Canadian Journal of Forest Research

Evaluation of the mixed effects model and quantile regression approaches for predicting tree height in larch (*Larix olgensis*) plantations in northeastern China

Journal:	<i>Canadian Journal of Forest Research</i>
Manuscript ID	cjfr-2021-0184.R1
Manuscript Type:	Article
Date Submitted by the Author:	17-Sep-2021
Complete List of Authors:	Xie, Longfei; Northeast Forestry University Widagdo, Faris; Northeast Forestry University, School of Forestry Miao, Zheng; Northeast Forestry University Dong, Lihu; Northeast Forestry University Li, Fengri; Northeast Forestry University
Keyword:	<i>Larix olgensis</i> , Height-diameter model, Mixed-effects models, Quantile regressions, Chapman-Richards function
Is the invited manuscript for consideration in a Special Issue? :	Not applicable (regular submission)

SCHOLARONE™
Manuscripts

Evaluation of the mixed effects model and quantile regression approaches for predicting tree height in larch (*Larix olgensis*) plantations in northeastern China

Longfei Xie, Faris Rafi Almay Widagdo, Zheng Miao, Lihu Dong*, and Fengri Li*

Key Laboratory of Sustainable Forest Ecosystem Management-Ministry of Education, School of Forestry, Northeast Forestry University, Harbin 150040, Heilongjiang, China

*Correspondence: Lihu Dong (lihudong@nefu.edu.cn) and Fengri Li (fengrili@nefu.edu.cn)

Draft

Abstract Tree height (H) is one of the most important tree variables and is widely used in growth and yield models, and its measurement is often time-consuming and costly. Hence, height-diameter (H-D) models have become a great alternative, providing easy-to-use and accurate tools for H prediction. In this study, H-D models were developed for *Larix olgensis* in Northeast China. The Chapman-Richards function with three predictors (diameter at breast height, dominant tree height, and relative size of individual trees) performed best. Nonlinear mixed effects (NLME) models and nonlinear quantile regressions (NQR9, 9 quantiles; NQR5, 5 quantiles; and NQR3, 3 quantiles) were further used and improved the generalized H-D model, successfully providing accurate H predictions. In addition, the H predictions were calibrated using several measurements from subsamples, which were obtained from different sampling designs and sizes. The results indicated that the predictive accuracy was higher when calibrated by using any number of height measurements for the NLME model and more than 3 height measurements for the NQR3, NQR5 and NQR9 models. The best sampling strategy for the NLME and NQR models involved sampling the medium-sized trees. Overall, the newly developed H-D models can provide highly accurate height predictions for *L. olgensis*.

Keywords *Larix olgensis*; Chapman-Richards function; Height-diameter model; Mixed effects models; Quantile regressions

1. Introduction

Korean larch (*Larix olgensis*) is native to Northeast China and has strong adaptability, fast growth at a young age, and high resistance against insects and diseases; it is mainly distributed in northern China, particularly in Inner Mongolia and Heilongjiang Province (State Forestry and Grassland Administration 2019; Yu 2008). Korean larch has been extensively used for various afforestation programs since it is fast growing and produces economically valuable timber. Therefore, it is necessary to maximize the ecological and economic benefits of this species by providing various management options based on a series of tree and stand attributes (Scolforo et al. 2019). Among the attributes, diameter at breast height (*DBH*) and total tree height (*H*) are two of the most important and commonly used variables in forest inventories since tree taper, volume, site index, and biomass are mostly determined by *DBH* and *H* (Cao and Wang 2015; Dong et al. 2020; Trim et al. 2020; Yang and Burkhardt 2020).

H can be measured artificially or extracted from light detection and ranging (LIDAR) data, which are costly and require time-consuming and labor-intensive analysis (Zang et al. 2016). Furthermore, the measurement of *H* is heavily affected by visual obstructions, which may result in higher bias than that in *DBH* measurements (Castano-Santamaria et al. 2013). Therefore, it is preferable to measure the *D* of all the trees and the *H* of some subsample trees within a forest inventory plot, as these data can be used to construct a local height-diameter (H-D) model. However, studies have found that the H-D relationship varies spatially, leading the local H-D models to have low applicability (Temesgen et al. 2008; Zheng et al. 2018). A generalized H-D model was found to be more effective and has higher applicability than local H-D models (Crecente-Campo et al. 2010; Zang et al. 2016). In generalized H-D models, stand-level variables need to represent the variation among stands, such as the dominant tree height (Özçelik et al. 2018; Sharma et al. 2018), quadratic mean diameter (Bronisz and Mehtatalo 2020), and basal area per hectare (Temesgen et al. 2008). Among these variables,

dominant tree height is the most commonly used stand variable, as it makes a more considerable contribution to H-D models, which describe the incorporative effects of stand development and site quality (Monserud 1984; Sharma et al. 2016; Sharma et al. 2018). In addition, competition indices are sometimes added to local H-D models to improve accuracy (Zang et al. 2016). To date, numerous H-D models have been developed for different species and areas worldwide (Huang et al. 1992; Sharma et al. 2018; Temesgen et al. 2008).

Various linear and nonlinear models have been used to simulate the H-D relationship (Adamec and Drápela 2017; Cysneiros et al. 2020), and the model parameters are generally estimated by ordinary least squares (OLS). Height and diameter data are usually collected from different stands, and unfortunately, the OLS approach was found to be inadequate for explaining the variations between stands (Pinheiro and Bates 2000; Zheng et al. 2018). For this reason, nonlinear mixed effects (NLME) models have gradually become the most popular method to construct H-D models (Bronisz and Mehtatalo 2020; Castano-Santamaria et al. 2013; Mehtatalo 2004). NLME models, which are considered an efficient tool for analyzing group data, provide a framework for evaluating the correlation structures of groups (Mehtatalo and Lappi 2020). Research has shown that the fit and performance of NLME models are better than those of OLS models (Crecente-Campo et al. 2014; Zang et al. 2016).

Nonlinear quantile regression (NQR), proposed by Koenker and Bassett (1978), is considered to be a flexible method to describe varying patterns of the relationship between dependent and independent variables when compared with the mean regression. It is aimed to estimate the complete conditional distribution of dependent variable and assess the effect of the independent variable at different quantiles (Özçelik et al. 2018). With the curves of different quantiles, the H-D curve of a specific plot can be obtained using two adjacent quantile curves, which could provide more accurate H predictions. As a result, each plot has its H-D curve, which is similar to the NLME approach. The calibration of NQR has been used

in other studies of tree diameter growth (Bohora and Cao 2014) and taper equation (Cao and Wang 2015), Bohora and Cao (2014) and Cao and Wang (2015) have used the calibration of NQR to develop tree diameter growth model and taper equation although the calibration was achieved by using one pre-measured observation, which is evidence that the calibration of NQR could provide an accurate prediction. In terms of modeling H-D model, Zang et al. (2016) and Zhang et al. (2020) ignored the difference between plots and used the NQR under specific quantile to predict H of new trees, which produced greater bias when compared with NLEM. In order to simulate the plot-specific H-D curve through different quantiles curves, Özçelik et al. (2018) proposed the new procedures that using curves from different quantiles to predict tree height from diameter and dominant tree height, solving the problem that the calibration of NQR in previous studies can only use one observation. At present, the calibration of NQR has not received widespread attention and use, and more data is needed to verify this method. In addition, the sampling strategies of samples used for calibration needs further testing.

Despite all their advantages, both NLME and NQR models need to be calibrated to be most accurate (Özçelik et al. 2018). For H-D NLME models, the random effect parameters are estimated for each group and calibrated with a subsample of H measurements (Mehtätalo and Lappi 2020). The predictive performance varies with the number of subsample measurements (Castano-Santamaria et al. 2013). The NQR approach also requires subsample measurements to construct a new H-D curve for each group and does so by combining two known curves of specific quantiles (Özçelik et al. 2018). If a subsample is not available, uncalibrated H predictions are allowed, consequently yielding a larger deviation.

The main objectives of this study are to (1) find the best function to describe the relationship between height and diameter; (2) explore the advantages and disadvantages of the NLME and NQR approaches for H-D models; (3) determine the best sampling designs for the

NLME and NQR approaches; and (4) determine an appropriate number of subsample measurements that considers both sampling cost and predictive accuracy.

2. Material and methods

2.1. Data

The data for this study were sampled from differently aged (8 ~ 64 year) pure plantations of Korean larch in Heilongjiang Province (from 121°11' E to 135°05' E and from 43°26' N to 53°33' N), Northeast China, which is among the provinces in China with the largest amount of forestry and has large areas planted with Korean larch (State Forestry and Grassland Administration 2019). Heilongjiang Province has a temperate continental monsoon climate, with spring seasons characterized by low temperatures and drought, summer seasons by warm temperatures and rain, autumn seasons by waterlogging and early frost, and winter seasons by a long cold period. The altitude and mean annual temperature range from 50 m to 1000 m above sea level and 3.3 °C, respectively. The average annual precipitation is 525 mm. A total of 524 plots were established within the last 30 years (1991–2019), and the plots ranged from 0.03 ha to 0.20 ha. As the GPS coordinates were not recorded for some of the plots measured earlier, a figure showing the locations of all the plots is not provided. In each sample plot, several basic attributes of all living trees, such as the *DBH* and total tree height (*H*), were measured and recorded. The *DBH* and *H* were measured with a diameter tape and a hypsometer (Vertex IV Ultrasonic by Haglöf Sweden), respectively. In total, 59,907 sample trees were used to develop the H-D model for Korean larch. The basic summary information and the relationship between *DBH* and *H* are given in Table 1 and Fig. 1.

2.2. Methods

2.2.1. Specifications of the generalized H-D model

Sixteen equations from several previous studies (Huang et al. 1992; Mehtatalo 2019; Zeide 1993) were used to model the H-D relationship for Korean larch. All candidate

equations were independently fitted to the whole dataset, and the equation with the best fit was determined by comparing three fitting statistics, i.e., the adjusted coefficient of determination (R_a^2), root mean square error (RMSE) and Akaike information criterion (AIC), which correspond to Equations 1, 2 and 3, respectively (Crecente-Campo et al. 2014). In addition, the predictive ability of the candidate equations was compared by using leave-one-out cross validation. However, all the candidate models yielded very similar fitting and validation performances, although slightly better statistical values were found for the Korf function (Equation 4). It was difficult to choose a suitable basic model of the H-D relationship for Korean larch from these statistics, and a comparison of generalized H-D models that included stand variables and competition indexes was performed for the five most suitable equations according to the fitting statistics. Finally, the generalized Chapman-Richards equation (Equation 5) resulted in the highest R_a^2 and the lowest RMSE and AIC values, which showed that this equation is much more flexible than previously reported (Burkhart and Tomé 2012). Hence, the Chapman-Richards equation was chosen as the basic model of the H-D relationship of Korean larch.

$$R_a^2 = 1 - \left(\frac{n-1}{n-p} \right) \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (1)$$

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

$$AIC = -2LL + 2p \quad (3)$$

$$H = 1.3 + \beta_1 e^{-\beta_2 DBH^{-\beta_3}} + \varepsilon \quad (4)$$

$$H = 1.3 + \beta_1 (1 - e^{-\beta_2 DBH})^{\beta_3} + \varepsilon \quad (5)$$

where n is the sample size; p is the number of model parameters; LL is the log likelihood value of the equations; y_i , $\hat{y}_i$ and $\bar{y}_i$ are the observed, predicted and mean observed tree heights, respectively; and ε is the error term.

The generalized H-D (GHD) model was constructed by introducing several stand-level variables as covariates to the basic model to diminish variability among sample plots (Zheng et al. 2018). In this study, the commonly used stand variables, such as stand basal area (BA , $\text{m}^2 \cdot \text{ha}^{-1}$), quadratic mean diameter (Dq , cm), average height of dominant trees (H_{dom} , m; the dominant trees were identified as the 100 trees per hectare with the largest DBH s), average diameter of dominant trees (D_{dom} , cm), stand age (Age , a), and stand density (SD , $\text{trees} \cdot \text{ha}^{-1}$, i.e., the number of trees per hectare), were utilized as candidate variables (Sharma et al. 2018). In addition, the basal area of trees with a diameter larger than that of the subject tree (BAL , $\text{m}^2 \cdot \text{ha}^{-1}$) and relative size in the stand (Rd , the ratio of DBH to Dq) were calculated as competition measures to evaluate the influence of neighboring trees (Sharma et al. 2018; Zang et al. 2016).

GHD models can generally provide predictions of H in a new plot k without any prior tree data. However, the predictions obtained in this way often have a relatively large bias. Temesgen et al. (2008) reported that tree height can be calibrated with a correction factor (CF) in GHD models, resulting in a smaller bias than in corresponding models without calibration. The detailed form of the CF was calculated by the following expression:

$$CF = \frac{\sum_{i=1}^{n_k} (\hat{y}_i - 1.3) \times (y_i - 1.3)}{\sum_{i=1}^{n_k} (\hat{y}_i - 1.3)^2} \quad (6)$$

where $\hat{y}_i$ is the H prediction calculated by the GHD model, y_i is the H observation of the i th tree in plot k , and n_k is the size of the measured subsample in plot k . Using the CF, the calibrated H predictions are

$$H = 1.3 + CF \times f(x) \quad (7)$$

where $f(x)$ is the prediction of the GHD model.

2.2.2. Nonlinear mixed effects model

The mixed effects model has been widely used for nested data structures and has been found to be an effective technique with high transferability and accuracy (Calama and Montero 2004; Ciceu et al. 2020; Crecente-Campo et al. 2010; Xie et al. 2020). The H-D data were obtained based on measurements from different sample plots, forming hierarchical data, which is very suitable for NLME models. A general one-level NLME model can be expressed as follows (Pinheiro and Bates 2000):

$$y_i = f(\Phi_i, x_i) + \varepsilon_i, i = 1, 2, \dots, m \quad (8)$$

$$\Phi_i = A_i\beta + B_ib_i \quad (9)$$

where M is the number of plots, y_i is the vector of H observations from the i th sample plot ($i = 1, 2, \dots, m$), and m is the number of sample plots. f is a nonlinear function of predictor variables x_i and a plot-specific parameter vector Φ_i , and ε_i is the vector of the within-plot error term in the i th plot. β is the fixed-effect parameter vector; b_i is a random effect parameter vector associated with the i th plot; and A_i and B_i are matrices of appropriate dimensions that depend on the plot, particularly the plot-specific values of some covariates.

ε_i and b_i follow a normal distribution:

$$b_i \sim N(0, D) \quad (10)$$

$$\varepsilon_i \sim N(0, \sigma^2 G_i^{0.5} \Gamma_i G_i^{0.5}) \quad (11)$$

where D is the variance–covariance matrix of random effect parameters; σ^2 is the residual variance; the matrix Γ_i accounts for within-plot autocorrelations of the residual errors and is considered an identity matrix; and the matrix G_i accounts for within-plot variance heteroscedasticity, and its diagonal elements are provided by the variance function. Within-plot heteroscedasticity in the residuals usually appears in the H-D model. Hence, the power variance function (Equation 8) is added into the model, which is the most commonly used variance function (Sharma et al. 2018).

$$\text{var}(\varepsilon_i) = \gamma \hat{h}_i^{2\delta} \quad (12)$$

where ε_i and $\hat{h}_i$ are the residuals and estimated H s calculated by fixed-effect parameters for the i th sample plot, respectively; γ and δ are the estimated parameters for the power function.

At least one measurement of H within each plot should be available when using the NLME model to make predictions. With the observed value of H in the new plot k , the random effect parameters could be calculated by the following equation (Pinheiro and Bates 2000):

$$b_k = DZ_k^T(Z_k DZ_k^T + R_k)^{-1}e_k \quad (13)$$

where R_k is the variance–covariance matrix of residuals for plot k ; Z_k is the matrix of the partial derivatives of the nonlinear function with respect to its random parameters; and e_k is the vector of residuals taken from plot k using the fixed-effects parameters of the mixed effects models.

2.2.3. Nonlinear quantile regression

NQR can predict H with different quantiles (q), and the general expression is:

$$\hat{y}_q = f(x, \beta_q) + \varepsilon_q \quad (14)$$

where $\hat{y}_q$ is the estimated H in the q th quantile; q is the quantile, and the values are 0.1, 0.2, 0.3, ..., 0.9; f is a nonlinear function of predictor variables, including DBH and other stand variables and parameters, denoted by β ; and ε_q is the residual. The parameters in Equation 14 with q th quantiles were estimated by minimizing the following equation (Bohora and Cao 2014):

$$S = \sum_{y_q \geq \hat{y}_q} q(y_q - \hat{y}_q) + \sum_{y_q < \hat{y}_q} (1 - q)(\hat{y}_q - y_q) \quad (15)$$

where S is the sum of the weighted absolute residuals of the q th quantile and y_q is the observed H in the q th quantile.

Three types of NQR models were constructed based on the combination of three quantiles (NQR3, $q = 0.1, 0.5, 0.9$), five quantiles (NQR5, $q = 0.1, 0.3, 0.5, 0.7, 0.9$) and nine quantiles (NQR9, $q = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$). Similar to predictions made with the NLME model approach, observed values of H are also needed for plot k in the NQR model to identify the H-D curve that passes through the two closest quantile regression curves (q th and $(q+1)$ th). In cases in which there was at least one H measured for plot k , the two closest quantile regression curves were determined by the absolute mean difference between the observed H and q -th predicted H of the trees measured in subsample plot k . In detail, the quantiles of the quantile regression curve that corresponded to the two smallest values of absolute mean difference are the required quantiles. Once the closest q -th and $(q+1)$ -th quantile curves were determined, a modified value of H was calculated by interpolation:

$$\hat{y}_{ki} = \alpha \hat{y}_q + (1 - \alpha) \hat{y}_{q+1} \quad (16)$$

where $\hat{y}_{ki}$ is the estimated H of the i th tree in plot k ; α is the interpolation ratio, which is obtained by minimizing $\sum_{i=1} (y_{ki} - \hat{y}_{ki})^2$; and y_{ki} is the observed H of the i th tree in plot k .

2.2.4. Model validation and evaluation

The leave-one-out cross-validation (LOOCV) method was used to validate the developed models in this study; this approach is widely used and considered a useful method to test the predictive performance of a model (Cawley and Talbot 2003). The specific steps of the LOOCV applied in this study were as follows: (1) the sampled plots were marked from 1 to 524; (2) the subdataset that contained plot 1 was selected as the validation dataset and the remaining datasets were used for model fitting; (3) based on the fitting result, the tree heights in plot 1 were predicted; (4) the plot number was modified, and steps (2) and (3) were repeated. Finally, the LOOCV predictions of tree heights in each sampled plot were obtained. Several statistics of mean absolute error (MAE, Equation 17), mean percent error (MPE,

Equation 18), mean absolute percent error (MAPE, Equation 19) and fit index (FI, Equation 20) were calculated to assess and compare the predictive performances of the models.

$$MPE = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{y}_{ij,-j}) / y_{ij}}{\sum_{i=1}^m n_i} \times 100 \quad (17)$$

$$MAE = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} |y_{ij} - \hat{y}_{ij,-j}|}{\sum_{i=1}^m n_i} \quad (18)$$

$$MAPE = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} |y_{ij} - \hat{y}_{ij,-j}| / y_{ij}}{\sum_{i=1}^m n_i} \times 100 \quad (19)$$

$$FI = 1 - \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{y}_{ij,-j})^2}{\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{ij,-j})^2} \quad (20)$$

where m is the number of sampled plots, n_i is the number of trees in the i th sampled plot, and y_{ij} and $\hat{y}_{ij,-i}$ are the observed H and the predicted H obtained by LOOCV, respectively.

2.2.5. Model performance for different sampling designs and sizes

When predicting tree height using NLME and NQR, new subset data, including the values of response and explanatory variables, are necessary for calibrating the predictions. For the GHD, NLME and NQR models, higher prediction accuracy was generated as the number of trees with measured H values per plot increased (Bohara and Cao 2014; Zheng et al. 2018; Özçelik et al. 2018). In addition, previous studies found that the prediction accuracy of H varied according to the number of trees sampled per plot (Crecente-Campo et al. 2010; Bronisz and Mehtatalo 2020). Hence, in this section, we assessed and compared the predicted performances of the developed H-D models with observed data from several trees and with data collected by different sampling designs per sampled plot. Specifically, six types of DBH -based sampling designs were evaluated for various l sample sizes ($l = 2, 3, \dots, 10$):

Type I: Selecting l trees randomly per plot.

Type II: Selecting l thickest trees per plot.

Type III: Selecting l thinnest trees per plot.

Type IV: Selecting l medium-sized trees (closest to the quadratic mean DBH) per plot.

Selecting $\frac{l}{2}$ thickest trees and $\frac{l}{2}$ thinnest trees per plot, where l is an even

Type V: number, and selecting $\frac{l-1}{2}$ thickest trees and $\frac{l-1}{2}$ thinnest trees plus adding one thinnest tree per plot when l is an odd number.

Type VI: Similar to type V, but adding one thickest tree per plot instead of one thinnest tree per plot when l is an odd number.

All calibrations were performed in R software 4.0.3 (R Core Team 2020). The first type of sampling design was repeated 500 times due to the effects of random sampling, while the remaining sampling designs were simulated once. MAE, MPE, MAPE and FI were used to evaluate the predicted performances of the models for different sampling designs and sizes per plot.

3. Results

3.1. Generalized height-diameter model

Using the reparameterization method and based on the introduction of candidate variables associated with stands and competition, the analysis results of the plot-specific Chapman-Richards function parameters suggested that the asymptote parameter should be expanded to a linear equation of H_{dom} , and the third parameter, β_3 , can be expressed as a linear equation of Rd . Thus, the GHD model, the mathematical expression of which is shown in Equation 21, was developed. As heteroscedasticity in residuals widely existed in the H-D model, the GHD model was refitted by weighted OLS. The estimates of parameters and residual variance are given in Table 2; all the estimates were statistically significant at a significance level of 5%. In addition, the predictive performances of the GHD models are listed in Table 3, and the difference between the uncalibrated and calibrated cases was clear

(the MAE and MAPE decreased by approximately 11% and 12%, respectively, and the FI increased by 1.48%).

$$H = 1.3 + (\beta_1 + \beta_2 H_{dom}) \times (1 - e^{-\beta_3 DBH})^{(\beta_4 + \beta_5 Rd)} \quad (21)$$

where $\beta_1 \sim \beta_5$ are model parameters, and the other definitions are the same as above.

3.2. Mixed effects height-diameter model

In the process of developing the mixed effects height-diameter model, the AIC and log likelihood values were compared for different combinations of random effect parameters to determine whether the parameters included the random effect parameters. However, the more random effect parameters that were included in the NLME models, the harder it was for the NLME models to converge, and Table 4 shows the fitting results for only the converged models. Please note that the variance-covariance matrix of random effect parameters was set to a diagonal matrix structure to speed up the calculation. As a result, the parameters β_1 and β_3 were identified as having both fixed- and random-effect components that were common to all trees and specific to trees within the same sample plot, respectively. In addition, the variance-covariance matrix of random effect parameters was set to a general positive-definite matrix structure for comparison, and the latter was proven to be more appropriate since it result in smaller AIC values.

Fig. 2 shows the residual plots against the fitted tree heights before and after weighted regression; this result indicated that the weighted regression improved the homogeneous variance by using a power function as a weighted function. Estimates of the fixed-effect parameters and residual variance and the variance-covariance matrix of random effect parameters are listed in Table 2. The estimated parameters were all statistically significant at a significance level of 5%. To evaluate the predictive performance of the NLME model, we used the whole dataset to calculate the random effect parameters of each sample plot by LOOCV. The results showed that NLME resulted in a small bias in H predictions, having

MAE, MPE, MAPE and FI values of only 0.9670 m, 1.0357%, 7.4845% and 0.9392, respectively (Table 3). However, greater bias was found for the uncalibrated NLME model (the corresponding statistics were 1.1046 m, -1.5110%, 8.8290% and 0.9223, respectively). The MAE and MAPE of the GHD model were 1.0851 m and 8.6548%, respectively, which indicated that the predicted bias for the individual GHD model was larger than that for the calibrated NLME model, while the GHD model had smaller biases on average since the corresponding MPE was only 0.0019%. The comparison of the four statistics indicated that the introduction of plot-level random effects improved the performance of the GHD model.

3.3. Nonlinear quantile regression height-diameter model

Table 2 shows the estimates of generalized H-D model parameters obtained in the q th quantile ($q=0.1, 0.2, \dots, 0.9$) from the NQR. The estimates of the parameters were all statistically significant at a significance level of 5%. The statistics representing prediction performance were calculated for uncalibrated and calibrated models created with all the datasets (Table 3). Please note that the parameter estimates in the 0.5 quantile of the NQR were used for predicting tree heights when the predictions were uncalibrated. As a result, the calibrated NQR3, NQR5 and NQR9 models had lower values of MAE (0.9551~0.9562 m), MPE (-0.0114~0.0052%), MAPE (7.5117~7.5222%) and higher values of FI (0.9414~0.9415) than the uncalibrated NQR model. In addition, Table 3 shows that only a slight difference was detected between the three types of quantile combinations, and NQR5 performed better than the other two models since the MAE and MAPE of NQR5 were the smallest, the MPE was closest to 0 and the FI was highest.

3.4. Influences of the sampling designs and sizes

Fig. 3 shows different responses of the developed models (i.e., GHD, NLME, NQR3, NQR5, and NQR9) to several calibrated patterns of sampling design and size. Since random sampling (type I) results in large variation, the error bars of the maximum and minimum

values of 500 simulations are also drawn, and the performance of the random sampling approach was proven to be unstable (Fig. 3). In addition to the random sampling for the GHD, NLME, NQR3, NQR5 and NQR9 models, the type II sampling design produced the worst MAE, MPE, MAPE and FI values for all the sample sizes, while type IV outperformed all of the other sampling designs used in this study. Compared with the case where the sample size was zero, namely, the uncalibrated model, the GHD and NLME models performed better based on the MAE, MPE, MAPE and FI values, whereas the calibrated NQR3, NQR5 and NQR9 models produced better statistical values than the corresponding uncalibrated models only when three trees per sample plot were measured. According to the change in MAE, MPE, MAPE and FI values as the number of sampled trees per plot increased (Table 5), the prediction accuracy increased for all the NQR models when the sample size was more than 3, while NLME provided a prediction accuracy that first increased when the sample size was smaller than 7 and then decreased when the sample size was more than 7. For the NQR methods, using 2 trees for calibration unexpectedly yielded larger MAE, MPE, MAPE and FI values than the uncalibrated predictions (sample size of 0), and the values of these statistics were different from those of the GHD and NLME models. Since the MPEs in Table 5 were all very close to zero, the prediction accuracies of different modeling approaches were compared by MAE, MAPE and FI values calculated for subsamples of different sizes and sampling designs that selected medium-sized trees. The changes in MAE, MAPE and FI (Table 5) indicated that the NLME model performed better when the subsample sizes were approximately less than 8 and that the NQR9 model performed better as the subsample sizes increased.

3.5. Effects of different estimation approaches on individual trees and plots

As the GHD models provided H predictions with relatively more bias when the sizes of the measured subsample were small, the next analysis did not consider the GHD models.

Boxplots of MPE and MAPE across *DBH* classes (taking intervals of 10 cm as a class) were used to compare models constructed by the NLME, NQR3, NQR5, and NQR9 methods in terms of estimation accuracy and stability (Fig. 4 for tree-level samples). It was found that variations in MPE and MAPE decreased with increasing tree size, which suggested a larger bias for small trees ($DBH < 10$ cm) and a smaller bias for large trees ($DBH \geq 40$ cm), as shown in Fig. 4. In the comparison of the four modeling methods, the MPE of the NQR models was slightly closer to 0 when the *DBH* was less than 30 cm. The change in the MAPE was similar to that in the MPE, which indicated that the NQR models outperformed the others slightly when predicting the heights of individual trees.

In addition, the predictive accuracy at the plot level according to the MPE and MAPE statistics is shown in Fig. 5. Overall, there was a similar range of variation for each sample plot's MAPE with different methods, whereas the range of MPE varied across the modeling methods in the following order: NQR3 < NQR5 < NQR9 < NLME (Fig. 5). The mean and median values of MPE obtained by the NQR approach were closer to 0 than those obtained by the NLME approach. Overall, all the methods generally produced small biases across whole plots and had relatively good predictive performances, where the NQR models seemed to be better than the NLME models in terms of plot-level prediction.

4. Discussion

4.1. Basic height-diameter model and generalized model

The height-diameter relationship is useful for forest managers, such as for estimating tree- and stand-level volumes and quantifying forest site productivity (Ciceu et al. 2020; MacPhee et al. 2018). Hence, determining the most appropriate modeling approaches to develop H-D models is crucial for providing accurate *H* predictions in forest plantations. Thus, we explored the advantages and disadvantages of two approaches for creating H-D models of planted Korean larch based on a large dataset. In this study, sixteen commonly used

functions (e.g., Naslund, Power, Wykoff, Logistic, Chapman-Richards, Korf) with two or three parameters were fitted separately, and the fitting performances of the models were then compared. The differences between the assessed basic models in terms of fitting performance and prediction ability was negligible, and the five functions with the best performances according to the statistics of R_a^2 , RMSE and AIC were used to perform a further comparison of generalized H-D models. The results indicated that the Chapman-Richards function that included stand variables and competition outperformed the five selected models, as it resulted in better R_a^2 , RMSE and AIC values. Similar to the study by Zang et al. (2016), the generalized Chapman-Richards function performed very well in modeling the H-D relationship in larch plantations. The Chapman-Richards functions have been widely used as H-D functions for many tree species and are considered a flexible function in modeling individual tree attributes (Burkhart and Tomé 2012; Sharma et al. 2018; Özçelik et al. 2018).

Previous studies proved that H and DBH relationships vary within stands due to existing differences, e.g., site conditions and competition status (Castedo-Dorado et al. 2006; Zheng et al. 2018). In this study, the asymptote parameter was best explained with the stand-related variable $Hdom$ since it can represent the degree of site quality; this finding was consistent with those of previous studies (Sharma et al. 2018; Zheng et al. 2018). The tree-level competition status that affects the H-D relationship was described by the relative size of a tree within a stand (Rd), and this variable was properly incorporated into the third parameter of the Chapman-Richards function. We also evaluated the effects of another competition index, BAL , and the stand-level competition status (BA), which were found to perform worse than Rd . Thus, only two stand variables ($Hdom$ and Rd) were incorporated into our generalized H-D model for Korean larch.

4.2. Advanced regression modeling

Based on Equation 21, both the NLME and NQR approaches were used to develop the advanced generalized H-D model. NLME and NQR considered the potential differences between plots and had advantages in dealing with group data compared with the OLS approaches. Among the two approaches, NLME is the most widely used method in terms of the H-D model, while NQR is less commonly used (Castano-Santamaria et al. 2013; Crecente-Campo et al. 2010; Özçelik et al. 2018). In our study, all the approaches performed very well, with the NLME models producing slightly better predictive statistics when the size of the subsample used for calibration was small (nearly less than 8), while the NQR9 approach had higher prediction accuracy when the size of the calibrated subsample was more than 8 (Table 5). However, this result was slightly inconsistent with the conclusions of Özçelik et al. (2018). Generally, quantile regression is a more flexible approach than NLME since the model parameters of different quantiles were different (Cao and Wang 2015). Thus, the plot-specific H-D curves were generated with different shapes. If the additional information is sufficient (namely, if there are more H measurements in a calibration plot), then the predictions generated by NQR may be more accurate. However, some parameters in the NLME model were considered fixed, and these parameters were kept at the default value so they remained consistent across plots. The NLME model with more random effect parameters required more calculation time and often failed to converge. In this sense, the NQR model, which obtained more stable estimates and rarely failed to converge, seemed to be simpler than the NLME model.

4.3. Sampling design and subsample size for calibration

Regardless of the approach used to develop the H-D model, the ultimate goal was to accurately predict the H of a new stand, saving the cost and time associated with investigation (Ciceu et al. 2020). Several studies have shown that not only the number of measured trees but also their diameter affect prediction accuracy (Crecente-Campo et al. 2010). Current

studies have reported that selecting extreme trees appears to be the most efficient sampling design for calibrating plot-specific NLME models (Bronisz and Mehtatalo 2020), while selecting the smallest trees yields the best performance when calibrating an H-D model that includes the dominant tree height (Crecente-Campo et al. 2010). In this study, different results were obtained, in which medium-sized trees (i.e., those closest to the quadratic mean DBH) were selected per plot, namely, type IV was found to be the best sampling design for the GHD and NLME approaches, although the differences among the sampling designs was small. The type IV sampling design provided better MAE, MPE, MAPE and FI values (Fig. 3), which suggested that heights corresponding to the trees of average size provided more additional information than those corresponding to the largest and smallest trees for calibrations in our study. In other words, the tree heights in the sampled plots were relatively concentrated, and it was most beneficial to use the average size of trees to calculate the random effect parameters of each sample plot.

In the calibration of the NQR models, the type IV sampling design was also determined to be the most efficient strategy, even though the performances of all the sampling designs differed only slightly, which suggested that the use of trees of an average size could more accurately determine the two quantile curves closest to the H-D curve of the new stand in the calibration process of NQR. This finding may be attributed to the fact that there are more trees of average height and fewer trees of extreme height per plot. Among all the sampling designs, type II produced the worst MAE, MPE, MAPE and FI values since the dominant tree height was considered to be a fixed variable in the H-D model; this result was similar to that reported for other studies (Castedo-Dorado et al. 2006; Crecente-Campo et al. 2010).

In practice, obtaining a large sample of tree height measurements is not generally justifiable because of the high cost of sampling. In this sense, determining the appropriate number of trees to measure appears to be essential, directly affecting the investigation's cost

and time. Generally, 4~9 sample trees have been used for calibration to maintain model predictability while reducing the costs and time associated with data collection (Adamec and Drápela 2017; Sharma et al. 2018). Based on the type IV sampling design, our results confirmed that the prediction accuracy of the calibrated NLME model first increased and then decreased with an increasing number of measured subsample trees, while continuing to increase for calibrated NQR models when the size of the measured subsample was not more than 10. To select the appropriate sample size for calibration, the calibrated H predictions for measured subsamples of different sizes were compared by a paired T test at a significance level of 5%. The results of all the paired T tests showed that there was a significant difference between the H predictions for different measured subsample sizes. Hence, in this study, seven sample trees was the most appropriate for calibrating the H predictions by NLME, as this number of trees resulted in the best MAE, MAPE and FI values. Interestingly, for the three types of NQR models, the prediction accuracy was worse when the size of the measured subsample was two or three than when the model was uncalibrated; this finding was different from that for the NLME approach (Fig. 3). Considering that NQR9 provided better predictive performance than the NLME model when the size of the subsample was more than 8, we recommend choosing a suitable approach to estimate tree height based on the subsample size used for calibration, namely, using the NLME approach when fewer than 7 tree measurements are available and using the NQR9 approach when greater numbers of tree measurements are available.

4.4. Comparison with an existing height-diameter model

Zang et al. (2016) developed H-D models for planted larch in northern and northeastern China and found that a mixed effects Chapman-Richards model with only DBH performed best. However, the larch species studied by Zang et al. (2016) included not only *L. olgensis* but also *Larix gmelinii*, *Larix kaempferi* and *Larix principis-rupprechtii*. We predicted the

tree heights in our dataset according to the H-D model and conducted an individual comparison with the newly developed H-D model in this study. As the best model in Zang et al. (2016) was a basic mixed effects model, six sampling designs were applied to calculate the random effect parameters for their model, and the type IV sampling design outperformed the others. Fig. 6 shows the comparison of H predictions between the newly developed models in this study and the predictions of Zang et al. (2016). For practicality, all the models were fitted using measurements of ten medium-sized trees per plot (type IV sampling design). The model by Zang et al. (2016) was found to slightly overestimate the H values of small trees and gradually shifted to underestimate the H values as the DBH increased. In contrast, the newly developed H-D models consistently provided more accurate H predictions. The basic mixed effects model in Zang et al. (2016) is inadequate for explaining plot-level differences when the random effect parameters were calculated by using few measurements of subsample trees (e.g., 10 trees). Although a large number of tree measurements may have improved the predictive accuracy of the H-D model of Zang et al. (2016), the time and cost of field work would have increased. Thus, the newly developed models performed better than the previous model developed by Zang et al. (2016).

5. Conclusion

To develop a sustainable forest management practice, the status and dynamics of the forest should be deeply understood; hence, various models are needed to reveal the forest characteristics. The H-D model is one of the important basic models in forestry. In this study, mixed effects and NQR H-D models were developed using DBH , H_{dom} , and R_d as predictors, providing an accurate tree height prediction for Korean larch in Northeast China. All the approaches performed well and were recommended for use. Both sampling designs and sizes can affect the prediction accuracy of the three approaches. The selection of medium-sized trees (i.e., the type IV sampling design) was the most suitable approach for the GHD, NLME

and three types of NQR models. The determination of the subsample size differed among the modeling approaches, i.e., if the subsample size was less than seven, the NLME model was preferred, whereas the NQR9 model seemed to perform better with larger subsample sizes.

Acknowledgements

This research was financially supported by the National Key R&D Program of China (No. 2017YFD0600402), the Fundamental Research Funds for the Central Universities (No. 2572020DR03), Provincial Funding for National Key R&D Program of China in Heilongjiang Province (No. GX18B041), and the Heilongjiang Touyan Innovation Team Program (Technology Development Team for High-efficient Silviculture of Forest Resources). The authors would also like to thank the faculty and students of the Department of Forest Management, Northeast Forestry University (NEFU), China, who collected and provided the data for this study.

References

- Adamec, Z. and Drápela, K. 2017. Comparison of parametric and nonparametric methods for modeling height-diameter relationships. *iForest* **10**(1): 1-8. doi:10.3832/ifor1928-009.
- Bohora, S.B. and Cao, Q.V. 2014. Prediction of tree diameter growth using quantile regression and mixed-effects models. *For. Ecol. Manage.* **319**: 62-66. doi:10.1016/j.foreco.2014.02.006.
- Bronisz, K. and Mehtatalo, L. 2020. Mixed-effects generalized height-diameter model for young silver birch stands on post-agricultural lands. *For. Ecol. Manage.* **460**. doi:10.1016/j.foreco.2020.117901.
- Burkhardt H.E., Tomé M. 2012. Modeling forest trees and stands. Springer, New York. 458 pp.

- Calama, R. and Montero, G. 2004. Interregional nonlinear height-diameter model with random coefficients for stone pine in Spain. *Can. J. For. Res.* **34**(1): 150-163. doi:10.1139/x03-199.
- Cao, Q.V. and Wang, J. 2015. Evaluation of Methods for Calibrating a Tree Taper Equation. *For. Sci.* **61**(2): 213-219. doi:10.5849/forsci.14-008.
- Castano-Santamaria, J., Crecente-Campo, F., Luis Fernandez-Martinez, J., Barrio-Anta, M., and Ramon Obeso, J. 2013. Tree height prediction approaches for uneven-aged beech forests in northwestern Spain. *For. Ecol. Manage.* **307**: 63-73. doi:10.1016/j.foreco.2013.07.014.
- Castedo-Dorado, F., Dieguez-Aranda, U., Barrio-Anta, M., Sanchez-Rodriguez, M., and von Gadow, K. 2006. A generalized height-diameter model including random components for radiata pine plantations in northwestern Spain [Article]. *For. Ecol. Manage.* **229**(1-3): 202-213. doi:10.1016/j.foreco.2006.04.028.
- Cawley, G.C. and Talbot, N.L.C. 2003. Efficient leave-one-out cross-validation of kernel Fisher discriminant classifiers. *Pattern Recognit.* **36**(11): 2585-2592. doi:10.1016/s0031-3203(03)00136-5.
- Ciceu, A., Garcia-Duro, J., Seceleanu, I., and Badea, O. 2020. A generalized nonlinear mixed-effects height-diameter model for Norway spruce in mixed-uneven aged stands. *For. Ecol. Manage.* **477**. doi:10.1016/j.foreco.2020.118507.
- Corona, P., Antonio Marziliano, P., and Scotti, R. 2002. Top-down growth modelling: a prototype for poplar plantations in Italy. *For. Ecol. Manage.* **161**(1): 65-73. doi:10.1016/S0378-1127(01)00481-9.
- Crecente-Campo, F., Tome, M., Soares, P., and Dieguez-Aranda, U. 2010. A generalized nonlinear mixed-effects height-diameter model for *Eucalyptus globulus* L. in northwestern Spain. *For. Ecol. Manage.* **259**(5): 943-952.

doi:10.1016/j.foreco.2009.11.036.

- Crecente-Campo, F., Corral-Rivas, J.J., Vargas-Larreta, B., and Wehenkel, C. 2014. Can random components explain differences in the height–diameter relationship in mixed uneven-aged stands? *Ann. For. Sci.* **71**(1): 51-70. doi:10.1007/s13595-013-0332-6.
- Cysneiros, V.C., Pelissari, A.L., Gaui, T.D., Fiorentin, L.D., Carvalho, D.C.d., Silveira Filho, T.B., and Machado, S.d.A. 2020. Modeling of tree height–diameter relationships in the Atlantic Forest: effect of forest type on tree allometry. *Can. J. For. Res.* **50**(12): 1289-1298. doi:10.1139/cjfr-2020-0060.
- Dong, L.H., Zhang, Y., Zhang, Z., Xie, L.F., and Li, F.R. 2020. Comparison of Tree Biomass Modeling Approaches for Larch (*Larix olgensis* Henry) Trees in Northeast China [Article]. *Forests* **11**(2): 20. doi:10.3390/f11020202.
- Guerra-Hernández, J., Arellano-Pérez, S., González-Ferreiro, E., Pascual, A., Sandoval Altellarrea, V., Ruiz-González, A.D., and Álvarez-González, J.G. 2021. Developing a site index model for *P. Pinaster* stands in NW Spain by combining bi-temporal ALS data and environmental data. *For. Ecol. Manage.* **481**: 118690. doi:10.1016/j.foreco.2020.118690.
- Huang, S., Titus, S.J., and Wiens, D.P. 1992. Comparison of Nonlinear Height-diameter Functions for Major Alberta Tree Species. *Can. J. For. Res.* **22**: 1297-1304. doi:10.1139/x92-172.
- Koenker, R.W. and Bassett, G., Jr. 1978. Regression Quantiles. *Econometrica* **46**(1): 33-50. doi:10.4135/9781526421036936680.
- Lumbres, R.I.C., Cabral, D.E., Parao, M.R., Seo, Y.O., and Lee, Y.J. 2012. Evaluation of height-diameter models for three tropical plantation species in the Philippines. *Asia Life Sciences* **21**(2): 455-468. Available from <Go to ISI>://WOS:000306379900010 [accessed].

- Lundqvist, B. 1957. On the height growth in cultivated stands of pine and spruce in Northern Sweden. *Medd Fran Statens Skogforsk* **47**: 1-64.
- MacPhee, C., Kershaw, J.A., Weiskittel, A.R., Golding, J., and Lavigne, M.B. 2018. Comparison of approaches for estimating individual tree height-diameter relationships in the Acadian forest region. *Forestry* **91**(1): 132-146. doi:10.1093/forestry/cpx039.
- Mehtatalo, L. 2004. A longitudinal height-diameter model for Norway spruce in Finland. *Can. J. For. Res.* **34**(1): 131-140. doi:10.1139/x03-207.
- Mehtatalo, L. 2019. *lmfor: Functions for Forest Biometrics*. R package version 1.4. <https://CRAN.R-project.org/package=lmfor>.
- Mehtätalo, L. and Lappi, J. 2020. *Biometry for Forestry and Environmental Data with Examples in R*. CRC Press, Boca Raton, FL, US.
- Monserud, R.A. 1984. Height growth and site index curves for inland Douglas-fir based on stem analysis data and forest habitat type. *Forest science*. **30**(4): 943-965. doi:10.1093/forestscience/30.4.943.
- Özçelik, R., Cao, Q.V., Trincado, G., and 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.* **419-420**: 240-248. doi:10.1016/j.foreco.2018.03.051.
- Pinheiro, J. and Bates, D. 2000. *Mixed-effects models in S and S-Plus*. Statistics and Computing. Springer-Verlag, New York, USA. pp. 1-528.
- R Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Scolforo, H.F., McTague, J.P., Burkhart, H., Roise, J., Campoe, O., and Stape, J.L. 2019. Eucalyptus growth and yield system: Linking individual-tree and stand-level growth models in clonal Eucalypt plantations in Brazil. *For. Ecol. Manage.* **432**: 1-16.

- doi:10.1016/j.foreco.2018.08.045.
- Sharma, R.P., Vacek, Z., and Vacek, S. 2016. Nonlinear mixed effect height-diameter model for mixed species forests in the central part of the Czech Republic. *J. For. Sci.* **62**(10): 470-484. doi:10.17221/41/2016-JFS.
- Sharma, R.P., Vacek, Z., Vacek, S., and Kučera, M. 2018. Modelling individual tree height–diameter relationships for multi-layered and multi-species forests in central Europe. *Trees* **33**(1): 103-119. doi:10.1007/s00468-018-1762-4.
- State Forestry and Grassland Administration. 2019. The ninth forest resource survey report (2014–2018). China forestry press, Beijing, China. pp. 1-451.
- Temesgen, H., Monleon, V.J., and Hann, D.W. 2008. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests [Article]. *Can. J. For. Res.* **38**(3): 553-565. doi:10.1139/X07-104.
- Trim, K.R., Coble, D.W., Weng, Y., Stovall, J.P., and Hung, I.K. 2020. A New Site Index Model for Intensively Managed Loblolly Pine (*Pinus taeda*) Plantations in the West Gulf Coastal Plain. *For. Sci.* **66**(1): 2-13. doi:10.1093/forsci/fxz050.
- Xie, L., Widagdo, F.R.A., Dong, L., and Li, F. 2020. Modeling Height-Diameter Relationships for Mixed-Species Plantations of *Fraxinus mandshurica* Rupr. and *Larix olgensis* Henry in Northeastern China. *Forests* **11**(6). doi:10.3390/f11060610.
- Yang, S.-I. and Burkhart, H.E. 2020. Evaluation of total tree height subsampling strategies for estimating volume in loblolly pine plantations. *For. Ecol. Manage.* **461**. doi:10.1016/j.foreco.2020.117878.
- Yu, B. 2008. Artificial cultivation techniques of *Larix olgensis*. *Heilongjiang Sci. Tec. Info.* **35**: 193.
- Zang, H., Lei, X., and Zeng, W. 2016. Height-diameter equations for larch plantations in northern and northeastern China: a comparison of the mixed-effects, quantile

regression and generalized additive models. *Forestry* **89**(4): 434-445.

doi:10.1093/forestry/cpw022.

Zeide, B. 1993. Analysis of Growth Equations. *For. Sci.* **39**(3): 594-616.

doi:10.1093/forestscience/39.3.594.

Zhang, B., Sajjad, S., Chen, K., Zhou, L., Zhang, Y., Yong, K.K., and Sun, Y. 2020.

Predicting Tree Height-Diameter Relationship from Relative Competition Levels

Using Quantile Regression Models for Chinese Fir (*Cunninghamia lanceolata*) in

Fujian Province, China. *Forests* **11**(2). doi:10.3390/f11020183.

Zheng, J., Zang, H., Yin, S., Sun, N., Zhu, P., Han, Y., Kang, H., and Liu, C. 2018. Modeling

height-diameter relationship for artificial monoculture *Metasequoia glyptostroboides*

in sub-tropic coastal megacity Shanghai, China. *Urban For. Urban Green.* **34**: 226-

232. doi:10.1016/j.ufug.2018.06.006.

Table 1 Summary statistics of the sample plots, sample trees.

	Variables	N	Mean	Std.	Minimum	Maximum
Tree-level	<i>DBH</i> (cm)	59907	13.14	5.94	0.70	43.00
	<i>H</i> (m)		14.14	5.23	1.70	32.50
	<i>BAL</i> (m ² ·ha ⁻¹)		13.72	8.05	0.00	42.17
	<i>Rd</i>		0.96	0.26	0.10	3.24
Stand-level	<i>Age</i> (a)	524	26	11	8	64
	<i>Dq</i> (cm)		13.53	5.29	4.00	32.70
	<i>Ddom</i> (cm)		19.91	5.40	6.10	37.90
	<i>Hdom</i> (m)		16.67	4.87	4.60	30.70
	<i>SD</i> (trees·ha ⁻¹)		1905	954	160	4450
	<i>BA</i> (m ² ·ha ⁻¹)		21.92	7.56	1.59	42.25

Std.: Standard Deviation; *BAL*: basal area of trees with larger diameter than the subjected tree; *Rd*: the ratio of *DBH* to *Dq*; *Dq*: quadratic mean diameter; *Ddom*: average diameter of dominant trees; *Hdom*: average height of dominant trees; *SD*: stand density; *BA*: stand basal area.

Table 2 Estimations of parameters and standard error (in parentheses) for generalized height-diameter (GHD) model by nonlinear mixed effects model (NLME) and nonlinear quantile regression (NQR) at nine quantiles (*q*).

Method	β_1	β_2	β_3	β_4	β_5	δ	σ	D matrix	
GHD	2.1202	0.9357	0.0621	0.7742	-0.2570	0.7419	0.5345	—	
	(0.0760)	(0.0042)	(0.0013)	(0.0092)	(0.0061)				
NLME	-0.0783	1.0180	0.0589	0.7361	-0.3549	0.4333	0.3986	$\begin{pmatrix} 0.7049^2 & -0.0084 \\ -0.0084 & 0.0167^2 \end{pmatrix}$	
	(0.1221)	(0.0060)	(0.0017)	(0.0127)	(0.0071)				
NQR (q)									
$q=0.1$	2.6355	0.8581	0.0702	1.032	-0.234	—	—	—	
	(0.1872)	(0.0096)	(0.0022)	(0.0185)	(0.018)				
$q=0.2$	2.2024	0.8873	0.0706	0.9629	-0.2703	—	—	—	
	(0.1098)	(0.0056)	(0.0016)	(0.0146)	(0.0108)				
$q=0.3$	1.9488	0.9129	0.0677	0.8915	-0.2817	—	—	—	
	(0.0947)	(0.0052)	(0.0015)	(0.0128)	(0.0086)				
$q=0.4$	1.7930	0.9320	0.0652	0.8303	-0.2817	—	—	—	
	(0.0832)	(0.0048)	(0.0015)	(0.0130)	(0.0078)				
$q=0.5$	1.6475	0.9519	0.0620	0.7668	-0.2744	—	—	—	
	(0.0815)	(0.0046)	(0.0015)	(0.0121)	(0.0072)				
$q=0.6$	1.5525	0.9650	0.0609	0.7217	-0.2694	—	—	—	
	(0.0682)	(0.0040)	(0.0014)	(0.0102)	(0.0064)				
$q=0.7$	1.5575	0.9777	0.0595	0.6770	-0.2574	—	—	—	
	(0.0791)	(0.0044)	(0.0016)	(0.0112)	(0.0076)				
$q=0.8$	1.7309	0.9924	0.0565	0.6218	-0.2335	—	—	—	
	(0.0794)	(0.0059)	(0.0021)	(0.0139)	(0.0077)				
$q=0.9$	2.0074	1.0101	0.0520	0.5617	-0.2178	—	—	—	
	(0.1212)	(0.0068)	(0.0023)	(0.0144)	(0.0091)				

Table 3 Assessment of models' predictive performance under uncalibrated or calibrated, with the GHD, NLME, three quantile (NQR3), five quantile (NQR5), and nine quantile regression (NQR9) models based on leave-one-out cross-validation.

Method	Type	MAE (m)	MPE (%)	MAPE (%)	FI
GHD	No calibration	1.0851	0.0019	8.6548	0.9266
	Calibration	0.9662	0.0273	7.6083	0.9403
NLME	No calibration	1.1046	-1.5110	8.8290	0.9223
	Calibration	0.9670	1.0357	7.4845	0.9392
NQR9, NQR5 and NQR3	No calibration	1.0842	-0.2012	8.6474	0.9264
NQR9	Calibration	0.9551	-0.0587	7.5192	0.9414
NQR5	Calibration	0.9551	0.0052	7.5117	0.9415
NQR3	Calibration	0.9562	-0.0114	7.5222	0.9414

Table 4 Results of model fitting with different combinations of random parameters and likelihood ratio test. LL is the log likelihood value and LRT is the value of likelihood ratio test.

Location of random effect parameters	Number of parameters	AIC	LL	LRT	<i>p</i>
None	6	211449.0	-105718.5		
β_1	7	201351.1	-100668.5		
β_2	7	201430.7	-100708.3		
β_3	7	199825.1	-99905.5	11626.0	< 0.0001
β_5	7	201045.3	-100515.7		
β_1 and β_2	8	201353.1	-100668.5		
β_1 and β_3	8	199374.4	-99679.2	452.6	< 0.0001
β_1 and β_4	8	199553.9	-99769.0		
β_1 and β_5	8	201047.4	-100515.7		
β_2 and β_3	8	199399.7	-99691.9		
β_2 and β_4	8	199562.2	-99773.1		
β_2 and β_5	8	201047.4	-100515.7		
β_3 and β_5	8	201047.3	-100515.7		
β_1 , β_2 and β_3	9	199375.9	-99678.9	0.6	> 0.05
β_1 , β_2 and β_4	9	199559.5	-99770.8		
β_1 , β_2 and β_5	9	201049.4	-100515.7		
β_1 , β_3 and β_4	9	199389.8	-99685.9		
β_1 , β_3 and β_5	9	199376.4	-99679.2		
β_2 , β_3 and β_4	9	199422.1	-99702.0		
β_2 , β_3 and β_5	9	199401.7	-99691.9		
β_1 , β_2 , β_3 and β_4	10	199392.0	-99686.0		
β_1 , β_2 , β_3 and β_5	10	199378.4	-99679.2	-0.6	> 0.05

Table 5 Evaluation criteria of GHD, NLME, NQR3, NQR5 and NQR9 models by using the sampling design of selecting the medium-sized trees and different sizes. The numbers in bold denote the best method.

Item	Size	GHD	NLME	NQR3	NQR5	NQR9
MAE (m)	0	1.0851	1.1046	1.0842	1.0842	1.0842
	2	1.1490	1.0315	1.1437	1.1561	1.1313
	3	1.0924	1.0173	1.0876	1.0960	1.0742
	4	1.0639	1.0102	1.0584	1.0653	1.0495
	5	1.0452	1.0091	1.0371	1.0444	1.0341
	6	1.0297	1.0051	1.0223	1.0263	1.0214
	7	1.0201	1.0033	1.0140	1.0157	1.0110
	8	1.0164	1.0070	1.0104	1.0118	1.0070
	9	1.0158	1.0161	1.0089	1.0116	1.0064
	10	1.0161	1.0183	1.0095	1.0128	1.0057
MPE (%)	0	0.0019	-1.5110	-0.2012	-0.2012	-0.2012
	2	0.2361	-0.4614	0.0111	0.2469	0.1650
	3	0.0368	-0.2532	-0.1685	0.0391	-0.3850
	4	0.1671	0.0793	-0.0152	0.1824	-0.2141
	5	0.1230	0.2500	-0.0379	0.1389	-0.2101
	6	0.1206	0.3265	-0.0118	0.1391	-0.2603
	7	-0.0107	0.3054	-0.1021	0.0009	-0.3435
	8	-0.0311	0.4216	-0.1257	-0.0208	-0.3014
	9	-0.0758	0.5016	-0.2094	-0.0696	-0.3652
	10	-0.0939	0.5748	-0.2343	-0.0869	-0.3928
MAPE (%)	0	8.6548	8.8290	8.6474	8.6474	8.6474
	2	9.0180	8.1880	9.0631	9.1560	8.9365
	3	8.5751	8.0624	8.6133	8.6701	8.5230
	4	8.3516	7.9681	8.3715	8.4198	8.3378
	5	8.1776	7.9285	8.1691	8.2197	8.1654
	6	8.0588	7.8892	8.0427	8.0677	8.0684
	7	8.0060	7.8736	7.9950	8.0015	7.9957
	8	7.9751	7.8810	7.9650	7.9683	7.9598
	9	7.9712	7.9427	7.9571	7.9707	7.9525
	10	7.9800	7.9703	7.9679	7.9896	7.9582
FI	0	0.9266	0.9223	0.9264	0.9264	0.9264
	2	0.9168	0.9328	0.9179	0.9158	0.9194
	3	0.9241	0.9342	0.9248	0.9235	0.9263
	4	0.9278	0.9349	0.9287	0.9277	0.9297
	5	0.9302	0.9348	0.9313	0.9302	0.9314
	6	0.9324	0.9354	0.9332	0.9327	0.9331
	7	0.9336	0.9355	0.9343	0.9341	0.9344
	8	0.9340	0.9344	0.9346	0.9344	0.9349
	9	0.9339	0.9325	0.9347	0.9343	0.9347
	10	0.9337	0.9325	0.9345	0.9340	0.9348

Figure captions

Figure 1 Scatter plot of total tree height (H) against diameter at breast height (DBH) (A) and the diameter at the breast height distribution for Korean larch (B).

Figure 2 Plots of standardized residuals against predictions of H for unweighted/weighted generalized height-diameter (GHD) model and nonlinear mixed-effects (NLME) model.

Figure 3 Mean absolute error (MAE), mean error percent (MPE), mean absolute error percent (MAPE) and fit index (FI) of calibrated NLME, NQR3, NQR5 and NOR9 models under several sampling designs and pre-measured subsample size. Type I: random sampling; Type II: sample the thickest trees; Type III: sample the thinnest trees; Type IV: sample the medium-sized trees; Type V: 50% thickest trees and 50% thinnest trees when the sample size is an even number, while adding one more thinnest tree for an odd number; Type VI: similar to Type V, but adding one more thickest tree was preferred when the sample size is an odd number. The error bar represented the maximum and minimum of the indices in 500 times of random simulations.

Figure 4 MPE (A) and MAPE (B) of the tree-level fitting results for NLME, NQR3, NQR5 and NQR9 approaches across several DBH classes. The filled signs in each box represented the mean values.

Figure 5 MPE (A) and MAPE (B) of the plot-level fitting results for NLME, NQR3, NQR5 and NQR9 approaches. The filled signs in boxes represented the mean values.

Figure 6 Comparison of H predictions of the newly developed model in this study and the mixed-effects basic height-diameter model of Zang et al. (2016). All H predictions were calibrated by selecting ten medium-sized trees per sampled plot.

Figure 1 Scatter plot of total tree height (H) against diameter at breast height (DBH) (A) and the diameter at the breast height distribution for Korean larch (B).

Figure 2 Plots of standardized residuals against predictions of H for unweighted/weighted generalized height-diameter (GHD) model and nonlinear mixed-effects (NLME) model.

Figure 3 Mean absolute error (MAE), mean error percent (MPE), mean absolute error percent (MAPE) and fit index (FI) of calibrated NLME, NQR3, NQR5 and NOR9 models under several sampling designs and pre-measured subsample size. Type I: random sampling; Type II: sample the thickest trees; Type III: sample the thinnest trees; Type IV: sample the medium-sized trees; Type V: 50% thickest trees and 50% thinnest trees when the sample size is an even number, while adding one more thinnest tree for an odd number; Type VI: similar to Type V, but adding one more thickest tree was preferred when the sample size is an odd number. The error bar represented the maximum and minimum of the indices in 500 times of random simulations.

Figure 4 MPE (A) and MAPE (B) of the tree-level fitting results for NLME, NQR3, NQR5 and NQR9 approaches across several DBH classes. The filled signs in each box represented the mean values.

Figure 5 MPE (A) and MAPE (B) of the plot-level fitting results for NLME, NQR3, NQR5 and NQR9 approaches. The filled signs in boxes represented the mean values.

Figure 6 Comparison of H predictions of the newly developed model in this study and the mixed-effects basic height-diameter model of Zang et al. (2016). All H predictions were calibrated by selecting ten medium-sized trees per sampled plot.