2018

Height-Diameter models for Eucalyptus sp plantations in Brazil

Ferraz Filho, A C
FapUNIFESP (SciELO)

Tieteelliset aikakauslehtiartikkelit
© Authors
CC BY http://creativecommons.org/licenses/by/4.0/
http://dx.doi.org/10.1590/01047760201824012466

https://erepo.uef.fi/handle/123456789/6721
Downloaded from University of Eastern Finland's eRepository
Antonio Carlos Ferraz Filho¹, Blas Mola-Yudego²³, Andressa Ribeiro¹, José Roberto Soares Scolforo⁴, Rodolfo Araújo Loos⁵, Henrique Ferraço Scolforo⁶

HEIGHT-DIAMETER MODELS FOR Eucalyptus sp. PLANTATIONS IN BRAZIL


HIGHLIGHTS

Height-diameter models are provided for eucalyptus plantations in Brazil.

A wide variety of stand conditions (age and tree densities) were covered in the database.

Mixed modeling was used due to the hierarchical structure of the data (trees in plots).

The models use diameter, age, and dominant height as predictor variables.

ABSTRACT

The use of height-diameter models have important applications in the management of eucalyptus plantations. The aim of this paper is to provide adequate height models based on the diameter and stand variables that could be easily used and applied to plantation’s management. The data was based on tests located in Espírito Santo and Bahia. A total of 312 plots, including 35 different treatments (different initial planting densities, number and severity of thinning operations, fertilization regimes, among others), and four different clones. This resulted in 2,407 conditions for 85,608 tree measurements. The constructed equations were based on mixed-effect models, in order to accommodate possible autocorrelations, to address variations between sites and to obtain locally calibrated estimates. Among the different model structures and combination of variables tested and presented, the best model showed a coefficient of determination of 0.953 and 0.988, for the fixed and fixed+random parts, respectively. The model included a random factor for measurements-within-tree and tree-within-plot. The bias of the fixed part was 0.208 cm (0.0659%). We believe the models in this study can have broad applications in management and inventory methods applied to plantations.

¹ Federal University of Piauí - Bom Jesus, Piauí, Brazil
² University of Eastern Finland - Joensuu, Finland
³ Swedish University of Agricultural Sciences - Uppsala, Sweden
⁴ Federal University of Lavras, Lavras, Minas Gerais, Brazil
⁵ Fibria Celulose S.A, Aracruz, Espírito Santo, Brazil
⁶ North Carolina State University, Raleigh, North Carolina, United States of America
INTRODUCTION

Eucalyptus plantations are distributed worldwide and are destined for the production of several goods, including: charcoal, pulp and paper, construction timber, firewood, honey, essential oil, ornamental, carbon sequestration, and solid wood products (Ferraz Filho et al., 2014; Magnano et al., 2016). In Brazil, eucalyptus plantations constitute most of the plantation areas, primarily due to the high growth rates and advanced silvicultural techniques applied in the management of these areas (Gonçalvez et al., 2008). Currently, Brazil holds about 7.5 million hectares of eucalyptus plantations, the majority destined for energy and pulping purposes, and in a smaller degree for solid wood products (IBGE, 2015).

This large area, as well as the important role of these plantations in economic, energetic and environmental considerations, underlines the importance of proper monitoring and measuring methods (Raimundo et al., 2017). In this context, height and diameter measurements are fundamental for the planning and management of these plantation enterprises, which are used for estimating timber volume, site index or carbon budgets, among others.

However, whereas diameter records of the trees can be retrieved at a low cost and high accuracy, tree height is a variable often considered more expensive and difficult to retrieve. For instance, field inventory of commercial eucalyptus plantations in Brazil usually advocate measuring the heights of 5 to 15 trees (plus the height of the dominant trees), from plots with a total of 20 to 80 trees (Guimarães et al., 2009; Moreira et al., 2015). The effective correlations between tree height and diameter it is thus used to formulate models to estimate the heights of the remaining trees in the stand, reducing the costs associated to forest inventory (Paulo et al., 2011).

Finding accurate and easy to apply height-diameter relations are therefore important tools in inventory design and monitoring and in forest growth simulators, where individual tree diameter or stand level growth is estimated based on site productivity and stand characteristics (e.g. Palahí et al., 2003; Adame et al., 2008; Ahmadi et al., 2013). However, the height–diameter relationship is neither constant between stands nor over time (Curtis, 1967), which it is often addressed by creating models including additional variables to account for the stand dynamics. In this sense, plantation’s age has often been used (e.g. Curtis, 1967; Zhang et al., 2016) as well as other surrogates like the quadratic median diameter of the stand (e.g. Mehtätalo, 2004a; Schmidt et al., 2011).

In general, studies dealing with height-diameter models in fast growing plantations have been common (e.g. Soares and Tome, 2002; Crescente-Campo et al., 2010; Hjelm et al., 2015). In Brazil, the fact that eucalyptus plantations are mainly destined for pulping or energy purposes, as well as the fast growth conditions of the country, means that the applicability of the proposed models is limited to young plantations (around 7 years or less) grown in high densities (around 1,111 trees per hectare). For instance, the following works all present height-diameter models for eucalyptus grown in Brazil using data from stands ages 8 years and younger: Ribeiro et al., 2010; Thiersch et al. 2013; Miranda et al., 2014; Oliveira et al. 2014; Mendonça et al., 2015; Souza et al., 2016. While there are some exceptions of studies dealing with older plantations (Azevedo et al., 2011; Oliveira et al., 2015) these are generally from small plantations that have not received active management operations (such as thinning or fertilization).

The present study aims at producing stem height estimations for eucalypt plantations in Brazil, covering different clonal material of the hybrid Eucalyptus grandis x E. Urophylla, based on a large pool of measurements along several years. The final models are meant to be easily implemented in forest inventory and management applications, and at the same time, result in locally calibrated estimates high accuracy even when data is scarce; thus contributing to a cost-efficient planning of future wood production.

MATERIAL AND METHODS

Origin of data

The data consisted of large-scale thinning tests located in two Brazilian states, Espirito Santo (grid reference 19° 42’ S 40° 12’ W) and Bahia (located in a 40km radius with central grid reference 17° 58’ S 39° 42’ W). A total of 312 plots in eight different experiments were evaluated, totalling about 35 different treatments. These treatments tested different initial planting densities (555 to 1,111 trees per hectare), number of thinning operations (zero to 2), ages of thinning (beginning at 2.5 up to 7 years), different remaining number of trees (150 to 450 trees per hectare), fertilization regime (with and without fertilization after thinning), and four different clones. Half of the experiments were installed in 1999 (last measurement age at 11.5 years) and the other half in 2003-2004 (last measurement age at 6-7 years). More information on the stand establishment can be obtained in Ferraz Filho (2013). The wide variability in stand characteristics ensured a representative database (Figure 1, Table 1) of different eucalyptus management options. This resulted in 85,608 measurements of 20,028 trees in 312 plots.
RESULTS

Several candidate models were derived from the model structures considered, including plot level random factors in the slope and intercept, and additional predictors, such as: mean diameter, median diameter, mean height, dominant height, density, tree clone and age. Models with random factors both in the slope and in the intercept were difficult to compute and could not be calculated, therefore, the final models were selected with the random factors applied to the intercept. The models included different transformations and the interactions between candidate parameters, as justified by the measurements and their relationships (Figure 2).

Among the predictors tested, age (Equation 3, model 1) as well as clonal material, age and dominant height were selected for the final version of the models (Equation 4, model 2) as additional predictors to describe the stand conditions, proving strong relationships, where $h$ is the measured height (m), $d$ is the dbh (cm), age is the age (years) and $h_{dom}$ is the dominant height of the plot. Subscripts $i$, $j$ and $k$ refer to tree $i$ growing in plot $j$ in measurement $k$. This data structure may present autocorrelations (measurements from the same tree, and measurements from trees in the same plots share obvious common allometric features, which does not fulfill the assumption for independent measurements). To address this hierarchical structure of data a mixed model approach was applied, including random factors at plot level.

$$\ln(h_j - 1.3) = \beta_0 + \beta_1 d_{ijk}^{-1}$$ \hspace{1cm} [1]

$$\ln(h_{ijk} - 1.3) = \beta_0 + \beta_1 \ln(d_{ijk})$$ \hspace{1cm} [2]

Different predictors were considered (Table 1) to improve the predictive power of the models and to correct potential lack of fit, including mean diameter, median diameter, dominant height, mean height, density, clone and age, among others. In this study we considered the dominant height as being the mean height of the 100 thickest trees per hectare that presented no visible defects. The predictors were selected in order to reflect stand characteristics that would affect the height-diameter relationship, to have a good predictive performance and to result in a parsimonious model, easy to measure and to apply. All parameter estimates had to be significant at the 0.05 level, and the resulting residuals had to indicate a non-biased model. To convert the logarithmic predictions of the models to the arithmetic scale, an empirical estimator for bias correction was used, based on the ratio of the mean height and the mean of the back-transformed predicted height from the model (Snowdon, 1991).

All models were evaluated quantitatively by examining the magnitude and distribution of the residuals for the variables (mean diameter, median diameter, mean height, dominant height, density and age) aiming at detecting obvious dependencies or patterns that indicated systematic discrepancies. To determine the accuracy of the predictions, absolute and relative biases and root mean square errors (RMSEs) were used. The calculations were made using the package nlme (Pinheiro et al., 2016) of the R software (R Core Team, 2014).

Data Methods

The models considered several structures, based on the linearized version of the Korf equation (Korf, 1939) and the allometric model. The general structures of the models considered were, where $h$ is the height (m) and $d$ the dbh (cm). $\beta_0$, ...$n$ are parameters to be estimated and subscripts $i$ and $j$ and $k$ refer to tree $i$ growing in plot $j$ in measurement $k$. This data structure may present autocorrelations (measurements from the same tree, and measurements from trees in the same plots share obvious common allometric features, which does not fulfill the assumption for independent measurements). To address this hierarchical structure of data a mixed model approach was applied, including random factors at plot level.
The parameter estimates were significant in all cases (Table 2). Model 1 presented $R^2 = 0.877$ and $R^2 = 0.955$ for the fixed part, and including the random factors, respectively. Model 2 showed a higher predictive power, with equivalent $R^2 = 0.953$ and $R^2 = 0.983$ (Figure 3). The inclusion of individual intercepts for each clone or subspecies did not improve the overall predictive performance, but reduced the absolute bias from 0.319 to 0.208 cm. The relative values were 0.0659% and 1.95% for bias and RMSE, respectively.

The bias of the fixed part of the model, as well as the final predictions, were examined by plotting the residuals as a function of the predicted variable and predictors of the model (Figure 4). The inclusion of additional variables related to the stand conditions such as the mean dominant height and age improved the distribution of the residuals, reduced the bias and increased the predictive power of the models. No obvious dependencies or patterns that indicate systematic trends among the residuals and the independent variables included were found in the fixed part of the final models selected. The resulting ratios for logarithmic correction were 1.001835 and 1.001059, and the coefficients of determination of the back-transformed data resulted in 0.962 and 0.988 for model 1 and model 2, respectively.

Although both models showed similar results, there were differences in the predictions (Figure 5), particularly concerning younger and older trees. The inclusion of random parameters permitted individual predictions for each plot and tree, which could be used to produce locally

### TABLE 2

<table>
<thead>
<tr>
<th>model</th>
<th>parameter</th>
<th>SE</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_0$</td>
<td>3.080</td>
<td>0.005</td>
<td>608.192</td>
</tr>
<tr>
<td></td>
<td>$\beta_1$</td>
<td>-7.316</td>
<td>0.026</td>
<td>-282.164</td>
</tr>
<tr>
<td></td>
<td>$\beta_2$</td>
<td>0.302</td>
<td>0.001</td>
<td>363.327</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{plot}$</td>
<td>0.077</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\sigma_{tree in plot}$</td>
<td>0.032</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\sigma_{res}$</td>
<td>0.064</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$R^2$ (fixed)</td>
<td>0.877</td>
<td></td>
<td></td>
</tr>
<tr>
<td>model 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$R^2$ (all)</td>
<td>0.955</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\beta_3$</td>
<td>0.585</td>
<td>0.0085</td>
<td>72.780</td>
</tr>
<tr>
<td></td>
<td>$\beta_4$</td>
<td>0.601</td>
<td>0.0094</td>
<td>72.780</td>
</tr>
<tr>
<td></td>
<td>$\beta_5$</td>
<td>0.601</td>
<td>0.0083</td>
<td>72.780</td>
</tr>
<tr>
<td></td>
<td>$\beta_6$</td>
<td>0.591</td>
<td>0.0093</td>
<td>72.780</td>
</tr>
<tr>
<td></td>
<td>$\beta_7$</td>
<td>-5.950</td>
<td>0.022</td>
<td>-271.160</td>
</tr>
<tr>
<td></td>
<td>$\beta_8$</td>
<td>-0.085</td>
<td>0.001</td>
<td>-61.487</td>
</tr>
<tr>
<td></td>
<td>$\beta_9$</td>
<td>0.929</td>
<td>0.003</td>
<td>317.367</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{plot}$</td>
<td>0.025</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\sigma_{tree in plot}$</td>
<td>0.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\sigma_{res}$</td>
<td>0.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$R^2$ (fixed)</td>
<td>0.953</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$R^2$ (all)</td>
<td>0.983</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
calibrated models with a minor number of new records, such as presented by Trincado and Burkhart (2006). However, most of the predictive improvements due to the random factors were explained by the fixed part when the dominant height was included (model 2).

**DISCUSSION**

The management and monitoring of forest plantations require accurate and reliable predictions of tree height. The data included in this study covers a wide range of different eucalyptus spacing and thinning practices, from very dense unthinned plots to plots thinned down to 150 trees per hectare at age 2.5 years. The broad conditions considered as well as the extensive amount of data (85,608 measurements) provides a solid basis for modelling and guarantees that the models can be applied to many different management scenarios, from short rotation stands destined for cellulose or energy to longer rotations for solid wood products.
Concerning the model structure, the approach taken considered a linear model. In general, height-diameter relations have been considered to approach a sigmoidal shape (Peng et al., 2001; Gómez-García et al., 2015; Costa et al., 2016). However, despite the non-linear nature of the relationship, the use of linearization has been used in many studies (Mehtätalo et al., 2015), since it presents many advantages given its simplicity. The hierarchical structure of the data, with multiple measurements for the same tree, and trees in the same plot, implied a potential autocorrelation for records within each group (Palahí et al., 2003). The fitting approach was based on a mixed model, including fixed and random factors, to accommodate this data structure, which is common in forest modelling (e.g. Palahí et al., 2003; Trasobares et al., 2004, Ferraz Filho et al., 2016).

The mixed model approach allows the variability detected between plots, locations and trees to be integrated in the predictions, after defining a common and fixed functional structure (Lindstrom and Bates, 1990), and helps address the problems derived from autocorrelation. By using their predictive potential, the proposed models can easily be calibrated to local conditions in new areas by retrieving a few new records describing the plot or grouping factor (see Mehtätalo, 2004b). These properties have made the use of mixed models a popular approach with height-diameter, with good predictive performance; both for commercial forest species, like Pine, Spruce and Birch in northern countries (Hökkä, 1997; Lappi, 1997, Mehtätalo, 2004a; Mehtätalo, 2005, Schmidt et al., 2011; Sharma and Breidenbach, 2015), in Central Europe (Sharma et al., 2016) or in Southern Europe (Calama and Montero, 2004); and in the case of plantations, for Eucalyptus (Castedo-Dorado et al., 2006; Crecente-Campo, 2010), loblolly pine (Coble; Lee, 2011), poplar (Hjelm et al., 2015), among others.

Concerning the variables, age and dominant height are commonly used variables to help explain the height-diameter relationship (e.g. Paulo et al., 2011; Retslaff et al., 2015; Zhang et al., 2016). Additional variables, such as fertilization and thinning regime, were considered when designing the model; Ferraz Filho (2013) showed that differences of up to 4.5 m were found in 11.5 year-old plots that underwent very early thinning and unthinned plots, with the thinned plots achieving higher values. However, it was observed that these additional variables effectively influenced the behaviour of dominant height, and did not contribute to explain the predicted height when included in the model, thus, we assumed that the sensitivity of the variable dominant height to the stand’s silvicultural management would be adequate to describe
the height-diameter relationship. Another advantage of using dominant height as a predictor variable is its response to climatic conditions and the possibility to explain this variation using regression models (e.g., Ferraz Filho et al., 2011; Scolforo et al., 2013). Other variables related to the plot’s metrics, such as diameter, were explained using age (Mehtatálo, 2005), and could be considered interchangeable.

The use of clonal material as a categorical variable proved significant when included in the model, which was expected since different clonal materials are expected to have different height growth behaviour (Calegario et al., 2005; Scolforo et al., 2017). In this case, however, the variable contributed to a far lesser extent than age or dominant height, and it was justified to present a model version with a common intercept, since it can expand the applicability of the model in those cases with limited information concerning the clonal material.

In this sense, one of the objectives of the models was their applicability, prioritizing simpler approaches and a limitation in the number of variables to be retrieved. More complex approaches can also address the complexities of the data and their analysis. For instance, alternative computer methods, such as Decision trees, Neural Networks and Random Forest (Binoti et al., 2013; Lima et al., 2017) are being tested to substitute traditional statistical models to describe tree development. While these methods are suggested to be more precise than regression models (Vendruscolo et al., 2015), they have the disadvantage of being more difficult to implement and have potential, under extrapolation procedures, to provide illogical estimates (Sabatia and Burkhart, 2014).

CONCLUSION

The variables dbh, age and dominant height proved to be the best variables to predict tree height using a mixed models approach. The models showed a good predictive performance, and their ease of application constitutes one of the main advantages of the present models, easily implementable in forest inventory procedures or growth simulators.

ACKNOWLEDGEMENTS

We would like to express our gratitude to the FinCEAL Plus Latin America and the Caribbean programme. We are also grateful to Fibria S.A. for providing the data.

REFERENCES


