UNIVERSITY OF KWAZULU-NATAL

MODELING ENVIRONMENTAL FACTORS AFFECTING THE GROWTH OF EUCALYPT CLONES

2009

MORRIES CHAUKE

MODELING ENVIRONMENTAL FACTORS AFFECTING THE GROWTH OF EUCALYPT CLONES

By

MORRIES CHAUKE

Submitted in fulfilment of the academic

requirements for the degree of

MASTER OF SCIENCE

in

Statistics

in the

School of Statistics and Actuarial Science

University of Kwazulu –Natal

Pietermaritzburg

2009

Dedication

To my father, Abel Chauke (1959-2003) and my mother, N'wa-July Elizabeth Chauke

Declaration

The research work described in this thesis was carried out in the School of Statistics and Actuarial Sciences, University of KwaZulu-Natal, Pietermaritzburg, under the supervision of Prof. Temesgen Zewotir and Dr. Principal Ndlovu, in collaboration with Dr Valerie Grzeskowiak.

I, Morries Chauke, declare that this thesis is my own, unaided work. It has not been submitted in any form for any degree or diploma to any other University. Where the work of others has been used, it is duly acknowledged.

March, 2009.

_____ _____

Mr Morries Chauke Date

_____ _____

Prof. Temesgen Zewotir Date

_____ _____

Dr Principal Ndlovu Date

Acknowledgements

Foremost, I express my thanks to my supervisors Prof. Temesgen Zewotir and Dr. Principal Ndlovu for their fatherly counsel both in and out of school work, and especially for inspiring me towards completing this thesis. In addition, I extend special thanks to Dr. Valerie Grzeskowiak whose thoughtful advice often served to give a sense of direction during my MSc studies. And I am deeply grateful to Sappi for trusting me to use their data and for funding my study.

I would like to take this opportunity to thank Tinyiko Moses Ngobeni, William Baloyi (BMW) and William Shivuri, for encouraging, supporting and believing in me from the day I stepped into the University of Natal, Pietermaritzburg.

I would like to thank my friends individually. However, because the list might be too long and by fear of leaving someone out, I will simply say thank you very much to you all. With all due respect to my friends, Collen is simply my best friend.

I cannot finish without saying how grateful I am with my family whose support was a source of encouragement when I was working on my masters. Particular thanks go to my brothers (Widas and Lucky). Lastly, and most importantly, I wish to thank my parents, Abel Chauke (1959-2003) and N'wa-July Elizabeth Chauke. They have always encouraged me to do my best in all matters of life.

Abstract

Tree growth is influenced by environment and genetic factors. The same tree growing in different areas will have different growth patterns. Trees with different genetic material, e.g. pine and *Eucalyptus* trees, growing under the same environmental conditions have different growth patterns. Plantation trees in South Africa are mainly used for pulp and paper production. Growth is an important economic factor in the pulp and paper industry. Plantations with fast growth will be available for processing earlier compared to a slow growth plantation. Consequently, it is important to understand the role played by environmental factors, especially climatic factors, on tree growth.

This thesis investigated the climatic effects on the radial growth of two *Eucalyptus* clones using growth data collected daily over five years by Sappi. The general linear model and the time series models were used to assess the effects of climate on radial growth of the two clones. It was found that the two clones have similar overall growth patterns over time, but differ in growth rates. The growth pattern of the two clones appears to be characterized by substantial jumps/changes in growth rates over time. The times at which the jumps/changes in growth rate occur are referred to as the "breakpoints". The piecewise linear regression model was used to estimate when the breakpoints occur. After estimating the breakpoints, the climatic effects associated with these breakpoints were investigated.

The linear and time series modeling results indicated that the contribution of climatic factors on radial growth of *Eucalyptus* clones was small. Most of the variation in radial growth was explained by the age of the trees. Consequently, this thesis also investigated the appropriate functional relationship between radial growth and age. In particular, this nonlinear growth models were used to model the radial growth process. The investigated growth curve models were those which included the maximum radius and the age at which the radial growth rate is largest as some of the parameters. The maximum growth rate was calculated from the estimated model of each clone. The results indicated that the two clones reach the maximum growth rate at different times. In particular, the two clones reach the maximum growth rates at around 368 and 376 days, respectively. Furthermore, the maximum radius was found to be different for the two clones.

Contents

Dedication ..... **i**

Declaration ..... **ii**

Acknowledgements ..... **iii**

Abstract ..... **iv**

1. Introduction **1**

2. The data and statistical models **4**

2.1 The data **4**

2.2 Data preprocessing **6**

2.2.1 Data cleaning 7

2.2.2 Data transformation 7

2.2.3 Data reduction 8

2.2.4 Categorical data 10

2.3 Statistical models **11**

2.3.1 The importance of statistical model **11**

2.3.2 The general linear model (GLM) 12

2.3.3 Polynomial regression models 16

2.3.4 Generalized linear models 17

2.4 ARIMA models **19**

2.4.1 Arima models overview **19**

2.4.2 Transfer function models 23

2.5 Nonlinear regression model **24**

2.5.1 Piecewise regression model 29

2.5.2 Growth curve models 30

2.6 Model diagnostics **33**

3. The effects of the continuous climatic variables on radial growth **35**

3.1 Linear effects of climatic variables on radial growth **36**

3.2 Polynomial climatic and age effects on radial growth **40**

3.3 Summary of the results of the multiple linear regression approach **47**

3.4 Results of the ARIMA modeling approach **48**

3.4.1 The effects of lagged temperature on radial growth 50

3.4.2 The effects of lagged rainfall on radial growth 52
3.4.3 The effects of lagged solar radiation on radial growth 55
3.4.4 The effects of lagged relative humidity on radial growth 57
3.4.5 The effects of lagged wind speed on radial growth 60
3.4.6 Summary on ARIMA modelling approach 63
3.5 The piecewise regression model **63**
3.5.1 The effect of climatic variables on breakpoints 67
3.5.3 Summary of results obtained from fitting piecewise models 70

4. The effects of climatic variables on daily radial increment 72

4.1 Summary of the results obtained from fitting the daily radial increment model **79**

5. Fitting a growth curve model to the radial growth of the two clones 80

5.1 The Logistic growth curve model **81**

5.2 The Gompertz growth curve **82**

5.3 The von Bertalanffy growth model **84**

5.4 Comparison of the three growth curve models **85**

6. Summary and conclusions 88

References 92

Appendix A: Additional figures 96

List of Figures

Figure 2.1.1 A typical 24 hour radial variation cycle 5

Figure 2.1.2 Daily radial stem growth of the two eucalypt clones over a five year period 6

Figure 2.2.1 Outliers detected for temperature and wind speed 7

Figure 2.2.2 A scree plot for five principal components 10

Figure 2.5.1 Sigmoidal growth curve with α : inal size, r = point of inflection and M = maximum growth rate.....31

Figure 3.1.1 Constance of variance check of the errors from the linear climatic effects model 37

Figure 3.1.2 Independence check of errors from the linear climatic effects model.....38

Figure 3.1.3 Normality check of the errors from the linear climatic effects model 38

Figure 3.1.4 Independence of the errors from the linear climatic and age effects model.....39

Figure 3.2.1 Independence of errors check from the polynomial effects model 41

Figure 3.2.2 The plot of studentized residuals over time for the final polynomial model.....43

Figure 3.2.3 The plot of studentized residuals vs. predicted values for the final polynomial model 43

Figure 3.2.4 Normality plot for the final polynomial model 43

Figure 3.2.5 Cook's index plot for the final polynomial model 44

Figure 3.2.6 The quadratic effect of temperature on the radial growth of GC1 46

Figure 3.2.7 The quadratic effect of solar radiation on the radial growth of GU1 46

Figure 3.2.8 The quadratic effect of wind speed on the radial growth of GU1 47

Figure 3.4.1a Autocorrelations of the once differenced temperature series 50

Figure 3.4.1b Partial autocorrelations of the once differenced temperature series.....51

Figure 3.4.2a Autocorrelations of the undifferenced rainfall series 53

Figure 3.4.2b Partial autocorrelations of the undifferenced rainfall series.....53

Figure 3.4.3a Autocorrelations of the solar radiation series 55

Figure 3.4.3b Partial autocorrelations of the solar radiation series.....56

Figure 3.4.4a Autocorrelations of the relative humidity series.....58

Figure 3.4.4b Partial autocorrelations of the relative humidity series 58

Figure 3.4.5a Autocorrelations of the wind speed series 61

Figure 3.4.5b Partial autocorrelations of the wind speed series 61

Figure 3.5.1 The observed and fitted curves for GC1 piecewise model by growth rate phases 65

Figure 3.5.2 The observed and fitted curves for GU1 piecewise model by growth rate phases 65

Figure 3.5.3 Piecewise model with two breakpoints and neighborhoods 68

Figure 4.1.1 Index plot of residuals for the increment duration and classified climatic effects model 74

Figure 4.1.2 The plot of studentized residuals vs. predicted values for the increment duration and classified climatic effects model 75

Figure 4.1.3 Histogram of studentized residuals for increment duration and classified climatic effects model 75

Figure 4.1.4 Index plot of Cook's distance for the classified climatic and duration effects model 76

Figure 4.1.5 The effect of wind speed on GC1 radial increment 76

Figure 4.1.6 The joint effect of rainfall and relative humidity on GC1 radial increment 77

Figure 4.1.7 The joint effect of temperature and rainfall on GU1 radial increment 77

Figure 4.1.8 The joint effect of temperature and relative humidity on GU1 radial increment 78

Figure 5.1.1 Observed vs. fitted values from fitting the square root function for GC1 and GU1 80

Figure 5.2.1 Observed and fitted curves for Logistic model 82

Figure 5.2.2 Observed and fitted curves for Gompertz model 84

Figure 5.2.3 Observed and fitted curves for von Bertalanffy model 85

Figure A.1 Observed and once differenced temperature for the five year period 96

Figure A.2 Observed and once differenced relative humidity for the five year period 96

Figure A.3 Observed and once differenced solar radiation for the five year period 96

Figure A.4 Observed and once differenced wind speed for the five year period 97

Figure A.5 Observed and once differenced GC1 radial growth for the five year period 97

Figure A.6 Observed and once differenced GU1 radial growth for the five year period 97
Figure A.7 Crosscorrelation between lagged temperature and GC1 radial growth.....98
Figure A.8 Crosscorrelation between lagged temperature and GU1radial growth.....98
Figure A.10 Crosscorrelation between lagged rainfall and GU1radial growth 99
Figure A.11 Crosscorrelation between lagged solar radiation and GC1radial growth 100
Figure A.12 Crosscorrelation between lagged solar radiation and GU1radial growth.....100
Figure A.13 Crosscorrelation between lagged relative humidity and GC1radial growth.....100
Figure A.14 Crosscorrelation between lagged relative humidity and GU1radial growth 101
Figure A.15 Crosscorrelation between wind speed and GC1 radial growth.....101
Figure A.16 Crosscorrelation between wind speed and GU1 radial growth 102

List of tables

Table 2.2.1 Means and standard deviations used to standardize the dendrometer data.....8

Table 2.2.2 Correlations among five standardized climatic variables 9

Table 2.2.3 Categories of the climatic variables obtained from the fifty years data 11

Table 2.4.1 The properties of the ACF and PCAF of an ARMA(p,q) process.....21

Table 2.5.1 Properties of growth curve models 33

Table 2.5.2 Growth models and partial derivatives 33

Table 3.1.1 ANOVA results for ANCOVA model to assess the difference between clones over time 35

Table 3.1.2 Type III tests for the ANCOVA model parameters to assess the difference between the two clones over time.....35

Table 3.1.3 The variance inflation factors for the linear climatic effects model (3.1.1) 36

Table 3.1.4 ANOVA results for the linear climatic effects model (3.1.1).....36

Table 3.1.5 ANOVA results for the linear climatic effects model 39

Table 3.2.1 The variance inflation factors for the variables in the final polynomial effects model 40

Table 3.2.2 The ANOVA results for the polynomial climatic and age effects model.....41

Table 3.2.3 ANOVA results for the final polynomial climatic and age effects model.....42

Table 3.2.6 Type III tests for the ANCOVA model parameters to assess the seasonality (month nested in year) effect 45

Table 3.2.7 Parameter estimates for the final polynomial climatic and age effects model 45

Table 3.4.1 The standard deviations and the optimal orders of differencing of the climatic variable series. 49

Table 3.4.2 The standard deviations and the optimal orders of differencing of the radial growth series. 50

Table 3.4.3a Autocorrelation check for white noise from fitting the MA (3) model to the once differenced temperature series.....51

Table 3.4.3b Maximum Likelihood estimation of the MA (3) model for the temperature series 52

Table 3.4.3c Crosscorrelation check between lagged temperature and radial growth series ..52

Table 3.4.4a Autocorrelation check for white noise from fitting the MA (1) model to the undifferenced rainfall series 54

Table 3.4.4b Maximum Likelihood estimation of the MA (1) model for the undifferenced rainfall series..... 54

Table 3.4.4c Crosscorrelation check between lagged rainfall and radial growth series 55

Table 3.4.5a Autocorrelation check for white noise from fitting the MA (2) model to the solar radiation series 56

Table 3.4.5b Maximum Likelihood estimation of the MA (2) model for the solar radiation series 57

Table 3.4.5c Crosscorrelation check between lagged solar radiation and radial growth series 57

Table 3.4.6a Autocorrelation check for white noise from fitting the MA (3) model to the relative humidity series..... 59

Table 3.4.6b Maximum Likelihood estimation of the MA (3) model for the relative humidity series 59

Table 3.4.6c Crosscorrelation check between lagged relative humidity and radial growth series 60

Table 3.4.7a Autocorrelation check for white noise from fitting the MA (2) model to the wind speed series 62

Table 3.4.7b Maximum Likelihood estimation of the MA(2) model for the wind speed series 62

Table 3.3.7c Crosscorrelation check between wind speed and radial growth series 63

Table 3.5.1 ANOVA results from fitting the four-breakpoint piecewise models..... 65

Table 3.5.2 Parameter estimates for the four- breakpoint piecewise models 66

Table 3.5.3 The effect of climatic factors in five distinct piecewise growth phases 67

Table 3.5.4 Comparison of climatic conditions before and at the breakpoint neighbourhoods 69

Table 3.5.5 Duncan comparison of means at breakpoint neighborhoods 70

Table 4.1.1 The ANOVA results for the classified effects model without increment duration 72

Table 4.1.2 ANOVA results for the increment duration and classified climatic effects model 74

Table 4.1.3 Type III tests for the classified climatic effects and duration effects model 78

Table 5.1.1 The ANOVA results for the logistic growth curve model 81

Table 5.1.2 Parameter estimates for the Logistic growth curve 82

Table 5.2.1 ANOVA results for the Gompertz growth curve model 83

Table 5.2.2 Parameter estimates for Gompertz growth curve 83

Table 5.3.1 ANOVA results for the von Bertalannfy growth curve model 84

Table 5.3.2 Parameter estimates for von Bertalanffy growth curve 85

Table 5.4.1 Parameter summary of growth curve models 86

Table 5.4.2 Estimated number of days required to achieve certain percentage of the maximum radius 87

1. Introduction

Woody plants are important and beneficial to mankind. In particular, eucalypts are a major source of forest products. The name *Eucalyptus* was given to an Australian tree in the late eighteenth century by the French botanist, Charles Louis L'Heritier De Brutelle (Turnbull, 1999). At that time, the appreciation of the potential of eucalypts to become a major source of forest products was minimal. Now at the end of the 21st century, eucalypts have become the most widely planted hardwood species in the world (Turnbull, 1999). Most plantation trees are established and managed for profit. Growth is an important economic factor. Plantations with fast growth will be available for processing earlier compared to slow growth plantations. South Africa is one of the best examples of the economic benefits that can flow from such plantations when the plantations are well managed (Turnbull, 1999).

Sappi is one of South Africa's leading suppliers of coated fine paper and chemical cellulose. The company conducts its business through two business units: Sappi Fine Paper and Sappi Forest products. Sappi Fine Paper South Africa produces and markets a range of coated, uncoated and specialty papers, as well as creped tissue and fibreboard in South Africa. Sappi Forest Products is a pulp and paper producer in Africa with a production capacity of 830,000 metric tons of paper, 600,000 metric tons of chemical cellulose and 1,090,000 metric tons of paper pulp per annum. Sappi Forest Products is also a timber grower and manages approximately 553,000 hectares of forestland, of which approximately 409,000 hectares is planted with primarily pine and eucalypts (Reuters, 2008).

Sappi Forests together with Usuntu Forests supplies all of the Sappi Forest Products and Sappi Fine Paper South Africa's domestic pulpwood requirements of approximately six million metric tons per annum. Together they manage about 553,000 hectares of land situated in Mpumalanga (44%), Kwazulu-Natal (44%) and Swaziland (12%). The company competes with Mondi Paper Company Limited, Borregaad ChemCell Atisholz, Tembec Inc., Buckeye Technologies Inc. and Rayonier Inc. More information about Sappi and Mondi forests can be found in Meadows (1999) among others.

If fast-growing eucalypts are to assemble their full potential, they must be well managed. Sufficient management requires good understanding of factors affecting tree growth (Pallardy and Kozlowski, 2007). Tree growth is a product of the interaction of genetic and environmental factors (Kozlowski, 1962). Trees with different genetic material will grow at

different growth rates. For example, eucalypts take about four months from planting before the seedlings are ready for transfer to the forest while pines take about seven months (Meadows, 1999). Furthermore, genetic effects on growth differ extensively between widely spread populations of trees. In particular, the closer two populations are to one another, the more obscure differences between them will become. On the other hand the growth differences between populations close to one another are slight (Pallard and Kozlowski, 2007). Extensive literature on genetic factors affecting the growth of trees can be found in Pallardy and Kozlowski (1997) among others.

Environmental factors such as temperature, sunlight, precipitation, soil moisture, soil nutrients, and length growing season can all affect tree growth. Spacing between trees tends to rob trees of nutrients and room to grow. Some trees are more sensitive and can be killed even by moderate levels of competition. It is therefore important to take spacing into account to avoid competition among trees. Less than optimal growing conditions can slow the growth rate of a tree or change its shape. If suitable climatic conditions for optimal growth are known, then trees will planted under such conditions. In Australia, a study by Downes, Beadle and Worledge (1999) used multiple linear regression methods to investigate the linear relationship between radial stem growth and climatic variations over a 12-month period. Their study indicated that 43% of the total variation in radial stem growth was explained by linear climatic variations. Further investigation to understand the effect of climate on stem radial growth for a longer period may obtain different results in terms of percentage of total variation in radial growth explained by the climatic effects. Moreover, there was no similar study conducted in South African *Eucalyptus* trees.

Consequently, the objectives of this thesis were: 1) to investigate linear and nonlinear relationships between radial growth of the two *Eucalyptus* clones (see Chapter 2) and the climatic variables; 2) to determine whether or not the growth rates of the two clones are different; and 3) if the growth rates are different, to investigate whether or not these differences are due to genetic and/or climatic factors; using the growth and climatic data collected over a 5-year period.

There are six chapters in this thesis. This chapter gave some background about the origin and importance of eucalypts in South Africa. This chapter also briefly reviewed related studies and gave the objectives of the thesis. Chapter 2 contains a review of literature on relevant statistical models, and a description of the data used in this thesis to achieve the objectives.

This chapter also provides some of the data preprocessing techniques that were used with the aim of obtaining good quality results. The results on the effects of the continuous climatic variables on radial growth are presented in Chapter 3. Chapter 4 contains the results on the effects of climatic variables on daily radial increment. Chapter 5 presents the results on growth curve modeling. Chapter 6 provides the summary and the conclusions of the study.

2. The data and statistical models

2.1 The data

The following information about the data was extracted from the dendrometer trial project database. The *Eucalyptus* fibre research trial, known as the "dendrometer trial", started at the site of Sappi Kwambonambi in costal Zulu-land in mid-2001. This trial was a joint effort of Sappi and CSIR, to conduct a range of experiments to understand the process of fibre development in plantation eucalypts. Two Sappi hybrids (GC1 and GU1) were established at the site. The fundamental objective of the research was to link the short-term variations in environmental and tree physiological characteristics. This research was designed to run over at least seven years in separate phases.

This site was chosen for its proximity to the nursery, security and accessibility; homogeneity of the soil; and availability. Furthermore, the site represents a typical site in the region, and a large amount of growth data for similar trees grown over a full rotation on the site was available. Thus, the possibility of the site exhibiting unexpected characteristics was minimal. A detailed soil survey was conducted prior to this experiment, and the soils at the site were found to be highly uniform.

The radial stem growth and weather data were key in this research, and thus an effort was made to obtaining high quality data. Daily stem radial growth was determined from hourly dendrometer measurements in the eighteen sampled trees from the two clones. Figure 2.1.1 shows a typical growth cycle (shrinkage, recovery, increment) phases. The shrinkage phase was defined as the period during which the tree decreased in radius. The recovery phase was defined as the period during which the radius increased until it reached its previous maximum. Increment was defined as the period during which the stem increases in size from the previous temporary maximum until the onset of the following shrinkage. Further literature on shrinkage, recovery and increment phases can be found in Downes, *et al*. (1999). The average radial growth was obtained by cumulating and averaging the daily radial increments of nine trees of each clone.

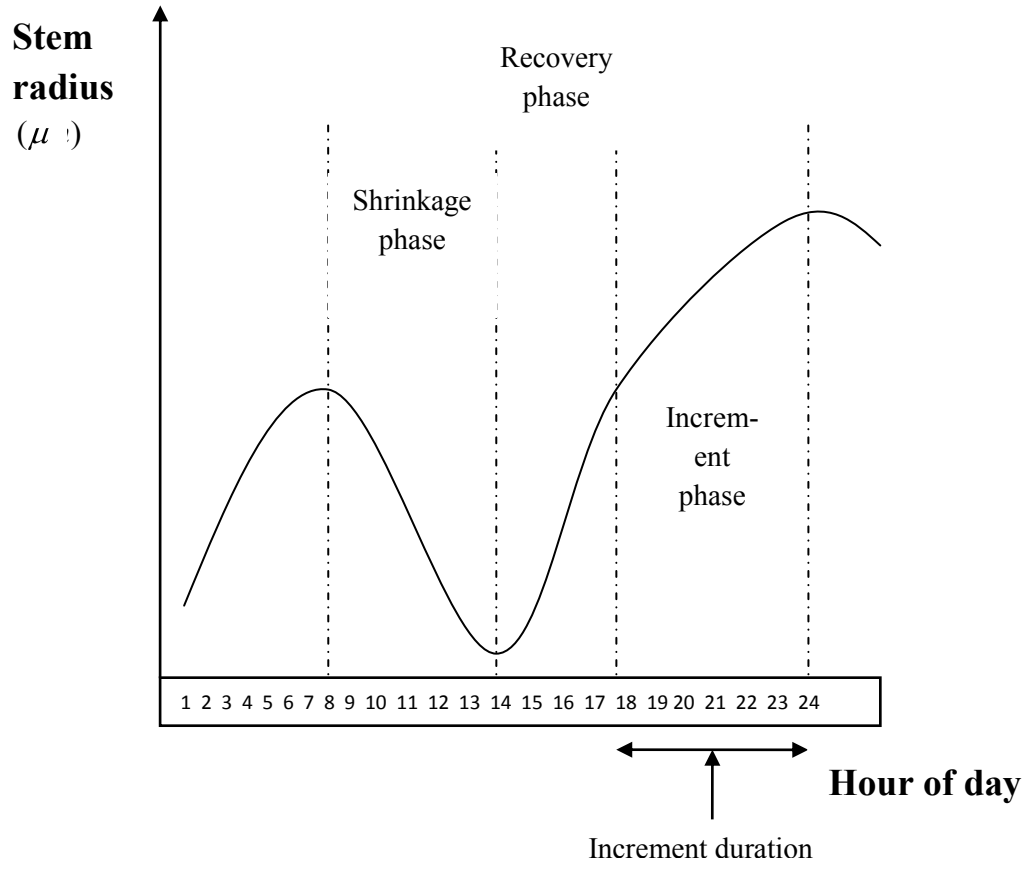

Figure 2.1.1 A typical 24 hour radial variation cycle

This radial growth data was collected during the five year period (April 2002-March 2007). This period was classified as Phases I, II, III and IV. Growth measurements for experimental Phase I started in April 2002 and were completed in August 2003. Experimental Phase II growth measurements began in September 2003 and were completed in July 2004. Experimental Phase III growth measurements ran through September 2004 to October 2005. The measurements for experimental growth Phase IV began in December 2005 and were completed in January 2007. Each experimental growth phase ended with the cutting of trees and sampling of new eighteen study trees for the next experimental phase: nine trees from GC1 and nine trees from GU1. Measurements from one experimental Phase to another were done on different trees of the same clones. Figure 2.1.2 displays the measured radial stem growth of the two clones over a five year period.

In addition to radial growth, an automatic weather station was installed at the nursery, a distance approximately 200m from the trial to record hourly temperature using a temperature sensor ($^{\circ}C$), relative humidity using an electronic relative humidity sensor (%), solar

radiation using a radiation sensor (mJ/hr), rainfall using a rain gauge (mm), and wind speed using a wind speed sensor (m/s).

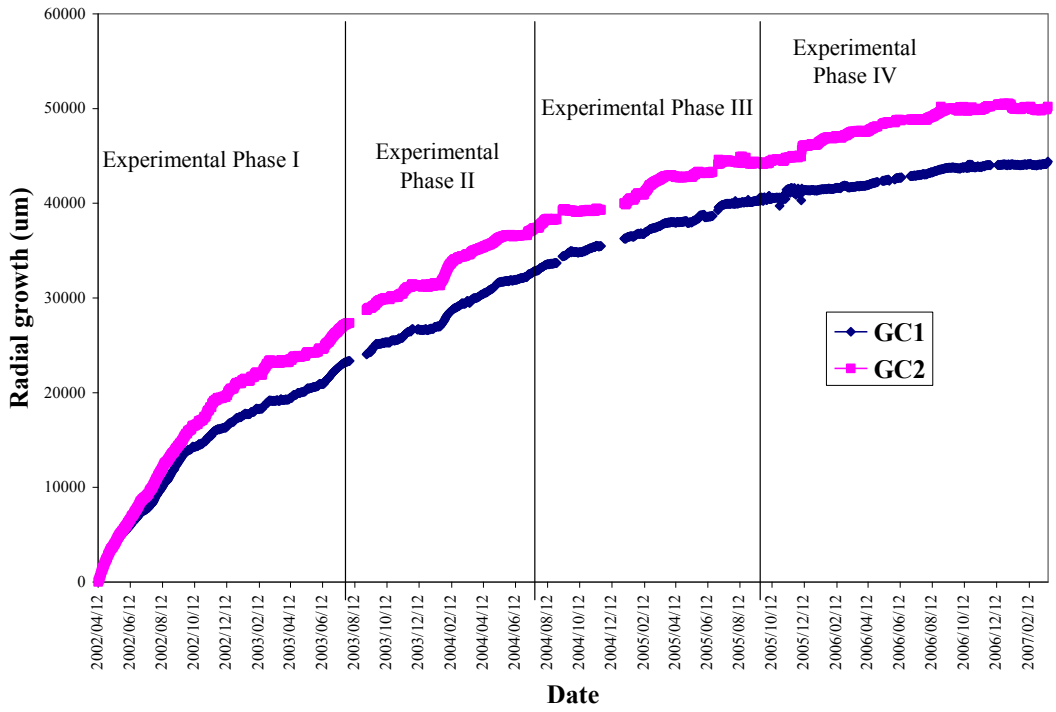

Figure 2.1.2 Daily radial stem growth of the two eucalypt clones over a five year period

2.2 Data preprocessing

Large scale data sets are susceptible to errors and inconsistencies. So if these unchecked data are used in raw form, they may not be ideal for good quality results. Data preprocessing need to be done prior to any analysis to improve the quality of such data sets. Data preprocessing techniques include data cleaning; data transformation; and data reduction among others. Data cleaning is ideal to reduce or correct inconsistencies in the data. Data transformation improves the accuracy and efficiency of different predictive techniques. Data reduction may involve combining highly correlated variables to reduce redundant features. The overall quality of the results may be improved substantially when these data preprocessing techniques are applied prior to analysis.

2.2.1 Data cleaning

Some observations may differ substantially from the majority of the observations in the sample due to reasons such as sensor failure, data transmission or improper data entry (Famili, *et al*. 1997). Such observations are referred to as outliers. Outliers may be influential. That is, they may alter the important results though their presence or absence. Data cleaning involves identifying such observations and taking remedial actions. The observed values of temperature and wind speed displayed in Figures 2.2.1 indicate that some observations were inconsistent with rest of the data. If observations are known to be erroneous, they can be dropped (Goharian and Grossman, 2003). A considerable time and effort was applied to check the correctness of the data. Some obvious errors were identified. After a number of discussions with Sappi and the dendrometer trial database administrator, erroneous data points were corrected and a true reflection of the situation provided to us. Further thorough exploration of the revised data was done. Except few outliers, no gross error was identified. Also, none of the identified outliers were erroneous, and thus they were retained.

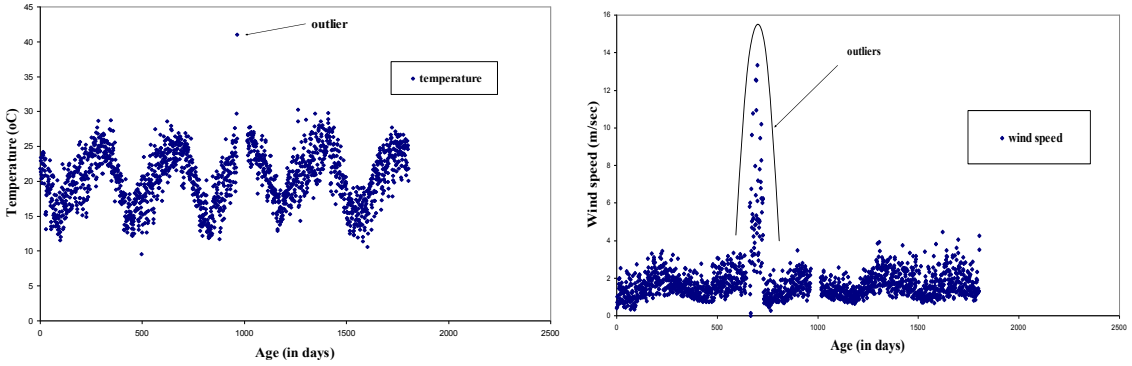

Figure 2.2.1 Outliers detected for temperature and wind speed

2.2.2 Data transformation

It is necessary to consider transforming the variables prior to analysis, because variables with large variances tend to have a larger effect than those with smaller variance (SAS, 2004). Standardization is used to reduce variability or dispersion in the data set. This makes the variable measurements have a mean zero and a variance of one, and thereby giving every variable data equal weight in the model (Nsofor, 2006). Standardization is done by

subtracting the mean from the observed variable values and dividing the difference by the standard deviation of the observed variable values. For example, the values of a variable x are standardized as follows:

$$x_i^* = (x_i - \bar{x})/s, i = 1,2\ldots,n$$

where x_i^* is the i^{th} standardized value, $\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$; and $s = \sqrt{\frac{1}{n-1}\sum(x_i - \bar{x})^2}$. Analyses are then carried out using standardized values instead of observed values. The use of standardized independent variables in polynomial regression also helps in solving the multicollinearity problem (Ravishanker and Dey, 2002). The means and standard deviations used to standardize the climatic data are displayed in Table 2.2.1.

Table 2.2.1 Means and standard deviations used to standardize the dendrometer data

Measure	**Temperature** ($^{\circ}C$)	**Rainfall (mm)**	**Solar radiation (m/joules)**	**Relative humidity (%)**	**Wind speed (m/sec)**
Mean	20.45	0.06	0.45	80.20	1.77
Standard deviation	3.93	0.22	0.28	10.37	1.09

2.2.3 Data reduction

When data are collected from a large number of variables, it is possible that some variables will be correlated. The high correlation between two variables could be a sign that they measure the same construct, and the multicollinearity problem arises when linear regression on highly correlated variables is performed. Principal component analysis (PCA) is a variable reduction procedure useful when data are obtained on many highly correlated variables. PCA attempts to reduce the many highly correlated variables to a smaller number of uncorrelated artificial variables (called principal components) that will account for most of the total variance in the observed values of the variables without any significant loss of information. A principal component is a linear combination of optimally-weighted observed variables. The i^{th} principal component created in a principal component analysis is given as

$$C_i = \gamma_{i1}(X_1) + \gamma_{i2}(X_2) + \cdots + \gamma_{ip}(X_p)$$

where b_{ij} =the weight for the variable X_j, $i, j = , 2, \ldots, p$.

Linear regression on the principal components of the independent variables does not suffer from the multicollinearity problem. Further discussions of principal component analysis are in Johnson and Wichern (1988); Jackson (1999); and Ravishanker and Dey (2002) among others.

Table 2.2.2 Correlations among five standardized climatic variables

	temperature	**rainfall**	**solar radiation**	**relative humidity**	**wind speed**
temperature	1.000	-0.029	0.298	-0.416	0.314
rainfall	-0.029	1.000	-0.143	0.243	0.114
solar radiation	0.298	-0.143	1.000	-0.034	0.174
relative humidity	-0.416	0.243	-0.034	1.000	-0.200
wind speed	0.314	0.114	0.174	-0.200	1.000

The correlations among the climatic variables presented in Table 2.2.2 are relatively small, indicating weak linear relationships among climatic variables. The PCA is therefore not expected to collapse the climatic variables into a smaller number of variables. We used the scree plot (Eigenvalue plot) to confirm if climatic variables do not have strong linear relationship. A scree plot is a line segment plot that shows the fraction of total variance in the data as explained by each principal component. The principal components are ordered by decreasing order of contribution to total variance. This plot when read left to right can show clear separation in fraction of total variance where most important components cease and the least important variables begin. The principal components before the break in the scree plot are retained. The scree plot displayed in Figure 2.2.2 indicated that there was one break between the eigenvalues. That is, the first principal component may be used to reduce the variables. The proportion of the total variation explained by the first principal component analysis is about 39%. This proportion is very small to consider the first principal component as a proxy of the five variables. If the five variables were strongly linearly related, then the principal components before the scree plot break would have explained a remarkable portion

of the total variation. The PCA result displayed in Figure 2.2.2 favours the weak linear relationship among the climatic variables.

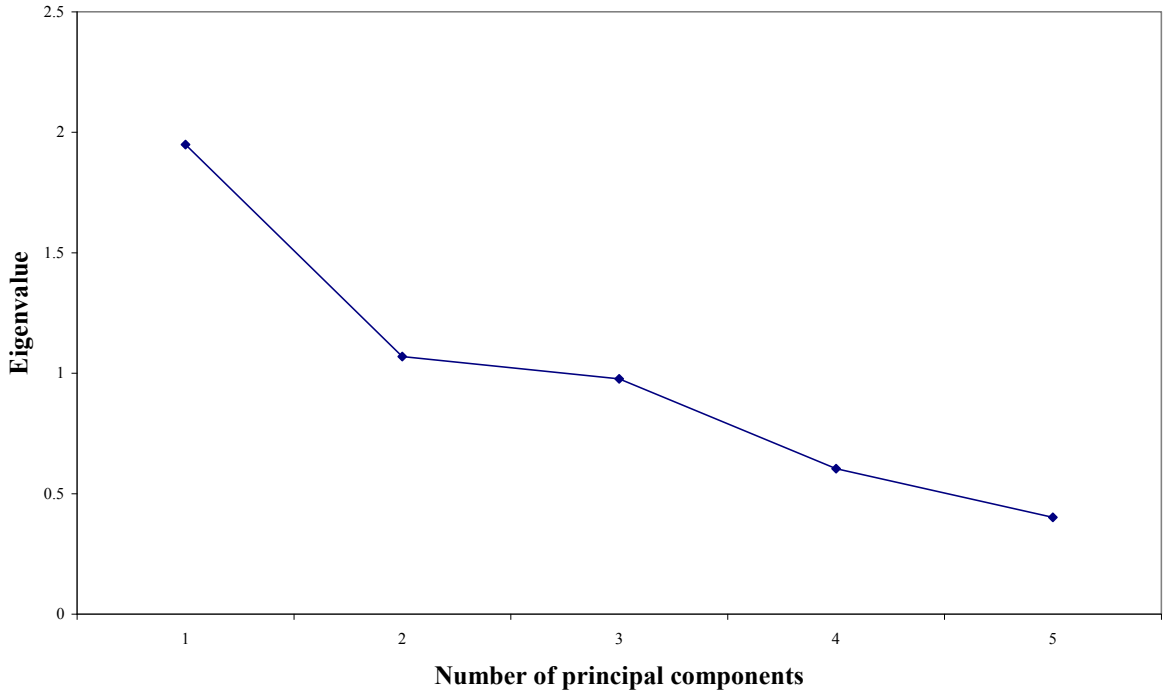

Figure 2.2.2 A scree plot for five principal components

2.2.4 Categorical data

The analysis of variance or covariance of radial growth with the climatic variables as the factors requires categorized climatic data. The climatic data were classified into three different levels (low, normal and high), using the fifty years climatic data from the same area where the Dendrometer trial took place. The categories of the fifty years climatic data are displayed in Table 2.2.3. In particular, the first (Q_1) and third (Q_3) quartiles, of the fifty years data, were used to classify the data for each climatic variable. If variable data point was less or equal to Q_1 , then variable factor level is *low*; if variable data point was greater than Q_3 , then the variable factor level is *high*; Otherwise variable factor level is *normal*.

Table 2.2.3 Categories of the climatic variables obtained from the fifty years data

Climatic variable	Category/level		
	Low **Less or equal to** Q_1	**Normal** Q_1 to Q_3	**High** **Greater than** Q_3
Temperature ($^{\circ}C$**)**	Less or equal to 17.29	17.29 to 23.26	Greater than 23.26
Rainfall (mm)	Less or equal to 0.0	0.0 to 0.017	Greater than 0.017
Solar radiation (Mj/hr)	Less or equal to 0.26	0.26 to 0.6	Greater than 0.60
Relative humidity (%)	Less or equal to 73.27	73.27 to 86.54	Greater than 86.54
Wind speed (m/s)	Less or equal to 1.17	1.17 to 2.05	Greater than 2.05

2.3 Statistical models

2.3.1 The importance of statistical model

Statistical models have been widely used to explain natural phenomenon in all disciplines, including business, economics, engineering, the physical sciences, and the social, health and biological sciences. The statistical model usually has a simple solution that can be evaluated using probability distributions (Freund and Wilson, 1998). A statistic model separates the systematic features of the data from random variation. The systematic features are represented by a regression function involving parameters which can be related to the structure of the sample and to important variables measured on each experimental or observational unit. The random variation unrelated to important variables is represented by a probability distribution depending on a number of parameters. Interpretation of the data and conclusions about the population can then be based on the regression function, with no attempt to interpret random variation (Aitkin, *et al*. 1989). Successful use of these models requires understanding of the theoretical underpinnings of the phenomena, statistical characteristics of the model, and the practical problems that may be encountered when using these models in real life situations (Freund and Wilson, 1998). The following subsections give an overview of the statistical models that were used in this study.

2.3.2 The general linear model (GLM)

Suppose that there are n data points, $y_1, \ldots, y_n$, that are to be explained using n values for each of the k explanatory variables $X_1, X_2, \ldots, X_k$. For $i = , 2, \ldots, n$ and $j = , 2, \ldots, k$, let X_{ij} be the i^{th} value of variable X_j which may be a continuous variable or categorical variable. The standard linear model for the data is given by

$$y_i = \beta_{\_} + \beta_{\_} X_{i1} + \beta_{\_} X_{i2} + \ldots + \beta_{\_} X_{ik} + \varepsilon_{\_}, \; i = , 2, ..., n.$$

where $\beta_0, \ldots, \beta$ are unknown fixed parameters to be estimated; and $\varepsilon_1, \ldots, \varepsilon_{\ }$ are unknown independent and identically distributed normal random variables with mean 0 and variance σ . In matrix notation, the preceding model can be written as

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon \end{bmatrix}$$

For convenience, this model can be equivalently expressed as

$$\mathbf{y} = \zeta\boldsymbol{\beta} + \qquad (2.3.1)$$

where **y** denotes the $n\times$ vector of the observed $y_i's$; $\mathbf{X}_{n\times k+)}$ is a matrix of X_{ij}'s; $\boldsymbol{\beta}_{(k+)\times}$ is the vector of unknown fixed parameters; and $\boldsymbol{\varepsilon}_{n\times}$ is the unknown vector of independent and identically distributed normal random errors. Additionally $E(\boldsymbol{\varepsilon}) = \ $ and $cov(\boldsymbol{\varepsilon}) = \tau\ \mathbf{I}$, where **I** is an $n\times$ identity matrix.

If $X_1, X_2, \ldots, X_k$ are continous variables, then (2.3.1) is the usual multiple linear regression model. If $X_1, X_2, \ldots, X_k$ are categorical, then (2.3.1) is an analysis of variance (ANOVA) model. If $X_1, X_2, \ldots, X_k$ are a mixture of both continous and categorical variables, then (2.3.1) is an analysis of covariance (ANCOVA) model (Milliken and Johnson, 1984). In short, multiple linear regression, ANOVA and ANCOVA models are special cases of the general linear model given above (Ravishanker and Dey, 2002; Milliken and Johnson,1984).

Estimation

The least squares estimators of the linear model parameters are obtained by minimizing the error sum of squares with respect to $\boldsymbol{\beta}$. That is, by minimizing

$$S(\boldsymbol{\beta}) = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

with respect to $\boldsymbol{\beta}$. Differentiating $S(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$, equating the derivative to zero and solving the equation for $\boldsymbol{\beta}$ gives the least squares estimator of $\boldsymbol{\beta}$ as

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

The estimator $\hat{\boldsymbol{\beta}}$ is multivariate normally distributed if and only if $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$, with variance-covariance matrix given by $var(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$. That is

$\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$. The unbiased estimator σ^2 is given by

$$s^2 = S(\hat{\boldsymbol{\beta}})/(n-p),\ p = k+1.$$

$\hat{\boldsymbol{\beta}}$ is statistically independent of $S(\hat{\boldsymbol{\beta}})$ (Seber and Wild, 2003), and $S(\hat{\boldsymbol{\beta}})/\sigma^2 \sim \chi^2_{n-p}$, where $p = k+1$.

An alternative method of estimating $\boldsymbol{\beta}$ and σ^2 is the maximum likelihood method. If $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$, then the likelihood function for $\boldsymbol{\beta}$ and σ^2 is given by

$$L(\boldsymbol{\beta}, \sigma^2 \mid \mathbf{y}) = (2\pi)^{-n/2}[\det(\sigma^2\mathbf{I})]^{-1/2}\exp\{-(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})/2\sigma^2\}.$$

The log likelihood function is given by

$$\log L(\boldsymbol{\beta}, \sigma^2 \mid \mathbf{y}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})/2\sigma^2.$$

$\log L(\boldsymbol{\beta},\sigma^2 \mid \mathbf{y})$ is maximized by choosing the values of $\boldsymbol{\beta}$ which minimize $(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})$ which is equivalent to minimizing the error sum of squares, and thus, the least squares estimator $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ is also the maximum likelihood estimator of $\boldsymbol{\beta}$ if $\boldsymbol{\varepsilon} \sim N(\mathbf{0},\sigma^2\mathbf{I})$ (Myers and Milton, 1991; Seber and Wild, 2003). Now substituting $\hat{\boldsymbol{\beta}}$ for $\boldsymbol{\beta}$ into the loglikelihood and differentiating with respect to σ^2 gives

$$\frac{\partial \log L(\boldsymbol{\beta},\sigma^2 \mid \mathbf{y})}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + (\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})/2\sigma^4.$$

Equating this derivative to zero and solving for σ^2 gives

$$\sigma^2 = (\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})/n$$

as $\hat{\sigma}^2$ is the maximum likelihood estimator of σ^2. This estimator is biased and it is therefore rarely used. The mean squared error

$$s^2 = (\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}})/(n-p)$$

is an unbiased estimator of σ^2 (Seber and Wild, 2003).

It can be shown that $(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})/\sigma^2 \sim \chi^2_p$ from which it follows that

$$\begin{aligned} F &= (\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})/ps^2 \\ &= \frac{(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})}{S(\hat{\boldsymbol{\beta}})}\frac{(n-p)}{p} \sim F_{p,n-p} \end{aligned}$$

A $100(1-\alpha)\%$ confidence region for $\boldsymbol{\beta}$ is given by $\{\boldsymbol{\beta}:(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}) \le ps^2 F^{\alpha}_{p,n-p}\}$, where $F^{\alpha}_{p,n-p}$ denotes the $(1-\alpha)\times 100^{th}$ percentile of the F distribution with p and n-p degrees of freedom. The hypothesis $H_0: \beta_1 = \beta_2 = \ldots = \beta_p = 0$ is tested using

$$F = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}/ps^2$$

which is distributed as $F_{p,n-p}$ when H_0 is true. Furthermore, the hypothesis $H_0: \mathbf{K}\boldsymbol{\beta} = \mathbf{0}$, where $\mathbf{K}$ $q \times p$ of rank q is tested using

$$F = \hat{\beta}'\mathbf{K}'[\mathbf{K}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{K}']^{-1}\mathbf{K}\hat{\beta}/qs^2$$

which is distributed as $F_{q,n-}$, when H_0 is true (Seber and Wild, 2003). We reject H_0 at α significance level of significance if $F > F^{\alpha}_{q,n-}$. Finally, for $i = ,2,\ldots,k$, the hypothesis $H_0 : \beta =)$ is tested using

$$t = \frac{\beta - }{se(\beta)} \sim t_{n-},$$

where $se(\beta)$ is the standard error of $\hat{\beta}$. We reject H_0 at α significance level if $|t| > t^{\alpha/2}_{n-}$, where $t^{\alpha/2}_{n-}$ denotes the upper $(1-\alpha/2)\times 00^{th}$ percentile of the Student *t* distribution with $n-$ degrees of freedom.

It is likely that the data can be adequately described by more than one model, and this makes it difficult to decide which model provides the best fit (Rust, *et al*. 1995). We therefore present some of the commonly used statistics for model selection.

The coefficient of determination is the percentage of total variability in the data that is accounted for by the linear model. It is given by

$$R^2 = \frac{SSR}{SST}$$

where $SSR = \mathbf{y}'\left[\mathbf{H} - \left(\frac{1}{n}\right)\mathbf{J}\right]\mathbf{y}$ and $SST = \mathbf{y}'\left[\mathbf{I} - \left(\frac{1}{n}\right)\mathbf{J}\right]\mathbf{y}$ are regression sum of squares and total sum of squares, respectively; $\mathbf{H}$ is the hat matrix defined as $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, $\mathbf{J}$ is an $n\times$ matrix of 1s and $\mathbf{I}$ is an $n\times$ identity matrix. A large value of R^2 implies that more variability in the data is captured by the model. However R^2 tends to increase as the number of explanatory variables increase, even if the added explanatory variables are not relevant.

The adjusted R^2 has a penalty for large number of explanatory values and hence it is used to compensate for the adequacy of R^2. It is given by

$$\bar{R}^2 = -\frac{SSE/(n-p)}{SST/(n-)},$$

where $SSE = \mathbf{y}'(\mathbf{I} - \mathbf{)y}$ is the error sum of squares. The mean squared error or the unbiased estimator of σ is another statistic that is used for model selection. Models with small mean squared error are preferred.

The linear regression frameworks can be extended to handle regression functions that are nonlinear in one or more explanatory variables. In particular, the linear form of regression does not describe curvilinear relationships. Curvilinear relationships are described using polynomial regression models (Hasani and Amidi, 1983; and Van Laar, 1991).

2.3.3 Polynomial regression models

A k^{th} order polynomial regression model with one independent variable, X, is written as

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta X_{ik} + \varepsilon \qquad (2.3.2)$$

where k is a positive integer; $X_{ij} = X^j$ for $j = 1, 2, \ldots, k$; the β_j's are the model parameters and ε is the random error term. Clearly, model 2.3.2 is the same as the general linear model. Hence the model is fitted to the data as described in Section 2.3.1. Any continuous regression function can be approximated over a limited range by a polynomial regression (Freund and Wilson, 1998).

An important issue in polynomial regression models is the specification of k, the degree of polynomial. The greater the value of k, the better the fit. However, as with most regressions, models with fewer parameters are preferred (Freund and Wilson, 1998). Furthermore, in polynomial regression model, the correlations among variables with different powers tend to be highly correlated and thus likely to result in multicollinearity (Freund and Wilson, 1998). Further discussions on polynomial regression can be found in Bradley and Srivastava (1979) among others.

Multicollinearity

Multicollinearity is a term used to describe a case when the inter-correlation of explanatory variables is high (Ravishanker and Dey, 2002). Multicollinearity does not invalidate the proposed model if the regression coefficients are interpreted collectively. However, multicollinearity is undesirable when interest is on the individual effects of explanatory variables because even if the overall model is significant, the individual parameters may be insignificant (Freund and Wilson, 1998).

In the presence of multicollinearity 1) variances of regression coefficients are unstable and thus making precise estimation is difficult; 2) it is not possible to separate the effect on one explanatory variable from others; 3) the relative magnitudes and even the signs of the coefficients may defy interpretation; and 4) the least squares estimators and their standard errors can be sensitive to small changes in the data (Neter, *et al*. 1996).

The variance inflation factor (VIF) is a statistic that can be used to identify multicollinearity. Accordingly, the VIF is function of the coefficient of multiple determination from the regression of each explanatory variable on all the other explanatory variables. The VIF is calculated using the formula

$$\mathrm{VIF}_i = \frac{1}{1 - R_i^2}$$

where R_i^2 is the coefficient of multiple determination in a regression of the i^{th} explanatory variable on all other explanatory variables, that is, VIF_i is the variance inflation factor associated with the i^{th} explanatory variable. If the i^{th} explanatory variable is independent of the other explanatory variables, then the variance inflation factor is one, while if the i^{th} explanatory variable can almost be perfectly predicted from other explanatory variables, the variance inflation factor approaches infinity. Multicollinearity is considered to be a problem when the variance inflation factor of one or more explanatory variables is large. Some researchers use a VIF of 5 and others use a VIF of 10 as a critical threshold (Ravishanker and Dey, 2002). The VIF of 10 was used as a cut-off point in this thesis.

2.3.4 Generalized linear models

Linear regression, ANOVA and ANCOVA models are special forms of the general linear model. The least squares criterion is used to obtain estimates of the parameters of such models. Certain assumptions must be satisfied in order to test the hypothesis about the model's parameters (Neter, *et al*. 1996). Apart from the assumption of independence, the hypothesis tests derived from linear model require normality of the response variable and homogeneity of the error variances. However, when the distribution of the response variable is not normal or if the error variances are not homogeneous, then the inferences from fitting the linear model can be invalid. Traditionally, transformation of the scale of the response variable is applied to insure that the assumptions required for the validity of hypotheses tests

are met. However, transformations are not always successful in achieving the desired end. The generalized linear model is an extension of the general linear model. The response variables under generalized linear model have distributions that belong to exponential family of distribution. The theory of these models is discussed in McCullagh and Nelder (2002); Meyer and Laud (2002); and Moeti (2007).

The generalized linear models are models for response variable which include response variables with distributions that belong to the exponential family. The density of these distributions has the form

$$f(y) = \exp\left[\frac{\theta y - b(\theta}{a(\phi} + c(y,\phi\right]$$

where θ is the parameter that determines the location of the distribution, ϕ is a parameter that scales the variability, $a(\phi$ and $c(y,\phi$ are functions such that $a(\phi = \; w$ and $c(y,\phi = \; (y,\phi \; w)$, where w is a known weight for each observation. The mean and variance for distribution are

$$\mu = \; (y) = \; ' \; \theta \quad \text{and } var(y) = \; ' \; \theta \; a(\phi \; .$$

The variance can also be expressed as a function of the mean (μ , and is thus given as

$$var(y) = \; (\mu \; a(\phi$$

where $V(\cdot)$ is called the variance function. The mean μ of the i^{th} response is related to the explanatory variable through a link function g as

$$g(\mu \;) = \mathbf{X}'\boldsymbol{\beta}, \; i = \; , 2, ..., n$$

where $\mathbf{X}_i$ and $\boldsymbol{\beta}$ are the known vector of explanatory variables and the vector of uknown parameters, respectively.

The parameter estimates are obtained using the maximum likelihood method. The log-likelihood of $\mathbf{y}$ is given as

$$l(\boldsymbol{\theta},\phi;\mathbf{y}) = \sum_{i=} \quad \frac{\vartheta_i - \nu(\theta_i)}{a(\theta_i)} + \varsigma(y_i,\phi).$$

The hypotheses tests are based on comparison of likelihoods or deviances of nested models. In particular, the goodness of fit of a generalized linear model can be measured by a scaled deviance defined as

$$D(\mathbf{y};\hat{\boldsymbol{\mu}}) = 2[l(\mathbf{y},\mathbf{y}) - (\hat{\boldsymbol{\mu}},\mathbf{y})]$$

where $l(\mathbf{y};\mathbf{y})$ is the maximum log-likelihood achievable for an exact fit in which the fitted values are equal to the observed values, and $l(\hat{\boldsymbol{\mu}};\mathbf{y})$ is the log-likelihood function evaluated at the estimated parameter β. The best models are the ones with the smallest deviances.

The linear models discussed so far are good for modeling the response as a function of the immediate effects of explanatory variables(s) that are measured over time. However, the effects of explanatory variable(s) may be delayed. The time series approach through autoregressive integrated moving average (ARIMA) models is used to assess the delayed or lagged effects of explanatory variables on the response variables. The theory of these models is discussed in Wei (1990); Kendall and Ord (1990); Diggle (1990); Janacek and Swift (1993); Walton (1997); Chatfield (2000); and Janacek (2001) among others.

2.4 ARIMA models

2.4.1 Arima models overview

A time series is a set of data $\{Y_t : t = ,..., N\}$ in which the subscript t indicates the time at which the datum Y_t was observed. Time series without seasonal effects are usually modeled using autoregressive integrated moving average (ARIMA) which are discussed below.

Let B be the back-shift operator defined as $B^i Y_t = Y_{t-}$. Furthermore, suppose that for some nonnegative integer d, the series $\{(1-3)^d Y_t\}$ is stationary, i.e. does not depend on time t. Then the model for $\{Y_t\}$ is said to be an autoregressive integrated moving average of order p,d,q, written ARIMA(p,d,q), given by

$$\phi(B)(1-3)^d Y_t = u + \vartheta B)Z_t$$

where

$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ and $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$ are polynomials in B of order p and q, respectively, with all roots outside the unit circle; and $\{Z_t\}$ is white noise. Fitting the ARIMA (p,d,q) model proceeds in three stages: (1) model identification; (2) model estimation; and (3) model diagnostics. These stages are discussed below.

Model identification

The first step in the identification stage of an ARIMA process is examination of a time series plot (graph of Y_t versus t) for the existence of trends, seasonality, cyclicity and outliers. In the absence of outliers, cyclicity and nonlinear trends, polynomial trends and seasonality are removed by differencing the series. Nonlinear trends may be linearized by transforming the series to linearity. For example, the differenced series $(1-B)^d Y_t$ is free of polynomial trend of degree *d*. If the series appears to be stationary, no differencing is called for (Kendall and Ord, 1983). If the series appears to be non-stationary, the series is successively differenced until the differenced series is stationary. The optimal order of differencing is *d* for which the standard deviation of $\{(1-B)^d Y_t\}$ is the smallest.

The next step is to examine the autocorrelations of the stationary series at lag *k*. For a stationary series $\{Y_t : t = 1, \ldots, N\}$ with sample mean $0\bar{Y}$, the sample autocovariance at lag k is given by

$$\gamma(k) = \frac{1}{N}\sum_{t=1}^{N-|k|}(Y_t - \bar{Y})(Y_{t+|k|} - \bar{Y}), \quad k = 0, \pm 1, \pm 2, \ldots$$

or the unbiased form

$$\gamma(k) = \frac{1}{N-|k|}\sum_{t=1}^{N-|k|}(Y_t - \bar{Y})(Y_{t+|k|} - \bar{Y}), \quad k = 0, \pm 1, \pm 2, \ldots$$

The corresponding sample autocorrelation function (ACF) is defind as

$$r_k = \gamma(k)/\gamma(0), k = 0, \pm 1, \pm 2, \ldots$$

and is the estimator of the true ACF $\rho_k = Corr(Y_t, Y_{t+k})$, defined as the correlation between observations at a given time *k* apart, $k = 0, \pm 1, \pm 2, \ldots$

The plot of r_k versus k is called a sample correlogram whose pattern is used to identify the model for the series. For example, for a moving average model of order *q* (MA(*q*)), written as $Y_t = \theta(B)Z_t$, the autocorrelations are zero after lag *q* while for an autoregressive model of

order p (AR(p)), written as $\phi(B)Y_t = Z_t$ the autocorrelations decay exponentially. Furthermore, for a mixed AR(p) and MA(q) model (ARMA(p,q)), written as $\phi(B)Y_t = \theta(B)Z_t$, the autocorrelations are expected to decay exponentially and may contain damped oscillations (Janacek, 2001).

For large N, the approximate sampling distribution of each r_k is normal, with mean zero and variance $1/N$. Thus, the limits $\pm .96/\sqrt{N}$ are used to check individual r_k for significant departure from zero. These limits are used as a rough guide to interpreting a correlogram. Diggle (1990) emphasizes that these limits should not be interpreted strictly.

In addition to the use of the autocorrelation function (ACF) for model identification, the autocorrelation between Y_t and Y_{t+} after their mutual linear dependency on the intervening variables $Y_{t+}, ..., Y_{t+k-}$ has been removed, called the partial autocorrelation function (PACF) is also used for model identification. The properties of the ACF and the PCAF for ARMA(p,q) are summarized in Table 2.4.1

Table 2.4.1 The properties of the ACF and PCAF of an ARMA(p,q) process

Process	ACF	PACF
AR (p)	Exponential decay	Zero after lag p
MA (q)	q spikes	Exponential decay
ARMA (p,q)	Exponential decay after lag (q-p)	Decay after lag (p-q)

Model estimation

Once the ARIMA(p,d,q) model has been identified, the method of maximum likelihood is used to estimate the model parameters. Suppose that $\{X_t = (1-B)^d Y_t\}$ is described by the ARMA(p,q) model

$$X_t = \mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + Z_t - \theta_1 Z_{t-1} - ... - \theta_q Z_{t-q} \qquad (2.4.1)$$

where $\{Z_t\}$ is Gaussian white noise with variance σ . Let $\boldsymbol{\varphi} = (\phi_1, \phi_2, \ldots, \phi_p)'$

and $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_q)'$. Then, the joint density of $\{Z_t\}$ is given by

$$f(\{Z_t\} \mid \boldsymbol{\varphi}, \mu\ \boldsymbol{\theta}, \sigma\ ) = \frac{1}{(2\pi\sigma\ )^{n/2}} \exp\left[- \frac{1}{2\sigma} \sum_{t=1}^{N} Z_t^2 \right].$$

Rewriting (2.4.1) as

$$Z_t = \mu + X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \ldots - \phi_{,} X_{t-,} + \theta_1 Z_{t-1} + \ldots + \theta_{,} Z_{t-}$$

and substituting for Z_t in $f(Z_t \mid \boldsymbol{\varphi}, \mu\ \boldsymbol{\theta}, \sigma\ )$ gives the log-likelihood function

$$l(\boldsymbol{\varphi}, \mu\ \boldsymbol{\theta}, \sigma\ ) = - \frac{N}{2} \log[2\pi\sigma\ ] - \frac{1}{2\sigma} S(\boldsymbol{\varphi}, \mu\ \boldsymbol{\theta}, \sigma\ )$$

where $S(\boldsymbol{\varphi}, \mu\ \boldsymbol{\theta}, \sigma\ ) = \sum_{t=1}^{N} Z_t^2$. The maximum likelihood estimators are the values of $(\boldsymbol{\varphi}, \mu\ \boldsymbol{\theta}, \sigma\ )$, denoted by $(\hat{\boldsymbol{\varphi}}, \mu\ \hat{\boldsymbol{\theta}}, \sigma\ )$ which maximize $l(\boldsymbol{\varphi}, \mu\ \boldsymbol{\theta}, \sigma\ )$.

Model diagnostics

After model identification and estimation, the model adequacy is checked. If the model fitted adequately describes the series $\{X_t\}$, and the fitted value conditional of the previous $(t - \ )$ value is $\hat{X}_{t|t-}$, then the residuals

$$\hat{Z}_t = X_t - \hat{X}_{t|t-}$$

are expected to behave like Gaussian white noise. In particular, the ACF of the white noise residuals $\{\hat{Z}_t\}$ is expected to be uniformly equal to zero (Janacek, 2001). Alternatively, the adequacy of the model can be checked using the portmanteau test (Box and Jenkins, 1970). This test uses all the sample ACF's as a unit to test the joint null hypothesis

$$H_0 : \rho\ = \rho\ = \cdots = \rho\ =)$$

with the test statistic

$$Q = N(N+2) \sum_{k=1}^{K} \frac{1}{N-k} r_k^2 \ ,$$

where K is the number of coefficients to test autocorrelation. Under $H_0 : \rho\ = \rho\ = \cdots = \rho\ =)$, the Q statistic approximately follows the Chi-squared distribution with $K - \ ) -$ degrees of freedom. The hypothesis that the residuals series is white noise is rejected if Q exceeds the percentage point of the chi-squared distribution (Janacek, 2001).

There may be several competing ARMA(*p,q*) that adequately describe the series $\{X_t\}$. The Akaike's information criterion (AIC) and the Bayesian information criterion (BIC) are two of main criteria of choosing the best model among the competing models. Suppose that an ARMA(*p,q*) model with $M = p + q +$ parameters is fitted to the series $\{X_t\}$. Let σ be the maximum likelihood estimate of the variance of the residuals $\{Z_t\}$. Then, the AIC is given by

$$\text{AIC}(M) = \ln \sigma \; + \; M / N .$$

The best ARMA(*p,q*) model for $\{X_t\}$ is one with the smallest AIC. Wei (1990) reports that the AIC tends to overestimate the order of the autoregression. The BIC defined as

$$\text{BIC}(M) = \ln \sigma \; + \; M \ln(N) / N$$

is the extension of AIC. Again the best ARMA(*p,q*) is one with the smallest BIC.

2.4.2 Transfer function models

The ARIMA models discussed in Section 2.4.1 are useful for modeling univariate series. That is, the models that relate a series to its own past. However a time series may not only be related to its own past, but may also be influenced by the present and/or past values of other time series. Regression like models for modeling a time series as a function of other time series are referred to as transfer function models (Kendall and Ord, 1983). The following is an overview on transfer function models.

Suppose that the explanatory series $\{X_t\}$ is white noise and that response series $\{Y_t\}$ is stationary. The linear model that includes all possible lags of the explanatory series is given by

$$Y_t = u + \beta_0 X_t + \beta_1 X_{t-1} + \ldots + \beta \; X_{t-} + \; , \quad t \geq \qquad (2.4.2)$$

The set of coefficients $\{\beta \; , \beta \; , \ldots , \beta \; \}$ is called the impulse response function. The relationship between $\{X_t\}$ and Y_t is determined from the cross-correlation function (CCF), given by

$$\begin{aligned} \rho_{XY}(k) &= \text{Corr}(X_{t-} \; , Y_t), \quad k = \; ,1,2,\ldots \\ &= \beta \; \sigma_{} \; / \sigma_{} \end{aligned}$$

where var(Y) = τ and var(X) = τ. Thus, ρ(k) =) for all k implies that β =) for all k, which implies that $\{X_t\}$ and Y_t are not linearly related. On the other hand if | ρ(k) |>) for some k then | β |>) for that k, which means Y_t has a linear relationship with X_{t-} .

If X_t is stationary but not necessarily white noise, the CCF and the impulse response function is contaminated by the autocorrelation structure of the explanatory series $\{X_t\}$ and thus making the identification of significant/insignificant β difficult. Suppose that X_t may be described by an ARMA process

$$\phi(B)X_t = \vartheta B)Z_t$$

where $\{Z_t\}$ is white noise. Then

$$Z_t = \frac{\phi(B)}{\theta B)} X_t .$$

Assuming that in (2.4.2) the constant term is zero, and that ε is negligible, the model can be rewritten as

$$Y_t = \}(B)X_t \qquad (2.4.3)$$

where $\beta(B) = 1 + \beta_1 B + \beta_2 B^2 + \ldots + \beta B^k$. Multiplying both sides of (2.4.3) by $\phi B)/\theta B)$ gives

$$V_t = \frac{\phi(B)}{\theta B)} Y_t = \frac{\phi(B)}{\theta B)} \beta B)X_t = \vartheta B)Z_t$$

so that the CCF of V_t and Z_t provides an indication of the terms of the impulse response function to retain in the model (2.4.3). The sample CCF after pre-whitening is computed to obtain estimates of β (Kendall and Ord, 1990).

2.5 Nonlinear regression model

Linear models are generally satisfactory approximations for most regression applications. However, there are occasions when a theoretically justified nonlinear regression model is more appropriate (Neter, et al. 1996). The following gives an overview of nonlinear regression models. A nonlinear regression model has the form

$$y_i = f(\mathbf{x}_i, \boldsymbol{\gamma}) + \varepsilon, \qquad i = ,2,\ldots,n \qquad (2.5.1)$$

where the y_i are the responses, $f(\mathbf{x}_i, \boldsymbol{\gamma})$ is a known function of the covariate vector $\mathbf{x}_i = (x_{i1}, x_{i2}, ..., x_{ik})'$ and the parameter vector $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_p)'$, the ε are random errors which are assumed to be independent and normally distributed with mean zero and constant variance.

The unknown vector $\boldsymbol{\gamma}$ can be estimated by minimizing the error sum of squares with respect to $\boldsymbol{\gamma}$. That is, by minimizing

$$S(\boldsymbol{\gamma}) = \sum_{i=1}^{n} [y_i - f(\mathbf{x}_i, \boldsymbol{\gamma})]^2$$

with respect to $\boldsymbol{\gamma}$. Equating the derivatives to zero and replacing the parameters γ by the least squares $\hat{\gamma}$, we obtain *p* normal equations.

$$\sum_{i=1} [y_i - f(\mathbf{x}_i, \hat{\boldsymbol{\gamma}})] \left[\frac{\partial f(\mathbf{x}_i, \hat{\boldsymbol{\gamma}})}{\partial \gamma} \right] = 0, \quad i = 1, ..., p$$

where $\hat{\boldsymbol{\gamma}} = (\hat{\gamma}_1, ..., \hat{\gamma}_p)'$. These equations are in general nonlinear and are usually difficult to solve. Hence, numerical iterative procedures are used to obtain a solution of the normal equations. To add to the difficulties, multiple solutions may be possible corresponding to multiple stationary values of $S(\hat{\boldsymbol{\gamma}})$ (Nash and Walker-Smith, 1987; Fox, 2002; Karim, 2002; Seber and Wild, 2003).

One of the most widely used method of solving the nonlinear equations is the Gauss-Newton method, also referred to as the linearization method (Gallant, 1975). The linearization method uses the results of the linear least squares method in a succession of stages. This method begins with the initial values of $\gamma_1, ..., \gamma_p$. These initial values are denoted by $\gamma_{10}, ..., \gamma_{p0}$. These values may be guesses or preliminary estimates from previous or related studies. However, a poor choice of starting values may result in slow convergence or no convergence at all (Neter, *et al*. 1996). Any continuous function *f* can be written as

$$f(x) = f(a) + (x-a) f'(a) + \frac{(x-a)^2}{2!} f''(a) + ... + \frac{(x-a)^p}{p!} f^{(p)}(a) + R_{p+1}$$

expanded by Taylor series where $f'(a) = df(x)/dx$ at $x = a$ and R_{p+1} is the remainder term. If the Taylor series expansion of $f(\mathbf{x}_i, \boldsymbol{\gamma})$ is carried out about the point $\boldsymbol{\gamma}_0 = (\gamma_{10}, ..., \gamma_{p0})'$

and $\boldsymbol{\gamma}$ is close to $\boldsymbol{\gamma}_0$, then linearization is accomplished by a Taylor series expansion of $f(\mathbf{x}_i, \boldsymbol{\gamma})$ about $\boldsymbol{\gamma}_0$ with only linear terms retained. Now $f(\mathbf{x}_i, \boldsymbol{\gamma})$ for the *n* cases by the linear terms in the Taylor series expansion is given by

$$f(\mathbf{x}_i, \boldsymbol{\gamma}) \approx f(\mathbf{x}_i, \boldsymbol{\gamma}_0) + \sum_{j=} \left[\frac{(\mathbf{x}_i, \boldsymbol{\gamma})}{\partial} \right]_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_0} (\gamma_j - \gamma_{j0}) \qquad (2.5.2)$$

For simplicity and convenience, we set:

$$f_i^0 = f(\mathbf{x}_i, \boldsymbol{\gamma}_0),$$
$$\theta_j = (\gamma_j - \gamma_{j0}),$$
$$D_{ij}^0 = \left[\frac{\partial f(\mathbf{x}_i, \boldsymbol{\gamma})}{\partial \gamma_j} \right]_{=0}$$

Taylor approximation (2.5.2) becomes

$$f(\mathbf{x}_i, \boldsymbol{\gamma}) \approx f_i^0 + \sum_{j=} \theta_j .$$

An approximation to the nonlinear regression model (2.5.1) is

$$y_i \approx f_i^0 + \sum_{j=1} \theta_j^0 + \varepsilon \qquad (2.5.3)$$

If we define $y_i^0 = y_i - f_i^0$, then the model (2.5.3) becomes the linear regression model approximation

$$y_i^0 \approx \sum_{j=1} \theta_j^0 + \varepsilon .$$

This linear regression model approximation can be represented in matrix form as

$$\mathbf{y}^0 \approx \mathbf{D}^0 \boldsymbol{\theta}^0 + \boldsymbol{\varepsilon}$$

which is in the form of the general linear regression model, where the **D** matrix of partial derivatives plays the role of the **X** matrix and $\boldsymbol{\theta}^0 = (\theta_1, \ldots, \theta_p)'$ parameters can be estimated by ordinary least squares given by

$$\hat{\boldsymbol{\theta}}^0 = (\mathbf{D}^0{}'\mathbf{D}^0)^- \mathbf{D}^0{}'\mathbf{y}^0 .$$

Thus, the vector $\hat{\boldsymbol{\theta}}^0$ minimizes the error sum of squares with respect to $\boldsymbol{\theta}^0$. We denote least squares criterion measure evaluated for $\hat{\boldsymbol{\theta}}^0$ by SSE^0. After the first iteration, the estimated regression coefficients are $\hat{\boldsymbol{\theta}}^1$, and the least squares criterion measure evaluated at this stage is denoted by SSE^1. If the Gauss-Newton method is working effectively in the first iteration, $SSE^1 < SSE^0$. The iterative process is continued until the differences between successive coefficient estimates $\hat{\boldsymbol{\theta}}^{l+1} - \hat{\boldsymbol{\theta}}^l$ and the difference between successive least squares criterion measures $SSE^{l+1} - SSE^l$ become negligible. There are alternative approaches for minimizing the error sum of squares in nonlinear regression models. Some of these approaches are: direct search; steepest descent and Marquardt compromise method. Further discussion of these methods can be found in Gallant (1987); Bates and Watts (1988); Small and Wang, (2003) and Seber and Wild (2003) among others.

If the errors ε follow a normal distribution, then the least squares estimator for $\boldsymbol{\gamma}$ is also the maximum likelihood estimator (Smyth, 2002; Fox, 2002; Seber and Wild, 2003).

The likelihood for the nonlinear regression model is

$$L(\boldsymbol{\gamma},\sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{ - \frac{\sum_{i=1}^{n} y_i - f(\mathbf{x}_i\boldsymbol{\gamma})]^2}{2\sigma^2} \right\}$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{ - \frac{1}{2\sigma^2} S(\boldsymbol{\gamma}) \right\} \quad \text{(Smyth, 2002)}$$

This likelihood is maximized when the error sum of squares $S(\boldsymbol{\gamma})$ is minimized. That is when $\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}}$ (least squares estimator). Furthermore, the estimate of the variance of errors ε is given by

$$s^2 = \frac{1}{n-p} S(\hat{\boldsymbol{\gamma}}) .$$

A $100(1-\ \%$ confidence region for $\boldsymbol{\gamma}$ is given by $\left\{\boldsymbol{\gamma}: \frac{S(\boldsymbol{\gamma})-S(\hat{\boldsymbol{\gamma}})}{S(\hat{\boldsymbol{\gamma}})} \leq \frac{p}{n-p} F_{p,n-p}^{\alpha}\right\}$, where $F_{p,n-}^{\alpha}$, denotes the $(1-\alpha \times 00^{th}$ percentile value of the *F* distribution with *p* and *n-p* degrees of freedom..

The hypothesis $H_0: \boldsymbol{\gamma} = \ _0$ versus $H_0: \boldsymbol{\gamma} \neq \ _0$ is tested using

$$F = \frac{S(\boldsymbol{\gamma}_0) - S(\hat{\boldsymbol{\gamma}})}{S(\hat{\boldsymbol{\gamma}})} \frac{n-}{p}.$$

We reject H_0 at the α level of significance if $F > \ ^{\alpha}_{p,n-}$.

Let

$$\mathbf{f}(\boldsymbol{\gamma}) = \ f_1(\boldsymbol{\gamma}), f_2(\boldsymbol{\gamma}),..., f_n(\boldsymbol{\gamma}))$$

and

$$C = \mathbf{F}_{\bullet}'\mathbf{F}_{\bullet}$$

where

$$\mathbf{F}_{\bullet} = \ _{\bullet}(\boldsymbol{\gamma}) = \frac{\partial \ \boldsymbol{\gamma})}{\partial \ '}$$

The null hypothesis about individual $\gamma_j's$ is: $H_0: \gamma_j = \ _{j0}$, and is tested using the statistic

$$t = \frac{\gamma_j - \gamma_{j0}}{\sqrt{s^2 \hat{c}_{jj}}} \sim t_{n-},$$

where $\hat{c}_{jj}$ denotes the *j*th diagonal element of estimated $C(\hat{C})$ matrix. This test rejects H_0 if $|t| > \ ^{\alpha}_{n-}$, where t^{α}_{n-} denotes the $(1-\alpha\ 2)\times 00^{th}$ percentile of the *t* distribution with $n-$ degrees of freedom at $\alpha\ 2$ significance level. The approximate confidence limits for γ_j are given by

$$\gamma_{\ } \pm \ ^{\alpha}_{-}s\sqrt{\hat{c}_{jj}}.$$

Among other models, the piecewise regression and Growth curve models are types of nonlinear regression model and they are discussed in the following sections.

2.5.1 Piecewise regression model

Consider a regression relationship between y and x in which the regression function $E[y \mid x] = f(x;\boldsymbol{\gamma})$ is given

$$E[y \mid x] = \begin{cases} f_1(x;\boldsymbol{\gamma}_1), & \alpha_0 < x \le \alpha_1 \\ f_2(x,\boldsymbol{\gamma}_2), & \alpha_1 < x \le \alpha_2 \\ \vdots & \vdots \\ f_r(x;\boldsymbol{\gamma}_r), & \alpha_{r-1} < x \le \alpha_r \end{cases}$$

where $E[y \mid x]$ is continous in x throughout the interval $[\alpha_0, \alpha_r]$; r, α_0, α_r, and the functional form of each $f_i(x;\boldsymbol{\gamma}_i)$ are assumed to be known; the parameters $\boldsymbol{\gamma}_1,...,\boldsymbol{\gamma}_r$ and the transition points or breakpoints $\boldsymbol{\alpha} = (\alpha_1,...,\alpha_{r-1})$ are to be estimated from the data. This model is called piecewise or multiphase regression. Other authors like Lerman (1980) refer to this model as segmented regression model. An extended overview concerning piecewise regression model can be found in Seber and Wild (2003).

Continuity of $E[y \mid x]$ at the breakpoints α_i implies that the submodels $f_i(x;\boldsymbol{\gamma}_i)$ are subject to the constraints

$$f_i(\alpha_i, \boldsymbol{\gamma}_i) = f_{i+1}(\alpha_i, \boldsymbol{\gamma}_{i+1}),\ i = 1, 2, \ldots, r-1. \tag{2.5.4}$$

When inferences are to be made, the y_j are assumed to be independent and normally distributed. The least squares estimators of $\boldsymbol{\gamma}_1,...,\boldsymbol{\gamma}_r,\boldsymbol{\alpha}$ denoted by $\hat{\boldsymbol{\gamma}}_1,...,\hat{\boldsymbol{\gamma}}_r,\hat{\boldsymbol{\alpha}}$, minimize the error sum of squares function

$$S(\boldsymbol{\gamma},\boldsymbol{\alpha}) = \sum_{i=1}^{r} \sum_{x_{i-1} < x_j \le x_i} [y_i - f_i(x,\boldsymbol{\gamma}_i)]^2$$

subject to the continuity constraints (2.5.4). Let S^0 denote the minimum attainable value of S under H_0, p denotes the dimension of the model parameter space (= number of unknowns minus number of constraints), $s^2 = \hat{S}/(n-p)$ denote the residual mean square and $F_{u,v}^{\alpha}$ denotes the $(1-\alpha) \times 100^{th}$ percentile of the F distribution with u and v degrees of freedom (Lerman, 1980). The likelihood ratio statistic for testing H_0 is approximated by $R = (S^0 - \hat{S})/s^2$. It can be shown that if the model is identified and if H_0 is true, the asymptotic distribution of R is χ^2_q (where q is the dimensional reduction in the model's

parameter space imposed by H_0) and $\hat{\boldsymbol{\gamma}}_1, ..., \hat{\boldsymbol{\gamma}}_r, \hat{\boldsymbol{\alpha}}$ are asymptotically normal with known variance-covariance structure. The approximate $100(1-$ $\%$ confidence region for the breakpoints, $\boldsymbol{\alpha}$, derived from *R* comprises values which satisfy

$$S(\boldsymbol{\alpha}) \leq C^m,$$

Where $C^m = \hat{\ } + {}^2(r-1)F^m_{q,n-}$,. Approximate standard errors for $\hat{\boldsymbol{\gamma}}_1, ..., \hat{\boldsymbol{\gamma}}_r, \hat{\boldsymbol{\alpha}}$ may be obtained from the information matrix conditional on the data partition implied $\hat{\boldsymbol{\alpha}}$ without the imposition of the continuity constraints (2.5.4). Further reading on inference for piecewise linear regression model can be found in Lerman (1980) and Küchenhoff(1996) among others.

2.5.2 Growth curve models

Long term size changes observed throughout the life of individuals are difficult to quantify. Mathematical models are employed to simplify and describe these changes in fewer parameters. Growth models express the lifetime growth course mathematically. To accomplish this purpose, the growth models must comply with some requirements. The estimated size must correspond to the actual observed size and the model parameters have biological meaning. For many types of growth data, the growth rate does not steadily decline, but rather increases to a maximum before steadily declining to zero. This is shown in the growth curve by an S-shaped or sigmoidal pattern (Figure 2.5.1). The theory of growth models is discussed in Ricklefs (1967); Frazer and Ehrhart (1985); Frazer, Gibbons and Greene (1990); Bozkurt and Erkmen (2001); Seber and Wild (2003); and Lei and Zhang (2004) among others.

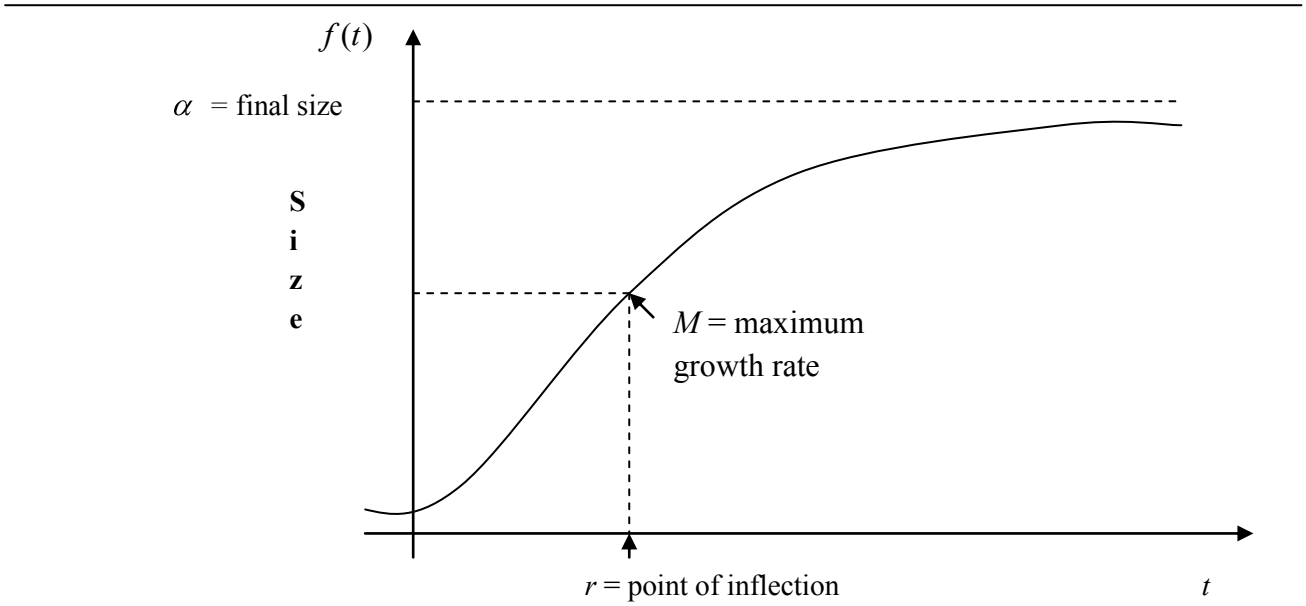

Figure 2.5.1 Sigmoidal growth curve with α = final size, r = point of inflection and M =maximum growth rate.

Most growth curves are special cases of the four parameter Richards model (Hassen, et al., 2004) which is derived as follows:

The growth rate is a proportional to the current size ($f = f(t)$), that is

$$\frac{df}{dt} = k(\alpha - f) \qquad (2.5.5)$$

where α is the upper asymptotic value and k is the scale parameter. Equation (2.5.5) can be transformed as follows

$$\frac{df^{1-\delta}}{dt} = k(\alpha^{1-\delta} - f^{1-\delta}) \qquad (2.5.6)$$

where $\delta \neq 1$ is the shape parameter. Equation (2.5.6) can be rewritten as

$$(1-\delta) f^{-\delta} \frac{df}{dt} = k(\alpha^{1-\delta} - f^{1-\delta}),$$

and hence

$$\frac{df}{dt} = \frac{k}{1-\delta} f\left[\left(\frac{f}{\alpha}\right)^{\delta-} - \right] \quad \delta \neq \qquad (2.5.7)$$

Equation (2.5.7) is the Richards growth rate equation. The integral forms of (2.5.7) describe size as an explicit function of age/time and can provide additional information about growth patterns. The following are special cases of the Richards growth model.

If in equation (2.5.5),

$$\frac{df}{dt} = \cdot \; y$$

where

$$y = \begin{cases} \alpha^{\delta} - f^{1-\delta}, \delta < 1 \\ f^{1-\delta} - \alpha^{\delta}, \quad \delta > 1 \end{cases},$$

then

$$f(t) = \chi\left] + \delta - \right) e^{-k(t-r)\frac{-1}{(1-\delta)}}, \; \delta \neq$$

where α is an upper asymptotic value of $f(t)$, k is the scale parameter governing growth rate, δ is the shape parameter of the growth curve and the parameter r locates the point of inflection on the t-axis. Differentiating (2.5.7) with respect to t obtains

$$\frac{d^2 f}{dt^2} = \frac{k}{1-\delta}\left[\delta\left(\frac{f}{\alpha}\right)^{\delta-} - \right]\frac{df}{dt}, \quad \delta \neq .$$

This derivative is zero at $f = \alpha\ \delta^{(\delta-)}$ when $\delta >$. The maximum growth rate is

$$m = k\alpha\delta^{(1-\delta)}.$$

The Richards growth model includes: the logistic model ($\delta =$); the Gompertz model (by taking the limit of f as $\delta \to$; and the von Bertalanffy model ($\delta = /3$). Table 2.5.1 provides the properties of these growth models. Most statistical programs require the partial derivatives of this growth curves to be specified. These derivatives are presented in Table 2.5.2.

Table 2.5.1 Properties of growth curve models

Properties	Logistic	Gompertz	von bertalanffy
Form, $f(t)$	α 1+ xp(– (t –))	α xp(– xp(– (t –)))	α 1+ xp(– (t – ·))^3
Inflection point	α 2	α ?	2α 3

$e =$ xp(1)

Table 2.5.2 Growth models and partial derivatives

Logistic model
$\partial (t)/\partial\alpha$: /(1+ xp(– (t –)))
$\partial (t)/\partial = \alpha$ t – ·) exp(– (t – ·)) /(1+ xp(– (t – ·)))^2
$\partial (t)/\partial = - k$ exp(– (t – ·)) /(1+ xp(– (t – ·)))^2
Gompertz model
$\partial (t)/\partial\alpha$: xp(– xp(– (t –)))
$\partial (t)/\partial = \alpha$ –) exp(– xp(– (t –))) exp(– (t –))
$\partial (t)/\partial = - k$ exp(– xp(– (t –))) exp(– (t –))
von Bertalanffy model
$\partial (t)/\partial\alpha$ = 1– xp(– (t – ·)))^3
$\partial (t)/\partial = 3\alpha$ t – ·)(1– xp(– (t – ·)))^2 exp(– (– (t – ·))
$\partial (t)/\partial = - ak$(1– xp(– (t – ·)))^2 exp(– (t – ·))

2.6 Model diagnostics

The model represents how the data were generated. It is important to assess the agreement between the model and the data. This assessment involves checking whether or not the model assumptions are satisfied, whether or not the model components should be refined, and assessing the sensitivity of the results to the model and the data. This process is called model diagnostics. The basic building blocks of model diagnostics are: 1) residual analysis; 2) goodness of fit tests; and 3) influence analysis. These diagnostics may result in reformulation of the model. The following is a discussion on these building blocks of model diagnostics.

Residuals are the deviation of the predicted values from the observed values. The residuals are plotted against the predicted values. This plot will display a random scatter of the residuals about mean zero if the fitted model is adequate. Departures of data from the model will be indicated by a pattern in the plot of residuals against the predicted values. This will imply that: 1) some explanatory variable(s) were inappropriately transformed or that some important explanatory variable(s) are not captured in the model; 2) the assumption of constant variance of errors is violated. The residual vs. Age plot is used to investigate the assumption of independence of errors. This plot will display a random scatter of the residuals about mean zero if the fitted model is adequate. The normal probability plots (QQ plot) are used to investigate the assumption of normality of errors. The QQ plot displays a straight line if the fitted model is adequate.

Residual analysis may reveal potential outliers and influential observations. An observation is said to be an outlier if it is inconsistent with the rest of the data. Such observations can be influential depending on their ability to alter important aspects of the results through their presence or absence. If their absence or presence do not alter results and do not violate the model assumptions, then such observations are not influential. Well developed and commonly used statistics to detect influential observations include Cook's distance, difference in fit values (DFFTS), etc. Further discussions on diagnostic measures can be found in Bowerman, O'Connell and Dickey (1986); Artkinson (1985); Chatterjee and Hadi (1988); Neter, *et al*. (1996) among others.

In data analysis, it is likely that there will be competing models. In such cases, the best model is selected based on the goodness of fit. Some of the commonly used criteria in assessment of the goodness of fit are: R^2; adjusted R^2; AIC; and BIC. These criteria were discussed in Subsection 2.3.2 as well as in Section 2.4.

3. The effects of the continuous climatic variables on radial growth

In this chapter, the statistical models reviewed in Chapter 2 were used to investigate the effects of climatic factors on radial growth of *Eucalyptus* clones. The statistical software used to investigate the effects of climatic factors on radial growth is SAS version 9.1.3. Figure 2.1.1 indicated that the radial growth of the two clones diverged over time. The following ANCOVA model was used to confirm that the radial growth of the two clones diverged significantly over time

$$RadialG = \beta_0 + \beta_1 age + \beta_2 clone + \beta_3 clone * age + \varepsilon$$

where *clone* is 0 if the clone is GC1 and 1 if the clone is GU1; *RadialG*is the radial growth of the two clones and the age is in days. The analyses of the data for the two clones will be carried out separately if the ($clone * age$) interaction effects are found to be significant. The ANOVA results for this model are presented in Table 3.1.1. These results indicate that this model was significant. Furthermore, the Type III test for the clone by age interaction effect presented in Table 3.1.2 indicates that the interaction effect was significant (p<0.0001). This implies that the radial growth of the two clones diverged significantly over time (also see Figure 2.1.1). It is therefore justifiable to conduct the analyses separately for the two clones.

Table 3.1.1 ANOVA results for ANCOVA model to assess the difference between clones over time

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	492149167942	164049722647	13745.9	<.0001
Error	3294	39312193254	11934484.898		
Corrected Total	3297	531461361197		R-Square=0.926	

Table 3.1.2 Type III tests for the ANCOVA model parameters to assess the difference between the two clones over time

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Time	1	475691896728	475691896728	39858.6	<.0001
Clone	1	1218649072.1	1218649072.1	102.11	<.0001
Time*clone	1	1096747056.3	1096747056.3	91.90	<.0001

3.1 Linear effects of climatic variables on radial growth

The objective of this section is to determine whether or not the radial growth of the two *Eucalyptus* clones can be adequately explained by the linear effects of the climatic variables. In particular, the fitted model for the radial growth of each clone as a linear function of climatic variables was

$$RadialG = \beta_0 + \beta_1 temp + \beta_2 rain + \beta_3 solar + \beta_4 wind + \beta_5 humidity + \varepsilon \quad (3.1.1)$$

where *RadialG* is the radial growth, *temp* = temperature, *rain* = rainfall, *solar* = solar radiation, *wind* = wind speed and *humidity* = relative humidity. It was very important to insure that multicollinearity among the climatic variables was minimised prior to the analysis of linear climatic effects on radial growth. The variance inflation factors presented in Table 3.1.3 were small, indicating that multicollinearity was not a problem. Recall that all the analyses are made using standardized climatic variables.

Table 3.1.3 The variance inflation factors for the linear climatic effects model (3.1.1)

Climatic variable	Variance inflation factor
Temperature	1.442
Rainfall	1.137
Solar radiation	1.163
Relative humidity	1.559
Wind speed	1.350

The ANOVA results are summarized in Table 3.1.4. The R-Square values of 0.34 and 0.32 for GC1 and GU1, respectively, indicated that only about 33% of variation in growth was explained by the linear climatic effects. The P-values (<0.0001) indicated that the models were significant. However, this test can only be valid if the regression assumptions are not violated by the data.

Table 3.1.4 ANOVA results for the linear climatic effects model (3.1.1)

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
GC1					
Model	5	78937256369	15787451274	159.59	<.0001
Error	1544	1.527367E11	98922705		
Corrected Total	1549	2.316739E11			
R^2=0.3407			Adj R^2=0.3386		
GU1					
Model	5	90505807350	18101161470	145.22	<.0001
Error	1544	1.924507E11	124644234		
Corrected Total	1549	2.829565E11			
R^2=0.3199			Adj R^2=0.3177		

The model assumptions of constant variance, independence and normality of errors were checked. The residual plots displayed in Figures 3.1.1-3.1.3 indicated that these assumptions were violated. The points in these residuals plots were not randomly scattered over time. This indicated that the assumption of constant variance and independence of errors were violated by the data. In particular, the residuals displayed some pattern over time and thus indicating that some important variable (s) are not captured in the model. Furthermore, the normality plot of residuals did not display a straight line, indicating that the normality assumption was violated by the data.

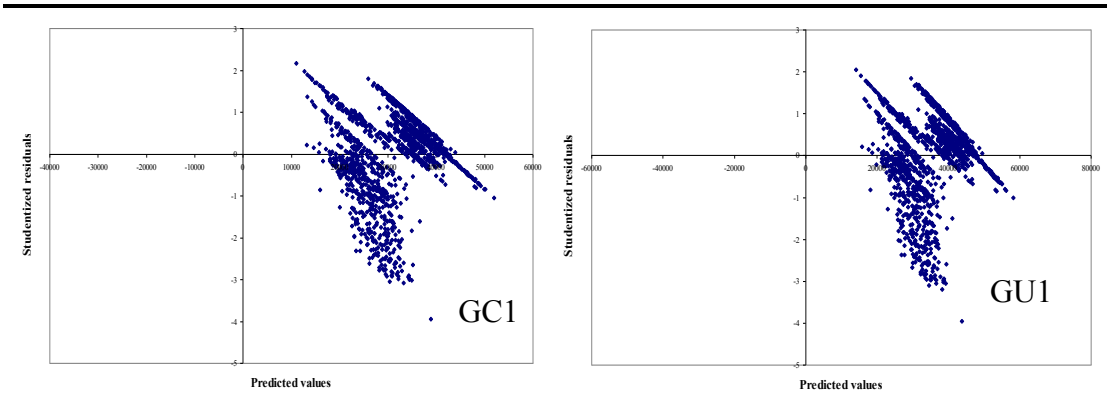

Figure 3.1.1 Constance of variance check of the errors from the linear climatic effects model

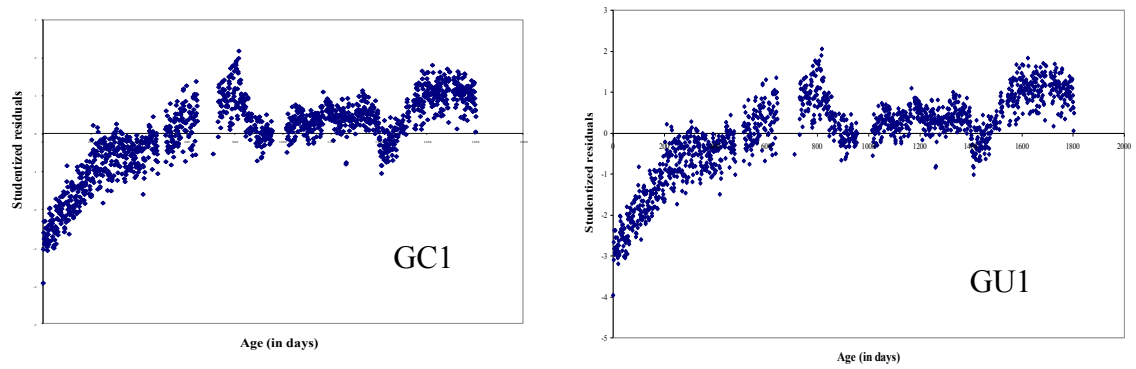

Figure 3.1.2 Independence check of errors from the linear climatic effects model

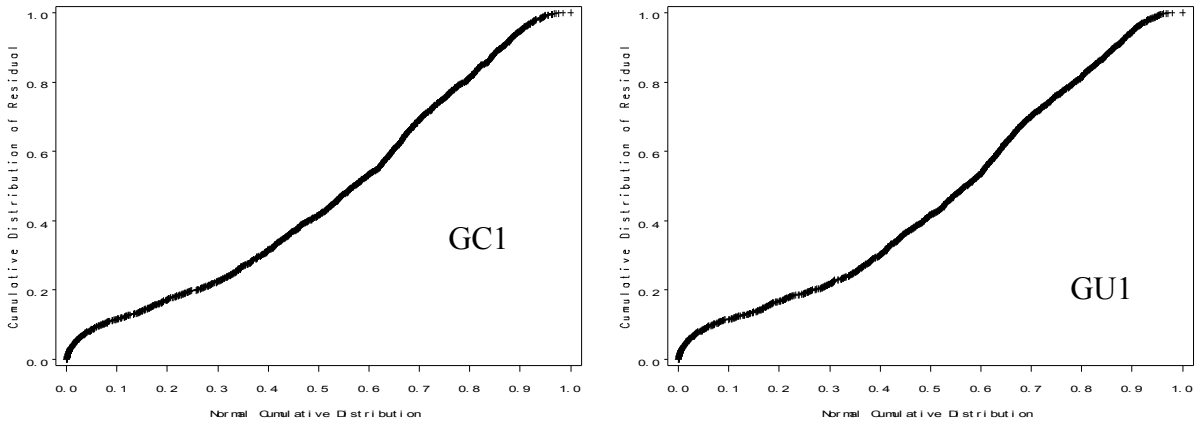

Figure 3.1.3 Normality check of the errors from the linear climatic effects model

Since the model assumptions were violated and the R-square indicated that the model was inadequate, no formal tests of hypotheses on parameter estimates were carried out on this model. The next step was to include the age effects in the model (3.1.1). The fitted linear climatic and age effects model was given by

$$radialG = \beta_0 + \beta_1 temp + \beta_2 rain + \beta_3 solar + \beta_4 wind + \beta_5 humidity + \beta_6 age + \varepsilon \qquad (3.1.2)$$

where age is in days. The results for this model are summarised in Table 3.1.5. The R-Square values of 0.93 for each clone indicate that about 93% of radial growth was explained by the linear climatic and age effects. The mean squared errors are small compared to those of the linear climatic effects model without the age effect (3.1.1). Therefore in terms of R-Square values and the mean squared errors, model (3.1.2) is an improvement of model (3.1.1).

Table 3.1.5 ANOVA results for the linear climatic effects model

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
GC1					
Model	6	2.158631E11	35977189023	3511.07	<.0001
Error	1543	15810778141	10246778		
Corrected Total	1549	2.316739E11			
R^2=0.9318			Adj R^2=0.9315		
GU1					
Model	6	2.631593E11	43859886455	3418.46	<.0001
Error	1543	19797185173	12830321		
Corrected Total	1549	2.829565E11			
R^2=0.9300			Adj R^2=0.9298		

However, the model assumptions of constant variance, independence and normality of errors were checked and it was found that these assumptions were violated by the data. In particular, the residuals displayed some quadratic pattern over time (Figure 3.1.1d). This suggested polynomial regression models for the radial growth. The following section thus investigated the polynomial effects of climatic variables and age on radial growth.

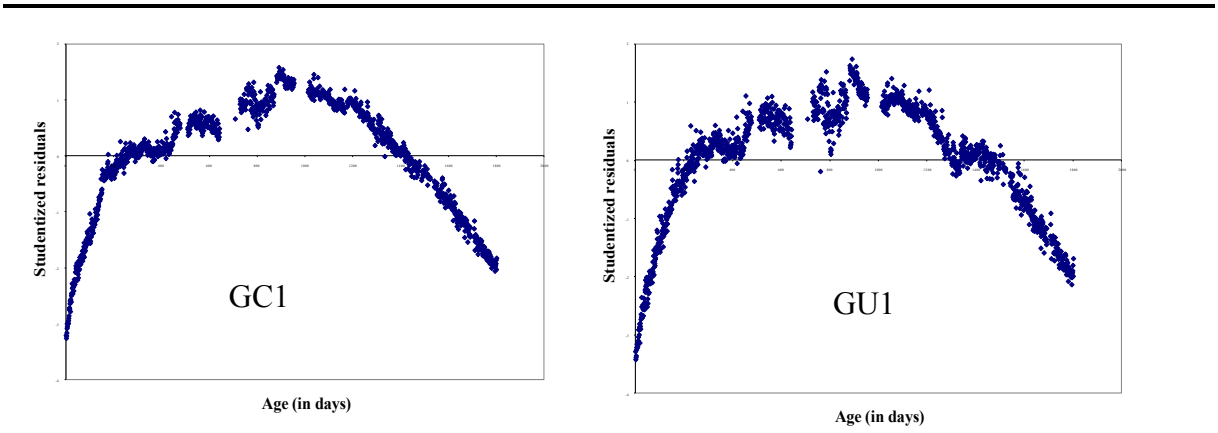

Figure 3.1.4 Independence of the errors from the linear climatic and age effects model

3.2 Polynomial climatic and age effects on radial growth

The objective of this section is to determine whether or not the radial growth of the two clones can be described adequately by the model

$$radialG = \beta_0 + funct(age) + poly3(temperature) + poly3(rain) + poly3(solar) + poly3(humidity) + poly3(wind) + \cdot$$

where $poly3(X) = \beta X + \beta X^2 + \beta X^3$ and $funct(age)$ = power age effect. We began fitting the model with all climatic variables up to order three. Several power functions (hyperbolic, power, logarithm, e.t.c.) were investigated to determine the appropriate functional relationship between age and radial growth. The highest R-squared value was obtained for $\sqrt{age}$. The age was therefore included in the model as $\sqrt{age}$. The stepwise procedure was used to select the most important variables in explaining the radial growth of the two clones. The radial growth of the two clones was affected by the linear effects of relative humidity and the quadratic effects of temperature and relative humidity. The other effects on the radial growth were different for the two clones. In particular, the significant variables with their corresponding variance inflation factors are presented in Table 3.2.1.

Table 3.2.1 The variance inflation factors for the variables in the final polynomial effects model

GC1		GU1	
Variable	**Variance inflation factor**	**Variable**	**Variance inflation factor**
$\sqrt{age}$	1.16111	$\sqrt{age}$	1.51065
relative humidity	2.77538	rainfall	1.10208
wind speed	1.25675	solar radiation	1.62325
(temperature)2	1.06422	relative humidity	1.38445
(relative humidity)2	1.57209	(temperature)2	1.07589
(relative humidity)3	2.85286	(solar radiation)2	7.18168
		(wind speed)2	1.02140
		(relative humidity)2	1.23193
		(solar radiation)3	6.85647

The results obtained from fitting the stepwise selected model are presented in Table 3.2.2. The R-Square values of 0.99 for each clone indicated that 99% of the total variation in radial growth was explained by the polynomial effects model.

Table 3.2.2 The ANOVA results for the polynomial climatic and age effects model

GC1					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	6	2.296189E11	38269812210	28734.4	<.0001
Error	1543	2055039022	1331846		
Corrected Total	1549	2.316739E11			
R^2=0.991			Adj R^2=0.991		
GU1					
Model	9	2.808076E11	31200844917	22360.0	<.0001
Error	1540	2148899650	1395389		
Corrected Total	1549	2.829565E11			
R^2=0.993			Adj R^2=0.993		

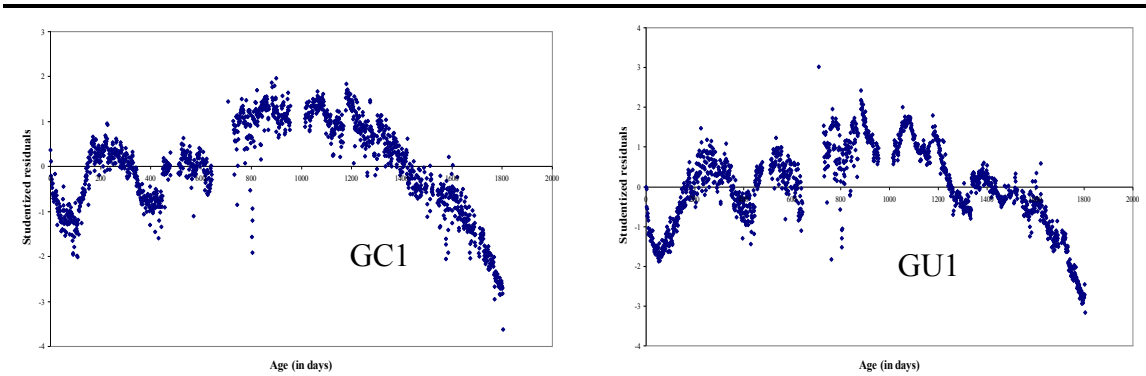

Figure 3.2.1 Independence of errors check from the polynomial effects model

The model assumptions were checked. The plot of studentized residuals versus age displayed in Figure 3.2.1 indicates that the errors in radial growth were not independent. In particular, this plot indicated that there was some seasonality that is not captured by the model. The stepwise selected model explained most of the total variation in terms of R-Square. The mean squared errors were smaller than those of the previous models. However, the plot of residuals over time indicated there is some seasonality that should be added to the model. The month

nested in year effect was added to the model to account for seasonality. In particular, the fitted refined polynomial climatic and age effects model for GC1 and GU1 was

$$radialG = \beta_0 + funct(age) + poly3(temperature) + poly3(rain) + poly3(solar) + poly3(humidity) + poly3(wind) + year(month) + \varepsilon$$

The ANOVA results for this model were presented in Table 3.2.3. The R-Square values of about 0.999 for both clones indicated that about 99.9% of the total variation in radial growth was explained by the fit. The mean squared errors were small compared to those of the previous models, and thus indicating that this is the better fit.

Table 3.2.3 ANOVA results for the final polynomial climatic and age effects model

GC1					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	62	231970830888	3741465014.3	186974	<.0001
Error	1571	31436729.908	20010.64921		
Corrected Total	1633	232002267618			
R^2=0.9999					
GU1					
Model	64	282904956175	4420389940.2	127344	<.0001
Error	1485	51547727.761	34712.274586		
Corrected Total	1549	282956503903			
R^2=0.9998					

This model was assessed for independence, constant variance and normality of residuals. The residual plots displayed in Figures 3.2.2 and 3.2.3 indicate that the assumptions of constant variance and independence of errors were not violated by the data. The normal probability plot in Figure 3.2.4 was suspicious, but the formal tests (Shapiro-Wilk, Kolmogorov-Smirnov, Anderson-Darling and Cramer-Von Mises) indicated that the normality assumption on error terms was not violated by the data. Furthermore, the Cook's index plot in Figure 3.2.5 indicates that there were no influential observations. This was the final model for predicting the effect of climate on radial growth.

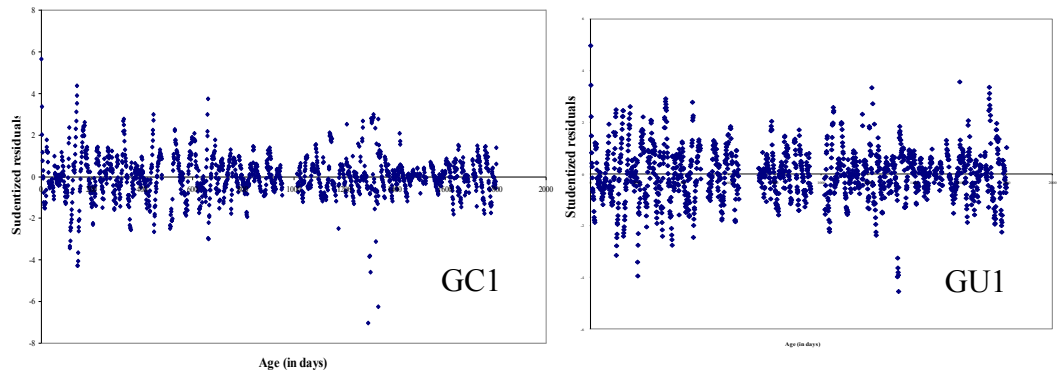

Figure 3.2.2 The plot of studentized residuals over time for the final polynomial model

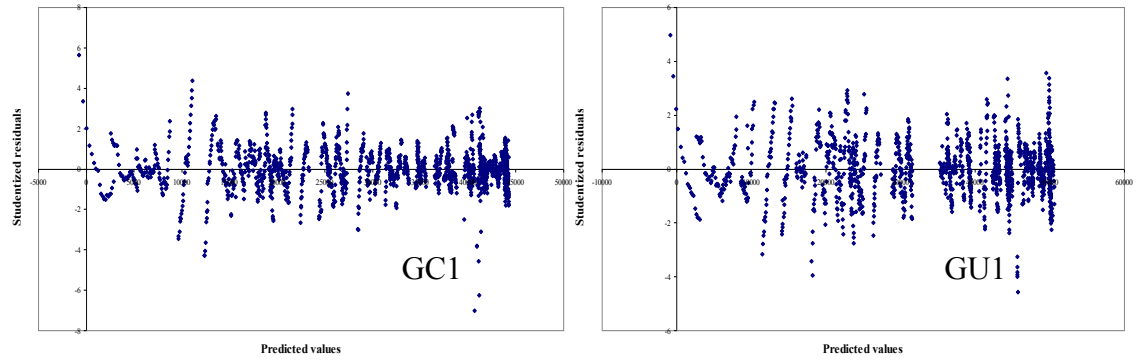

Figure 3.2.3 The plot of studentized residuals vs. predicted values for the final polynomial model

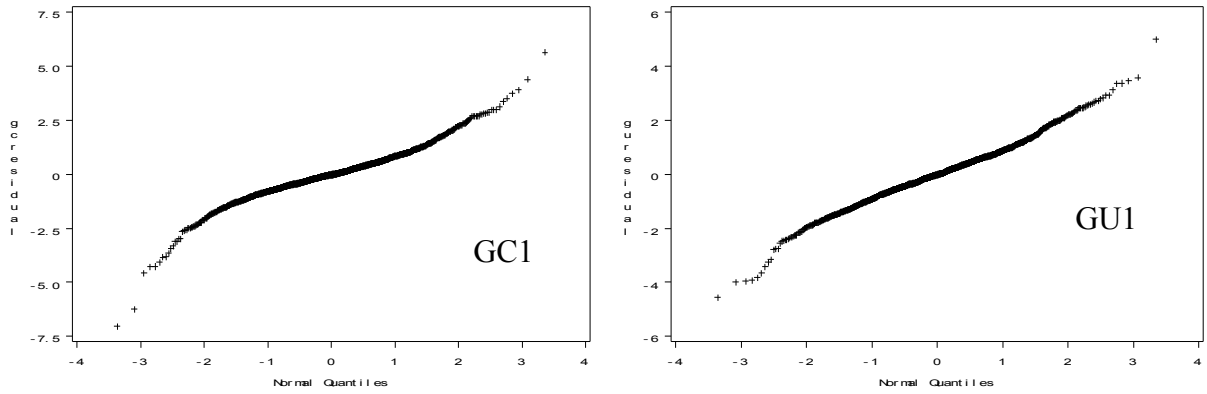

Figure 3.2.4 Normality plot for the final polynomial model

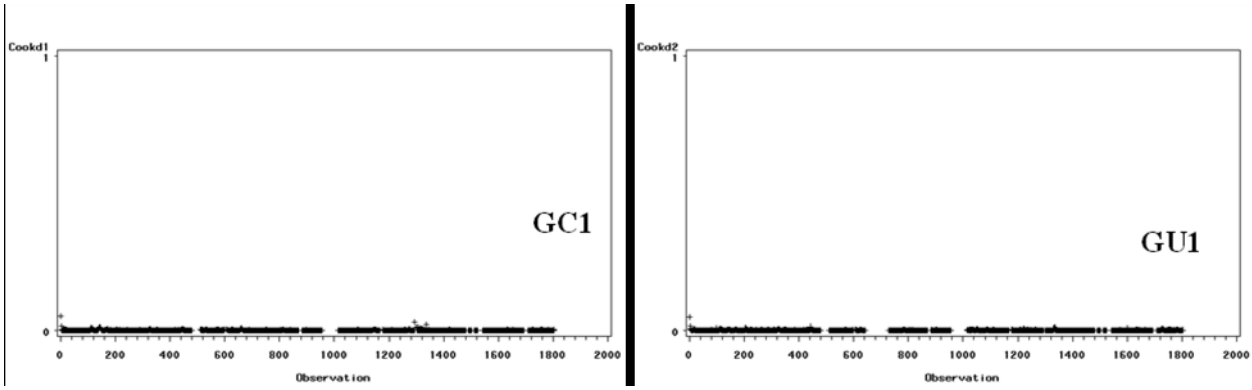

Figure 3.2.5 Cook's index plot for the final polynomial model

The Type III statistics for testing the significance of parameter estimates in Table 3.2.6 indicated that the month nested in year effects on radial growth were significant (p<0.0001). This means that the effects of some months within some years on the radial growth of the two clones were significantly different. The parameter estimates for interpretation of the continuous variables in the final model are presented in Table 3.2.7. The polynomial effects of climate on growth were interpreted with the help of graphs. The quadratic effect of temperature on the radial growth of GC1 is displayed on Figure 3.2.6. This figure indicated that the radial growth of GC1 increased when temperature was below $20.45^{\circ}C$. On the other hand, the radial growth of GC1 decreased when temperature was above $20.45^{\circ}C$.

The quadratic effects solar radiation and wind speed on the radial growth of GU1 are displayed in Figures 3.2.6 and 3.2.7, respectively. The radial growth of GU1 decreased when solar radiation was below 0.45 mJ/hr. On the other hand, the radial growth of GU1 increased when solar radiation was above 0.45 mJ/hr. Furthermore, the radial growth of both clones increases with an increase in relative humidity. In particular, for 1% increase in relative humidity, the radial growths of GC1 and GU1 increased by 23.11μ and 38.65μ, respectively. The radial growth of GU1 decreases with an increase in rainfall (see Table 3.2.7).

Table 3.2.6 Type III tests for the ANCOVA model parameters to assess the seasonality (month nested in year) effect

Parameter	DF	Type III SS	Mean Square	F Value	Pr > F	Parameter	DF	Type III SS	Mean Square	F Value	Pr > F
Year (month)	58	2157349473	37195681	1858.79	<.0001	**Year (month)**	57	2199111103	38580897	1111.45	<.0001
$\sqrt{age}$	1	114326256	114326256	5713.27	<.0001	$\sqrt{age}$	1	139618387	139618387	4022.16	<.0001
Humidity	1	243561	243561	12.17	0.0005	**humidity**	1	744348	744348	21.44	<.0001
(temperature)^2	1	189466	189466	9.47	0.0021	**(solar radiation)^2**	1	188258	188258	5.42	0.0200
						Rainfall	1	616996	616996	17.77	<.0001
						(wind speed)^2	1	194925	194925	5.62	0.0179

Table 3.2.7 Parameter estimates for the final polynomial climatic and age effects model

GC1					GU1				
Parameter	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Parameter**	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
$\sqrt{age}$	1076.05	14.24	75.59	<.0001	$\sqrt{age}$	1206.75	19.03	63.42	<.0001
humidity	23.11	6.62	3.49	0.0005	**humidity**	38.65	8.35	4.63	<.0001
(temperature)^2	-12.09	3.93	-3.08	0.0021	**(solar radiation)^2**	14.77	6.34	2.33	0.0200
					Rainfall	-24.85	5.89	-4.22	<.0001
					(wind speed)^2	26.47	11.17	2.37	0.0179

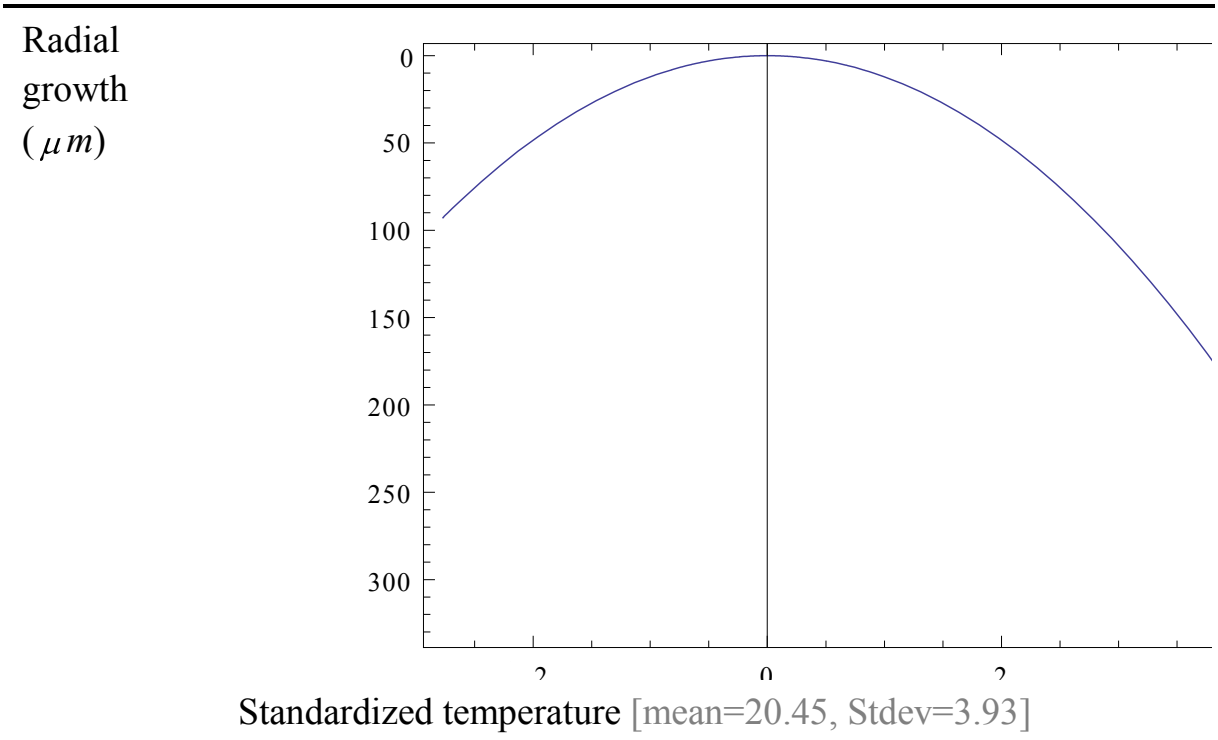

Figure 3.2.6 The quadratic effect of temperature on the radial growth of GC1

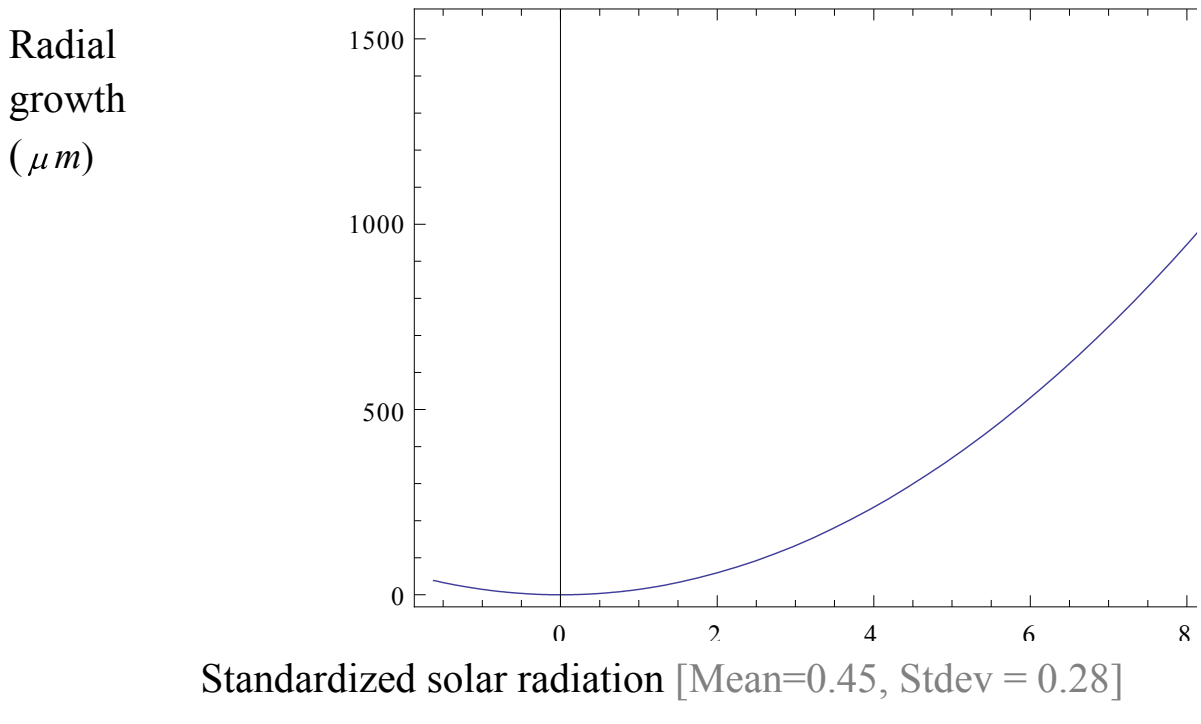

Figure 3.2.7 The quadratic effect of solar radiation on the radial growth of GU1

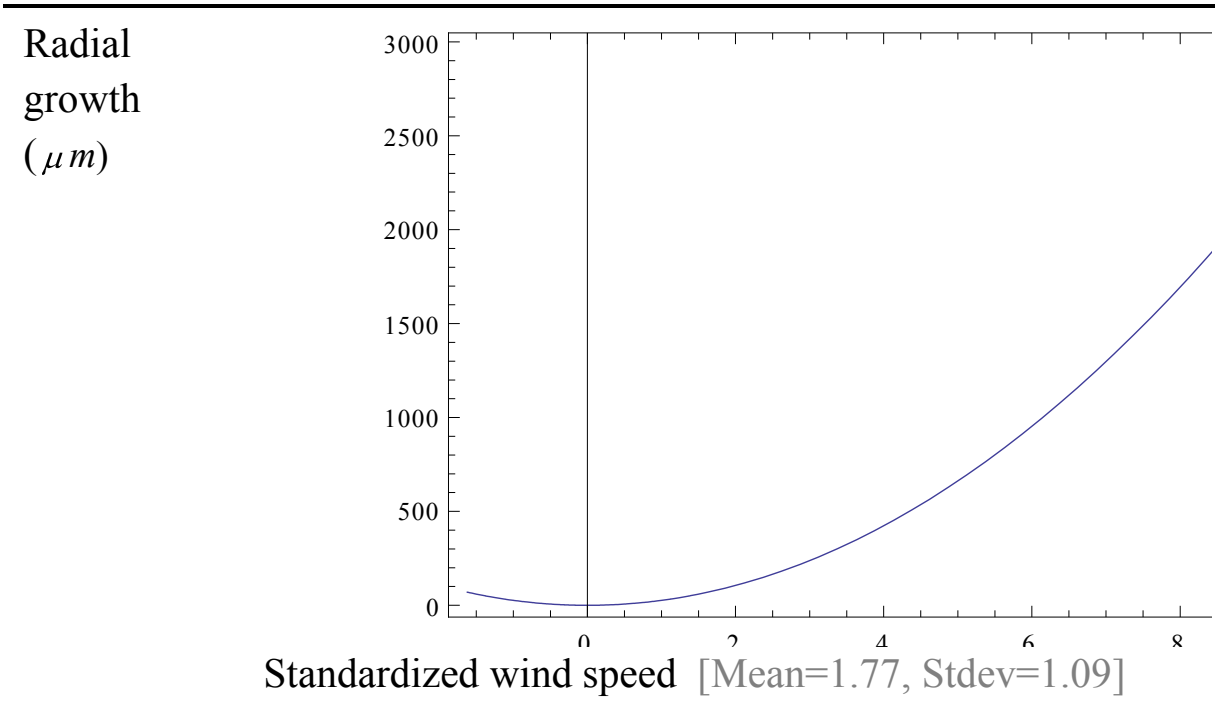

Figure 3.2.8 The quadratic effect of wind speed on the radial growth of GU1

3.3 Summary of the results of the multiple linear regression approach

The radial growth of the two clones was found to be significantly different over time. The analyses were then carried out separately for the two clones. The first model to be fitted was the linear regression of the radial growth of the two clones on the five climatic variables. The R-Square values indicated that the linear climatic effects did not adequately explain the variability in radial growth. The linear effect of age was added to the first model. The inclusion of the linear age effect in the model increased the R-Square from 0.33 to 0.93, indicating that the age effect explained a substantial amount of the total variation in radial growth. However, this model did not satisfy the regression assumptions. In particular, the plot of residuals versus age suggested polynomial transformations on the explanatory variables. Polynomials up to order three of each climatic variable were investigated. Furthermore, the square root of age appeared to be more appropriate than age in explaining the relationship between radial growth and age. In particular, the square root of time explained 93% of the total variability in radial growth of the two clones. The stepwise procedure was then used to select the best models for the two clones. The resulting R-Square values were 0.99 for each clone. The stepwise selected climatic variables were different for the two clones. The regression assumptions of the selected models were checked, and it was found that regression

assumptions were violated by the data. In particular, the plot of residuals versus age indicated the presence of seasonal effects in the radial growth of the two clones. Further investigations indicated that the seasonal effects were month nested in years effects. That is, the effects of some months within year were not the same. Thus, the final models included month nested in years effects and the R-Square values of the models were found to be 0.999. The fitted models were the best for predicting the climatic effects on radial growth of the two clones up to the age of five years. In summary the model suggested that the climatic factors to consider when planting GC1 at Kwambonambi are relative humidity and temperature. On the other hand relative humidity, solar radiation, rainfall and wind speed should be considered when growing GU1 at Kwambonambi. In particular, the radial growth of GC1 increased when temperature was below $20.45^{\circ}C$. On the other hand, the radial growth of GC1 decreased when temperature was above $20.45^{\circ}C$. The radial growth of GU1 decreased when solar radiation was below 0.45 mJ/hr. On the other hand, the radial growth of GU1 increased when solar radiation was above 0.45 mJ/hr. The radial growth increased as the wind speed increased above 1.77 m/sec. Furthermore, the radial growth of both clones increased with an increase in relative humidity. In particular, for 1% increase in relative humidity, the radial growths of GC1 and GU1 increased by 23.11 μ ! and 38.65 μ !. For a 1mm in rainfall, the radial growth of GU1 decreased by 24.85 μ !, respectively.

3.4 Results of the ARIMA modeling approach

The analyses thus far considered immediate effects of climatic variables on radial growth. However, the growth responses of trees often lag considerably behind climatic changes (Kramer and Kozlowski, 1979). For example, the effect of temperature from yesterday may be significant on radial growth today. The regression models considered thus far do not account for such effects. So, the aim of this section is to assess whether or not the delayed effects were significant on radial growth and thereby fit the model that includes the lagged effects. The ARIMA models (Section 2.4.1) were used to investigate the lagged effects of climatic variables on radial growth of the two clones. In particular, the climatic variables were lagged up to twenty four days.

It was most likely that each climatic variable series was not white noise, and hence the direct cross-correlation between the climatic variable series and the nonstationary radial growth series may give misleading indications of the relationship between the radial growth series and climatic variable series as explained in Section 2.4. This problem was solved by cross-

correlating the prewhitened series of both radial growth and the climatic variable. This means that we have to first fit an ARIMA model to the climatic variable series which was sufficient to reduce the residuals to white noise. The stationary radial growth series was then filtered with the same model, and the filtered stationary radial growth series was cross-correlated with the white noise residuals from fitting the climatic variable series. Identification of the appropriate ARIMA model and the estimation of the model for the climatic variable series are as were described in Section 2.4.

The order of differencing to remove trends in each of the climatic variable series was determined using the minimum standard deviation method described in Section 2.4. The results are displayed in Table 3.4.1.

Table 3.4.1 The standard deviations and the optimal orders of differencing of the climatic variable series.

Order of differencing	Temperature	Rainfall	Solar radiation	Wind speed	Relative humidity
$d = 0$	3.92	**0.22**	0.28	1.09	10.36
$d = 1$	**2.27**	0.26	**0.22**	**0.93**	**8.14**
$d = 2$	3.38	0.42	0.37	1.51	12.98

The smallest standard deviations were for once differenced temperature, solar radiation, wind speed and relative humidity series. Hence the optimal order of differencing for these climatic variables is $d = 1$. The rainfall series had no trend. The observed and once differenced series for these climatic variables are displayed in Figures A.1 - A.4 in the appendices. Figures for temperature, solar radiation and relative humidity indicate that there is some seasonality in the climatic variables' series. Differencing to remove seasonality was done after differencing to remove trend. In particular, the seasonality in the climatic variables was approximately a year (371 days).

The order of differencing to remove trends in the radial growth of the two clones was also determined using the minimum standard deviation criterion. The results are displayed in Table 3.4.2. The smallest standard deviations were for the once differenced radial growths of the two clones. The once differenced radial growth series are displayed in Figures A.5 and A.6 in the appendices. These figures indicated that a few observations were inconsistent with the rest of the data. The analysis was done with and without these observations and the results did not change. The final analyses were therefore carried out with these observations.

Table 3.4.2 The standard deviations and the optimal orders of differencing of the radial growth series.

Order of differencing	GC1	GU1
$d = 0$	11879.27	13114.11
$d = 1$	**81.63**	**76.08**
$d = 2$	121.38	100.19

3.4.1 The effects of lagged temperature on radial growth

In this section the effects of the lagged temperature on radial growth were investigated. The autocorrelation and partial autocorrelation plots in Figures 3.4.1a and b suggest that the detrended and deseasonalized temperature series could be explained by an MA (3) model. In particular, the MA (3) model for the temperature series was estimated to be

$$(1-3)(1-3^{371})Y_t = 1-).3718\,B-).4393\,B^2-).1414\,B^3)Z_t \qquad (3.4.1)$$

where B is the Backward shift operator defined in Section 2.4, Y_t is the original temperature series and Z_t is white noise.

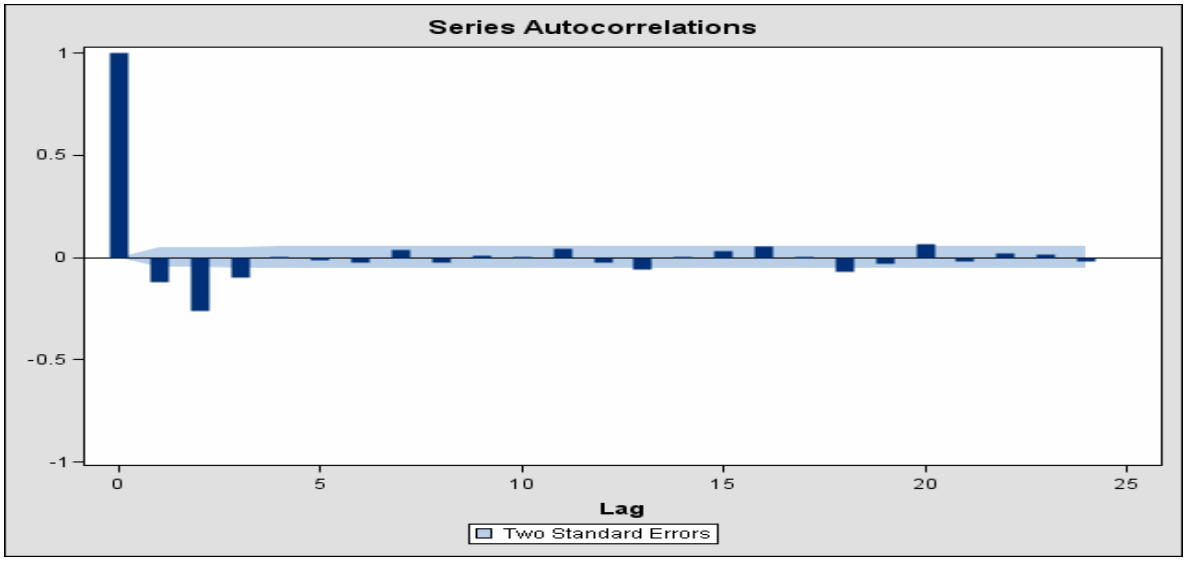

Figure 3.4.1a Autocorrelations of the once differenced temperature series

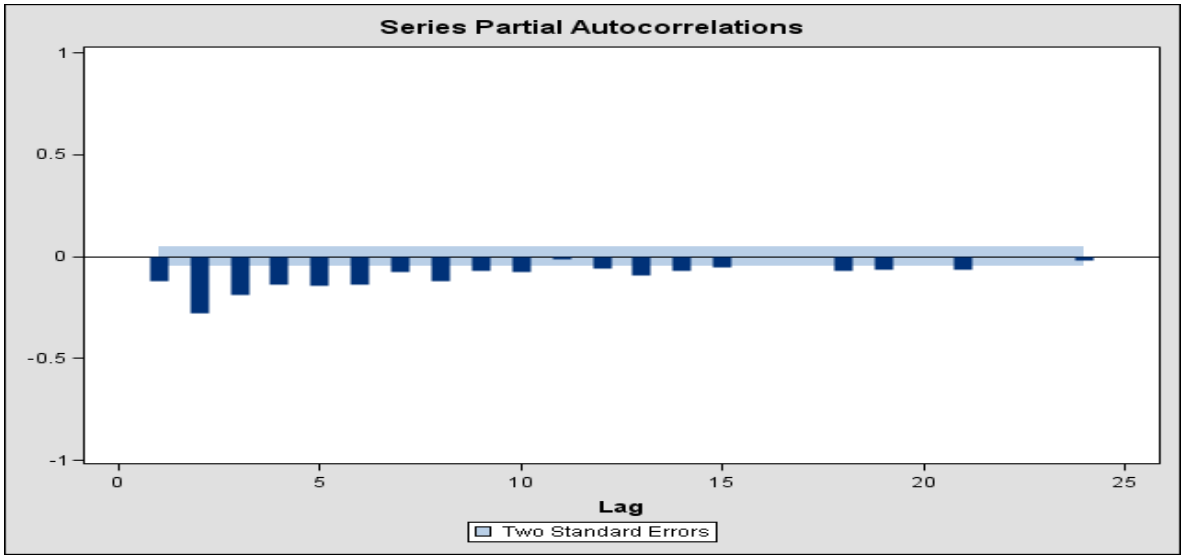

Figure 3.4.1b Partial autocorrelations of the once differenced temperature series

The residual check for white noise (Table 3.4.3a) indicated that the MA (3) provided a good fit to the temperature series since the *p* values are large. Furthermore, all the MA (3) parameters were significant (except for the intercept) as shown in Table 3.4.3b.

Table 3.4.3a Autocorrelation check for white noise from fitting the MA (3) model to the once differenced temperature series

To Lag	Chi-Square	DF	Pr > ChiSq
6	3.69	3	0.2972
12	7.60	9	0.5750
18	13.79	15	0.5413
24	21.16	21	0.4489
30	23.28	27	0.6697
36	25.97	33	0.8031
42	29.61	39	0.8613
48	31.98	45	0.9279

Table 3.4.3b Maximum Likelihood estimation of the MA (3) model for the temperature series

Parameter	Estimate	Standard Error	Approx Pr > |t|
Constant	0.00262	0.00351	0.4562
MA1,1	0.37183	0.02738	<.0001
MA1,2	0.43931	0.02668	<.0001
MA1,3	0.14140	0.02750	<.0001

After filtering the stationary series of both temperature and radial growth with model (3.4.1), the crosscorrelations between the filtered series were calculated. The *p* values presented in Table 3.4.3c were large, indicating that there was no linear relationship between the radial growth series and the lagged temperature series. The plots of crosscorrelations between radial growths of GC1 and GU1 and temperature are displayed in Figures A7 and A8 in the appendices, respectively. These figures have no significant spikes, and thus confirming that there was no linear relationship between the radial growth series and the lagged temperature series.

Table 3.4.3c Crosscorrelation check between lagged temperature and radial growth series

	GC1			GU1		
To Lag	Chi-Square	DF	Pr > ChiSq	Chi-Square	DF	Pr > ChiSq
5	2.92	6	0.8190	3.02	6	0.8064
11	4.44	12	0.9741	5.29	12	0.9477
17	7.42	18	0.9861	9.22	18	0.9545
23	10.34	24	0.9930	12.27	24	0.9767

3.4.2 The effects of lagged rainfall on radial growth

In this section the effects of the lagged rainfall on radial growth were investigated. The rainfall series had nine extreme observations which made it difficult to stationarize the rainfall series. Removing these observations was the remedial measure undertaken to solve this problem. The autocorrelation and partial autocorrelation plots in Figures 3.4.2a and b

suggested that the deseasonalized rainfall series could be explained by an MA (1) model. In particular, the MA (1) model for the rainfall series was estimated to be

$$(1- 3^{371})Y_t = \ + 1.159\, Z_t \qquad (3.4.2)$$

where Y_t is the original rainfall series.

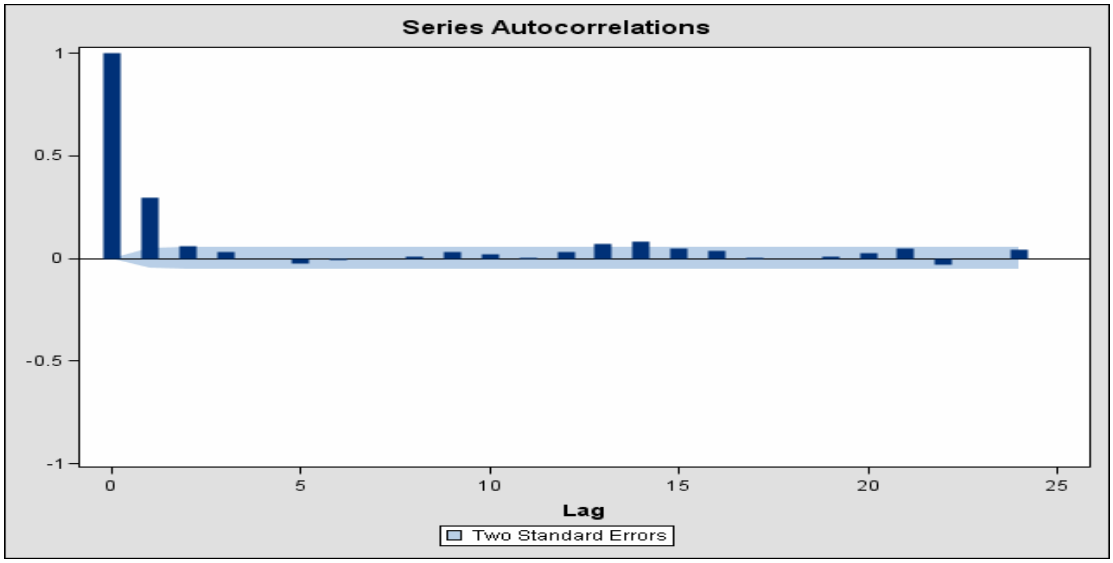

Figure 3.4.2a Autocorrelations of the undifferenced rainfall series

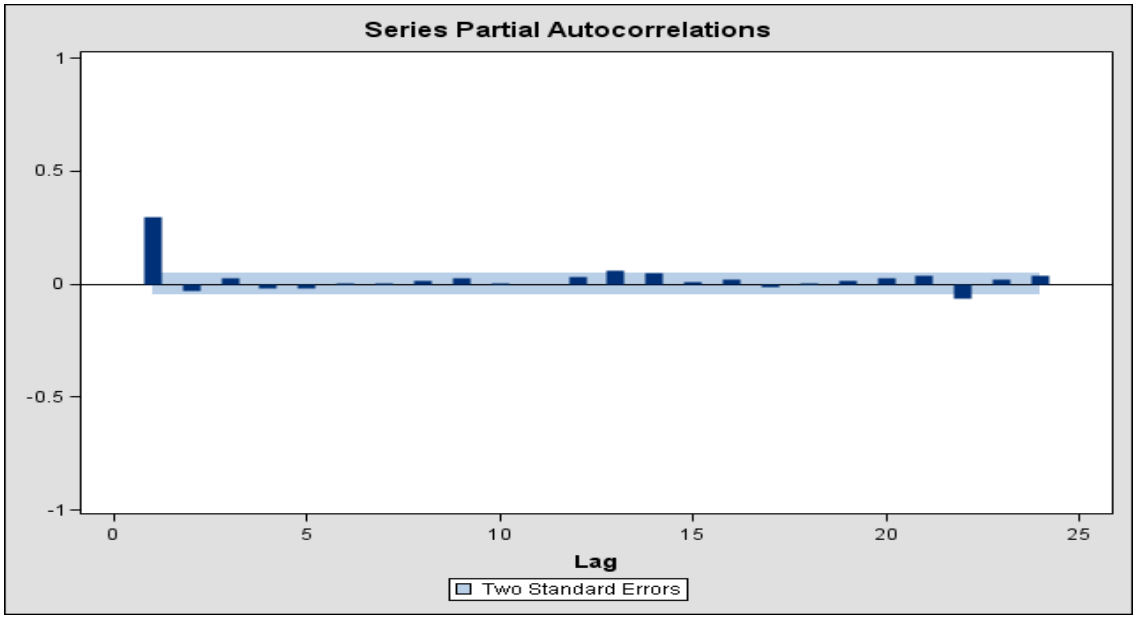

Figure 3.4.2b Partial autocorrelations of the undifferenced rainfall series

The residual check for white noise (Table 3.4.4a) indicated that the MA (1) provided a reasonable good fit to the rainfall series since the *p* values are large (>0.05). Furthermore, all the MA (1) parameters were significant (except for the intercept) as shown in Table 3.4.4b.

Table 3.4.4a Autocorrelation check for white noise from fitting the MA (1) model to the undifferenced rainfall series

To Lag	Chi-Square	DF	Pr > ChiSq
6	5.99	5	0.3071
12	12.58	11	0.3217
18	27.27	17	0.0542
24	31.65	23	0.1077
30	40.31	29	0.0790
36	50.82	35	0.0409
42	56.39	41	0.0552
48	62.25	47	0.0673

Table 3.4.4b Maximum Likelihood estimation of the MA (1) model for the undifferenced rainfall series

Parameter	Estimate	Standard Error	Approx Pr > |t|
Constant	-0.0.00831	0.01121	0.4583
MA1,1	-0.15855	0.02719	<.0001

After filtering both the stationary radial growth and the rainfall series with model (3.4.2), the crosscorrelations between the filtered series were calculated. The crosscorrelation check between the rainfall and radial growth series is presented in Table 3.4.4c. The *p* values were large for GC1, indicating that there was no linear relationship between radial growth of GC1 and lagged rainfall. Furthermore, the crosscorrelations plot in Figure A9 had no spikes, confirming that there was no linear relationship between the rainfall series and the GC1 radial growth series. On the other hand, the *p* values for GU1 were small, indicating that there was a significant linear relationship between the radial growth of GU1 and the lagged rainfall. In

particular, Figure A10 had spikes at lags 0 and 1 with positive correlations. However, the correlations were very small (0.11 and 0.13) indicating that the positive linear relationship between lagged GU1 and lagged rainfall was weak.

Table 3.4.4c Crosscorrelation check between lagged rainfall and radial growth series

	GC1			GU1		
To Lag	**Chi-Square**	**DF**	**Pr > ChiSq**	**Chi-Square**	**DF**	**Pr > ChiSq**
5	11.56	6	0.0725	44.12	6	<.0001
11	19.28	12	0.0820	53.97	12	<.0001
17	23.43	18	0.1745	60.32	18	<.0001
23	25.72	24	0.3673	67.51	24	<.0001

3.4.3 The effects of lagged solar radiation on radial growth

In this section the effects of the lagged solar radiation on radial growth were investigated. The autocorrelation and partial autocorrelation plots in Figures 3.4.3a and b suggested that the detrended and deseasonalized solar radiation series could be explained by an MA (2) model. In particular, the MA (2) model for the solar radiation series was estimated to be

$$(1-3)(1-3^{371})Y_t = 1-).776B-).134B^2)Z_t \qquad (3.4.3)$$

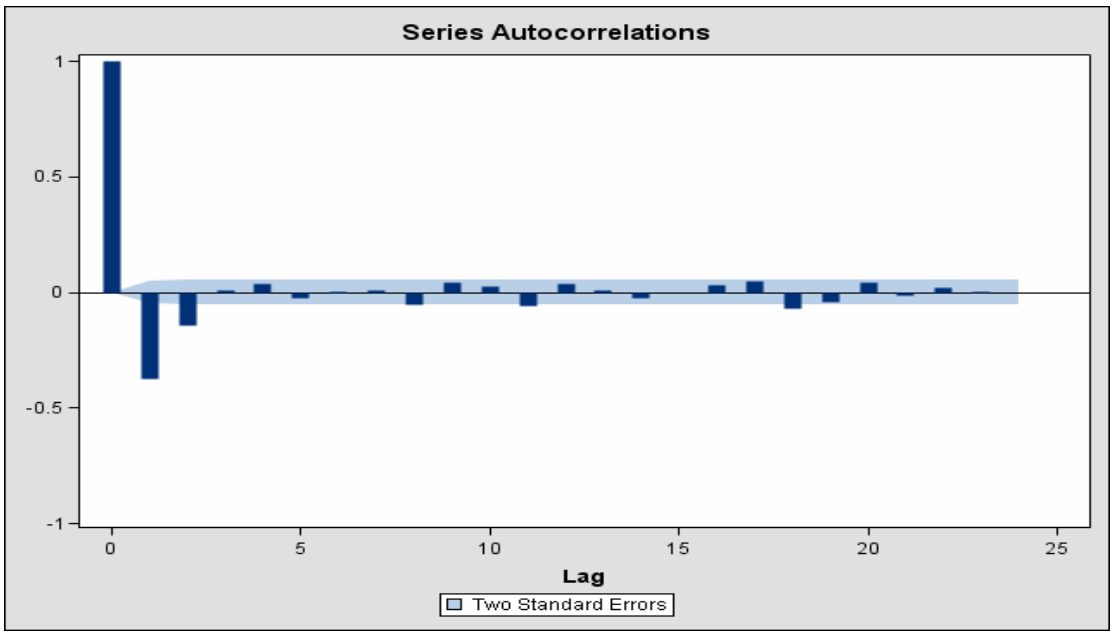

Figure 3.4.3a Autocorrelations of the solar radiation series

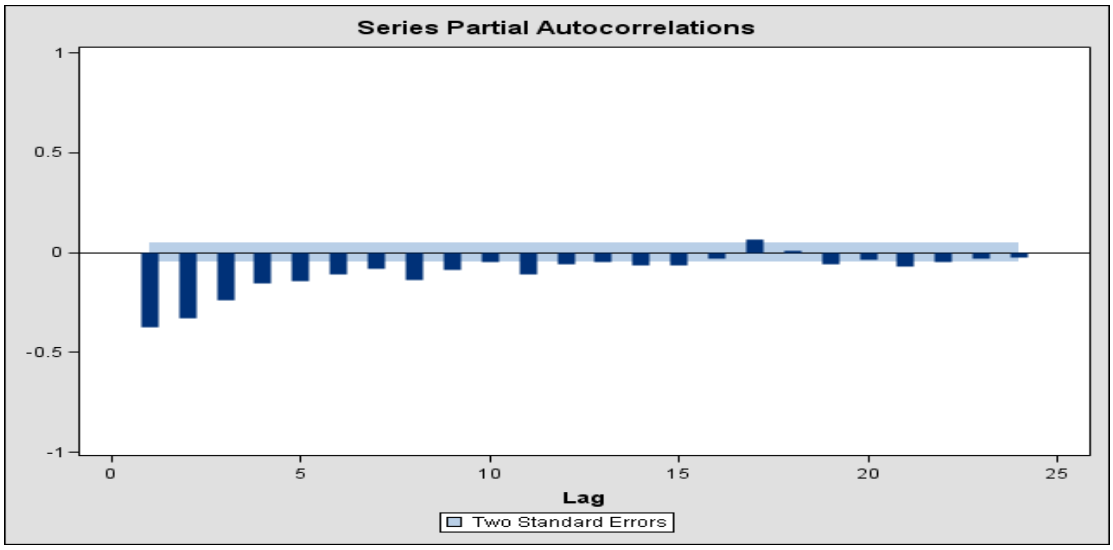

Figure 3.4.3b Partial autocorrelations of the solar radiation series

The residual check for white noise (Table 3.4.5a) indicated that the MA (2) provided a reasonable good fit to the solar radiation series because the *p* values are large (>0.05). Furthermore, all the MA(2) parameters were significant (except for the intercept) as shown in Table 3.4.5b.

Table 3.4.5a Autocorrelation check for white noise from fitting the MA (2) model to the solar radiation series

To Lag	Chi-Square	DF	Pr > ChiSq
6	2.78	4	0.5954
12	6.89	10	0.7353
18	16.39	16	0.4261
24	24.12	22	0.3411
30	32.49	28	0.2550
36	35.45	34	0.3996
42	41.36	40	0.4112
48	46.04	46	0.4706

Table 3.4.5b Maximum Likelihood estimation of the MA (2) model for the solar radiation series

Parameter	Estimate	Standard Error	Approx Pr > |t|
Constant	0.00044	0.00095	0.6460
MA1,1	0.77596	0.02946	<.0001
MA1,2	0.13353	0.02953	<.0001

After filtering both the stationary radial growth and the solar radiation series with model (3.4.3), the crosscorrelations between the filtered series were calculated. The crosscorrelation check between these series is presented in Table 3.4.5c. The *p* values were small (<0.05) for both clones, indicating that there was a linear relationship between the lagged solar radiation and the radial growths of the two clones. In particular, Figures A11 and A12 indicated that the linear relationship between lagged solar radiation and the radial growths of the two clones was negative. Furthermore, there were no spikes in Figure A11 indicating that the linear relationship between the lagged solar radiation and GC1 radial growth was weak. On the other hand, Figure A12 had significant spikes at lags 1 and 2. The correlations at lags 1 and 2 were -0.08041 and -0.08046, respectively. These correlations indicate that the radial growth of GU1 was negatively correlated to the previous two days solar radiation.

Table 3.4.5c Crosscorrelation check between lagged solar radiation and radial growth series

	GC1			GU1		
To Lag	Chi-Square	DF	Pr > ChiSq	Chi-Square	DF	Pr > ChiSq
5	24.17	6	0.0005	37.28	6	<.0001
11	36.43	12	0.0003	45.43	12	<.0001
17	47.04	18	0.0002	54.40	18	<.0001
23	49.57	24	0.0016	57.92	24	0.0001

3.4.4 The effects of lagged relative humidity on radial growth

In this section the effects of the lagged relative humidity on radial growth were investigated. The autocorrelation and partial autocorrelation plots in Figures 3.4.4a and b suggested that

the detrended and deseasonalized relative humidity series could be explained by an MA (3) model. In particular, the MA (3) model for the relative humidity series was estimated to be

$$(1-3)(1-3^{371})Y_t = 1-).460B-).276B^2-).111B^3)Z_t \qquad (3.4.4)$$

where Y_t is the original relative humidity.

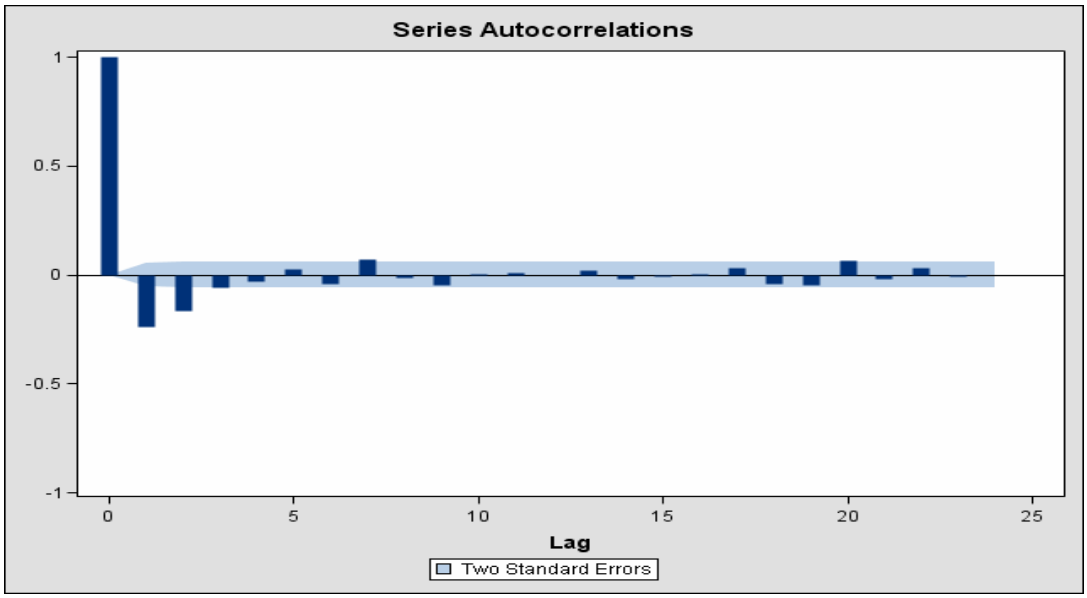

Figure 3.4.4a Autocorrelations of the relative humidity series

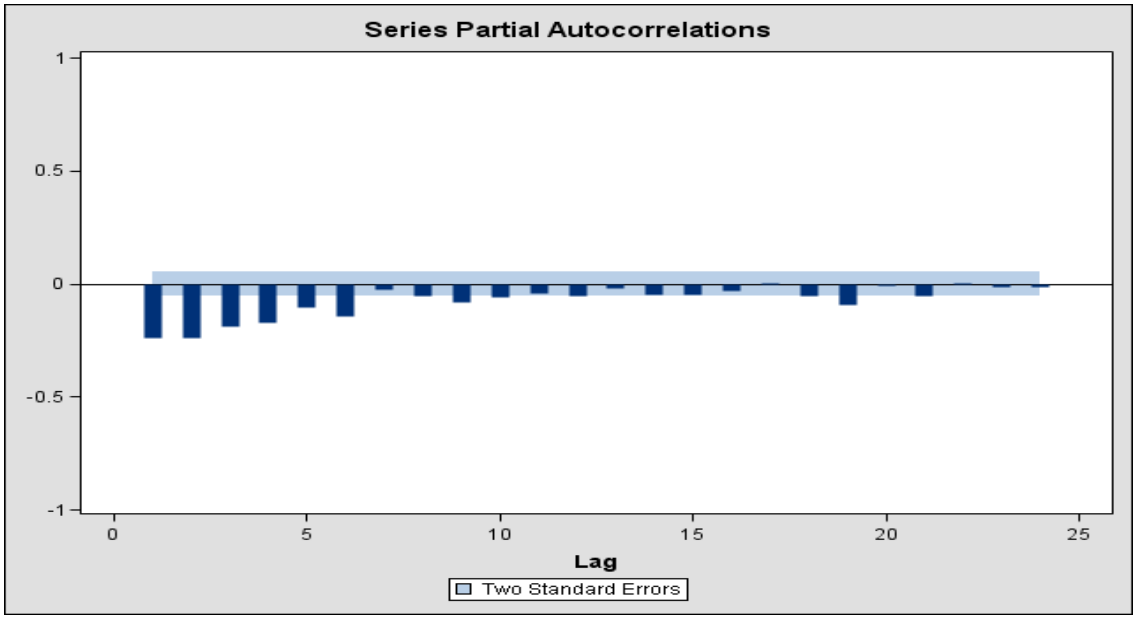

Figure 3.4.4b Partial autocorrelations of the relative humidity series

The residual check for white noise (Table 3.4.6a) indicated that the MA (3) provided a reasonable good fit to the relative humidity series because the *p* values were large. Furthermore, all the MA(3) parameters were significant (except for the intercept) as shown in Table 3.4.6b.

Table 3.4.6a Autocorrelation check for white noise from fitting the MA (3) model to the relative humidity series

To Lag	Chi-Square	DF	Pr > ChiSq
6	3.57	3	0.3115
12	10.99	9	0.2761
18	15.67	15	0.4045
24	21.98	21	0.4008
30	27.31	27	0.4470
36	36.50	33	0.3092
42	38.36	39	0.4988
48	41.77	45	0.6095

Table 3.4.6b Maximum Likelihood estimation of the MA (3) model for the relative humidity series

Parameter	Estimate	Standard Error	Approx Pr > |t|
Constant	-0.01455	0.04036	0.7185
MA1,1	0.46005	0.02888	<.0001
MA1,2	0.27569	0.03092	<.0001
MA1,3	0.11139	0.02901	0.0001

After filtering both the stationary radial growth and the relative humidity series with model (3.4.4), the cross correlations between the filtered series were calculated. The results in Table 3.4.6c for GC1 and GU1 indicated that there was a linear relationship between radial growths of both clones and lagged relative humidity. The radial growth increased with an increase in

lagged relative humidity (see Figures A13 and A14). In particular, Figure A13 had spikes at lags 15 and 24, indicating that relative humidity from the previous 15 and 24 days were positively correlated with today's GC1 radial growth. On the other hand, Figure A14 indicated that relative humidity from today, 15 days and 18 days are positively correlated with today's GU1 radial growth. However, the largest correlations were 0.09 for both clones, and thus indicating that the linear relationship between lagged relative humidity and radial growth of the two clones was weak.

Table 3.4.6c Crosscorrelation check between lagged relative humidity and radial growth series

	GC1			GU1		
To Lag	**Chi-Square**	**DF**	**Pr > ChiSq**	**Chi-Square**	**DF**	**Pr > ChiSq**
5	2.80	6	0.8333	18.89	6	0.0043
11	4.29	12	0.9776	34.27	12	0.0006
17	26.42	18	0.0905	62.10	18	<.0001
23	40.51	24	0.0188	94.71	24	<.0001

3.4.5 The effects of lagged wind speed on radial growth

In this section the effects of the lagged wind speed on radial growth were investigated. The wind speed series also had extreme observations which were truncated to achieve stationarity. In particular, wind speed less than four was used to investigate the appropriate ARIMA model to describe the wind speed series. The autocorrelation and partial autocorrelation plots in Figures 3.4.5a and b suggested that the detrended and deseasonalized wind speed series could be explained by an MA (2) model. In particular, the MA (2) model for the wind speed series was estimated to be

$$(1-3)(1-3^{371})Y_t = 1-).602B-).291B^2)Z_t \qquad (3.4.5)$$

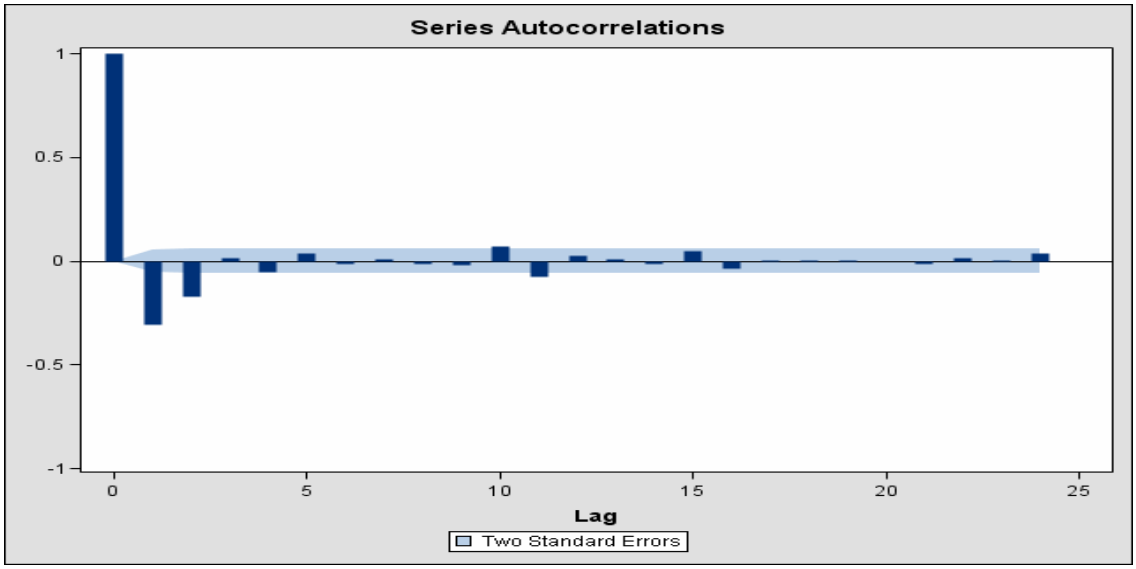

Figure 3.4.5a Autocorrelations of the wind speed series

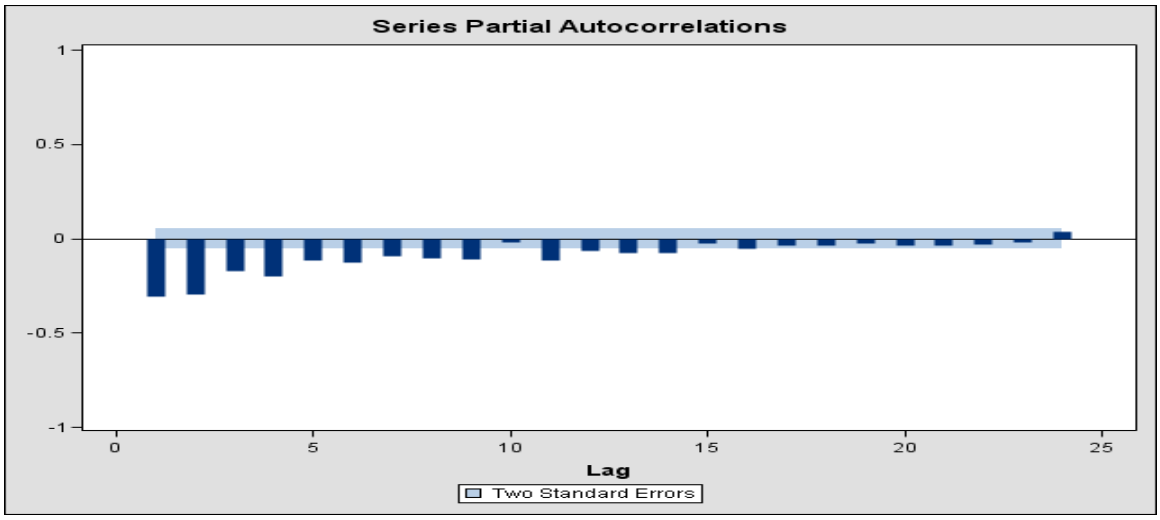

Figure 3.4.5b Partial autocorrelations of the wind speed series

The residual check for white noise (Table 3.4.7a) indicated that the MA (2) adequately explains the wind speed series. Furthermore, all the MA (2) parameters were significant (except for the intercept) as shown in Table 3.4.7b.

Table 3.4.7a Autocorrelation check for white noise from fitting the MA (2) model to the wind speed series

To Lag	Chi-Square	DF	Pr > ChiSq
6	5.10	4	0.2774
12	13.40	10	0.2022
18	16.52	16	0.4173
24	20.91	22	0.5265
30	31.37	28	0.3007
36	35.13	34	0.4145
42	39.22	40	0.5052
48	46.90	46	0.4353

Table 3.4.7b Maximum Likelihood estimation of the MA(2) model for the wind speed series

Parameter	Estimate	Standard Error	Approx Pr > |t|
Constant	0.00071	0.00552	0.8983
MA1,1	0.60163	0.02729	<.0001
MA1,2	0.29133	0.02730	<.0001

After filtering both the stationary radial growth and the wind speed series with model (3.4.5), the cross correlations between the filtered series were calculated. The results in Table 3.5.7c for GC1 and GU1 indicated that there was some linear relationship between lagged wind speed and radial growth. In particular, the lagged wind speed was associated with both increase and decrease in radial growth of the two clones (see Figures A15 and A16). However, the largest correlations were -0.076 and -0.11 for GC1 and GU1, respectively. This indicated that the linear relationship between lagged wind speed and radial growths of the two clones was weak.

Table 3.3.7c Crosscorrelation check between wind speed and radial growth series

	GC1			GU1		
To Lag	**Chi-Square**	**DF**	**Pr > ChiSq**	**Chi-Square**	**DF**	**Pr > ChiSq**
5	7.86	6	0.2485	29.49	6	<.0001
11	10.61	12	0.5624	36.63	12	0.0003
17	26.68	18	0.0851	74.98	18	<.0001
23	38.81	24	0.0286	89.70	24	<.0001

3.4.6 Summary on ARIMA modelling approach

This section investigated the lagged effects of climatic variables on the radial growth. The climatic and radial growth series were prewhitened to obtain the most appropriate relationships between the climatic variable series and the radial growth series. The seasonality effect in the series of the climatic variables was taken into consideration during modeling. Rainfall and wind speed had extreme values which were removed prior to identifying the appropriate ARIMA models to describe the series. The results indicated that there was no linear relationship between lagged temperature and radial growth. There was no linear relationship between lagged rainfall and the GC1 radial growth. On the other hand, the radial growth of GU1 increases if high rainfall was received in the previous six days. The radial growth of the two clones increase with an increased in lagged relative humidity. The lagged wind speed is associated with both increase and decrease in radial growth of the two clones. The crosscorrelation plots between the climatic variables' and radial growths series indicated that the linear relation was weak. Therefore, no further modeling was done on lagged effects.

3.5 The piecewise regression model

Analyses carried thus far were investigating the overall effects of climatic parameters on radial growth. These analyses indicated that the impacts of climatic variables on radial growth were minimal. This section explores a different way to assess the impact of climatic variables on radial growth. The observed radial growths of the two clones indicated that the growth rates experienced substantial changes over time (see Figure 2.1.2). These changes were referred to as breakpoints. If the breakpoints could be estimated, the climatic variables could be revisited to assess whether or not they were related to the breakpoints. The breakpoints were estimated using piecewise linear regression. Models with different number

of breakpoints were assessed, and the model with four breakpoints was found to be the best fit. The four breakpoints model was given by

$$RadialG = \begin{cases} b_0 + b_1 age + \varepsilon & if \quad age \le \alpha \\ b_0 + b_1 age + b_2 (age - \alpha \quad + \varepsilon & if \quad \alpha \quad < age \le \alpha \\ b_0 + b_1 age + b_2 (age - \alpha \quad + b_3 (age - \alpha \quad + \varepsilon & if \quad \alpha \quad < age \le \alpha \\ b_0 + b_1 age + b_2 (age - \alpha \quad + b_3 (age - \alpha \quad + b_4 (age - \alpha \quad + \varepsilon & if \quad \alpha \quad < age \le \alpha \\ b_0 + b_1 age + b_2 (age - \alpha \quad + b_3 (age - \alpha \quad + b_4 (age - \alpha \quad + b_5 (age - \alpha \quad + \varepsilon & if \quad age > \alpha \end{cases}$$

where b_0 was the average stem radius of the clone at plantation day (14.84 for GC1 and 3.87 for GU1); α , α , α and α are the breakpoint ages of the clones; ε *s* are the random errors with mean zero and constant cariance. The parameter b_1 is the growth rate before the first breakpoint. The rest of the b_i'*s* can be interpreted as the difference between the growth rates. Starting parameters $(b_0, b_1, b_2, b_3, b_4, b_5, \alpha , \alpha , \alpha , \alpha)$ are needed by the PROC NLIN procedure in SAS to begin fitting the model.

We began by examining the graph of the observed values for GC1 and GU1 versus age, and visually found that the breakpoints occur somewhere around 100, 500, 900 and 1300 days. The two models (one for GC1 and one for GU1), were fitted using PROC NLIN and these graphically found breakpoints and the above values of b_0 as the initial values. The rest of the parameters $(b_1, b_2, b_3, b_4, b_5)$ were assigned zero values as their starting points. The following are the results obtained.

The ANOVA results presented in Table 3.5.1 indicate that this model was significant (p<0.0001). This implies that this model explained a significant amount of variation in the data. The R-Square of 0.9999 for each clone indicates that about 99.99% of variation in radial growth was explained by age with four breakpoints or five distinct growth phases. The observed vs. fitted values displayed in Figures 3.5.1 and 3.5.2, for GC1 and GU1, respectively, indicate that growth was adequately described by the four breakpoints piecewise model.

Table 3.5.1 ANOVA results from fitting the four-breakpoint piecewise models

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
GC1					
Model	9	2.324E11	2.582E10	314014	<.0001
Error	1639	1.3476E8	82222.6		
Corrected Total	1648	2.325E11			
R^2=0.9999					
GU1					
Model	9	2.833E11	3.148E10	167603	<.0001
Error	1639	3.0781E8	187803		
Corrected Total	1648	2.836E11			
R^2=0.9999					

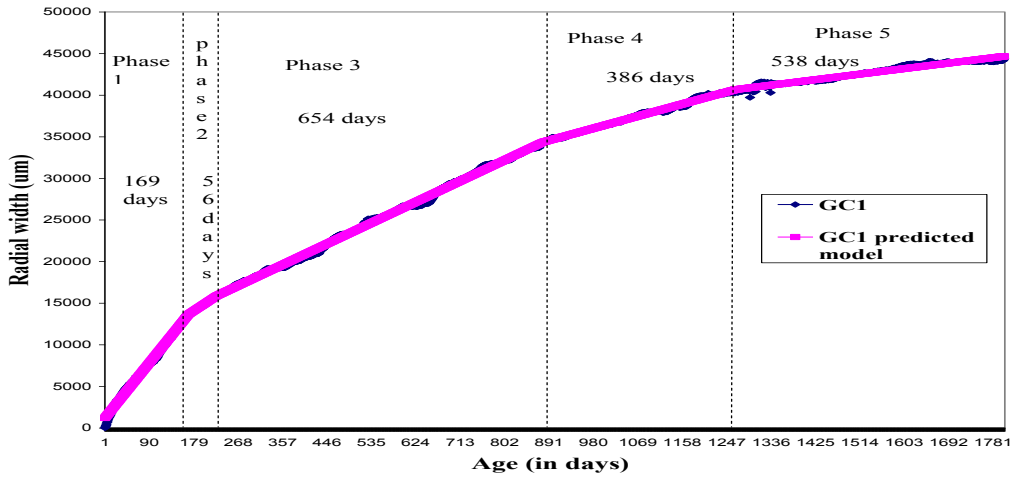

Figure 3.5.1 The observed and fitted curves for GC1 piecewise model by growth rate phases

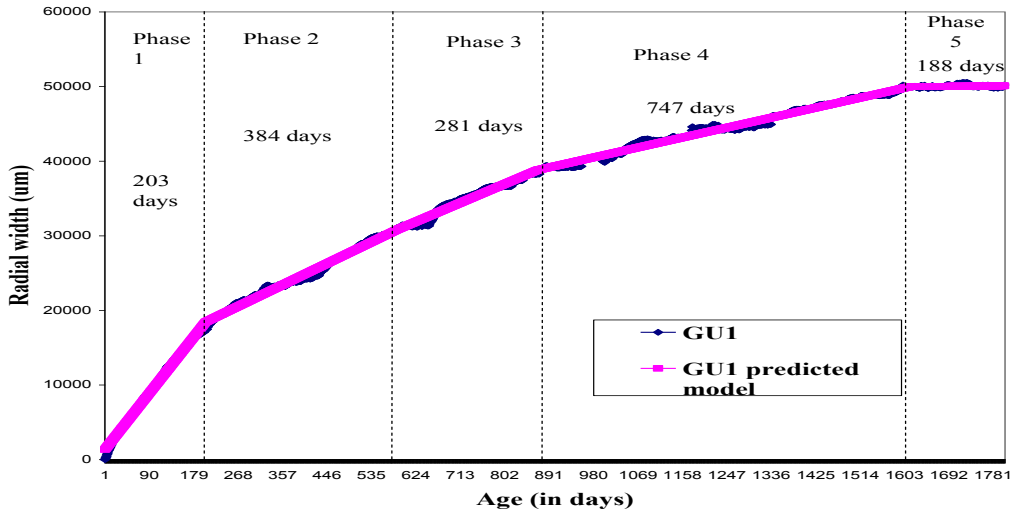

Figure 3.5.2 The observed and fitted curves for GU1 piecewise model by growth rate phases

The parameter estimates for the four breakpoints piecewise models are presented in Table 3.5.2 The parameter estimates indicated that the breakpoints occurred around days: 169, 225, 879, and 1265 for GC1; and around: 203, 587, 868 and 1615 for GU1. The table also provides the estimates of growth rate differences. It was observed that the growth rate increased sharply in piecewise Phase 1 (before the first breakpoint). The growth rate then started diminishing in subsequent piecewise Phases. In particular, the growth rate in piecewise Phase *i* was higher than that in piecewise Phase *i*+1. The daily growth rate for GC1 was 74.33 μ ! in piecewise Phase 1. It then decreased to 39.51 μ ! in piecewise Phase 2, 28.25 μ ! in piecewise Phase 3, 16.37 μ ! in piecewise Phase 4 and 7.43 μ ! in piecewise Phase 5. On the other hand, the daily growth rate for GU1 was 84.73 μ ! in piecewise Phase 1. This growth rate then started diminishing to 32.22 μ ! in Phase 2, 28.44 μ ! in piecewise Phase 3, 14.96 and 0.51 μ !in Phase 5.

Table 3.5.2 Parameter estimates for the four- breakpoint piecewise models

Parameter	GC1				GU1			
	Estimate	Approx Std Error	Approximate 95% Confidence Limits		Estimate	Approx Std Error	Approximate 95% Confidence Limits	
b_0	1124.3	44.44	1037.2	1211.5	1257.1	61.21	1137.1	1377.2
b_1	74.33	0.46	73.43	75.22	84.73	0.52	83.71	85.76
b_2	-34.82	2.41	-39.56	-30.09	-52.51	0.56	-53.61	-51.41
b_3	-11.26	2.37	-15.91	-6.60	-3.78	0.39	-4.54	-3.02
b_4	-11.88	0.15	-12.18	-11.58	-13.48	0.34	-14.14	-12.81
b_5	-8.94	0.16	-9.26	-8.62	-14.45	0.58	-15.59	-13.30
α	168.6	2.5374	163.6	173.6	202.8	1.4379	199.9	205.6
α_	224.8	7.1507	210.7	238.8	586.6	19.0462	549.2	623.9
α	878.8	3.3683	872.2	885.4	867.7	4.7005	858.5	876.9
α	1264.9	4.4253	1256.2	1273.5	1614.8	5.0180	1605.0	1624.7

The third breakpoint was not significantly different for the two clones. However, the rest of the breakpoints were different for the two clones. Consequently, the phase durations were different. The comparison of the two clones in terms of growth rates also indicates that the growth rates were significantly different. The growth rate for GU1 was significantly higher than that of GC1 in piecewise Phase 1. However, the growth rate for GU1 was not always greater than that of GC1 (see Table 3.5.2). The growth rate for the GU1 clone is almost

negligible at the final growth phase. This might indicate that the GU1 clone matures early compared to the GC1 clone.

3.5.1 The effect of climatic variables on breakpoints

The objective of this section was to investigate whether or not climatic variables were related to the breakpoints.

In particular, the effects of linear and polynomial climatic variables were investigated on the five distinct piecewise phases. The linear and polynomial climatic effects were presented in Table 3.5.3. Empty cells this table indicated that climatic variables were not significant. In particular, rainfall was not significant in any of the five distinct phases for both clones. The effects of temperature, solar radiation, relative humidity and wind speed did not always differ from one phase to the other. However the growth rate was different among these three piecewise phases. This indicated that the phases did not have any systematic relationship with the climatic conditions.

Table 3.5.3 The effect of climatic factors in five distinct piecewise growth phases

Piecewise Phase	Temperature	Rainfall	Solar radiation	Wind speed	Relative humidity
GC1					
1	Quadratic (-)		Cubic (+)		
2					
3			Cubic (+)		Cubic (+)
4				Linear (-)	
5	Cubic (-)		Cubic (+)	Quadratic (+)	Quadratic (-)
GU1					
1	Quadratic (-)		Quadratic (-)	Quadratic (-)	Linear (+)
2	Linear (+)		Cubic (-)	Cubic (+)	Quadratic (-)
3	Quadratic (-)		Cubic (-)	Cubic (+)	Quadratic (-)
4			Quadratic (-)	Linear (-)	Cubic (-)
5			Cubic (+)		

The results in Table 3.5.3 indicated that the climatic variables have no relationship with the piecewise phases. In the following, we assess the change in the climatic conditions around the breakpoint neighborhood. The breakpoint neighborhood was defined as the period between the lower and upper limits of the breakpoint 95% confidence interval (see Figure 3.5.3). The

first approach used was to compare the climatic conditions before and on the breakpoint neighborhoods. If the climatic conditions before and on the breakpoint neighborhood were the same, then it would be concluded that the breakpoint was not caused by climatic changes. On the other hand, if the climatic conditions at the breakpoint neighborhoods differs from the conditions before the breakpoint neighborhood it would be concluded that the climatic changes contributed to breakpoints.

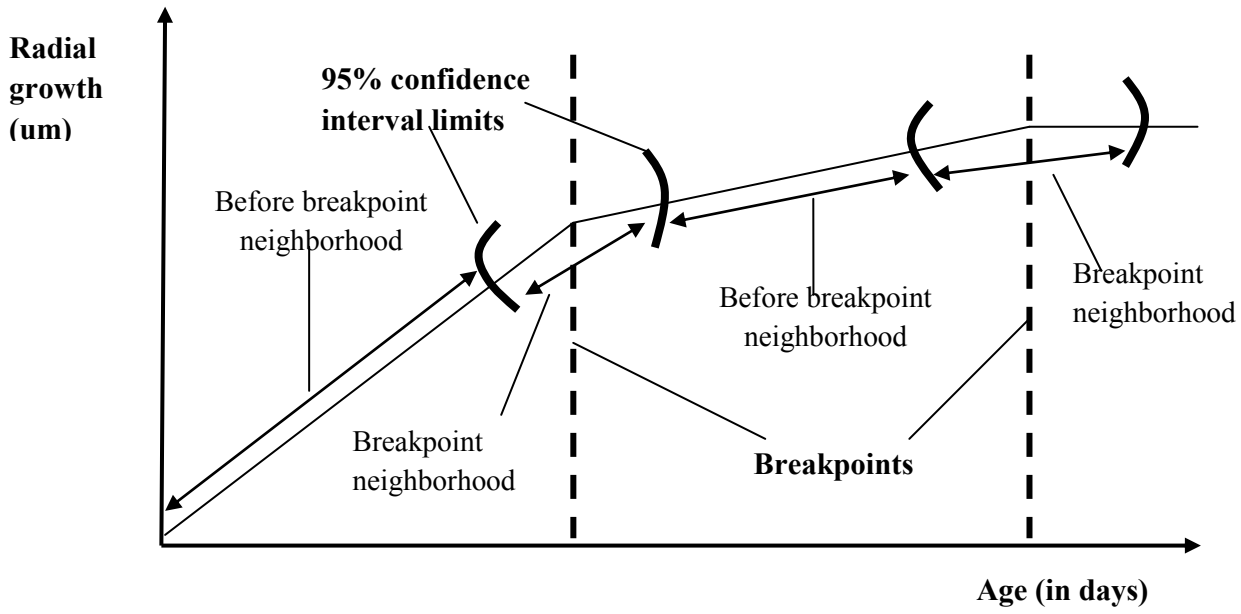

Figure 3.5.3 Piecewise model with two breakpoints and neighborhoods

It was observed in Table 3.5.4 that the climatic variables before and on the breakpoints neighborhoods were not always different. This confirmed that the breakpoints were not caused by these climatic changes.

Table 3.5.4 Comparison of climatic conditions before and at the breakpoint neighbourhoods

Breakpoint	GC1			GU1		
	Mean of the phase before breakpoint neighborhood	Mean on breakpoint neighborhood	Difference	Mean of the phase before breakpoint neighborhood	Mean on breakpoint neighborhood	Difference
Temperature						
1	18.06	20.10	-2.04	18.51	21.22	-2.72*
2	20.48	20.42	0.06	20.08	22.03	-1.95
3	20.31	15.72	4.59*	20.15	18.80	1.34
4	20.77	20.97	-0.19	20.62	17.91	2.71*
Rainfall						
1	0.08	0.03	0.05	0.07	0.14	-0.06
2	0.09	0.07	0.02	0.07	0.07	0.00
3	0.08	0.09	-0.01	0.11	0.02	0.09
4	0.05	0.03	0.02	0.05	0.02	0.03
Solar radiation						
1	0.49	0.55	-0.06	0.52	0.65	-0.13
2	0.70	0.77	-0.07	0.64	0.82	-0.17
3	0.63	0.28	0.36	0.59	0.29	0.30
4	0.35	0.35	0.00	0.28	0.32	-0.04
Relative humidity						
1	81.58	85.58	-4.00*	81.50	86.86	-5.37*
2	80.95	80.89	0.06	86.19	81.77	4.41*
3	87.28	93.22	-5.94*	89.71	89.72	-0.01
4	73.37	68.54	4.83*	74.00	78.96	-4.96*
Wind speed						
1	1.30	1.74	-0.45	1.43	1.80	-0.37
2	2.06	2.19	-0.13	1.65	2.06	-0.41
3	1.98	1.54	0.44	2.52	1.62	0.90
4	1.48	1.76	-0.28	1.64	1.71	-0.07

* significant at 5% level of significant

The second approach was to assess if the breakpoints were due to identical climatic conditions at the breakpoint neighborhoods (refer to Figure 3.5.3). If the breakpoints occurred under the same climatic conditions, then it would be concluded that the same climatic conditions enforced the occurrence of breakpoints. The Duncan's multiple range tests in ANOVA were used to compare the means of the climatic variables on breakpoint neighbourhoods. The grouping of means on the breakpoint neighborhoods was presented in Table 3.5.5 where means of the same letter are not significantly different. The results in Table 3.5.5 indicated that the breakpoints occurred under different climatic conditions except rainfall. However, as it can be seen in Table 3.5.4 rainfall identical before breakpoint neighbourhood and on the breakpoint neighbourhood. Thus we can conclude that the breakpoints were not caused by climatic changes.

Table 3.5.5 Duncan comparison of means at breakpoint neighborhoods

Breakpoint	GC1			GU1		
	Duncan Grouping		Mean on the breakpoint neighbourhood	Duncan Grouping		Mean on the breakpoint neighbourhood
Temperature						
1		A	19.96		A	21.52
2		A	20.43		A	22.00
3		C	16.01		C	18.80
4		A	20.99		C	17.91
Solar radiation						
1	C	B	86.03		A	86.76
2	C		80.88		B	81.76
3		A	92.55		A	89.72
4		D	68.67		B	78.96
Relative humidity						
1		C	0.52		B	0.62
2		A	0.78		A	0.82
3		E	0.28		C	0.29
4		E	0.35		C	0.32
Rainfall						
1		A	0.03		A	0.14
2		A	0.07		A	0.07
3		A	0.09		A	0.03
4		A	0.03		A	0.02
Wind speed						
1	B	A	1.74	B	C	1.80
2		A	2.19	B	A	2.06
3	B		1.54	B	C	1.62
4	B	A	1.76	B	C	1.71

3.5.3 Summary of results obtained from fitting piecewise models

Four substantial changes in growth rate were identified for each clone, indicating that growth occurs in five distinct phases. The changes occurred at different times for the two *Eucalyptus* clones. Consequently, the growth phase durations were different. However, the overall growth pattern was identical for the two clones. i.e. the growth rate increases sharply in Phase 1, it then starts to diminish, with the growth rate in Phase i greater than that of Phase i+1 for both clones. The growth rate of GU1 was higher than that of GC1 in Phase 1. Furthermore, Phase 1 duration for GU1 was longer than that of GC1. In particular, the growth rate of GU1 accelerated for the first 203 days and then started diminishing; on the other hand the growth rate of GC1 accelerates for the first 169 days and then starts diminishing. The margins of the changes in growth rates are different for the two clones. Statistical methods used indicated

that these changes have no detectable systematic relationship with climatic changes. These breakpoints were possibly related to genetic and/or age factors.

4. The effects of climatic variables on daily radial increment

In all the regression models (linear and polynomial) considered so far, we were dealing with the daily radial growth. In other words, we were dealing with the cumulated daily increments up to and including the current day. Unlike the previous chapters, this chapter dealt with the factors affecting the daily radial increments. In particular, the objective of this chapter was to determine whether or not the radial increment was different under different levels (low, normal, high) of climatic variables and thereby determine the combination of these classified climatic variables for optimal daily radial increments. Classification of climatic variables was explained in Section 2.2.4.

The first part of this analysis investigated the main and joint effects of climatic variables on daily radial increment without the increment duration. In particular, the ANOVA model without the increment duration is given as

$$\begin{aligned} increment = {} & \beta_0 + \beta_1 temperature + \beta_2 rain + \beta_3 solar + \beta_4 wind + \beta_5 humidity + \beta_6 temperature * rain \\ & \beta_7 temperature * solar + \beta_8 temperature * wind + \beta_9 temperature * humidity + \\ & \beta_{10} rain * solar + \beta_{11} rain * wind + \beta_{12} rain * humidity + \beta_{13} solar * wind + \\ & \beta_{14} solar * humidity + \beta_{15} wind * humidity + \end{aligned}$$

Table 4.1.1 The ANOVA results for the classified effects model without increment duration

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
GC1					
Model	50	721376.94	14427.54	2.15	<.0001
Error	1488	9981594.79	6708.06		
Corrected Total	1538	10702971.73			
R^2=0.0674					
GU1					
Model	50	1512365.258	30247.305	5.92	<.0001
Error	1488	7604428.370	5110.503		
Corrected Total	1538	9116793.628			
R^2=0.1659					

The ANOVA results for this model are presented in Table 4.1.1. The R-Square values of 0.07 and 0.17 for GC1 and GU1, respectively, indicated that only about 7% and 17% of variation in radial increments for GC1 and GU1 were explained by the classified climatic effects model.

Since the R-Square indicated that the model was inadequate, no further inferences were carried out in this model. The next step was to include the increment duration effects in the classified climatic effects model. The fitted increment duration and classified climatic effects model was given by

$$\begin{aligned} increment = {} & \beta_0 + \beta_1 temperature + \beta_2 rain + \beta_3 solar + \beta_4 wind + \beta_5 humidity + \beta_6 temperature * rain \\ & \beta_7 temperature * solar + \beta_8 temperature * wind + \beta_9 temperature * humidity + \\ & \beta_{10} rain * solar + \beta_{11} rain * wind + \beta_{12} rain * humidity + \beta_{13} solar * wind + \\ & \beta_{14} solar * humidity + \beta_{15} wind * humidity + \beta_{16} duration + \end{aligned}$$

where *duration* is Increment duration. The increment duration was only recorded for three years and eight months. So, unlike other models in this study, this model predicts for three years and eight months. The results for this model were summarised in Table 4.1.2.

The R-Square values of 0.31 and 0.47 for GC1 and GU1, respectively, indicated that: 1) 31% of total variation in daily radial increment of the GC1 clone was explained by the fit; and 2) 47% of the total variability in daily radial increment of the GU1 clone was explained by the fit. The mean squared errors were small compared to those of the classified climatic effects model without the increment duration effects. Therefore in terms of R-Square values and the mean squared errors, this model was an improvement of the classified climatic effects model.

Table 4.1.2 ANOVA results for the increment duration and classified climatic effects model

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
GC1					
Model	51	1352651.917	26522.587	7.71	<.0001
Error	893	3070224.592	3438.101		
Corrected Total	944	4422876.509			
R^2=0.3058					
GU1					
Model	51	1547307.584	30339.364	13.19	<.0001
Error	766	1761977.965	2300.232		
Corrected Total	817	3309285.549			
R^2=0.4676					

This model was assessed for independence, constant variance and normality of residuals. The residual plots displayed in Figures 4.1.1 and 4.1.2 indicated that the assumptions of constant variance and independence of errors were not violated by the data. The joint histogram of studentized residuals in Figure 4.1.3 indicated that the assumption of normality of errors was not violated by the data. Some studentized residual values appear to be inconsistent with the rest of the residuals. This indicates that there may be some outliers. These outliers would be a be a problem if they are found to be influential. However, the Cook's index plot in Figure 4.1.4 indicated that there were no influential observations.

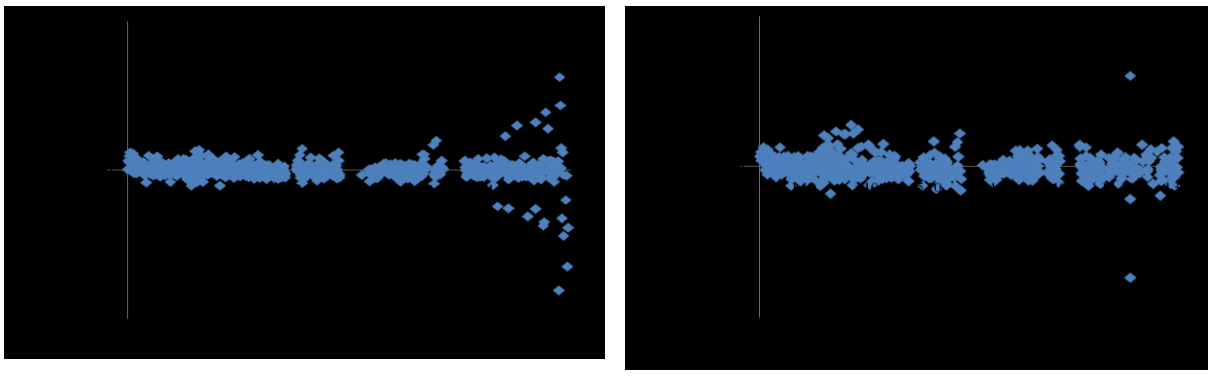

Figure 4.1.1 Index plot of residuals for the increment duration and classified climatic effects model

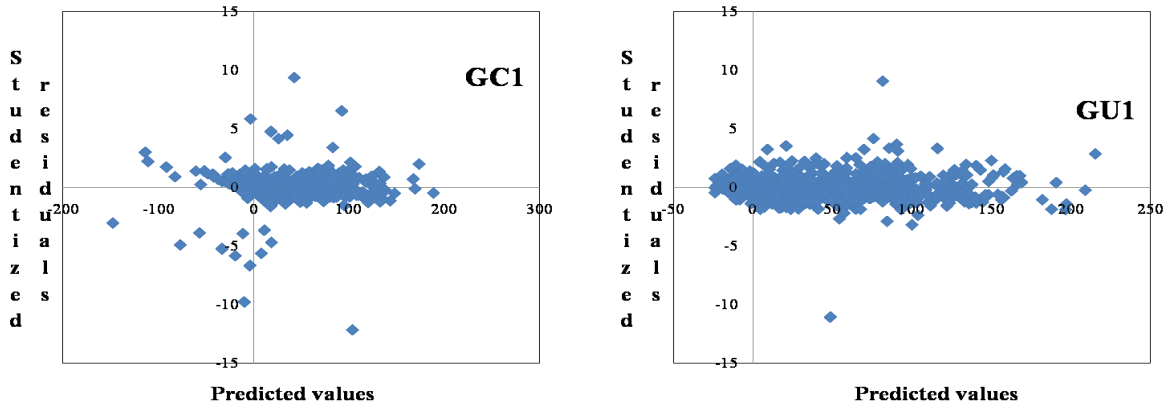

Figure 4.1.2 The plot of studentized residuals vs. predicted values for the increment duration and classified climatic effects model

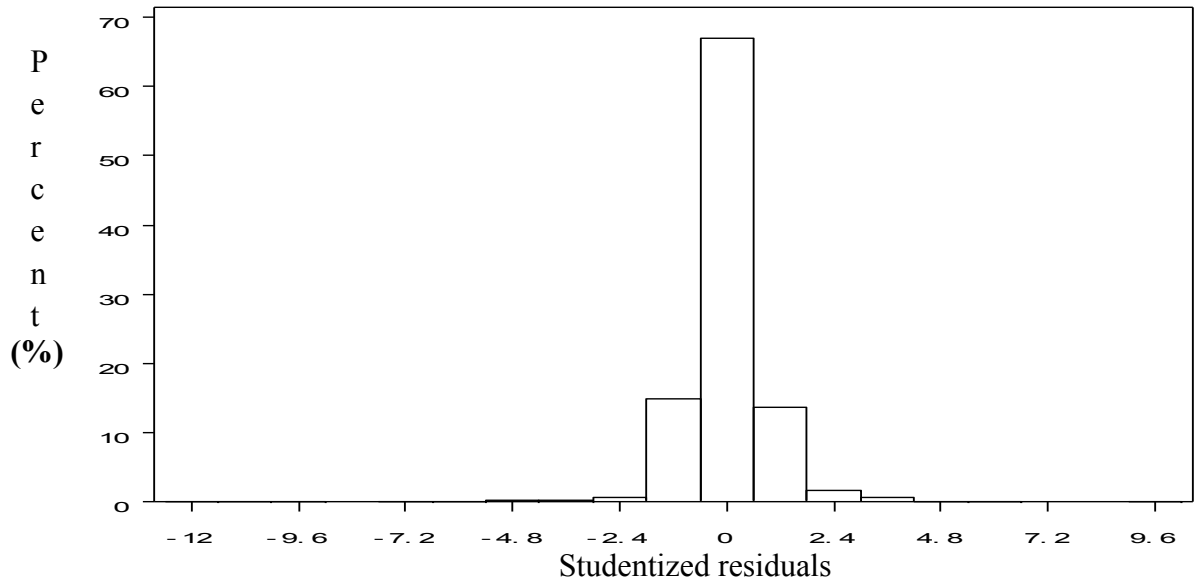

Figure 4.1.3 Histogram of studentized residuals for increment duration and classified climatic effects model

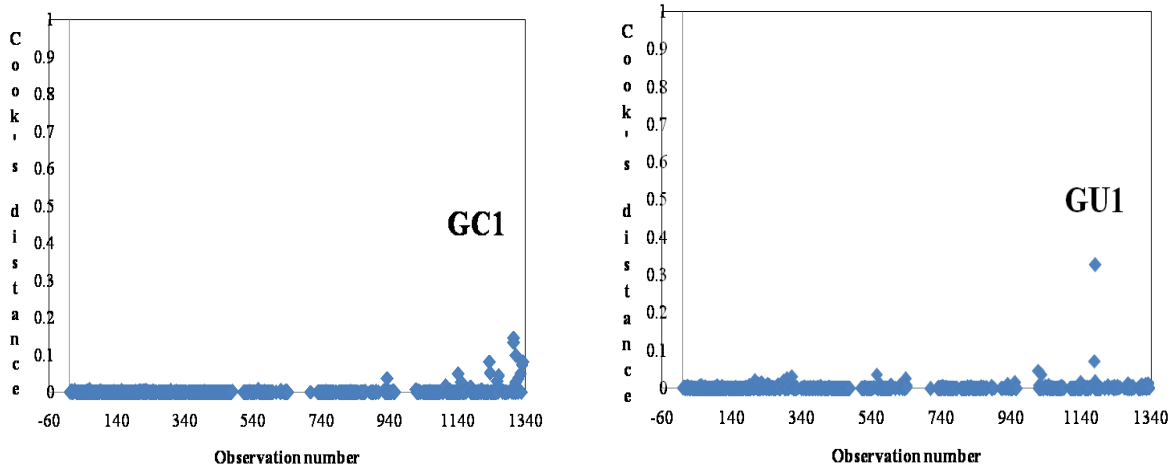

Figure 4.1.4 Index plot of Cook's distance for the classified climatic and duration effects model

The Type III tests for this model are presented in Table 4.1.3. The parameter estimate for increment duration was found to be positive and significant ($p<0.05$) indicating that longer increment durations are preferable for more increment. The radial increment for the GC1 decreases with an increase in wind speed (Figure 4.1.5). So, low wind speed was desired for optimal radial increment for GC1. Moreover, the combinations of rainfall and relative humidity for optimal radial increment for GC1 were high rainfall and high relative humidity (Figure 4.1.6). On the other hand, the combinations of temperature and rainfall for optimal radial increment for GU1 were low temperature and high rainfall (Figure 4.1.7). Furthermore, high temperature and high relative humidity were desired for optimal radial increment for GU1 (Figure 4.1.8).

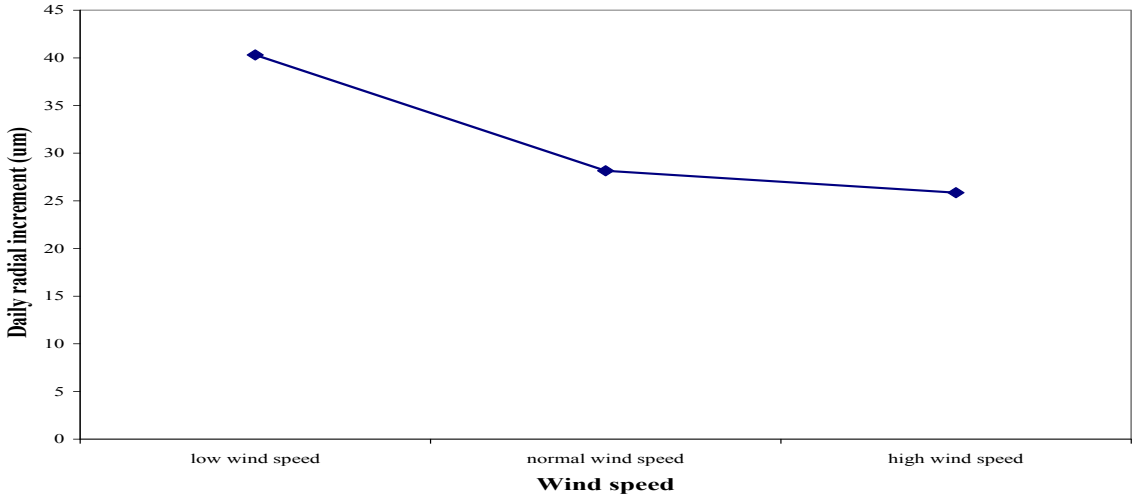

Figure 4.1.5 The effect of wind speed on GC1 radial increment

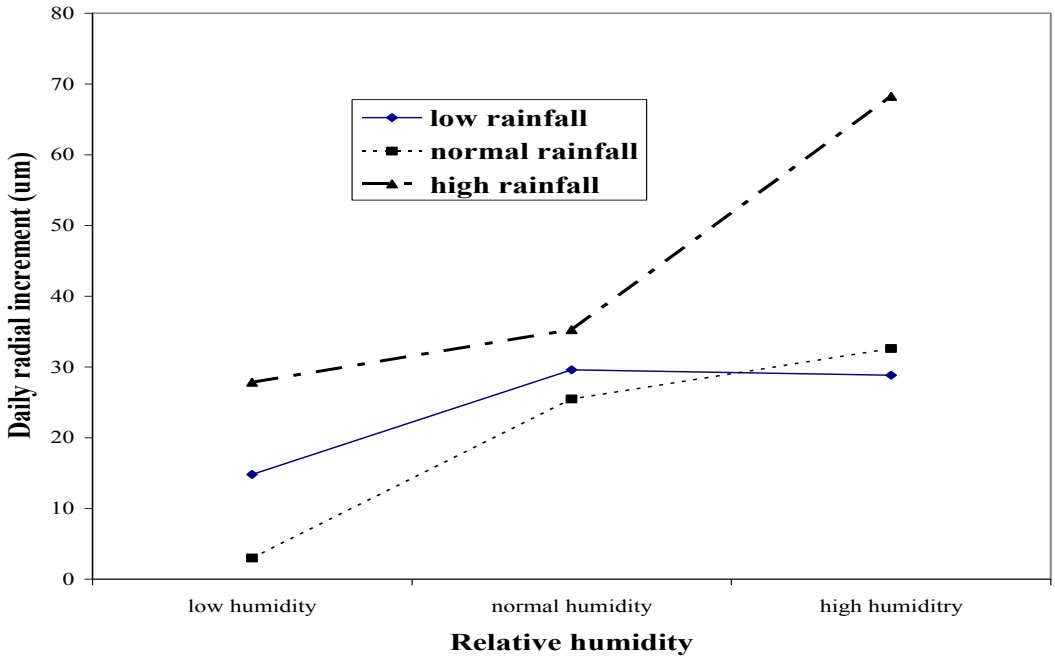

Figure 4.1.6 The joint effect of rainfall and relative humidity on GC1 radial increment

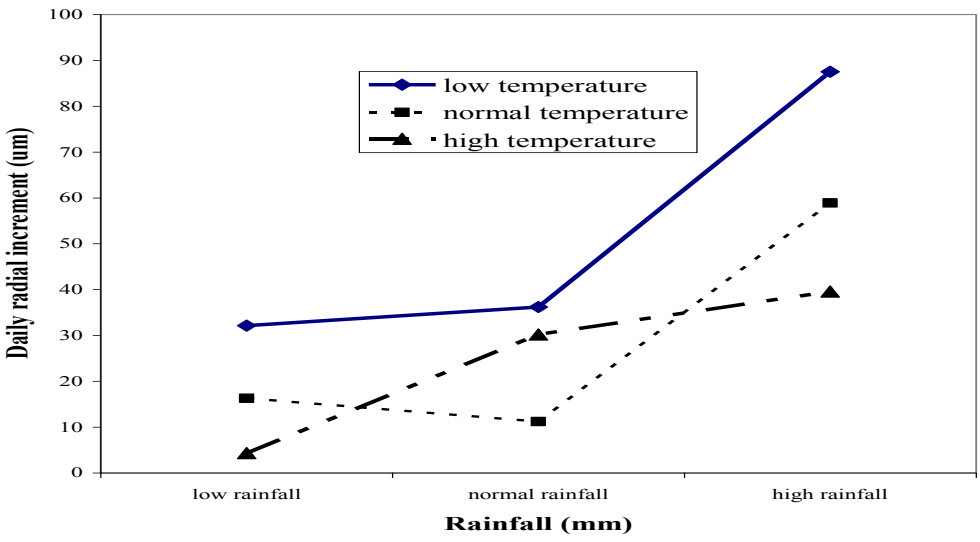

Figure 4.1.7 The joint effect of temperature and rainfall on GU1 radial increment

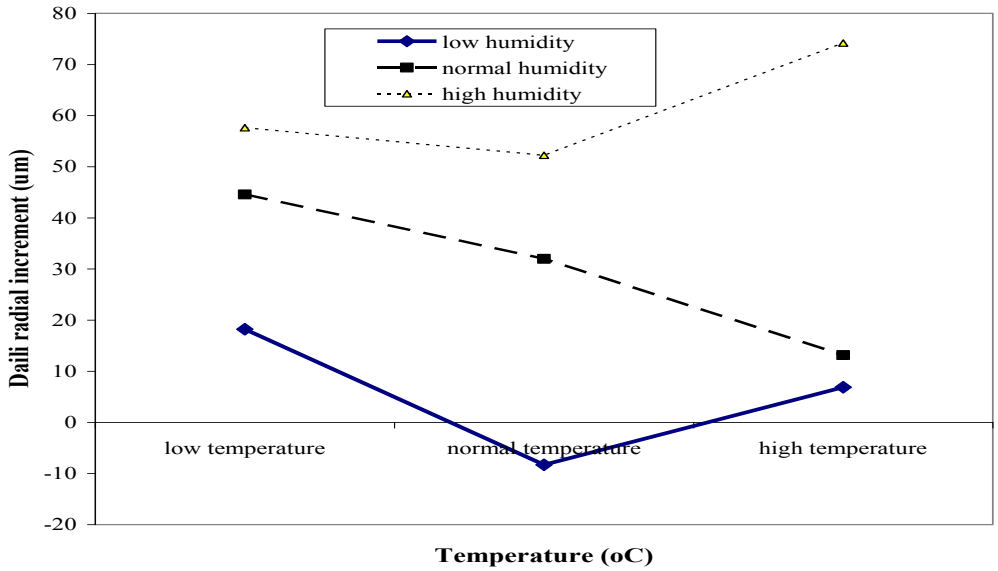

Figure 4.1.8 The joint effect of temperature and relative humidity on GU1 radial increment

Table 4.1.3 Type III tests for the classified climatic effects and duration effects model

Parameter estimate	GC1			GU1		
	Mean Square	F Value	Pr > F	Mean Square	F Value	Pr > F
Increment duration	636103.82	185.02	<.0001	730444.26	317.55	<.0001
Temperature	4378.37	1.27	0.2804	1908.42	0.83	0.4366
Rainfall	20973.28	6.10	0.0023	7973.55	3.47	0.0317
Temperature*rainfall	4961.97	1.44	0.2177	6818.25	2.96	0.0191
Solar radiation	4828.82	1.40	0.2460	3988.33	1.73	0.1773
Temperature*solar radiation	3969.65	1.15	0.3295	5178.39	2.25	0.0620
Rainfall*solar	5632.67	1.64	0.1625	3179.46	1.38	0.2383
Wind speed	19441.54	5.65	0.0036	2138.19	0.93	0.3952
Wind speed*temperature	1580.75	0.46	0.7653	2397.64	1.04	0.3843
Wind speed*rainfall	6089.90	1.77	0.1325	1849.29	0.80	0.5228
Wind speed*solar radiation	6322.41	1.84	0.1193	1544.71	0.67	0.6118
Relative humidity	9693.60	2.82	0.0602	9309.47	4.05	0.0178
Relative humidity* temperature	7409.70	2.16	0.0722	7385.76	3.21	0.0125
Relative humidity* rainfall	15430.77	4.49	0.0014	815.34	0.35	0.8410
Relative humidity* solar radiation	3019.30	0.88	0.4764	3938.18	1.71	0.1453
Wind speed* relative humidity	7144.00	2.08	0.0818	4994.07	2.17	0.0706

4.1 Summary of the results obtained from fitting the daily radial increment model

The climatic variables were classified into three levels and assessed on daily radial increment. Two models were fitted for the radial daily increment. The first model consisted of classified climatic variables. This model provided small R-Square values, indicating that some important variables(s) were not captured in the model. The increment duration was then added in the classified climatic effects model. The increment duration was only recorded for about three years and eight months. So, unlike other models in this study, this model predicts for three years and eight months. The results of this model indicated that longer increment durations desirable for more increment. Low wind speed was desired for optimal radial increment for the GC1. The combination of low relative humidity and high rainfall was also desired for optimal radial increment for the GC1. On the other hand, the combinations of: 1) low temperature and high rainfall; 2) high temperature and high relative humidity were desired for optimal radial increment of the GU1.

All the analyses carried out did not clearly quantify the contribution of climate on radial growth. Most of the variation in radial growth was explained by age (see Sections 3.2 and 3.5). Therefore the following chapter investigated the functional relationship between age and radial growth.

5. Fitting a growth curve model to the radial growth of the two clones

This chapter investigates the functional relationship between radial growth and age. The observed radial growth of the two clones over the five year period displayed in Figure 2.1.2 indicated that the relationship between radial growth and age was not linear. It was therefore sensible to investigate the goodness of fit of several nonlinear growth models. The radial growth-age model is of the form

$$radialG = f(age) + \cdot$$

where $f(age)$ is the mean growth function of age or the radial growth function of age. The mean response function can take different forms. The hyperbolic, power and logarithm models were investigated. Among these models, the power function of age appeared to explain more variability in radial growth in terms of R-Square criterion. In particular, the square root of age ($\sqrt{age}$) explained 99% of the total variation in radial growth. However, the plotted observed and fitted in Figures 5.1.1 indicate that the growth model with predictor $\sqrt{age}$ fits well in the first 753 days and 666 days of GC1 and GU1, respectively. The fitted values then start differing from the observed values and thus indicating that the growth model with predictor $\sqrt{age}$ did not adequately explain relationship between radial growth and age. This suggested further investigation of growth models to explain the relationship between age and radial growth.

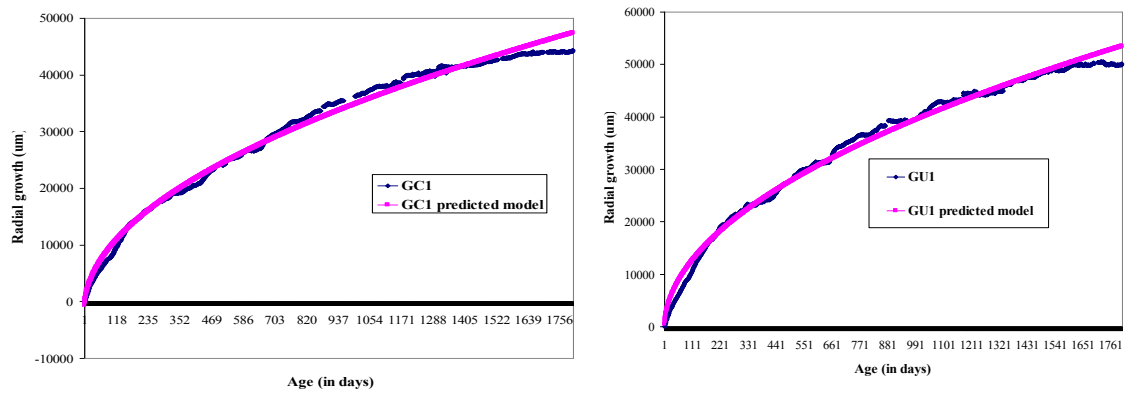

Figure 5.1.1 Observed vs. fitted values from fitting the square root function for GC1 and GU1

The next approach was to investigate the goodness of fit of nonlinear growth curve models. The growth curve model models the periodic changes in the underlying growth process. The maximum lifetime size is one of the parameters of the growth curve. The maximum growth rate and the age at which this maximum growth rate is reached can be calculated from the estimated parameters. Once the model parameters are estimated, they are used to estimate the growth at a particular age as a fraction of the estimated maximum radius. Once the size of interest is determined, the duration of completion is also estimated. Many growth curve models were explored. The Logistic, Gompertz and von Bertalanffy growth curve models provide parameters that allow for biological interpretation (see Subsection 2.5.2). Thus, these three models were used to model the relationship between age and radial growth of the two clones. The following are the results on these three growth curve models.

5.1 The Logistic growth curve model

The ANOVA results for this model are presented in Table 5.1.1. These results indicate that this model was significant. The R-square values of 0.999 indicate that the model explained 99.9% of the total variation in radial growth.

Table 5.1.1 The ANOVA results for the logistic growth curve model

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
GC1					
Model	3	1.776E12	5.921E11	290549	<.0001
Error	1646	3.3543E9	2037856		
Corrected Total	1649	1.78E12			
R^2=0.999					
GU1					
Model	3	2.291E12	7.637E11	194933	<.0001
Error	1646	6.4485E9	3917656		
Corrected Total	1649	2.297E12			
R^2=0.999					

The parameter estimates for the Logistic growth curve are presented in Table 5.1.2. All parameter estimates for this model were found to be significant. The estimated maximum radius were 44170.3 and 49463.7 μ m for GC1 and GU1, respectively. Half of the maximum growth sizes were estimated to be achieved in about 473 days for GC1 and 443 for GU1. These are the days at which the maximum growth rates were reached. Thus this model

implies that the GU1 reaches half of its estimated maximum radius earlier than the GC1. The predicted versus the fitted values are potted in Figure 5.2.1. The upper and lower 95% of the predicted values (also displayed in Figure 5.2.1) indicate that this model can accommodate a different data set. Thus the Logistic model was a good fit, but predicted poorly at early stage.

Table 5.1.2 Parameter estimates for the Logistic growth curve

Parameter estimate	GC1				GU1			
	Estimate	**Approx Std Error**	**Approximate 95% Confidence Limits**		**Estimate**	**Approx Std Error**	**Approximate 95% Confidence Limits**	
α	44170.3	83.6635	44006.2	44334.4	49463.7	112.3	49243.3	49684.0
r	472.3	2.2758	467.8	476.7	443.0	2.7597	437.6	448.4
k	0.00312	0.000014	0.00309	0.00314	0.00313	0.000018	0.00310	0.00317

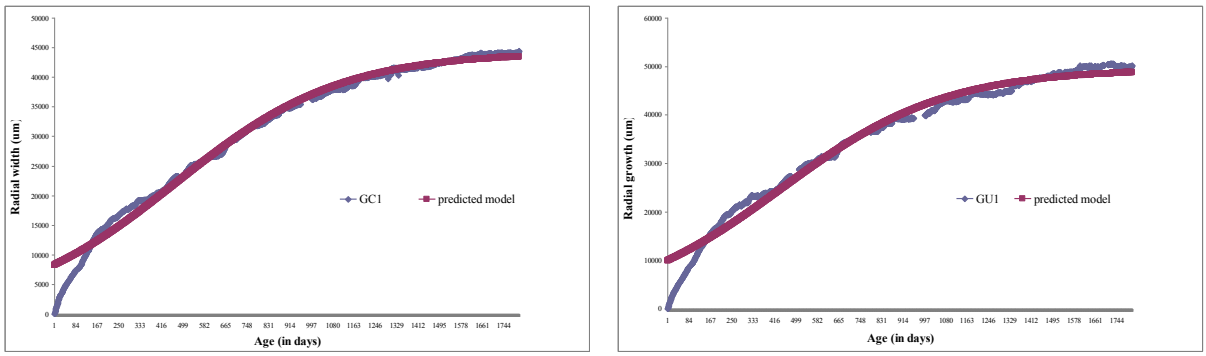

Figure 5.2.1 Observed and fitted curves for Logistic model

5.2 The Gompertz growth curve

The ANOVA results for the Gompertz growth curve model are presented in Table 5.2.1. This model was found to be significant and explains about 99.9% of total variation in radial growth. the MSE's of this model are less than those of the logistic model, indicating that the Gompertz model was a better fit.

Table 5.2.1 ANOVA results for the Gompertz growth curve model

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
GC1					
Model	3	1.778E12	5.926E11	491702	<.0001
Error	1646	1.9836E9	1205109		
Corrected Total	1649	1.78E12			
R^2=0.999					
GU1					
Model	3	2.294E12	7.645E11	319477	<.0001
Error	1646	3.9389E9	2393021		
Corrected Total	1649	2.297E12			
R^2=0.999					

The parameter estimates for the Gompertz model are presented in Table 5.2.2. The estimated maximum growth sizes are 45721.6 μ ı for GC1 and 51102.8 μ ı for GU1. About 37% of the estimated maximum growth sizes is estimated to be achieved in about 307days for GC1 and 281 for GU1. It is at these points that the growth rate is expected to be most rapid. This model also indicates that the GU1 clone reaches its maximum growth rate earlier than the GC1 clone. The predicted versus the fitted values are potted in Figure 5.2.2. These plots indicate that the Gompertz model was a good fit, but predicted poorly at early stage.

Table 5.2.2 Parameter estimates for Gompertz growth curve

Parameter estimates	GC1				GU1			
	Estimate	**Approx Std Error**	**Approximate 95% Confidence Limits**		**Estimate**	**Approx Std Error**	**Approximate 95% Confidence Limits**	
α	45721.6	96.6666	45532.0	45911.2	51102.8	130.4	50847.1	51358.5
r	306.7	1.5817	303.6	309.8	280.6	1.9310	276.8	284.4
k	0.00218	0.000015	0.00215	0.00221	0.00222	0.000019	0.00218	0.00225

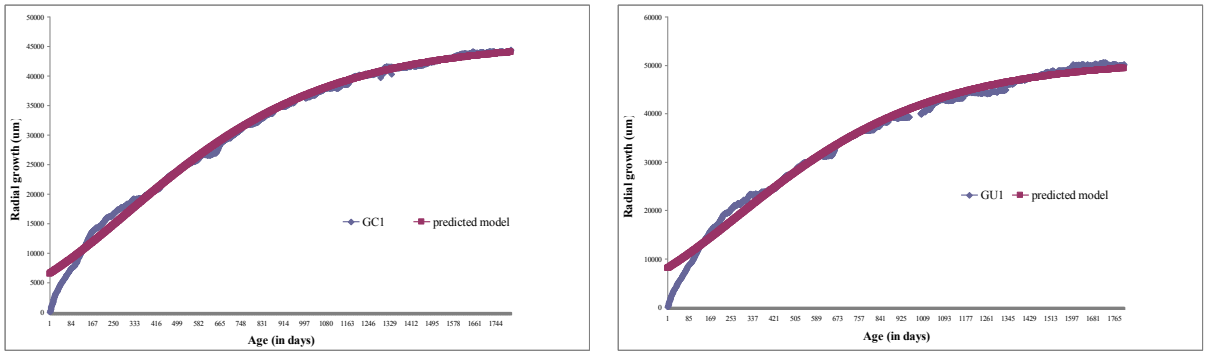

Figure 5.2.2 Observed and fitted curves for Gompertz model

5.3 The von Bertalanffy growth model

The ANOVA results for the von Bertalanffy model are presented in Table 5.3.1. This table indicated that this model was significant and accounts for 99.9% of the total variation in radial growth. The MSE's of this model were smaller than that of those of the Logistic and Gompertz models, indicating that the von Bertalanffy model was a better fit.

Table 5.3.1 ANOVA results for the von Bertalannfy growth curve model

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
GC1					
Model	3	1.778E12	5.927E11	627788	<.0001
Error	1646	1.554E9	944105		
Corrected Total	1649	1.78E12			
R^2=0.999					
GU1					
Model	3	2.294E12	7.648E11	410741	<.0001
Error	1646	3.0649E9	1862017		
Corrected Total	1649	2.297E12			
R^2=0.999					

The parameter estimates for the von Bertalanffy model are presented in Table 5.3.2. The estimated maximum radius for GC1 is 46694.2 μ ı. About 30% of this estimated maximum radius is reached in 368 days. On the other hand, the estimated maximum radius of the GU1

is 52029.6 μ ı. 30% of this size is reached in 376 days. These are the days in which the growth rate is expected to be most rapid. This model indicates that the GC1 reaches its maximum growth rate earlier than the GU1. The predicted versus the fitted values are potted in Figure 5.2.3. These plots indicate that the von Bertalanffy model was a good fit, but predicted poorly at early stage.

Table 5.3.2 Parameter estimates for von Bertalanffy growth curve

Parameter estimates	GC1				GU1			
	Estimate	Approx Std Error	Approximate 95% Confidence Limits		Estimate	Approx Std Error	Approximate 95% Confidence Limits	
α	46694.2	99.0468	46499.9	46888.4	52029.6	130.9	51772.9	52286.3
r	367.9	3.7837	360.5	375.3	375.1	4.7378	365.8	384.4
k	0.00186	0.000012	0.00184	0.00188	0.00191	0.000015	0.00188	0.00194

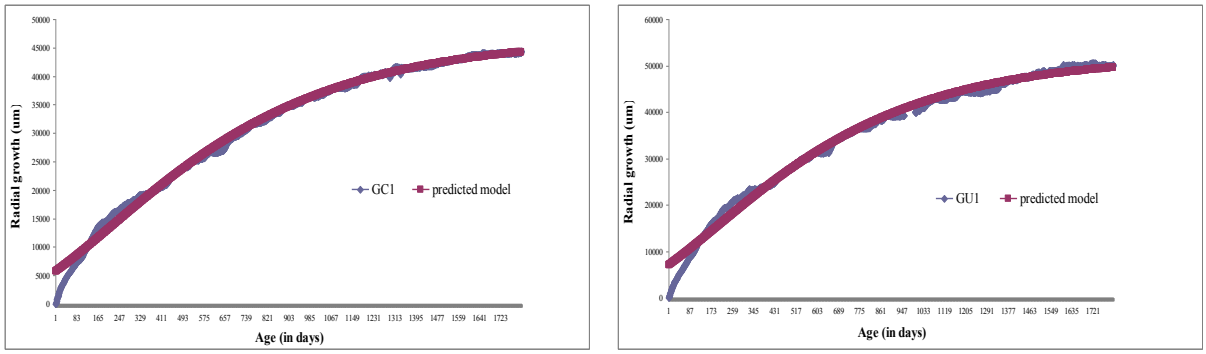

Figure 5.2.3 Observed and fitted curves for von Bertalanffy model

5.4 Comparison of the three growth curve models

The three growth curves provided estimates of the asymptote, inflection points and the scale parameters governing growth rate to the end of the growth circle. These parameters were used to estimate the maximum growth rates. The calculated maximum growth rate and the estimated parameters for all the models are summarized in Table 5.4.1. The largest values of the maximum radius were provided by the von Bertalanffy model, followed by the Gompertz model and then the Logistic model. All three growth curves indicate that the expected maximum radius of the GU1 is greater than that of the GC1. In particular, the estimated lifetime size of the GU1 is 6% (on average) larger than that of the GC1. The estimated periods of maximum growth rate were different for the three models. The largest estimated

maximum growth rates were provided by the Gompertz model, followed by the Logistic model and then the von Bertalanffy model. The Logistic and Gompertz models indicate that the GC1 achieves is maximum growth rate earlier than the GC1. However, the von Bertalanffy model indicated that the GC1 reaches its maximum growth rate earlier than the GU1. All the models indicate that the maximum growth rate of the GU1 is greater than that of the GC1. In particular, the maximum growth rate of the GU1 is 5% higher than that of the GC1. These three models explained a great deal of variation in radial growth in terms of R-Square. The von Bertalanffy model provided the smallest values of MSE, indicating that this model was the best among the three. The von Bertalanffy model was followed by the Gompertz model, and then the Logistic model.

Table 5.4.1 Parameter summary of growth curve models

Model	Carrying capacity	Maximum growth rate	Age of inflection point (days)	R-Square (%)	MSE
GC1					
Logistic	44170.3	34.45283	473	99.8	2037856
Gompertz	45721.6	36.66768	307	99.9	1205109
von Bertalanffy	46694.2	25.73369	368	99.9	944105
GU1					
Logistic	49463.7	38.70535	443	99.7	3917656
Gompertz	51102.8	41.73527	281	99.9	2393021
von Bertalanffy	52029.6	28.67409	376	99.9	1862017

The estimated maximum radius can be used to estimate the durations at which different sizes were achieved. For our research we chose to estimate the durations required to complete 50%, 75%, 90%, 95% and 99% of the maximum radius. The sizes as proportions of the estimated maximum radius and the durations (in days and in years) for the best fit (von Bertalanffy model) are summarized in Table 5.4.2. This model indicated that the two clones achieve half of the estimated lifetime size in about one year and five months. However, it takes about six years for them to complete the other half of the estimated maximum radius. This indicates that the growth rate is most rapid at young age. The growth rate slowed down as the trees grew older. The time needed to complete these percentages of the estimated maximum radius is always less for GU1 as compared to that of the GC1, indicating that the

GU1 grows faster than the GU1. However, the estimated time age of maturity is about eight years for both clones.

Table 5.4.2 Estimated number of days required to achieve certain percentage of the maximum radius

% of maximum radius	von Bertalanffy model
GC1	
50	481 (1.3 years)
75	919 (2.5 years)
90	1442 (4.0 years)
95	1825 (5.0 years)
99	2697 (7.4 years)
GU1	
50	473 (1.3 years)
75	911 (2.5 years)
90	1434 (3.9 years)
95	1817 (5.0 years)
99	2689 (7.4 years)

6. Summary and conclusions

The main objective of this thesis was to investigate the linear and nonlinear relationships between the radial growth of two *Eucalyptus* clones and climatic variables. In particular, the objective was to identify the climatic variables for optimal radial growth. This would help in choosing suitable climatic conditions to grow eucalypts. The data used in this thesis were collected over a 5-year period by Sappi. The data were preprocessed prior to the analyses with the aim of getting good quality results. The growth curves of the two clones were found to be different as well as not parallel, suggesting that the analyses should be carried out separately for the two clones. The general linear model (multiple linear regression, ANOVA and ANCOVA) model was used to investigate the immediate effects of climatic variables on radial growth.

Firstly, a multiple linear regression model of the growth of each clone on linear climatic effects was fitted to the data. The result was an R-Square value of about 0.35 for each clone, indicating that the model does not adequately explain the variations in radial growth. Furthermore, analysis of the model residuals indicated that some variables, not in the model, explain a large proportion of the total variation in radial growth. When the linear age effect was included in the model for both clones, the R-Square values increased to about 0.93. However, the model residuals appeared to violate the regression model assumptions. In particular, the plot of residuals versus age suggested a power transformation of age. The appropriate transformation turned out to be the square root of age.

Secondly, a multiple regression model of the radial growth of each clone on the square root of age, and polynomial effects of each climatic variable up to order three, was fitted to the data. Due to the large number of explanatory variables in the model, the stepwise selection procedure was used to select the important variables. The selected models for the two clones were different up to common square root of age, linear relative humidity and wind speed effects. The R-Square values for both models were approximately 0.99. The plots of residuals versus age suggested the presence of seasonality effects in the radial growth. Thus, the effects of month nested in year were included in the models for the two clones. The radial growth of the two clones was significantly affected by the month nested in year effects. These models are useful for predicting radial growth for five years. These models indicated that the radial growth of the two clones increased with different magnitudes per one percent increase in

relative humidity. The ideal temperature for optimal radial growth of the GC1 clone was 20.45 $^{\circ}C$. On the other hand, the radial growth of the GU1 was minimal under average wind speed (1.77m/s) and average solar radiation (mJ/hr). Moreover, the radial growth of GU1 decreased per 1mm increase in rainfall.

The response of radial growth to lagged climatic effects was investigated using the ARIMA models. There was no linear relationship between lagged temperature and radial growths of the two clones. The climatic variables were lagged up to twenty four days. There was no linear relationship between lagged rainfall and the GC1 radial growth. On the other hand, the radial growth of GU1 increases if high rainfall was received in the previous six days. The radial growth was positively correlated with the lagged relative humidity. The lagged wind speed was associated with both increase and decrease in radial growth of the two clones. The crosscorrelation plots between the climatic variables' and radial growths series indicated that the linear relationships were weak. Therefore, no further modeling was done on lagged effects.

The secondary objective was to determine whether or not the growth rates of the two clones are different. The growth rates of the two clones appeared to be changing at different times during the growth process. The piecewise linear regression model was used to identify age values at which changes in growth rates occur. These change points were referred to as breakpoints. Four breakpoints were identified for the two clones, indicating that the growth of these eucalypts occurred in five distinct piecewise phases over the five year period. The breakpoints for the two clones were found to be different. In particular, the breakpoints occurred around 169 days, 225 days, 879 days and 1265 days for the GC1 clone. On the other hand, the breakpoints for the GU1 clone occurred around 203 days, 587 days, 868 days and 1615 days. However, the pattern of change of the growth rate was identical for the two clones, i.e. the growth rate increases sharply in piecewise phase 1, it then starts to diminish, with the growth rate in piecewise phase i greater than that of piecewise phase i+1 for both clones. The investigation of whether or not the climatic variables were related to the breakpoints was done. Different statistical methods indicated that breakpoints were not related to climate, but were possibly due to genetic and/or age factors. The growth rate accelerated for the first 169 days (approximately six months) for the GC1 clone and 203 days (approximately seven months) for GU1 clone. The growth rate then started slowing down as

the trees got older. In particular, the daily growth rate for GC1 was 74.33 μ ı in piecewise phase 1. It then slowed down to 39.51 μ ı in piecewise phase 2, 28.25 μ ı in piecewise phase 3, 16.37 μ ı in piecewise phase 4 and 7.43 μ ı in piecewise phase 5. On the other hand, the daily growth rate for GU1 was 84.73 μ ı in piecewise phase 1. This growth rate slowed down to 32.22 μ ı in piecewise phase 2, 28.44 μ ı in piecewise phase 3, 14.96 in piecewise phase 4 and 0.51 μ ı in piecewise phase 5.

Moreover, the climatic variables were classified into three levels and assessed on daily radial increment. The increment duration was included in the classified climatic effects model. The increment duration was only recorded for about three years and eight months. So, unlike other models in this study, this model predicts for three years and eight months. The results of this model indicated longer increment durations resulted in more radial increment for the two clones. Low wind speed was desired for optimal radial increment for the GC1. The combination of low relative humidity and high rainfall was also desired for optimal radial increment for the GC1. On the other hand, the combinations of: 1) low temperature and high rainfall; 2) high temperature and high relative humidity are desired for optimal radial increment of the GU1.

All the analyses carried out so far indicated that most of the variability in radial growth was explained by the age. This suggested further investigation of the functional relationship between radial growth and age. Since the observed growth of the two clones appeared to be nonlinear, it was sensible to model growth with nonlinear growth curve models. In particular, the goodness of fit of the special cases of Richards growth curve model: Logistic, Gompertz and von Bertalanffy growth curve models was investigated. These models include the maximum radius of the clone as one of the parameters. The maximum growth rates reached by the two clones were estimated from the fitted models. The models generally fitted the data well, but poorly fitted the data during the early stages of growth. The von Bertalanffy growth curve model had the smallest values of MSE among the other models, indicating that it was the best fit. The estimated maximum stem radiuses were 46694.2 μ ı and 52029.6 μ ı for GC1 and GU1, respectively. The approximated age of these maximum radiuses was seven and half years.

In conclusion, we developed three models (1) polynomial climatic and age effects model; (2) four breakpoints piecewise model; and (3) von Bertalanffy growth model. All these models

were useful and they served different purposes. The polynomial and climatic effects model was the best for predicting the impact of climatic variables on radial growth of five year old eucalypts. The four breakpoints piecewise model was the best for estimating the periods of changes in radial growth rate and compare growth pattern between the two clones for five years. Finally, the von Bertalanffy model was the best for estimating the maximum radius and the maximum growth rate.

There are limitations associated with this study. Firstly, differencing was used to achieve stationarity in time series modeling. However, differencing only removes polynomial trend, it does not remove nonlinear trend. Consequently, areas of further research include investigation of more sophisticated methods to remove nonlinear trend in time series. Secondly, we used the mean daily response of each clone. The assumption for using the mean response is that the replications within the clone are homogeneous. It is, however, important to check the assumption and if the variability within the clone is heterogeneous, then the future direction of this study is to use a longitudinal mixed model.

References

Aitkin, M., Anderson, D., Francis, B. and Hinde, J. (1989). Stistical modeling in GLIM. Clarendon Press, Oxford

Atkinson, A.C. (1985). Plots, Transformations, and Regression: An introduction to graphical methods of diagnostic regression analysis. Clarendon press, Oxford.

Bates, M.D. and Watts, G.D. (1988). Nonlinear regression analysis and its applications. Wiley, New York

Bowerman, L.B., O'Connelll, T.R., and Dickey, A.D. (1986). Linear statistical methods: An applied approach. Duxbury press, Boston.

Box, G.E.P. and Jenkins, G.M. (1970). Time series analysis: Forecasting and control. San Francisco, Holden Day.

Bozkurt, H. and Erkmen, O. (2001). Predictive modeling of Yesinia enterocilica in Turkish feta cheese during storage. *Journal of food engineering*, 47, 81-87.

Bradley, R.A. and Srivastava, S.S. (1979). Correlation in polynomial regression. *Journal of the American Statististical Association*, 33, 11-14.

Chatfield, C. (2000). Time-series forecasting. Chapman and Hall, London.

Chatterjee, S. and Hadi, S.A. (1988). Sensitivity analysis in nonlinear regression. Wiley, New York.

Diggle, P.J. (1990). Time series: A biostatistical introduction. Clarendon press, Oxford.

Downes, G., Beadle, C. and Worledge, D. (1999). Daily stem growth patterns in irrigated Eucalyptus globules and E. nitens in relation to climate. *Trees: Springer-Verlag*, 14, 102-111.

Famili, A., Shen, W., Weber, R., and Simoudis, E. (1997). Data preprocessing and intelligent data analysis. *Journal of intelligent analysis*, 1, 1-28.

Fox, J. (2002). Nonlinear regression and nonlinear least squares: Appendix to an R and S-PLUS comparison to applied regression. Retrieved in April 2008 from www.cran.r-project.org.

Frazer, N.B. and Ehrhart, L.m. (1985). Preliminary growth models for green, *Chelonian mydas*, and *Loggerhead*, *Caretta caretta*, turtles in the wild. *Copea*, 1, 73-79.

Frazer, N.B., Gibbons, J.W., and, Green, J.L. (1990). Exploring Fabens' growth interval model with data on a long-lived vertebrate, *Trachemys Scripta*. *Journal of zoology*,1, 112-118.

Freund, R.J. and Wilson, W.J. (1998). Regression analysis: Statistical modeling of a response variable. Academic press, London.

Gallant, A.R. (1975). Nonlinear regression. *The American Statistician*, 29, 73-81.

Gallant, A.R. (1987). Nonlinear statistical models. North Carolina state University, Wiley, New york.

Goharian, N. and Grossman, D. (2003). Data preprocessing. Retrieved in April 2008 from www.cmapspublic.ihmc.us.

Hasani, H. and Amidi, M.Z.A. (1983). A new approach to polynomial regression and its application to physical growth of human height. Shahid Beheshti University, Iran.

Hassen, A., Wilson, D.E, Rouse, G.H. and Tail, R.G. (2004). Linear and non-linear growth curves to describe body weight changes of young Angus bulls and Heifers. Retrieved in September 2008 from www.ans.iastate.edu.

Jackson, J.E. (1999). A user's guide to principal components. Wiley, New York.

Janacek, G. (2001). Practical time series. Arnold press, London.

Janacek, G. and Swift, L. (1993). Time series: Forecasting, Simulation, Applications. Ellis Horwood, New York.

Johnson , R.A. and Wicherm, D.W. (1988). Applied multivariate statistical analysis (2nd ed). Prentice Hall, New Jersey.

Karim, M.K. (2002). Non-linear models. University of Dhaka, Bangladesh.

Kendall, M. and Ord, J.K. (1990). Time series (3rd ed). Arnold press, London.

Kozlowski, T.T. (1962). Tree growth. Ronald press, New York.

Kozlowski, T.T. and Pallardy, S.G. (1997). Growth control in woody plants. Academic press, San Diego.

Kramer, P.J. and Kozlowski, T.T. (1979). Physiology of woody plants. Academic press, London.

Küchenhoff,H. (1996). An exact algorithm for estimating breakpoints in segmented generalized linear models. Retrieved in May 2008 from www.epub.ub.uni-muenchen.de.

Lei, Y.C. and Zhang, S.Y. (2004). Features and partial derivatives of Bertalanffy-Richards growth curve model in forestry. *Nonlinear Analysis: Modeling and Control*, 9, 65-73.

Lerman, P.M. (1980). Fitting segmented regression models by grid search. *Applied statistics*, 29, 77-84.

McCullagh, P. and Nelder, J.A. (2002). Generalized linear models (2nd ed). Chapman & Hall, New York.

Meadows, D.G. (1999). Growing the forest in South Africa. *Tappi journal*, 82, 60-66.

Meyer, M.C. and Laud, P.W. (2002). Predictive variable selection in generalized linear models. *Journal of the American Statistical Association*, 97, 859-879.

Milliken, G.A. and Johnson, D.E. (1984). Analysis of messy data: Volume I: Designed experiments. Chapman and Hall, New York.

Moeti, A. (2007). Factors affecting the health status of people of Lesotho. Unpublished thesis, University of KwaZulu-Natal, Pietermaritzburg.

Myers, R.H. and Milton, J.S. (1991). A first course in the theory of linear statistical models. PWS-KENT, Boston.

Nash, C.J. and Walker-Smith, M. (1987). Nonlinear parameter estimation: An integrated system in basic. Marcel Dekker, Canada.

Neter, J., Kutner, M.H, Nachtsheim, C.J. and Wasserman, W. (1996). Applied linear statistical models (4^{th} ed). McGraw Hill, Boston.

Nsofor, G.C. (2006). A comparative analysis of predictive data-mining techniques. University of Tennessee, Knoxville

Pallardy, S.G. and Kozlowski, T.T. (1997). Physiology of woody plants. Academic press, San Diego.

Pallardy, S.G. and Kozlowski, T.T. (2007). Growth control of woody plants. Academic press, San Diego.

Ravishanker, N. and Dey, DK. Dipak. (2002). A first course in linear linear model theory. Chapman and Hall, New York.

Reuters (2008). Sappi Limited (NYSE Arca). Available in http://www.reuters.com/finance/storks/companyProfile?symbol=SPP.P accessed in September 2008.

Ricklefs, R.E. (1967). A graphical method of fitting equations to growth curves. *Ecology*, 48, 978-980.

Rust, R.T., Simester, D., Brodie, R.J. and Nilikant, V. (1995). Model selection criteria: An investigation of relative accuracy, posterior probabilities, and combinations of criteria. *Journal of management sciences*, 41, 222-233.

SAS Institute Inc. (2004). SAS ® 9.1.3. Cary, NC.

Seber, G.A.F. and Wild, C.J. (2003). Nonlinear regression. Wiley, New York.

Small, G.C. and Wang, J. (2003). Numerical methods for estimating nonlinear estimating equations. Oxford University press, New York.

Smyth, G.K. (2002). Nonlinear regression. Wiley, Chichester.

Turnbull, J.W. (1999). Eucalypt plantations. Kluwer academic, Netherlands.

Van Laar, A. (1991). Forest biometry. University of Stellenbosch, Stellenbosch.

Walton, A.N. (1997). Forecasting the monthly electricity consumption of municipalities in Kwazulu-Natal. Unpublished thesis, University of Natal, Pietermaritzburg.

Wei, W.W.S. (1990). Time series analysis: Univariate and multivariate methods. Addison-Wesley, New York.

Appendix A: Additional figures

Appendix figures

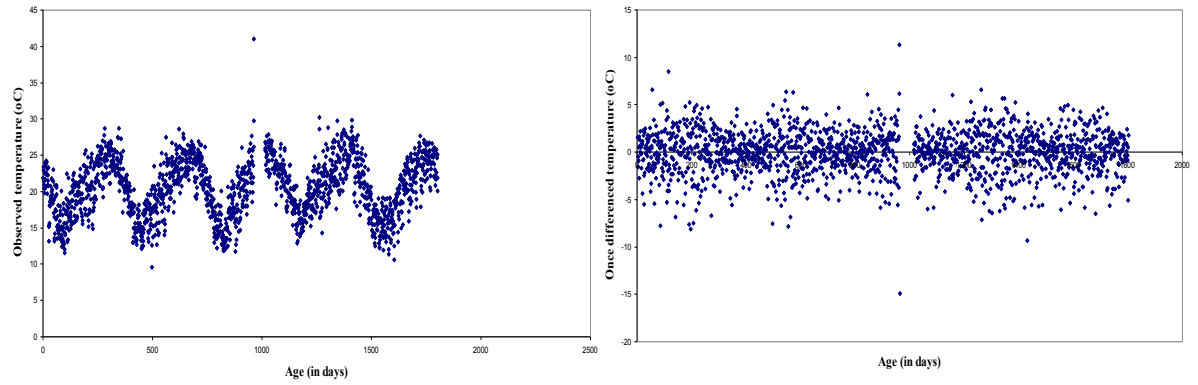

Figure A.1 Observed and once differenced temperature for the five year period

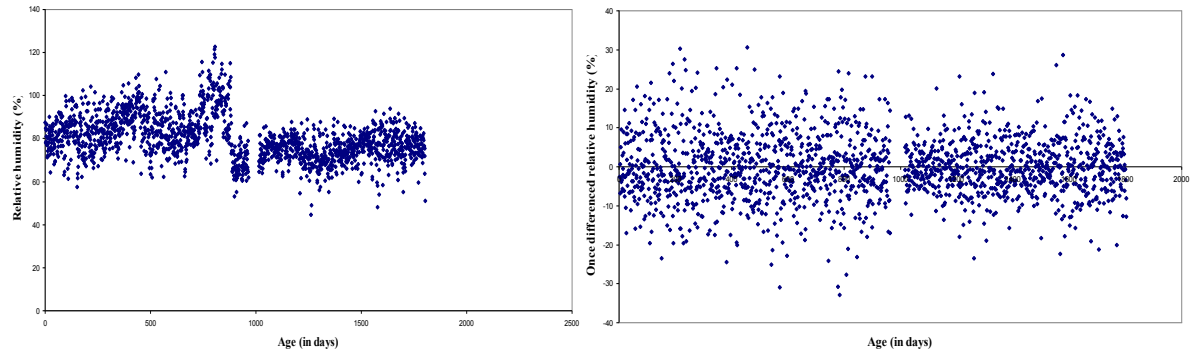

Figure A.2 Observed and once differenced relative humidity for the five year period

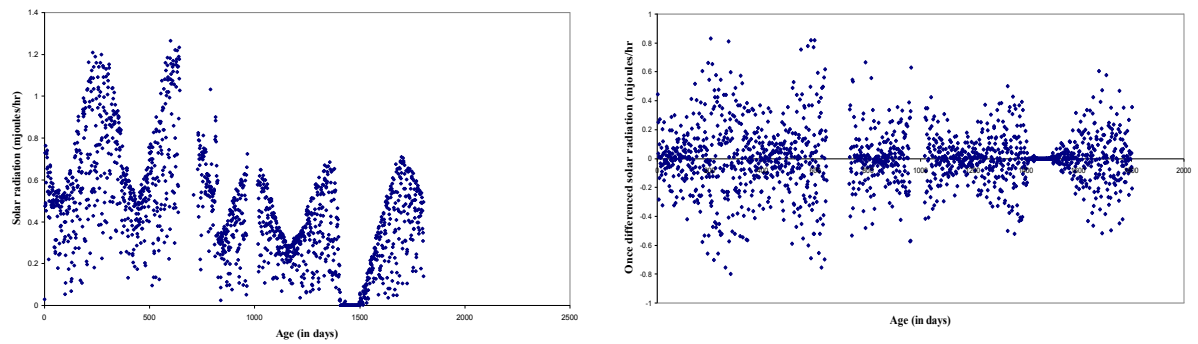

Figure A.3 Observed and once differenced solar radiation for the five year period

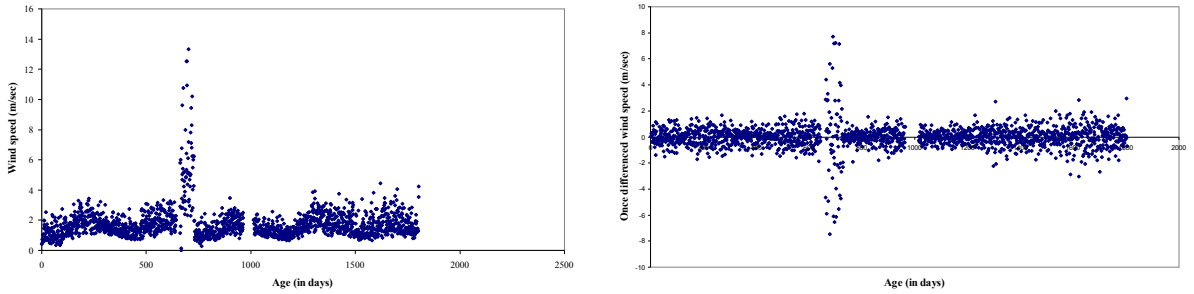

Figure A.4 Observed and once differenced wind speed for the five year period

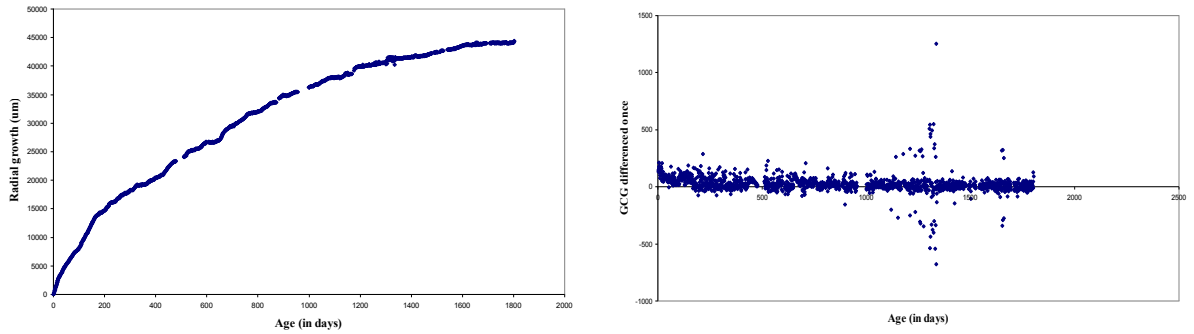

Figure A.5 Observed and once differenced GC1 radial growth for the five year period

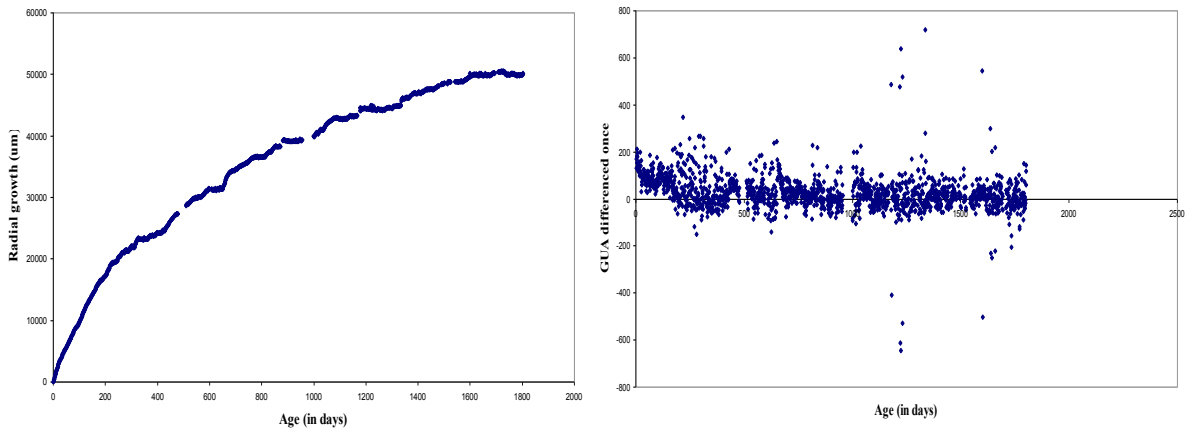

Figure A.6 Observed and once differenced GU1 radial growth for the five year period

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	-11.392417	-.01988	|.|
1	9.424500	0.01644	|.|
2	-0.876511	-.00153	|.|
3	6.582358	0.01148	|.|
4	-15.750790	-.02748	|*|
5	12.616414	0.02201	|.|
6	-12.500548	-.02181	|.|
7	-5.454740	-.00952	|.|
8	-10.830178	-.01890	|.|
9	-5.598615	-.00977	|.|
10	0.787877	0.00137	|.|
11	-3.725519	-.00650	|.|
12	-15.568878	-.02716	|*|
13	-14.545362	-.02538	|*|
14	-4.336742	-.00757	|.|
15	-10.636184	-.01856	|.|
16	-8.095522	-.01412	|.|
17	-5.776681	-.01008	|.|
18	-8.339240	-.01455	|.|
19	-11.952284	-.02085	|.|
20	-3.486940	-.00608	|.|
21	-5.557767	-.00970	|.|
22	-19.707006	-.03438	|*|
23	-5.168331	-.00902	|.|
24	-11.605027	-.02025	|.|

Figure A.7 Crosscorrelation between lagged temperature and GC1 radial growth

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	-28.854183	-.04327	|*|
1	-4.921205	-.00738	|.|
2	-6.453796	-.00968	|.|
3	-2.227547	-.00334	|.|
4	-2.459301	-.00369	|.|
5	5.382898	0.00807	|.|
6	-7.667605	-.01150	|.|
7	-17.294446	-.02593	|*|
8	-9.478754	-.01421	|.|
9	-9.187638	-.01378	|.|
10	-5.935898	-.00890	|.|
11	-11.706854	-.01755	|.|
12	-11.282280	-.01692	|.|
13	-19.824776	-.02973	|*|
14	-7.451203	-.01117	|.|
15	-19.365598	-.02904	|*|
16	-9.266458	-.01390	|.|
17	-13.619282	-.02042	|.|
18	-8.985016	-.01347	|.|
19	-9.237522	-.01385	|.|
20	-13.474079	-.02020	|.|
21	-9.138614	-.01370	|.|
22	-11.357272	-.01703	|.|
23	-19.699761	-.02954	|*|
24	-18.534063	-.02779	|*|

Figure A.8 Crosscorrelation between lagged temperature and GU1radial growth

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	1.226733	0.06244	|.|*
1	0.078631	0.00400	|.|
2	1.232132	0.06271	|.|*
3	-0.276659	-.01408	|.|
4	-0.117158	-.00596	|.|
5	0.131656	0.00670	|.|
6	1.164081	0.05925	|.|*
7	0.346580	0.01764	|.|
8	-0.624280	-.03177	|*|

9	0.373647	0.01902	.|.
10	0.141199	0.00719	.|.
11	0.262407	0.01336	.|.
12	-0.457449	-.02328	.|.
13	0.557349	0.02837	.|*
14	0.405288	0.02063	.|.
15	0.322531	0.01642	.|.
16	0.577097	0.02937	.|*
17	-0.075081	-.00382	.|.
18	-0.375106	-.01909	.|.
19	-0.387318	-.01971	.|.
20	-0.395068	-.02011	.|.
21	-0.330189	-.01681	.|.
22	0.045187	0.00230	.|.
23	0.251900	0.01282	.|.
24	0.697706	0.03551	.|*

Figure A.9 Crosscorrelation between lagged rainfall and GC1radial growth

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	1.836946	0.11178	.|**
1	2.132887	0.12979	.|***
2	0.492636	0.02998	.|*
3	0.309120	0.01881	.|.
4	0.078470	0.00477	.|.
5	0.325068	0.01978	.|.
6	0.745285	0.04535	.|*
7	-0.021263	-.00129	.|.
8	-0.635334	-.03866	*|.
9	0.823891	0.05013	.|*
10	0.351759	0.02140	.|.
11	-0.327854	-.01995	.|.
12	-0.345389	-.02102	.|.
13	0.738346	0.04493	.|*
14	0.345425	0.02102	.|.
15	-0.125520	-.00764	.|.
16	0.593732	0.03613	.|*
17	0.229631	0.01397	.|.
18	0.090438	0.00550	.|.
19	-0.852643	-.05188	*|.
20	-0.029682	-.00181	.|.
21	-0.525665	-.03199	*|.
22	0.468699	0.02852	.|*
23	0.364840	0.02220	.|.
24	0.564265	0.03434	.|*

Figure A.10 Crosscorrelation between lagged rainfall and GU1radial growth

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	-2.138125	-.04918	*|.
1	-2.931412	-.06742	*|.
2	-2.161382	-.04971	*|.
3	-1.635086	-.03761	*|.
4	-3.234394	-.07439	*|.
5	-0.977005	-.02247	.|.
6	-2.101175	-.04833	*|.
7	-1.909253	-.04391	*|.
8	-1.814179	-.04173	*|.
9	-1.744750	-.04013	*|.
10	-0.908115	-.02089	.|.
11	-0.990774	-.02279	.|.
12	-1.900203	-.04370	*|.
13	-1.487137	-.03420	*|.
14	-1.290766	-.02969	*|.
15	-1.581654	-.03638	*|.
16	-1.509590	-.03472	*|.
17	-1.319927	-.03036	*|.

18	-1.737016	-.03995	*|.
19	-0.414571	-.00953	.|.
20	-0.072747	-.00167	.|.
21	0.345372	0.00794	.|.
22	-0.055985	-.00129	.|.
23	0.153593	0.00353	.|.
24	0.384320	0.00884	.|.

Figure A.11 Crosscorrelation between lagged solar radiation and GC1radial growth

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	-3.114695	-.05856	*|.
1	-4.262597	-.08014	**|.
2	-4.279653	-.08046	**|.
3	-3.534375	-.06645	*|.
4	-3.239958	-.06092	*|.
5	-2.126109	-.03997	*|.
6	-2.622059	-.04930	*|.
7	-2.141527	-.04026	*|.
8	-1.848311	-.03475	*|.
9	-0.837388	-.01574	.|.
10	-0.657954	-.01237	.|.
11	-0.317397	-.00597	.|.
12	-0.707708	-.01331	.|.
13	-1.104385	-.02076	.|.
14	-0.430701	-.00810	.|.
15	-1.327665	-.02496	.|.
16	-2.223045	-.04180	*|.
17	-3.020664	-.05679	*|.
18	-2.145063	-.04033	*|.
19	-1.303939	-.02452	.|.
20	-0.743833	-.01399	.|.
21	0.301247	0.00566	.|.
22	-0.084648	-.00159	.|.
23	-0.026328	-.00050	.|.
24	-0.582274	-.01095	.|.

Figure A.12 Crosscorrelation between lagged solar radiation and GU1radial growth

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	32.190638	0.02625	.|*
1	-2.728516	-.00222	.|.
2	29.961368	0.02443	.|.
3	-5.973143	-.00487	.|.
4	-1.468213	-.00120	.|.
5	-35.451334	-.02891	*|.
6	13.915270	0.01135	.|.
7	-7.084994	-.00578	.|.
8	31.128902	0.02538	.|*
9	-2.450235	-.00200	.|.
10	21.489697	0.01752	.|.
11	6.417024	0.00523	.|.
12	12.897182	0.01052	.|.
13	83.075832	0.06774	.|*
14	2.566575	0.00209	.|.
15	106.218	0.08661	.|**
16	37.277213	0.03040	.|*
17	76.255342	0.06218	.|*
18	61.927799	0.05050	.|*
19	57.485817	0.04687	.|*
20	46.574956	0.03798	.|*
21	21.798917	0.01778	.|.
22	76.033887	0.06200	.|*
23	26.750227	0.02181	.|.
24	93.740184	0.07644	.|**

Figure A.13 Crosscorrelation between lagged relative humidity and GC1radial growth

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	109.123	0.07725	.|**
1	100.103	0.07086	.|*
2	59.969431	0.04245	.|*
3	21.391043	0.01514	.|.
4	42.167173	0.02985	.|*
5	34.612981	0.02450	.|.
6	47.524639	0.03364	.|*
7	64.100120	0.04538	.|*
8	72.495782	0.05132	.|*
9	65.279999	0.04621	.|*
10	50.947685	0.03607	.|*
11	71.296944	0.05047	.|*
12	54.585392	0.03864	.|*
13	78.037569	0.05524	.|*
14	33.895839	0.02399	.|.
15	121.172	0.08578	.|**
16	86.955840	0.06156	.|*
17	100.995	0.07149	.|*
18	116.663	0.08258	.|**
19	81.923778	0.05799	.|*
20	100.925	0.07144	.|*
21	75.364099	0.05335	.|*
22	87.478371	0.06193	.|*
23	78.379792	0.05548	.|*
24	88.675088	0.06277	.|*

Figure A.14 Crosscorrelation between lagged relative humidity and GU1radial growth

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	-1.350652	-.01272	.|.
1	3.318351	0.03126	.|*
2	-4.631243	-.04362	*|.
3	2.499685	0.02355	.|.
4	-3.086764	-.02908	*|.
5	3.750273	0.03533	.|*
6	-3.513561	-.03310	*|.
7	0.516598	0.00487	.|.
8	1.401646	0.01320	.|.
9	-0.673468	-.00634	.|.
10	-0.994321	-.00937	.|.
11	2.535904	0.02389	.|.
12	-8.032039	-.07566	**|.
13	3.334022	0.03140	.|*
14	4.014828	0.03782	.|*
15	-4.276983	-.04029	*|.
16	0.600599	0.00566	.|.
17	-4.546909	-.04283	*|.
18	0.242854	0.00229	.|.
19	-1.552197	-.01462	.|.
20	5.725814	0.05393	.|*
21	-3.274711	-.03085	*|.
22	-2.298749	-.02165	.|.
23	6.904802	0.06504	.|*
24	0.665142	0.00627	.|.

Figure A.15 Crosscorrelation between wind speed and GC1 radial growth

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
0	-11.417624	-.11454	**|.
1	8.079135	0.08105	.|**
2	-1.135850	-.01139	.|.
3	3.841596	0.03854	.|*
4	-0.245623	-.00246	.|.
5	0.616632	0.00619	.|.
6	-3.786711	-.03799	*|.
7	0.037575	0.00038	.|.

8	3.493225	0.03504	.	* .
9	0.931729	0.00935	.	. .
10	0.164326	0.00165	.	. .
11	-4.888077	-.04904	*	. .
12	4.626747	0.04641	.	* .
13	-9.350433	-.09380	**	. .
14	8.944456	0.08973	.	** .
15	-9.137880	-.09167	**	. .
16	-0.828241	-.00831	.	. .
17	1.692103	0.01697	.	. .
18	-2.425998	-.02434	.	. .
19	2.660806	0.02669	.	* .
20	-1.855653	-.01862	.	. .
21	7.557290	0.07581	.	** .
22	-0.472216	-.00474	.	. .
23	5.672448	0.05690	.	* .
24	-2.537782	-.02546	*	. .

Figure A.16 Crosscorrelation between wind speed and GU1 radial growth