicted using the CART method. CART models have been previously applied to estimate different variables in forestry, such as tree mortality (Dobbertin & Biging 1998, Fan et al. 2006), site index (Aertsen et al. 2010) and the form factor of sample trees (Räty & Kangas 2008). The interpretation of the regression tree obtained with the CART model is very easy and intuitive, which makes this method a suitable tool for classifying forest stands (Gómez et al. 2012). Moreover, once the regression tree is achieved, no additional measurements are required, thus reducing the costs of data collection.

In this study, the comparison of mixed model, regression tree, and the local model indicated that LME performed better than CART in predicting tree height based on dbh. The CART method also provided results significantly worse than those obtained with the OLS method (local model), except for parameters of the residual distribution. Moisen & Frescino (2002) compared several modeling techniques to predict various forest characteristics, confirming the worse performances of CART compared to OLS methods. Aertsen et al. (2010) compared CART with multiple linear regression (MLR), generalized additive models (GAM) boosted regression trees (BRT) and artificial neural networks (ANN), confirming that CART is more user-friendly than the other four modeling techniques.

Regarding the comparison between LME and OLS models, we found that predictions obtained with LME only slightly differs from those of the local model. Crecente-Campo et al. (2010) found that calibrated LME provides significantly better results

than the global model, thanks to the calibrated conditional prediction of the local model. Similar results were achieved by Castaño-Santamaría et al. (2013). We concluded that the mixed model with calibrated conditional prediction is best suited for the regionalization of the global model, e.g., up to the level of the forest stand. We recommend to calibrate the model based on 6 to 9 tree height measurements in a given stand, depending on the required level of accuracy.

Acknowledgements

The authors thank two anonymous reviewers whose comments helped to improve an earlier version of the manuscript. We would also like to thank Jan Kadavý for his valuable comments.

References

- Aakala T, Fraver S, Amato AW, Palik BJ (2013). Influence of competition and age on tree growth in structurally complex old-growth forests in Northern Minnesota, USA. *Forest Ecology and Management* 308: 128-135. - doi: [10.1016/j.foreco.2013.07.057](https://doi.org/10.1016/j.foreco.2013.07.057)
- Adame P, Del Rio M, Cañellas I (2008). A mixed nonlinear height-diameter model for Pyrenean oak (*Quercus pyrenaica* Willd.). *Forest Ecology and Management* 256: 88-98. - doi: [10.1016/j.foreco.2008.04.006](https://doi.org/10.1016/j.foreco.2008.04.006)
- Aertsen W, Kint V, Van Orshoven J, Ozkan K, Muys B (2010). Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. *Ecological Modelling* 221: 1119-1130. - doi: [10.1016/j.ecolmodel.2010.01.007](https://doi.org/10.1016/j.ecolmodel.2010.01.007)
- Akaike H (1973). Information theory and an extension of the maximum likelihood principle. In: Proceedings of the "2nd International Symposium on Information Theory" (Petrov BN, Csáki F eds). Budapest (Hungary) 2-8 September 1973. Akademiai Kiado, Budapest, Hungary, pp. 268-281.
- Anderson MG, Ferree CE, Olivero AP, Zhao F (2010). Assessing floodplain forests: using flow modeling and remote sensing to determine the best places for conservation. *Natural Areas Journal* 30: 39-52. - doi: [10.3375/043.030.0105](https://doi.org/10.3375/043.030.0105)
- Breiman L, Friedman JH, Olshen RA, Stone CJ (1984). Classification and regression trees. Chapman and Hall/CRC, Boca Raton, FL, USA, pp. 368.
- Budhathoki CB, Lynch TB, Guldin JM (2008). Nonlinear mixed modeling of basal area growth for shortleaf pine. *Forest Ecology and Management* 255: 3440-3446. - doi: [10.1016/j.foreco.2008.02.035](https://doi.org/10.1016/j.foreco.2008.02.035)
- Calama R, Montero G (2005). Multilevel linear mixed model for tree diameter increment in stone pine (*Pinus pinea*): a calibrating approach. *Silva Fennica* 39: 37-54. - doi: [10.14214/sf.394](https://doi.org/10.14214/sf.394)
- Castaño-Santamaría J, Crecente-Campo F, Fernández-Martínez JL, Barrio-Anta M, Obeso JR (2013). Tree height prediction approaches for uneven-aged beech forests in northwestern Spain. *Forest Ecology and Management* 307: 63-73. - doi: [10.1016/j.foreco.2013.07.014](https://doi.org/10.1016/j.foreco.2013.07.014)
- Crecente-Campo F, Tomé M, Soares P, Diéguez-

- Aranda U (2010). A generalized nonlinear mixed-effects height-diameter model for *Eucalyptus globulus* L. in northwestern Spain. *Forest Ecology and Management* 259: 943-952. - doi: [10.1016/j.foreco.2009.11.036](https://doi.org/10.1016/j.foreco.2009.11.036)
- De'ath G, Fabricius KE (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81: 3178-3192. - doi: [10.1890/0012-9658\(2000\)081\[3178:CARTAP\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2)
- Dobbertin M, Biging GS (1998). Using the non-parametric classifier CART to model forest tree mortality. *Forest Science* 44: 507-516. [online] URL: <http://www.ingentaconnect.com/content/saf/1998/00000044/00000004/art00005>
- Drápela K (2011). Regresní modely a možnosti jejich využití v lesnictví [The regression models and options of their utilization in the forestry]. Mendel University in Brno, Brno, Czech Republic, pp. 236. [in Czech]
- Erikäinen K (2003). Predicting the height-diameter pattern of planted *Pinus kesiya* stands in Zambia and Zimbabwe. *Forest Ecology and Management* 175: 355-366. - doi: [10.1016/S0378-1127\(02\)00138-X](https://doi.org/10.1016/S0378-1127(02)00138-X)
- Fan Z, Kabrick JM, Shifley SR (2006). Classification and regression tree based survival analysis in oak-dominated forests of Missouri's Ozark highlands. *Canadian Journal of Forest Research* 36: 1740-1748. - doi: [10.1139/x06-068](https://doi.org/10.1139/x06-068)
- Ferguson IS, Leech JW (1978). Generalized least squares estimation of yield functions. *Forest Science* 24: 27-42. [online] URL: <http://www.ingentaconnect.com/content/saf/1978/00000024/00000001/art00007>
- Fox JC, Ades PK, Bi H (2001). Stochastic structure and individual-tree growth models. *Forest Ecology and Management* 154: 261-276. - doi: [10.1016/S0378-1127\(00\)00632-0](https://doi.org/10.1016/S0378-1127(00)00632-0)
- Gómez C, Wulder MA, Montes F, Delgado JA (2012). Modeling forest structural parameters in the Mediterranean pines of central Spain using QuickBird-2 imagery and classification and regression tree analysis (CART). *Remote Sensing* 4: 135-159. - doi: [10.3390/rs4010135](https://doi.org/10.3390/rs4010135)
- Gregoire TG, Schabenberger O, Barrett JP (1995). Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent plot measurements. *Canadian Journal of Forest Research* 25: 137-156. - doi: [10.1139/x95-017](https://doi.org/10.1139/x95-017)
- Huang S, Price D, Titus SJ (2000). Development of ecoregion-based height-diameter models for white spruce in boreal forest. *Forest Ecology and Management* 129: 125-141. - doi: [10.1016/S0378-1127\(99\)00151-6](https://doi.org/10.1016/S0378-1127(99)00151-6)
- Kangas A, Haara A (2012). Comparison of non-spatial and spatial approaches with parametric and nonparametric methods in prediction of tree height. *European Journal of Forest Research* 131: 1771-1782. - doi: [10.1007/s10342-012-0631-8](https://doi.org/10.1007/s10342-012-0631-8)
- Legendre P (1993). Spatial autocorrelation: trouble or new paradigm? *Ecology* 74: 1659-1673. - doi: [10.2307/1939924](https://doi.org/10.2307/1939924)
- Levakovic A (1935). Analytical form of growth laws. *Glasnik za sumske pokuse* 4: 189-282.
- Lu J, Zhang L (2012). Evaluation of structure specification in linear mixed models for modeling the spatial effects in tree height-diameter relationships. *Annals of Forest Research* 56: 137-148. [online] URL: <http://afjournal.org/index.php/af/article/view/49>
- Meyer HA (1940). A mathematical expression for height curves. *Journal of Forestry* 38: 415-420. [online] URL: <http://www.ingentaconnect.com/content/saf/jof/1940/00000038/00000005/art00013>
- Mehntälö L (2004). A longitudinal height-diameter model for Norway spruce in Finland. *Canadian Journal of Forest Research* 34: 131-140. - doi: [10.1139/x03-207](https://doi.org/10.1139/x03-207)
- Michailov I (1943). Zahlenmäßiges Verfahren für die Ausführung der Bestandeshöhenkurven [Numerical estimation of stand height curves]. *Forstwissenschaftliches Centralblatt und Tharandter Forstliches Jahrbuch* 6: 273-279. [in German]
- Miguel S, Guzmán G, Pukkala T (2013). A comparison of fixed- and mixed-effects modelling in tree growth and yield prediction of an indigenous Neotropical species (*Centropomus tomentosum*) in a plantation system. *Forest Ecology and Management* 291: 249-258. - doi: [10.1016/j.foreco.2012.11.026](https://doi.org/10.1016/j.foreco.2012.11.026)
- Moisen GG, Frescino TS (2002). Comparing five modelling techniques for predicting forest characteristics. *Ecological Modelling* 157: 209-225. - doi: [10.1016/S0304-3800\(02\)00197-7](https://doi.org/10.1016/S0304-3800(02)00197-7)
- Näslund M (1947). Functions and tables for computing the cubic volume of standing trees: pine, spruce and birch in Southern Sweden and in the whole of Sweden. Reports of the Forest Research Institute of Sweden 36, Stockholm, Sweden, pp. 81.
- Petterson H (1955). Barrskogens volymproduktion [Volume production of coniferous forests]. *Meddelanden från Statens skogsforskningsinstitut, Stockholm, Sweden*, pp. 391. [in Swedish]
- Petráš R, Pajtík J (1991). Sústava česko-slovenských objemových tabuliek drevín [Volume tables for the Czech and Slovak Republic]. *Lesnický časopis - Forestry Journal* 37: 49-56. [in Slovak]
- Räty M, Kangas A (2008). Localizing general models with classification and regression trees. *Scandinavian Journal of Forest Research* 23: 419-430. - doi: [10.1080/02827580802378826](https://doi.org/10.1080/02827580802378826)
- R Core Team (2015). R: a language and environment for statistical computing. R foundation for Statistical Computing, Vienna, Austria. [online] URL: <http://www.R-project.org>
- Rejwan C, Collins LJ, Brunner LJ, Shuter BJ, Ridgway MS (1999). Tree regression analysis on the nesting habitat of smallmouth bass. *Ecology* 80: 341-348. - doi: [10.1890/0012-9658\(1999\)080\[0341:TRAOTN\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1999)080[0341:TRAOTN]2.0.CO;2)
- Robinson GK (1991). That BLUP is a good thing: the estimation of random effects. *Statistical Science* 6: 15-32. - doi: [10.1214/ss/1177011926](https://doi.org/10.1214/ss/1177011926)
- Schmidt M, Kiviste A, Von Gadow K (2011). A spatially explicit height-diameter model for Scots pine in Estonia. *European Journal of Forest Research* 130: 303-315. - doi: [10.1007/s10342-010-0434-8](https://doi.org/10.1007/s10342-010-0434-8)
- Sharma M, Parton J (2007). Height-diameter equations for boreal tree species in Ontario using a mixed-effects modeling approach. *Forest Ecology and Management* 249: 187-198. - doi: [10.1016/j.foreco.2007.05.006](https://doi.org/10.1016/j.foreco.2007.05.006)
- Shifley SR, Fan Z, Kabrick JM, Jensen RG (2006). Oak mortality risk factors and mortality estimation. *Forest Ecology and Management* 229: 16-26. - doi: [10.1016/j.foreco.2006.03.033](https://doi.org/10.1016/j.foreco.2006.03.033)
- Sironen S, Kangas A, Maltamo M, Kangas J (2003). Estimating individual tree growth with nonparametric methods. *Canadian Journal of Forest Research* 33: 444-449. - doi: [10.1139/x02-162](https://doi.org/10.1139/x02-162)
- Soares P, Tomé M (2002). Height-diameter equation for first rotation eucalypt plantations in Portugal. *Forest Ecology and Management* 166: 99-109. - doi: [10.1016/S0378-1127\(01\)00674-0](https://doi.org/10.1016/S0378-1127(01)00674-0)
- Trincado G, VanderSchaaf CL, Burkhart HE (2007). Regional mixed-effects height-diameter models for loblolly pine (*Pinus taeda* L.) plantations. *European Journal of Forest Research* 126: 253-262. - doi: [10.1007/s10342-006-0141-7](https://doi.org/10.1007/s10342-006-0141-7)
- Van Laar A, Akça A (2007). *Forest mensuration*. Springer, Dordrecht, Netherlands, pp. 389. [online] URL: <http://books.google.com/books?id=ybjdxL2LHOQC>
- Vonesh EF, Chiinchilli VM (1997). *Linear and nonlinear models for the analysis of repeated measurements*. Marcel Dekker, New York, NY, USA, pp. 560.
- Weiss J (2007). *Environmental studies 562: Statistics for Environmental Science*. Lecture notes, web site. [online] URL: <http://www.unc.edu/courses/2007spring/enst/562/001/docs/lectures.htm>
- Yang Y, Huang S (2011). Comparison of different methods for fitting nonlinear mixed forest models and for making predictions. *Canadian Journal of Forest Research* 41: 1671-1686. - doi: [10.1139/x11-071](https://doi.org/10.1139/x11-071)
- Zhang L, Bi H, Cheng P, Davis CJ (2004). Modeling spatial variation in tree diameter-height relationships. *Forest Ecology and Management* 189: 317-329. - doi: [10.1016/j.foreco.2003.09.004](https://doi.org/10.1016/j.foreco.2003.09.004)
- Zhang L, Gove JH, Heath LS (2005). Spatial residual analysis of six modeling techniques. *Ecological Modelling* 186: 154-177. - doi: [10.1016/j.ecolmodel.2005.01.007](https://doi.org/10.1016/j.ecolmodel.2005.01.007)
- Zhang L, Ma Z, Guo L (2008). Spatially assessing model errors of four regression techniques for three types of forest stands. *Forestry* 81: 209-225. - doi: [10.1093/forestry/cpn014](https://doi.org/10.1093/forestry/cpn014)