# FITTING THREE PARAMETER GROWTH CURVES USING A NONLINEAR MIXED EFFECTS MODELLING APPROACH

***Sileshi F. Melesse*** [1]
School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Private
Bag X01, Scottsville 3209, Pietermaritzburg, South Africa.
e-mail: *Melesse@ukzn.ac.za*
***and***
***Temesgen Zewotir***
School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Private
Bag X54001 Durban 4000, Westville, South Africa.

***Key words:*** Dendrometers, loess, longitudinal study, maximum likelihood, random effects, variance-covariance.

***Summary:*** A nonlinear mixed modelling approach was applied to model individual tree diameter increment using three nonlinear growth functions for Eucalyptus tree plantations. The objective of the study is to develop a stem radial increment model for two clones of Eucalyptus tree, and compare their growth potential. Three nonlinear growth curves (Gompertz, logistic and asymptotic regression) were fitted to stem radius data. Estimations of parameters were made using the approximate likelihood functions. The estimators obtained from these approximate likelihood function are a combination of least squares estimators for nonlinear mixed effects models and maximum likelihood estimators from linear mixed effects models. The asymptotic regression model with three random effects for $\phi_1$, $\phi_2$ and $\phi_3$ appears well suited to represent the random effect covariance structure. The heterogeneous variance model that varies with tree age is found to be suitable model that characterize the within tree error variability. Clone has a significant effect on the asymptote of the asymptotic regression curve. The analysis suggests that GU clone on the average has a larger stem radial measurement than the GC clone during the entire juvenile stage.

## 1. Introduction

A cross sectional study may allow the comparison between sub-populations that happen to differ in age, but it does not provide any information about how individuals change over time. The assessment of within subject changes in response over time can only be achieved within a longitudinal study. A distinctive feature of longitudinal data is that observations within the same individual are correlated. Failure to account for the effect of correlation can result in an erroneous estimation of the variability

---

[1] Corresponding author.

*AMS: 62G20, 62F10, 62J02, 97K80*

of parameter estimates and hence in misleading inference. This interdependence can be modelled using mixed models. The current data set consisted of repeated measurements of the same subjects over time, therefore, a mixed effects models approach was adopted (Pinheiro and Bates, 1995; Verbeke and Molenberghs, 1997; Verbeke and Molenberghs, 2000; Meng and Huang, 2010; Fitzmaurice, Laird and Ware, 2011) in the analysis of the longitudinal data. Models for the analysis of such data recognize the relationship between serial observations on the same entity. Most of the work on methods of repeated measures data has focused on data that can be modelled by an expectation function that is linear in its parameters (Laird and Ware, 1982). Nonlinear mixed-effects models involve both fixed and random effects, in which some, or all, of the fixed and random effects occur nonlinearly in the model function.

Numerous nonlinear mixed-effects models have been proposed (Sheiner and Beal, 1980; Mallet, Mentré, Steimer and Lokiec, 1988; Lindstrom and Bates, 1990; Vonesh and Carter, 1992; Davidian and Gallant, 1992; Wakefield, Smith, Racine-Poon and Gelfand, 1994). Some of these like Davidian and Giltinan (1995) and Vonesh and Chinchilli (1996) offered overviews along with general theoretical developments and some examples of nonlinear mixed models. Lindstrom and Bates (1990) proposed a general nonlinear mixed-effects model for repeated measures data and define estimators for its parameters. These estimators are a combination of the least square estimators for nonlinear fixed effects models and maximum likelihood (or restricted maximum likelihood) estimators for linear mixed-effects models. Pinheiro and Bates (2000) presented a slight generalization of the nonlinear mixed models proposed by Lindstrom and Bates (1990). This generalization allows the incorporation of "time varying" covariates in the fixed effects or the random effects for the model. This general formulation is implemented in R Statistical software (R Core Team, 2013). The implementation in R allows the use of nested random effects and also permits the within group errors to be correlated and/or to have unequal variances. This general formulation was considered in this study. Nonlinear mixed models can be viewed as an extension of the linear mixed-model of Laird and Ware (1982) in which the conditional expectation of the response given the random effects is allowed to be a nonlinear function of the coefficients. It can also be regarded as an extension of nonlinear models for independent data (Bates and Watts, 1988) in which random effects are integrated in the coefficients to allow them to vary by group.

Nonlinear mixed-models are becoming increasingly popular (Wolfinger, 1999). They are applied in many fields of study such as agriculture, forestry, biology, ecology and biomedicine. According to Pinheiro and Bates (2000), the main reason why one would want to use nonlinear mixed-models is interpretability, parsimony and validity beyond the observed range of data. By increasing the order of the polynomial model, it is possible to get increasingly accurate approximations to the true, usually nonlinear, regression function, within the range of the data. However, these higher order polynomial models may result in multicollinearity problems and they also provide no theoretical considerations about the underlying mechanism producing the data. Nonlinear models, on the other hand, are often mechanistic, i.e., based on a model for the mechanism producing the response. As a consequence, the model parameters in nonlinear models generally have a natural physical interpretation. Even when derived empirically, nonlinear models usually incorporate known, theoretical characteristics of the data, such as asymptotes and monotonicity, and in these cases, can be considered as semi-mechanistic models. A nonlinear model generally uses fewer parameters than a competitor linear model, such as a polynomial, giving more parsimonious description of the data. Nonlinear models also provide more reliable predictions for the response variable outside the observed range of the data than, say, polynomial models would (Pinheiro and Bates, 2000). The objectives of this study were to develop models for stem radial increment of two Eucalyptus (*E. grandis* × *E. urophylla* and *E. grandis* × *E. camaldulensis* ) hybrid clones and compare their growth potential by comparing the parameters of the model.

# 2. Materials and Methods

## 2.1. Description of the data

The data used in this study are secondary data from Sappi Forest Research Center in Tweedie. Sappi started the dendrometer trial project in July 2001. A dendrometer trial focused on the growth of a *Eucalyptus grandis* × *E. urophylla* (GU) and an *E. grandis* × *E. camaldulensis* (GC) hybrid clones and the trial established on Sappi landholdings near the town KwaMbonambi, in the KwaZulu-Natal province of South Africa. The eucalyptus fiber research experiment was initiated in July 2001 and a huge database acquired. The results presented in this study are based on the data collected from April 2002 when trees were 39-weeks-old until August 2003 when they were 107-weeks-old. Using dendrometers, repeated measurements of stem radius were obtained, during this time, for a sample of 18 trees, nine from each clone. Nine trees per plot were selected from each clone for intensive monitoring of radial growth (Drew, 2004; Drew, Downes, Grzeskowiak and Naidoo, 2009). From the 18 sampled trees (nine per clone), longitudinal data of 1242 weekly radial measurements were obtained. The response variable investigated in this study was the weekly stem radius, which is of interest because it can be used to understand the underlying processes of fibre development in fast-growing Eucalyptus plantations. Some studies have been made from the data extracted from the same Sappi database. These are the studies by Drew et al. (2009) and the study by Melesse and Zewotir (2013a) and Melesse and Zewotir (2013b). However, none of these studies have considered the longitudinal aspect of the data.

## 2.2. The general nonlinear mixed model

The nonlinear mixed model can be viewed as a two stage model. In the first stage the $j^{th}$ observation on the $i^{th}$ individual is modelled as:

$$y_{ij} = f(\phi_{ij}, X_{ij}) + \varepsilon_{ij}, \quad i = 1,2,\ldots,M \quad \text{and} \quad j = 1,2,\ldots,n_i, \tag{1}$$

where $f$ is a nonlinear function of an individual specific parameter vector $\phi_{ij}$ and the predictor vector $X_{ij}$, and $\varepsilon_{ij}$ the normally distributed within-group error term. $M$ is the total number of individuals and $n_i$ is the number of observations on the $i^{th}$ individual. In the second stage the individual specific parameter vector is modelled as:

$$\phi_{ij} = A_{ij}\beta + B_{ij}b_i, \quad b_i \sim N(0, \psi),$$

where $\beta$ is a $p$-dimensional vector of fixed population parameters, and $b_i$ is a $q$-dimensional random effects vector associated with the $i^{\text{th}}$ individual (not varying with $j$), with variance covariance matrix

$\psi$. The matrices $A_{ij}$ and $B_{ij}$ are design matrices for the fixed and random effects respectively. It is further assumed that observations made on different individuals are independent and that the within group errors $\varepsilon_{ij}$ are independently distributed as $N(0,\sigma^2)$ and independent of the $b_i$. Three growth functions, Gompertz, logistic and the asymptotic regression curves were chosen, to replace the function $f$, in Equation (1). This is mainly because these models have meaningful parameters from a forestry perspective. Moreover, they may also be more reliable for predictions which comprise extrapolations beyond the range of data compared to conventional polynomial models. All three curves are considered to have three parameters and their description is given as follows.

## The Gompertz function

The three parameter Gompertz function can be expressed as:

$$y = \alpha \exp(-\beta \times \exp(-\gamma t)), \quad \alpha > 0, \quad \beta > 0, \quad \gamma > 0. \tag{2}$$

The limiting value as $t$ approaches infinity is $\alpha$. The starting value of $y$ at $t = 0$ is $\alpha \exp(-\beta)$, and with the restrictions on the parameters $0 < \alpha \exp(-\beta) < \alpha$. The representation of the Gompertz function in R Statistical Software is

$$y = \phi_1 \exp(-\phi_2 \times \phi_3^t). \tag{3}$$

The relationship between the parameters of the models in Equations (2) and (3) is that $\alpha = \phi_1$ , $\beta = \phi_2$ and $\exp(-\gamma) = \phi_3$. The representation in R is used, because the R Statistical Software is employed for fitting the model. The parameters of this model have physical interpretation. $\phi_1$ refers to the asymptotic stem radius. The starting value of the stem radius at $t = 0$ is $\phi_1 \times \exp(-\phi_2)$ with the restrictions on the parameters $0 < \phi_1 \exp(-\phi_2) < \phi_1$ the exponent of the negative of the shape parameter. This indicates the parameter $(-ln(\phi_3))$ models the shape of the function. Differentiating Equation (3), with respect to $t$, we have:

$$\frac{dy}{dt} = \phi_1(-ln(\phi_3))\phi_2\ \phi_3^t\ \exp(-\phi_2 \times \phi_3^t) = (-ln(\phi_3))\ y\ \phi_2\ \phi_3^t \tag{4}$$

$$\frac{d^2y}{dt^2} = (-ln(\phi_3))^2\ y\ \phi_2\ \phi_3^t(\phi_2\ \phi_3^t - 1). \tag{5}$$

From Equation (5), there is a point of inflection when

$$t = \frac{-ln(\phi_2)}{ln(\phi_3)}. \tag{6}$$

The relative growth rate as a function time $(t)$ is

$$\frac{1}{y}\frac{dy}{dt} = (-ln(\phi_3))\ \phi_2\ \phi_3^t. \tag{7}$$

The nonlinear mixed model corresponding to the Gompertz function in Equation (3), with the random effects for all three parameters, is:

$$y_{ij} = \phi_{1i}\ \exp(-\phi_{2i} \times \phi_{3i}^{t_{ij}}) + \varepsilon_{ij}, \tag{8}$$

$$\phi_i = \begin{bmatrix} \phi_{1i} \\ \phi_{2i} \\ \phi_{3i} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} b_{1i} \\ b_{2i} \\ b_{3i} \end{bmatrix} = \beta + b_i$$

where $b_i \sim N(0, \psi)$, and $\varepsilon_{ij} \sim N(0, \sigma^2)$, and where

$y_{ij}$ : The stem radius at time j for $i^{th}$ tree ($\mu m$)
$t_{ij}$ : The age at time $j$ for the $i^{th}$ tree (weeks).
The fixed effects, $\beta$ represent the mean value of the individual parameters, $\phi_i$, in the population of eucalyptus tree and the random effects, $b_i$, represent the deviations of the $\phi_i$, from their mean values.

## Three parameter logistic function

The second growth curve introduced is the three parameter logistic function. This model can be expressed as

$$y = \frac{\phi_1}{1 + \exp[-(t - \phi_2)/\phi_3]}. \tag{9}$$

The parameters of the logistic model have a physical interpretation. $\phi_1$ refers to the asymptotic stem radius. $\phi_2$ refers to the time at which the tree reaches half of the asymptotic stem radius. $\phi_3$ is the time elapsed for the tree to reach between half and three fourths of its asymptotic stem radius. The first derivative of Equation (9) with respect to $t$ is given by

$$\frac{dy}{dt} = \frac{\frac{\phi_1}{\phi_3}\ \exp[(t + \phi_2)/\phi_3])}{[\exp(\phi_2/\phi_3) + \exp(1/\phi_3)^t]} \tag{10}$$

$$\frac{d^2y}{dt^2} = \frac{\phi_1\ \exp[\phi_2/\phi_3](1/\phi_2)^2 \exp[(1/\phi_3)^t](\exp[\phi_2/\phi_3] - \exp[1/\phi_3]^t)}{(\exp[\phi_2/\phi_3] + \exp[1/\phi_3]^t)^2}. \tag{11}$$

From Equation (11) the inflection point of the logistic curve is at the point $t = \phi_2$ . The relative growth rate as a function of time $t$ is

$$\frac{1}{y}\frac{dy}{dt} = \frac{\frac{1}{\phi_3}\ \exp[\phi_2/\phi_3]}{(\exp[\phi_2/\phi_3] + \exp[1/\phi_3]^t)}. \tag{12}$$

The nonlinear mixed model corresponding to the logistic function in Equation (9), with the random effects for all parameters, is

$$y_{ij} = \frac{\phi_{1i}}{1 + \exp[-(t_{ij} - \phi_{2i})/\phi_{3i}]} + \varepsilon_{ij}$$

$$\phi_i = \begin{bmatrix} \phi_{1i} \\ \phi_{2i} \\ \phi_{3i} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} b_{1i} \\ b_{2i} \\ b_{3i} \end{bmatrix} = \beta + b_i$$

$$b_i \sim N(0, \psi), \quad \varepsilon_{ij} \sim N(0, \sigma^2),$$

where $y_{ij}$ is the stem radius for tree $i$ at $t_{ij}$ weeks after planting. The fixed effects, $\beta$, represent the mean value of the individual parameters, $\phi_i$, in the population of eucalyptus tree and the random effects, $b_i$, represent the deviations of the $\phi_i$, from their mean values.

**Asymptotic regression function**

The asymptotic regression (Ratkowsky, 1983; Pinheiro and Bates, 2000) is one of the convex functions that have an upper asymptote. Convex functions such as asymptotic regression have been applied in ecology to model species area relationships (Tjørve, 2003). The asymptotic regression curve is given by the formula

$$y = \phi_1 + (\phi_2 - \phi_1)(\exp[-(\exp(\phi_3)t)]) \tag{13}$$

where $\phi_1$ is the asymptote as $x$ approaches infinity. $\phi_2$ is the value of $y$ when $x$ is zero, and $\phi_3$ is the logarithm of the rate constant. Differentiating Equation (13) with respect to $t$ gives

$$\frac{dy}{dt} = (\phi_1 - \phi_2)\exp[\phi_3](\exp[(-\exp(\phi_3))t]) \tag{14}$$

$$\frac{d^2y}{dt^2} = (\phi_2 - \phi_1)\exp[2\phi_3](\exp[(-\exp(\phi_3))t]). \tag{15}$$

From Equation (15), it is clear that the asymptotic regression curve has no point of inflection. The relative growth rate as a function of time $t$ is given by

$$\frac{1}{y}\frac{dy}{dt} = \frac{(\phi_1 - \phi_2)\exp[\phi_3](\exp[(-\exp(\phi_3))\,t])}{\phi_1 + (\phi_2 - \phi_1)\exp[(-\exp(\phi_3))\,t]}. \tag{16}$$

The corresponding nonlinear mixed effects model for the radial measure $y_{ij}$ and tree $i$ at $t_{ij}$ weeks after planting is

$$y_{ij} = \phi_{1i} + (\phi_{2i} - \phi_{1i})(\exp[-(\exp(\phi_{3i})t_{ij})])$$

$$\phi_i = \begin{bmatrix} \phi_{1i} \\ \phi_{2i} \\ \phi_{3i} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} b_{1i} \\ b_{2i} \\ b_{3i} \end{bmatrix} = \beta + b_i,$$

where $b_i \sim N(0,\psi)$, and $\varepsilon_{ij} \sim N(0,\sigma^2)$.

## 3. Results and Discussions

**Exploratory Data Analysis**

The stem radius growth data is based on repeated measurements obtained for a sample of 18 trees, nine from each clone. The Loess smoothing technique by Cleveland (1979) was used to study the functional relationship between the mean radial growth and tree age for each clone.

A sharp increase in the estimated mean response profile for tree stem radial growth was observed from the time of initial measurements (i.e. at 39 weeks) up to the age of 70 weeks (Figure 1). Thereafter the increase slowed down for both clones. These curves suggest that the relationship between radial growth and tree age may be curvilinear (rather than linear). It also appears that the average profile of the GU clone was higher than that of the GC clone with the difference becoming more apparent after the age of 50 weeks. An initial examination of the individual growth suggested

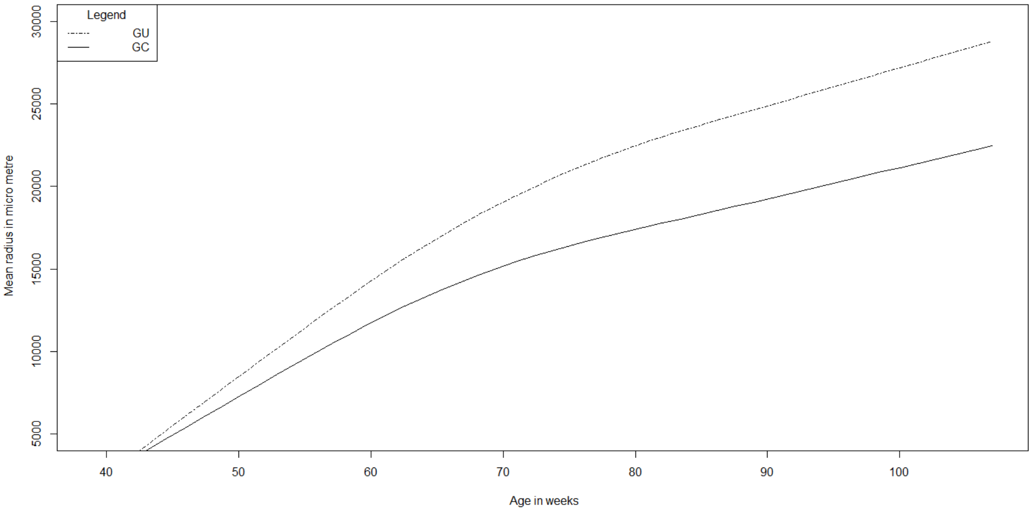


**Figure 1**: Loess smoothed curves of stem radial measure (in micro meters) against time for both E. grandis clones (GU and GC).

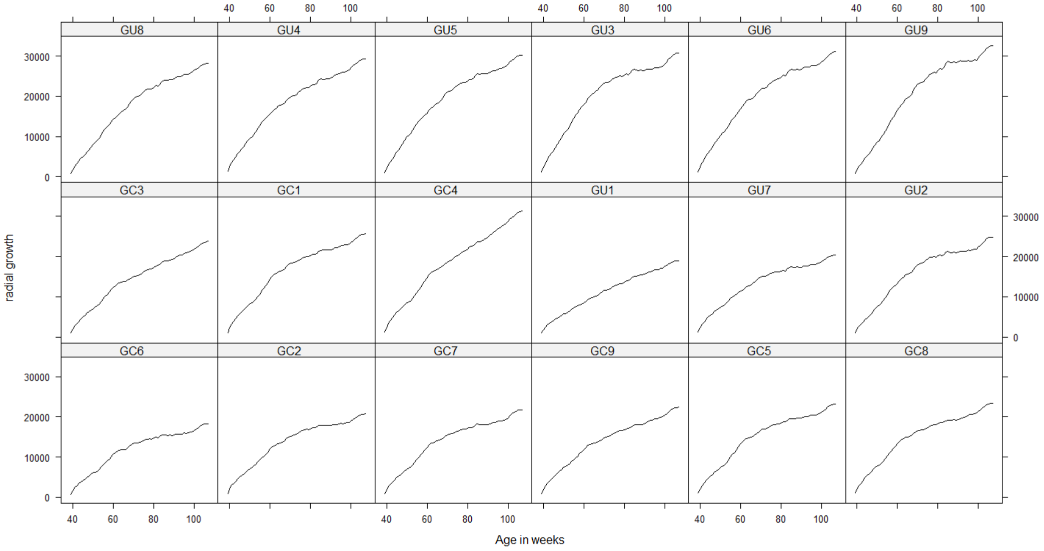


**Figure 2**: Profile plot of stem radial measure (in micro meters) against tree age for the trees sampled within each clone, GU and GC.

that the average stem radial growth for all trees during their juvenile stage (up to 107 weeks of age) might follow some three parameter nonlinear growth curves (Figure 2). This is mainly because the four parameter growth curves (for instance, four parameter logistic curves) have lower asymptote and none of our curves in Figure 2 have lower asymptote. Consequently, an attempt to fit three parameter nonlinear growth curves was made.

### Fitting and Exploring the Models

The parameters in the nonlinear mixed model are estimated by using maximum likelihood or by restricted maximum likelihood. A specified nonlinear mixed model can be fitted by maximizing an approximation to the likelihood integrated over the random effects. Several approximations to the integral are available. Some of these methods are the LME approximation method suggested by Lindstrom and Bates (1990); the method by Sheiner and Beal (1980) and Vonesh and Carter (1992) that takes first order Taylor expansion of the model function, $f$, around the expected value of the random effects; a modified Laplacian approximation (Tierney and Kadane, 1986) and Gaussian quadrature (Davidian and Gallant, 1992). Pinheiro and Bates (1995) analysed several approximations to log-likelihood of nonlinear mixed effects model and concluded that Lindstrom and Bates' (1990) approximation usually gives accurate results. This approximation is implemented in R statistical software. Model fitting was performed using the nonlinear mixed effects (NLME) package of the R statistical software.

Variance - Covariance Modelling: The first question to be addressed in the modelling process was choosing which parameters should be treated as random effects and which were purely fixed effects. One way of exploring for the presence of a random effect is the construction of confidence intervals for the parameters of the model under consideration. Because a large number of repeated measurements are made on each tree, the data have sufficient observations to have meaningful parameter estimates in the fits for individual trees. A separate fit for each tree was made and inter-tree variability is assessed using the individual confidence intervals. Approximate 95% confidence intervals for parameters of the model in Equation (8) for each tree did not overlap. This suggested that the random effects for all three parameters might be necessary.

An alternative approach was to fit different prospective models and compare nested models using the likelihood ratio tests or information criterion statistics, such as the Akaike Information Criterion (*AIC*) (Sakamoto, Ishiguro and Kitagawa, 1986). This alternative approach was considered for the parameters of the model in Equation (8). The model was fitted with each of $\phi_1$ ,$\phi_2$ and $\phi_3$ as mixed effect, called model I. The resulting AIC was 17777.77. From a reduced form of model I, with only $\phi_1$ and $\phi_2$ as mixed, an AIC of 18027.40 was obtained. This is model II. The model with $\phi_1$ and $\phi_3$ as mixed effects was also considered. The resulting AIC was 17982.21. This is model III. Finally, the model with $\phi_2$ and $\phi_3$ as mixed effects, was fitted and the resulting AIC was equal to 19041.99. This represented model IV. The AIC of the models that considered each of $\phi_1$, $\phi_2$, and $\phi_3$ at a time as fixed effects were 19041.99, 17982.21 and 18027.40, respectively. All these values are larger than the AIC of model I. This gave a clear indication that the elimination of any of these random effects has huge impact on the quality of the fit. The comparison of model I with any of the three reduced models (Models II, III and IV) using the likelihood ratio test, produced a p-value which was less than 0.0001 for all comparisons. It was concluded that the covariance structure should not be simplified by deleting any of the random effects of the model in Equation (8). This was consistent with the conclusions of the individual fits analysis discussed using the approximate confidence intervals. A model with random effects for all three parameters was, therefore, preferable. A similar analysis was made for the logistic function in Equation (9) and the asymptotic regression function in Equation (13). The results are summarized in Table 1. The resulting AIC values in Table 1 suggest that the elimination of any of the random effects has a huge impact on the quality of the fit. The comparison of a model with all three parameter as mixed with any of the reduced models using the likelihood ratio test produced a p-value, which was less than 0.0001 for all comparisons. This is true for all three nonlinear growth (Gompertz, Logistic and Asymptotic regression) curves. Moreover, it was observed that the AICs for asymptotic regression models are smaller than the AICs of either Gompertz or logistic models. For all models presented in Table 1, variation of and the covariation between the repeated measures are partitioned into between individual and within- individual components. Furthermore, it was assumed that the within individual components (residuals) are assumed to be independent with constant variance. One way of checking the assumption of constant variance is to plot the residuals against the fitted value of the model. The plots of residuals against fitted values (not shown here) for all models in Table 1 show a clear violation of the assumption of homogeneity of variance. That means these plots showed a clear pattern of variability for the within-group error. The residuals also fluctuated with tree age and the variance of residuals was not the same for the two clones.

**Table 1**: The comparison of different models using Akaike Information Criterion (AIC).

| | AIC | | |
|---|---|---|---|
| Model Parameters | Gompertz model | Logistic model | Asymptotic regression model |
| $\phi_1$, $\phi_2$, and $\phi_3$ are mixed | 17777.77 | 18478.85 | 17405.92 |
| $\phi_1$ and $\phi_2$ are mixed | 18027.40 | 18608.32 | 17751.62 |
| $\phi_1$ and $\phi_3$ are mixed | 17982.21 | 18692.69 | 17850.49 |
| $\phi_2$ and $\phi_3$ are mixed | 19041.99 | 19481.16 | 18342.28 |

The within group heterogeneity was modelled using different variance functions and different correlation structures as discussed by Pinheiro and Bates (2000). The model with the different variance of residuals for each time point appeared to be the best fit among those models for which convergence was achieved. The last aspect of the modelling process is to allow for the autocorrelation structure in the residuals. However, the attempt was not successful due to convergence problem. Therefore, models with heterogeneous variance for the within-individual variability were considered as better models for all three functions (Gompertz, Logistic and Asymptotic regression). The significance of clone for fixed effects was also assessed by comparing the models with clone effect and without clone effect, using the likelihood ratio statistics for all three functions. For the Gompertz model, clone had significant effect on the asymptote ($\phi_1$) of the model (p-value was equal to 0.015). For the remaining fixed effect parameters of the Gompertz model, no significant effect of clone is observed (Table 2). Clone had also significant effects on the asymptote ($\phi_1$) of the logistic model( p-value was equal to 0.021) and $\phi_1$ of the asymptotic regression model (p-value = 0.01).

The fitted Gompertz model, the estimated rate of growth and relative growth rate can be obtained

**Table 2**: ANOVA table for the fitted nonlinear mixed effects models.

| Model Parameters | Estimate | Standard error | Degree of freedom | $t$-value | $P$-value |
|---|---|---|---|---|---|
| 1. **Gompertz model** | | | | | |
| Asymptote-intercept ($\phi_1$) | 25938.38 | 1252.68 | 1095 | 20.71 | 0.000 |
| Asymptote-clone(slope) | -4326.60 | 1765.02 | 1095 | -2.45 | 0.015 |
| Intercept-( $\phi_2$) | 29.40 | 2.75 | 1095 | 10.68 | 0.000 |
| clone(slope) ( $\phi_2$) | -3.33 | 1.84 | 1095 | -1.81 | 0.070 |
| Intercept ( $\phi_3$) | 0.94 | 0.002 | 1095 | 421.16 | 0.000 |
| 2. **Logistic model** | | | | | |
| Asymptote-intercept( $\phi_1$) | 24263.93 | 1108.26 | 1096 | 21.89 | 0.000 |
| Asymptote-clone (slope) | -3395.49 | 1469.13 | 1096 | -2.31 | 0.021 |
| Inflection point ($\phi_2$) | 56.67 | 0.51 | 1096 | 111.91 | 0.000 |
| Scale parameter ($\phi_3$) | 10.97 | 0.38 | 1096 | 29.67 | 0.000 |
| 3. **Asymptotic regression** | | | | | |
| Asymptote-intercept( $\phi_1$) | 31023.28 | 1656.56 | 1095 | 18.73 | 0.000 |
| Asymptote-clone(slope) | -5498.73 | 2175.38 | 1095 | -2.53 | 0.010 |
| Intercept- ($\phi_2$) | -69191.06 | 5700.30 | 1095 | -12.14 | 0.000 |
| clone(slope) ($\phi_2$) | 15644.78 | 5523.57 | 1095 | 2.83 | 0.005 |
| Scale parameter ($\phi_3$) | -3.52 | 0.07 | 1095 | -54.34 | 0.000 |

from the estimated parameters in Table 2. From the model in Equation (4) the estimated rate of growth in stem radius for the two clones shows that GU clone has faster rate of growth than the GC clone. By applying Equation (6) to the parameter estimates of Table 2, the result revealed that the stem radius (for both clones) reached the inflection point about 54 weeks after the first measurement was taken. The approximate 95% confidence interval for the inflection point is between 46.55 and 60.47 weeks. Applying Equation (7), the relative growth rate for both clones can be estimated. The estimated relative growth rate is a decreasing function of time for both clones. This shows that the two clones grew in a similar manner.

According to the logistic model, clone has a significant effect on the asymptotes of the logistic curve ( p-value = 0.021 ). The parameter estimate for $\phi_2$ suggests that the average stem radius reached the inflection point at about 57 weeks (Table 2). The estimated rate of growth in stem radius for the two clones can be obtained by using Equation (10) and the parameter estimates presented in Table 2. The logistic model also shows that the rate of growth for GU is faster than that of GC clone. Applying Equation (12) and the estimates in Table 2, the logistic model also shows that the relative growth rate is a decreasing function of tree age.

For the asymptotic regression model, the dependence of each parameter on clone is modelled and tested. Clone has a significant effect on $\phi_1$ (p-value = 0.01) and $\phi_2$ (p-value = 0.005) of the model . From Equation (14) and the estimates in Table 2 , the estimated rate of growth in stem radius shows that the rate of growth for GU clone is faster than that of GC clone. Using Equation (16) and the results in Table 2 , the relative growth rate for GU and GC clones is a decreasing function of time.

For tree age between 40 weeks to 50 weeks, the relative growth of GU is higher than that of GC clone. However, as tree age increases the relative growth for the two clones get close to each other.

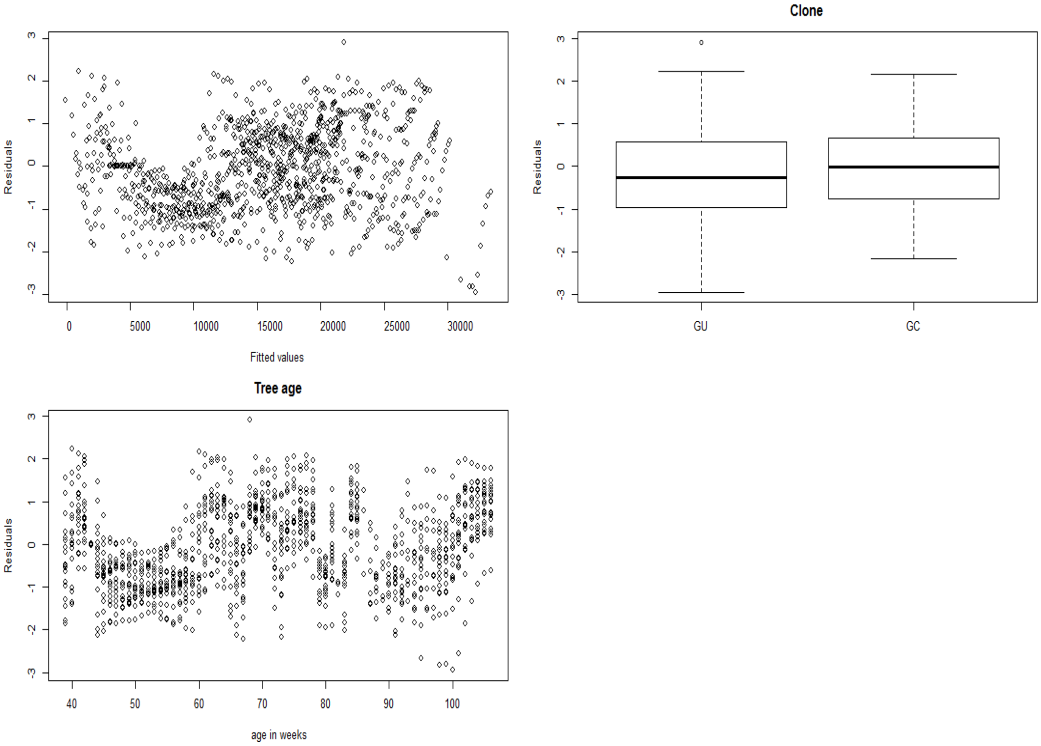
Statistical significance of the fixed effect parameters of the final nonlinear mixed model was also determined by evaluating the 95% asymptotic confidence intervals of the estimated parameters (Table 3). The null hypothesis that the parameter $H_0 : \phi_j = 0$ was rejected when the 95% asymptotic confidence interval of $\phi_j$ did not include zero. Clone had a significant negative slope for the asymptote, for all three growth models. This indicates the fact that the asymptote for the GU clone was larger than that of GC clone. This observation is in agreement with results shown in Figure 1. For the logistic and Gompertz models the 95% confidence interval for the slope of clone includes zero for parameters $\phi_2$ and $\phi_3$ (Table 3). This indicates that there was no significant difference between the two clones with regard to these parameters. With reference to the asymptotic regression model, clone had a significant positive effect on $\phi_2$. It is important to note that the parameters for different nonlinear growth curves may have different meaning. For instance the parameter $\phi_2$ is the inflection point for the logistic curve while it is the value of the response (stem radius) at time zero for the Gompertz and asymptotic regression curves. According to the asymptotic regression model, the overall average stem radius at the end of the juvenile stage of the tree is 31023.28 with the 95% confidence interval [27780.17, 34266.39] for GU clone. The estimate for GC clone is 25524.55 with the corresponding 95% confidence interval [18022.62, 33026.48].

**Table 3**: Summary of the fixed effects parameter estimate together with 95% confidence interval for the fitted three parameter nonlinear growth curves. $LCL^a$ = lower confidence limit $UCL^b$ = upper confidence limit.

| Fixed effects | $LCL^a$ | Estimated | $UCL^b$ |
|---|---|---|---|
| 1. **Gompertz model** | | | |
| Asymptote-intercept ($\phi_1$) | 23485.968 | 25938.382 | 28390.796 |
| Asymptote-clone(slope) | -7782.049 | -4326.600 | -871.152 |
| Intercept-( $\phi_2$) | 24.010 | 29.396 | 34.782 |
| clone(slope) ( $\phi_2$) | -6.931 | -3.330 | 0.271 |
| Intercept ( $\phi_3$) | 0.934 | 0.938 | 0.943 |
| 2. **Logistic model** | | | |
| Asymptote-intercept( $\phi_1$) | 22093.31 | 24263.96 | 26434.61 |
| Asymptote-clone (slope) | -6772.98 | -3395.51 | -518.04 |
| Inflection point ($\phi_2$) | 55.68 | 56.67 | 57.66 |
| Scale parameter ($\phi_3$) | 10.24 | 10.98 | 11.71 |
| 3. **Asymptotic regression** | | | |
| Asymptote-intercept( $\phi_1$) | 27780.17 | 31023.28 | 34266.39 |
| Asymptote-clone(slope) | -9757.55 | 5498.73 | -1239.91 |
| Intercept- ($\phi_2$) | -80350.75 | -69191.06 | -58031.47 |
| clone(slope) ($\phi_2$) | 4831.09 | 15644.78 | 26458.47 |
| Scale parameter ($\phi_3$) | -3.65 | -3.52 | -3.39 |

**Model Checks and Diagnosis**

The AICs for asymptotic regression models are smaller than the corresponding models fitted for the other two growth curves (Table 1). Among the three fitted models, the asymptotic regression model performs better than the other two models. The adequacy of models was assessed for all three models and we present the results of the asymptotic regression in this part of the paper. The plot of the standardized residuals against the fitted values, tree age and clone is shown for the asymptotic regression model (Figure 3). There was no clear indication for departure from the nonlinear mixed model assumptions. The model with different variance for each time point adequately fits the within-group heteroscedasticity. The assumption of normality for the within group errors was assessed using the normal probability plot of residuals (Figure 4). Close examination of the behaviour of the two plots (Zewotir and Galpin, 2004) showed that the normality assumption was plausible.


Clone

Tree age

**Figure 3**: Model validation graphs for the extended model with different variance for each time point for asymptotic regression model.

The Shapiro-Wilk normality test $(W = 0.9974, \text{p-value} = 0.07)$ also suggest there is no violation in the assumption of normality. The investigation of the marginal normality of the corresponding random effects was also made. The assumption of normality appeared reasonable for all three random effects. The p-values reported by Shapiro Wilk test are 0.4, 0.16 and 0.1 respectively for random effects associated with $\phi_1$, $\phi_2$ and $\phi_3$ of the asymptotic regression model. The adequacy of models

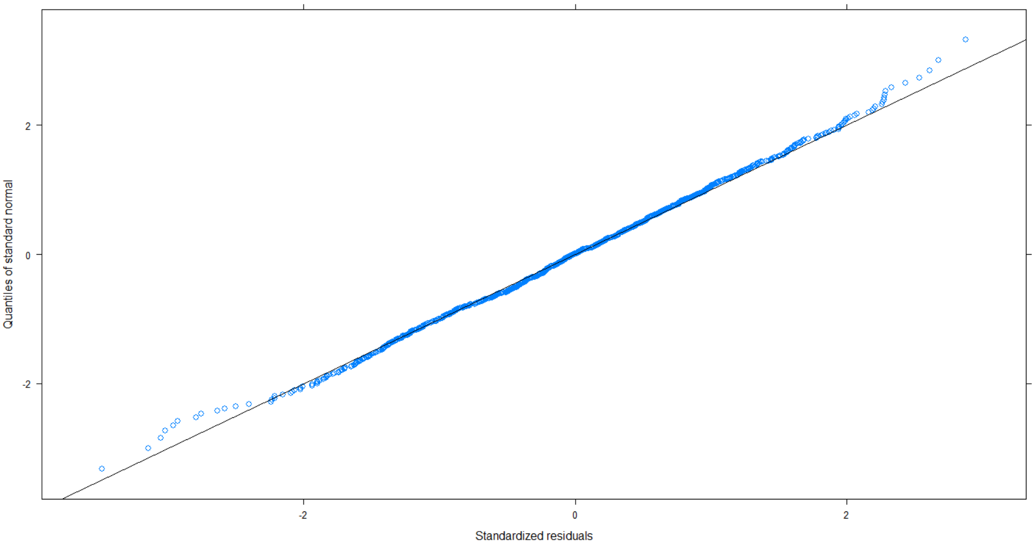


**Figure 4**: Normal probability plot of residuals for asymptotic regression model.

fitted by all three functions (Gompertz, logistic and asymptotic regression) at individual tree level was checked. The asymptotic regression model gave the best overall fit. The plot of the augmented predictions, by tree, was used as an assessment for adequacy the asymptotic regression model (Figure 5). The predicted values closely matched the observed radial growth measurements, demonstrating the acceptability of the model. The plot of the augmented predictions for the Gompertz model is presented in Figure 6. This clearly shows that the asymptotic regression model performed better than the Gompertz. Moreover, the linear regression between the observed and fitted values of the asymptotic regression model , which had an $R^2 = 0.9936$, suggested that the overall model fit was good (Figure 7).

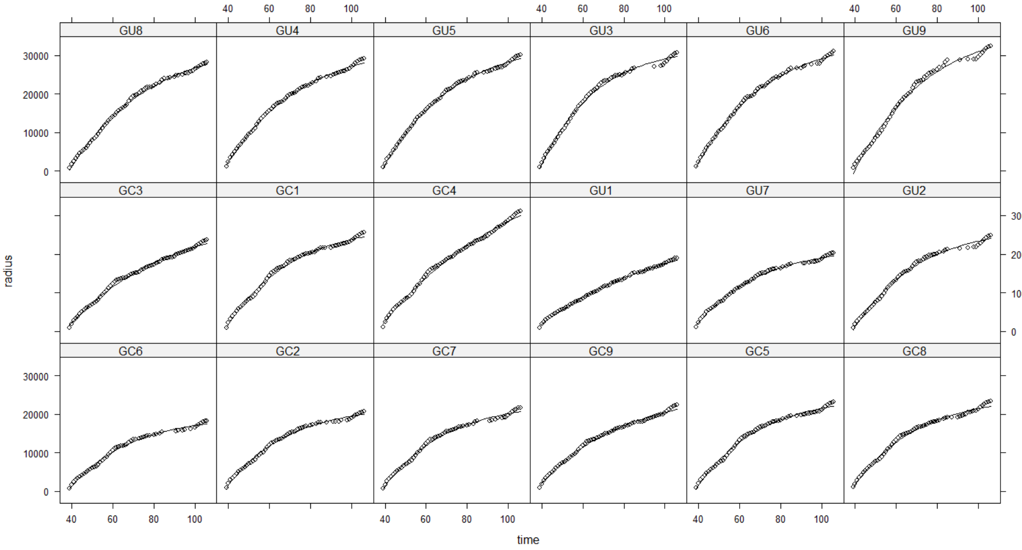


**Figure 5**: Plots of the fitted model and observed values for each tree for the asymptotic regression model.

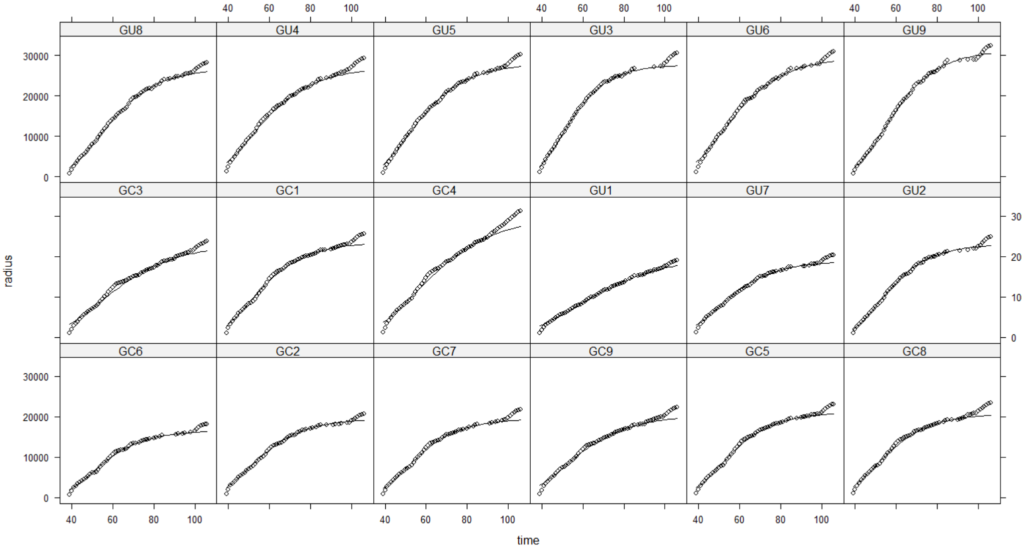


**Figure 6**: Plots of the fitted values of the Gompertz model and observed values for each tree.

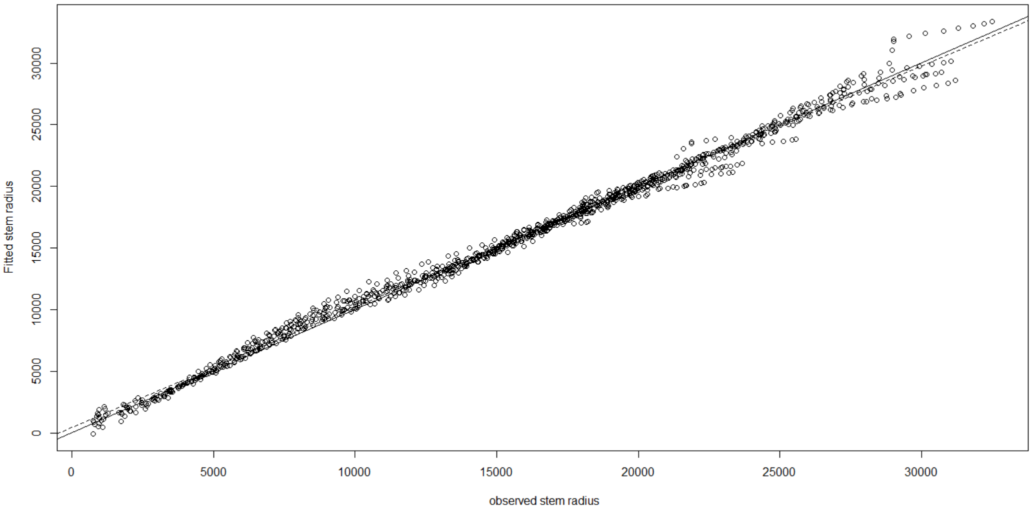


**Figure 7**: Scatter plot of the fitted versus observed stem radius. The dashed line is the estimated regression line between the observed and fitted values. (Fitted= 446.4+0.976 observed) and the solid line is the 1:1 line.

## 4. Summary and Conclusion

Based on descriptive and graphical exploratory analysis, an appropriate nonlinear tree growth function was identified. The nonlinear growth curves were fitted to individual trees under consideration and the presence of random effects for each parameter of the nonlinear growth curves were assessed graphically. Following the graphical assessment, the selections of random effects were made, by fitting different prospective models and comparing these nested models using likelihood ratio tests or information criterion statistics. These resulted in the significance of all three random effects for all growth curves. Model validation graphs showed that the within-group errors were heteroscedastic. The extended nonlinear mixed effects models with heteroscedastic, correlated within group error were fitted. The models with the heterogeneous variance that varies with tree age were found to be the best fitting models. The effect of clone on all parameters of the nonlinear growth curves was studied. Clone has a significant effect on the asymptotes of all growth curves considered. This work clearly demonstrated that the asymptotic regression curve fits better than the Gompertz and logistic curves. These analyses also confirmed that the GU clone has a larger stem radial measure than the GC clone during the entire juvenile growth stage.

Although only one clone from each hybrid cross was tested in this study, the faster growth characteristics of the GU clone points to improved genetics of this hybrid cross and to its potential to better exploit available resources, making it an economically viable hybrid cross as reported elsewhere (Galloway, 2003). In addition to being able to describe the data well, the nonlinear growth curves used in this study have meaningful parameters from a biological perspective. The asymptote represents the growth limitations within a growing season. The inflection point represents the time point at which the maximum growth rate within the season is reached. The shape-parameter models the shape of the function. These analyses also suggested that the inflection point and the shape of the Gompertz and the logistic curves were not affected by clone. This finding implies that the stem radial growth of the Eucalyptus trees was also controlled by factors other than the genetic component. The nonlinear mixed model used in this study is a convenient tool in modelling the longitudinal tree growth. It is a method that allowed the fit of nonlinear mixed models to stem radial growth data, with three parameters presented as sums of fixed and random components. It also offered an easy way to produce individual and population growth curves and other longitudinal characteristics such as growth velocity. The properties of the model and the relationship between the parameters are expected to hold for other sets of the juvenile tree data.

## Acknowledgements

The authors are grateful to Sappi Forests Research for providing the dataset on which the statistical analyses were conducted and to Dr. Valerie Grzeskowiak and Dr. Nicky Jones for several important comments and suggestions.

## References

BATES, D. M. AND WATTS, D. G. (1988). *Nonlinear Regression Analysis and Its Applications*. John Wiley and Sons: New York.

CLEVELAND, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, **74** (368), 829–836.

DAVIDIAN, M. AND GALLANT, A. R. (1992). Smooth nonparametric maximum likelihood estimation for population pharmacokinetics, with application to quinidine. *Journal of Pharmacokinetics and Biopharmaceutics*, **20** (5), 529–556.

DAVIDIAN, M. AND GILTINAN, D. M. (1995). *Nonlinear Models for Repeated Measurement Data*, volume 62. Chapman and Hall / CRC Press: New York.

DREW, D. M. (2004). Dendrometer trial phases one technical report. Technical Report EFR092T, Division of Water, Environment and Forestry Technology, CSIRO.

DREW, D. M., DOWNES, G. M., GRZESKOWIAK, V., AND NAIDOO, T. (2009). Differences in daily stem size variation and growth in two hybrid eucalypt clones. *Trees*, **23** (3), 585–595.

FITZMAURICE, G. M., LAIRD, N. M., AND WARE, J. H. (2011). *Applied Longitudinal Analysis*. John Wiley and Sons: New York.

GALLOWAY, G. (2003). Comparison of the performance of various sources of e. grandis and operational clones under operational conditions. Technical report, Sappi Forests Research. File Note ESST002aK.

LAIRD, N. M. AND WARE, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, **38** (4), 963–974.

LINDSTROM, M. J. AND BATES, D. M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics*, **46**, 673–687.

MALLET, A., MENTRÉ, F., STEIMER, J., AND LOKIEC, F. (1988). Nonparametric maximum likelihood estimation for population pharmacokinetics, with application to cyclosporine. *Journal of Pharmacokinetics and Biopharmaceutics*, **16** (3), 311–327.

MELESSE, S. F. AND ZEWOTIR, T. (2013a). The effect correlated climatic factors on the radial growth of eucalypt trees grown in coastal Zululand of South Africa. *African Journal of Agricultural Research*, **8** (14), 1233–1244. Available online at http://www.academicjournals.org/AJAR.

MELESSE, S. F. AND ZEWOTIR, T. (2013b). Path models approach to the study the effect of climatic factors and tree age on radial growth of juvenile Eucalyptus hybrid clones. *African Journal of Agricultural Research*, **8** (22), 2685–2695. Available online at http://www.academicjournals.org/AJAR.

MENG, S. X. AND HUANG, S. (2010). Incorporating correlated error structures into mixed forest growth models: prediction and inference implications. *Canadian Journal of Forest Research*, **40** (5), 977–990.

PINHEIRO, J. C. AND BATES, D. M. (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics*, **4** (1), 12–35.

PINHEIRO, J. C. AND BATES, D. M. (2000). *Mixed-effects Models in S and S-PLUS*. Springer: New York.

R CORE TEAM (2013). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna: Austria. URL: `http://www.R-project.org/`.

RATKOWSKY, D. A. (1983). *Nonlinear Regression Modelling : A Unified Approach*. Marcel Dekker: New York.

SAKAMOTO, Y., ISHIGURO, M., AND KITAGAWA, G. (1986). *Akaike Information Criterion Statistics*. D. Reidel Publishing Company: Dordrecht, The Netherlands.

SHEINER, L. B. AND BEAL, S. L. (1980). Evaluation of methods for estimating population pharmacokinetic parameters. i. michaelis-menten model: routine clinical pharmacokinetic data. *Journal of Pharmacokinetics and Biopharmaceutics*, **8** (6), 553–571.

TIERNEY, L. AND KADANE, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, **81** (393), 82–86.

TJØRVE, E. (2003). Shapes and functions of species–area curves: a review of possible models. *Journal of Biogeography*, **30** (6), 827–835.

VERBEKE, G. AND MOLENBERGHS, G. (1997). Linear Mixed Models in Practice: a SAS-oriented approach Lecture Notes in Statistics 126.

VERBEKE, G. AND MOLENBERGHS, G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer: New York.

VONESH, E. F. AND CARTER, R. L. (1992). Mixed-effects nonlinear regression for unbalanced repeated measures. *Biometrics*, **48** (1), 1–17.

VONESH, E. F. AND CHINCHILLI, V. M. (1996). *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. Marcel Dekker: New York.

WAKEFIELD, J. C., SMITH, A. F. M., RACINE-POON, A., AND GELFAND, A. E. (1994). Bayesian analysis of linear and non-linear population models by using the Gibbs sampler. *Journal of the Royal Statistical Society. Series C, Applied statistics*, **43** (1), 201–221.

WOLFINGER, R. D. (1999). Fitting nonlinear mixed models with the new NLMIXED procedure. *In Proceedings of the 24th Annual SAS Users Group International Conference (SUGI 24)*. pp. 278–284.

ZEWOTIR, T. AND GALPIN, J. S. (2004). The behaviour of normal plots under non-normality for mixed models: theory and methods. *South African Statistical Journal*, **38** (2), 115–138.

Manuscript received, 2014-03-16, revised, 2015-03-30, accepted, 2015-04-28.