AN APPROACH TO GROWTH AND YIELD MODELS FOR INDIVIDUAL CHINA-FIR (CUNNINGHAMILA LANCEOLATA) TREES IN SOUTHEAST CHINA


1School of Economics and Management, Ningxia University
No. 489 Helanshan West Road, Yinchuan 750021, PR China
(phone: +86-952-509-3001)

2College of Forestry, Beijing Forestry University
No. 35 Qinghua East Road, Haidian District, Beijing 100083, PR China

3Research Institute of Desertification Control, Ningxia Academy of Agriculture and Forestry Science, No. 590 Huanghe East Road, Yinchuan 750002, PR China

*Corresponding author
e-mail: sunyj@bjfu.edu.cn; phone: +86-10-6233-8197

(Received 12th Jul 2019; accepted 14th Nov 2019)

Abstract. As the most commonly grown afforestation species in southeast China, China-fir (Cunninghamia lanceolata) shows a huge ecological service function. Generalized individual-tree growth models were developed for C. lanceolata. Data was obtained from 61 plantation-grown China-fir trees in 17 single-species plots located in four sites by stem analysis. The best base models were chosen from five theoretical growth equations for modeling increases in diameter at breast high, tree, height and stem volume using ordinary nonlinear least squares regression; selection criteria were the smallest absolute mean residual, root mean square error and the largest adjusted coefficient of determination. To account for autocorrelation in the data with repeated measures, we developed a nested multi-level nonlinear mixed-effects (NLME) model, constructed on the selected base model; the NLME models incorporated random effect for tree, plot and site. The best random-effects combinations for the NLME models were identified by Akaike’s information criterion, Bayesian information criterion and -2 logarithm likelihood. Heteroscedasticity was reduced and autocorrelation was also addressed. For diameter and height growth, the NLME models including the exponential function and ARMA(1,1) performed best, and the NLME models including the power function and ARMA(1,1) performed best for stem growth. The NLME models were considered to be the best approach to analyze the variation of tree growth and yield.

Keywords: individual-tree model, autocorrelation, heteroscedasticity, nonlinear mixed-effects models, Cunninghamia lanceolata

Introduction

China-fir (Cunninghamia lanceolata (Lamb.) Hook.) is the most commonly grown afforestation species in southeast China. According to the National Continuous Forest Inventory, about 10.96 million ha and 726.09 million m³ of China-fir is distributed in more than ten provinces in China in 2015, making it the number one species in terms of harvest volume. As a result, a sustainable forest management plan for the species needs to be established.

The availability of growth information, such as growth models for individual tree characteristics including diameter at breast height (d), total tree height (h) and stem volume (v), enables a more detailed description than stand-level models (Burkhart, 2003; Vanclay, 2003). This important asset in sustainable forest management allows the selection of tree species for estimation of cutting cycles as well as for logging or
These basic growth models simplify the study of the variability of individual-tree increments and the estimation of yields directly; they also serve as the main tool used to estimate dynamic variability of biomass and carbon storage indirectly. In addition, they allow silvicultural treatments to be simulated and enable comparison of alternative thinning regimes (Mabvurira and Miina, 2002).

Individual-tree growth models are established based on tree growth increments as the dependent variable and the comparable factors, such as age, as the independent variables. They can simulate growth for individual trees (Sharma et al., 2016) and provide an invaluable tool for forest management planning at any level (Uzoh and Oliver, 2008).

The ordinary non-linear least squares (ONLS) regression, is the most commonly used statistical method in forest modeling (Grégoire et al., 1995). Individual-tree growth models have been fitted to growth increment data collected repeatedly over time on the same tree. The hierarchical nature of the data results in spatial and temporal correlation among observations made in the same sampling unit (i.e., site, plot and tree). However, the stochastic structure is often ignored and independence of observations is assumed (Gutzwiller and Riffell, 2007; Hao et al., 2015). Furthermore, when data are autocorrelated they cannot be considered to be independent samples of the basic tree population (Fox et al., 2001). Therefore, a violation of the ONLS regression assumption of independent residuals exists and results in biased estimates of the standard error of parameter estimation (Dumont et al., 2018).

Nonlinear mixed-effects (NLME) models incorporate both fixed and random effects. Fixed effects are parameters associated with an entire population or with certain repeatable levels of experimental factors, while random effects are associated with individual experimental units drawn at random from a population (Zhu et al., 2019). Random effects account for spatial and temporal correlation by defining the covariance structure of the model’s random components and by using this structure during parameter estimation. Therefore, the evolution of NLME models provides an efficient statistical method for explicitly modeling hierarchical stochastic structure (Sharma et al., 2016; Abiyu et al., 2018; Huff et al., 2018). Because of their flexibility, NLME models have become increasingly used to forest growth and yield modeling (Fu et al., 2013; Matos et al., 2013; Xu et al., 2014; Huff et al., 2018).

The main objective of this study was to develop individual-tree growth models for *C. lanceolata* in Fujian Province, southeast China. The data were derived from 61 trees in 17 sample plots located in four sites. A multi-level nonlinear mixed modeling approach that included both fixed and random components was applied to the hierarchical structure of the data. This diminished the level of variance among the sampling units, which were included as covariates. In developing the growth models, we considered the use of nested three-level models. The first, second, and third levels are the site, the plots nested within the site, and the tree, nested within the plot within site, respectively. Our preliminary analysis showed that the NLME models with random effects effectively removed the heteroscedasticity and autocorrelations in repeated-measure data, and therefore could be important tools for sustainable management of the China-fir species within the study area. The predictive ability of the model developed here and the applicability of the NLME model were demonstrated using separate validation data.
Materials and methods

Data

Pure China-fir stands located in Jiangle County (26°26′–27°04′N, 117°05′–117°40′E), Fujian Province, in southeast China with a red soil type were used in this study. Average mean precipitation, frost-free length, and temperature were about 1699 mm, 287 days, and 18.7°C, respectively.

The sampling period took place over 2016-2017 during the summer. The data were obtained from 61 trees in 17 single-species plots of plantation-grown China-fir that located in four sites, identified as the Qiantan, Shuinan, Yuhua and Yuandang areas, which were numbered I, II, III and IV, respectively (Fig. 1).

Figure 1. Seventeen sample plots located in four regions in Fujian Province, southeast China

Sample plots were square and varied in size from 400 to 600 m². The increments of \( d \), \( h \) and \( v \) per year were measured or calculated, using the stem analysis approach. The data were randomly divided into two groups; 75% of the points were used for model fitting, and 25% were used for model validation. The fitting data and the validation data included 23 trees from eight plots, and eight trees from four plots, respectively. To achieve more confidence in the results, the fitting and validation data were chosen three times and the model estimation and prediction procedures were also iterated. However, the procedures and results were listed only once in this paper, including summary statistics for both fitting and validation data (Table 1). The increments data of \( d \), \( h \) and \( v \) are graphically depicted in Fig. 2.

Methods

Individual-tree growth equations

Five theoretical nonlinear growth equations (NLs), the Richards, Weibull, Korf, Logistic and Schumacher equations, were selected as candidates for modeling individual-tree growth. These equations are widely used for the simulation of individual-tree growth, particularly the Richards and Korf equations (Zhang et al., 2011; Xu et al., 2014). Table 2 shows the mathematical expressions of the equations.

The five equations mentioned above are all S-shaped growth equations with inflection points and asymptotes. One characteristic of the Richards, Weibull and Korf equations is that the coordinates of inflection points are variable multiples of asymptotic values; in contrast, the equivalent of the Logistic and Schumacher equations are fixed.
multiples. The five equations were initially fit by ONLS regression without considering random parameters. Different initial values for the parameters were tried to ensure that a global minimum was achieved. The best performing function was selected as the base model by applying three statistical criteria; absolute mean residual (AMR), root mean square error (RMSE), and adjusted coefficient of determination (adj-$R^2$) (Zhang et al., 2011). The function with the smallest AMR and RMSE and the largest adj-$R^2$ provides the best fit. Similarly, the adjusted coefficient of determination is used as an unbiased estimator in both multiple regression and canonical redundancy analysis. The formulas of the fit statistics are:

$$AMR = \frac{\sum_{i=1}^{M} \sum_{j=1}^{M_i} \sum_{k=1}^{M_{ij}} |y_{ijk} - \hat{y}_{ijk}|}{n_{ijk}} \quad (Eq.1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{M} \sum_{j=1}^{M_i} \sum_{k=1}^{M_{ij}} (y_{ijk} - \hat{y}_{ijk})^2}{n_{ijk} - r}} \quad (Eq.2)$$

$$adj-R^2 = 1 - \left[ \frac{\sum_{i=1}^{M} \sum_{j=1}^{M_i} \sum_{k=1}^{M_{ij}} n_{ijk} (y_{ijk} - \hat{y}_{ijk})^2}{\sum_{i=1}^{M} \sum_{j=1}^{M_i} \sum_{k=1}^{M_{ij}} n_{ijk} (y_{ijk} - \bar{y})^2} \right] \quad (Eq.3)$$

where $M$ is the number of sites, $M_i$ is the number of plots within the $i$th site, $M_{ij}$ is the number of trees within the $j$th plot of the $i$th site, $n_{ijk}$ is the number of observations (increments), $r$ is the number of parameters in the model, $y_{ijk}$ is the value of increments of $k$th tree taken from the $j$th plot in the $i$th site, $\hat{y}_{ijk}$ is the increment prediction, and $\bar{y}$ is the average of observations. The accuracy of the models was tested against the modeling data and against independent data from the same site and plot (Adame et al., 2008).

**Table 1. Summary statistics for increment data for Cunninghamia lanceolata in Fujian Province, in southeast China**

<table>
<thead>
<tr>
<th>Site</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plot</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Tree</td>
<td>10</td>
<td>14</td>
<td>21</td>
<td>16</td>
</tr>
<tr>
<td>$d$ (cm)</td>
<td>13.4–27.7 (18.9)</td>
<td>15.8–22.4 (19.5)</td>
<td>7.3–26.6 (15.0)</td>
<td>5.1–10.1 (7.3)</td>
</tr>
<tr>
<td>SD</td>
<td>4.4</td>
<td>2.7</td>
<td>6.6</td>
<td>1.8</td>
</tr>
<tr>
<td>$h$ (m)</td>
<td>11.5–22.3 (18.0)</td>
<td>15.0–25.0 (19.8)</td>
<td>7.0–24.0 (14.8)</td>
<td>4.1–7.6 (5.9)</td>
</tr>
<tr>
<td>SD</td>
<td>3.9</td>
<td>3.6</td>
<td>5.4</td>
<td>1.4</td>
</tr>
<tr>
<td>Age (years)</td>
<td>25–38 (30)</td>
<td>23–32 (28)</td>
<td>15–35 (25)</td>
<td>6–8 (7)</td>
</tr>
<tr>
<td>SD</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Elevation (m)</td>
<td>192–242 (226)</td>
<td>227–332 (293)</td>
<td>219–286 (249)</td>
<td>210–276 (243)</td>
</tr>
<tr>
<td>SD</td>
<td>21</td>
<td>51</td>
<td>24</td>
<td>36</td>
</tr>
<tr>
<td>Slope (°)</td>
<td>15–45 (30)</td>
<td>30–41 (37)</td>
<td>22–44 (29)</td>
<td>34–37 (36)</td>
</tr>
<tr>
<td>SD</td>
<td>13</td>
<td>5</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

Note: $d$, diameter at breast height; $h$, total tree height; SD, standard deviation.
Table 2. Mathematical expressions for the five theoretical growth equations

<table>
<thead>
<tr>
<th>Equation</th>
<th>Expression</th>
<th>Inflection point</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richards</td>
<td>$y = a(1 - \exp(-bx))^c$</td>
<td>$1/(b \ln c)$, $a(1 - 1/c)^c$</td>
<td>$a, b &gt; 0$</td>
</tr>
<tr>
<td>Weibull</td>
<td>$y = a(1 - \exp(-bx^c))$</td>
<td>$(c-1)/bc$, $a(1 - \exp(1-c)/c)$</td>
<td>$a, b, c &gt; 0$</td>
</tr>
<tr>
<td>Korf</td>
<td>$y = a \exp(-b/x^c)$</td>
<td>$(c+1)/bc$, $a\exp((c-1)/c)$</td>
<td>$a, b, c &gt; 0$</td>
</tr>
<tr>
<td>Logistic</td>
<td>$y = a/(1 + \exp(b-cx))$</td>
<td>$b/c$, $a/2$</td>
<td>$a, c &gt; 0$</td>
</tr>
<tr>
<td>Schumacher</td>
<td>$y = a \exp(-b/x)$</td>
<td>$b/2$, $ae^{-2}$</td>
<td>$a, b &gt; 0$</td>
</tr>
</tbody>
</table>

Source of equations: Duan et al., 2003; Zhang et al., 2011

Multi-level NLME models

Available data were based on the measured annual increment of the sample plots located at different sites. The nested sampling structure created a high degree of correlation among observations taken from the same tree, plot and site. To alleviate this, a mixed-effects model approach has been proposed by other authors (Laird and Ware, 1982; Palmer et al., 1991; Grégoire et al., 1995). A general expression for a NLME model can be defined as (Lindstrom and Bates, 1990; Vonesh and Chinchilli, 1997):

$$y_{ijkl} = f(\theta_{ijkl}, x_{ijkl}) + \varepsilon_{ijkl} \quad i = 1,...,M, \quad j = 1,...,M_i, \quad k = 1,...,M_0, \quad l = 1,...,n_{ijkl} \quad \text{(Eq.4a)}$$

where $x_{ijkl}$ is the age, $\theta_{ijkl}$ is a parameter vector, $r \times 1$, $f$ is a nonlinear function of the predictor variables and the parameter vector, and $\varepsilon_{ijkl}$ is the within-group error including the within-group variance and correlation (Davidian and Giltinan, 1995) which is assumed to be normally distributed with zero expectation and a positive-definite variance-covariance structure $R_{ijkl}$ (Meng and Huang, 2009):

$$\varepsilon_{ijkl} \sim N(0, R_{ijkl}) \quad \text{(Eq.4b)}$$
Moreover, $\phi_{ijkl}$ can be expressed as:

$$\phi_{ijkl} = A_{ijkl} + B_{ijkl}^\mu + B_{ijkl}^\beta + B_{ijkl}^\gamma$$

(Eq.4c)

$$\mu_i \sim N(0, \psi_i), \mu_{ij} \sim N(0, \psi_{ij}), \mu_{ijk} \sim N(0, \psi_{ijk})$$

(Eq.4d)

where $\lambda$ is the $p \times 1$ vector of fixed population parameters (where $p$ is the number of fixed parameters in the model). $\mu_i, \mu_{ij}$ and $\mu_{ijk}$ are the $q_1 \times 1, q_2 \times 1$ and $q_3 \times 1$ vectors of random effects associated with the first, second and the third levels, respectively (where $q_1, q_2, q_3$ are the numbers of random parameters of different levels in the model). $\mu_i, \mu_{ij}$ and $\mu_{ijk}$ are assumed to be normal (or Gaussian) with a mean of 0 and have the variance-covariance matrices $\psi_i, \psi_{ij}$ and $\psi_{ijk}$, which are the $q_1 \times q_1, q_2 \times q_2$ and $q_3 \times q_3$ variance-covariance matrices of random effects in three levels, respectively. $A_{ijkl}, B_{ijkl}^\mu, B_{ijkl}^\beta$ and $B_{ijkl}^\gamma$ are design matrices of size $r \times p$, $r \times q_1$, $r \times q_2$ and $r \times q_3$ for the fixed and random effects of different levels specific to each sampling unit, respectively.

Determining the variance-covariance structure

The variance-covariance matrices $\psi_i, \psi_{ij}$ and $\psi_{ijk}$ are positive-definite and symmetric, which is to say that all their eigenvalues must be strictly positive (Pinheiro and Bates, 2000). A hypothetical $2 \times 2$ variance-covariance matrix is shown as follows (Calama and Montero, 2004; Fu et al., 2013):

$$\begin{bmatrix}
\sigma_u^2 & \sigma_{uw} \\
\sigma_{uw} & \sigma_w^2
\end{bmatrix}$$

(Eq.5)

where $\sigma_u^2$ and $\sigma_w^2$ are the variances for the random effects $u$ and $w$, respectively, and $\sigma_{uw} = \sigma_{wu}$ is the covariance among random effects $u$ and $w$.

Determining the structure of $R$

The matrix $R_{ijkl}$ can be proposed to include both correlation effects and weighting factors to account for within-tree heteroscedasticity and autocorrelation (Davidian and Giltinan, 1995; Calama and Montero, 2004; Meng and Huang, 2009). In this study, a general expression for the matrix is given by (Calama and Montero, 2004; Crecente-Campo et al., 2010):

$$R_{ijkl} = \sigma^2 G_{ijkl}^{0.5} I_{ijkl} G_{ijkl}^{0.5}$$

(Eq.6)

where $G_{ijkl}$ is an $n_{ijkl} \times n_{ijkl}$ diagonal matrix within-tree heteroscedasticity variances and $I_{ijkl}$ is an $n_{ijkl} \times n_{ijkl}$ matrix showing the within-tree autocorrelation structure of errors, and $\sigma^2$ is a scaling factor for the error dispersion (Calama and Montero, 2004; Crecente-Campo et al., 2010).
In individual-tree growth models, the variance is often found to be dependent of the means, and the variance will generally increase with increasing mean tree diameter. To remove this effect, we modeled the variance as an exponential function or power function (Pinheiro and Bates, 2000). Autocorrelation structures were used to address the within-tree correlations observed in the data. A method was selected from among three commonly used approaches: first-order autoregressive structure \([AR(1)]\), a combination of first-order autoregressive and moving average structures \([ARMA(1,1)]\), and the compound symmetry structure \([CS]\) (Pinheiro and Bates, 2000).

\[
V\left(\varepsilon_{ijkl}\right) = \sigma^2 \exp\left(2\delta x_{ijkl}\right) \quad \text{(Eq.7)}
\]

\[
V\left(\varepsilon_{ijkl}\right) = \sigma^2 x_{ijkl}^{2\delta} \quad \text{(Eq.8)}
\]

\[
AR(1) = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix} \quad \text{(Eq.9)}
\]

\[
ARMA(1,1) = \sigma^2 \begin{bmatrix} 1 & \gamma & \gamma \rho \\ \gamma & 1 & \gamma \\ \gamma \rho & \gamma & 1 \end{bmatrix} \quad \text{(Eq.10)}
\]

\[
CS = \begin{bmatrix} \sigma^2 + \sigma_i & \sigma_i & \sigma_i \\ \sigma_i & \sigma^2 + \sigma_i & \sigma_i \\ \sigma_i & \sigma_i & \sigma^2 + \sigma_i \end{bmatrix} \quad \text{(Eq.11)}
\]

where \(\delta\) is an estimated parameter, \(\rho\) is the autoregressive parameter, \(\gamma\) is a moving average component, \(\sigma^2\) is the residual variance, and \(\sigma_i\) is the residual covariance (Leak, 1996).

**Mixed parameter evaluation**

A crucial issue in fitting mixed-effects models is deciding which parameters should be considered random effects and which can be treated as fixed effects. A common approach is to start with random effects for all parameters and then to examine the fitted object to decide which, if any, of the random effects can be eliminated from the model (Pinheiro and Bates, 2000). Therefore, different combinations of model parameters were tested to ascertain their contribution to predictions of individual-tree growth (Adame et al., 2008).

**Parameter estimation**

The parameters in the equation were estimated by maximum likelihood (ML) using the Lindstrom and Bates (LB) algorithm implemented in the R nlme function (Lindstrom and Bates, 1990; Pinheiro and Bates, 2000). Several articles provide details related to the LB algorithm and nlme function (Lindstrom and Bates, 1990; Pinheiro and Bates, 2000).
Predicting the random effects parameters is more problematic during model application and prediction than during the fitting process. In this case, they were estimated by the EBLUPs (Vonesh and Chinchilli, 1997), using the increment data.

\[
\hat{b}_{ijk} \approx D Z_{ijk}^{T} \left( R_{ijkl} + Z_{ijk} D Z_{ijk}^{T} \right)^{-1} \hat{e}_{ijk}
\]

(Eq.12)

where \( D \) is the estimated variance-covariance matrix for the random effects \( \hat{b}_{ijk} \), \( R_{ijkl} \) is the estimated variance-covariance matrix for the error term, \( Z_{ijk} \) is the estimated partial derivatives matrix, which contains the values of partial derivatives with respect to random effects parameters for the new observation.

**Model evaluation**

To compare the candidate models and to determine the fixed-versus random-effects parameters for each candidate model, Akaike’s information criterion, Bayesian information criterion and -2 logarithm likelihood (Zhao et al., 2012) were used. AIC is the log-likelihood penalized by the number of parameters. BIC is similar but penalizes more for models with more parameters. Based on the AIC, BIC and -2 LL statistics, one or more candidate models were selected for further evaluation (Yang et al., 2009; Hao et al., 2015). The appropriate variance function and autoregressive structure for the NLME models were determined by the likelihood ratio test (LRT) (Pinheiro and Bates, 2000; Fang and Bailey, 2001). The predictions were calculated by Equation 4 based on the fixed- and random-effects parameters. All models presented in this paper were calibrated using the nlme function in the R statistical environment (Ihaka and Gentleman, 2004; Pinheiro et al., 2005).

**Results**

**Function selection**

The R nls function was used to evaluate the parameter estimates and model fit statistics of the five equations (Table 2). For \( d \) and \( h \) growth models, the Korf equation has slightly better predictive ability than the others do. For \( v \) growth models, the Richard’s equation performs better than the others perform. Therefore, the Korf equation was selected as the basic nonlinear model for estimating \( d \) and \( h \) growth models and the Richards equation was selected as the \( v \) growth model. The final base models are given by:

\[
d_{i} = a \exp\left(-b/h_{i}^{c}\right) + \varepsilon_{i}
\]

(Eq.13)

\[
h_{i} = a \exp\left(-b/h_{i}^{c}\right) + \varepsilon_{i}
\]

(Eq.14)

\[
v_{i} = a\left(1-\exp\left(-b h_{i}^{c}\right)\right)^{c} + \varepsilon_{i}
\]

(Eq.15)

where \( a, b \) and \( c \) are fixed effects parameters, and \( \varepsilon_{i} \) is the residual noise term.
Construction of NLME models

AIC, BIC and −2 LL fit statistics were compared among different combinations of random effects parameters (Table 3). The models represented by Eqs. 16–17 based on the Korf equation and incorporating the nested effects of site as well as plot and tree on \( a \) and \( b \), yielded the smallest AIC, BIC and −2 LL. Meanwhile, the model represented by Eq. 18 based on the Richards equation with the nested effects of site, plot and tree on \( a \) and \( b \) performed better.

\[
d_{ij} = \left( \alpha + u_i + w_i + u_{ijk} \right) \exp\left( -\left( \beta + w_i + w_j + w_{ijk} \right) x_{ij} \right) + \epsilon_{ij}
\]

(Eq.16)

\[
h_{ij} = \left( \alpha + u_i + w_i + u_{ijk} \right) \exp\left( -\left( \beta + w_i + w_j + w_{ijk} \right) x_{ij} \right) + \epsilon_{ij}
\]

(Eq.17)

\[
v_{ij} = \left( \alpha + u_i + w_i + w_{ijk} \right) \left( 1 - \exp\left( -\beta x_{ij} \right) \right)^{\gamma + u_i + w_j + w_{ijk}} + \epsilon_{ij}
\]

(Eq.18)

where \( \alpha \), \( \beta \) and \( \gamma \) are fixed-effects parameters, \( u_i \) and \( w_i \) are random-effects parameters generated by site, \( u_{ij} \) and \( w_{ij} \) are random-effects parameters generated by plot, and \( u_{ijk} \) and \( w_{ijk} \) are random-effects parameters generated by interaction by tree.

Table 3. Evaluation indices of each NLME model

<table>
<thead>
<tr>
<th>Mixed parameters</th>
<th>AIC</th>
<th>BIC</th>
<th>-2 LL</th>
<th>AIC</th>
<th>BIC</th>
<th>-2 LL</th>
<th>AIC</th>
<th>BIC</th>
<th>-2 LL</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( a )</td>
<td>1370.53</td>
<td>1400.00</td>
<td>1356.53</td>
<td>1648.74</td>
<td>1678.78</td>
<td>1634.74</td>
<td>-3364.98</td>
<td>-3334.92</td>
<td>-3378.98</td>
</tr>
<tr>
<td>( b )</td>
<td>1461.37</td>
<td>1490.85</td>
<td>1447.37</td>
<td>1661.18</td>
<td>1691.22</td>
<td>1647.18</td>
<td>-3244.45</td>
<td>-3254.40</td>
<td>-3294.45</td>
</tr>
<tr>
<td>( c )</td>
<td>1336.96</td>
<td>1366.43</td>
<td>1322.96</td>
<td>1619.78</td>
<td>1649.82</td>
<td>1605.78</td>
<td>not converge</td>
<td>not converge</td>
<td>not converge</td>
</tr>
<tr>
<td>( a, b )</td>
<td>768.90</td>
<td>823.64</td>
<td>742.90</td>
<td>1307.83</td>
<td>1363.62</td>
<td>1281.83</td>
<td>-3951.46</td>
<td>-3895.64</td>
<td>-3977.46</td>
</tr>
<tr>
<td>( a, c )</td>
<td>not converge</td>
<td>not converge</td>
<td>not converge</td>
<td>-3956.04</td>
<td>-3900.23</td>
<td>-3982.04</td>
<td>not converge</td>
<td>not converge</td>
<td>not converge</td>
</tr>
<tr>
<td>( b, c )</td>
<td>864.86</td>
<td>919.60</td>
<td>838.86</td>
<td>1330.43</td>
<td>1386.22</td>
<td>1304.43</td>
<td>-3743.03</td>
<td>-3687.22</td>
<td>-3769.03</td>
</tr>
<tr>
<td>( a, b, c )</td>
<td>not converge</td>
<td>not converge</td>
<td>not converge</td>
<td>not converge</td>
<td>not converge</td>
<td>not converge</td>
<td>not converge</td>
<td>not converge</td>
<td>not converge</td>
</tr>
</tbody>
</table>

Note: \( a \), \( b \), and \( c \) are fixed parameters; \( d \), \( h \), and \( v \) are growth models; AIC, Akaike’s information criterion; BIC, Bayesian information criterion; -2 LL, -2 logarithm likelihood

NLME models with heteroscedasticity and autocorrelation

The power function or the exponential function were used as the variance functions and the AR(1), ARMA(1,1) or CS were used as the autocorrelation structures to fit growth models incorporating different random effects (Table 4). The selected models had the smallest AIC, BIC and -2 LL. Thus, the final models of the three nested effects of site, plot and tree for \( d \), \( h \), \( v \) growth are, respectively:

\[
\text{Eq.16+Eq.7+Eq.10}
\]

(Eq.19)

\[
\text{Eq.17+Eq.7+Eq.10}
\]

(Eq.20)

\[
\text{Eq.18+Eq.8+Eq.10}
\]

(Eq.21)
**Parameter estimates**

We can conclude the general growth models for China-fir individual-tree in southeast China become:

\[
d_{ijkl} = (21.1384 + u_i + u_j + u_{ijk})\exp\left(-\frac{(12.4464 + w_i + w_j + w_{ijk})}{x_{ijkl}^{1.0904}}\right) + e_{ijkl} \quad \text{(Eq.22)}
\]

\[
h_{ijkl} = (32.7823 + u_i + u_j + u_{ijk})\exp\left(-\frac{(7.5768 + w_i + w_j + w_{ijk})}{x_{ijkl}^{0.7240}}\right) + e_{ijkl} \quad \text{(Eq.23)}
\]

\[
v_{ijkl} = (0.2527 + u_i + u_j + u_{ijk})(1 - \exp(-0.0982x_{ijkl}))^{(5.3620 + u_i + u_j + u_{ijk})} + e_{ijkl} \quad \text{(Eq.24)}
\]

where Table 5 lists the random effect matrix, residual error, heteroscedasticity and autocorrelation of \(d\), \(h\) and \(v\) growth NLME models.

**Table 4. Performance criteria for the best NLME models**

<table>
<thead>
<tr>
<th>Factors</th>
<th>Variance function</th>
<th>Correlation structure</th>
<th>AIC</th>
<th>BIC</th>
<th>-2 LL</th>
<th>LRT</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(d)</td>
<td>None</td>
<td>None</td>
<td>768.90</td>
<td>823.64</td>
<td>742.90</td>
<td>2.1744</td>
<td>0.1403</td>
</tr>
<tr>
<td></td>
<td>Exponent</td>
<td>None</td>
<td>768.73</td>
<td>827.67</td>
<td>740.73</td>
<td>760.1422</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>ARMA(1, 1)</td>
<td>10.58</td>
<td>73.74</td>
<td>-19.42</td>
<td>604.9958</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Exponent</td>
<td>ARMA(1, 1)</td>
<td>-183.20</td>
<td>-115.83</td>
<td>-215.20</td>
<td>15.2631</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>None</td>
<td>1307.83</td>
<td>1363.62</td>
<td>1281.83</td>
<td>8.0797</td>
<td>0.0045</td>
</tr>
<tr>
<td></td>
<td>Exponent</td>
<td>None</td>
<td>1301.75</td>
<td>1361.83</td>
<td>1273.75</td>
<td>604.9958</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>ARMA(1, 1)</td>
<td>703.78</td>
<td>768.15</td>
<td>673.78</td>
<td>15.2631</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>Exponent</td>
<td>ARMA(1, 1)</td>
<td>699.32</td>
<td>767.99</td>
<td>667.32</td>
<td>15.2631</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>None</td>
<td>-3956.04</td>
<td>-3900.23</td>
<td>-3982.04</td>
<td>-3982.04</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Power</td>
<td>None</td>
<td>-4811.76</td>
<td>-4751.66</td>
<td>-4839.76</td>
<td>857.7244</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>ARMA(1, 1)</td>
<td>-4644.57</td>
<td>-4580.17</td>
<td>-4674.57</td>
<td>165.1945</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Power</td>
<td>ARMA(1, 1)</td>
<td>-5694.74</td>
<td>-5626.05</td>
<td>-5726.74</td>
<td>1052.2020</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Note: \(d\), \(h\) and \(v\) are growth models; AIC, Akaike’s information criterion; BIC, Bayesian information criterion; -2 LL, -2 logarithm likelihood; LRT, likelihood ratio test; ARMA, Auto-Regressive and Moving Average

**Model prediction**

The predictive ability of Eqs. 13–15 was evaluated using prediction procedures and Eqs. 1–3 on both fitting and validation data. The performance of the NLME models was evaluated using cross-validation procedures for both fitting and validation data; the random effects were predicted with EBLUPs (Eq. 12), using the measurement data.

Table 6 lists the three fit statistics for Eqs. 13–15 and 22–24. Eqs. 22–24 were the best predictors, with increases in adj-R² and decreases in RMSE for both fitting and validation data. In Fig. 3, the residuals of Eqs. 13–15 and 22–24 are plotted against the fitted values; the fitted values are plotted against the observed values (Fig. 4). Based on the above analysis, we can conclude that the NLME models, such as Eqs. 22–24, incorporating nested multi-level effects of site, plot and tree, have sufficiently high predictive power. Therefore, the NLME models incorporating the random effects of site, plot and tree were the best model for predicting \(d\), \(h\) and \(v\) growth of individual China-fir trees in the single-species plantations of the study area.
Table 5. The factors of d, h and v growth NLME models

<table>
<thead>
<tr>
<th>Factors</th>
<th>d</th>
<th>h</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random effect matrix</td>
<td>$\mu_i = u_{w, j} \cdot N(0, \psi = 2.82E-08, -0.9936)$</td>
<td>$\mu_i = u_{w, j} \cdot N(0, \psi = 6.95E-04, -0.0008)$</td>
<td>$\mu_i = u_{w, j} \cdot N(0, \psi = 3.71E-05, 0.0009)$</td>
</tr>
<tr>
<td>Residual error</td>
<td>$\epsilon_{ijkl} \cdot N(0, R_{ijkl} = 1.4817G_{ikl}^3G_{jkl}^3)$</td>
<td>$\epsilon_{ijkl} \cdot N(0, R_{ijkl} = 1.1042G_{ikl}^3I_{ijkl}G_{jkl}^3)$</td>
<td>$\epsilon_{ijkl} \cdot N(0, R_{ijkl} = 0.0302G_{ikl}^3I_{ijkl}G_{jkl}^3)$</td>
</tr>
<tr>
<td>Heteroscedasticity</td>
<td>$V(e_{ijkl}) = 1.4817 \exp(1.072x_{ijkl})$</td>
<td>$V(e_{ijkl}) = 1.1042 \exp(-0.0512x_{ijkl})$</td>
<td>$V(e_{ijkl}) = 0.0302x_{ijkl}^{0.076}$</td>
</tr>
<tr>
<td>Correlation structure</td>
<td>ARMA(1, 1) = 1.4817 $\times$ [1 0.9157 0.5409]</td>
<td>ARMA(1, 1) = 1.1042 $\times$ [1 1.0406 -0.1789]</td>
<td>ARMA(1, 1) = 0.0302 $\times$ [1 -0.140 -1.387]</td>
</tr>
</tbody>
</table>

Note: d, h, and v are growth models

Table 6. Performance criteria for each model

<table>
<thead>
<tr>
<th>Factor</th>
<th>Model</th>
<th>Fitting data</th>
<th>Validation data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AMR</td>
<td>RMSE</td>
</tr>
<tr>
<td>d</td>
<td>Eq.13</td>
<td>2.3513</td>
<td>2.9111</td>
</tr>
<tr>
<td></td>
<td>Eq.22</td>
<td>0.3782</td>
<td>0.5336</td>
</tr>
<tr>
<td>h</td>
<td>Eq.14</td>
<td>1.7484</td>
<td>2.2899</td>
</tr>
<tr>
<td></td>
<td>Eq.23</td>
<td>0.5436</td>
<td>0.7402</td>
</tr>
<tr>
<td>v</td>
<td>Eq.15</td>
<td>0.0308</td>
<td>0.0482</td>
</tr>
<tr>
<td></td>
<td>Eq.24</td>
<td>0.0039</td>
<td>0.0068</td>
</tr>
</tbody>
</table>

Note: d, h, and v are growth models
Discussion

Of the five theoretical growth equations tested, the Korf equation provided the best fit for \( d \) and \( h \) growth data of China-fir when evaluated based on AMR, RMSE and adj-R\(^2\), while the Richards equation showed the best performance for \( v \) growth. The results were consistent with the ones of the other two groups of data (Table 7). In general, the Korf and Richards equations have been widely used for forest growth and yield simulation models (Birch, 1999; Li et al., 2000; Malik et al., 2005; Gea-Izquierdo et al., 2008; Rugmini and Jayaraman, 2009; Bontemps and Duplat, 2012; Xu et al., 2014). ONLS regression is commonly used to build forest growth and yield models, but its usefulness is limited because tree data typically violate the assumption of independence and identically distributed errors (West et al., 1984; Grégoire et al., 1995). NLME models are useful tools for analyzing repeated measures data and spatially correlated data (Pinheiro and Bates, 2000; Fang and Bailey, 2001). A model can be constructed with a unique variance-covariance structure that eliminates the influence of random effects (site, plot and tree effects in this study) (Zhu et al., 2019). The two primary challenges in fitting NLME models are determining the mixed parameter and calculating the random effects (Uzoh and Oliver, 2008; Yang et al., 2009; Dumont et al., 2018; Huff et al., 2018).
The Richards and Korf equations have been widely used as the base NLME model for modeling forest growth and yield. For example, Fang and Bailey (2001) successfully used a modified Richards equation with an NLME model to build a slash pine (Pinus elliottii Engelm.) dominant height growth model. In their study, the Korf equation was chosen as the base model of the NLME model to fit loblolly pine (Pinus taeda Linn) diameter relationships (Cheng and Gordon, 2012); in that case, the nested three-level individual-tree NLME models (Eqs. 22–24) provided the best fit. Parameters $a$, $b$, $c$ are the asymptotic values, the values associated with the growth rate of the tree and the values associated with the curve shape of both the Richards and Korf equations, respectively (Zhang and Sun, 2010). Therefore, the random effects (site, plot and tree) mainly influence the growth rate and the maximum value of the tree for $d$ and $h$ growth, and the maximum value and curve shape for $v$ growth.

### Table 7. Performance criteria of each model for the other two groups of data

<table>
<thead>
<tr>
<th>Factor</th>
<th>Model</th>
<th>Fitting data</th>
<th>Validation data</th>
<th>Fitting data</th>
<th>Validation data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AMR</td>
<td>RMSE</td>
<td>adj-$R^2$</td>
<td>AMR</td>
</tr>
<tr>
<td>$d$</td>
<td>NL</td>
<td>2.4651</td>
<td>3.1384</td>
<td>0.7864</td>
<td>2.6182</td>
</tr>
<tr>
<td></td>
<td>NLME</td>
<td>0.4103</td>
<td>0.5602</td>
<td>0.9932</td>
<td>0.5048</td>
</tr>
<tr>
<td>$h$</td>
<td>NL</td>
<td>1.8253</td>
<td>2.6288</td>
<td>0.8634</td>
<td>1.9182</td>
</tr>
<tr>
<td></td>
<td>NLME</td>
<td>0.5814</td>
<td>0.8215</td>
<td>0.9865</td>
<td>0.7082</td>
</tr>
<tr>
<td>$v$</td>
<td>NL</td>
<td>0.0508</td>
<td>0.0573</td>
<td>0.7739</td>
<td>0.0604</td>
</tr>
<tr>
<td></td>
<td>NLME</td>
<td>0.0055</td>
<td>0.0178</td>
<td>0.9782</td>
<td>0.0095</td>
</tr>
</tbody>
</table>

Note: $d$, $h$, and $v$ are growth models

Sometimes, no prior information is available from which the random parameters can be predicted. In this case, the mixed-effects model with the random parameters set to 0 is not the same as the population average model and will give biased predictions. Instead, the population average model, fit without random effects, should be used.

### Conclusions

China-fir trees play an important role in ecological service functions in Fujian Province of southeastern China. In order to estimate accurately the ecological values, five theoretical growth equations were evaluated for estimating the $d$, $h$ and $v$ growth for China-fir trees grown in monospecific plantations. The equations can be evaluated for both a biological perspective and for statistical meaning. All five equations and the Korf and Richards equations in particular have commonly and successfully been used to model individual-tree growth. A multi-level NLME models based on the Korf and Richards equations, with variance functions and correlation structures, were used to estimate $d$, $h$ and $v$ growth of individual trees; this approach was necessitated by the hierarchical structure of the experimental design and the autocorrelated increment data. The results showed that the NLME models (Eqs. 22–24) with random effects provided a better fit than the other models (Eqs. 13–15) (Table 6, Figs. 3 and 4). For the other two groups of data, the adj-$R^2$ of the NLME models were more than 0.9 and better than the results of the NLMs (Table 7). Therefore, we recommend using a nonlinear mixed-effects model to estimate individual-tree growth and yield models.
Acknowledgements. This work was supported by World Bank Loan Project for Sand Prevention and Control on East Coast of the yellow river in Ningxia [grant numbers P121289]; the Independent Subject of Ningxia Academy of Agriculture and Forestry Sciences “Study of Carbon Sequestration Ability and Potential of Planted-forest in the arid aeolian sand region, Ningxia Province”; the National Natural Science Foundation of China [grant numbers 71663041]; and the Theoretical Economics First-class Discipline of Ningxia University [grant numbers NXYLXK2017B04].

REFERENCES


