Comparison of Two Nonlinear Curves to Study the Stem Radial Growth of *Eucalyptus* Tree

SILESHI F. MELESSE AND TEMESGEN ZEWOTIR
School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, South Africa
E-mail of corresponding author: Melese@ukzn.ac.za
University of KwaZulu-Natal, Private Bag X01, Scottsville, PMB (South Africa)

Melesse, S. F. and Zewotir, T. 2017. Comparison of Two Nonlinear Curves to Study the Stem Radial Growth of *Eucalyptus* Tree. *Baltic Forestry* 23(2): 438-448.

Abstract

Because of growing wood consumption, pulp and paper demands, plantations of the fast-growing tree species, managed under short rotations, have a mounting significance for the sustainability of industrial wood raw material. Consequently, the efficient utilization of fast growing plantations can have an enormous impact on productivity. The study is based on longitudinal data obtained from Sappi land holdings in coastal Zululand of eastern South Africa. During the first two years of growth, repeated measurements of stem radius were made using dendrometer attached to 18 trees. The objective of the study is to suggest a growth model for two clones of Eucalyptus tree, and compare their growth potential. Two nonlinear growth curves (Logistic and asymptotic regression with an offset) were fitted to the stem radius data. The asymptotic regression model with an offset having random effects for all its parameters appears well suited to represent the data. It was also observed that the asymptotic regression with offset fits better than the logistic regression model. All the three parameters of the logistic and asymptotic regression with an offset were considered as mixed effects. A significant difference was observed between the asymptotes of the two clones. In that GU clone grew faster than the GC clone, indicating better genetic potential for rapid growth.

Key words: asymptotic regression with an offset, dendrometer trial, Loess, logistic regression, longitudinal study, nonlinear mixed effects.

Introduction

Eucalyptus trees are woody plants that are essential and beneficial to humans. They are major sources of forest products in many countries of the world. At present, eucalyptus is used for fuel wood, transmission poles, plywood, timber, pulp building materials and other wood products. In many parts of the world, most plantation trees are established and administered for profit. Therefore, understanding the growth of eucalyptus trees has an immense economic significance. Due to increasing wood consumption and the development of pulp and paper producing industries, plantations of fast growing tree species, managed with short rotations, have a growing importance for the sustainability of industrial wood raw material. Eucalyptus trees are fast growing and can regrow from stumps after harvesting which makes multiple productions possible. As a result of this versatile nature of eucalyptus, many countries promote the planting of eucalyptus species.

At the end of the twentieth century, *eucalyptus* had become the most widely planted hardwood species in the world (Turnbull 1999). According to (FAO 2001), eucalyptus is the most commonly planted genus in Africa covering 22.4% of all planted area, followed by Pinus (20.5%), Hevea (7.1%), Acacia (4.3 %) and Tectona (2.6%). Ownership of eucalyptus plantations in Africa ranges from small farmers whose purpose is subsistence to large conglomerates for industrial wood supply. For small-holder farmers in Africa eucalyptus supports rural livelihoods in terms of fuel wood, poles and building materials. Industrial plantations are 52% publicly owned, 34% privately owned and 14% other or unspecified, while of the non-industrial plantations, 62% are publicly owned, 9% privately owned and 29% other or not specified (FAO 2001). However, South Africa is an exception. Most of the plantations in South Africa (about 72%) are owned by companies and small growers (CIFOR 2000). The cultivation of eucalyptus trees is very important in South Africa. It is the base for supplying timbers for both pulp production and solid wood applications. Therefore, eucalyptus tree cultivation is one the most important sources of revenue for the country. The cultivation of eucalyptus trees is very important in South Africa. It is the base for supplying timbers for both pulp production and solid wood applications. Therefore, eucalyptus tree cultivation is one the most important sources of revenue for the country. In South Africa, five commercially grown species make up the majority in eucalyptus plantations that is *Eucalyptus grandis*, *E. nitens*, *E. smithii*, *E macarthuri* and *E. dunnii*. Research into the site requirements of different species has been extensive and now well understood (Schönau and Gardner 1991, Darrow 1996, Herbert 2000). *E. gra*ndis is the most widely planted eucalyptus for industrial wood production. For instance, about 73.8 percent of the total area for commercial forestry is covered by *E. grandis* and its hybrid (Owen 2000).

Modelling juvenile tree growth is important in forest management to determine timber yield and long-term response of forest structure and dynamics to selective logging (Finegan et al. 1999, Chazdon et al. 2010, Herault et al. 2010). The greatest potential for improving growth rates is during juvenile development (Watt et al. 2004). Appropriate juvenile development modelling is crucial for simulation models (Gang et al. 2011). For given species, the average juvenile tree growth is conventionally expressed as a function of age. However, such average growth models assume a constant and negligible variability of the given species trees. The juvenile growth data have shown a high variability growth even with the same hybrid clone of trees. The objectives of this study were to fit two nonlinear growth curves in the framework of nonlinear mixed models to stem radius data of two eucalyptus hybrid clones (*E. grandis E. urophylla* and *E.grandis E. Camaldulensis)*. Specifically, the objective of the study is to compare the performance of the logistic growth curve with that of the asymptotic regression with an offset curve in fitting the data.

Data

A dendrometer trial, which focused on the growth of *Eucalyptus grand*is *E. urophylla* (GU) and an *E.* gradis *E. camaldulensis* (GC) hybrid clone, was established on Sappi landholdings at KwaMbonambi in the coastal Zululand area in the eastern part of South Africa. The trial was designed to run over at least nine years with separate growth monitoring phases. Each phase ended with the destructive sampling of study trees in order to measure several wood characteristics. The results presented in this study are based on the data collected only during the first of these phases of growth. The first phase ran from April 2002 when trees are 39 weeks–old until August 2003 when trees were 107 weeks old. One dendrometer was mounted on the north side of each tree at breast height (1.3 m) when trees were nine months old. Using dendrometers repeated measurements of stem radius were obtained, during this time, for a sample of 18 trees, nine from each clone. A detailed soil survey of the site conducted in August 2001 showed that the site was very uniform, thus minimizing the potential manifestation of unexpected or anomalous growth characteristics. Planting began on 16 July 2001. The site preparation prior to planting, in April 2001, included the treatment of tree stumps from the previous rotation with herbicide (to prevent coppicing) and slash from the previous harvest was burned. Each rooted cutting was established in pre-prepared soil pits between existing stumps, together with approximately 2 liters of water. The two clones were planted in alternating columns of 7 24 trees at a 3 m 2.5 m spacing. Within each column of trees for a particular clone, three plots of 12 trees (43), each with two surrounding tree rows were demarcated. The plots were established in pairs, so that in any phase of the research a GU and a GC plot could be measured simultaneously. Nine trees per plot were selected from each clone for intensive monitoring of radial growth (Drew 2004, Drew et al. 2009, Melesse and Zewotir 2013 a, Melesse and Zewotir 2013 b, Melesse and Zewotir 2015).

From the 18 sampled trees (nine per clone), longitudinal data of 1,242 weekly stem 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 fiber development in fast-growing *Eucalyptus* plantations. In addition, the study of young trees may be very important in the selection of a more productive tree species. Some studies have been made (Drew 2004, Drew et al. 2009, Melesse and Zewotir 2013 a, Melesse and Zewotir 2013 b, Melesse and Zewotir 2015) from the data extracted from the same Sappi database. With the exception of Melesse and Zewotir (2015), none of these studies consider the longitudinal aspect of the data. Moreover, this study tries to show that the asymptotic regression model with an offset curve (which is not considered by any of the above studies) can be an alternative to the most widely known logistic growth curve.

Methods

Cross sectional study may allow comparison among subpopulations 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 of parameter estimates and hence in misleading inference (Melesse and Zewotir 2015). This interdependence can be modeled 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 (Verbeke and Molenberghs 1997, 2000, Fitzmaurice et al. 2004, Meng and Huang 2010) in the analysis of the longitudinal data. Models for the analysis of such data recognize the relationship between serial observations on the same unit. Most of the work on methods of repeated measures data has focused on data that can be modeled 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. Several different nonlinear mixed effects models have been proposed (Sheiner and Beal 1980, Mallet et al.1988, Lindtorm and Bates1990, Vonesh and Carter 1992, Davidian and Gallant 1992, Wakefield et al. 1994). Pinheiro and Bates (2000) presented a general formulation for non-linear mixed models which was proposed by Lindstrom and Bates (1990). We considered here the general formulation given by Pinheiro and Bates (2000). This model can be viewed as an extension of linear mixed model of Laird and Ware (1982) in which the conditional expectation of the response given the random effects is allowed to be nonlinear function of the coefficients. It can also be regarded as an extension of nonlinear model for independent data (Bates and Watts 1988) in which random effects are integrated in the coefficients to allow them to vary by group.

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

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

where f is a nonlinear function of an individual – specific parameter vector ϕ_{ij}, the predictor vector X_{ij} and ε_{ij} 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 modeled as

$$\phi_{ij} = A_{ij}\,\beta + B_{ij}\,b_i \qquad b_i \sim N(0, \psi) \qquad (2)$$

where β is a p-dimensional vector of fixed population parameters, and b_i is a q-dimensional random effects vector associated with the i^{th} individual (not varying with j), with variance covariance matrix ψ. 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 ε_{ij} are independently distributed as $N(0, \sigma^2)$ and independent of the b_i.

We can express (1) and (2) in matrix form as

$$y_i = f(\phi_i, X_i) + e_i$$
$$\phi_i = A_i\, b + B_i\, b_i \qquad (3)$$

for i= 1, 2, M, where

$$y_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ . \\ . \\ . \\ y_{in_i} \end{bmatrix} \qquad \phi_i = \begin{bmatrix} \phi_{i1} \\ \phi_{i2} \\ . \\ . \\ . \\ \phi_{in_i} \end{bmatrix} \qquad \varepsilon_i = \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ . \\ . \\ . \\ \varepsilon_{in_i} \end{bmatrix}$$

$$f_i(\phi_i, X_i) = \begin{bmatrix} f(\phi_{i1}, X_{i1}) \\ f(\phi_{i2,} X_{i2}) \\ . \\ . \\ . \\ f(\phi_{in_i}, X_{in_i}) \end{bmatrix} \qquad X_i = \begin{bmatrix} X_{i1} \\ X_{i2} \\ . \\ . \\ . \\ X_{in_i} \end{bmatrix}$$

$$A_i = \begin{bmatrix} A_{i1} \\ A_{i2} \\ . \\ . \\ . \\ A_{in_i} \end{bmatrix} \qquad B_i = \begin{bmatrix} B_{i1} \\ B_{i2} \\ . \\ . \\ . \\ B_{in_i} \end{bmatrix}$$

The parameters in the model are estimated either by maximum likelihood, or by restricted maximum likelihood. The maximum likelihood estimation in (1) is based on the marginal density of **y**

$$p(y / \beta, \sigma^2, \psi) = \int p(y / b, \beta, \psi, \sigma^2)\, p(b)\, db \qquad (4)$$

where $p(y / \beta, \sigma^2, \psi)$ is the marginal density of y, $p(y / b, \beta, \psi, \sigma^2)$ is the conditional density of y given the random effects b and $p(b)$ is the marginal distribution of b. In general, the integral in (4) does not have a closed -form expression when the model function f is nonlinear in random effects. Different approximations have been proposed for estimating it. Some of these methods are the approximation suggested by Lindstrom and Bates (1990) LME method, the method by (Sheiner and Beal 1980, 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). Pinheriro and Bates (1995) analyzed several approximations to log-likelihood of non-linear mixed effects model and conclude that Lindstrom and Bates' approximation usually gives accurate results.

Two growth functions, three parameter logistic and asymptotic regression with an offset were selected, to replace the function *f* in model (1). This is because the parameters of these growth functions have meaningful interpretation from forestry point of view. Moreover, the nonlinear functions are more reliable for predictions with the possibility of extrapolation beyond the range of data compared to conventional polynomials. Both growth curves are considered to have three parameters and their description is given as follows.

Logistic function

This function can be expressed as

$$f(t,\phi) = \frac{\phi_1}{\left\{1 + \exp\left[-\frac{(t-\phi_2)}{\phi_3}\right]\right\}} \quad (5)$$

The parameters of this function have physical interpretation. ϕ_1 refers to the asymptotic stem radius. ϕ_2 refers is the time at which the tree reaches half of the asymptotic stem radius. ϕ_3 is the time elapsed for tree to reach between half and three fourth of its asymptotic stem radius.

The nonlinear mixed model corresponding to the logistic function (5), with the random effects for all parameters, is

$$y_{ij} = \frac{\phi_{1i}}{1 + \exp\left[-\left(t_{ij} - \phi_{2i}\right)\big/\phi_{3i}\right]} + \varepsilon_{ij} \quad (6)$$

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

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

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

Asymptotic regression with an offset

The asymptotic regression with an offset (Pinheiro and Bates 2000) is one of the convex function that have an upper asymptote. It is considered as an alternative formulation to asymptotic regression. The asymptotic regression with an offset curve is given by the formula

$$y = \phi_1 \left\{1 - \exp^{\left[-\exp(\phi_2) \times (t-\phi_3)\right]}\right\} \quad (6)$$

As in the case of logistic function ϕ_1 is the asymptote as t approaches infinity. ϕ_2 is the logarithm of the rate constant, corresponding to half-life

$t_{0.5} = \frac{\log(2)}{\exp(\phi_2)}$. The value of t at which the response variable will be zero.

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

$$y_{ij} = \phi_{1i} \left\{1 - \exp^{\left[-\exp(\phi_{2i}) \times (t-\phi_{3i})\right]}\right\} \quad (5)$$

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

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

where: y_{ij} – the stem radius at time j for i^{th} tree (μm); t_{ij} – the age at time j for the i^{th} tree (weeks).

The fixed effects, β represent the mean value of the individual parameters, ϕ_i, in the population of eucalyptus tree and the random effects, b_i, represent the deviations of the ϕ_i from their mean values.

Results

Exploratory Data Analysis

The data on the growth of stem radius is based on repeated measurements obtained, for a sample of 18 trees, nine from each clone. The Loess smoothing technique by Cleveland (1979) is used to study the functional relationship between the mean radial growth and tree age for each clone. Figure 1, shows a sharp increase in the estimated mean response profile of the stem radius from the beginning (that is, 39 weeks) up to the age of 70 weeks, and thereafter the increase slows down for both clones. An initial examination of the individual tree growth suggested that the average stem radial growth for all trees over the juvenile age (up to 107 weeks of age) follow some nonlinear growth curves (Figure 2). Consequently, two nonlinear growth curves namely the logistic and asymptotic regression with offset curves were fitted.

have meaningful parameter estimates in the individual fits. Approximate 95% confidence intervals on parameters of models in equations (6) and (7) for each tree did not overlap. This suggests that the random effects for all three parameters may be necessary.

An alternative approach was to fit nested models and compare them using the likelihood ratio tests or information criterion statistics, such as the Akaike

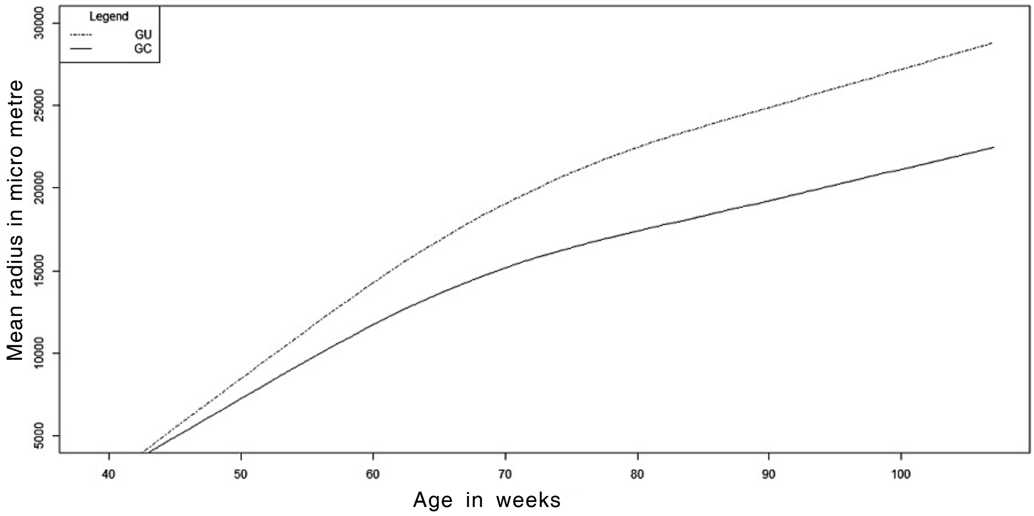

Figure 1. Loess smoothed curves of stem radial measure (in micro meters) against time for both the *E. grandis* hybrid clones (GU and GC)

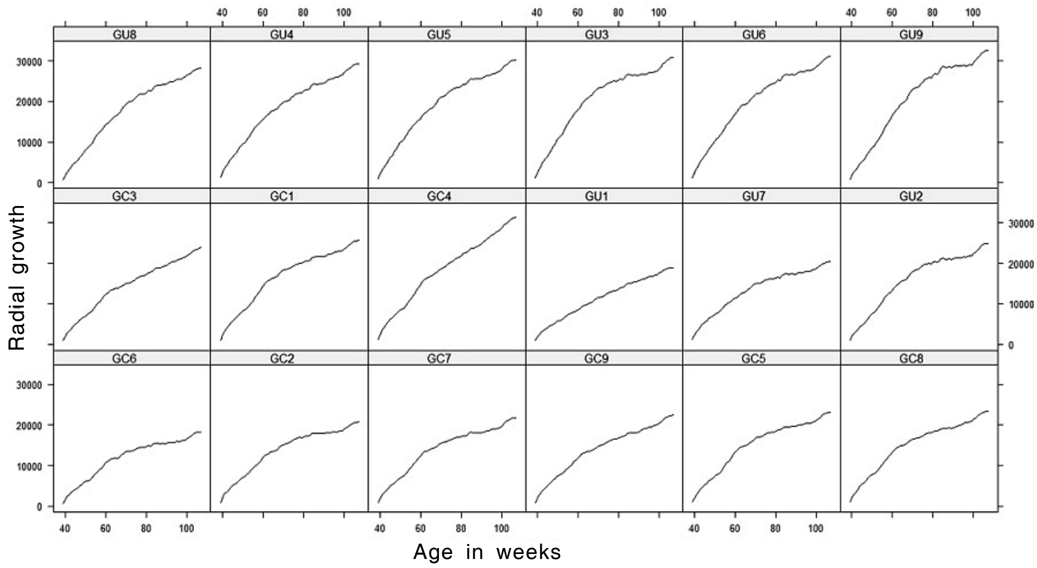

Figure 2. Profile plot of stem radial measure (in micro meters) against tree age for each of the sampled trees per clone, GU and GC

Variance – Covariance Modeling

It is necessary to consider the questions of determining which parameters in the model should have a random component and whether the variance covariance matrix of the random effects can be structured in the simpler form with fewer parameters. A separate fit for each tree was made and inter-tree variability is assessed using the individual confidence intervals. Because several repeated measurements are considered for each tree, the data have sufficient observation to

Information Criterion (AIC) (Sakamoto et al. 1986). This alternative approach is considered for the parameters of both models (6) and (8). The models were fitted with each of ϕ_1, ϕ_2, and ϕ_3 as mixed effect, called model I. For instance, for logistic function, the modeling process is described as follows. The resulting AIC is 18478.85. From a reduced form of model I, with only ϕ_1 and ϕ_2 as mixed, we get an AIC of 18608.32. This is model II. The model with ϕ_1 and ϕ_3 as mixed effects was also fitted. The resulting AIC is 18692.69. This is model III. Finally, the model with ϕ_2 and ϕ_3 as mixed effects, is considered and the resulting AIC is equal to 19481.16. This is model IV. The AIC of the models that consider each of ϕ_1, ϕ_2, and ϕ_3 at a time as fixed effects are respectively 19481.16, 18692.69 and 18608.32. All these values are larger than the AIC of model I. The comparison of model I with any of the three reduced models (Models II, III and IV) using likelihood ratio test, produced a p-value which is less 0.0001 for all comparisons. We conclude that the covariance structure should not be simplified by deleting any of the random effects of model (6). A similar analysis was made for asymptotic regression with an offset function in Equation (7). The results are summarized in Table 1. The resulting AIC values suggest that the covariance structure should not be simplified. The comparison of a model with all three parameters as mixed with any of the reduced models using the likelihood ratio test has a *p*-value, which is less than 0.0001 for all comparisons. This is true for both logistic and asymptotic regression with an offset curves. Moreover, it was observed that the AICs for logistic model are larger than the AICs of asymptotic regression model with an offset. In order to check the assumption of constant variance, the plot of residuals against the fitted values (not shown here) was used.

The plots residuals versus fitted values, clone and tree age (not shown here) for models in Table 1, show the clear violation of the constant variance assumption. That means a clear pattern of variability for the within group error is observed. The residuals also fluctuate with tree age and the variance of residuals is not the same for the two clones.

The within group heterogeneity was modeled using different variance functions and different correlation structures as presented by Pinheiro and Bates (2000). For logistic curve, the model with the different variance of residuals for each time point appears to be the best fit among those models for which convergence is achieved. For asymptotic regression with an offset curve, the model for which the variance is an exponential function of tree age appears to be best fit among those models for which convergence is achieved. The last attempt in the modelling process was to allow for the autocorrelation structure in the residuals. However, this effort was not successful due to convergence problem. The significance of clone for fixed effects was assessed by comparing models with and without clone effect, by making use of the likelihood ratio test for both curves. For the logistic curve, clone had significant effect on the asymptote (ϕ_1) of the model (*p*-value=0.021). For the rest of the fixed effects parameters of the logistic model, there is no significant effect of clone (Table 2).

Table 1. Comparison of Logistic and Asymptotic regression with an offset models using Akaike Information Criterion (AIC)

Model Parameters	AIC	
	Logistic model	Asymptotic regression model with an Offset
ϕ_1, ϕ_2 and ϕ_3 are mixed	18478.85	17348.3
ϕ_1 and ϕ_2 are mixed	18608.32	17508.73
ϕ_1 and ϕ_3 are mixed	18692.69	17748.75
ϕ_2 and ϕ_3 are mixed	19481.16	18274.66

Table 2. Fixed effects parameter estimates for the fitted nonlinear mixed effects models

Model Parameters	Estimate	Standard error	Degree of freedom	*t*-value	*p*-value
1.Logistic model					
Asymptote-intercept (ϕ_1)	24263.93	1108.26	1096	19.27	0.000
Asymptote-clone (slope)	-3395.49	1469.13	1096	-2.31	0.021
Inflection point (ϕ_2)	56.67	0.51	1096	111.91	0.000
Scale parameter (ϕ_3)	10.97	0.38	1096	29.67	0.000
2.Asymptotic regression with an offset					
Asymptote-intercept (ϕ_1)	29453.86	1528.35	1096	19.57	0.000
Asymptote-clone (slope)	-2456.40	1430.80	1096	-1.72	0.081
Natural log of rate (ϕ_2)	-3.51	0.068	1096	-51.82	0.000
Input for which the response is zero (ϕ_3)	37.37	0.276	1096	135.59	0.000

Statistical significance of the fixed effect parameters of the last nonlinear mixed model was assessed 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% confidence interval of ϕ_j does not include zero. Clone has significant negative slope for the asymptote of logistic curve. After averaging the data, a general nonlinear model (without random effects) was fitted to the summarized data. The results obtained by general nonlinear models compared with the fixed effects of nonlinear mixed models for both growth curves considered in this paper. The results are presented in Table 4. Although the two approaches are different (the mean

Table 3. Summary of the fixed effects parameter estimate together with 95% confidence interval for the two nonlinear curves. LCL^a = lower confidence limit, UCL^b = upper confidence limit

Fixed effects	LCL^a	Estimated	UCL^a
1.Logistic Model			
Asymptote-intercept (ϕ_1)	22093.31	24263.96	26434.61
Asymptote-clone (slope)	-6272.98	-3395.51	-518.04
Inflection point (ϕ_2)	55.68	56.67	57.66
Scale parameter (ϕ_3)	10.24	10.98	11.71
2.Asymptotic regression with an offset			
Horizontal Asymptote-intercept (ϕ_1)	26460.41	29453.86	32447.30
Horizontal Asymptote-slope (ϕ_1)	-5258.80	-2456.40	345.99
Natural log of rate (ϕ_2)	-3.64	-3.51	-3.38
Input for which response is zero (ϕ_3)	36.83	37.37	37.91

model does not take into account the covariance structure of the data) the results obtained are more or less similar.

Table 4. Comparison of parameter estimates for fixed effect of nonlinear mixed models (NMM) and generalized nonlinear models (GNM) without random effects

Model Parameters	Fixed effects estimates for NMM	Estimate for GNM
1.Logistic model		
Asymptote-intercept (ϕ_1)	24263.93	25560
Asymptote-clone (slope)	-3395.49	-3850
Inflection point (ϕ_2)	56.67	58.27
Scale parameter (ϕ_3)	10.97	11.40
2.Asymptotic regression with an offset		
Asymptote-intercept (ϕ_1)	29453.86	30680
Asymptote-clone (slope)	-2456.40	-4390
Natural log of rate (ϕ_2)	-3.51	-3.54
Input for which the response is zero (ϕ_3)	37.37	38.04

Model Checks and Diagnosis

The adequacy of fitted models was assessed by plotting the standardized residuals against the fitted values, tree age and clone. This was done for both logistic and asymptotic regression with an offset curves. However, the results of the asymptotic regression with an offset is presented in this part of the paper. Figure 3 shows the plot of the standardized residuals against the fitted values, tree age and clone. There is a huge improvement of the test graphs. There is no strong sign for the departure from nonlinear mixed model assumptions. The model for which the variance is an exponential function of tree age adequately fits the within-group heteroscedasticity. The normality of the within group errors was assessed using the normal probability plot of residuals (Figure 4). The examination of the normal probability plot (Zewotir and Galpin 2004) showed that there is no clear indication for the violation of the assumption of normality. The Shapiro-Wilk test of normality (W=0.99768, p-value=0.1148) also indicates the assumption of normality is reasonable. The normal probability plots of the random effects were also assessed. The assumption of normality seems reasonable for all three random effects. Moreover, the p-values produced by Shapiro –Wilks test of normality are 0.37, 0.01 and 0.55 respectively for the random effects associated with ϕ_1, ϕ_2 and ϕ_3 of the asymptotic regression with an offset curve.

The adequacy of asymptotic regression with an offset model, at individual tree level, was checked. The plot of the augmented predictions, by tree, was employed as an assessment for adequacy of the growth model (Figure 5). The predicted values meticulously matched the observed radial growth measurements confirming the suitability of the model. The corresponding plot for logistic model is presented in Figure 6. This evidently shows that the asymptotic regression with an offset model performed better than the logistic model. In addition, the linear regression

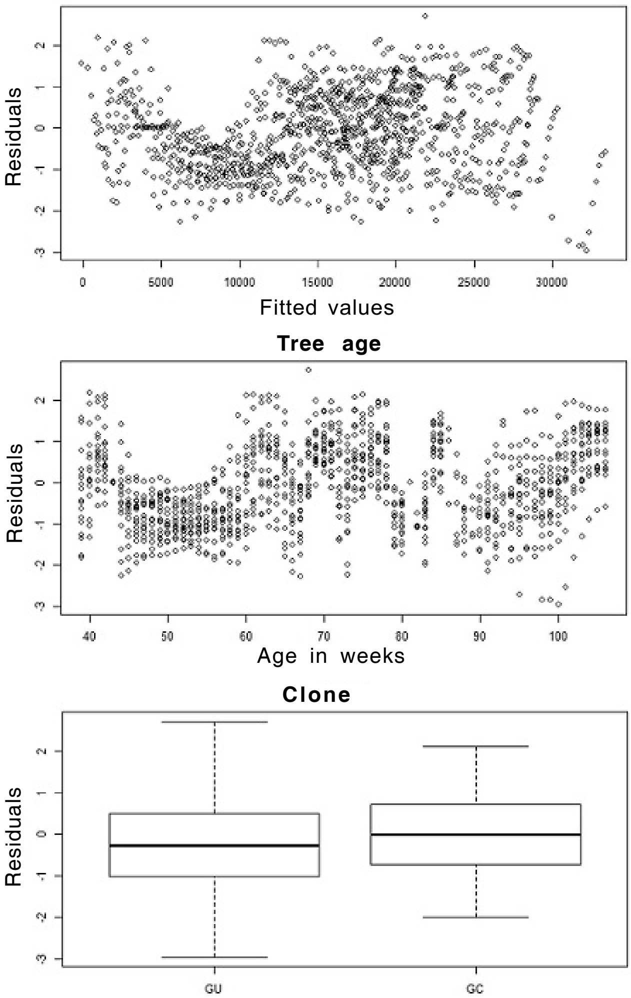

Figure 3. Model validation graphs for the extended model with variance as exponential function of time (Asymptotic regression with an offset)

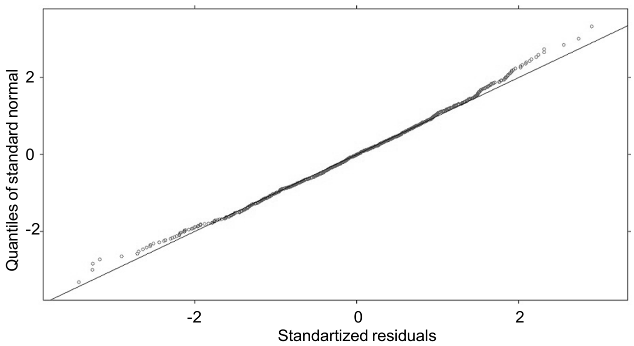

Figure 4. Normal probability plot of residuals for the asymptomtic regression with an offset

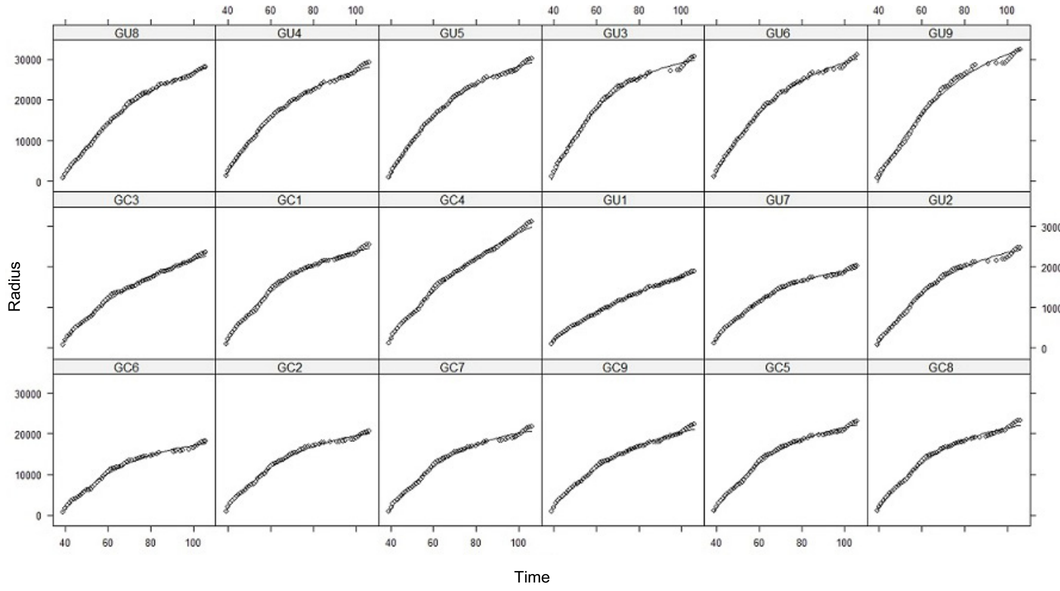

Figure 5. Plots of the fitted model and observed values for each tree using the asymptotic regression model with an offset

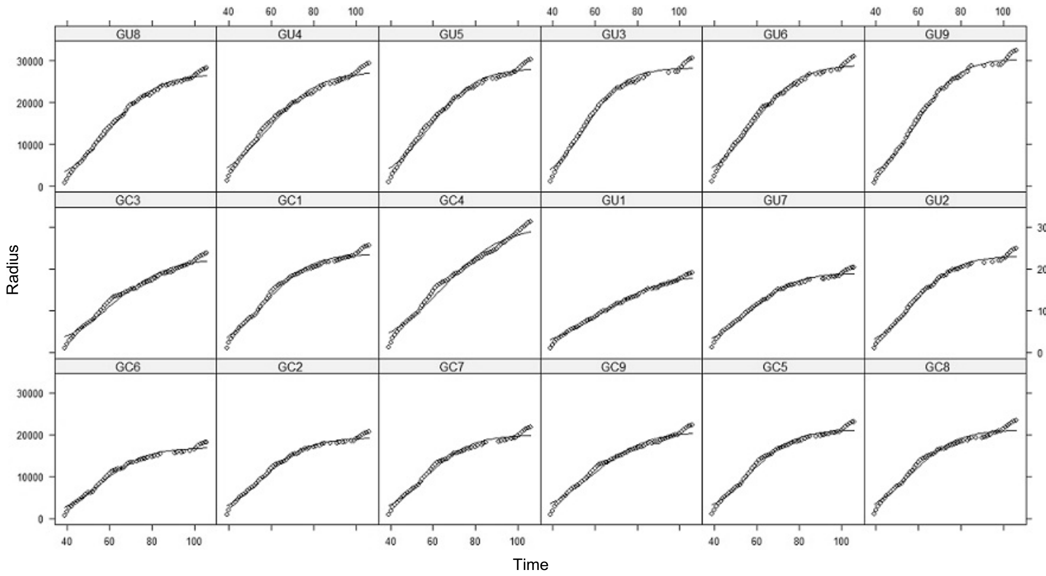

Figure 6. Plots of the fitted model and observed values for each tree using the logistic growth curve

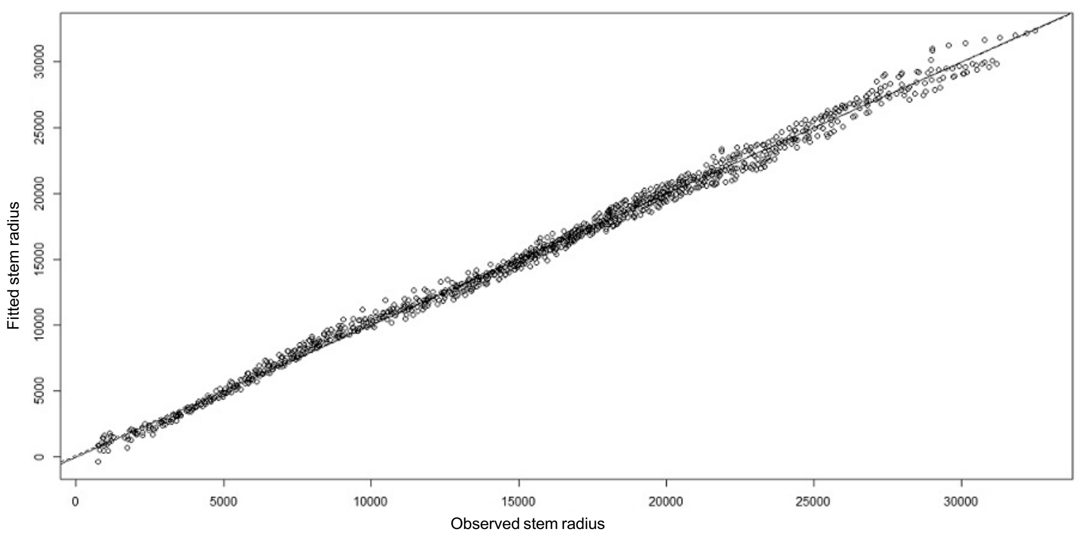

Figure 7. Scatter plot of the fitted versus observed average stem radius (The asymptotic regression line with an offset). The dashed line is the estimated regressioin line between the observed and fitted values (Fitted = 67.354+0.996 observed) and the solid line is the 1:1 line

between the observed and fitted values, indicates a strong linear relationship between the observed and fitted values. The coefficient determination was found to be $R^2 = 0.9954$, suggested that the overall model fit was adequate (Figure 7).

Discussion and Conclusion

The exploratory data analyses suggest that relationship between radial growth and age may be curvilinear (not linear). It also appears that the average profile of the GU clone is higher than that of the GC clone with the difference becoming very apparent as age of the tree increases. Based on descriptive and graphical exploratory analyses, two possible nonlinear growth functions were identified. These 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 evaluated graphically.

After the graphical evaluation, the selections of random effects were made, by fitting different nested models and comparing these models using likelihood ratio tests or information criterion statistics. These resulted in the significance of all three random effects for both growth curves. This is consistent with the conclusions of the individual fits analysis discussed using the approximate confidence intervals. This gives a clear indication that the elimination of any of these random effects has huge impact on the quality of the fit. Therefore, a model with random effects for all three parameters is considered.

Model testing graphs show that the within-group errors are heteroscedastic. The extended non-linear mixed effects models with heteroscedastic, correlated within group error were fitted. The model with the heterogeneous variance that varies with tree age was found to be the best fitting model for logistic growth curve while the model for which the variance is an exponential function of tree age is more appropriate for asymptotic regression with an offset. The effect of clone on all parameters of the logistic curve was studied. Clone has significant effect on the asymptote of the logistic curve. The parameter estimate suggests that the average stem radius of each tree reached the inflection point about 57 weeks since first measurement started. Another 11 weeks after the inflection point was reached (i.e., 68 weeks after first measurement was taken), the average stem radius reached about 75% of the growth asymptote for each experimental tree. The overall average stem radius at the end of the juvenile stage of the tree is 24263.96 and 20868.45 for GU clone and GC clones respectively. Clone has a significant negative slope (Table 2), which indicates the asymptote for GU clone is larger than that of GC.

Clone also has a marginal significance the asymptote of the asymptotic regression with an offset curve. This analysis suggests that the GU clone has larger stem radial measure than the GC clone during the entire juvenile stage (Melesse and Zewotir 2015). This is in agreement with results obtained in Figure 1. This is an indication that GU has a better genetic potential for growth than the GC clone. Clone had also marginal significance on the asymptote (ϕ_1) of asymptotic regression with an offset (p-value=0.08). It is important to note that the parameters for the two nonlinear growth curves may have different meaning. For instance, the parameter ϕ_1 is the inflection point for logistic regression while it the logarithm of the rate constant, corresponding to half-life $t_{0.5} = \frac{\log(2)}{\exp(\phi_2)}$ for asymptotic regression with an offset. According to the asymptotic regression with an offset model, the overall average stem radius at the end of the juvenile stage of the tree is 29453.86 with 95% approximate confidence interval [26460.41, 32447.30].

Even though only one clone from each hybrid cross was tested in this study, the faster growth features of the GU clone points to enhanced genetics of this hybrid cross and to its potential capacity to better exploit available resources, making it an economically feasible hybrid cross as reported elsewhere (Galloway 2003, Melesse and Zewotir 2015). Besides being able to describe the data well, the parameters of the two nonlinear growth curves used in this study are also biologically meaningful. The asymptote represents the growth limitations within a growing season for both logistic and asymptotic regression with offset curves. This parameter is affected by clone for both curves used in this study. The inflection point (for logistic curve) represents the time point when the maximum growth rate within a season is reached. The scale parameter (for logistic curve) provides the time when the growth within a season starts to level off. This analysis also suggested the inflection point and the scale of the logistic curve were not affected by clone. This finding implies that the stem radial growth of the eucalyptus trees was also controlled by factors other than genetic factors. It was also observed that the asymptotic regression with an offset model offered a better fit to the data than the logistic model. The nonlinear mixed model applied in this study is a suitable tool in modelling the longitudinal tree growth. The properties of the models and the association between the parameters are likely to hold for other sets of the juvenile tree data.

Acknowledgements

The authors are grateful to Dr. Valerie Grzekowiak 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.

Chazdon, R. L., B. Finegan, R. S. Capers, B. Salgado-Negret, F. Casanoves, V. Boukili, and Norden, N. 2010. Composition and dynamics of functional groups of trees during tropical forest succession in northeastern Costa Rica. *Biotropica* 42(1): 31-40.

CIFOR. 2000. Investing in the future: The private sector and sustainable forest management. A case study of progress towards sustainable forest management in the South African plantation forest industry. Center for international Forest Research, Bogor, Indonesia 21 pp.

Cleveland, W.S. 1979. Robust locally weighted regression and smoothing scatter plots. *Journal of the American Statistical Association* 74(368): 829-836.

Darrow, W.K. 1996. Species trials of cold–tolerant eucalypts in summer rainfall zone of South Africa: Results at six year of age. Institute for Commercial Research Bulletin Series 9/96, Pietermaritzburg, South Africa.

Drew, D. M. 2004. Dendrometer trial phase one technical report. Report No. EFR092T. Division of Water, Environment and Forestry Technology, CSIR, Pretoria, South Africa.

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.

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.

FAO (Food and Agricultural Organization of the United Nations) 2001. Global forest resources assessment 2000: main report. FAO Forestry Paper 140. FAO. Rome, Italy, 479 pp.

Finegan, B., Camacho, M. and Zamora, N. 1999. Diameter increment patterns among 106 tree species in a logged and silviculturally treated Costa Rican rain forest. *Forest Ecology and Management* 121(3): 159-176.

Fitzmaurice, G.M., Laird, N.M. and Ware, J.H. 2004. Applied longitudinal analysis. New York: John Wiley and Sons.

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.

Gang, Y., Weiqun, C., Dongxiang, L., Kang, M. Z., Letort, V., Meijun, X. and Xinyuan, H. 2011. Growth and development simulation based on functional-structural model GreenLab for poplar (Salicaceae). *African Journal of Agricultural Research* 6(13): 3071-3077.

Herault, B., Ouallet, J., Blanc, L., Wagner, F. and Baraloto, C. 2010. Growth responses of neotropical trees to logging gaps. *Journal of Applied Ecology* 47(4): 821-831.

Herbert, M. A. 2000. Site requirements and species matching: Eucalypts and Wattles. In: Own, D.L.(Ed.)., South African Forestry Handbook. South African Institute of Forestry, Pretoria, South Africa, p. 85-94.

Laird, N. M., and Ware, J. H. 1982. Random-effects models for longitudinal data. *Biometrics*: 963-974.

Lindstrom, M. J. and Bates, D. M. 1990. Nonlinear mixed effects models for repeated measures data. *Biometrics*: 673-687.

Mallet, A., Mentré, F., Steimer, J. L. 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. 2015. Fitting three parameter growth curves using a nonlinear mixed effects modelling approach. *South African Statistical Journal* 49(2): 223-240.

Melesse, S. F. and Zewotir, T. 2013a. The effect of correlated climatic factors on the radial growth of eucalypt trees grown in coastal Zululand of South Africa. *African Journal of Agricultural* 8(14): 1233-1244.

Melesse, S. F. and Zewotir, T. 2013 b. Path models-approach to the study of 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.

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.

Owen, D. L. 2000. South African Forestry Handbook. 4th Edition. Volume 1. Southern African Institute of Forestry, Pretoria, South Africa, 416 pp.

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. New York: Springer.

Sakamoto, Y., Ishiguro, M. and Kitagawa, G. 1986. Akaike information criterion statistics. Dordrecht, The Netherlands: D. Reidel.

Schönau, A.P.G. and Gardner, R.A.W. 1991. Eucalypts for colder areas in Southern Africa. In: *Schönau*, A.P.G. (Ed.), Intensive Forestry: The Role of Eucalypts. IUFRO working group P2.02- 01, Durban, South Africa, p. 467–479.

Sheiner, L.B. and Beal, S.L. 1980. Evaluation of Methods for Estimating Population Pharmacokinetic Parameters. I. Michaelis-Menton Model: Routine Clinical Pharmacokinetic Data. *Journal of Pharmacokinetic and Bio-pharmaceutics* 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.

Turnbull, J.W. 1999. Eucalypt plantations. *New Forests* 17: 37-52.

Verbeke, G. and Molenberghs, G. 1997. Linear mixed models in practice: a SAS oriented approach. Lecture notes in Statistics 126. New York: Springer-Verlag.

Verbeke, G. and Molenberghs G. 2000. Linear mixed models for longitudinal data. New York: Springer-Verlag.

Vonesh, E. F. and Carter, R. L. 1992. Mixed –Effects Nonlinear Regression for Unbalanced Repeated Measures. *Biometrics* 48: 1-18.

Wakefield, J. C., Smith, A. F. M., Racine-Poon, A and Gelfand, A.E. 1994. Bayesian Analysis of Linear and

Nonlinear Population Models using Gibbs Sampler. *Applied Statistics* 43**:** 201-221.
Watt, M. S., Kimberley, M. O., Richardson, B., Whitehead, D. and Mason, E. G. 2004. Testing a juvenile tree growth model sensitive to competition from weeds, using *Pinus radiata* at two contrasting sites in New Zealand. *Canadian Journal of Forest Research* 34(10): 1985-1992.
Zewotir, T. and Galpin, J. S. 2004. The behavior of normality under non-normality for mixed models. *South African Statistical Journal* 38(2): 115-138.

Received 21 November 2013
Accepted 24 February 2017