CALIBRATION OF A MIXED-EFFECT STEM TAPER MODEL FOR TECTONA GRANDIS

Lanssanova RL^{1, *}, Machado SA¹, Orso GA², Pelissari AL², Figueiredo Filho A¹ & Silva FA²

¹Department of Forest Engineering, Midwestern State University, Irati, Brazil, BR 153, Km 7, Riozinho, 84500-000 ²Postgraduate Program of Forestry Engineering, Department of Forest Engineering, Federal University of Paraná, Curitiba, Paraná, 80060-000 Brazil

*luciano.lanssanova@jna.ifmt.edu.br

Submitted July 2019; accepted March 2020

The assumption of independence between observations has been frequently violated in forest literature. A part of it is related to the fact that parameters estimated by least squares and predictions made by these models are impartial in the presence of autocorrelation. Since, the absence of autocorrelation between observations is one of the basic assumptions of regression analysis, this study aimed to assess the calibration of a mixed effect model to estimate diameter and volume of the stem of *Tectona grandis* trees. The log volumes of 509 trees were calculated using relative method, and initially the variable-exponent taper model was fitted (Kozak 2004). For mixed effect modelling, the trees were considered as random effect. A Bayesian calibration was performed on the diameters of 18 trees which were not part of the training data set, and for these trees nine height combinations were tested along the stem. The calibration results were assessed using root mean square error and graphic analysis of residuals autocorrelation. The calibration led to precise estimates of diameter and volume along the stem. The use of the diameter at breast height (DBH) as prior information on stem taper was efficient in reducing residual autocorrelation.

Keywords: Bayesian estimator, autocorrelation, stem volume, teak, random effects

INTRODUCTION

Tectona grandis plantations are an important source of good quality timber for industries, as its resistance and workability are similar to mahogany or cedar. About 2.0 to 2.5 million cubic meters of teak are annually harvested from natural and seminal stands. Due to the declining supply from natural forests, the market for cultivated teak timber is promising (Kollert & Kleine 2017).

Accurate tree volume estimates are a crucial task in forest research. Their obtainment for any merchantable limit remains a challenge for forest managers because it can hardly be obtained in the field (Arias-Rodil et al. 2015, Bouriaud et al. 2019). Thus, research related to volume modelling and stem taper are fundamental to gather reliable and precise information of forest products (Macfarlane & Weiskittel 2016). To achieve, it is often necessary to use sophisticated modelling techniques such as taper equations (Fonweban et al. 2011).

Taper equations provide more useful information compared to standard volume functions, since they can (1) estimate diameter at any point along the longitudinal profile of the stem, (2) estimate the merchantable height at any given diameter, and (3) calculate the volume of individual sections of any length at any height (Kozak 2004). Seeking this multiplicity of responses, many researchers developed several taper equations in the last few years, intended to precisely estimate diameters, heights and volumes along the stem of trees (Schoepfer 1966, Max & Burkhart 1976, Kozak 1988, Kozak 2004).

To fit stem taper models, several diameter measurements are taken along the bole, following a natural and correlated hierarchical structure (Garber & Maguire 2003). The mixedeffects models were then introduced in forest mensuration as a new approach to model the longitudinal profile of stems (Cao & Wang 2011, Gomez-García et al. 2013). In addition to the fixed effects that are common to all individuals, they include parameters with random effects that are unique to individuals or groups within the data set (Calama & Montero 2006, Cao & Wang 2011, Gomez-García et al. 2013, Ruslandi et al. 2017, Ferraz-Filho et al. 2018). With these techniques, it is possible to develop more accurate models and obtain more precise responses of forest productivity.

Forest literature has shown that mixed-effects models can improve predictive power and reduce residual bias, leading to smaller errors than traditional fixed effects models (Calama & Montero 2006, Cao & Wang 2011, Bouriaud et al. 2019). This kind of modelling allows autocorrelation to be, at least partially, explained by the inclusion of random effects. However, the magnitude and distribution of these random effects are rarely reported (Bouriaud et al. 2019). The mixed effects approach can also predict a specific response for a new tree through fixed parameters and random effects, if at least one diameter measurement is available (Arias-Rodil et al. 2015).

The mixed-effects modelling can provide both mean and subject-specific responses, based on the random effect considered. Using the trees as a random effect, a mean response will be obtained if the fitted fixed parameters are used, and a specific response will be obtained for each measured tree with the inclusion of random effect. The individual responses are acquired from a calibration process based on a Bayesian estimate approximation (Vonesh & Chinchilli 1996). It requires the measurement of the diameter in certain stem positions for the trees in which the random effect will be predicted, in order to compare with the diameters estimated by fixed parameters of the model. These are then used in the residual matrix, required by the calibration process. Thereby it is necessary to investigate different points of measurement of diameters along the stem, to obtain prior information for Bayesian estimator.

Since the absence of autocorrelation between observations is one of the basic assumptions of regression analysis, this study aimed to assess the calibration of a mixed-effects model to estimate diameters and volumes along the stem of *Tectona grandis*. The study was designed with the hypothesis that the use of the diameter at breast height (DBH) as prior information on taper equations of the calibration process reduces residual autocorrelation, and provides accurate diameter and volume estimates.

MATERIALS AND METHODS

The study was carried out in five 25-yearold *Tectona grandis* seminal stands, located in Brasnorte, Mato Grosso State, Brazil. The Brasnorte region has a tropical rainforest climate, with temperature varying from 4 to 40 °C, and mean annual precipitation of 2250 mm, distributed between November and March (Alvares et al. 2014).

The log volumes of 509 trees were calculated using relative height method, taking diameter measurements along the stem at 0, 1, 2, 3, 4, 5, 15, 25, 35, 45, 55, 65, 75, 85 and 95% of total stem height, and the volume of each section was calculated via Smalian method. The variableexponent taper equation from Kozak (2004) was then fitted using the nonlinear least squares (nls) function available in the R software (R Core Team 2017). The significance of the fitted parameters was tested with a T-test at 5% probability level.



where β_i = parameters to be estimated, d = diameter at breast height (cm), h = total height (m), = aboveground height at relative sections (m), X = h_i/h ; = diameter at a given height (cm).

The final form of the stem taper model was chosen based on the percentage standard error (S_{yx} %), pseudo-R² (the squared Pearson's correlation between measured and predicted values) and Akaike information criterion (AIC) (Burnham & Anderson 2002).

For the mixed effect model, single tree was considered as a random effect. Combinations of three parameters were set as random, and the remaining were set as fixed, using the nonlinear mixed effect (nlme) package (Pinheiro & Bates 2000). To solve the resulting dependency between observations originated from the consecutive diameter measurements along the stem of the trees, a first order autoregressive structure was used with the *correlation* function, available in the *nlme* package (Garber & Maguire 2003, Li & Weiskittel 2010, Rodríguez et al. 2013, Schröder et al. 2014).

Since Kozak's variable-exponent model cannot be directly integrated, the estimated log volumes at every measured section were calculated by numerical integration with integrate R function (Kozak 2004). The indicated combination of random parameters was chosen by comparing the estimates of diameter and volume of both mixed model and the basic formulation of the taper model with fixed effects. To compare the models, the relative root mean square error (RMSE%) was used as a measure of precision gain.

To assess the calibration process of mixed modelling, the indicated combination of random parameters was used, based on the aforementioned precision measure. The calibration of diameters was performed upon a set of 18 trees which were not part of the training data set, and the prediction of random effects was obtained from the Bayesian estimator (Vonesh & Chinchilli 1996, Trincado & Burkhart 2006, Meng & Huang 2009, Yang et al. 2013):

 $\mu_{ij} = \widehat{D}Z_i^T (Z_i \widehat{D}Z_i^T + \widehat{R}_i)^{-1} \operatorname{res}_{ij}$

where μ_{ij} = vector of random effect parameters, \widehat{D} = variance-covariance matrix for the random effect parameters, Z_i = partial derivatives matrix with respect to the random effect parameters, \widehat{R}_i = residual variance matrix, res_{ij} = model residuals at the i-th position of the j-th tree, defined as the difference between the observed diameter at a given bole height and the predicted diameter at that same height with fixed parameters of the equation.

Since the Bayesian estimator requires measurement of diameters at several positions of the stem to obtain residual matrix, the calibration was performed testing nine combinations of diameters along the stem of new trees (Vonesh & Chinchilli 1996):

i) 0 and 4% of total height, ii) 1 and 4% of total height, iii) 1 and 15% of total height, iv) 1 and 35% of total height, (v) 1% of total height and 1.30 m, vi) 4 and 35% of total height, vii) 4 and 15% of total height, viii) 1.30 m and ix) 1.30 m and 35% of total height.

The diameter and volume estimates along the bole for the 18 trees based on calibration process were assessed using RMSE%, as well as graphics analysis of the autocorrelation and the average stem profiles.

RESULTS

The parameter estimates were significant at p < 0.05. The values of pseudo-R², Syx (%) and

AIC statistics were 0.95, 8.55 and 31,050.41, respectively. The autocorrelation between residuals for the evaluated model is shown in Figure 1, which presents the scatter between ε_i and ε_{i-1} , ε_{i-2} , ε_{i-3} , ..., ε_{i-n} . A model fitted with only fixed parameters retains a residual autocorrelation pattern and diverges from an ideal situation where the pairs of residuals are centered in the origin (Gujarati & Porter 2008).



Figure 1 Residual autocorrelation of the diameter measurements along the bole for standard fixed-effect model

The combination of parameters β_0 , β_4 and β_8 provided a difference of 45.03% (3.85 percentage points) and 73.50% (10.43 percentage points) to diameter and volume estimates, respectively, in RMSE statistics, with respect to Kozak's (2004) model in fixed formulation.

The RMSE% of the diameter and volume estimates along the relative height combinations tested in the calibration process is shown in Table 1. The combination of 4 and 35% of total height for the calibration of mixed model resulted in the lowest values of RMSE%, both for diameter and volume estimates along the stem, while the combination of 1 and 4% of total height provided the biggest values.

Comparing the mean of observed and estimated values (Figures 2 and 3), similarities were observed among the combination profiles of relative heights tested in the calibration. In fact, only few profiles (1–4% and 1%–1.30 m) showed biased estimates at the top of the stem, underestimating the diameters and volumes. Autocorrelation analysis among residuals of different height combinations used in

Tested heights for — calibration	RMSE (%)	
	Diameter	Volume
0 and 4%	7.71	12.58
1 and 4%	15.74	20.64
1 and 15%	8.97	13.60
1 and 35%	6.15	7.23
1 and 1.30 m	11.56	14.76
4 and 15%	10.44	20.82
4 and $35%$	5.69	6.61
1.30 m	6.48	9.65
$1.30~\mathrm{m}$ and 35%	7.75	7.32

Table 1Root mean squared error (RMSE%) of diameter and volume
estimates along the bole at different heights tested for mixed-
effect model calibration

the calibration process (Figure 4) showed reduced error in combinations 1-5%, 4-35%, 1.30 m and 1.30 m-35%.

The RMSE% of the calibration process at relative heights along the stem (Figure 5) used only the diameter at 1.30 m when calibrating

new individuals. The lowest stem portions of the 18 trees used for calibration contained the most noble products, where the RMSE% statistic remained below 6 and 9% for both diameter and volume, respectively.



Figure 2 Mean stem profiles of observed and estimated diameter at different heights tested in the calibration; from (a) to (i): 0 and 4%, 1 and 4%, 1 and 15%, 1 and 35%, 1% and 1.30 m, 4 and 15%, 4 and 35%, 1.30 m, 1.30 m and 35%



Figure 3 Mean profiles of observed and estimated volume at different heights tested in the calibration. From (a) to (i): 0 and 4%, 1 and 4%, 1 and 15%, 1 and 35%, 1% and 1.30 m, 4 and 15%, 4 and 35%, 1.30 m, 1.30 m and 35%



Figure 4Residual autocorrelation of diameter and volume estimates at different heights tested in the
mixed-effect model calibration. From (a) to (i): 0 and 4%, 1 and 4%, 1 and 15%, 1 and 35%, 1%
and 1.30 m, 4 and 15%, 4 and 35%, 1.30 m, 1.30 m and 35%

DISCUSSION

The fit statistics and precision measures showed good results in general, with a low value of standard error and high pseudo-R² value. For the latter, a value of 0.95 was obtained, indicating that the equation was able to properly explain the diameter estimates along the stem. Although this statistic is not a proper measure to assess nonlinear models, it is easy to interpret and is often



Figure 5 RMSE (%) for (a) diameter and (b) volume of the stem in the calibration process, using only the height of 1.30 m to calibrate new individuals

used to evaluate models in forest mensuration (Spiess & Neumeyer 2010, Schröder et al. 2014). Besides, many computer programs calculate R² for non-linear fits, which unintentionally confirms its recurrent use (Spiess & Neumeyer 2010).

The Kozak's (2004) variable-exponent model fit presented residual autocorrelation (Figure 1) due to its consecutive measurements in each tree. It is reasonable to expect that observations of each tree are spatially correlated, and that the assumption of independence of residuals is violated (Rojo et al. 2005). Gujarati and Porter (2008) pointed that, although the parameter estimates of least squares remain unbiased and normally distributed, they lose efficiency when residual autocorrelation is present. Although the parameter estimates are significant, the standard error of each parameter might be inappropriate and biased, leading to biased estimates of the response variable (Calama & Montero 2006, Tang et al. 2016).

In this research, a random effect was introduced in three parameters of the taper model. However, it is possible to associate a random effect in every parameter estimate, but the excessive number of parameters might hinder the convergence when fitting the model. Previous research limited including two or three random parameters, aiming the fit convergence (Bouriaud et al. 2019).

Fonweban et al. (2011) used mixed-effects modelling to evaluate the stem profiles of *Picea sitchensis* and *Pinus sylvestris* using three random parameters, showing 45 to 63% accuracy improvement in traditional method.

The mixed-effects modelling applied in Kozak's (2004) variable-exponent model was

also tested by Schröder et al. (2014), which emphasised that, despite the practical use of least squares in taper equations modelling, its predictive power is surpassed when using mixedeffect models.

This kind of modelling can provide both a mean response curve and subject-specific curves for groups or individuals where the random component is applied (Schabenberger & Pierce 2001, Westfall 2016). However, for a mean curve estimated by fixed parameters, the curve is conditioned to random effects. Therefore the model can capture part of the dependence between observations, considering the autocorrelation in each individual in its covariance structure (Verbeke et al. 2014), such as the tree in the study.

For the remaining parameters β_1 , β_2 , β_3 , β_5 , β_6 , and β_7 , where no random component was included, the mixed modelling improved their efficiency, since the variance has decreased, thereafter resulting in precision gain.

Despite the combination of 4 and 35% of total height provided the lowest values of RMSE (%) in calibration of the mixed model (Table 1), the use of the DBH (1.30 m) resulted in good values when estimating diameter and volume, with RMSE below 10%. This implied that, in field work, the measurement of diameter at many points of the stem, as prior information to predict the random effects, would not be necessary. Moreover, field measurements would not be time-consuming or costly, making the calibration technique feasible.

When evaluating the residual autocorrelation, Garber & Maguire (2003) pointed out that even for calibrated models, when predicting the random effects of new individuals, the residual autocorrelation may not be completely corrected. It is reasonable to assume that equally spaced measurements are spatially correlated along the stem of a tree, which reflects in the residual estimates.

The diameter and volume estimates at the lowest part of the stem presented an average error, inferior to 9% (Figure 5), and reduced the residual autocorrelation (Figure 4). This confirmed the assumption that including random effects of a new tree, based on the Bayesian estimator, improved the predictive power of the model, mainly the lowest part of the stem (Trincado & Burkhart's 2006). Calibrating a mixed model with Bayesian estimator can be more accurate and practical than techniques that incorporate auxiliary variables into stem profile models (Dean 2003, Henning & Radtke 2006).

In the current increasing intensity of forest management and decrease of merchantable diameter limits, more accurate techniques are required to determine the diameters and volumes along the bole of trees through calibration process, and to assess the productivity with respect to different regimes (Garber & Maguire 2003). The precise individual estimates and the fulfillment of the assumptions of regression analysis by mixed-effects models provide benefits to future forest managements.

CONCLUSION

The Bayesian calibration process provided accurate individual estimates of diameters and volumes along the bole of *Tecnona grandis* as a promising technique to predict random effects in taper models.

The DBH, as prior information stem taper to obtain residual matrix, was efficient in reducing residual autocorrelation. Hereafter, there will be no need to take more measurements along the bole in forest inventories.

REFERENCES

- ARIAS-RODIL M, CASTEDO-DORADO F, CÁMARA-OBREGÓN A & DIÉGUEZ-ARANDA U. 2015. Fitting and calibrating a multilevel mixed effects stem taper model for maritime pine in NW Spain. *Plos One* 10: 12. https:// doi.org/10.1371/journal.pone.0151297.
- BOURIAUD O, STEFAN G & SAINT-ANDRÉ L. 2019. Comparing local calibration using random effects estimation and Bayesian calibrations: a case study with a mixed effect

stem profile model. *Annals of Forest Science* 76: 3–12. doi: https://doi.org/10.1007/s13595-019-0848-5

- BURNHAM KP & ANDERSON DR. 2002. Model Selection and Multimodel Inference. Springer-Verlag, New York.
- CALAMA R & MONTERO G. 2006. Stand and tree-level variability on stem form and tree volume in *Pinus pinea* L.: a multilevel random components approach. *Investigación Agraria: Sistemas Recursos Forestales* 15: 24–41. doi:10.5424/srf/2006151-00951.
- ALVARES CA, STAPE JL, SENTELHAS PC, GONÇALVES JL & SPAROVEK G. 2014. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift* 22: 711–728. doi: 10.1127/0941-2948/2013/0507.
- CAO QV & WANG J. 2011. Calibrating fixed- and mixedeffects taper equations. *Forest Ecology and Management* 262: 671–673. https://doi.org/10.1016/j. foreco.2011.04.039.
- DEAN C. 2003. Calculation of wood volume and stem taper using terrestrial single-image close-range photogrammetry and contemporary software tools. *Silva Fennica* 37: 359–380. https://doi.org/10.14214/ sf.495.
- FERRAZ FILHO AC, MOLA-YUDEGO B, RIBEIRO A, SCOLFORO JRS, LOOS RA & SCOLFORO HF. 2018. Height-diameter models for *Eucalyptus* sp. plantations in Brazil. *Cerne* 24: 9–17. https://doi.org/10.1590/010477602018 24012466.
- FONWEBAN J, GARDINER B, MACDONALD E & AUTY D. 2011. Taper functions for Scots pine (*Pinus sylvestris* L.) and Sitka spruce (*Picea sitchensis* (Bong.) Carr.) in Northern Britain. *Forestry* 84: 49–60. https://doi.org/10.1093/ forestry/cpq043.
- GARBER SM & MAGUIRE DA. 2003. Modeling stem taper of three central Oregon species using nonlinear mixed effects models and autoregressive error structures. *Forest Ecology and Management* 179: 507–522. https:// doi.org/10.1016/S0378-1127(02)00528-5.
- GOMEZ-GARCÍA E, CRECENTE-CAMPO F & DIÉGUEZ-ARANDA U. 2013. Selection of mixed-effects parameters in a variable-exponent taper equation for birch trees in northwestern Spain. *Annals of Forest Science* 70: 707– 715. https://doi.org/10.1007/s13595-013-0313-9.
- GUJARATI DN & Porter DC. 2008. *Basic Econometrics*. McGraw-Hill: New York.
- HENNING JG & RADTKE PJ. 2006. Detailed stem measurements of standing trees from ground-based scanning lidar. *Forest Science* 52: 67–80. https://doi.org/10.1093/ forestscience/52.1.67.
- KOLLERT W & KLEINE M. (2017). The Global Teak Study Analysis, Evaluation and Future Potential of Teak Resources. International Union of Forest Research Organizations, Vienna.
- KOZAK A. 1988. A variable-exponent taper equation. Canadian Journal of Forest Research 18: 1363–1368. https://doi.org/10.1139/x88-213.
- KOZAK A. 2004. My last words on taper equations. *Forestry Chronicle* 80: 507–515. https://doi.org/10.5558/ tfc80507-4.
- LI R & WEISKITTEL AR. 2010. Comparison of model forms for estimating stem taper and volume in the primary conifer species of the North American Acadian Region. *Annals of Forest Science* 67: 302. https://doi. org/10.1051/forest/2009109.

- MACFARLANE DW & WEISKITTEL AR. 2016. A new method for capturing stem taper variation for trees of diverse morphological types. *Canadian Journal of Forest Research* 46: 804–815. https://doi.org/10.1139/ cjfr-2016-0018.
- MAX TA & BURKHART HE. 1976. Segmented polynomial regression applied to taper equations. *Forest Science* 22: 283–289. https://doi.org/10.1093/ forestscience/22.3.283.
- MENG SX & HUANG S. 2009. Improved Calibration of Nonlinear Mixed-Effects Models Demonstrated on a Height Growth Function. *Forest Science* 55: 238–248. https://doi.org/10.1093/forestscience/55.3.238.
- PINHEIRO JC & BATES DM. 2000. *Mixed-Effects Models in S and S-PLUS.* Springer-Verlag, New York.
- R CORE TEAM. 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/.
- RODRÍGUEZ F, LIZARRALDE I & BRAVO F. 2013. Additivity on nonlinear stem taper functions: a case for Corsican pine in northern Spain. *Forest Science* 59: 464–471. https://doi.org/10.5849/forsci.12-023.
- ROJO A, PERALES X, SÁNCHEZ-RODRÍGUEZ F, GONZÁLEZ-ALVAREZ JG & VON GADOW K. 2005. Stem taper functions for maritime pine (*Pinus pinaster* Ait.) in Galicia (Nortwestern Spain). *European Journal of Forest Research* 124: 177–186. https://doi.org/10.1007/ s10342-005-0066-6.
- RUSLANDI WP, CROPPER J & PUTZ FE. 2017. Tree diameter increments following silvicultural treatments in a dipterocarp forest in Kalimantan, Indonesia: a mixed-effects modelling approach. *Forest Ecology and Management* 396: 195–206. https://doi. org/10.1016/j.foreco.2017.04.025.
- SCHABENBERGER O & PIERCE FJ. 2001. Contemporary Statistical Models for The Plant and Soil Sciences. CRC Press, Boca Raton.

- SCHOEPFER W. 1966. Automatisierung dês Massen, Sorten und Wertberechung Stender Waldbestande Schriftenreihe Bad. [S.I.]: Wurtt-Forstl.
- SCHRÖDER T, ARNONI EC, VALÉRIO AF & LISBOA GS. 2014. Taper equations for *Pinus elliottii* Engelm. in Southern Paraná, Brazil. *Forest Science* 61: 311–319. https://doi.org/10.5849/forsci.14-054.
- SPIESS AN & NEUMEYER N. 2010. An evaluation of R² as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. *BMC Pharmacol* 10: 6. https://doi. org/10.1186/1471-2210-10-6.
- TANG X, CRUZADO CP, FEHRMANN L, ÁLVAREZ-GONZÁLEZ JG, LU Y & KLEINN C. 2016. Development of a compatible taper function and stand-level merchantable volume model for Chinese fir plantations. *PLoS ONE* 11: 1. https:// doi.org/10.1371/journal.pone.0147610.
- TRINCADO G & BURKHART HE. 2006. A generalized approach for modeling and localizing stem profle curve. *Forest Science* 52: 670–682. https://doi.org/10.1093/ forestscience/52.6.670.
- VERBEKE G, FIEUWS S & MOLENBERGHS G. 2014. The analysis of multivariate longitudinal data: a review. *Stat Methods Med Res* 23: 42–59. https://doi. org/10.1177/0962280212445834.
- VONESH EF & CHINCHILLI VM. 1996. Linear and nonlinear models for the analysis repeated measurements. CRC Press, Boca Raton.
- YANG Y, HUANG S, TRINCADO G & MENG SX. 2009. Nonlinear mixed-effects modelling of variable-exponent taper equations for lodgepole pine in Alberta, Canada. *European Journal of Forest Research* 128: 415–429. https://doi.org/10.1007/s10342-009-0286-2.
- WESTFALL JA. 2016. Strategies for the use of mixed-effects models in continuous forest inventories. *Environ Monit Assess* 188: 245. https://doi.org/10.1007/ s10661-016-5252-0.