

Growth of *Pinus radiata* shelterbelts in New Zealand

Monique MacKenzie, Leith Knowles and Mark Dean

Introduction

Shelterbelts are a common feature of the New Zealand landscape, particularly on exposed lowland plains such as Manawatu, Wairarapa, Marlborough, Canterbury and Southland. Surveys have confirmed that the most common species used is radiata pine, either planted as the sole species, or as the primary shelter-providing species in mixed species designs (Morey, 1988). Although shelterbelt management systems may vary widely, a common theme in many shelterbelts is that they are managed for both shelter, and timber production (Horvath et al, 1998). At maturity (age 25-35 yrs) these belts are often felled, and the logs are utilised for sawn timber, pulp, and other industrial uses. Predicting the growth and yield of shelterbelts is therefore an important element of any financial analysis of farming systems containing shelterbelts.

Utilisation of shelterbelts for timber production was first described by Cockayne (1914) in Canterbury. More recently Smail (1971), Tombleson (1986; 1988; 1989), and Tombleson and Inglis (1988) have elaborated on the management requirements such as pruning to produce higher quality sawn timber from shelterbelts. Current research is aiming to construct an overall farm-level modelling system to evaluate alternative shelterbelt systems, including a range of shelter planting intensities, designs, and management systems (Horvath et al 1998). Predicting the growth of radiata pine under different management systems, grown over a range of sites, is an important component of this research.

Materials and Methods

Forest Research has maintained an active programme of installing and measuring permanent sample plots (PSP's) in shelterbelts throughout New Zealand since 1983. These are usually designed to contain at least 15 sample trees in a linear plot. Each tree is tagged, and measured annually or bi-annually for diameter at breast height. A sample of trees covering the diameter range are also measured at the same time for height, and crown height. Since 1995, a 100% sample has been measured for height and crown height to strengthen the data set.

Plots have been installed into two types of shelterbelts:

- a) Existing shelterbelts which reflect normal practice for the district (termed 'grow' plots).
- b) New shelterbelts which have been established as part of replicated research trials, usually containing a range of either genetic material, cuttings vs seedlings, or silvicultural treatment such as pruning (termed 'research' plots).

The analysis reported here is confined to one and two row shelterbelts, with and without pruning, and with and without supplementary species. Trimmed shelterbelts were excluded from the analysis. Details of the PSP's are contained in table 1.

Table 1: Sample plot details

Region	Number of plots		Total no. of Observations	Total no. of increment pairs
	No of 'grow' plots	No of 'research' plots		
Northland		11	64	26
Auckland		2	6	6
Bay of Plenty		108	858	762
Hawkes Bay		47	343	279
Taranaki		16	100	56
Manawatu		8	52	38
Wairarapa		21	135	86
Nelson		8	39	32
Canterbury		25	178	101
Total		246	1771	1386

The development of a complete growth model requires the following functions and sub-models to be derived from the data:

- mean top height
- Change in stocking (mortality)
- basal area increment
- diameter distribution
- default starting values for diameter

Note that in table 1, the total number of observations were used in the formulation of the mean top height curve, while the number of increments pairs were used in the basal area increment function where the time between pairs of measurements was required to be between 0.8 and 2.0 years. Due to the size of the data set, it was not separated into model and validation sets.

General Modelling Approach:

Exploratory data analysis (EDA) was undertaken on chosen plots, all regional sets, combined North and South Island sets, and an overall New Zealand set. This exploratory work involved standard tools and also included the use of Generalized Additive Models (GAMs). The shapes of the partial relationships suggested by the GAM's were then checked for biological feasibility and were then fitted using two techniques: Nonlinear regression and Nonlinear mixed models. Invariably the mixed models were fitted using different error structures to detect the presence of serial correlation that is often present when stands are repeatedly measured.

Definition of Derived Variables:

Green Crown Length = (Mean Height - Crown Height)

Average Taper = (Mean Diameter at Breast Height - 1.4) / Mean Height

Part One: Mean Top Height Model

1.1 Abstract:

The change in mean top height over time was modelled using a Bertalanffy-Richards curve fitted within a Nonlinear mixed modelling framework with autoregressive errors. Different asymptote and rate parameters were obtained for the North and South islands given that more localised regional models were not all found to be necessary. The overall fit of the models appeared satisfactory with reasonable deviations from the fitted curves.

1.2 Data Specifications:

The data consisted of 1775 observations collected over 259 plots nationwide. Summary statistics describing the data are listed in the table below:

Variable	N	Mean	Std Dev	Minimum	Maximum
North Island	1558				
Age		8.306	4.434	2.050	27.000
Mean Top Height (m)		11.057	5.992	2.00	34.000
Height Count		14.307	4.795	3.000	38.000
Latitude		38.777		36.000	41.200
Longitude		175.902		174.200	176.700
South Island	217				
Age		13.205	7.131	3.000	32.250
Mean Top Height (m)		12.074	5.866	3.000	28.000
Height Count		20.184	5.979	4.000	30.000
Latitude		43.047		41.700	44.300
Longitude		172.479		171.100	173.700

Table 1-1. Summary Statistics for the national data set used in the height age curve analysis.

The range of the data with respect to Stems per kilometre, Age and Mean top height is also illustrated in figures 1-1, 1-2, and 1-3 below:

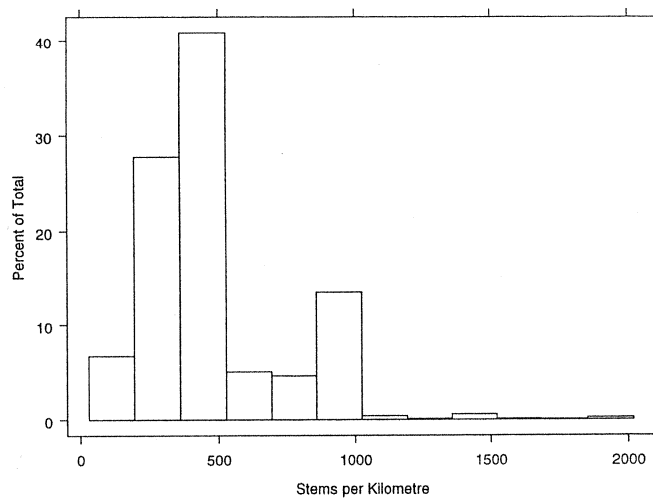


Figure 1-1. Distribution of Mean Top Height measurements versus Stems per Kilometre.

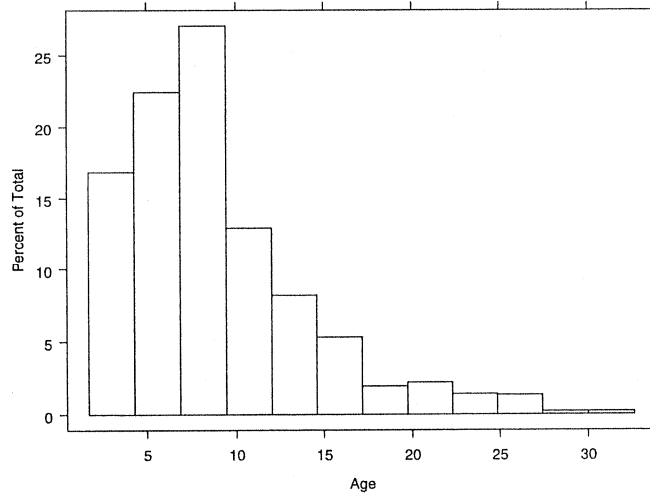


Figure 1-2. Distribution of Mean Top Height measurements versus Age of the shelterbelt (years).

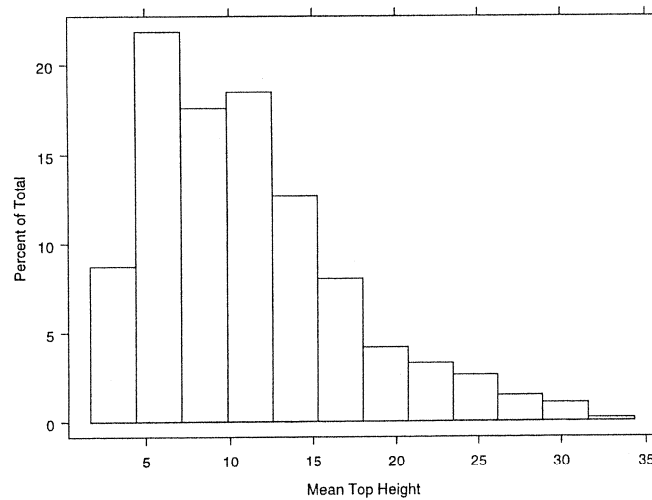


Figure 1-3. Distribution of Mean Top Height measurements (metres).

1.3 Exploratory Data Analysis:

The initial analysis involved plotting mean top height over age for the data at three levels. Firstly the relationship was looked at nationally, secondly the plots within the North and South Islands were examined and finally the plots within each of nine regions were explored. This was performed in order to determine the amount of variation apparent at each level and subsequently which level the models would be applied at. The scatterplots representing the growth information at each level are pictured in figures 1-4, 1-5 and 1-6 below.

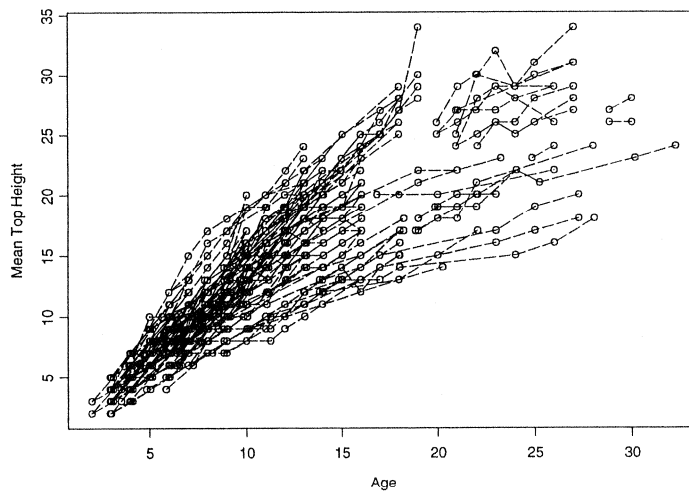


Figure 1-4 Mean Top Height versus the Age (years) of the Shelterbelts for all plots in the national data set.

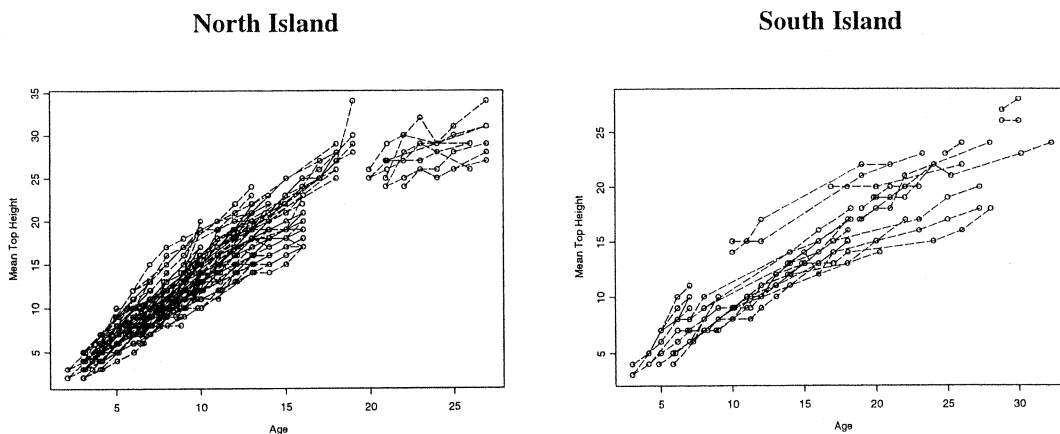


Figure 1-5. Mean top height versus Age (years) of the Shelterbelts for the North and South Islands.

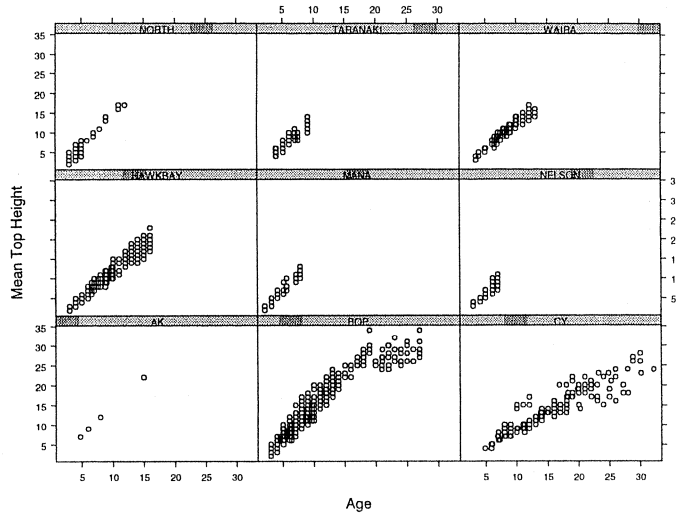


Figure 1-6 Mean Top Height versus Age (years) for different regions across New Zealand. Where AK=Auckland, BOP=Bay of Plenty, CY=Cantabury, HAWKBAY=Hawkes Bay, MANA=Manawatu, NORTH=Northland, WAIRA=Wairarapa

Figures 1-4, 1-5 and 1-6 displayed great variability in Mean Top Height at all levels of the data which appeared to increase with age. The relationship between Mean Top Height and age differed in both shape and magnitude across islands and regions, however this variability in shape and magnitude was not found to be statistically significant across all regions(see section 1.5.1). This variability suggested that a flexible function should be used to model the change in mean top height over time.

1.4 Modelling Methodology:

1.4.1 Previous Work:

The previous height age work for Shelterbelts resulted in a Bertalanffy-Richards Model relating Mean Top Height and Age (Auclair et al, 1991):

$$\text{Mean Top Height} = a \cdot (1 - \exp(-b \cdot (\text{Age} - t_0)))^{1/c}$$

where a is an asymptote parameter, b is a rate parameter, t₀ is the age of zero size, c is the shape parameter and age refers to height adjusted age.

This equation was fitted separately for the North and South Island (using a data set only 15% of the size of the existing data set) which resulted in a different set of parameters for each. A similar approach using dummy variables was also performed for the current data set (see section 1.5.2).

1.4.2 Model Specification:

The specification of the model commenced by re-fitting the original Bertalanffy-Richards model using standard Nonlinear regression for both the North and South Islands with satisfactory results. The North Island model was based on considerably more data than the South island (1558 observations and 219 plots compared to 217 observations and 40 plots respectively) which was reflected in the standard errors around the parameter estimates.

North Island Results:

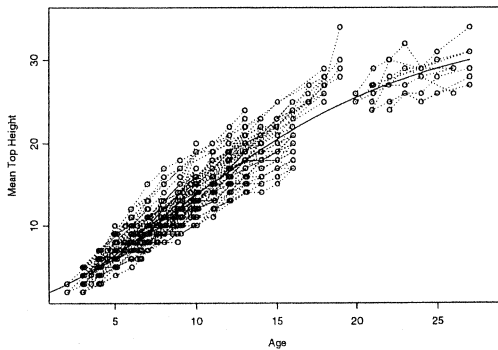


Figure 1-7. Re-fitted growth curve for the North Island data.

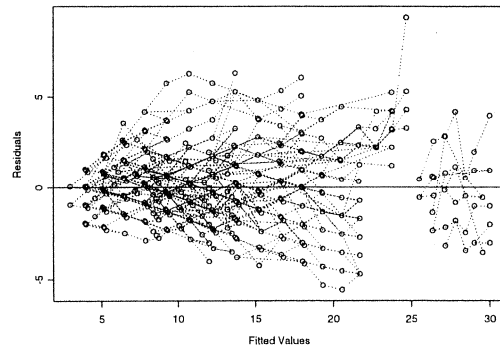


Figure 1-8. Residual Plot for the re-fitted model for the North Island data.

South Island Results:

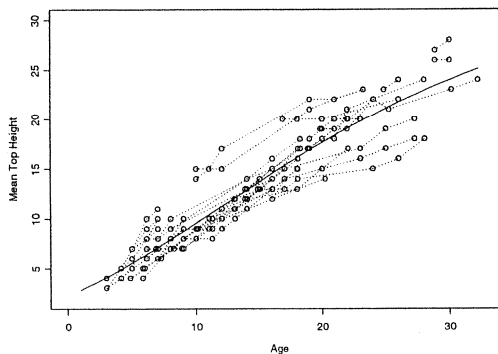


Figure 1-9. Re-fitted growth curve for the South Island data.

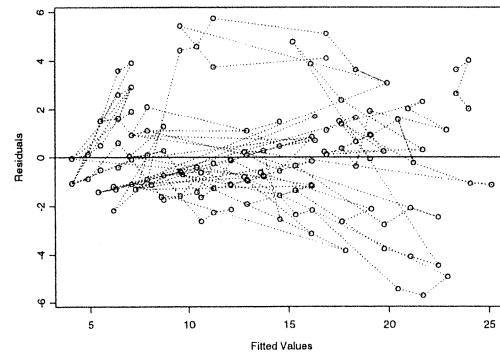


Figure 1-10. Residual Plot for the re-fitted model for the South Island data

The output for the re-fitted North and South island models:

Parameter	NI	SI	SE (NI)	SE (SI)
a	34.5597	35.3098	1.2818	11.0592
b	0.1067	0.0535*	0.0117	0.0394
t0	-7.5437*	-9.3118*	4.1278	17.2429
c	0.1810*	0.3376*	0.1196	0.6339

Table 1-2. Parameter Estimates and corresponding standard errors for the North Island (NI) and South Island (SI) respectively, where *= not statistically significant at the 95% level.

Fit statistics for the re-fitted models:

North Island:	Regression SS/Total SS	=	98.20%
	Residual Mean Square	=	2.84407
South Island:	Regression SS/Total SS	=	97.98%
	Residual Mean Square	=	3.69933

The re-fitted Mean top Height-Age curves for both the North and South islands differed in both their shape and magnitude (figs.1-7, 1-9) but did appear to reflect the relationship between Mean Top Height and age for the respective sets. This was suggested by their corresponding residual plots and fit statistics. For example, 90% of the residuals were in the range [-2.81, 2.80] (m) for the North island, while 90% of the residuals were in the range [-2.65, 3.85] (m) for the South Island.

Two of the four parameters for the north island model and three of the four parameters for the south island model were not statistically significant at the 95% level. It was considered this may have been due to the appropriateness of the nonlinear least squares fitting procedure given that fitting the parameters using restricted maximum likelihood (within a nonlinear mixed modelling framework) resulted in all parameters being statistically significant (see section 1.5.2).

The form and subsequent performance of the re-fitted models was compared with various alternatives. The Bertalanffy-Richards model was compared with a Generalized Additive Model (GAM) and various other standard growth curves. The GAM did not drastically improve the fit but improved the RMS to 2.797 and 3.6213 for the North and South Islands respectively and used equivalent degrees of freedom. The resulting GAM growth curves are shown in figures 1-11 and 1-12.

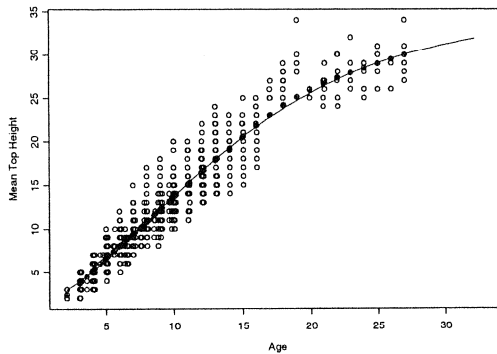


Figure 1-11. Growth curves produced using GAM's (bold points) and the Bertalanffy-Richards (solid line) over the raw North Island data.

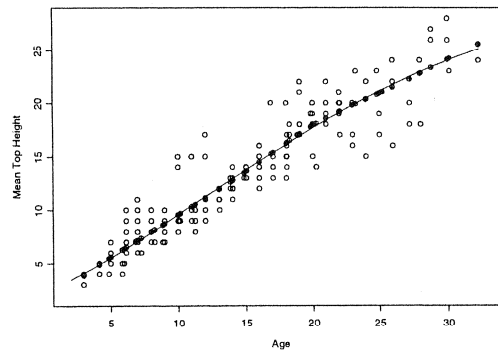


Figure 1-12. Growth curves produced using GAM's (bold points) and the Bertalanffy-Richards (solid line) over the raw South Island data..

Given that a GAM barely improved the fit of the Bertalanffy-Richards model (fitted using Nonlinear regression) the Bertalanffy Richards model still seemed suitable. Nevertheless the model form was further examined by fitting additional growth model forms including the Schumacher, Weibull and the Von-Bertalanffy curves. It was found that the Bertalanffy-Richards produced the best results overall (with reference to their respective residual mean squares) with no convergence problems encountered. Therefore it was decided that the Bertalanffy-Richards model would continue to be used.

1.5 Nonlinear Mixed Models:

The fitting of the Bertalanffy-Richards model was also undertaken using Nonlinear Mixed models. Nonlinear mixed models were used given that the presence of serial correlation within the errors was considered highly likely (and was indeed found to exist). Therefore, this alternative model framework which recognises and incorporates this error information was fitted and compared with the Bertalanffy-Richards model fitted using standard Nonlinear regression.

It was suggested that regional Mean Top Height equations models may be more appropriate (data permitting) than models fitted to the North and South islands. Therefore models at both the island and regional levels were fitted and evaluated. Specific regional parameters were not all found to be statistically significantly different from one another using either standard nonlinear regression or nonlinear mixed models.

1.5.1 Regional Models:

A separate Bertalanffy-Richards model was fitted to Shelterbelts for each region using standard Nonlinear regression and Nonlinear mixed models (with independent and autoregressive errors of the first order (AR(1))) and the results for the parameters were as follows:

Standard Nonlinear Regression:

Region	a	b	t0	c	N
Northland	29.85	0.1549	-23.02	0.0084*	64
Auckland	57.88	0.042	-0.7037	0.7454	6
Bay of Plenty	32.87	0.140	-25.87	0.007	858
Hawke's Bay	27.95	0.1364	-29.53	0.0054	343
Wairarapa	19.48	0.200	-20.024	0.005	135
Nelson	19.818*	0.197*	-9.10*	0.054*	39
Canterbury	44.798*	0.0264*	1.274*	1.0119*	178

Table 1-7. Parameter Estimates for Nonlinear regression for each region. Note: Taranaki and Manawatu regions did not converge. * = not significant at the 95% level.

Nonlinear Mixed Models with independent errors

Region	a	b	t0	c
Northland	27.8674	0.1544*	-10.509*	0.0472*
Auckland	57.8866	0.0425	-0.7037	0.7454
Bay of Plenty	35.954	0.1159	-8.2667	0.1461*
Hawke's Bay	24.2797	0.1473	-22.5023*	0.0098*
Taranaki	20.7337*	0.1737*	-5.6906*	0.0933*
Wairarapa	20.3412	0.1841	-10.7802*	0.0316*
Nelson	21.0683	0.1797	-5.867	0.1149
Canterbury	23.6074	0.0899	-0.7455*	0.5689*

Table 1-8. Parameter Estimates for Nonlinear mixed model with independent errors for each region. Note: The Manawatu region did not converge. * = not significant at the 95% level.

Nonlinear mixed model with AR(1) errors:

Region	a	b	t0	c
Northland	26.7861	0.1565*	-11.7022*	0.0366*
Bay of Plenty	35.1384	0.1209	-8.6309	0.1305*
Hawke's Bay	24.0039	0.1486	-19.1459*	0.0154*
Taranaki	20.7337*	0.1736*	-5.6906*	0.0933*
Wairarapa	20.3393	0.1838	-10.5137*	0.0334*
Canterbury	24.8039	0.0805	-1.1259*	0.5981*

Table 1-9. Parameter Estimates for Nonlinear mixed model with AR(1) errors for each region. Note: Auckland and Nelson regions did not converge. * = not significant at the 95% level.

Great variation was apparent in both the magnitude of the parameter estimates and their significance in tables 1-7, 1-8 and 1-9 above which was thought to be largely due to the limited amount of data within each region. This often resulted in convergence problems and regional parameter estimates with large standard errors. As a consequence of the variance around the parameter estimates, significant differences were not always evident between regions. This led to the conclusion that models at the island level were more appropriate. Perhaps given additional data regional models may have been justified.

Overall, it was noted that the Nonlinear mixed models usually produced slightly different parameter estimates to those obtained using Nonlinear regression. For example, Nonlinear mixed models produced slightly higher asymptote parameter estimates for the Bay of plenty, Wairarapa and Nelson regions and slightly lower estimates for the Northland and Hawkes Bay regions. A large difference was seen in the parameter estimates for the Canterbury region, however using Nonlinear regression the asymptote parameter was not significant at the 95% level. On average the Nonlinear mixed models found less parameters to be significant at the 95% level which may be attributed to the alternative (more appropriate) fitting procedure.

1.5.2 North and South Island Models

Standard Nonlinear Regression:

The island specific models involved fitting the original Bertalanffy-Richards model with dummy parameters (b1 and b2) in order to obtain different asymptote and rate parameters for each island. The shape and t0 parameters were fitted without dummy parameters given the small differences in these parameters for the two islands and the correlation amongst all four parameters. The model was as follows:

$$\text{Mean Top Height} = (a+b1*\text{indic})*(1-\exp(-(b+b2*\text{indic})*(age-t0))))^{**1/c}$$

where a, b, t0 and c are Bertalanffy-Richards parameters, b1 and b2 are the dummy parameters, and indic = dummy variable which indicates the island the plot is situated in (0=North Island, 1=South Island).

The dummy parameter approach results in the a parameter corresponding to the North island asymptote, while the South island asymptote corresponds to the a+b1 value. Similarly the rate parameter is simply b for the North island, while the South island rate is equal to b+b2. This dummy parameter model was fitted using standard Nonlinear regression and the results were as follows:

Parameter	Estimate	Standard Error
a	37.7544	1.7847
b1	-11.2634	1.1576
b	0.0840	0.0103
b2	0.0005*	0.0026
t0	-3.3983	1.5526
c	0.3872	0.1095

Table 1-3. Parameter estimates and corresponding standard errors for the Dummy parameter model fitted for the national data set. * = not significant at the 95% level.

The corresponding fit statistics for the model are:

$$\begin{aligned} \text{Regression SS/Total SS} &= 98.14\% \\ \text{Residual Mean Square} &= 2.9973 \end{aligned}$$

It is worth noting that this approach failed to find the b2 variable to be significant at the 95% level suggesting that the rate parameter was not significantly different between the two islands (remembering that this parameter indicates the rate at which

the asymptote is approached rather than a growth rate itself) however this was not seen when Nonlinear mixed models were used.

Nonlinear Mixed Models:

Using a Nonlinear mixed model involved fitting the following equation:

$$\text{Mean Top Height} = (a + ra + b1 * \text{indic}) * (1 - \exp(-(b + b2 * \text{indic}) * (\text{age} - t0)))^{**} (1/c)$$

where a, b, t0, c are model parameters, b1 and b2 are dummy parameters, indic=indicates the location of the plot (north island=0, south island=1) and ra indicates a random effect.

The random effect in the model is an estimate of the difference between the mean parameter estimate and that of a particular plot. This is obtained for every plot and reflects the variance of the distribution about parameter a. Parameter a was specified as a random effect due to its variation both within and between islands. Parameter b was also initially specified as a random effect however, when fitted it was not found to significantly improve the fit of the model (when the AIC statistics were compared). The equation written above was fitted assuming both independent errors and autoregressive errors of the first order. The results were as follows:

Nonlinear mixed model with identity errors:

Parameter	Estimate	Standard Error
a	40.3587	1.7784
b1	-15.1382	1.7549
b	0.0737	0.0063
b2	0.0141	0.0025
t0	-1.7729	0.4829
c	0.5074	0.0540

Table 1-4. Parameter estimates and corresponding standard errors for the Nonlinear mixed model with dummy variables assuming independent errors.

Fit Statistics: -2 Res Log Likelihood = 4510.593, AIC = -2257.30

Nonlinear Mixed model with AR(1) errors:

Parameter	Estimate	Standard Error
a	43.0308	2.7426
b1	-17.5237	2.5863
b	0.0640	0.0077
b2	0.0175	0.0036
t0	-1.1265	0.4797
c	0.5913	0.0657

Table 1-5. Parameter estimates and corresponding standard errors for the Nonlinear mixed model with dummy variables assuming AR(1) errors.

Fit Statistics: -2 Res Log Likelihood = 4259.548, AIC = -2132.77

The Nonlinear mixed model fitted with autoregressive errors of the first order resulted in a better model fit and larger standard errors than those obtained assuming independent errors around the parameter estimates. The larger standard errors are to be expected given that serial correlation tends to bias the estimator of the variance-covariance matrix and produce underestimates. Therefore, recognising the serial correlation results in more realistic (and invariably larger) variance estimates around the parameters. The Nonlinear mixed models consistently indicated greater (and thus significant) differences in their asymptote and rate parameters between the two islands with larger estimates for the shape parameters. In addition they invariably produced a better fit which is indicated in their AIC statistics and in the resulting residuals (table 1-6).

Graphical Comparison of the Standard Nonlinear Regression and Nonlinear mixed model Results:

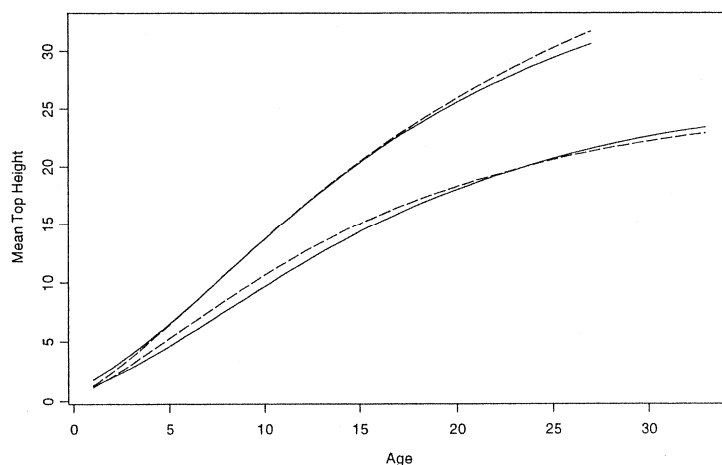


Figure 1-13. Growth Curves for the North island (larger asymptote) and South island (smaller asymptote). The solid line represents the standard Nonlinear regression line, the broken line represents the Nonlinear mixed model with AR(1) errors.

Summary of the Residuals for the Bertalanffy-Richards Models:

Model	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
NLS	-9.35	-0.83417	0.127185	0.97669	5.391645
Mixed (ID)	-9.03005	-0.91101	0.054986	1.088994	5.58864
Mixed (AR1)	-9.144	-0.90277	0.052204	1.097232	5.479031

Table 1-6. Summary statistics for the residuals produced by the three Mean Top Height-Age models. NLS= Nonlinear least squares model, Mixed(ID)=Nonlinear mixed model with independent errors, and Mixed (AR1)=Nonlinear mixed model with AR(1) errors.

Fit and Residual Plots for the Nonlinear Regression Model

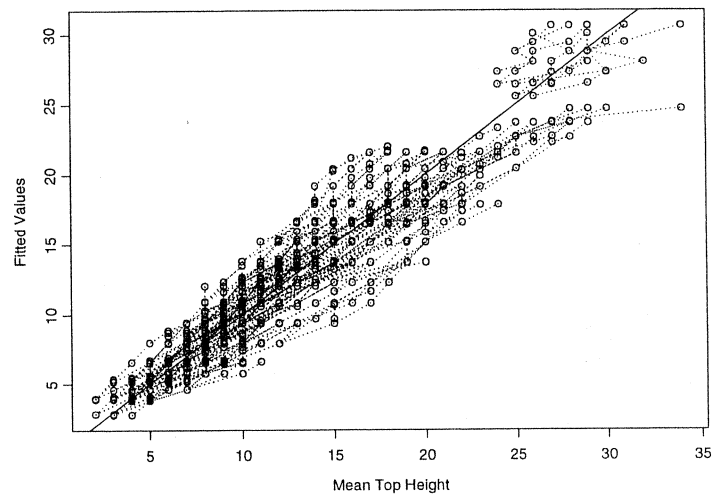


Figure 1-14. Predicted versus Observed values for Mean Top Height for the Standard Nonlinear Regression model.

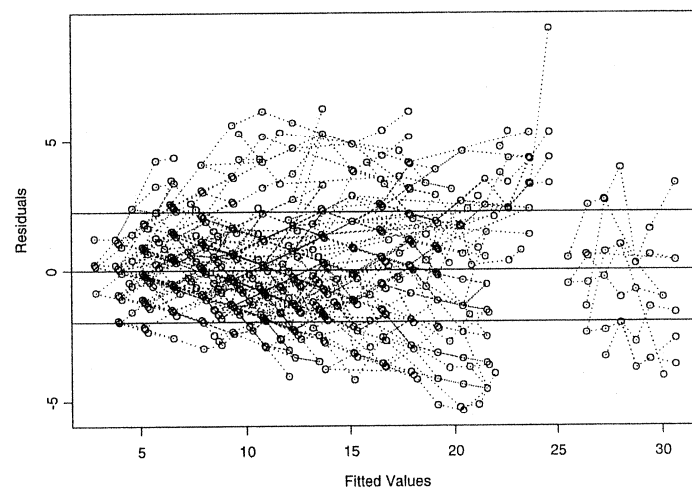


Figure 1-15. Residual Plot for the Nonlinear regression model where the outer horizontal lines encompass 80% of the residuals .

1.6 Conclusions

The Mean Top Height curves for each plot within each region, across region and within North and South islands showed large differences in shape and magnitude. Of the four model forms tested, the Bertalanffy-Richards form appeared to best accommodate these differences. The Bertalanffy-Richards model was fitted to plots within region and the parameter estimates across regions were compared. It was found that in the absence of sufficient information in some regions (and hence large standard errors around the parameters) significantly different parameters were not obtained. Therefore the Mean Top Height/Age curves labelled 'North Island' and 'South Island' were fitted.

The Bertalanffy-Richards model was fitted to the North and South island data using two Nonlinear modelling frameworks: Standard Nonlinear regression and Nonlinear mixed models. It was found that the Nonlinear mixed models resulted in slightly different parameter estimates which resulted in a better fit of the Bertalanffy-Richards model. The Nonlinear mixed models also allowed the serial correlation inherent in the data to be taken into account and correctly modelled via autoregressive errors of the first order.

Part Two: The Basal Area Increment Function

2.1 Abstract:

Various basal area increment models were formulated for shelterbelts with a range of site conditions and pruning treatments. The form of the basal area increment models were established using GAM's, fitted using Nonlinear regression and finally fitted as Nonlinear mixed models. Model fitting was undertaken at the regional level with satisfactory results.

2.2 Data Specifications:

Brief summary statistics for the variables implicated in the modelling are listed in table 2-1, while illustrations of the data available with respect to certain variables are shown in figures 2-1, 2-2, and 2-3 below.

Variable	Mean	S. D	Minimum	Maximum
Age (years)	8.970	4.889	2.080	28.810
Green Crown Length (m)	6.848	4.723	1.400	28.100
Average Taper (cm/m)	2.696	0.539	0.539	5.000
Stems per kilometre	463.028	232.017	50.000	1067.000
Mean Height (m)	10.789	5.933	1.800	29.400

Table 2-1. Data summary for variables used in the basal area increment modelling.

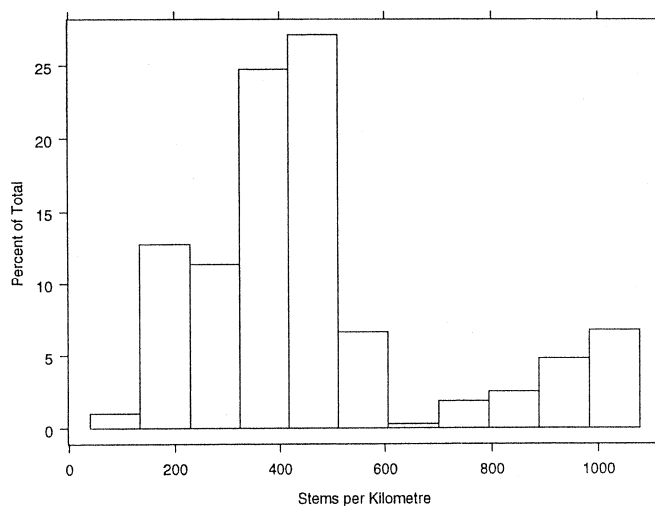


Figure 2-1. Distribution of basal area increments versus Stems per Kilometre.

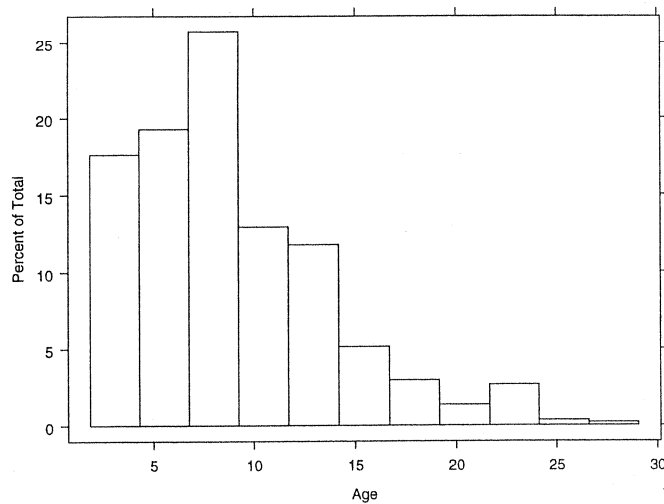


Figure 2-2. Distribution of Basal area increments versus Age of the shelterbelt (years).

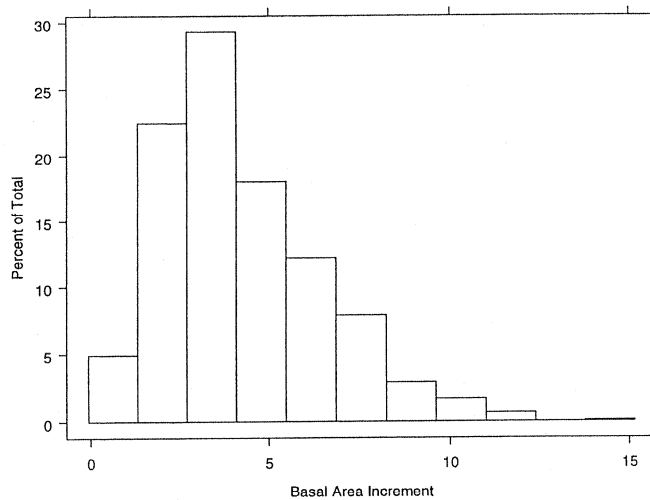


Figure 2-3. Distribution of Basal area increments (m²/km).

2.2 Exploratory Data Analysis

Exploratory Analysis initially involved scatterplots which investigated the relationship between each explanatory variable and the response. This was performed at both the national and regional levels and a sample of these (with smoothed local regression models fitted) at the regional level are illustrated in figures 2-4, 2-5 and 2-6.

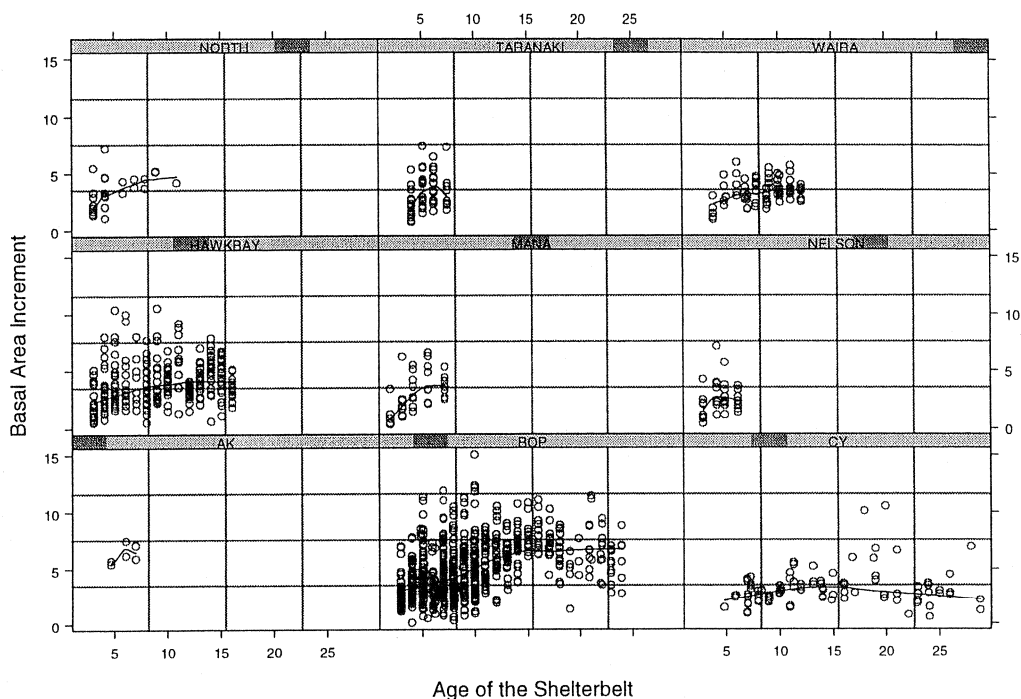


Figure 2-4. Scatterplots of Basal Area increment versus age with a local regression model fitted for each region. AK= Auckland, BOP= Bay of Plenty, CY= Canterbury, HAWKBAY= Hawkes Bay, MANA= Manawatu, NORTH= Northland and WAIRA= Wairarapa, NELSON= Nelson/Marlborough.

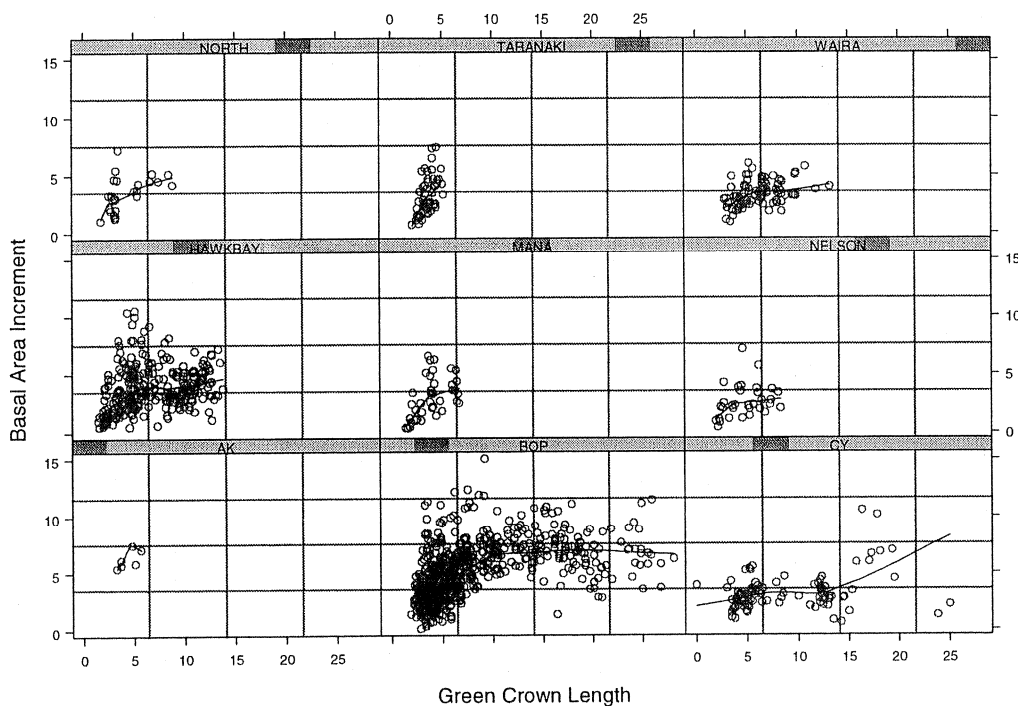


Figure 2-5. Scatterplots of Basal Area increment versus Green Crown length with a local regression model fitted for each region. AK= Auckland, BOP= Bay of Plenty, CY= Canterbury, HAWKBAY= Hawkes Bay, MANA= Manawatu, NORTH= Northland and WAIRA= Wairarapa, NELSON= Nelson/Marlborough.

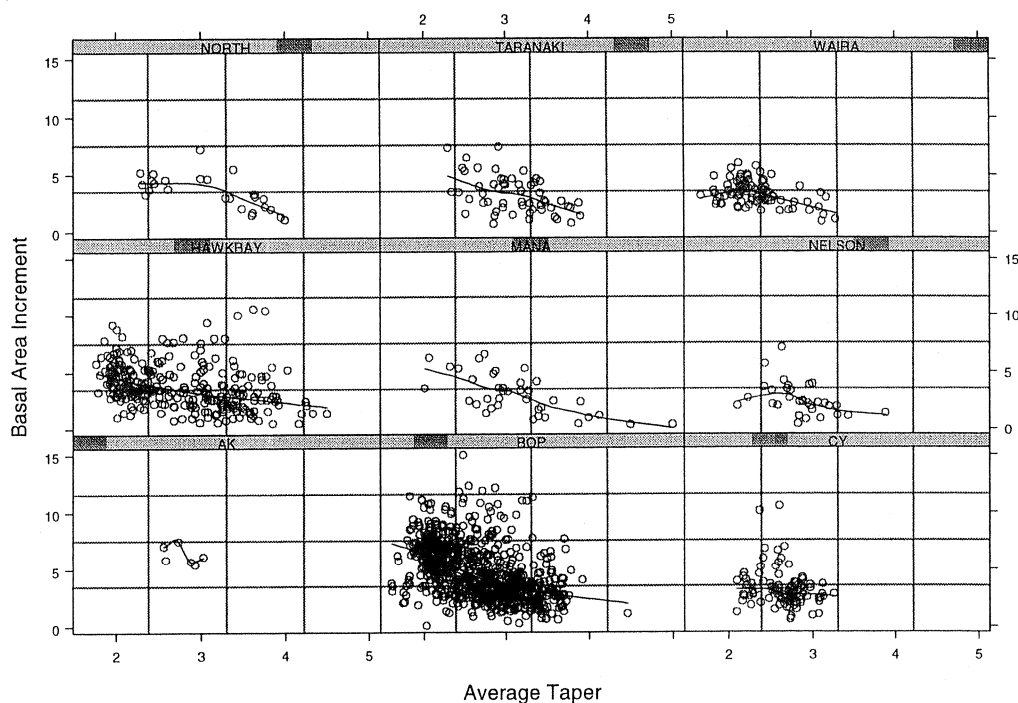


Figure 2-6. Scatterplots of Basal Area increment versus Average Taper with a local regression models fitted for each region. AK= Auckland, BOP= Bay of Plenty, CY= Canterbury, HAWKBAY= Hawkes Bay, MANA= Manawatu, NORTH= Northland and WAIRA= Wairarapa, NELSON= Nelson/Marlborough.

Figure 2-4 shows that the basal area increment increased quickly until reaching approximately age 15 and then levelled off soon after. This relationship varied in both shape and magnitude between the regions. Figure 2-5 shows a similar shaped relationship, between annual increment and green crown length. Again both the shape and magnitude varied between regions. Note that the relationship suggested between increment and green crown length for the Canterbury region may not be reliable given the amount of data available. Figure 2-6 illustrates a moderately decreasing increment with an increase in average taper.

In addition to various scatterplots, the relationships between each variable and basal area increment were explored at the national level using Generalized Additive models. This enabled both the shape and contribution of the relationship between each variable and the response to be determined.

2.4 Generalized Additive Models: Bivariate GAM Fits:

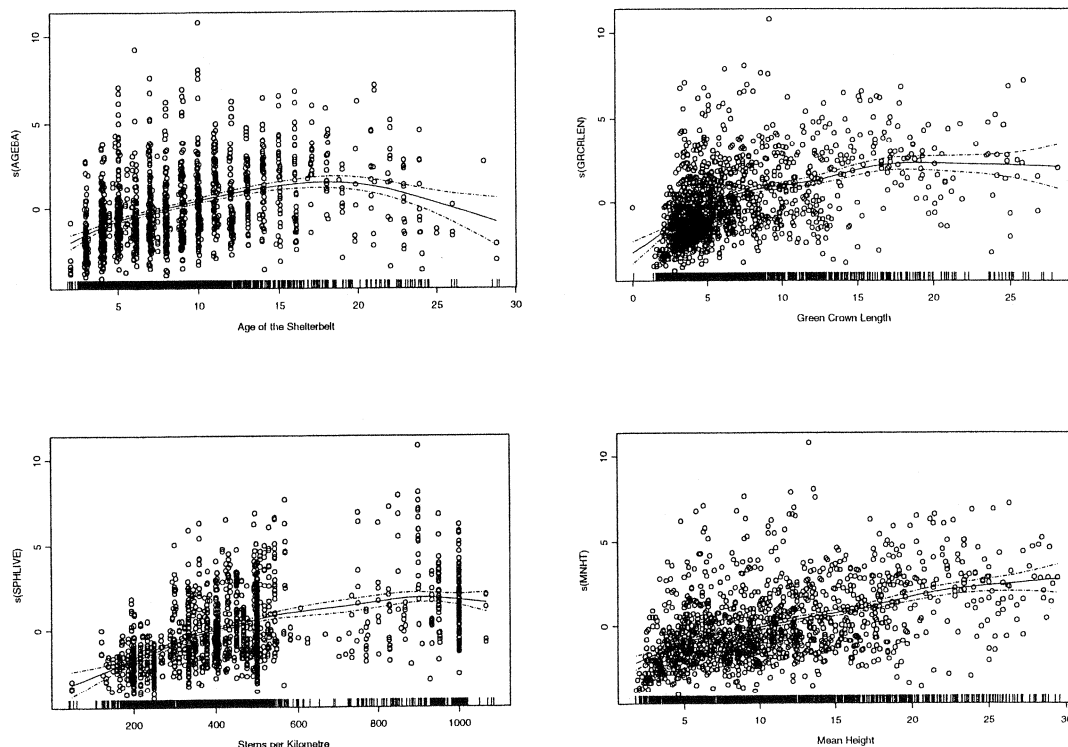
Table 2-2 represents the results of the bivariate fits for each variable and the response. It was found that the stems per kilometre variable explained the largest amount of information which was closely followed by height variables. The variables appear in descending order of the deviance explained.

Variable	Null Deviance	Residual Deviance	% Deviance Explained
Stems per Km	7080.541	5094.558	28.04846
Mean Height	6857.949	5174.432	24.5484
Mean Top height	6378.7	4892.235	23.30357
Green Crown length	7074.844	5428.285	23.27343
Average Taper	6857.949	5646.646	17.66276
Age	7080.541	5897.93	16.70227
Site Index	6885.617	5850.097	15.03888
Mean Crown Height	7074.844	6257.091	11.5586
Altitude	6481.508	5997.074	7.474094

Table 2-2. Results of Bivariate GAM fits for Basal Area Increment. % Deviance Explained is indicative only where %Deviance Explained = 1-(Residual Deviance/Null Deviance)*100.

In addition to the variables listed in table 2-2, previous annual increment, crown length per kilometre and basal area were also considered. For instance previous annual increment was found to explain over 57% of the deviance however it was considered that this variable would potentially bias the results. Crown length per kilometre was also considered however it was found that fitting the two variables separately (green crown length and stems per kilometre) resulted in a better fit. Basal Area was not included as it is often not available to the user but is predicted by the model. Green crown surface area for the stand was also considered as a potential variable however information about the width of the crowns was not available and is also not easily estimated.

The form of the Bivariate GAM Fits for each explanatory variable



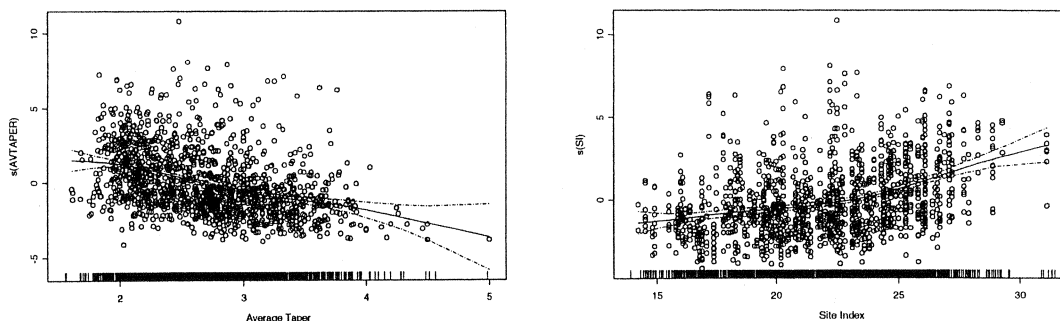


Figure 2-7. GAM plots for the bivariate fits. The x-axis represents the explanatory variable while the y axis represents the fitted smooth function.

Figure 2-7 illustrates the relationships between each variable and the response when fitted separately. Basal Area increment appears to be related similarly to green crown length, stems per kilometre and mean height. For these variables basal area increment appeared to be related in a sigmoidal-type fashion while the increment appeared to peak at about 16 years and then decrease shortly after this point. This increment- age relationship is slightly misleading (upon closer inspection) given that the data containing the 25-30 year age information was almost entirely from the Canterbury region. This lead to an apparent decrease in increment while the relationships within region were of sigmoidal origin and appeared to approach an asymptote.

2.5 Distribution of the Basal Area Increment variable:

Before the increment models were formulated the distribution of the increment variable was investigated. This was undertaken to ensure the assumption of normality or the response (often assumed in regression) was reasonable. It was found that the response was not strictly normally distributed which is illustrated by the density and quantile plots in figures 2-5 and 2-6 and this non-normality was confirmed by the Kolmogorov-Smirnov test.

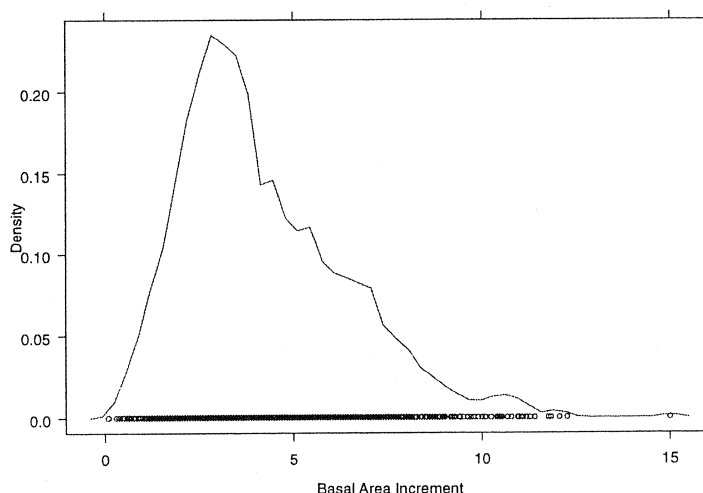


Figure 2-8. Density distribution of basal area increment (m^2/km).

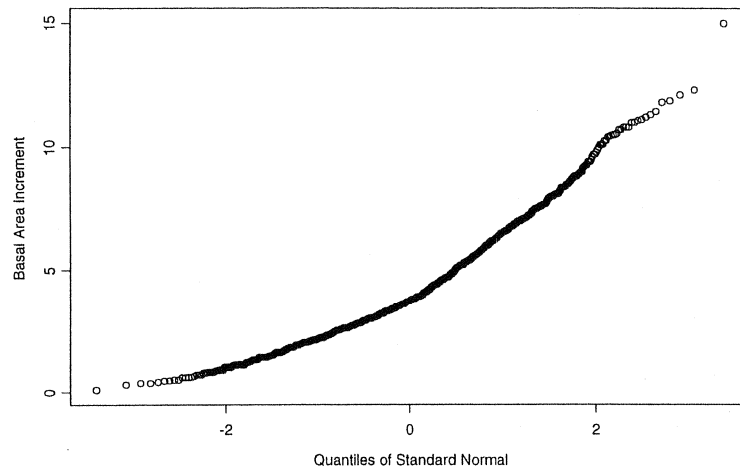


Figure 2-9. Quantile plot for basal area increment.

Figure 2-8 illustrates the skewed distribution of basal area increment while the non-linear curve in figure 2-9 confirms the non-normality of the variable. Given that the annual increment was found to be non-normal, basal area increment was square root transformed. This resulted in an approximately normal distribution (see figures 2-10 and 2-11) however the transformed distribution was still considered non-normal when the Kolmogorov-Smirnov test was undertaken.

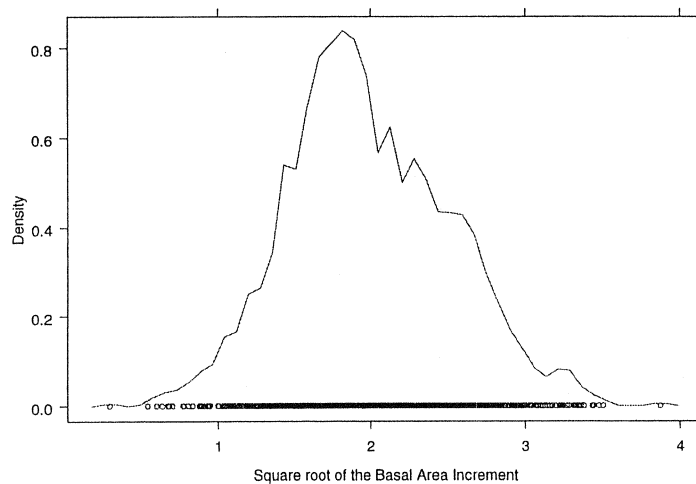


Figure 2-10. Density distribution for the square root transformed basal area increment.

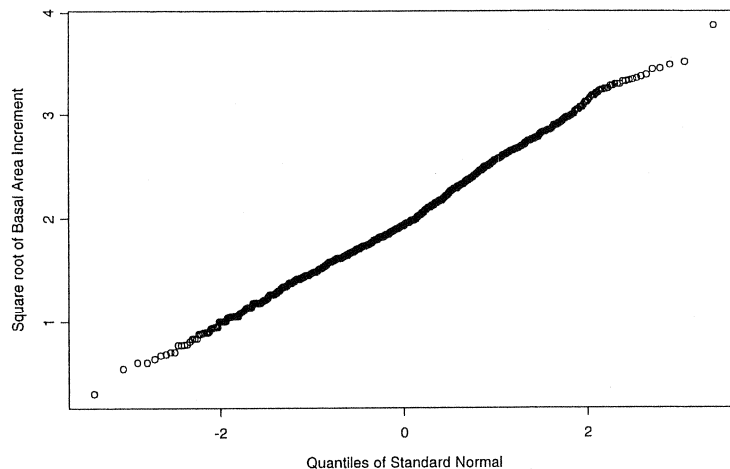


Figure 2-11. Quantile plot of square root transformed basal area increment.

Figures 2-10 and 2-11 show a more symmetrical distribution which corresponded to an approximately linear quantile plot suggesting a normal distribution. Due to these results, the form of the response was considered when the modelling was undertaken. This involved fitting all of the formulated models with and without the transformed response. It was found that only small differences resulted when either distributional form of the response was used. This illustrates the robustness of the normality assumption for the formulated models.

2.6 Detection of Collinearity / Concurvity

As a part of the modelling process, the correlation between all of the candidate variables was investigated in order to avoid multicollinearity. As with linear regression, when two explanatory variables are highly correlated (and both used in the model to estimate the response) the resulting fit is often unstable and commonly produces faulty coefficients and/or large standard errors around the parameter estimates.

As expected, mean height and mean top height were very highly correlated (correlation coefficient of 0.99) and while it was found that mean height explained more variation than mean top height about the increment it was also too highly correlated with age and green crown length to be included in the variable selection procedure. It was also realised that green crown length and age were very highly correlated (correlation coefficient of 0.90) however it was desirable to have both variables in the model since dropping either variable had detrimental effects on model performance (which are discussed later). Therefore given that mean height increases with age (and this is the reason the two variables are correlated) an interaction type term was included in the basal area increment modelling which consisted of mean height/age.

2.7 Variable Selection

Stepwise variable selection was performed using GAM's and the Akaike's Information Criterion (AIC), however not all variables were available for selection. For example, since mean crown height was used to formulate green crown length this variable was not supplied for selection. Average taper was also omitted from the formulation of one set of models but was included in another set of models for comparison purposes. The variable selection procedure resulted in smooth terms for all variables supplied (see Model 1) which suggested that each variable should be related to both the untransformed and the square root transformed response in a nonlinear fashion to obtain the best fit.

2.8 Model Specification

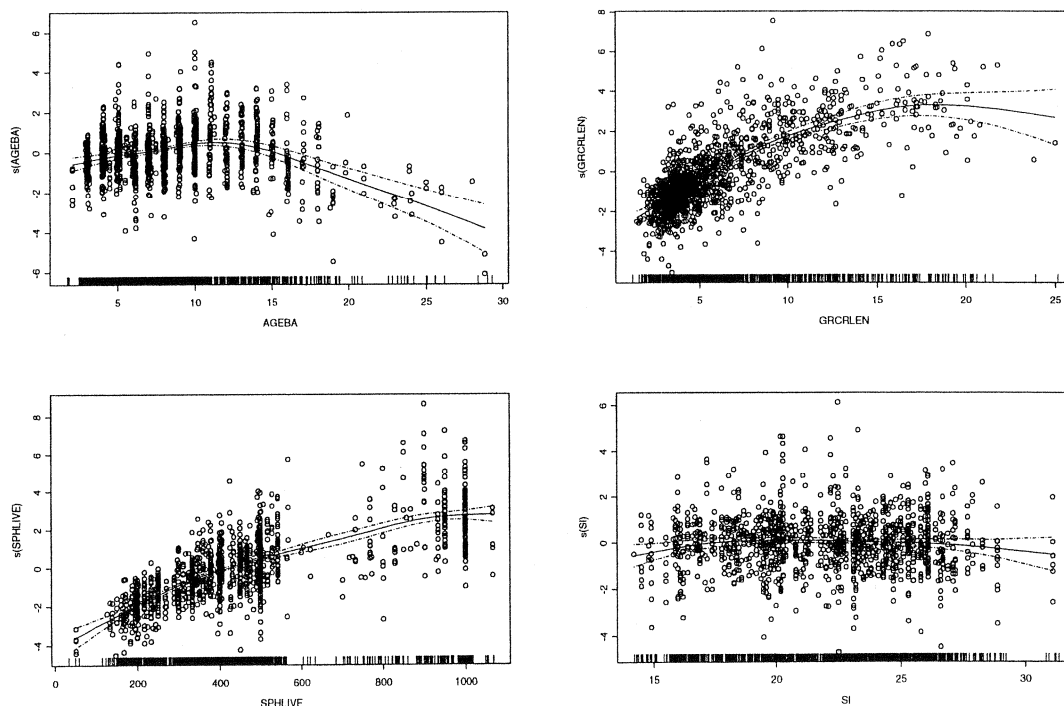
Model 1. Basal Area Increment = $s(\text{age}) + s(\text{green crown length}) + s(\text{stems per km}) + s(\text{average taper}) + s(\text{site index}) + s(\text{altitude}) + s(\text{mean height/age})$

(where s represents a smooth function fitted using a spline.)

Coefficients:

(Intercept)	$s(\text{AGEBA})$	$s(\text{GRCLEN})$	$s(\text{SPHLIVE})$	$s(\text{mnht/age})$	$s(\text{SI})$
-9.103343	-0.01532186	0.3530378	0.005900572	3.387307	-0.0004898435
	$s(\text{ALTITUDE})$	$s(\text{AVTAPER})$			
	-0.0002610189	1.631891			

Partial Regression plots for Model 1:



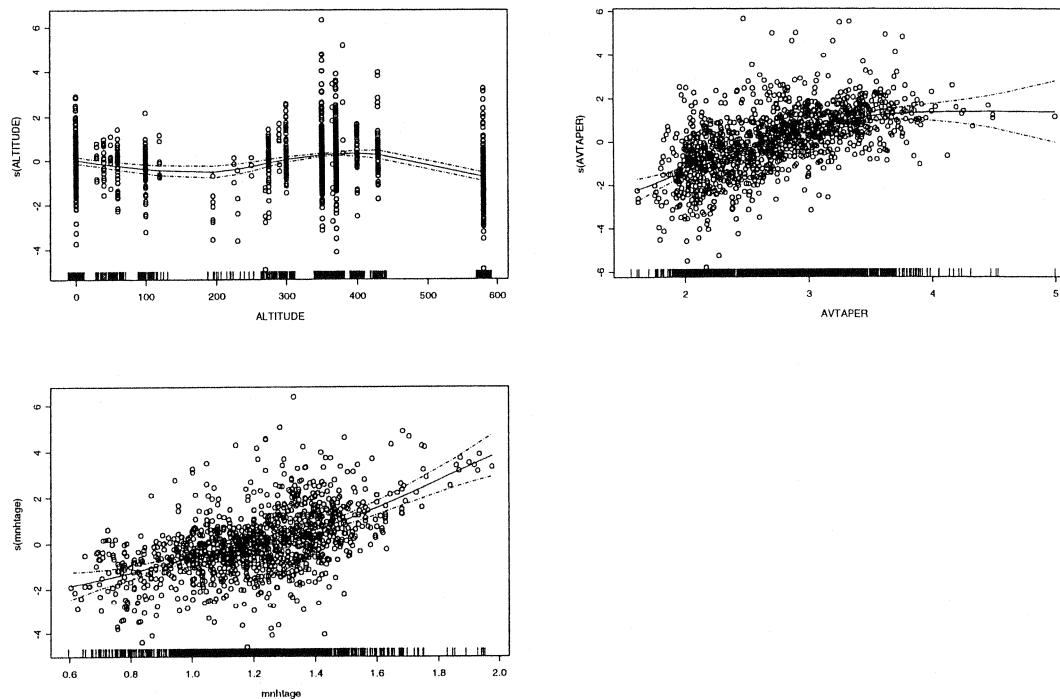


Figure 2-12. Partial regression plots for GAM: Model 1, where AGEBA=Age, GRCRLN=Green Crown Length, SPHLIVE=Stems per Km, AVTAPER=Average Taper, SI=Site Index, mnhtage=Mean height/Age

Fit and residual plots for Model 1:

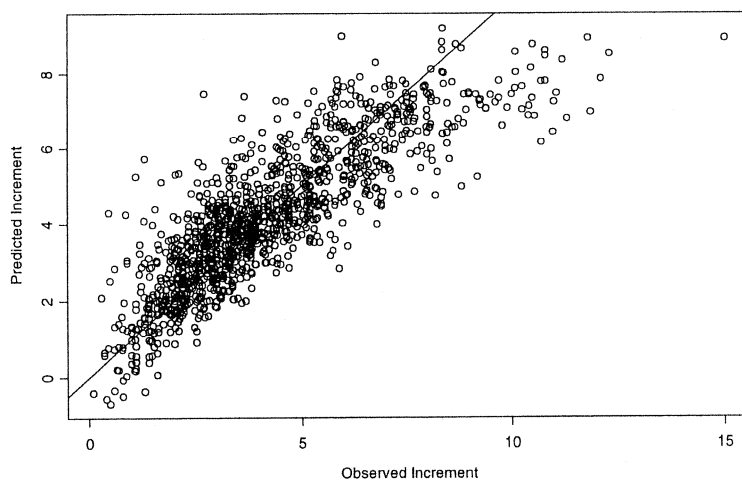


Figure 2-13. Plot of observed versus predicted values for Model 1.

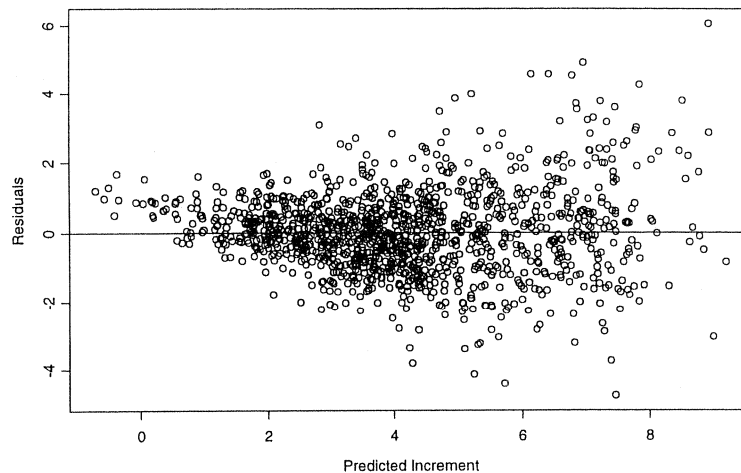


Figure 2-14. Residual plot for Model 1.

Fit statistics for Model 1:

AIC=1849.197, RMS = 1.45, Null deviance= 6102.423, Residual Deviance=1764.893
(Deviance Explained=71.07%)

Discussion:

The modified shape of the partial regression plots (figure 2-12) when compared to the bivariate fits (figure 2-7) confirm (as expected) that the variables are not purely additive but are related to one another to some extent. The pointwise standard error bars around the curves also suggest that the relationships are non-linear in every case. The fit and residual plots show a reasonable fit overall, however the model tends to underpredict basal area increment at large observed values.

It was considered that average taper should not perhaps be included in the modelling on the basis that average taper was a site effect which would be taken into account when nonlinear mixed models were used (see section 2.3.2). In addition it was thought that the site index estimates may be unreliable given that they were formulated using height age curves derived from forest data and were not necessarily based on shelterbelt information. Therefore two sets of models were formulated: one model set contained both the average taper and/or site index variables, and the other set did not. The model specification process for both sets of models involved a Generalized additive model (GAM) to indicate the form of the partial additive relationships between the response and the explanatory variables. This was followed by the substitution of known functions for the smooth splines. The results for the first set of models (absent of average taper information) were as follows:

Model 2: Basal area Increment = s(Age)+s(Green Crown Length)+s(Stems per Km)+s(mean height / age)

Coefficients:

(Intercept)	s(AGEBA)	s(GRCRLN)	s(SPHLIVE)	s(mnhtage)
-2.056658	-0.1375251	0.304319	0.004517932	2.770525

Partial Regression plots for Model 2:

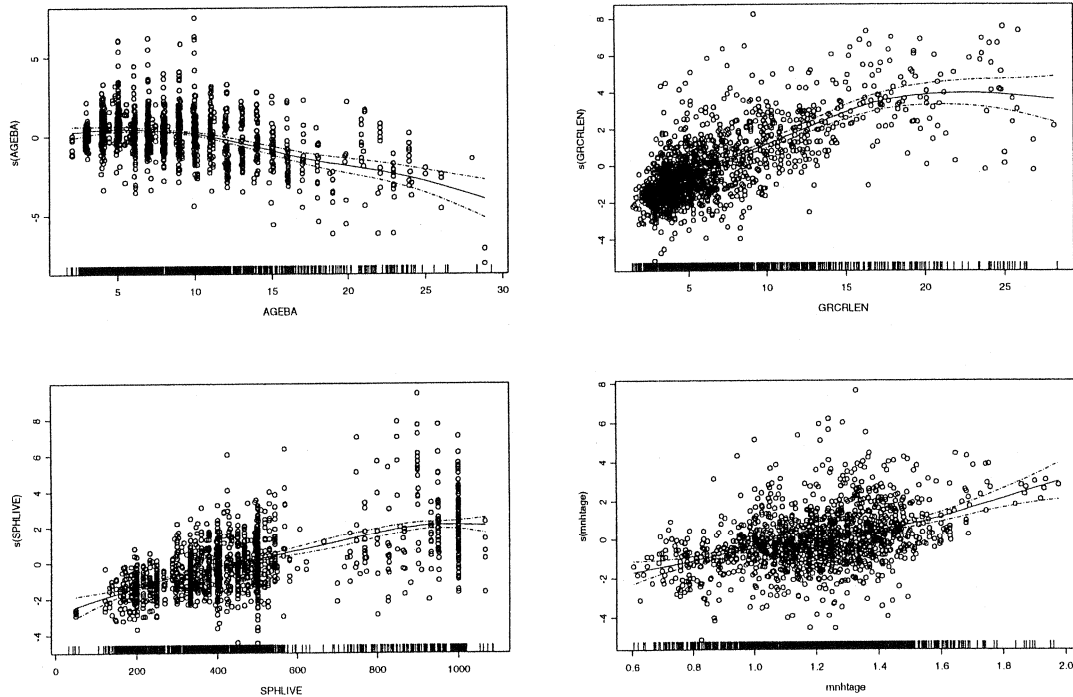


Figure 2-15. Partial regression plots for Model 2, where AGEBA=Age, GRCRLN=Green Crown Length, SPHLIVE=Stems per Km, AVTAPER=Average Taper, SI=Site Index, and mnhtage=Mean height/Age

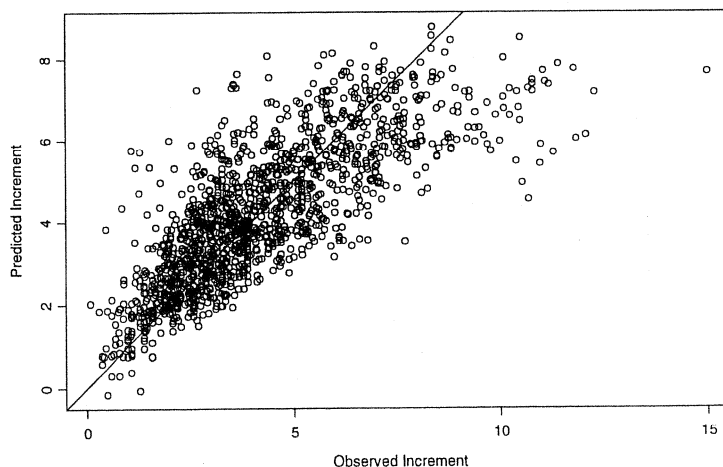


Figure 2-16. Observed versus predicted values for Model 2.

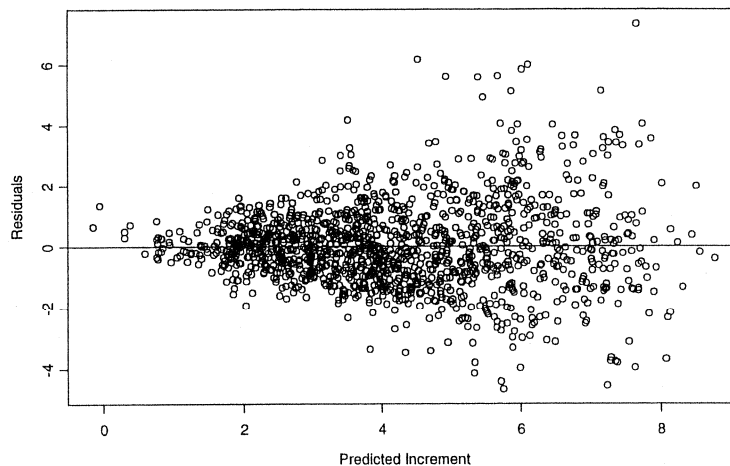


Figure 2-17. Residual plot for Model 2

Fit statistics for Model 2:

AIC = 2708.446, RMS = 1.93 , Null Deviance = 6703.323, Residual Deviance = 2642.758 (Deviance Explained = 60.57%)

Discussion:

The partial regression plots for the Generalized Additive Model (figure 2-15) suggested a decreasing function for age, a sigmoidal-type model for green crown length (note the absence of an inflection point) and a moderately increasing function for stems per km and mean height/age. The results show that once again the model is under-predicting the observed increment for the larger increment values (see figure 2-16) which is also evident in the residual plot (see figure 2-17).

The fit statistics show that the previous GAM (Model 1, which included site index and average taper) performed better given that it explained 10.5% more of the deviance in the model. In addition the AIC statistic which is penalised for the number of variables used in the model increased by 46% suggesting a substantially worse fit even after considering that less variables were fitted in Model 2.

Linear functions were substituted for the age, stems per kilometre and mean height/age variables, while a Von- Bertalanffy model was fitted for the green crown length variable. Various nonlinear models were fitted for the age variable however little benefit was found given the extra parameters required to be estimated. The results were as follows:

Model 2b: Basal Area Increment: $a \cdot \text{age} + (a2 \cdot (1 - \exp(-k \cdot (\text{green crown length} - t0)))) + r \cdot \text{stems per km} + s \cdot (\text{mean height} / \text{age})$

Coefficients:

A	A2	K	T0	R	S
-0.132861720	4.171088135	0.094196092	6.863862458	0.004456547	3.066846023

Fit and Residual Plots for Model 2b:

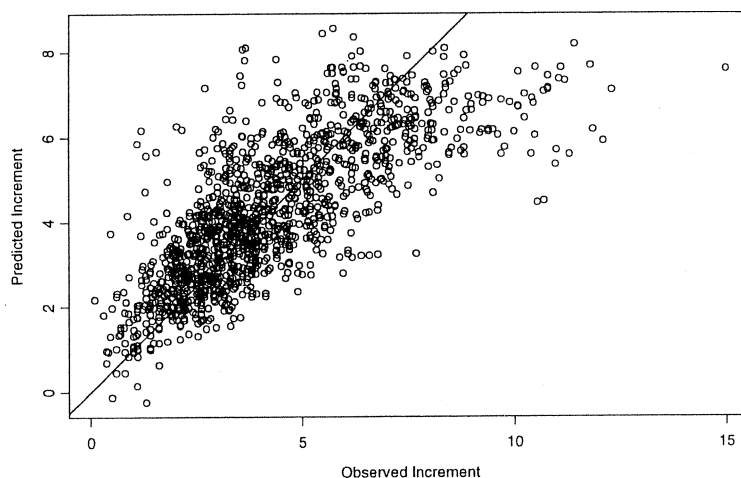


Figure 2-18. Observed versus predicted values for Model 2b.

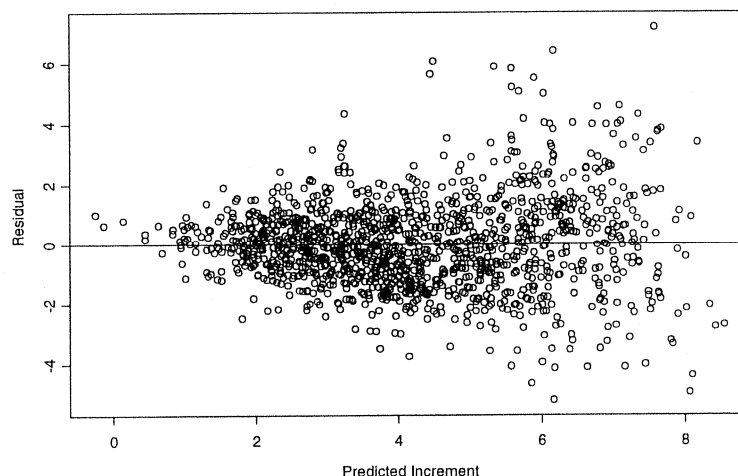


Figure 2-19. Residual plot for Model 2b.

Fit Statistics for Model 2b:

RMS = 2.03, Regression SS/Total SS * 100 = 91.06%

Discussion:

Model 2b has also shown to underpredict the larger increment values however, the performance of model 2b appeared to closely resemble the performance of the GAM

(RMS of 2.03 (Model 2b) as compared to 1.93 for the GAM (Model 1)) despite the substitution of three linear functions for smooth functions. The close performance of models 2 and 2b was also indicated in the coefficients for models 2 and 2b. The coefficients for age and stems per km were very similar when both modelling approaches were used.

In addition to Models 1, 2 and 2b, it was considered that a model similar to that formulated for Douglas Fir (Fight et al, 1995) may be useful for Pinus radiata shelterbelts given that the relationships may have been similar. This was fitted and is represented as model 3.

Model 3: Basal Area Increment = $(a/(1+b*age))*((1-\exp(-c*stems\ per\ km*green\ crown\ length))^d)$

Coefficients:

A	B	C	D
9.9052	-0.1879	0.00029	1.02322

Fit and Residual Plots for Model 3:

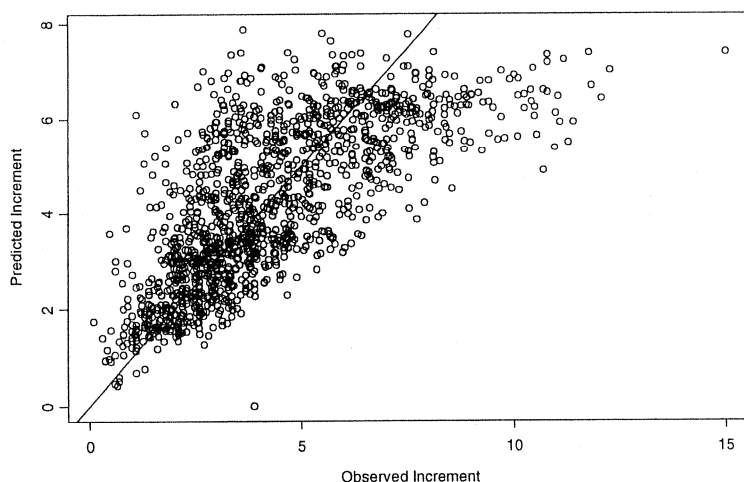


Figure 2-20. Observed versus predicted values for Model 3.

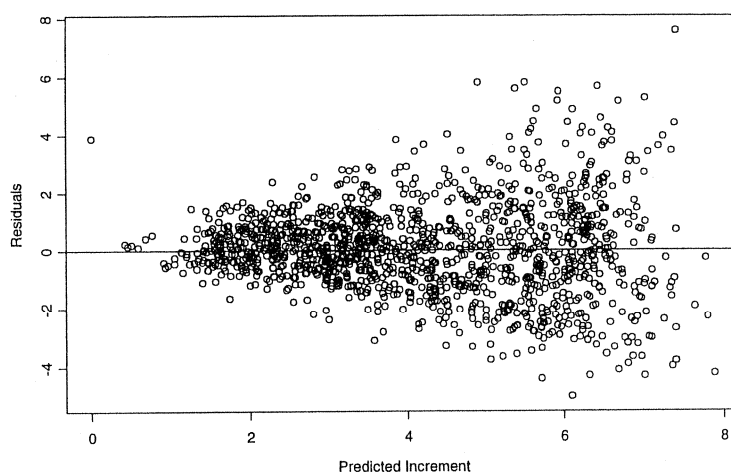


Figure 2-21. Residual plot for Model 3.

Fit Statistics for Model 3:

RMS = 2.37, Regression SS/Total SS = 89.5%

Discussion:

The results obtained from Model 3 indicate no improvement on the previously formulated additive model (Model 2b) despite the use of the similar variables and two less parameters. The fit statistics were slightly less for Model 3 (when compared to Model 2b) and no convergence problems were found with either model. Model 3 concludes the model set absent of the average taper variable.

Model set where average taper information is included:

Model 4: Basal area Increment = $s(\text{Age}) + s(\text{Green Crown Length}) + s(\text{Stems per Km}) + s(\text{Average Taper}) + s(\text{Mean Height / Age})$

Note: Site Index was initially included in model 4 however it was not significant at the 95% level.

Coefficients:

(Intercept)	s(AGEBA)	s(GRCLEN)	s(SPHLIVE)	s(AVTAPER)	s(Mnht/age)
-11.28	-0.008917368	0.3039957	0.005892332	2.003137	4.462683

Partial Regression Plots for Model 4:

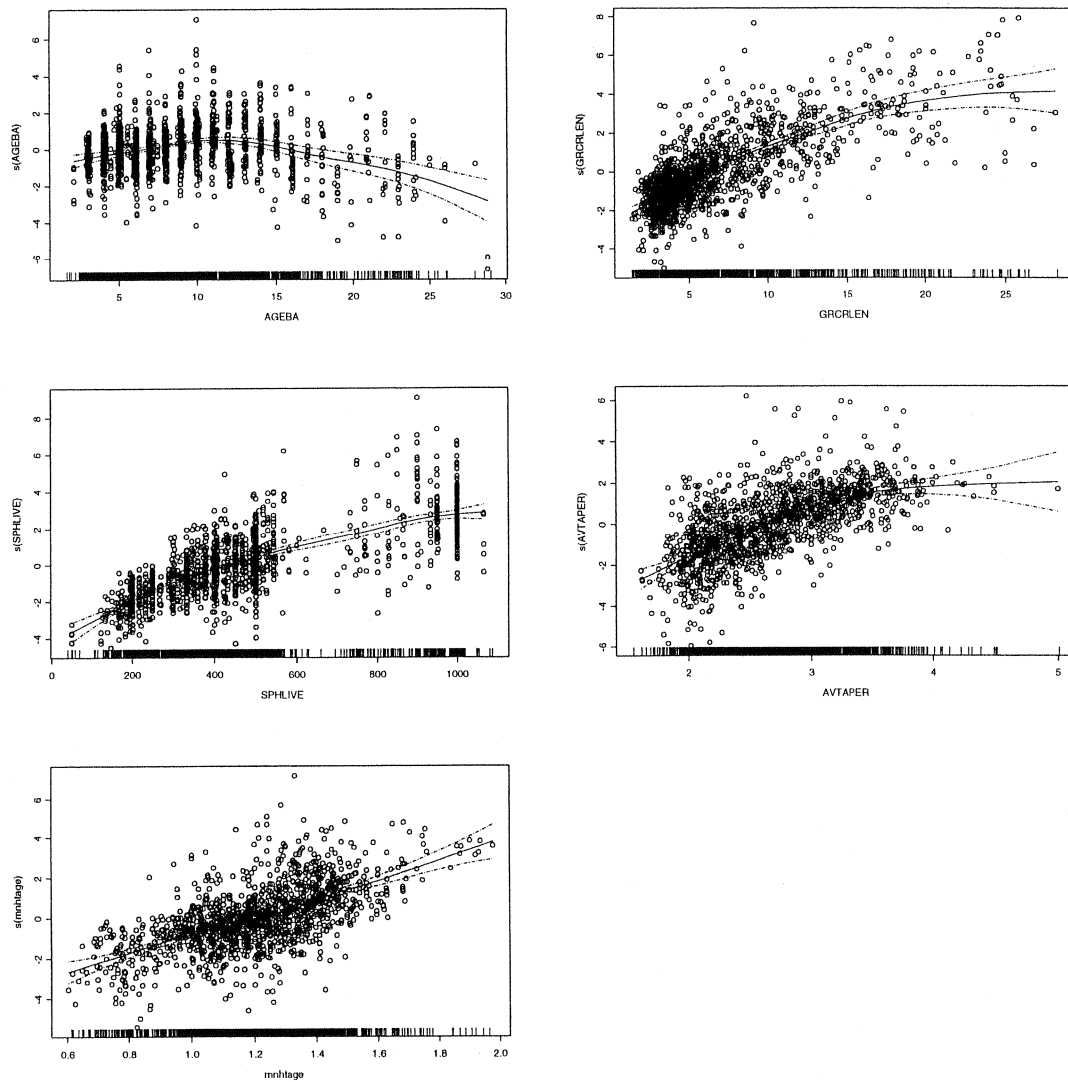


Figure 2-22. Partial regression plots for model 4 where AGEBA=Age, GRCTRLN=Green Crown Length, SPHLIVE=Stems per km, AVTAPER=Average Taper, mnhtage=Mean height / age.

Fit and Residual Plots for Model 4:

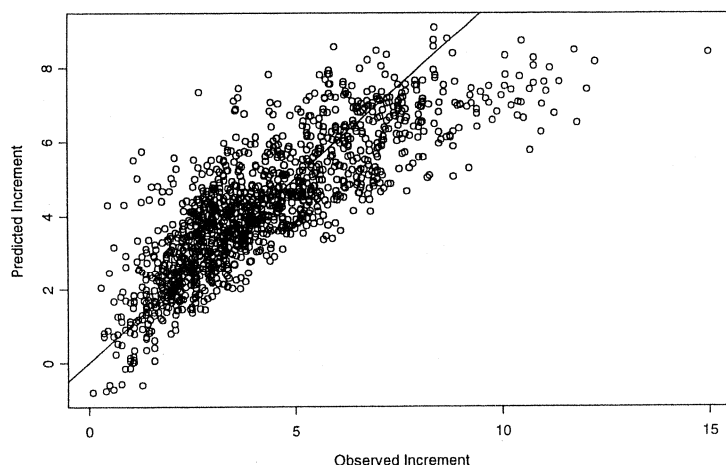


Figure 2-23. Observed versus predicted values for Model 4.

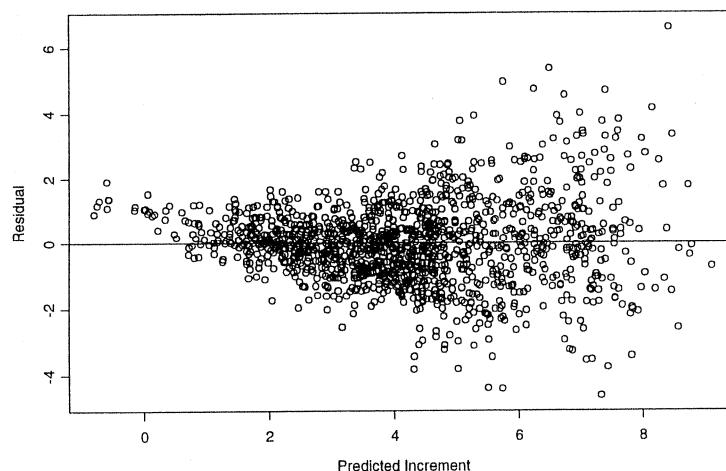


Figure 2-24. Residual plot for Model 4.

Fit Statistics for Model 4:

AIC = 2263.758 RMS = 1.61, Null Deviance = 6703.323, Residual Deviance = 2196.133 (Deviance Explained = 67.23%)

Discussion:

The partial regression plots (figure 2-22) show similar relationships to those seen previously in models one and two, while the fit illustrated in figures 2-23 and 2-24 appears similar to that seen in models 1, 2 and 3. The RMS for this model suggests a slightly better fit (1.61 as opposed to 1.93 for Model 2) while the AIC dropped substantially by about 19.6% when average taper information was included despite the estimation of three additional parameters.

Before substitute functions for model 4 were sought, the form of the average taper variable was modified. In the field the number of stems per kilometre is seen to affect

the average taper of the shelterbelt. Therefore rather than consider the average taper and stems per km variables separately, they were multiplied together and then related to the response. In addition, given that the model tended to underpredict for the larger increments it was thought that the proportion of the green crown length as a proportion of the mean height may provide a better more relative measure of productivity rather than an absolute value already used.

Model 5: Basal Area Increment = $s(\text{Age}) + s(\text{Green Crown Length}) + s(\text{Average Taper} * \text{Stems per km}) + s(\text{Green Crown Length} / \text{Mean Height}) + s(\text{Mean Height} / \text{Age})$

Coefficients:

(Intercept)	s(AGEBA)	s(GRCLEN)	s(av.sph)	s(gcl/mnht)	s(mnht/age)
-3.360108	-0.0776773	0.3069803	0.002008618	-0.6704519	3.46844

Partial Regression plots for Model 5:

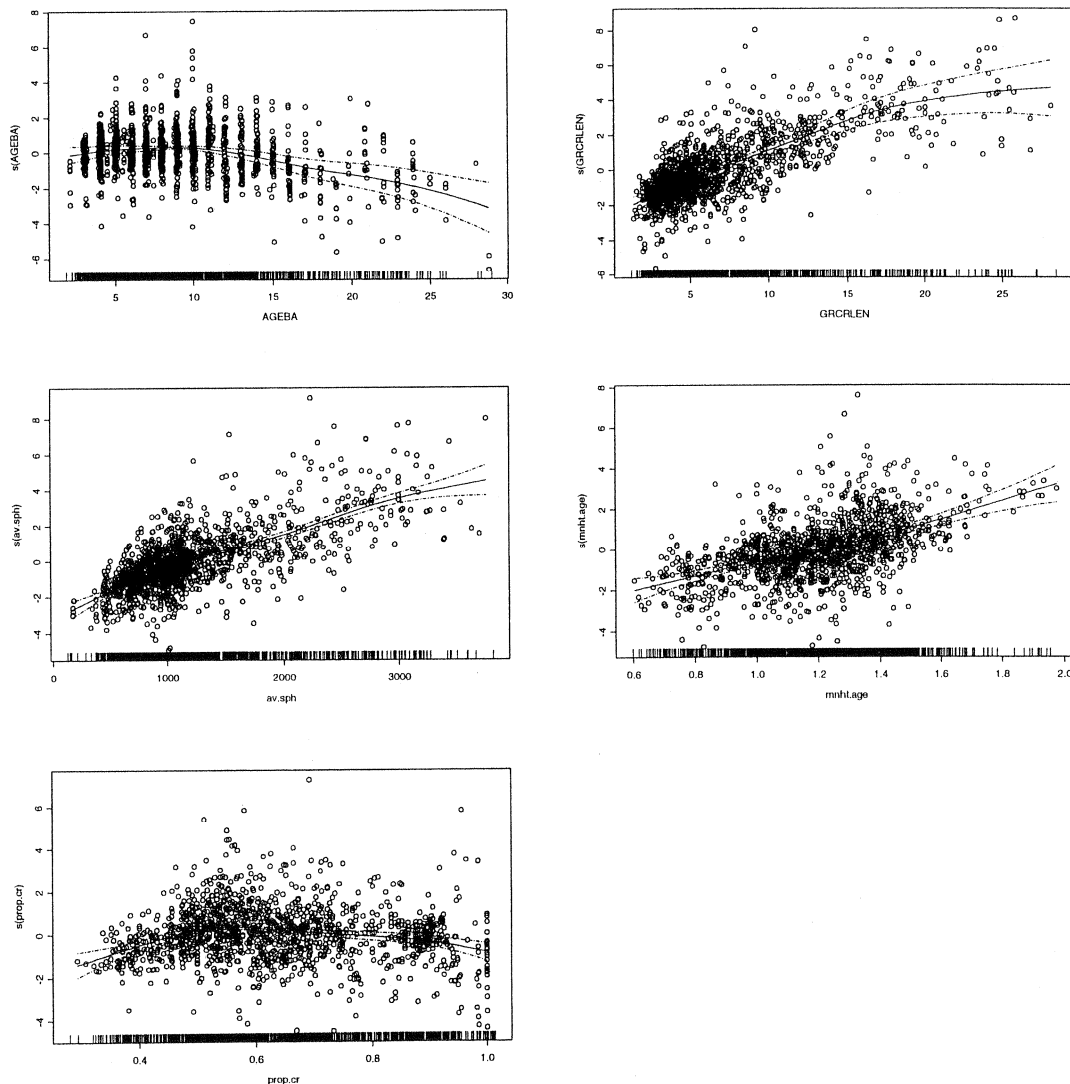


Figure 2-25. Partial Regression plots for Model 5 where AGEBA=Age, GRCLEN=Green Crown Length, avsph=Average taper/Stems per km, propr=Crown Length/Mean height, mnhtage=Mean height / age.

Fit and Residual Plots for Model 5:

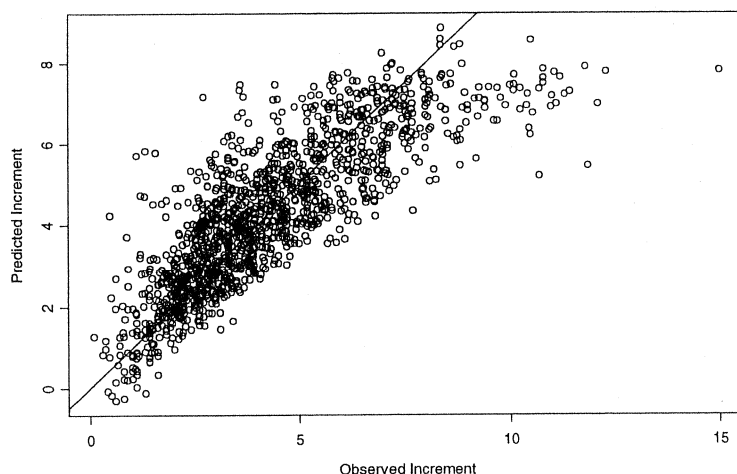


Figure 2-26. Observed versus Predicted values for Model 5.

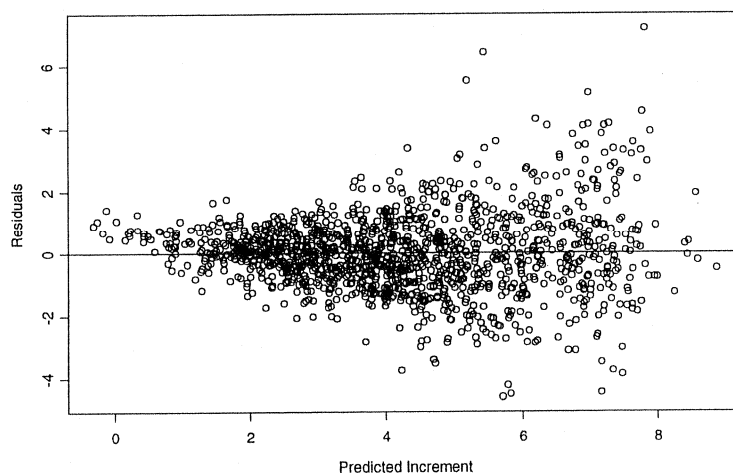


Figure 2-27. Residual plot for Model 5.

Fit Statistics for Model 5:

AIC = 2209.462 , RMS = 1.57, Null Deviance = 6703.323, Residual Deviance = 2143.354, (%VE = 68.02)

Discussion:

The partial regression plots in figure 2-25 show very similar relationships to previous models as do the fit and residual plots in figure 2-26, 2-27. However, the fit appears to have slightly improved. For example, the RMS dropped to 1.57 from 1.61 for Model 4. This slightly improved fit was confirmed by the AIC statistic as it decreased by a further 2.4%. Model 5 was substituted with linear functions for age, average taper * stems per Km, green crown length/mean height and mean height/age with a Von-Bertalanffy model fitted for Green crown length. Various non-linear functions were also investigated for the crown length/mean height variable however the

improvement in fit was negligible after the extra parameters necessary were estimated. The results were as follows:

Model 5b: Basal Area Increment = $a * \text{age} + (a2 * (1 - \exp(-(k * \text{green crown length})))) + (b * \text{Average taper} * \text{stems per km}) + (c * (\text{green crown length} / \text{mean height})) + s * (\text{mean height} / \text{age})$

Note: The necessity for a t_0 parameter for the Von Bertalanffy equation for the green crown length variable was also investigated and it was found that when a t_0 parameter was specified the parameter was not significant at the 95% level.

Coefficients:

a	a2	k	b	c	s
-0.3706394	18.0917	0.07506513	0.002012649	-4.461563	1.044587

Fit and Residual Plots for Model 5b:

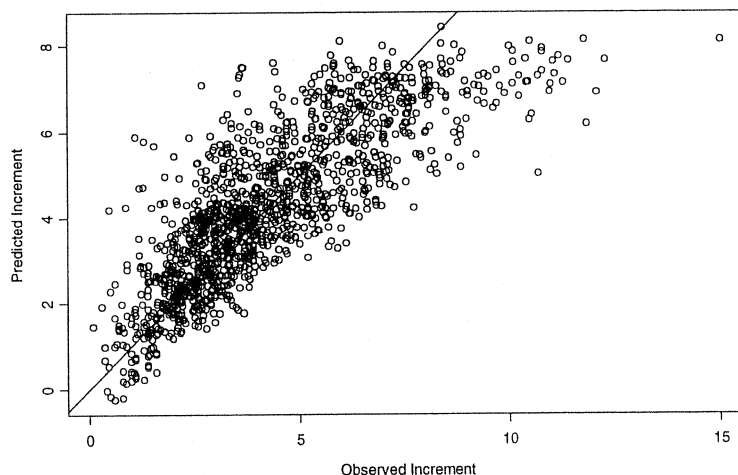


Figure 2-28. Observed versus predicted values for Model 5b.

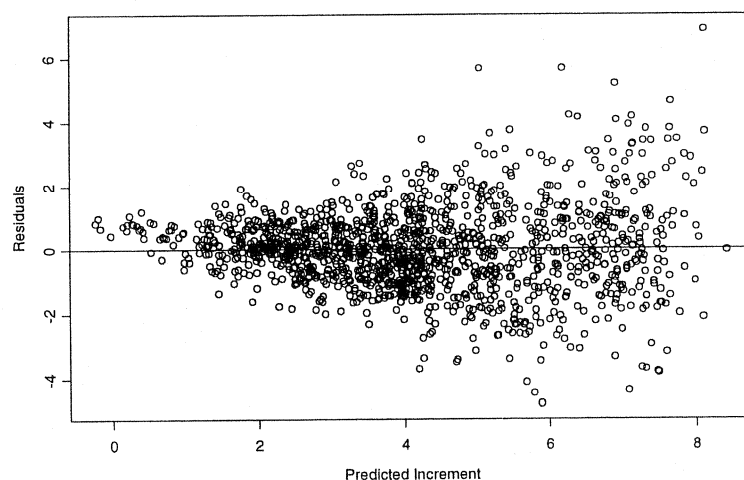


Figure 2-29. Residual plot for Model 5b.

Fit Statistics for Model 5b:

RMS = 1.689, Regression SS/Total SS = 92.5%

Discussion:

Model 5b did not perform as well as the GAM version (Model 5) and did not improve on Model 4 (however a GAM was also used for Model 4). In this instance it was difficult to find functions that significantly improved on the linear functions that justified the extra parameters estimated. However, Model 5b has produced satisfactory results, which are more closely fitting than models 2b and 3, which is not surprising given average taper information is included in Model 5b.

2.9 Non-Linear Mixed Modelling Framework:

Given that serial correlation was thought to be present in the data and the fact that nonlinear mixed models consider the data as a random sample of shelterbelts within a population of shelterbelts (where the inference is about the population), nonlinear mixed models were also used to estimate the coefficients in the two best fitting models.

Nonlinear mixed models using both independent and non-independent error structures were investigated for two basal area increment models. While it was realised that sufficient data was available to derive separate coefficients for the North and South Island it was desirable that regional estimates be derived if possible. This objective was achieved for the two best performing increment models (Models 3 and 5b).

Mixed models absent of Average Taper information:

Model 2b: Basal Area Increment: $a \cdot \text{age} + ((a2 + ra2) \cdot (1 - \exp(-k \cdot (\text{green crown length} - t0)))) + r \cdot \text{stems per km} + s \cdot (\text{Mean height/age})$
where $ra2$ is a random effect parameter.

Note: the $t0$ parameter exhibited the most variation and resulted in a better fit when a mixed model was used however it was thought that a random effect for $a2$ was more appropriate.

National Model Coefficients:

Parameter	Nonlinear regression	Mixed Model (ID)	Mixed Model (AR1)
a	-0.132861720	-0.11801266	-0.07995123
a2	4.171088135	2.92047227	2.18486071
k	0.094196092	0.10117820	0.11369091
t0	6.863862458	8.79093029	8.88424121
r	0.004456547	0.00434502	0.00429864
s	3.066846023	3.50401090	3.13036098

Table 2-2. National model coefficients for Mixed Model 2b.

Graphical Representation of Mixed Model 2b:

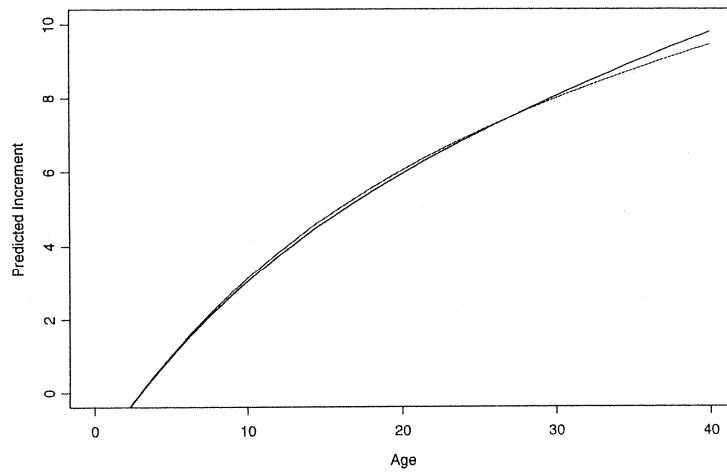


Figure 2-30. Graphical Representation of Mixed Model 2b: where the green curve is the standard nonlinear regression curve and the black curve is the mixed model with ar(1) errors.

Fit and Residual Plots for Mixed Model 2b:

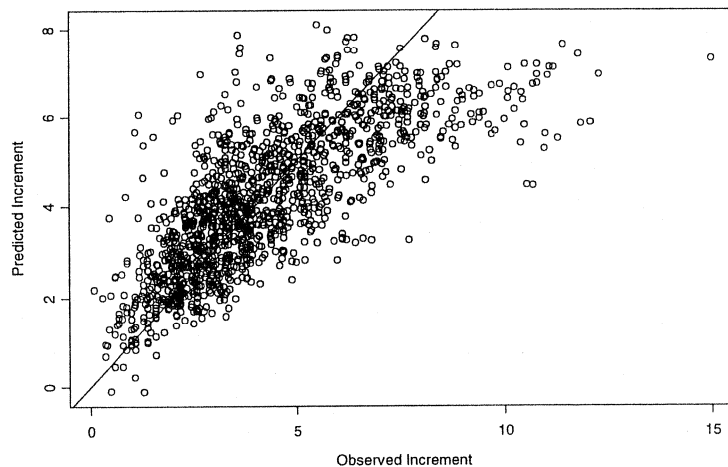


Figure 2-31. Observed versus predicted increment values for Mixed Model 2b with ar(1) errors.

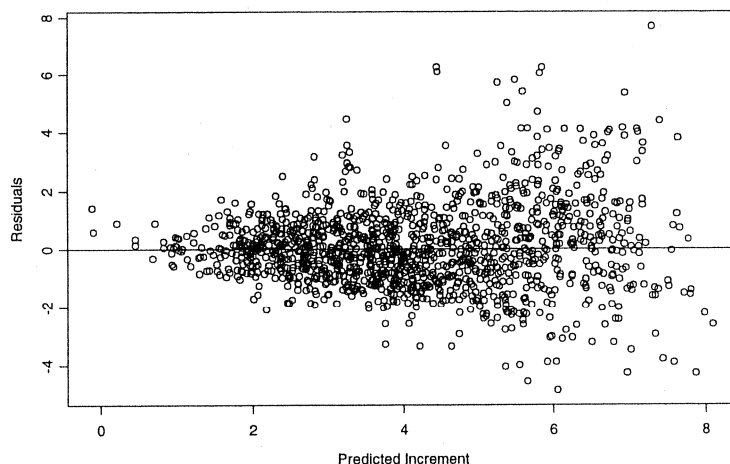


Figure 2-32. Residual plot for Mixed Model 2b with ar(1) errors.

Fit Statistics for Mixed Model 2b

AIC = -2437.02 (ID), -2345.65 (AR1)

Discussion:

Table 2-2 shows that the mixed models produced very different coefficients to those obtained using standard nonlinear regression however the resulting curves for the two approaches appeared very similar (figure 2-30). A possible reason for these differences may lie in the presence of serial correlation in the errors which is not correctly handled using nonlinear least squares. The specification of autoregressive errors (of the first order) improved the fit which was reflected in the AIC statistics, however the absolute fit (for models absent of average taper information) was improved upon with model 3. It is interesting to note that previously (when nonlinear regression was performed) Model 2b showed a better fit than Model 3. This suggests that Model 3 better represents any one shelterbelt given that mixed models indirectly fit the model for each shelterbelt, instead of fitting a single model to a data scatter.

Model 3: Basal Area Increment = $(a/1+b*Age)*(1-\exp(-c* Stems\ per\ km* Green\ crown\ length**d))$

National Coefficients:

Parameter	Nonlinear regression	Mixed Model (ID)	Mixed Model (AR1)
a	9.905231265	8.89376387	8.74102091
b	-0.187952896	-0.15995905	-0.15447496
c	0.000294633	0.00026764	0.00027547
d	1.023215868	1.17500977	1.17056529

Table 2-3. National model coefficients for Mixed Model 3.

Graphical Representation of Model 3:

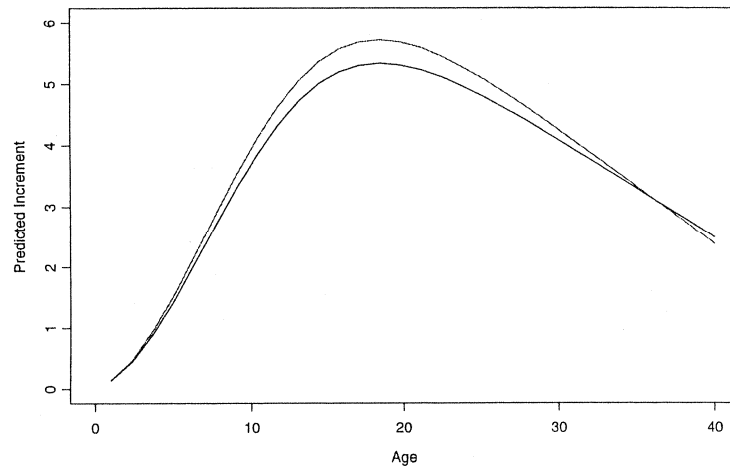


Figure 2-33. Graphical Representation of Model 3: where the green curve is standard nonlinear regression and the black curve is the mixed model with independent errors.

Fit and Residual Plots for National Mixed Model 3:

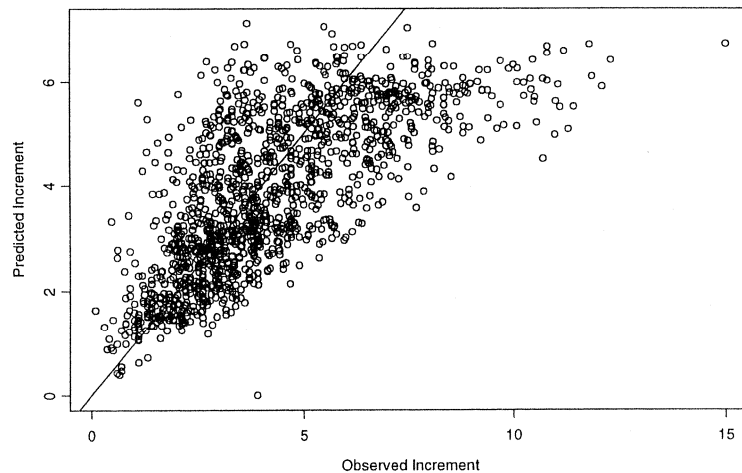


Figure 2-34. Observed versus predicted values for National Mixed Model 3 with independent errors.

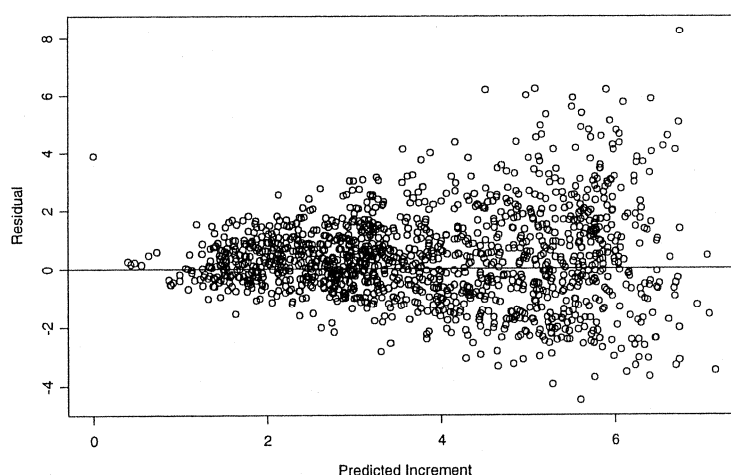


Figure 2-35. Residual plot for National Mixed Model 3 with independent errors.

Fit Statistics for National Mixed Model 3:

AIC = -2267.65 (ID), -2267.86(AR1)

Discussion:

The parameters estimated using the nonlinear mixed models were very similar to those estimated using nonlinear least squares which may be attributed to the almost undetectable levels of serial correlation inherent in the errors. Due to this low level of serial correlation in the data, the AIC statistic illustrated a worse fit when ar(1) errors were specified suggesting that an extra parameter was not justified. The overall fit of Mixed model 3 (with independent errors) was more satisfactory than the mixed model 2b (as mentioned earlier) therefore regional estimates were obtained for the asymptote parameter a, (by averaging the shelterbelt estimates within region) and the results of which are as follows:

Regional Mixed Model 3:

$$\text{Basal Area Increment} = (a/1+b*\text{Age})*(1-\exp(-c*\text{Stems per km} * \text{Green crown length}**d))$$

where the a parameter was unique for each region

Regional Coefficients for Mixed model 3:

Region	Asymptote parameter estimate
Auckland	12.1925
Bay of Plenty	9.9189
Canterbury	7.2015
Hawkes Bay	8.4485
Manawatu	8.5001
Nelson/Marlborough	7.1987
Northland	8.8810
Taranaki	8.9325
Wairarapa	7.0912

Table 2-4. Regional Coefficients for Mixed Model 3.

Graphical Representation of Regional Mixed Model 3:

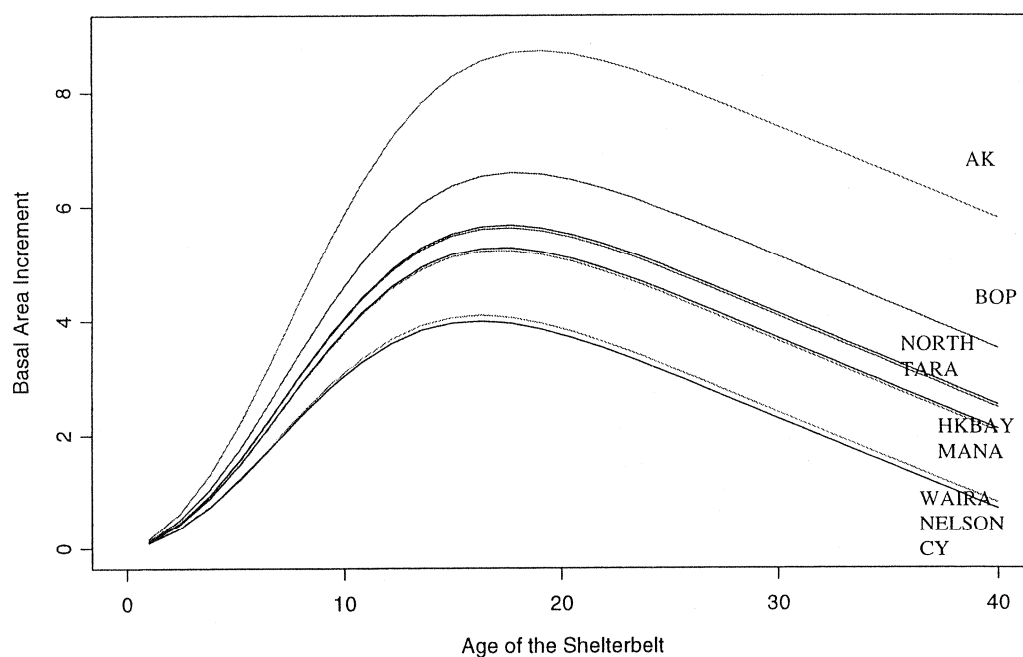


Figure 2-36. Graphical representation of Mixed Model 3 with regional parameter estimates, where AK=Auckland, BOP=Bay of Plenty, North=Northland, TARA=Taranaki, HKBAY=Hawkes Bay, MANA=Manawatu, WAIRA=Wairarapa, CY=Canterbury, NELSON=Nelson/Marlborough.

Fit and Residual Plots for Regional Mixed Model 3:

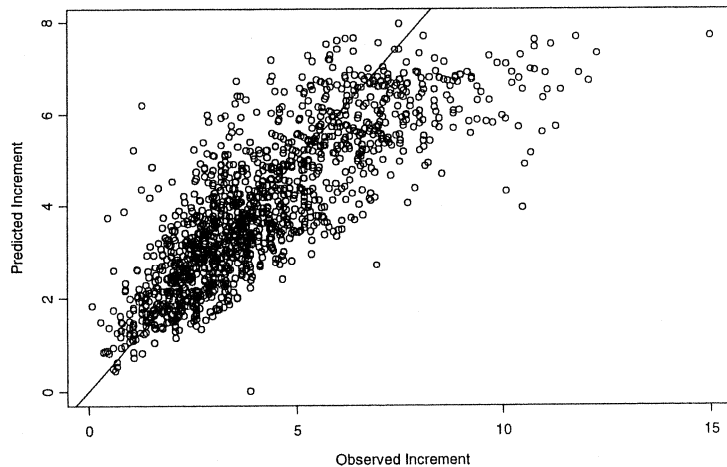


Figure 2-37. Observed versus predicted values for Mixed model 3 with regional estimates.

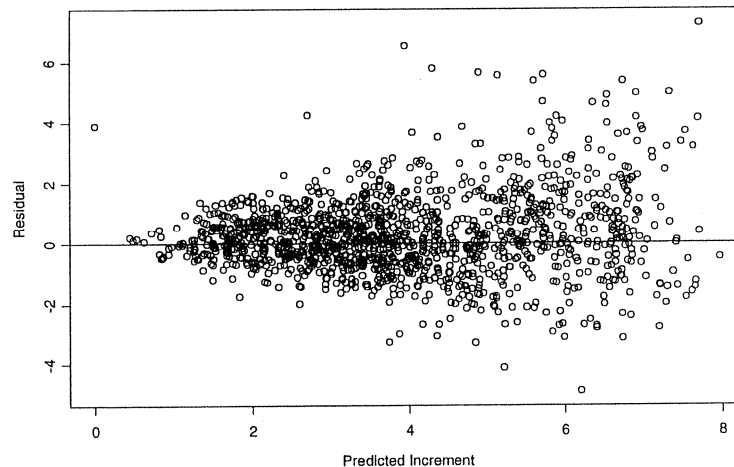


Figure 2-38. Residual plot for Mixed Model 3 with regional estimates.

Discussion:

The regional estimates listed in table 2-4 show that both the Auckland and Bay of Plenty regions exhibited estimates above the national average (9.90) while the estimate for the Wairarapa region was the smallest nationally. As could be expected the overall fit improved on the national model. This is illustrated in figures 2-37 and 2-38.

Mixed models using Average Taper information:

Model 5b was also fitted at the national and regional levels given the improved fit of this particular model. The coefficients from the models at the two levels are summarised in tables 2-5 and 2-6.

Model 5b: Basal Area Increment = $a * \text{age} + (a2 * (1 - \exp(-(k * \text{Green crown length})))) + (b * \text{Average taper} * \text{stems per km}) + c * (\text{green crown length} / \text{mean height}) + s * (\text{mean height} / \text{age})$

where the a2 parameter was considered a random effect

Coefficients for National Mixed Model 5b:

Parameter	Nonlinear regression	Mixed Model (ID)	Mixed Model (AR1)
a	-0.3706394	-0.38193237	-0.37272184
a2	18.0917	16.48224354	16.23503012
k	0.07506512	0.10237264	0.10153370
b	0.00201265	0.00175047	0.00176338
c	-4.46156271	-4.00160601	-4.04685101
s	1.04458712	0.39048901	0.45823943

Table 2-5. National Coefficients for Mixed Model 5b.

Fit and Residual Plots for National Mixed Model 5b:

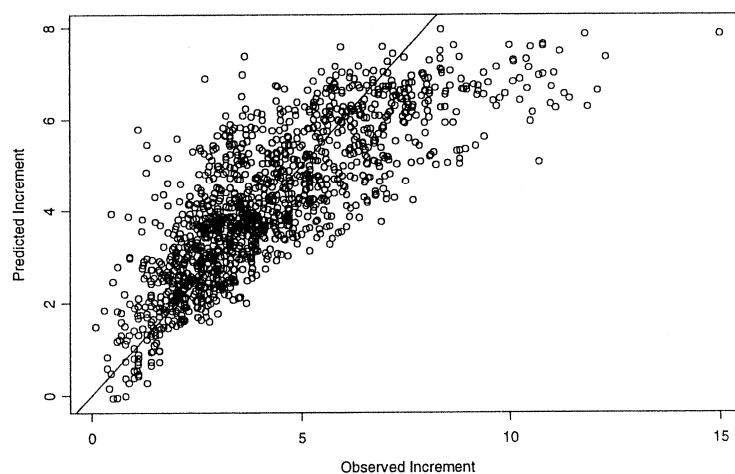


Figure 2-39. Observed versus predicted values for Mixed Model 5b with ar(1) errors.

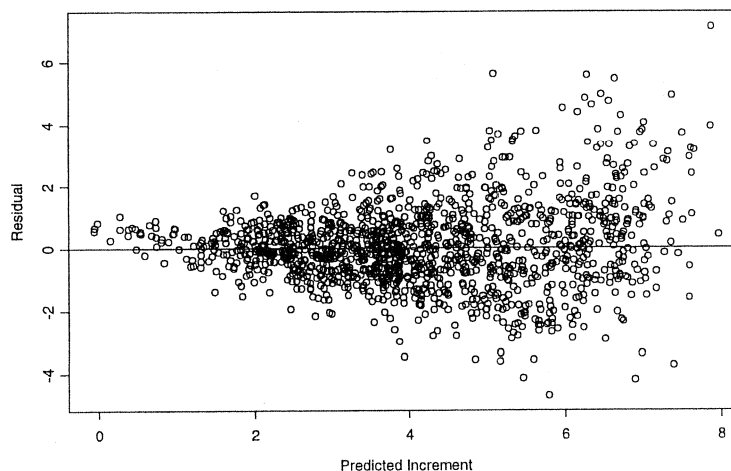


Figure 2-40. Residual plot for Mixed Model 5b with ar(1) errors.

Fit statistics for National Mixed Model 5b

AIC = -2223.96 (ID), -2222.73 (AR1)

Discussion:

The coefficients listed in table 2-5 show only small differences between those estimated using nonlinear least squares and restricted maximum likelihood (nonlinear mixed models) which may be attributed to the low level of serial correlation apparent in the errors. As with mixed model 3, the specification of non-independent errors only slightly changed the AIC statistic however for Model 5b the AIC increased suggesting that the non-independent errors were necessary. As with Model 3, regional estimates were obtained by averaging the plot estimates within each region and the results are as follows:

Regional Mixed Model 5b:

Basal Area Increment = $a * \text{Age} + (a2 * (1 - \exp(-(k * \text{Green crown length})))) + (b * \text{Average taper} * \text{stems per km}) + c * (\text{green crown length} / \text{mean height}) + s * (\text{mean height} / \text{age})$

where parameter a2 was unique for each region

Regional Coefficients for Mixed Model 5b:

Region	Asymptote parameter estimate
Auckland	18.7583
Bay of Plenty	16.9008
Canterbury	15.8120
Hawkes Bay	15.9667
Manawatu	15.4318
Nelson/Marlborough	14.8446
Northland	16.0294
Taranaki	15.8897
Wairarapa	14.8813

Table 2-6. Regional coefficients for Mixed Model 5b.

Graphical Representation of the Mixed Model 5b with regional coefficients:

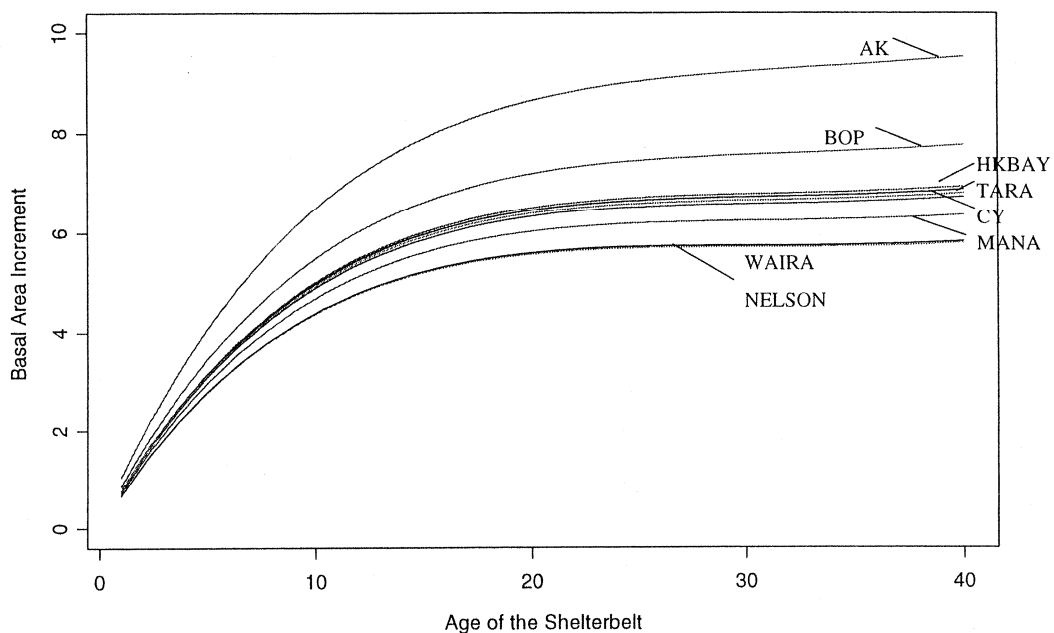


Figure 2-41. Graphical representation of Mixed model 5b with regional coefficients where AK=Auckland, BOP=Bay of Plenty, North=Northland, TARA=Taranaki, HKBAY=Hawkes Bay, MANA=Manawatu, WAIRA=Wairarapa, CY= Canterbury, NELSON= Nelson/Marlborough.

Fit and Residual Plots for Mixed Model 5b with regional coefficients:

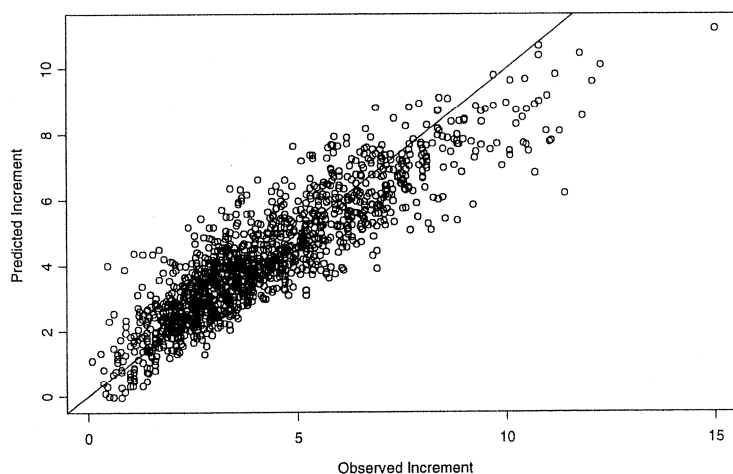


Figure 2-42. Observed versus predicted values for Mixed Model 5b with regional estimates and ar(1) errors.

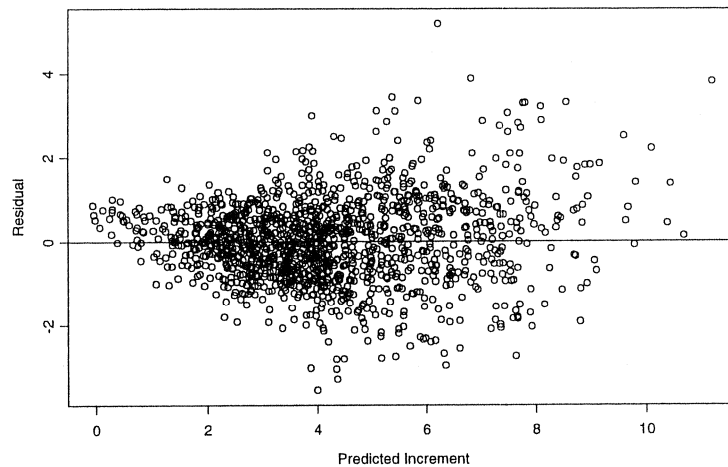


Figure 2-43. Residual Plot for Mixed Model 5b with regional estimates and ar(1) errors.

Discussion:

Once again, both the Auckland and Bay of Plenty regions produced estimates larger than the national average while the Nelson/Marlborough region exhibited the smallest asymptote estimate. The fit considerably improved when the regional estimates were used and resulted in the best fitting model formulated using the national data set.

It was considered that the number of rows in a shelterbelt would have some significant effects on the basal area increment modelling. However, while some nonlinear models showed statistically significant differences in some parameters between the one and two belt data, these differences appeared to be misleading and due to the unbalanced nature of the two data sets. For example, the one row belt data suggested a much smaller asymptote parameter (a_2) for the relationship between green crown and basal area increment however this asymptote parameter was not well supported by the available data (see figure 2-44). In addition, the residuals produced by Models 3 and 5b did not appear to be related to the number of rows in a belt (see figure 2-45).

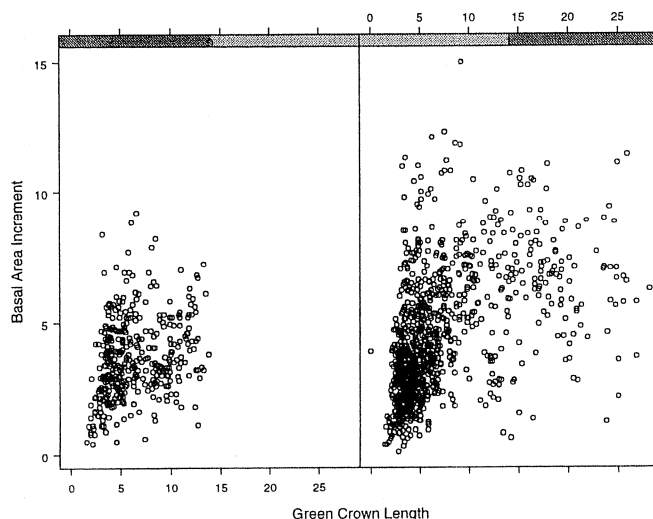


Figure 2-44. Green crown length versus basal area increment for shelterbelts with one row (left hand column headed with 0) and two rows (right hand column headed with 1).

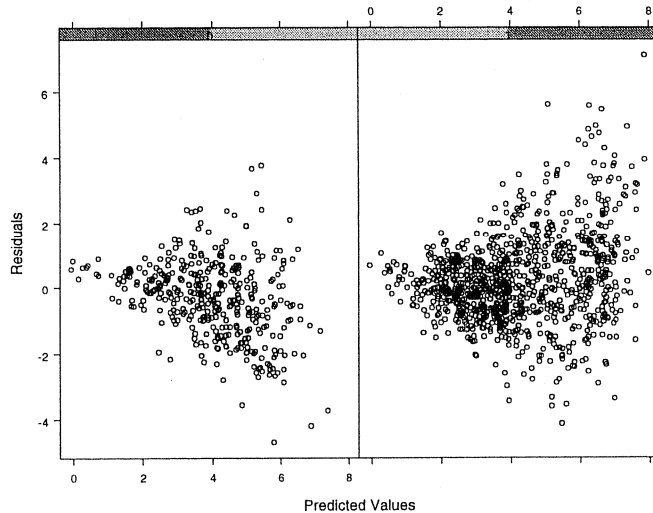


Figure 2-45. Predicted and Residual values from Mixed model 5b with AR(1) errors for shelterbelts with one row (left hand column headed with 0) and two rows (right hand column headed with 1)

Given the unbalanced nature of the data for the one and two row belts, the models were not adjusted for the number of rows in a belt.

2.10 Conclusions:

Two sets of basal area increment models were formulated using the national shelterbelt data set, one set omitted and the other included average taper information. The first set of models which excluded average taper information were fitted and assessed and it was found that Model 3 performed the best. This was not apparent when nonlinear regression was used but became evident when the coefficients were more appropriately estimated using nonlinear mixed models. Not surprisingly, the set of models which followed (which included the average taper variable) invariably produced better results than those fitted without average taper information. These models were also fitted using nonlinear least squares (nonlinear regression) and restricted maximum likelihood (nonlinear mixed models). The results show that Model 5b produced satisfactory results while the first order autoregressive errors made little difference to the results. Overall, the prediction of basal area increment using Model 5b is preferred, however all of the formulated models were found to underpredict basal area increment for a small subset of larger observed values.

Part Three: Estimating Average Taper for the Basal Area Increment Model:

3.1 Abstract:

Predicting average taper was of interest given that the model implementation requires both diameter and height information as starting values for a simulation run. The prediction of either height or diameter (when one component is assumed to be known) can be undertaken by estimating average taper from known variables eg. age and spacing. Predicting average taper was performed for two data sets: data for trees aged 2-10 and data for trees of all ages. It was thought that the modelling would solely be undertaken using the younger data set (given that either height or diameter information is more likely to be unavailable for younger trees), however if the user decided to extend the range of the model to older trees a further model was formulated. In all cases average taper was defined as quadratic mean diameter at breast height (cm)/ Mean Height (m) -1.4.

3.2 Exploratory Data Analysis:

Exploratory analysis involved simple scatterplots of average taper versus age within regions, and average taper versus spacing (metres) within the North and South islands (due to the range of data available). The results are illustrated in fig's 3-1 to 3.-4. The plots are illustrated in different colours to more easily distinguish the two data sets.

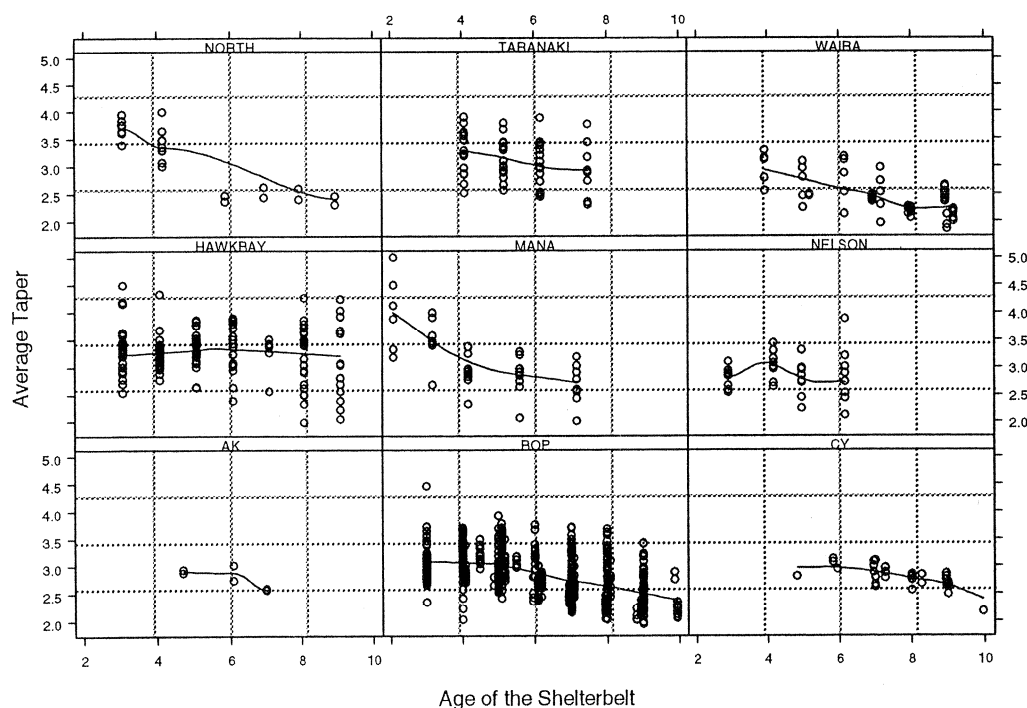


Figure 3-1. Scatterplot of Average taper versus age with a local regression model fitted for each region for trees aged 2-10 years. Note: AK=Auckland, BOP=Bay of Plenty, CY=Canterbury, HAWKBAY=Hawkes Bay, WAIRA=Wairarapa, NORTH=Northland, NELSON=Nelson/Marlborough and MANA=Manawatu.

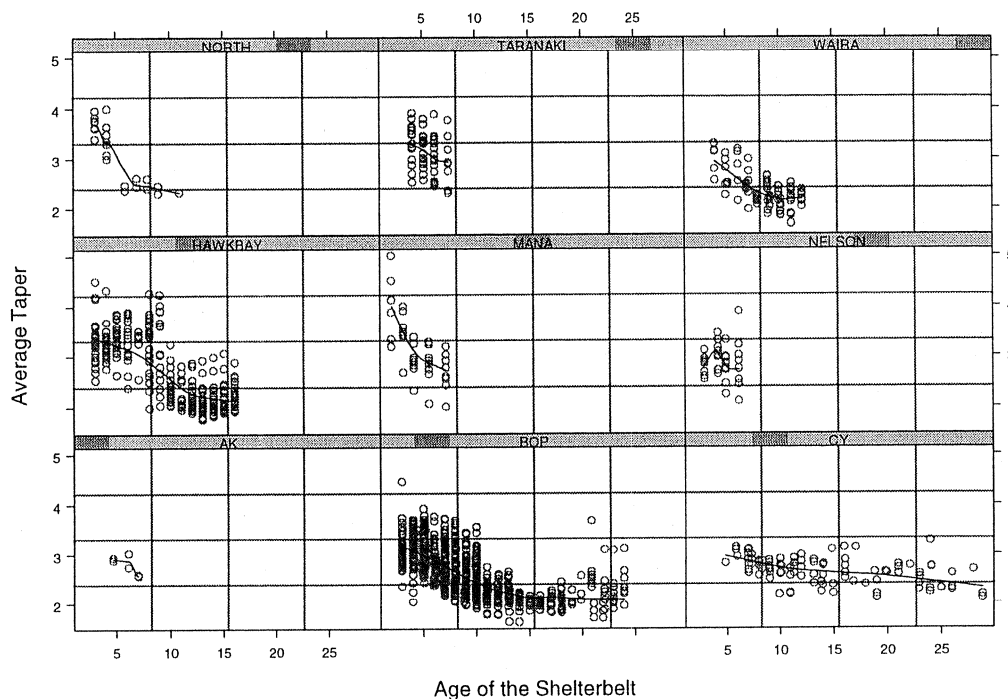


Figure 3-2. Scatterplot of Average taper versus age with a local regression model fitted for each region trees of all ages (years) where AK=Auckland, BOP=Bay of Plenty, CY=Canterbury, HAWKBAY=Hawkes Bay, WAIRA=Wairarapa, NORTH=Northland, NELSON=Nelson/Marlborough and MANA=Manawatu,.

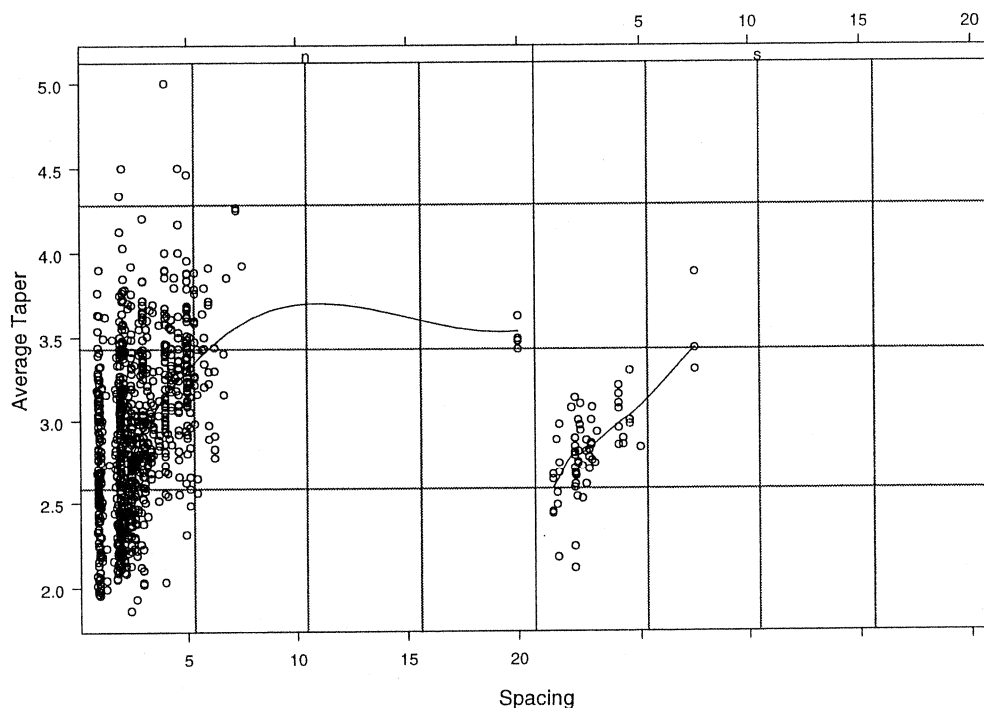


Figure 3-3. Scatterplot of Average taper versus spacing with a local regression model fitted for the North and South Islands for trees aged 2-10 years where n=North Island, s=South Island.

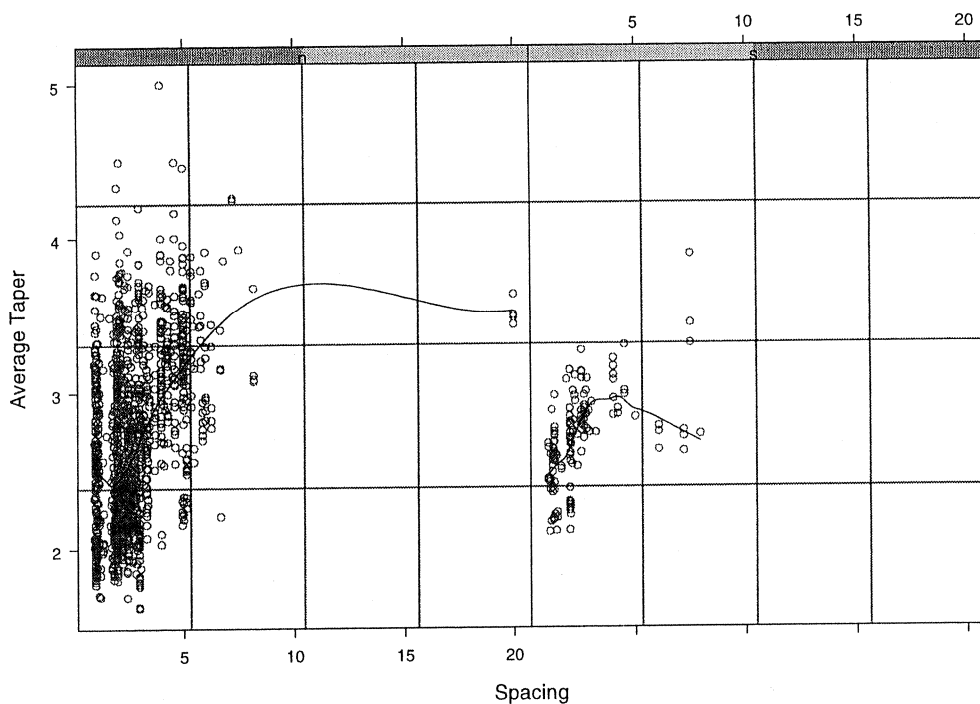


Figure 3-4. Scatterplot of Average taper versus spacing with a local regression model fitted for the North and South Islands for trees of all ages (years) where n=North Island, s=South Island.

Figure 3-1 shows that average taper decreased with age nonlinearly with great variation in the shape of the relationship between regions. Figure 3-2 represents the extended relationship which shows that while average taper decreased as the trees grew older, at approximately age 15 the relationship levelled off. Figures 3-3 and 3-4 illustrates that average taper increased until a spacing (distance in metres between trees in a single row) of about 10 metres was reached after which it levelled off (however little information was available for the wider spacings). Note: There was little correlation between age and spacing for either the younger or older data set (Pearson Correlation coefficients of 0.00063 and 0.067 respectively) which allowed both variables to be fitted together with average taper for both data sets.

3.3 Predicting Average Taper using Generalized Additive Models:

Generalized Additive models were fitted separately for Average taper with age and spacing (for both data sets) in order to establish the general shape of the relationships and the relative potential for each explanatory variable to explain the response. It was found that age offered more information about average taper given that it explained 46.7% and 21.8% of the null deviance (for the younger and older data sets respectively) while spacing explained only 21% and 21.3% of the null deviance (for the younger and older data sets respectively). The shape of the relationships between age, spacing and average taper for both data sets are illustrated in figures 3-5 and 3-6.

Fit statistics:

For trees aged 2-10:

AIC = 121.7872, RMS = 0.136, Null Deviance: 204.717, Residual Deviance: 119.332, Deviance explained: 41.7%

For trees of all ages:

AIC = 161.193, RMS = 0.116, Null Deviance = 402.144, Residual Deviance = 159.112, Deviance Explained = 60.43%.

3.3.2 Discussion:

The fit plots and the relevant statistics show that age and spacing seem to explain average taper satisfactorily for both data sets. For the younger data set the relationship between age and average taper is approximately linear while spacing seems to be related to average taper in a sigmoidal fashion (however the data at the larger spacings is limited).

For the older data set, average taper is related to both age and spacing in a nonlinear fashion. The nonlinear relationship for the spacing variable was initially considered spurious given the small amount of data at the larger spacings (Model 3-3) however, upon closer examination a linear model was not found to be sufficient and a nonlinear model was fitted (Model 3-4).

3.4 Nonlinear Regression Models:

Substitution for GAM's: trees aged 2-10

Model 3-2. Average Taper = $b * \text{Age} + a * (1 - \exp(-(k * \text{Spacing})))$

Coefficients:

	Estimate	Std. Error
b	-0.109909	0.00699327
a	3.676480	0.04506960
k	2.157300	0.09879680

Fit and Residual Plots for Model 3-2:

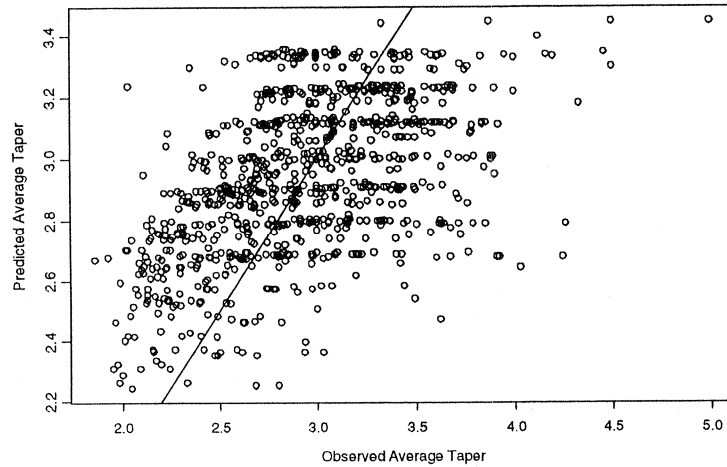


Figure 3-13. Observed versus predicted values for Model 3-2 for trees aged 2-10.

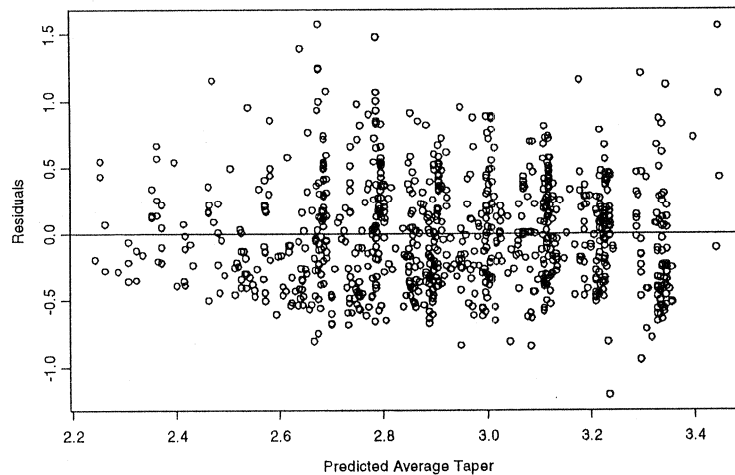


Figure 3-14. Residual plot for Model 3-2 for trees aged 2-10.

Fit Statistics:

$$\text{RMS} = 0.169 \text{ (GAM} = 0.136\text{)}$$

Discussion:

Model 3-2 appeared to fit reasonably closely to Model 3-1 given the similarity in the residual plots and the fit statistics. Model 3-2 however, appears to underpredict average taper slightly more than Model 3-1 for the larger average taper values.

Nonlinear Substitution for GAM's: trees of all ages

Note: This model was designed to represent those relationships indicated in the GAM Model 3-1, however at this stage it was thought that the limited data at the larger spacings would not warrant a nonlinear function.

Model 3-3: Average taper = $1 / (a + b * \text{age}) + c * \text{Spacing}$

Coefficients:

	Value	Std. Error
a	0.2995740	0.003856800
b	0.0144718	0.000503345
c	0.1114670	0.005921770

Fit and Residual plots for Model 3-3:

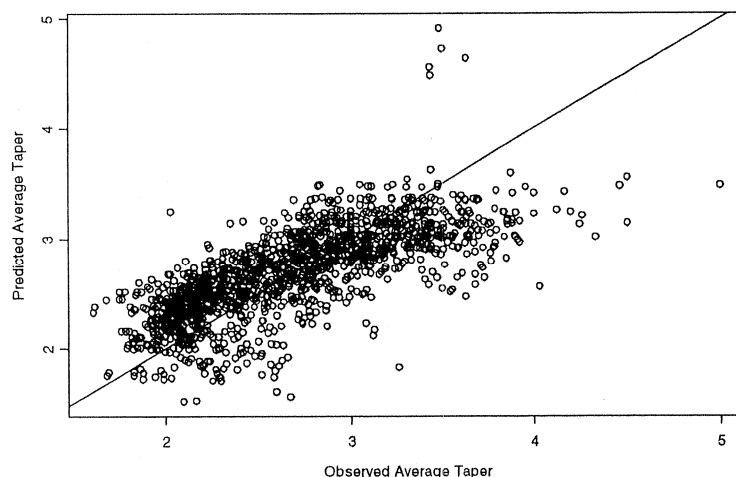


Figure 3-15. Observed versus fitted values for Model 3-3.

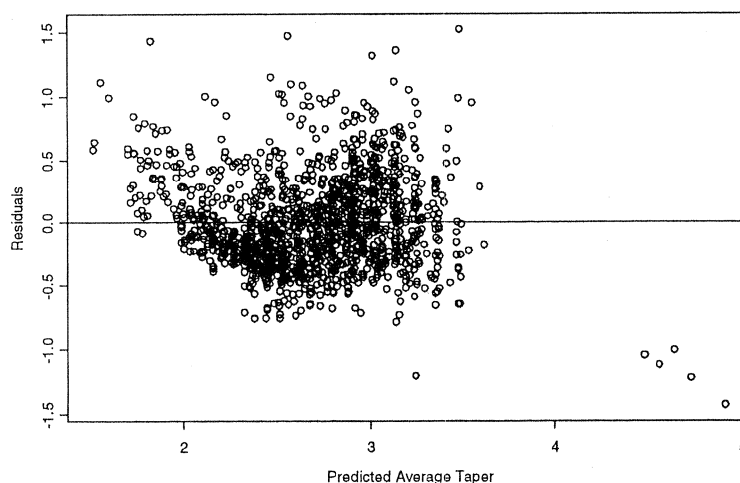


Figure 3-16. Residual plot for model 3-3.

Fit Statistics for Model 3-3:

RMS = 0.138 (GAM = 0.116)

Discussion:

It is evident from the fit and residual plots that a small group of observations were not well explained by Model 3-3 and it was found that these data points corresponded to the larger spacing values. Therefore a further model was formulated which specified a sigmoidal-type model for the spacing variable which was also considered biologically reasonable.

Model 3-4: Average Taper = $1/(a + b * \text{Age}) + a2 * (1 - \exp(-(k * (\text{Spacing} - t0))))$

Coefficients

	Value	Std. Error
a	0.3631840	0.0170397
b	0.0717659	0.0128065
a2	2.9727500	0.2890510
k	0.1092380	0.0245075
t0	-4.3941900	0.7170840

Fit and Residual plots for Model 3-4:

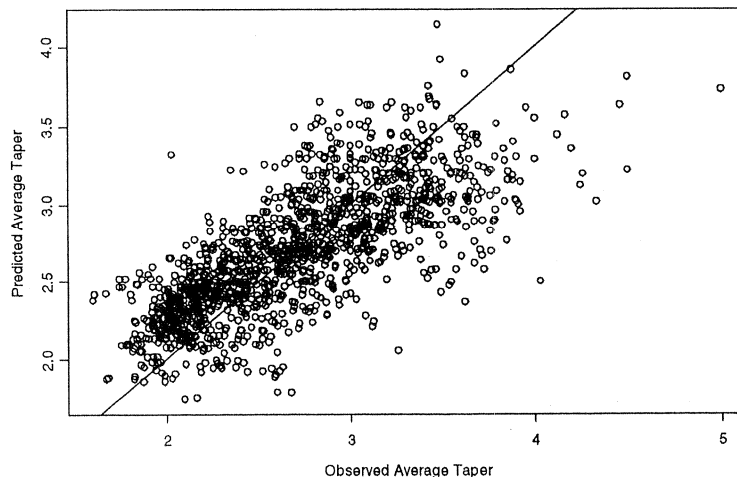


Figure 3-17. Observed versus fitted values for Model 3-4.

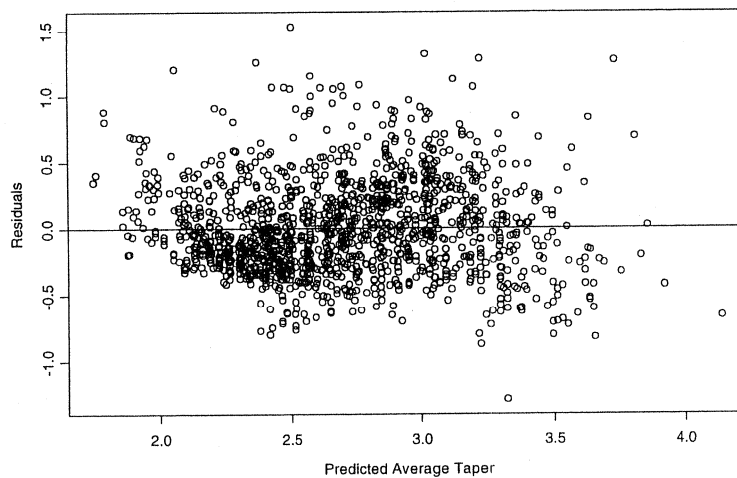


Figure 3-18. Residual plot for Model 3-4.

Fit statistics for Model 3-4:

RMS = 0.127, (GAM = 0.116)

Discussion:

The fit of Model 3-4 compared to Model 3-3 shows that even after the estimation of two additional parameters(given the nonlinear relationship with spacing) the RMS of the previous model still decreased by 7.5%. This improvement in fit is also reflected in the fit and residual plots which show very reasonable predictions. Regional models were also attempted for each region, however these were not always achievable given the different amounts of information available for each region.

3.5 Conclusions:

Average stem taper for the shelterbelt plots was analysed using Generalized Additive models (GAM's) which was followed by nonlinear regression. The absence of sufficient numbers of repeated measures for the younger data set meant that nonlinear mixed models were not suitable. The prediction of average taper using age and spacing has achieved good results for both the younger data set (<10 years) and the older data set (>10 years). However, the models for both data sets slightly underpredicted for a subset of larger average taper values.

Part Four: Obtaining Plot Asymptote estimates for the Basal Area Increment models using Average Taper and Regional Asymptote Estimates.

4.1 Introduction:

Plot estimates for the asymptote parameter used in the basal area increment mixed models 2b and 5b were available for all of the plots used in the modelling data set. However plot estimates are often desired for additional young plots not previously used in the modelling where a shelterbelt owner is interested in the performance of trees on a particular site in relation to their regional estimate.

4.2 Exploratory Data Analysis:

It was considered that the difference between the plot asymptote estimates and the relevant regional estimates for Model 3 and Model 5b might be adequately explained by the average taper measurements for those plots. The relationship between the regional estimates, plot estimates and average taper was investigated for plots up to the age of 10 using simple scatterplots (figure 4-1) and then fitted using Generalized Additive Models and Nonlinear Regression.

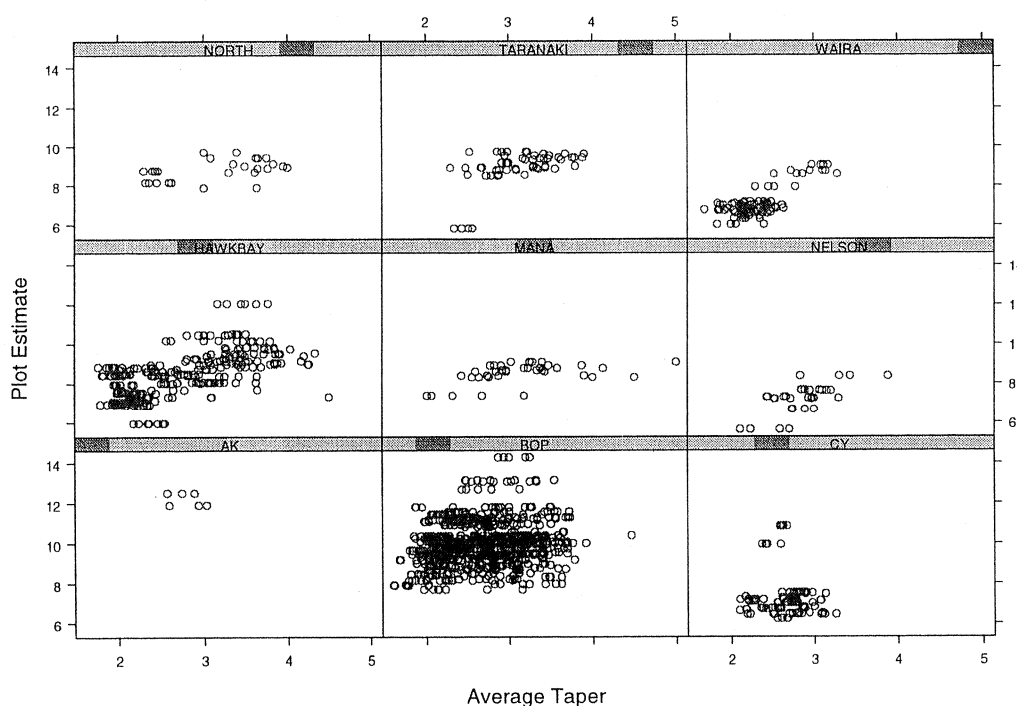


Figure 4-1. Scatter plots of plot estimates versus average taper within each region for mixed model 2b.

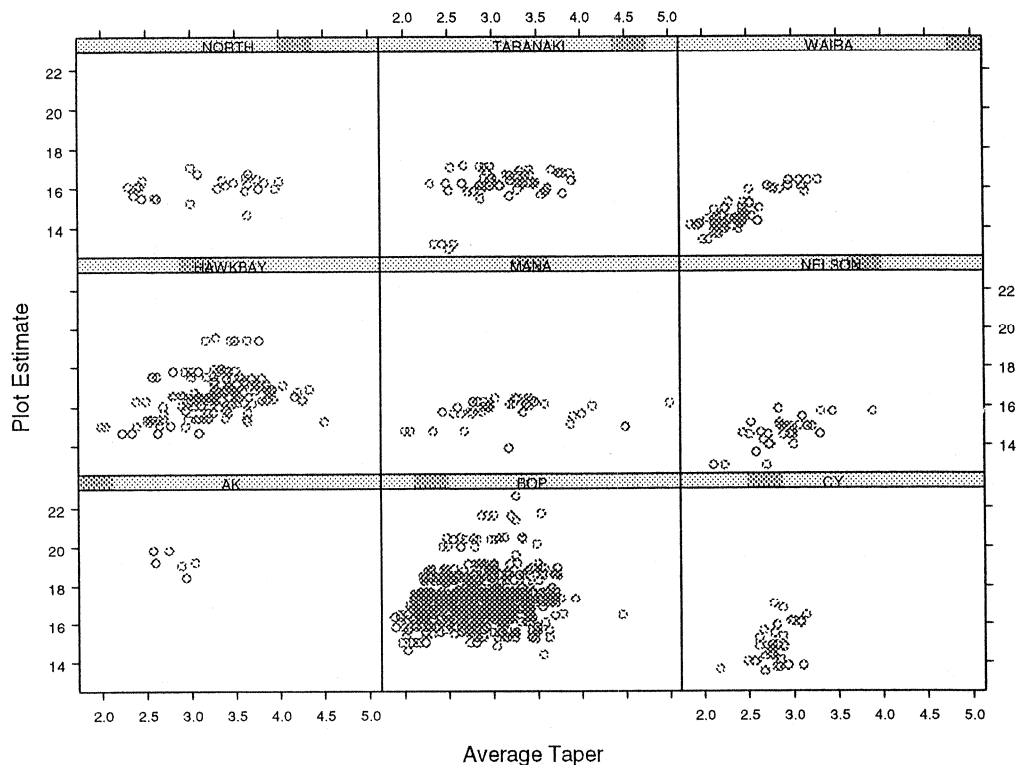


Figure 4-2. Scatter plots of plot estimates versus average taper within each region for mixed model 5b.

Figures 4-1 and 4-2 show moderately increasing relationships between the plot estimates and the average taper values for each plot. The relationship between spacing and plot estimates was also investigated but was not found to be significant at the 95% level. The shape of the relationship between average taper and the plot estimates was also investigated using smoothers (figure 4-3).

4.3 Predicting plot estimates using GAM's for Mixed Model 2b:

Model 4-1: Plot estimate = s (Regional Estimate) + s (Average Taper)

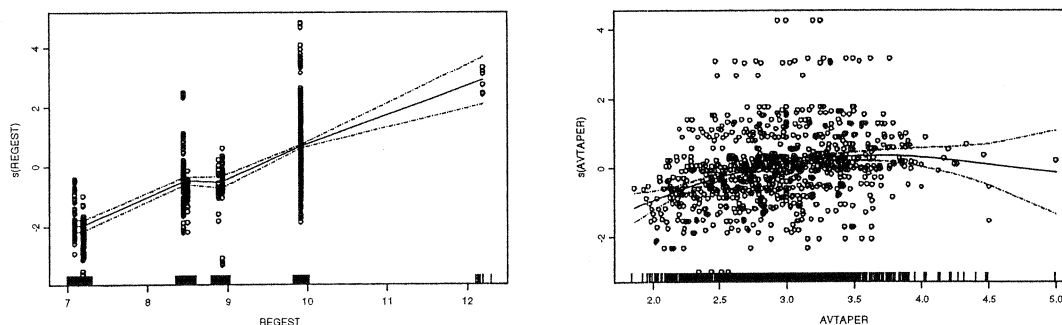


Figure 4-3. Partial Regression plots for Regional estimates (REGEST) and Average Taper (AVTAPER) for model 4-1.

Fit plots for Model 4-1:

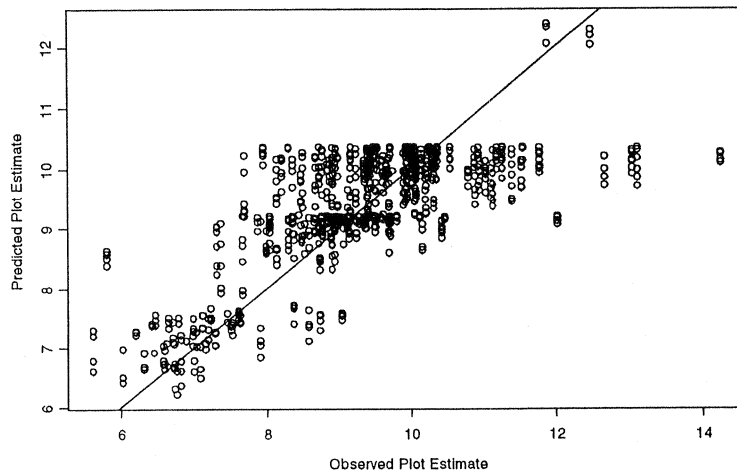


Figure 4-4. Observed versus predicted values for model 4-1.

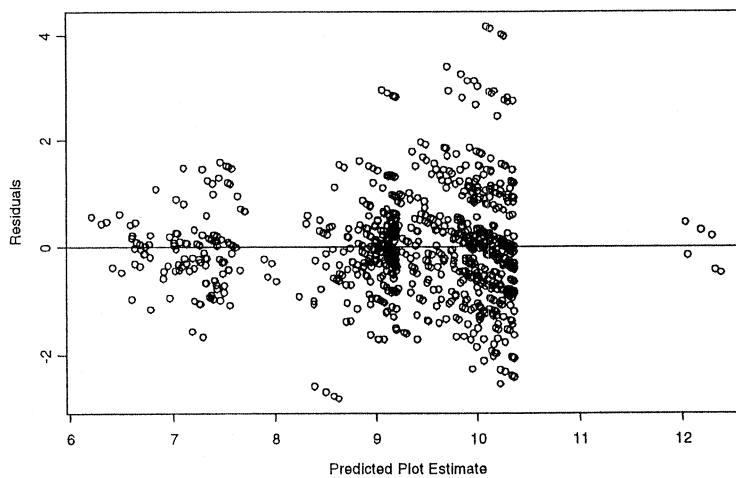


Figure 4-5. Residual plot for Model 4-1.

Fit Statistics for Model 4-1:

RMS = 0.9764186 (52.14% Deviance Explained)

4.4 Nonlinear Substitution for Model 4-1:

Model 4-2: Plot estimate = Regional estimate + (a *Average Taper *exp(k *Average Taper))

Coefficients:

a	k
0.0047096891	0.8685620245

Fit and Residual plots for Model 4-2:

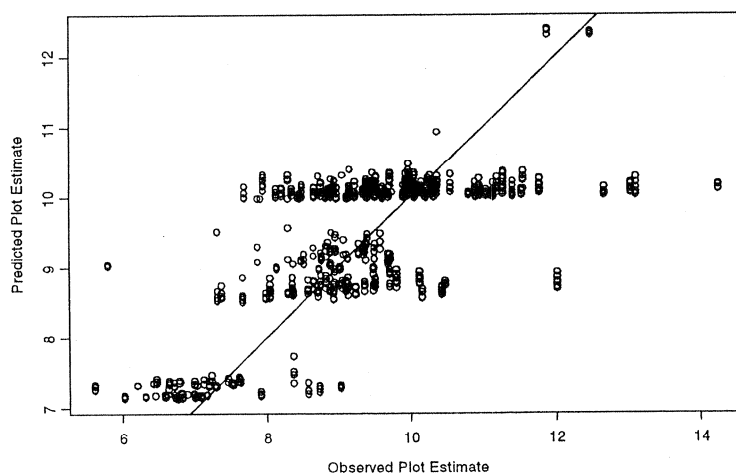


Figure 4-6. Observed versus predicted values for model 4-2.

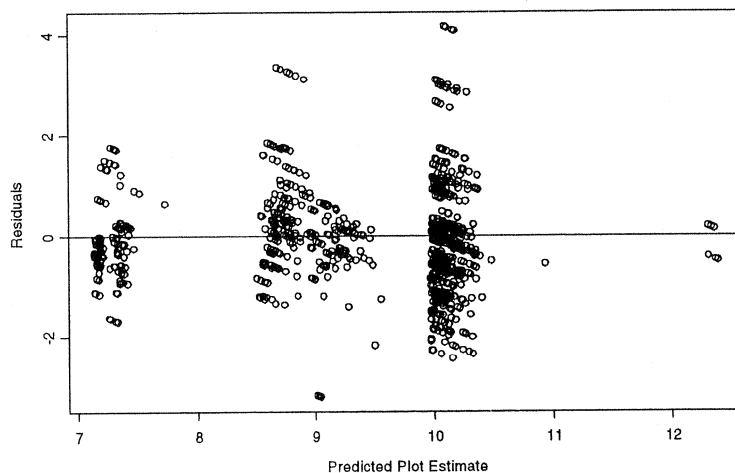


Figure 4-7. Residual plot for Model 4-2.

Fit statistics for Model 4-2:

RMS = 1.074222 (GAM = 0.9764186), Regression SS/Total SS = 98.78%

Discussion:

The nonlinear model 4-2 appeared to fit similarly to Model 4-1 which was reflected in the plots and fit statistics. Both the Generalized Additive models and the nonlinear models seemed to fit satisfactorily.

4.5 Predicting plot estimates using GAM's for Mixed Model 5b:

Model 4-3: Plot estimate = s (Regional Estimate) + s (Average Taper)

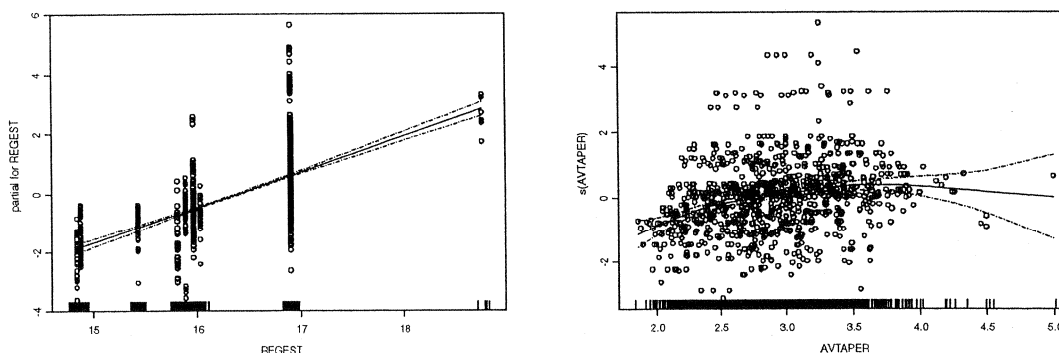


Figure 4-8. Partial Regression plots for Regional estimates (REGEST) and Average Taper (AVTAPER) for model 4-3.

Fit plots for Model 4-3:

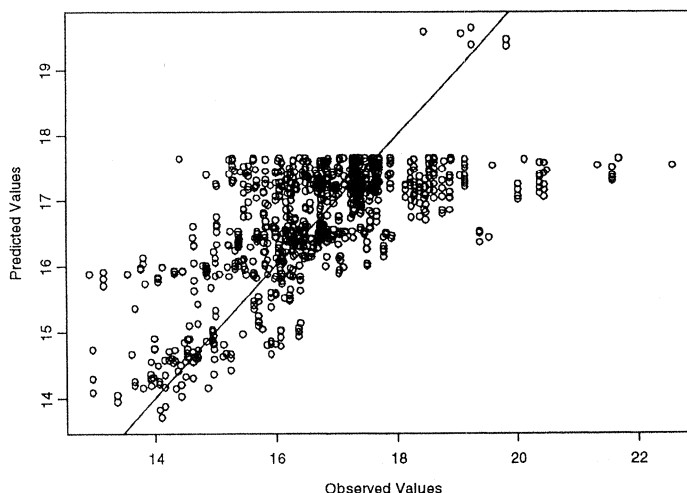


Figure 4-9. Observed versus predicted values for model 4-3.

Residual Plot for Model 4-3:

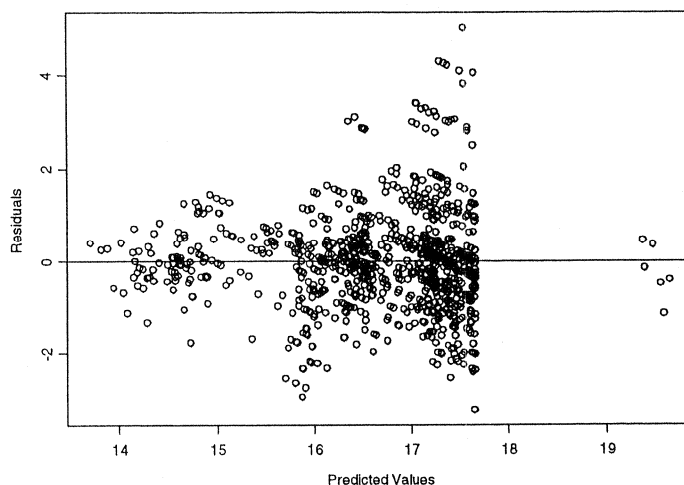


Figure 4-10. Residual plot for Model 4-3.

Fit Statistics for Model 4-3:

RMS = 1.158444 (42.49% Deviance Explained)

4.6 Nonlinear Substitution for Model 4-3:

Model 4-4: Plot estimate = Regional estimate + (a * Average Taper * exp(k * Average Taper))

Coefficients:

a = 0.0037937768, k = 0.9539926985

Fit and Residual plots for model 4-4:

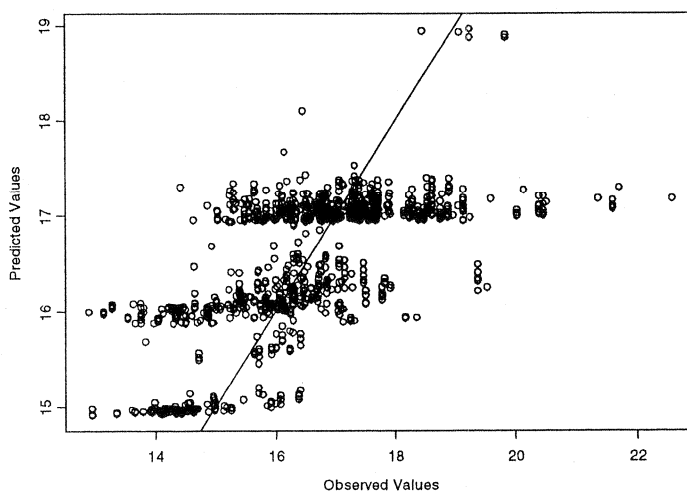


Figure 4-11. Observed versus predicted values for model 4-4.

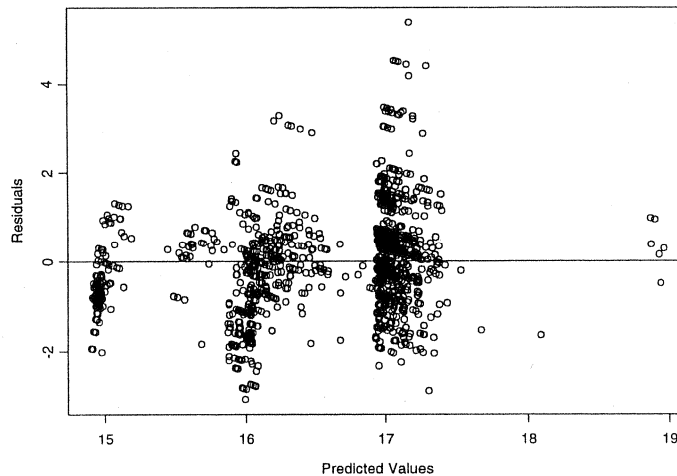


Figure 4-12. Residual plot for Model 4-4.

Fit Statistics for Model 4-4:

RMS = 1.30523(GAM =1.158444), Regression SS/Total SS = 99.52%

Discussion:

The nonlinear model 4-4 appeared to fit similarly to Model 4-3 which was reflected in both the plots and fit statistics. It appeared that both the GAM and the nonlinear models seemed to predict plot estimates satisfactorily.

Conclusions:

For the best performing basal area increment models (Models 3 and 5b) plot estimates were obtained for the appropriate asymptote value. These are necessary to calibrate the model for local use. The selected nonlinear model (Model 4-4) appeared to fit similarly to the Generalized Additive model (Model 4-3) which was reflected in both the plots and fit statistics and was found to predict plot estimates satisfactorily.

Part Five: Modelling the rise in green crown

5.1 Introduction:

In order to make accurate basal area predictions the length of the green crown needs to be estimated when crown height measurements are unavailable. The estimation of crown height was attempted for the shelterbelt data however it was found that only approximately 17.9% of the observations contained crown height information that was not a result of pruning. In addition, given that the aim was to model the natural rise of the crown, if a pruning operation was performed at the end of the annual period then that data point was not included in the modelling.

After taking into consideration the requirements above, the data set was reduced to only 28 observations, nevertheless the modelling was attempted. It was initially thought that crown height at the end of an annual period would be related to the age of the stand, the crown height at the beginning of the annual period and the spacing of the shelterbelt at the beginning of the annual period. However, only the age variable was found to offer a significant amount of information about crown height at the end of the annual interval.

5.2 Annual Model for Crown Height:

Model 5-1: Crown Height at end of annual period = $s(\text{Age at beginning})$

Coefficients:

(Intercept)	$s(\text{LAGAGE})$
1.180552	0.3733498

GAM regression plot for Model 5-1:

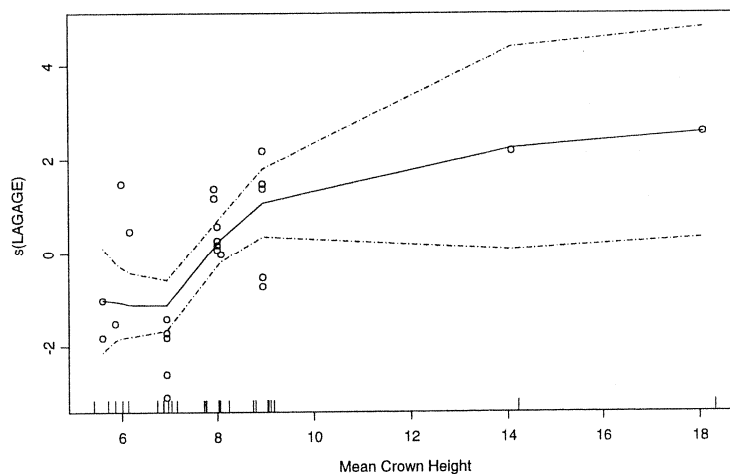


Figure 5-1. Fitted smooth function for GAM model 5-1.

Fit plot for Model 5-1:

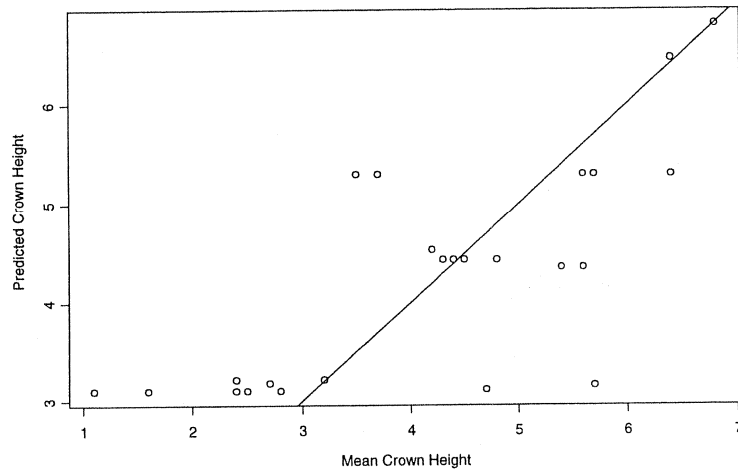


Figure 5-2. Observed versus predicted values for GAM model 5-1.

5.3 Conclusions:

Attempts were made to model crown height change over time in shelterbelts. The results did not provide a convincing fit given the small amount of data used in the modelling. While figure 5-1 suggested a sigmoidal relationship between mean crown height and age, the asymptote shape was only supported by two data points and was not considered reliable. Therefore due to insufficient data, the green crown height was not able to be modelled and the growth model assumes no natural rise in the green crown of shelterbelts.

Part Six: Predicting Mortality

6.1 Introduction

Mortality in shelterbelts is most often associated with wind and the exposed nature of the site, therefore a substantial amount of shelterbelt mortality may be considered to be catastrophic. Most often the younger trees are seen to blow over and be related to major weather events such as storms. This has proved difficult to model.

The mortality modelling for the shelterbelt data involved two approaches: Logistic regression which modelled the proportion of the stocking lost annually (Approach 1) and modelling the absolute stocking at the end of the interval given the stocking at the beginning (Approach 2). The latter approach produced satisfactory results.

Previous Work

Previous work related to the mortality function for shelterbelts suggested that mortality was uniform over time (Auclair et al, 1991) and is as follows:

$$dN / dt = -2.8N / 1000$$

To investigate the effectiveness of this function it was refitted with the available data and was found to perform unsatisfactorily with an R^2 value of 0.000739. An alternative function was therefore sought.

6.2 Exploratory Data Analysis

The exploratory analysis initially involved a simple correlation analysis of all of the available variables. This identified the following stand level variables as being significantly correlated with mortality: Mean height, Mean top height, Mean crown length, Basal area increment, Green Crown Length, Mean top diameter, Basal Area, Site index, and age. Additional variables were also found to be highly correlated with mortality but were found to be formulated using the variables already listed while mean top height and mean height were also too highly correlated to be fitted together. A high level of correlation was also found between the age, basal area, and mean top diameter variables which also resulted in the latter two variables being excluded.

In addition to the correlation analysis, simple scatter plots were produced for the candidate variables. The results are as follows:

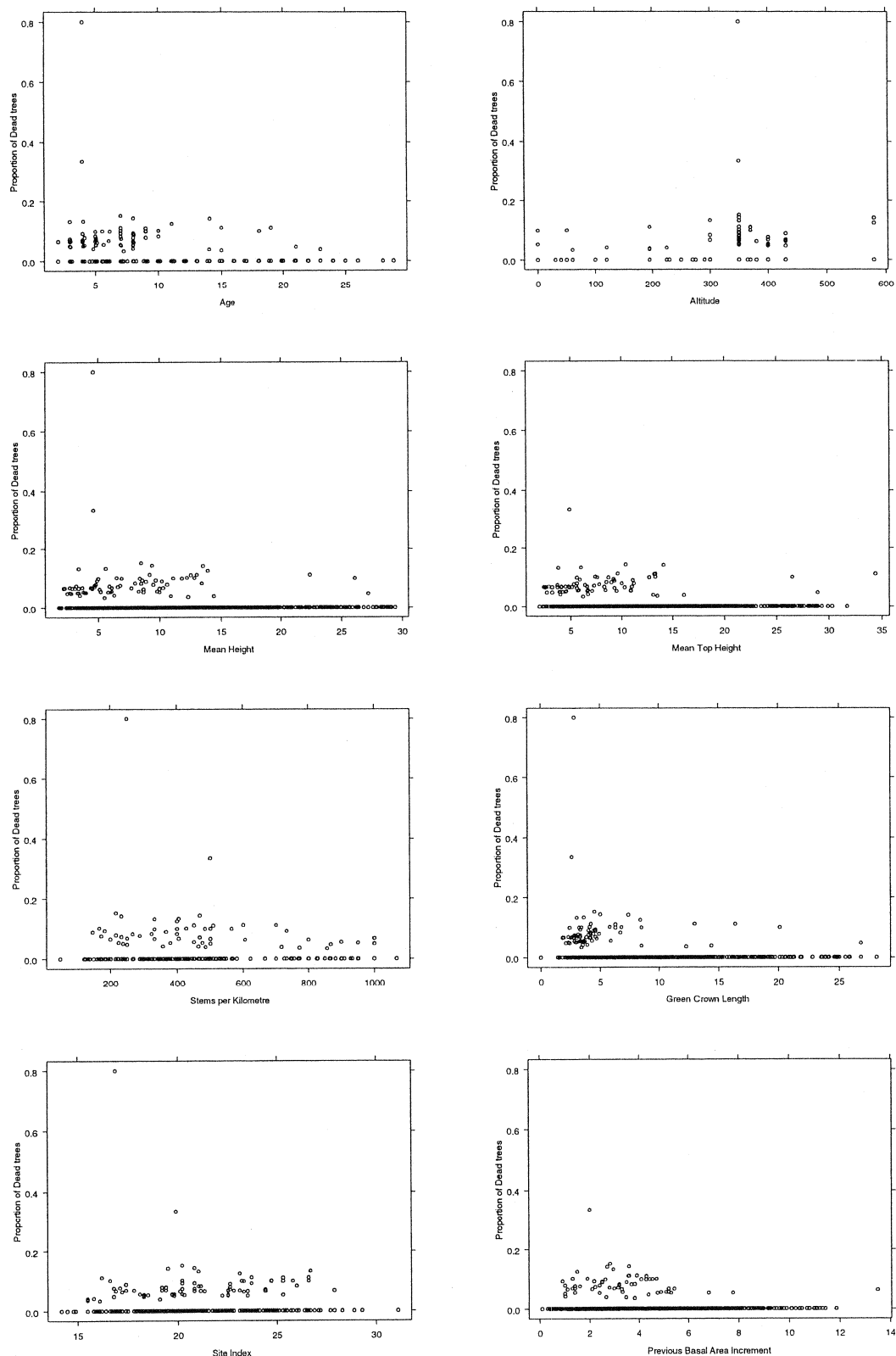


Figure 6-1. Scatterplots of the proportion of trees per kilometre that died during an annual interval versus the candidate explanatory variables.

Figure 6-1 shows reasonably weak relationships between the proportion of trees per kilometre that died annually and all of the candidate variables. It appears that the

younger trees are more prone to mortality (which also corresponds to the smaller trees with a shorter crown length) while mortality also appears to be weakly related to altitude where it appears that the plots situated at higher altitudes exhibited higher levels of mortality.

6.3 Variable Selection:

Variable selection involved fitting bivariate logistic GAMS to explain the probability of mortality during an annual interval. This was then followed by a stepwise regression using the AIC statistic. This provided information about which variables were most closely related to mortality and the form of the individual relationships.

Fit Statistics for Bivariate Logistic GAM's:

Variable	Null Deviance	Residual Deviance	% VE
Basal Area Increment	90.10576	61.32579	31.94021
Age	90.10576	81.73738	9.287286
Mean Top Height	78.43806	70.75814	9.791063
Mean Height	89.08064	80.96304	9.112642
Altitude	86.37899	83.7714	3.018778
Basal Area	90.10576	81.10981	9.983768
Green Crown Length	88.11575	82.10622	6.820041
Stems per Km	90.10576	86.26794	4.259239
Site Index	89.56461	85.82086	4.179943

Table 6-1. Statistics for Bivariate GAM fits using a logit link and binomial error.

Tables 6-1 shows that the basal area increment variable for the previous annual interval explained the most variation of all the available variables. Basal area, age, mean top height, and mean height variables followed the increment variable in terms of variation explained. While it was realised that using the previous increment was perhaps a potential bias (given that it is often only available as a prediction to the user) it was still considered, given that in the absence of climate information, shelterbelt mortality seemed difficult to explain.

Therefore, the variables in tables 6-1 were selected for the mortality model by stepwise selection (using the AIC criterion) and resulted in the following logistic mortality model:

Model 6-1: $\text{logit}(\text{Probability of Mortality}) = s(\text{age}) + s(\text{mean top height}) + s(\text{altitude}) + s(\text{green crown length}) + s(\text{site index}) + s(\text{stems per kilometre}) + s(\text{basal area increment for the previous year})$.

Model 6-1 was fitted and assessed using the standard statistics (AIC, - 2 log likelihood etc) and the resulting predicted probabilities were then converted into the number of dead stems per kilometre by simply multiplying the predicted probability of mortality by the number of stems per kilometre present at the start of the interval.

6.4 Results:

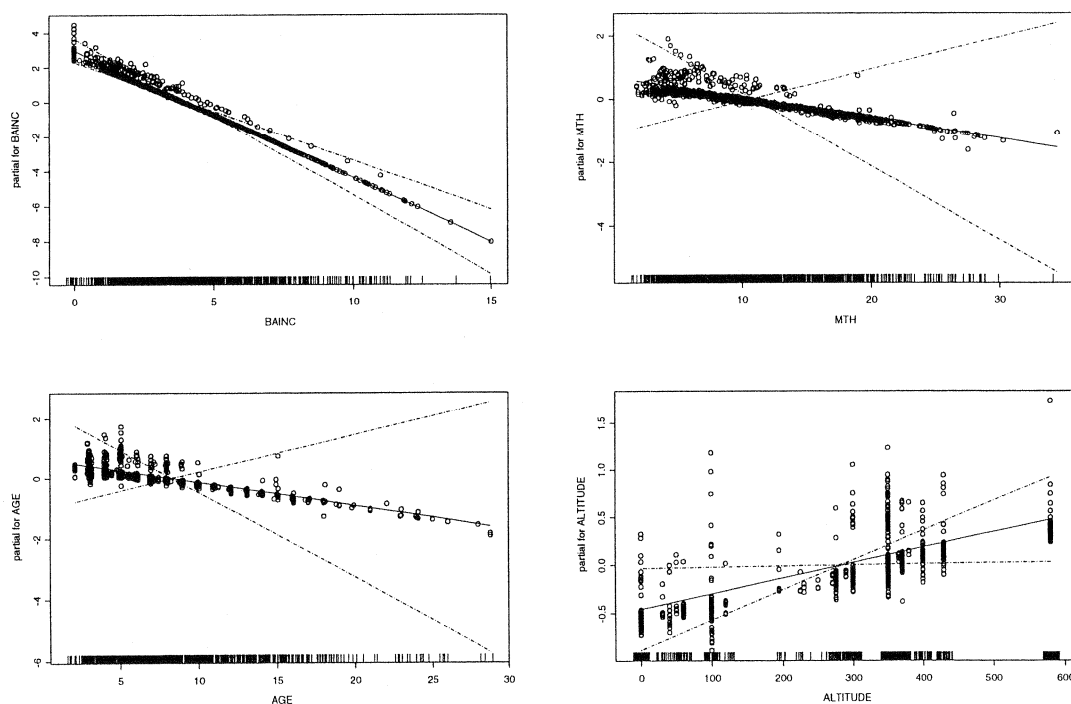
The results of the logistic regression modelling using both standard logistic regression and Generalized additive models are listed below:

Coefficients:

Variable	Coefficient - LR	Coefficient - GAM
Basal Area Incr.	-0.7357825	-0.7838508
Mean Top Height	-0.06466	-0.3453091
Age	-0.07749	0.2448011
Altitude	0.00161	0.002984032
Green Crown Length	0.14918	0.1741999
Site Index	0.03369	0.06941503
Stems per Km	0.000643	0.00074228

Table 6-2. Model Coefficients for Model 6-1 where LR= Logistic regression and GAM= logistic Generalized Additive model .

Partial Regression plots for Logistic Regression Model 6-1:



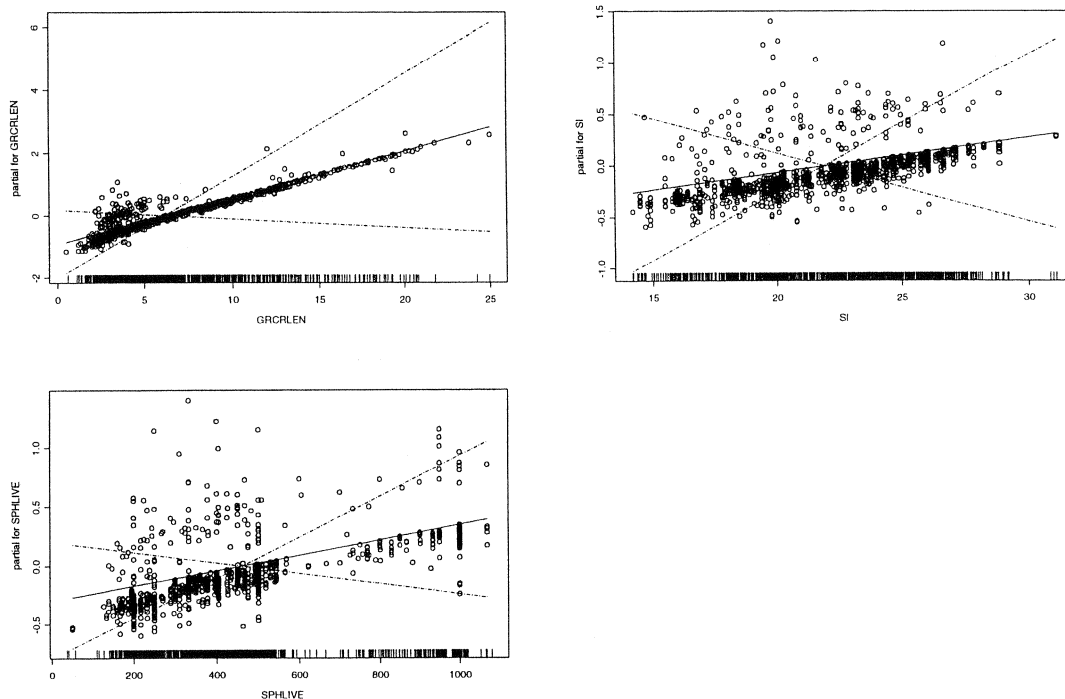
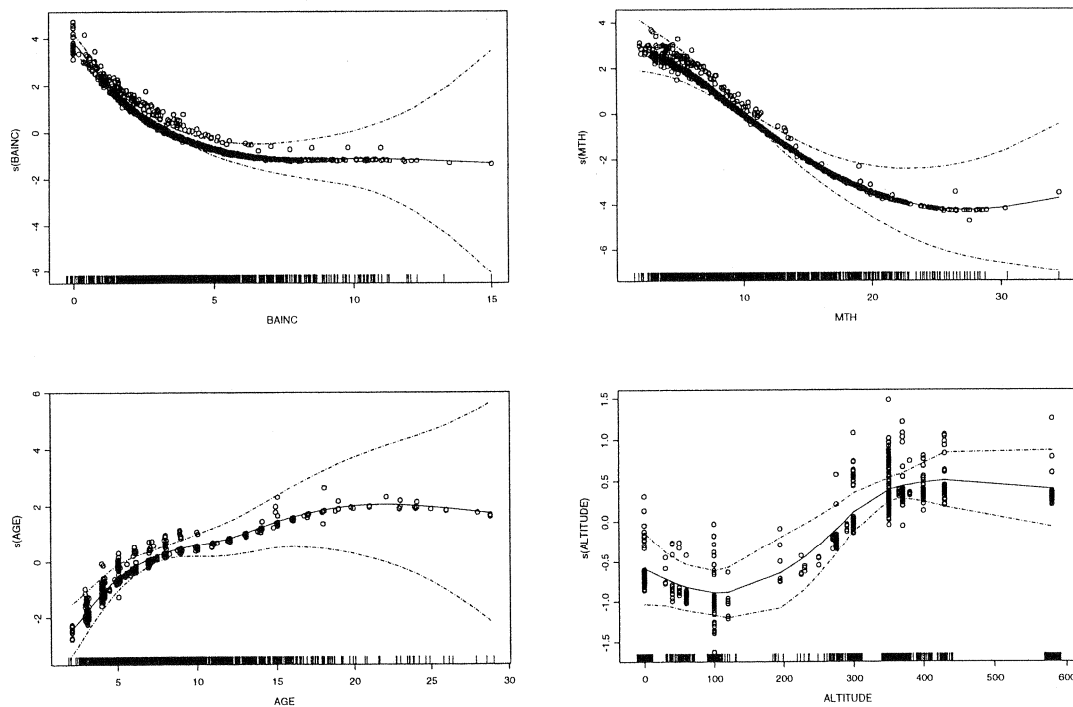


Figure 6-2. Partial Regression plots for Model 6-1.

Partial Regression plots for Generalized Additive Model 6-1:



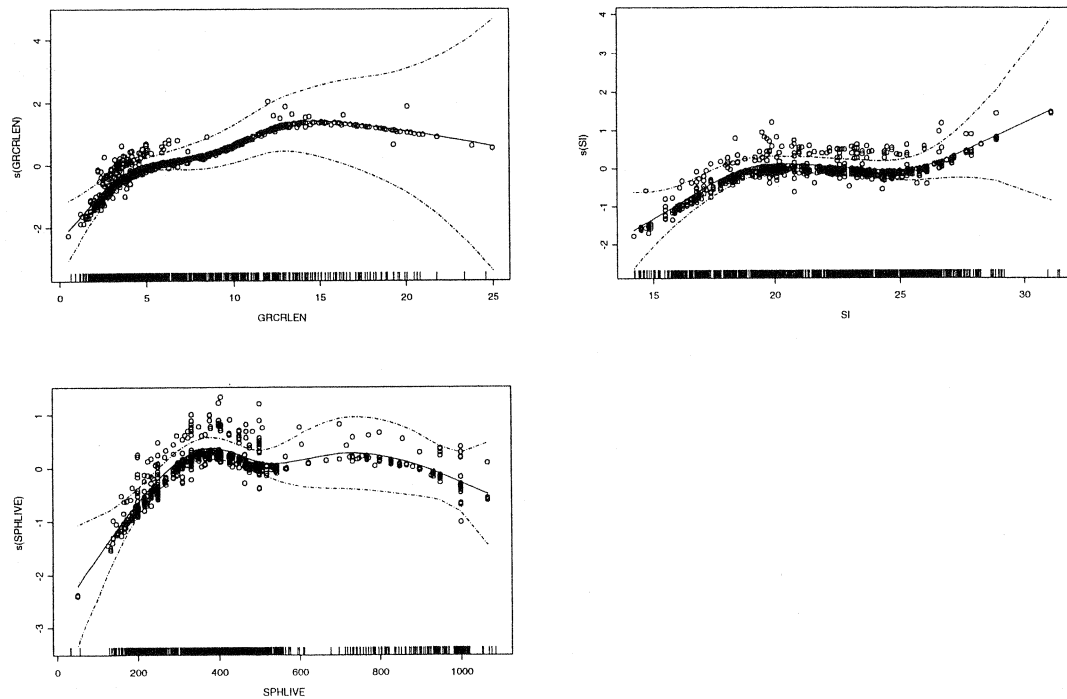


Figure 6-3. Partial Regression plots for GAM Model 6-1.

Fit Plots for GAM Model 6-1:

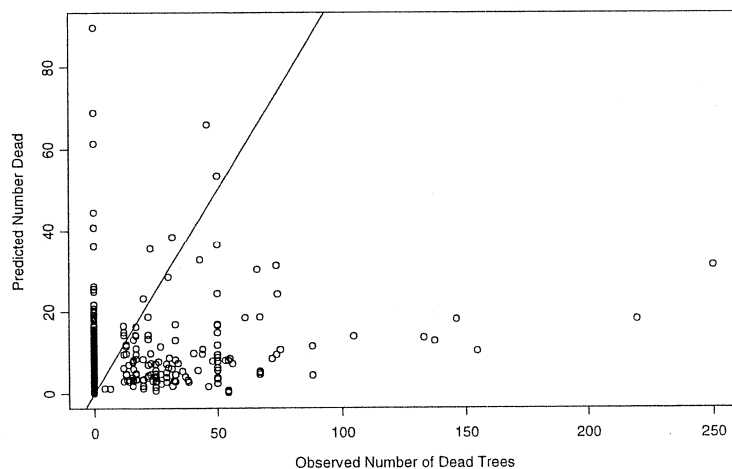


Figure 6-4. Observed versus Predicted Values for GAM Model 6-1.

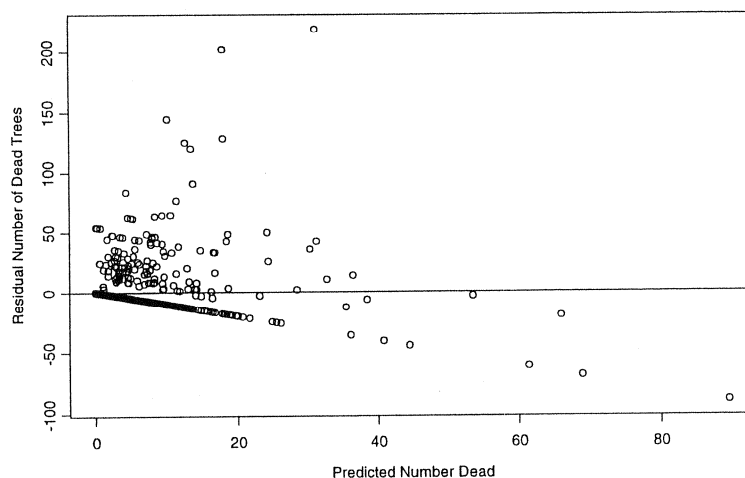


Figure 6-6. Residual plot for GAM Model 6-1.

Discussion:

The fit of the model as represented in figures 6-4 and 6-5 was found to be most inadequate given the drastic overall underprediction of mortality. Since even the GAM's did not perform satisfactorily, this suggests the information that describes why the trees die is simply not available. It was thought that the poor performance of the model may be attributed to the fact that wind and/or exposure information (which is often considered the main cause of mortality given the exposed places shelterbelts are planted) was not included in the modelling process. The incorporation of such information was considered, however it was realised that most users would not have such information available to them.

Therefore an alternative approach was suggested and used to model mortality which simply describes the absolute stocking per kilometre at the end of an annual interval given the stocking and the age at the beginning of the interval.

Alternative Approach:

At the suggestion of Mr Andy Gordon (pers comm) the following model form which has proven useful in growth models was tried.

Model 6-2: Stems per Km at end of interval = $(1 / (-k * ((1 / \sqrt{\text{age at the end of the interval}}) - 1 / (\sqrt{\text{Age}}))) + (1 / \sqrt{\text{Stems at the beginning of the interval}})) ** 2$;

Coefficients:

$$k = 0.0037518057$$

Fit and Residual plots:

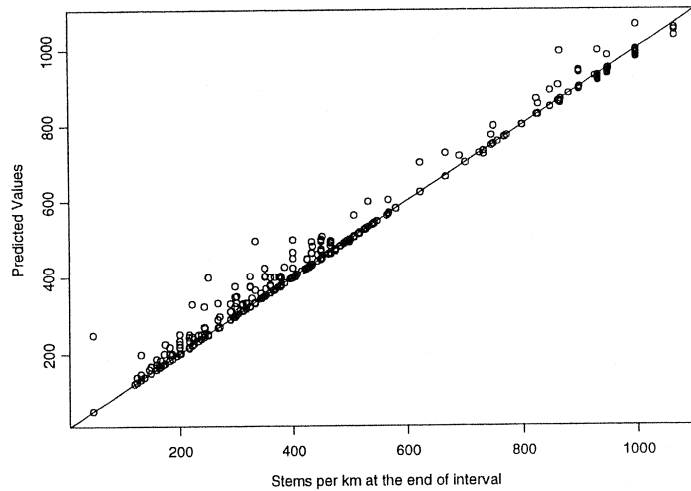


Figure 6-5. Observed versus predicted values for Model 6-1.

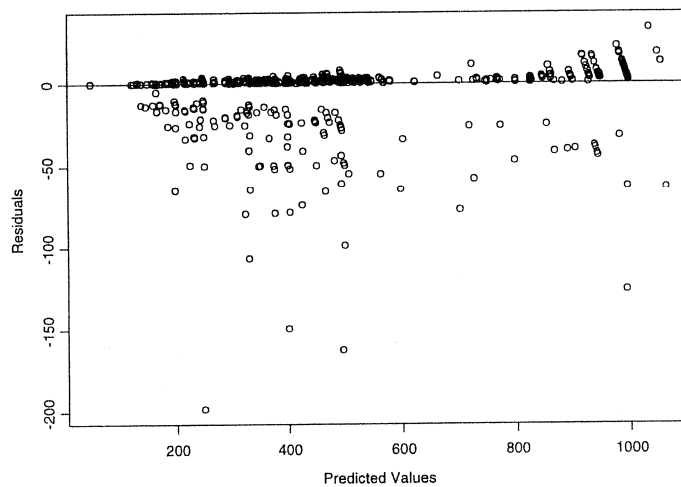


Figure 6-7. Residual plot for Model 6-1.

Fit Statistics for Model 6-2:

RMS = 225.78 %, Regression SS/Total SS = 99.9%

6.5 Conclusions:

Various models to predict mortality were fitted against the data set (including Logistic additive models, Poisson additive models and nonlinear regression models). Because mortality is erratic and was only found to be weakly related to the standard plot measurements, a satisfactory estimate of mortality was found difficult to obtain. The selected model based on a form previously used for growth demonstrated an acceptable fit which was reflected in both the fit plots and the fit statistics.

Part Seven: Estimating Diameter Distributions

7.1 Introduction:

A diameter distribution function was estimated so that for any given mean basal area it was possible to show how that basal area was distributed across a stand. This information is used in deriving log grades, for example, where the distribution parameter could be significant.

7.2 Exploratory Data Analysis

The exploratory analysis involved extracting the individual diameter measurements for each tree within each plot, converting them into basal area measurements and examining their distributions. The shape of the basal area distributions at different age classes are illustrated in figure 7-1 below.

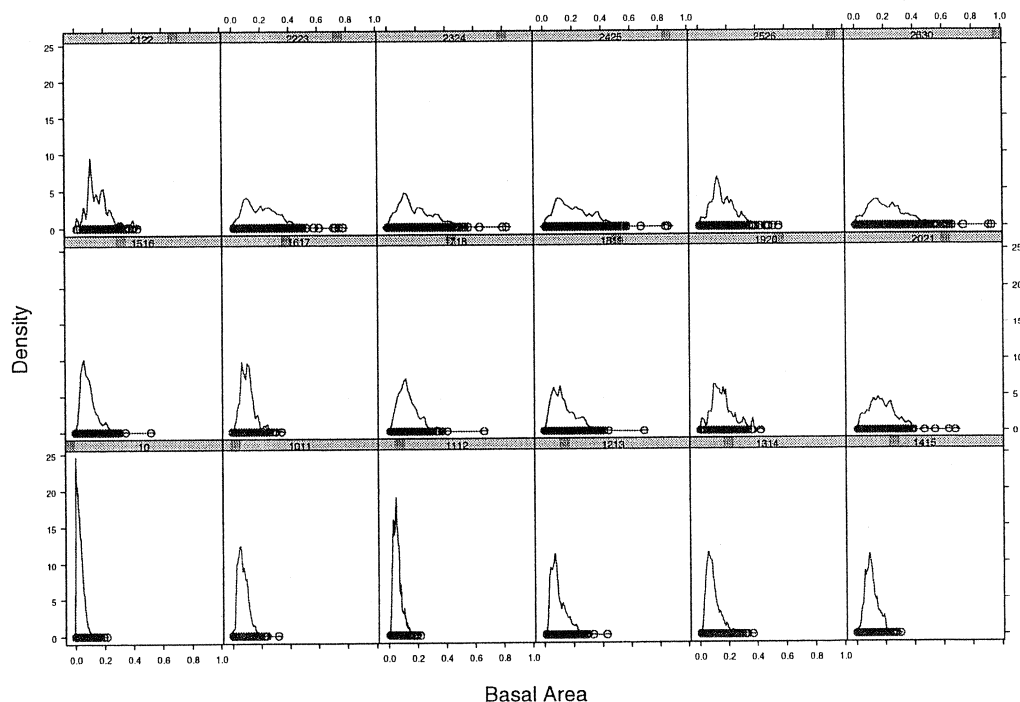


Figure 7-1. Basal Area distributions at different age classes for all of the shelterbelt data, where 1213 refers to trees aged between 12 and 13 and 2021 refers to trees aged between 20 and 21 etc.

Figure 7-1 shows that as the mean basal area increases so does the variation about the mean. This commonly resulted in substantial right skewness in the older age class distributions (which is also evident in figure 7-1).

Data Specifications:

Age Class	No' of observations	Mean Basal Area
10	19420	0.0278
1011	1755	0.0669
1112	1043	0.0561
1213	765	0.0940
1314	1176	0.0869
1415	587	0.1006
1516	717	0.0954
1617	364	0.1063
1718	475	0.1384
1819	455	0.1404
1920	193	0.1623
2021	277	0.1832
2122	172	0.1708
2223	469	0.2145
2324	330	0.2070
2425	287	0.2251
2526	194	0.1821
2630	539	0.2321

Table 7-1. Summary information for data used in estimating diameter distributions.

7.3 Comparing various distributions with the actual data

Before the fitting of any one particular distribution was carried out, the basal area data was fitted with various distributions for each age class. This enabled the most appropriate distribution to be chosen given the data available. The graphical results are illustrated below while the distribution parameters are listed in Appendix 14.

Note: The green line represents the empirical distribution, the red line indicates a normal distribution, the pink line represents a Lognormal distribution, the orange line the exponential distribution and the blue line represents the Weibull distribution.

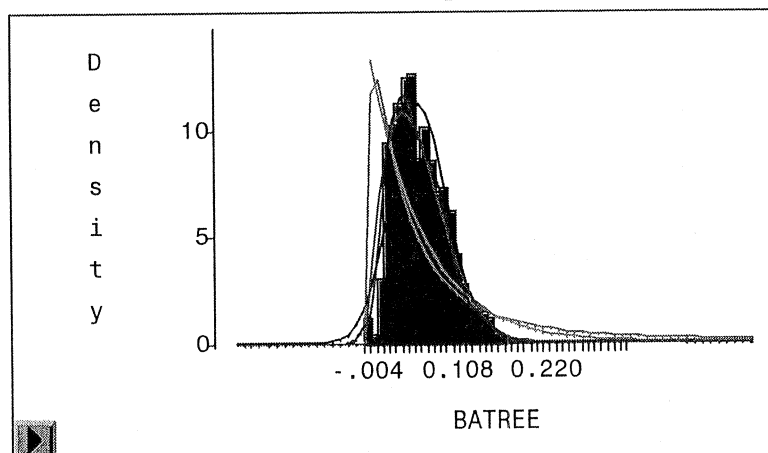


Figure 7-2 . Basal Area Distribution for 10-11 year old trees.

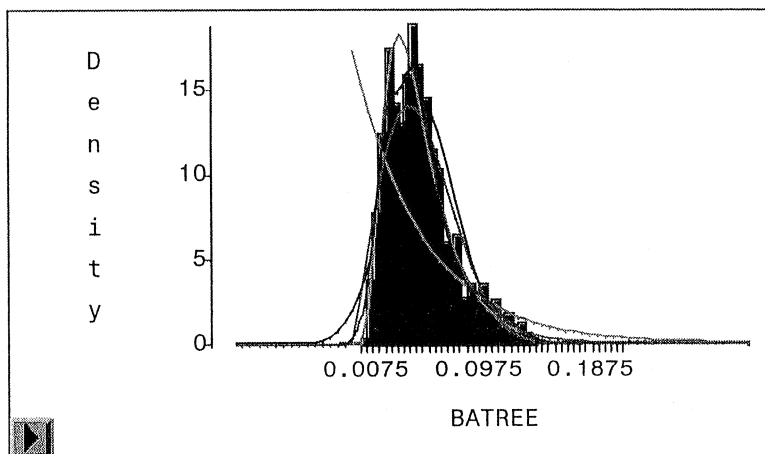


Figure 7-3 . Basal Area Distribution for 11-12 year old trees.

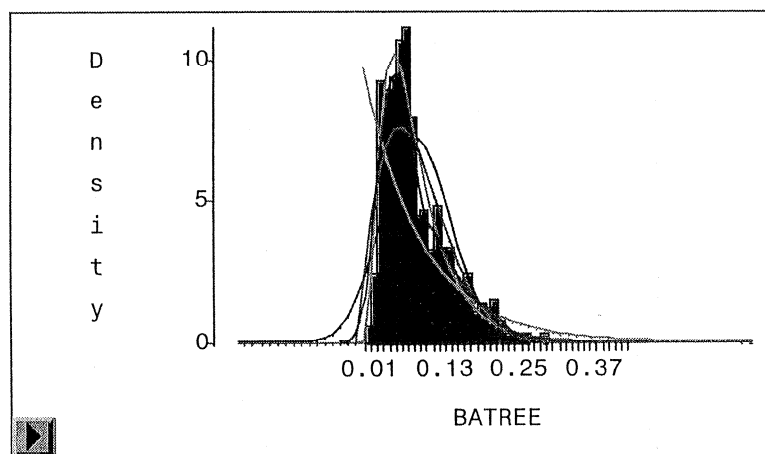


Figure 7-4. Basal Area Distribution for 12-13 year old trees.

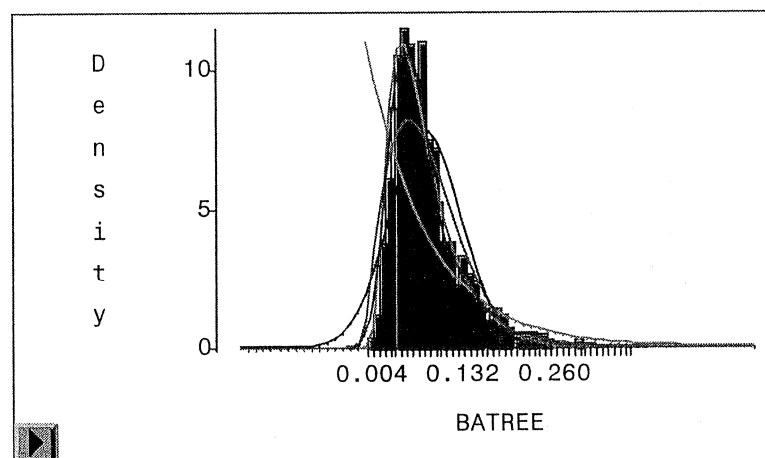


Figure 7-5. Basal Area Distribution for 13-14 year old trees.

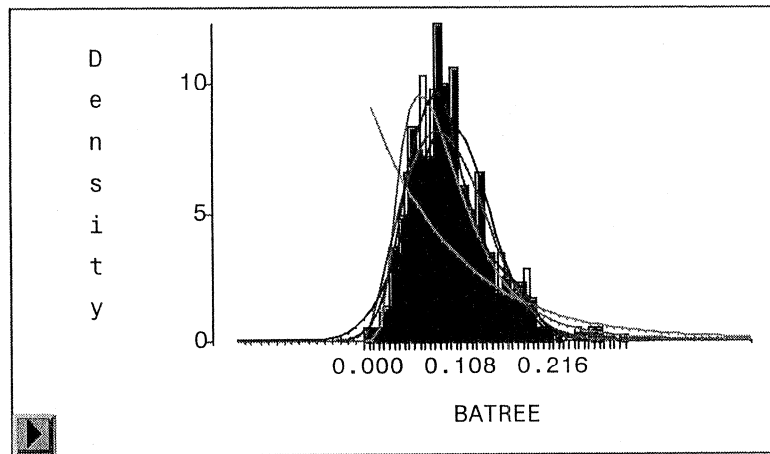


Figure 7-6. Basal Area Distribution for 14-15 year old trees.

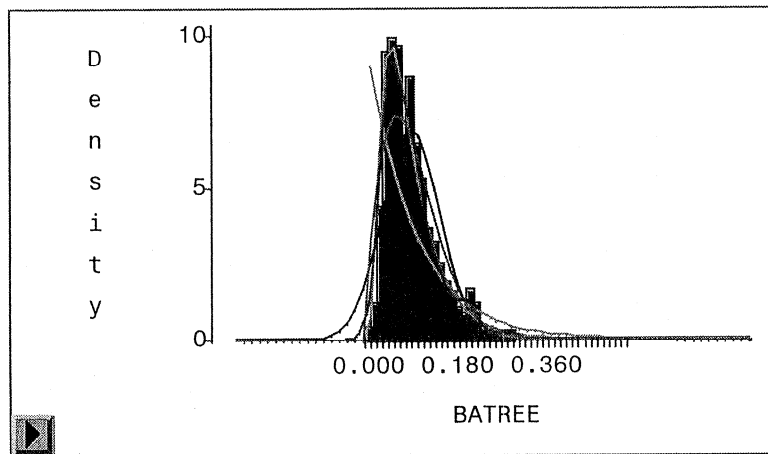


Figure 7-7. Basal Area Distribution for 15-16 year old trees.

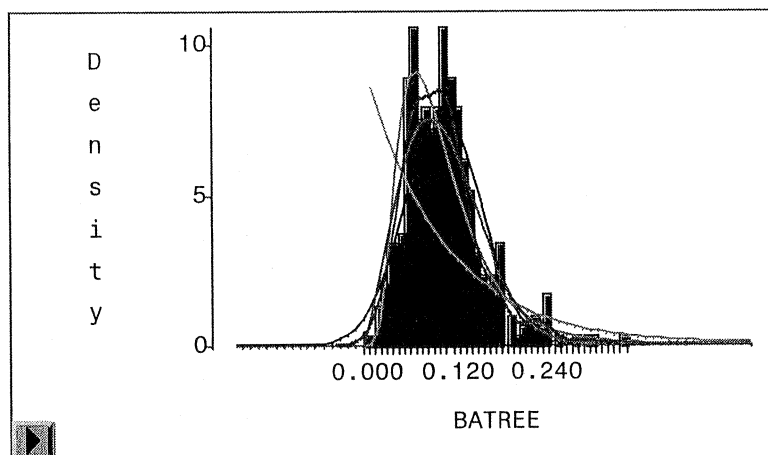


Figure 7-8 . Basal Area Distribution for 16-17 year old trees.

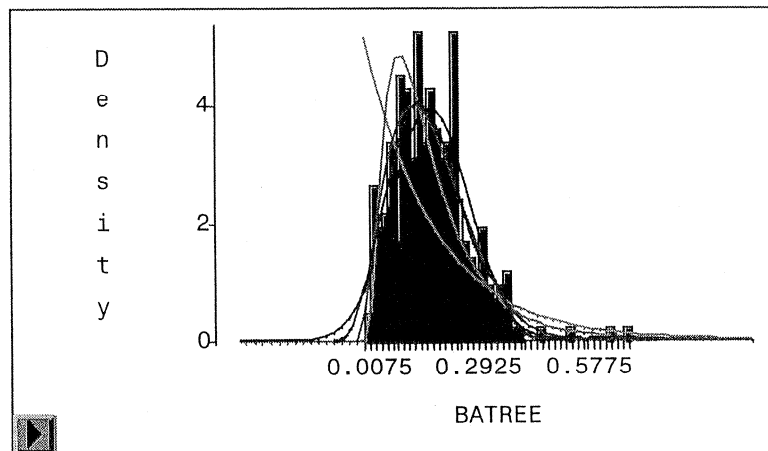


Figure 7-12. Basal Area Distribution for 20-21 year old trees.

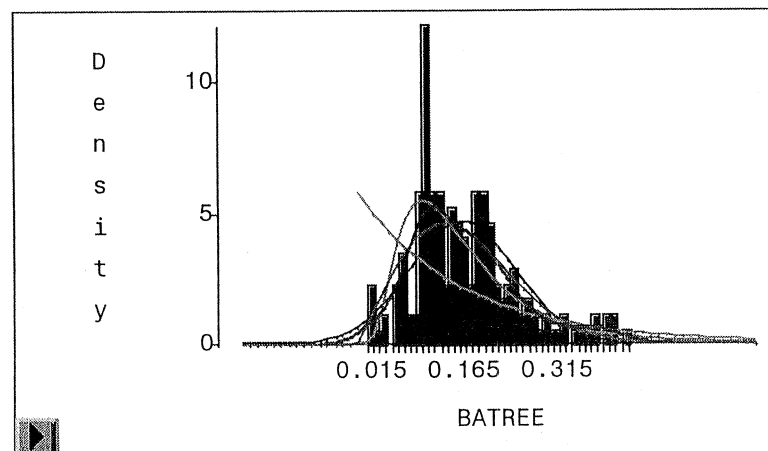


Figure 7-13. Basal Area Distribution for 21-22 year old trees.

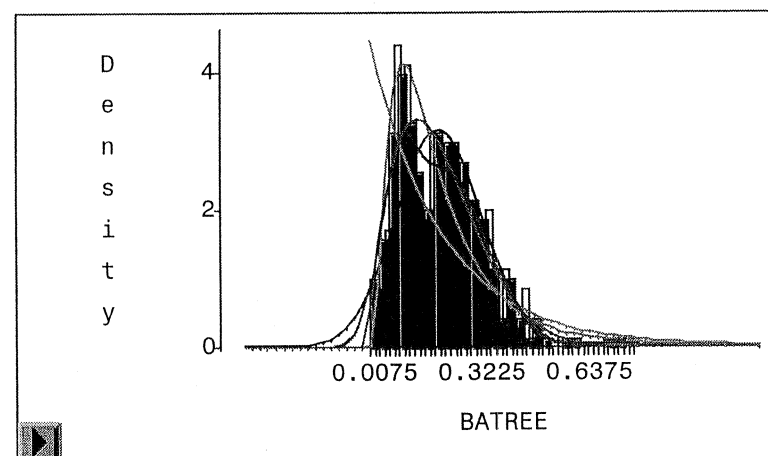


Figure 7-14. Basal Area Distribution for 22-23 year old trees.

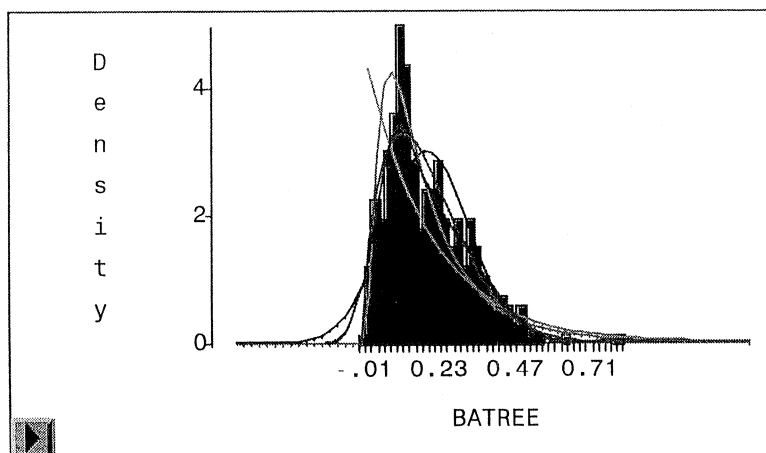


Figure 7-15. Basal Area Distribution for 23-24 year old trees.

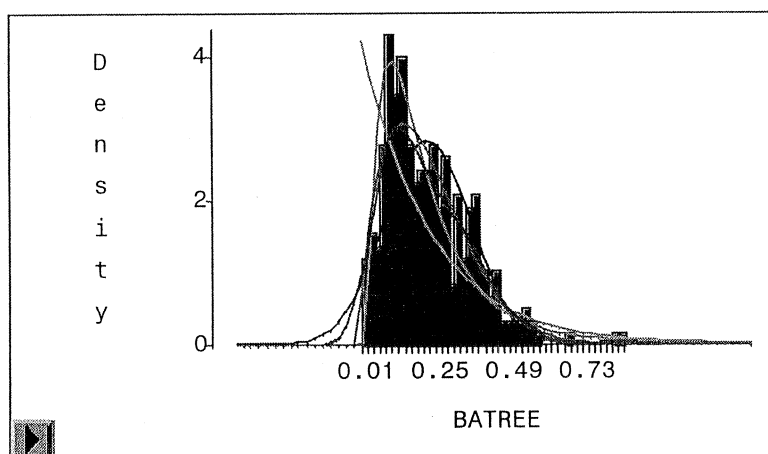


Figure 7-16. Basal Area Distribution for 24-25 year old trees.

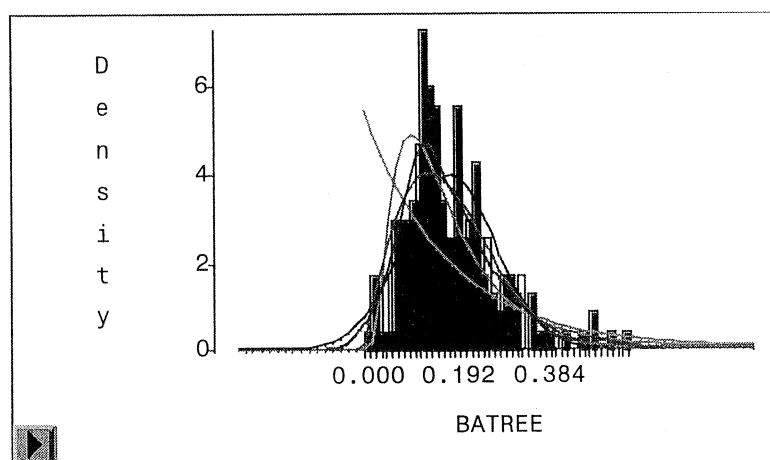


Figure 7-17. Basal Area Distribution for 25-26 year old trees.

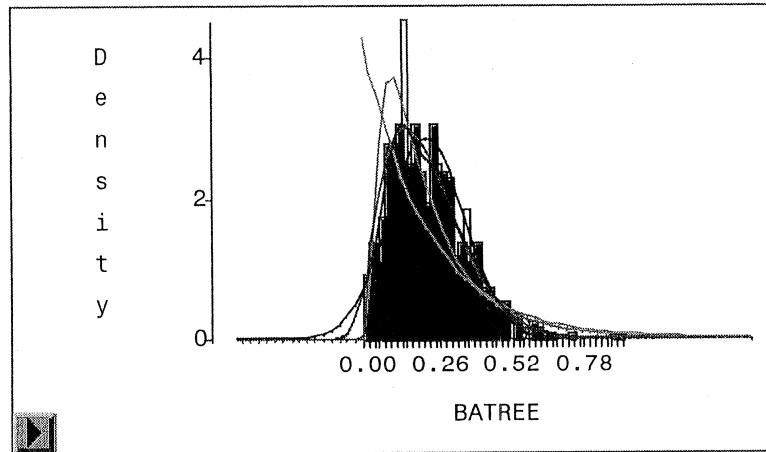


Figure 7-18. Basal Area Distribution for 26-30 year old trees.

7.4 Distribution Results:

The figures above show that the empirical distribution was most closely matched by the Lognormal distribution for the younger age groups (1213, 1314, 1516) similarly matched by the Lognormal and the Weibull distributions for some of the intermediate age classes (1112, 1415, 1617, 1718) and more closely matched by the Weibull distribution for mostly the older age groups (1011, 1819, 1920, 2021, 2122, 2223, 2324, 2425, 2526, 2630). This was confirmed by the results of the fit and the Kolmogorov-Smirnov tests.

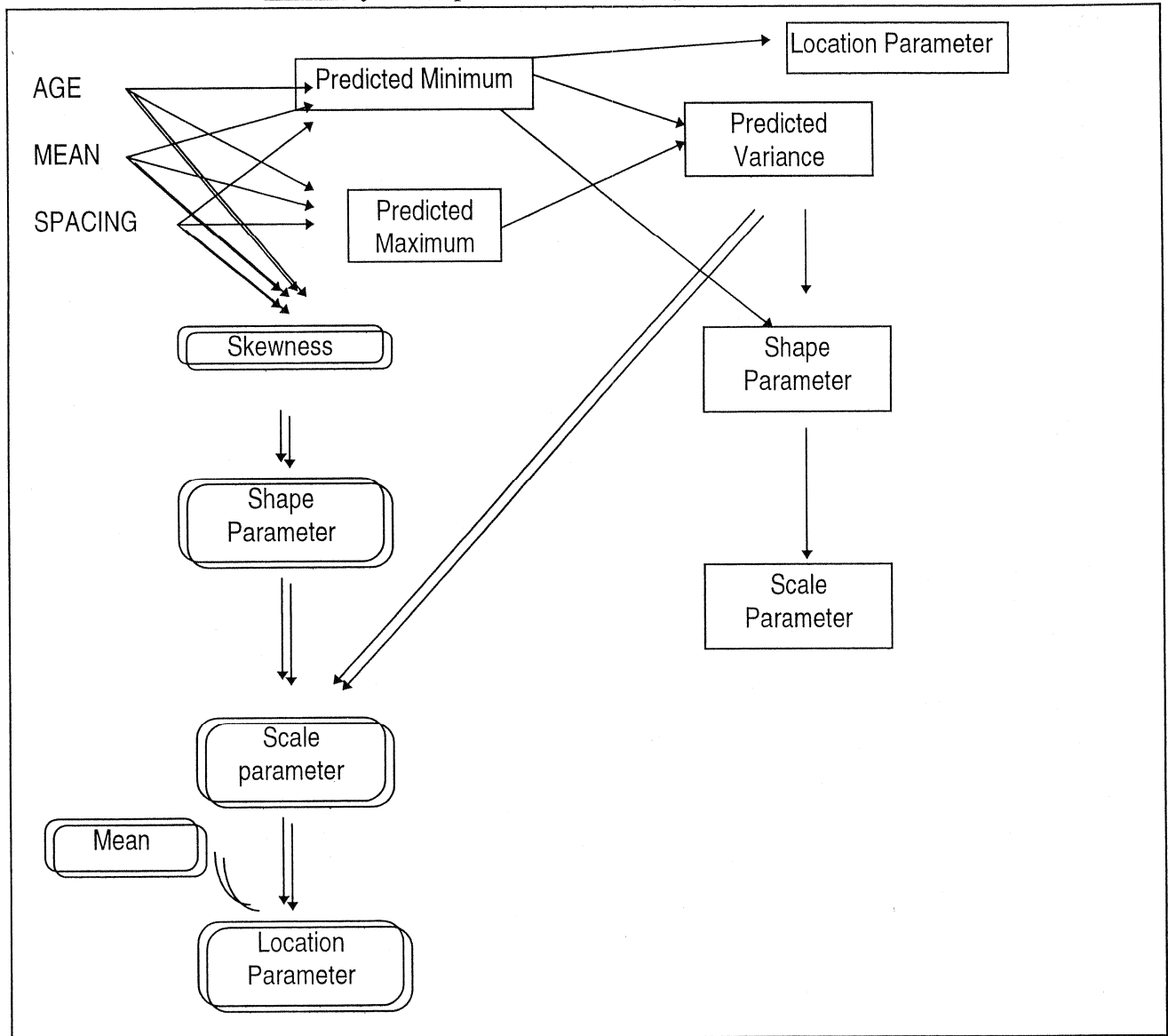
Relevant parameters for the Normal, Lognormal, Exponential and Weibull distributions were estimated using maximum likelihood for each age class and are listed in Appendix 14. In addition, each estimated distribution was tested against the data using the Kolmogorov-Smirnov test which identified which distributions were well fitted.

In summary, the Lognormal satisfied the Kolmogorov-Smirnov test for age classes 13-14 and 15-16 while the Weibull distribution satisfied the Kolmogorov-Smirnov test for the older age classes: 20-21, 23-24, 25-26 and 26-30.

Given the distribution fit results and the fact that harvest usually involves older trees, the Weibull distribution was chosen and fitted using three methods. Firstly the parameters were estimated using maximum likelihood (Appendix 14), secondly they were estimated in a standard way using moments and thirdly parameters were obtained using a method recently published involving skewness information (Lindsay et al, 1996). A summary of the pathways for the two methods of estimating moments is as follows:

7.5 Weibull Distribution Results:

Summary of the parameter estimate pathways



Note: The shadowed arrows and boxes represent those pathways relevant for the Lindsay et al.'s recently developed skewness method, and the single arrows and boxes represent those for the standard method.

7.5.1 Weibull Estimation using Moments:

Estimation of Weibull parameters using moments was performed using two approaches: a standard method and a technique using skewness information (Lindsay et al, 1996). For all plots at all ages the two methods were carried out and compared.

The standard procedure required estimates of the mean, minimum, maximum and subsequently the variance of the basal area distributions, while the new method also required an estimate of skewness to estimate the three parameters: the location, scale and shape parameters. Parameter estimates for the Weibull distribution were obtained

for each age for each plot (using both techniques) and the variation within the coefficients over time for selected plots was investigated. This was performed to ascertain if different parameter values were necessary for each age class. A sample of the parameter results and the resulting distributions from the two methods from a selection of regions are illustrated below.

Auckland

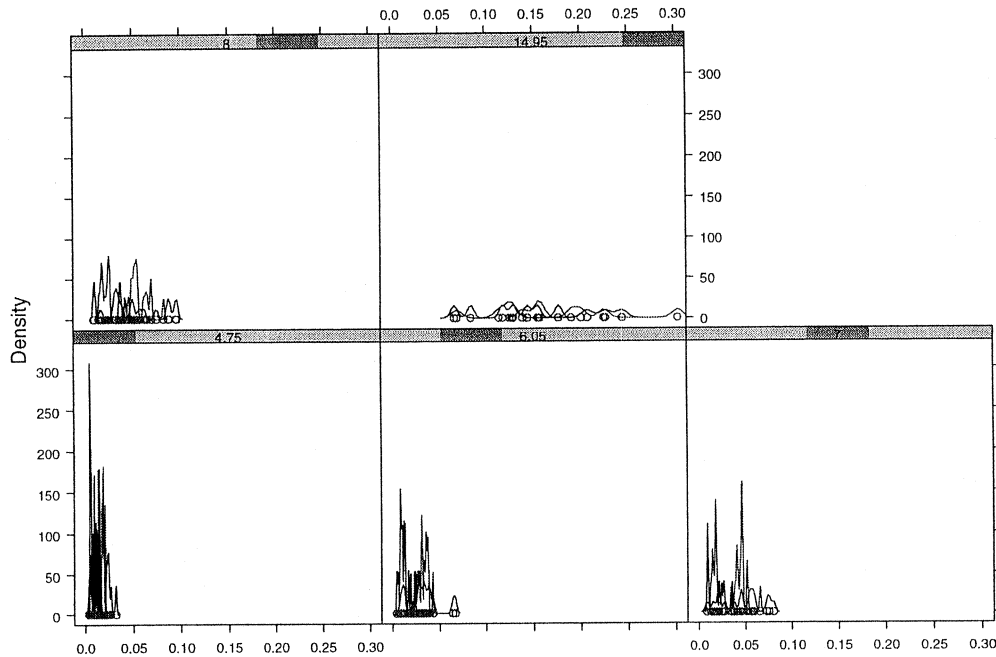


Figure 7-19. Distribution of the basal area within a stand situated in the Auckland region for each age. The green line represents the observed distribution while the blue line represents the distribution obtained using the standard moments method, while the red line indicates that obtained using skewness information.

Figure 7-19 shows that the distribution obtained using the standard approach seemed to fit more closely than that obtained using skewness information. Overall, the predicted distribution underestimated the basal area of the stand for the older trees. The results of the parameter estimates obtained for the shape and scale parameters using the standard method are illustrated below.

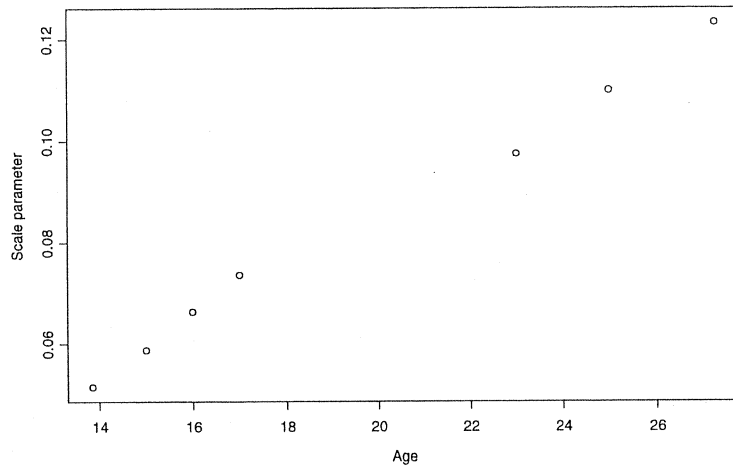


Figure 7-24. A Scatterplot representing the change in the scale parameter over time within a stand situated in Canterbury.

Figure 7-23 and 7-24 also show that scale parameter increased with the mean in basal area however once again the shape parameter only changed fractionally over time.

Hawkes Bay:

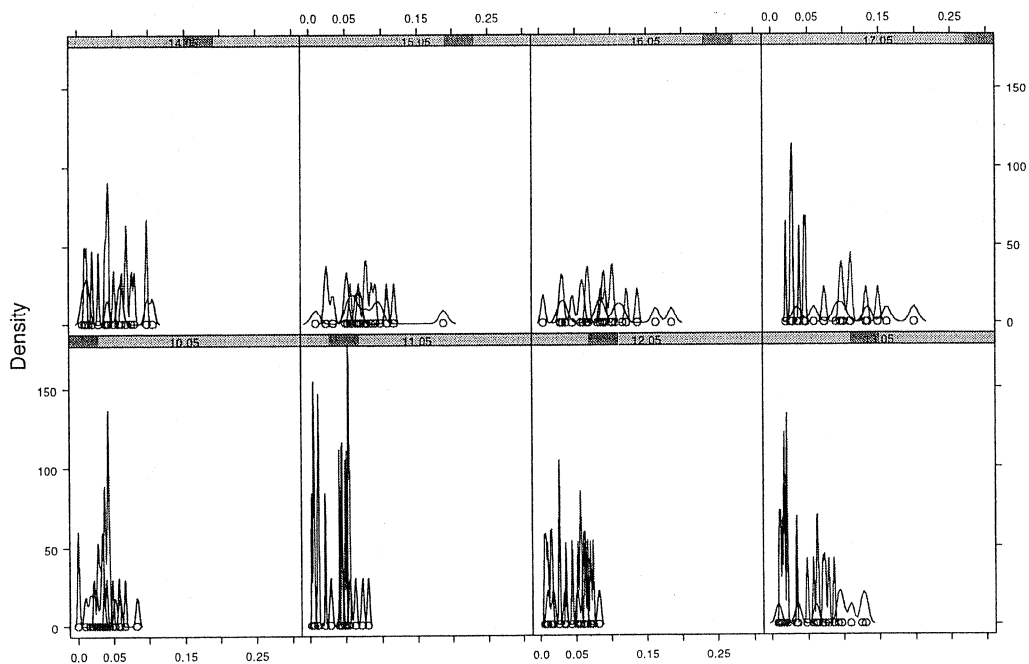


Figure 7-25. Distribution of the basal area within a stand situated in the Hawkes Bay region for each age. The green line represents the observed distribution while the blue line represents the distribution obtained using the standard moments method, while the red line indicates that obtained using skewness information.

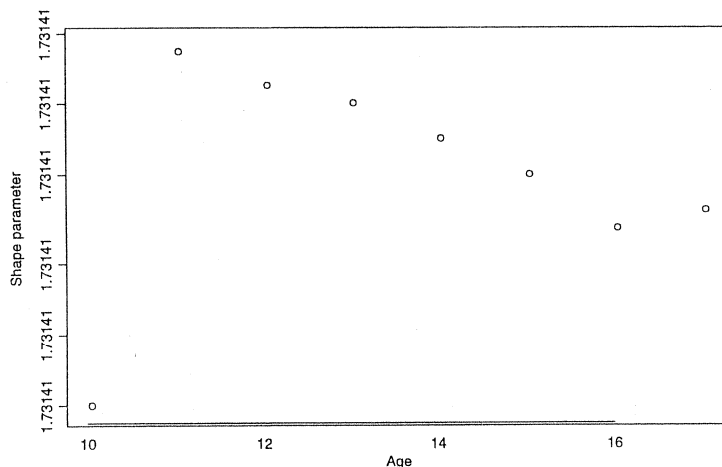


Figure 7-26. A Scatterplot representing the change in the shape parameter over time within a stand situated in Hawkes Bay.

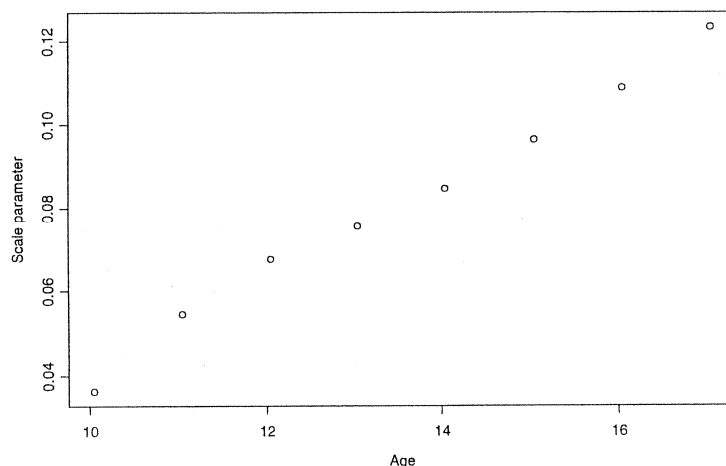


Figure 7-27. A Scatterplot representing the change in the scale parameter over time within a stand situated in Hawkes Bay.

Figure 7-25 shows that the distribution estimated using skewness information produced worse fitting results at all ages than those obtained using the standard method. In addition, the shape parameter estimated for each age (using the standard approach) (figure 7-26) within each plot showed little variation suggesting that there may be no need for separate shape parameters for each age. Given the small amount of variation in the shape parameter only one was estimated for all ages.

Given that little variation in the shape parameter was seen within plots across ages, one shape parameter was deemed sufficient for all ages. This required variance estimates (using both approaches) which also involved estimates of the minimum and maximum.

Estimation of Variance for Weibull parameter estimation using Moments:

The variance is required to estimate the Weibull parameters using moments and it was considered that the variance would be related to many different predictors (age, maximum, minimum, mean, stems per kilometre and basal area). This was investigated visually and is represented using scatterplots below:

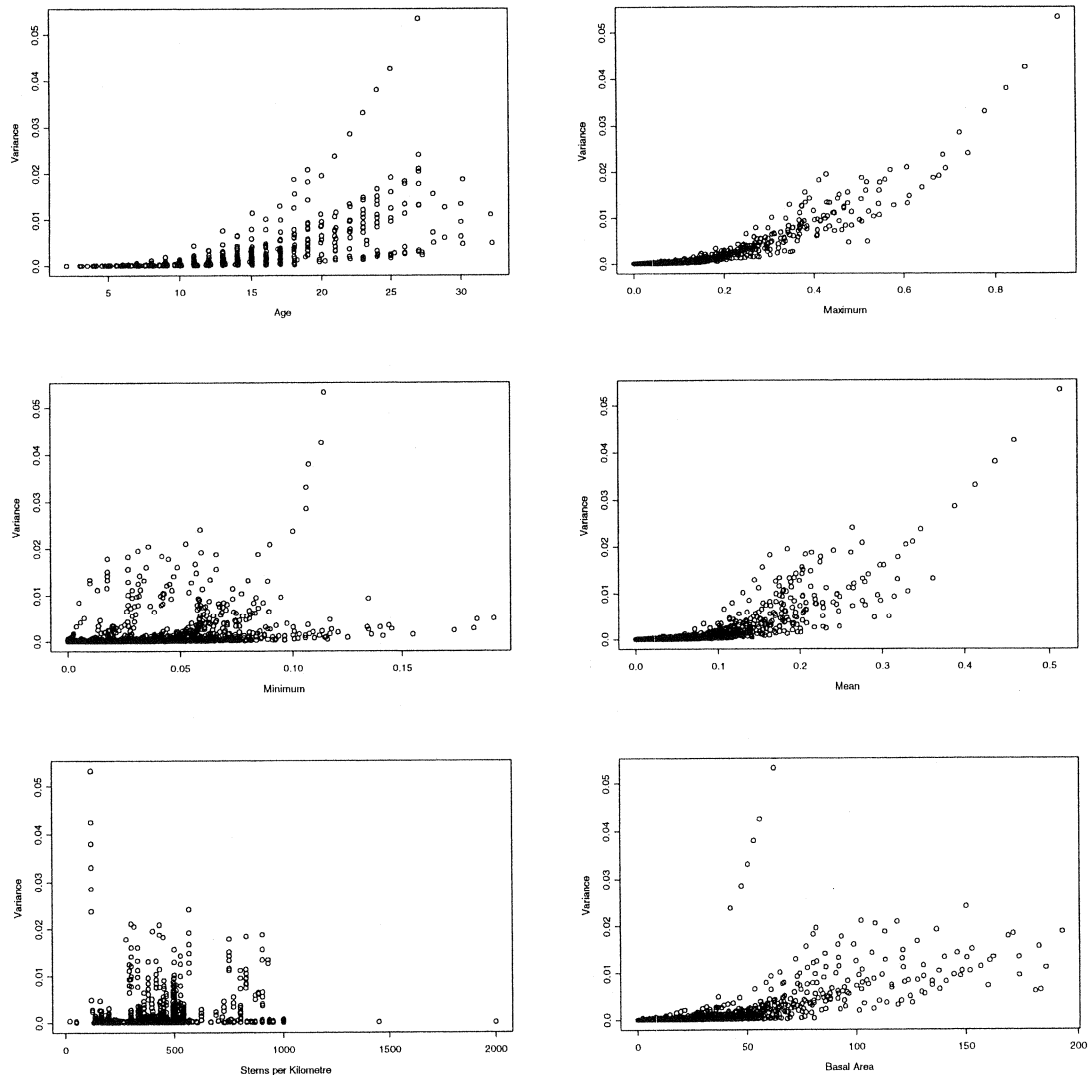


Figure 7-28. Scatter plots representing the relationship between variance and commonly used explanatory variables.

As expected the variance appeared to be related to all of the variables above (only weakly for stems per kilometre perhaps). However, since the mean basal area is predicted from the model and age and basal area were too highly correlated to be fitted together only the minimum, maximum and stocking variables were used to predict variance.

As shown in the earlier chapters, the variables were first fitted with GAM's and then with nonlinear or linear functions suggested by the GAM's. The results are as follows:

Estimation of variance for each age given age, minimum value, maximum value and spacing.

GAM Model 7-1: $\text{Variance} = s(\text{age}) + s(\text{minimum}) + s(\text{maximum}) + s(\text{spacing})$

Partial Regression plots for Model 7-1:

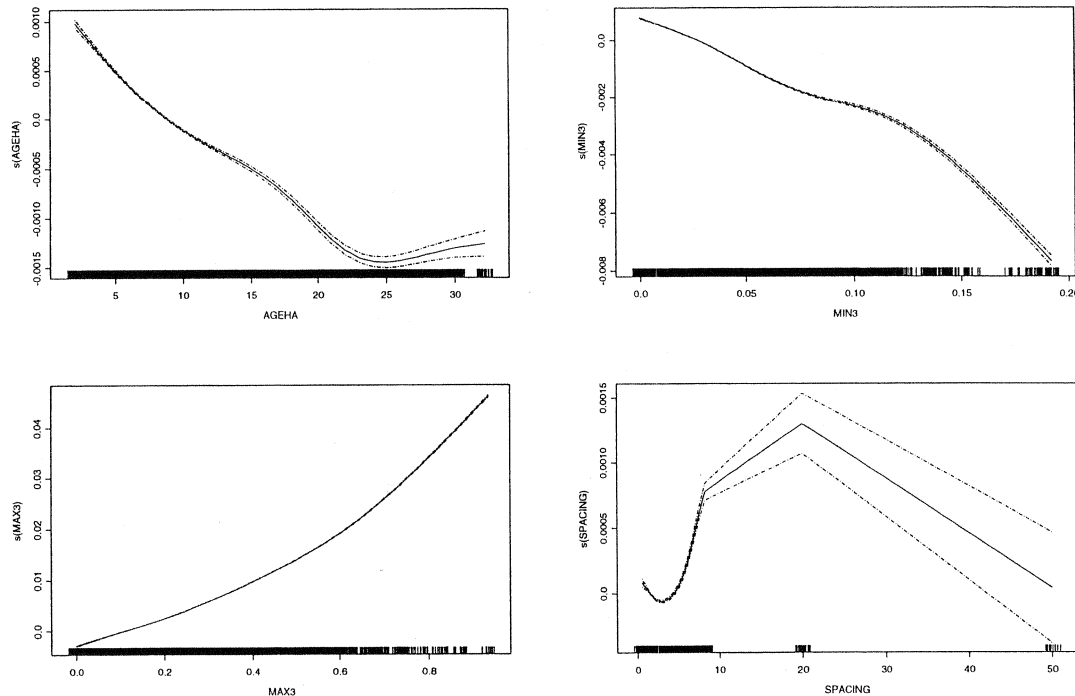


Figure 7-29. Partial Regression plot for GAM estimating the Variance.

Fit plot for Model 7-1:

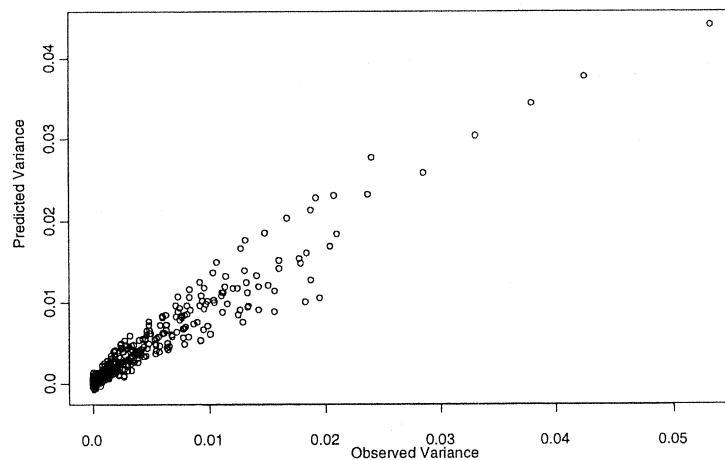


Figure 7-30. Observed versus predicted values for GAM predicting variance.

Fit Statistics:

Null Deviance = 0.3849797, Residual Deviance = 0.0193596, Deviance Explained = 94.97%

Given the satisfactory fit of the model, the maximum and minimum were estimated (since they will not always be available to the user) and the predicted values used to estimate the variance.

Estimation of the minimum value for each age given age, mean basal area, and spacing.

GAM: Model 7-2: Minimum basal area = $s(\text{age}) + s(\text{basal area}) + s(\text{Spacing})$

Partial Regression plots for Model 7-2:

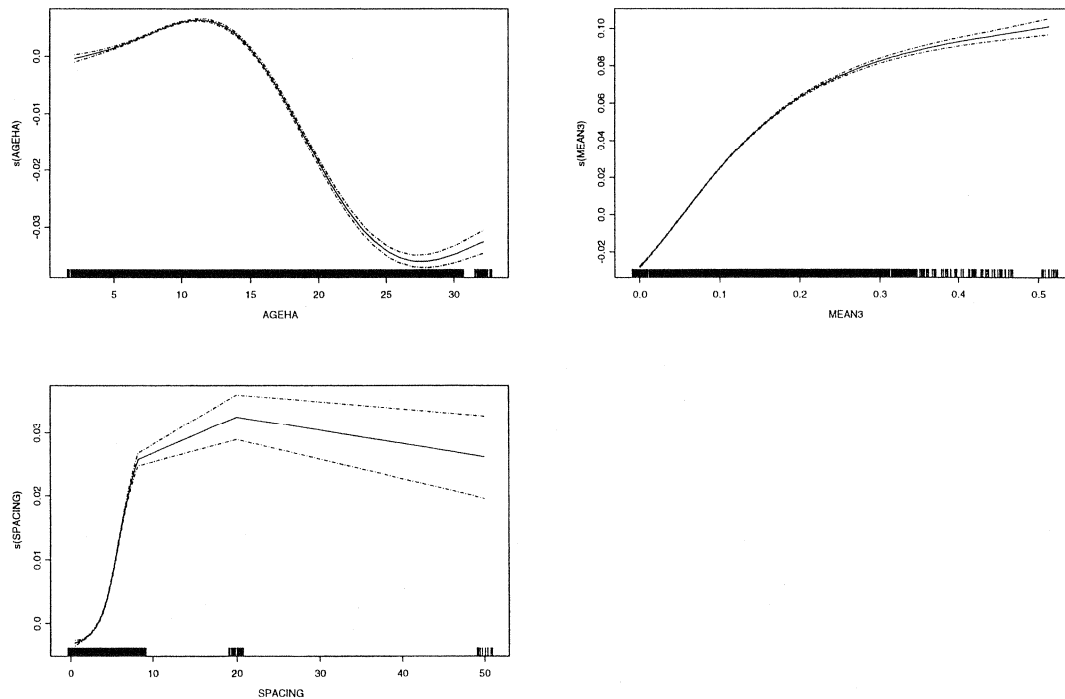


Figure 7-31. Partial Regression plots for model 7-2.

Fit Statistics:

Null Deviance = 20.38433, Residual Deviance = 4.657806, Deviance Explained: 77.2%

Nonlinear model for Minimum:

Model 7-3: Minimum value = $a * \text{Age} + a2 * (1 - \exp(-k * \text{Mean Basal Area})) + f * \text{Spacing}$

Coefficients:

$a = -0.002132557$, $a2 = 0.144616063$, $k = 7.300088319$, $f = 0.001502470$

Fit Statistics:

RMS = 0.00189796, Regression SS/Total SS = 85.8%

Estimating the Maximum Value using GAM's:

GAM : Maximum Value = $s(\text{Age}) + s(\text{Mean Basal Area}) + s(\text{Spacing})$

Partial Regression plots:

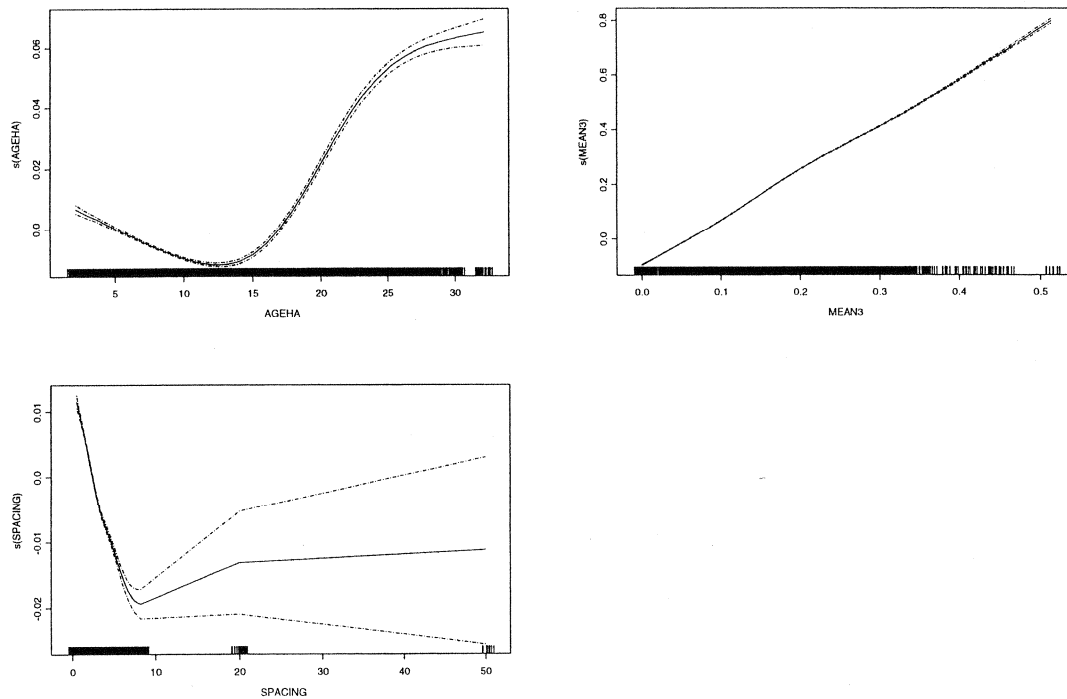


Figure 7-32. Partial regression plots for GAM predicting the maximum value (Model 7-3).

Fit Statistics:

Null Deviance = 436.225, Residual Deviance = 22.17412, Deviance Explained: 94.9%

Nonlinear model for maximum:

Model 7-4: Maximum value = $(a / (1 + \exp(f - g * \text{Age}))) + d * \text{Spacing} + e * \text{Mean Basal Area}$

Coefficients:

$a = 0.08065080$, $f = 22.60628254$, $g = 1.17548118$, $d = -0.00211949$, $e = 1.62365160$.

Fit Statistics

RMS = 0.00080371, Regression SS/Total SS = 96.7%

Nonlinear estimation of Variance using fitted values of Maximum and Minimum values:

Model 7-5: Variance = exp(a + j* Predicted Maximum)+g* Spacing +(1-exp(-exp(a2 - h* Predicted minimum value)))

Coefficients:

a = -6.78088189, j = 4.42528076, g = -0.00045303, a2 = -9.73392685,
h = -44.16808035

Fit Statistics:

RMS = 0.00000268607, Regression SS/Total SS = 99.9%

Estimating the shape parameter using the standard moments:

Predicted Variance = (((Mean - predicted minimum)**2)/(gamma(gamma(1+1/c))))*(gamma(1+2/c) - (gamma(gamma(1+1/c))))

Coefficients:

c = 1.285201339

Fit Statistics:

RMS = 0.00000069664, Regression SS/Total SS = 94.12%

Estimating the shape parameter using skewness information:

Skewness = (gamma(1+3/c3)- 3 *(gamma (1+1/c3) *gamma (1+2/c3)) +2*(gamma (1+1/c3))* (gamma(1+1/c3))*(gamma(1+1/c3))) / ((gamma(1+2/c3)- (gamma(1+1/c3)) * (gamma(1+1/c3)))**0.5)

Coefficients:

c3 = 2.318017223

Fit Statistics:

RMS = 0.490270, Regression SS/Total SS = 1.1%

7.6 Results of fitting the Weibull distribution using moments (standard approach) and estimates of minimum, maximum, and variance.

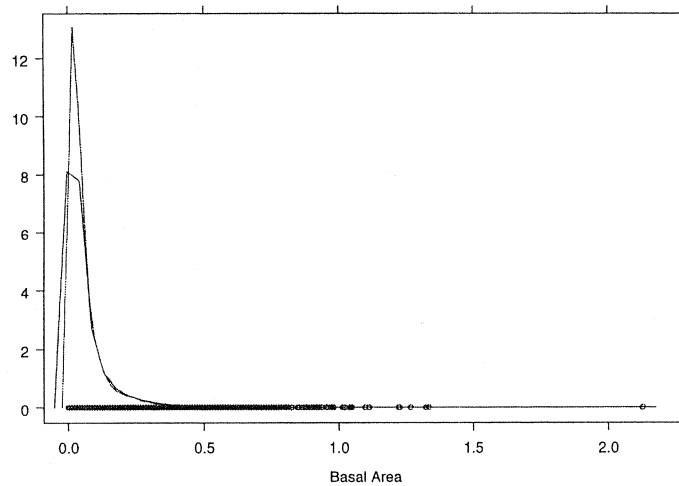


Figure 7-33. Observed and predicted basal area values for plots at all ages. Note: the green line is the observed distribution across all ages, the red line is the estimated distribution using the real mean and the estimated shape parameter.

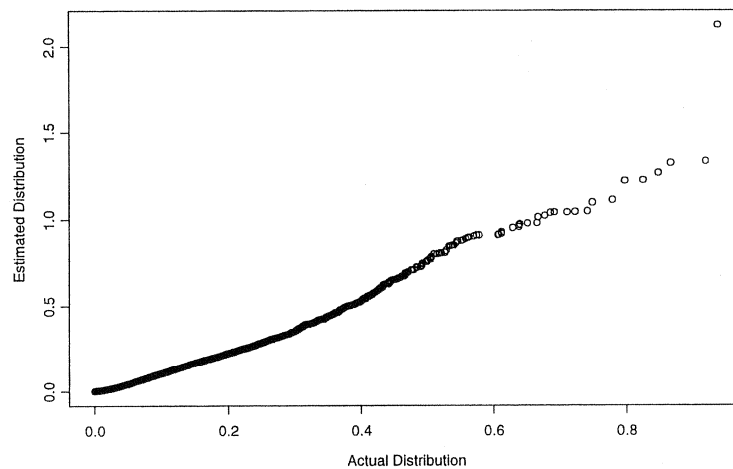


Figure 7-34. Quantile-quantile plot for the observed and predicted distributions for all plots at all ages.

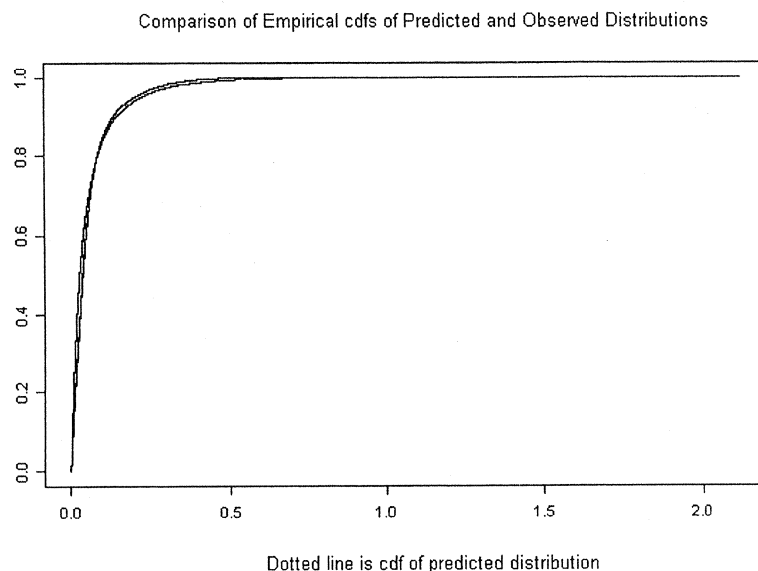


Figure 7-35. Cumulative density function of the predicted and observed distributions

Discussion:

The observed and predicted distributions for all of the data at all ages were compared using a distribution plot (figure 7-33), QQ- Plot (figure 7-34) and the Kolmogorov-Smirnov test. Figure 7-34 illustrates approximate linearity indicating that both the observed and predicted distributions were of a similar shape. While the cumulative distributions for both the predicted and observed also appeared similar (figure 7-35).

Figure 7-35 also highlighted the inadequacies of the fit given that the predicted distribution ceased at a value of approximately 0.7 while the actual distribution continued past this point. This underprediction of basal area was also evident in the density plot (figure 7-33) which was large enough for the Kolmogorov-Smirnov test to find the predicted and observed distributions to be significantly different from one another.

Conclusions:

The shape of the basal area distributions of shelterbelts at different ages was explored and various diameter distributions were fitted and compared. It was found that the Weibull distribution was most suitable for the older trees therefore this distribution was fitted using two methods. While both methods involved fitting the Weibull distribution using moments, one of the methods also used skewness information in the estimation of the shape parameter. This recently developed method was only found to improve the fit of the distribution in a small number of cases (and decrease the fit in others) therefore the standard approach was adopted. This standard approach involved the estimation of the minimum, maximum basal area values for shelterbelts in order to arrive at estimates for the location, scale and shape parameters. These estimates were obtained for each age for each shelterbelt, however it was found that one shape parameter for shelterbelts of all ages was sufficient. The results for the

shelterbelts at all ages showed a reasonable fit with a tendency for underprediction of basal area for a subset of values.

Part Eight: Predicting basal area remaining after a thinning

8.1 Introduction:

The sample plot data set available for shelterbelts contained very few shelterbelts which had been thinned during the measurement period. To obtain such data, a dummy thinning exercise was performed in order to quantify the percentage of basal area that is removed when a percentage of the stems are thinned. While it is likely that most shelterbelts will be unthinned during their life, it is considered necessary to include a function which predicts the effects of thinning on the remaining shelterbelt basal area to provide modelling flexibility.

Trees were chosen to be thinned based primarily on their form (vigour was a secondary consideration) and it was decided that for the first twenty percent of trees chosen they should not leave a gap as large as 4 metres apart. This involved going out to shelterbelt sites with relevant stockings (those closely spaced) and measuring the diameters and heights of 150 trees. 150 trees in total were chosen from three sites near Rotorua where PSP's were already located. For these sites, successively larger thinnings were applied and the appropriate trees chosen. The basal area remaining was then computed for each percentage of stems remaining and models fitted.

8.2 Exploratory Data Analysis:

The shelterbelt data was taken from 150 trees from three main sites and the diameters and heights recorded are plotted below:

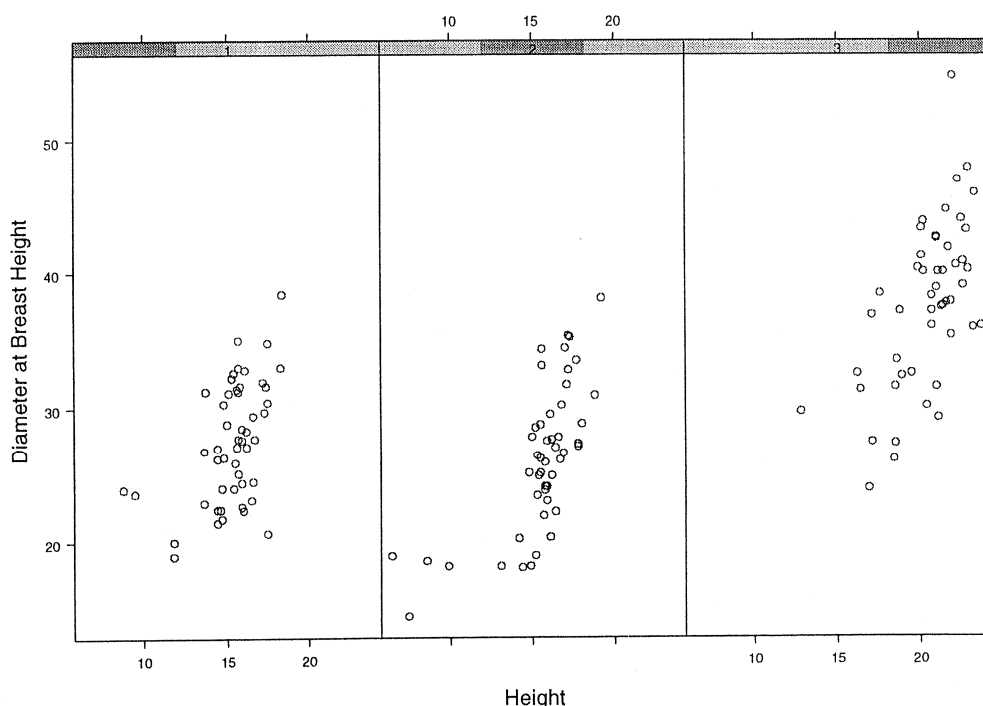


Figure 8-1. Scatterplot of Diameter at Breast Height versus Height for the shelterbelt thinning data.

Figure 8-1 shows that the third site exhibited larger taller trees while the first two sites appeared quite similar. For all sites larger diameters were associated with larger heights.

The relationship between diameter and height (across all sites) and the percentage of stems chosen for thinning was also examined and is illustrated in figure 8-2 below.

Note: An indicator variable was used to describe which trees were chosen for thinning. For example, if the indicator variable for the tree was zero then it was not chosen for thinning, if the variable was equal to one, then it was chosen for the first 10 percent to be thinned, if the variable was equal to two then it was included in the first 20 percent to be thinned and so on. This was performed until about 60% of the stems were chosen for thinning.

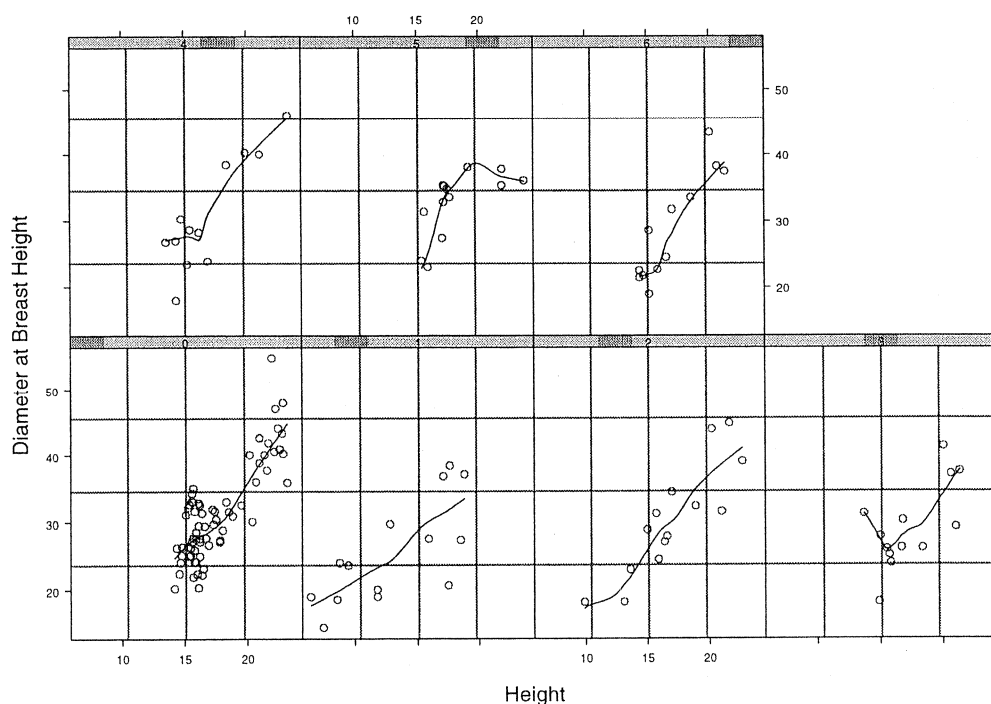


Figure 8-2. Scatterplots of Diameter at breast Height versus Height measurements with local regression models for the trees chosen for each percentage of stems thinned.

Figure 8-2 shows great variation in the diameter to height relationships across all trees. There seems to only be small differences between the trees immediately chosen for thinning and those only considered for thinning when over half the stand had already been chosen. This may be due to the criteria used for thinning (given they were chosen on form) however for completeness, any differences in their diameter to height relationships between groups were investigated.

8.3 Model Specification:

The relationship between the percentage of basal area remaining given the percentage of stems remaining after thinning was explored using a scatterplot (figure 8-3),

Generalized Additive Models and then a Nonlinear function. The results are illustrated below:

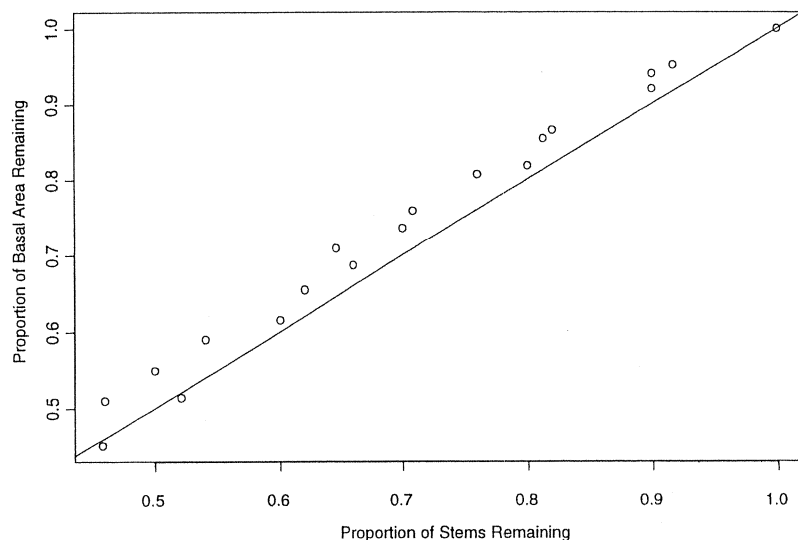


Figure 8-3. Scatterplot of Percent of basal area remaining versus the percent of stems remaining.

Figure 8-3 shows that the basal area remaining was consistently higher than the percent of stems remaining which was to be expected given the nature of thinning. The poor formed trees which were chosen for thinning appeared to have also been at least slightly smaller in diameter.

The form of this relationship plotted in figure 8-3 was further investigated using GAM's:

GAM Model:

Model 8-1 : Percent of Basal Area Remaining = $s(\text{Percent of Stems Remaining})$

Coefficients:

intercept = 0.0494309, $s(\text{Percent of Stems Remaining}) = 0.9716628$

Regression Plot for Model 8-1:

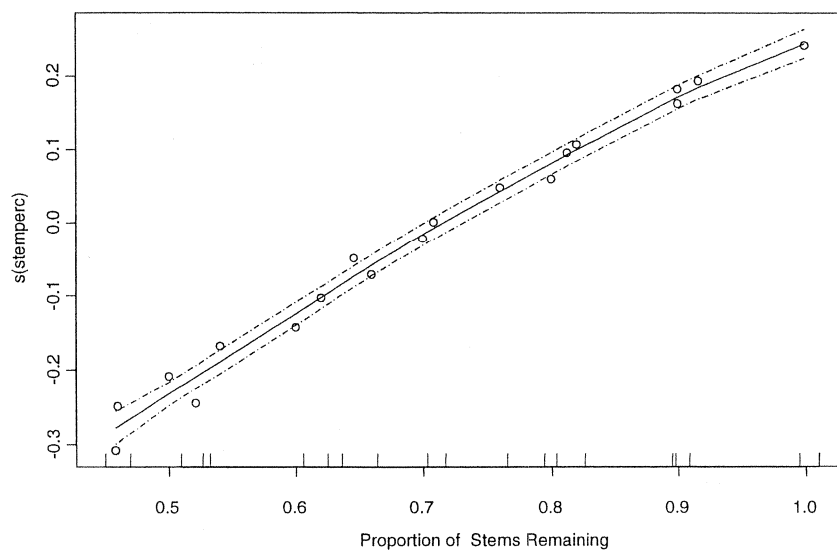


Figure 8-4. Regression plot for GAM 8-1.

Fit and Residual Plots:

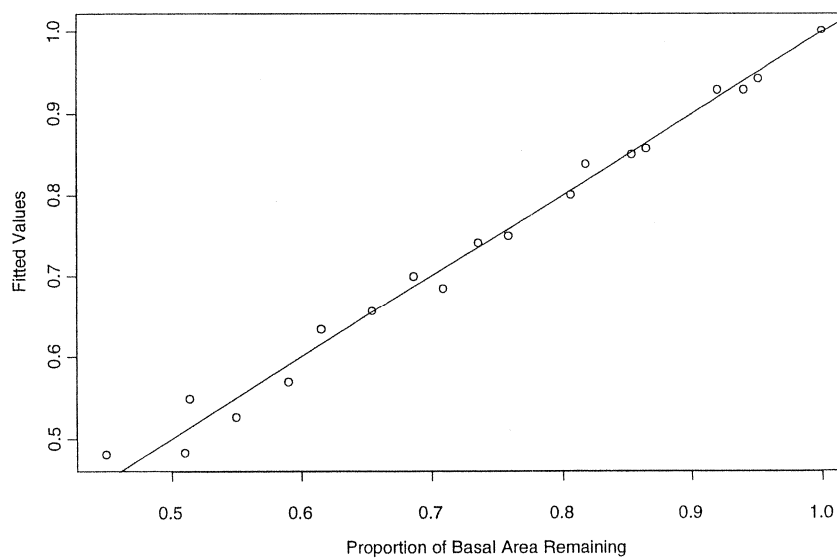


Figure 8-5. Observed versus predicted values for Model 8-1.

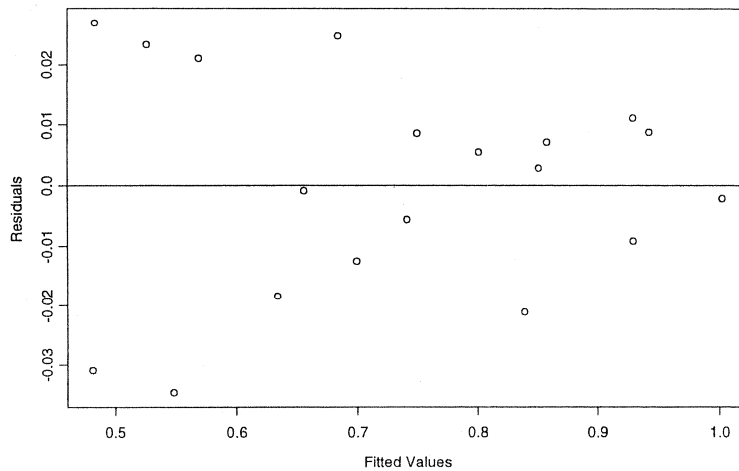


Figure 8-6. Residual plot for Model 8-1.

Figure 8-4 shows a slightly curved relationship which must pass through the origin and 100 percent. This suggested a simple nonlinear function with a single parameter.

8.4 Nonlinear Thinning Function:

Model 8-2: Percent of Basal Area Remaining = Percent of Stems remaining **c

Coefficients:

$$c = 0.8872487593$$

Fit and Residual plots:

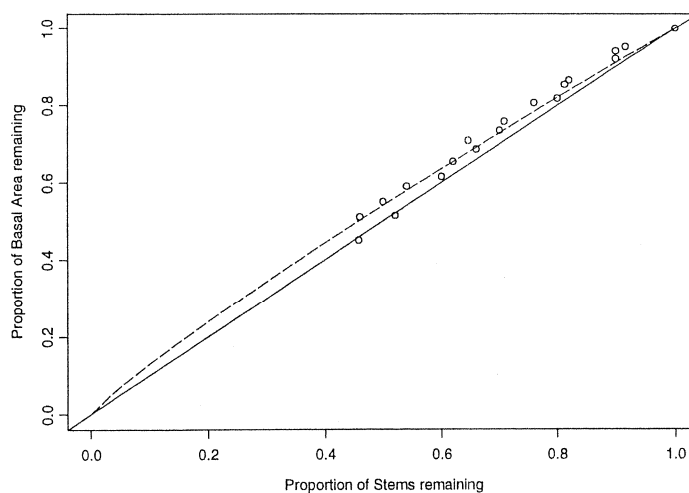


Figure 8-7. Scatterplot of the proportion of stems remaining and the proportion of basal area remaining with the Model 8-2 overlayed.

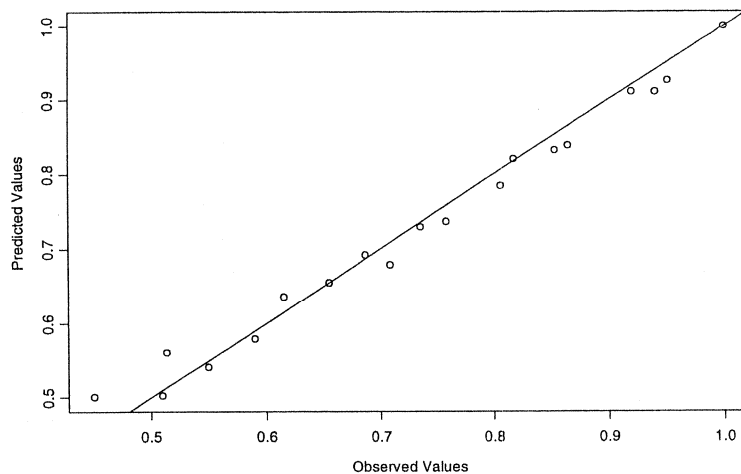


Figure 8-8. Observed versus predicted values for Model 8-2.

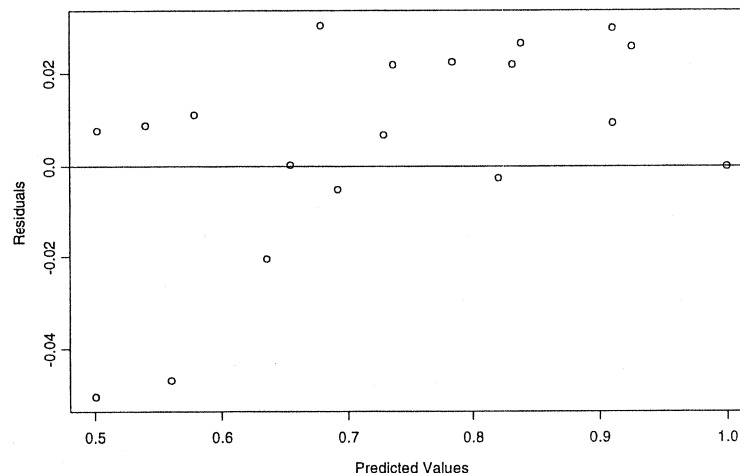


Figure 8-9. Residual plot for Model 8-2.

Fit Statistics:

$$\text{RMS} = 0.0003701403$$

8.5 Conclusions:

In order to predict the proportion of basal area remaining after a thinning (given the proportion of stocking remaining) this relationship was explored. This exploration also involved the use of a Generalized Additive model (GAM) which was followed by the fitting of a nonlinear model. The nonlinear model (Model 8-2) appeared to fit rather closely to both the GAM (Model 8-1) and the observed data. An alternative linear model was fitted to serve as a comparison to the nonlinear model 8-2 and the fit obtained was slightly better. However despite the better fit, the linear model was not considered suitable given that it would not pass through the points 0,0 and 100,100 therefore the nonlinear model was chosen.

Part Nine: Validation of the Complete Growth Model:

9.1 Introduction:

The performance of the various functions that comprise the complete growth model were examined using a similar process to that embedded within STANDPAK. This was undertaken to ensure that the Mean top Height, Basal area (with regional and plot asymptotes) and Mortality functions all worked together to provide reasonable predictions when supplied with single starting values for each plot.

9.2 Methods:

The validation process involved running all of the functions simultaneously after supplying the relevant initial values (Mean Top Height, Basal area, and Stocking). These were the only observed values entered in and the plots were grown over time (and compared at each re-measurement value) until the most recent measurement for each plot was reached.

9.3 Results for the Mean Top Height Curve:

9.3.1 Fit and Residual Plots:

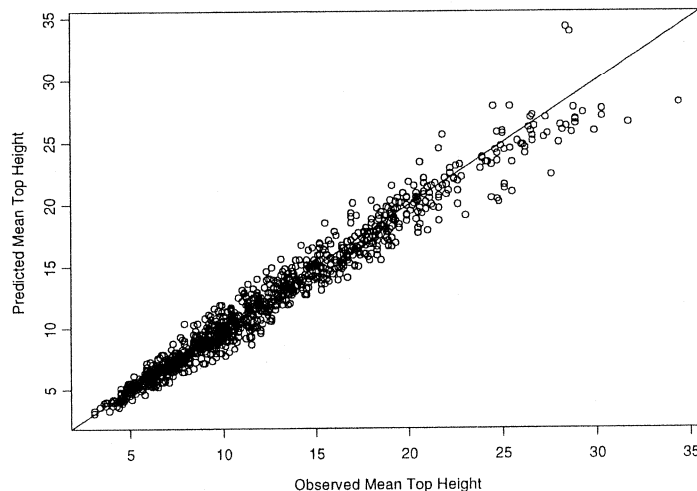


Figure 9-1. Plot of observed and predicted Mean Top Height versus age for all of the shelterbelt data.

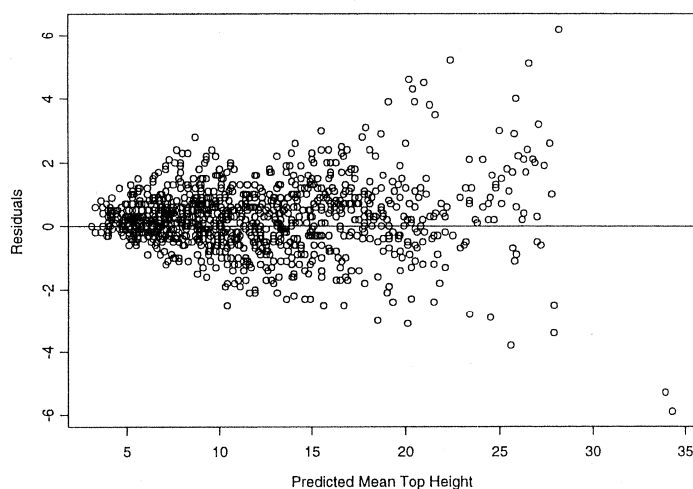


Figure 9-2. Residual plot for the Mean Top Height model for all of the shelterbelt data.

9.3.2 Residual Summary Statistics:

	Mean	SD	Min	Max	N
Absolute Errors (m)					
North Island	0.259	1.084	-5.900	6.200	974
South Island	0.170	1.067	-3.800	3.900	96
Combined	0.252	1.082	-5.900	6.200	1140

Table 9-1. Residual Summary statistics for Mean Top Height Function.

Figures 9-1 and 9-2 show the Mean Top Height function to fit satisfactorily with slight underprediction at larger Mean Top Height values. The mean residual for the North Island was considerably higher than that for the South Island, however the standard deviations around the mean were very similar for both islands.

The performance of the Mean Top Height function was also examined at the regional level to identify any particularly ill-fitting regions (Figures 9-3 and 9-4). It was found that Mean Top Height was underpredicted the worst on average for the Nelson/Marlborough region with an error rate of 11.5% while Taranaki was overpredicted on average with an error rate of 7.5%. (Note: Error rate = Mean Absolute Residual / Mean predicted value) The best fitting regions were Canterbury and the Bay of Plenty with error rates of 0.15% and 1.8% respectively.

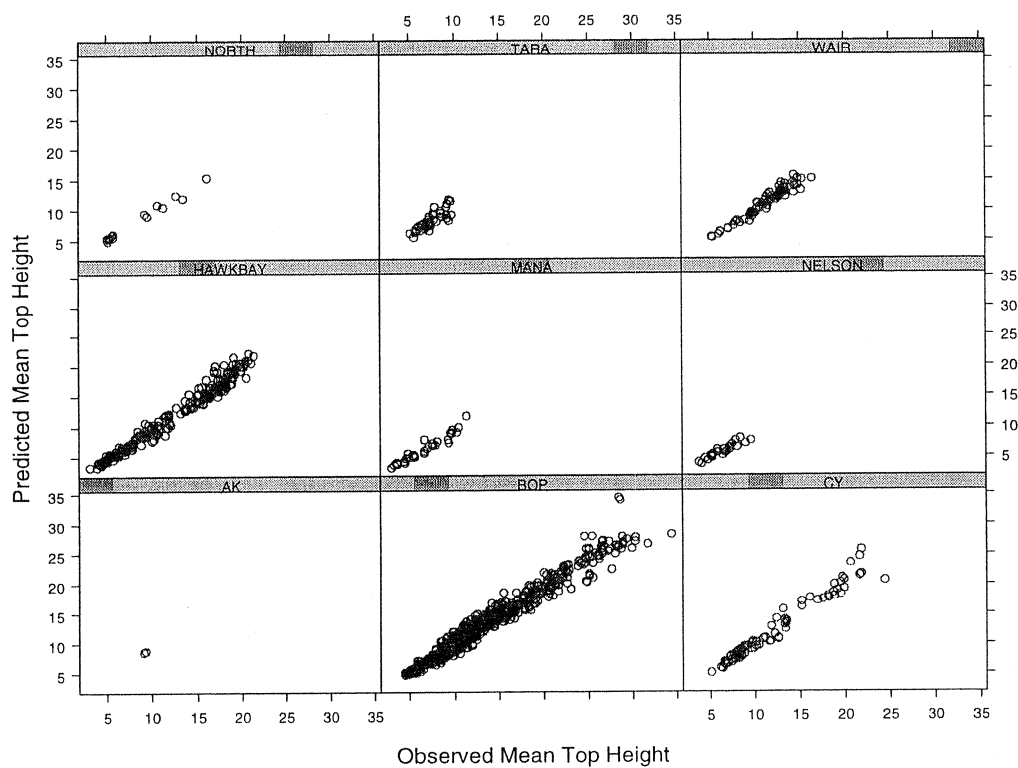


Figure 9-3. Plot of Observed and Predicted Mean top height versus age for all of the Shelterbelt data within region.

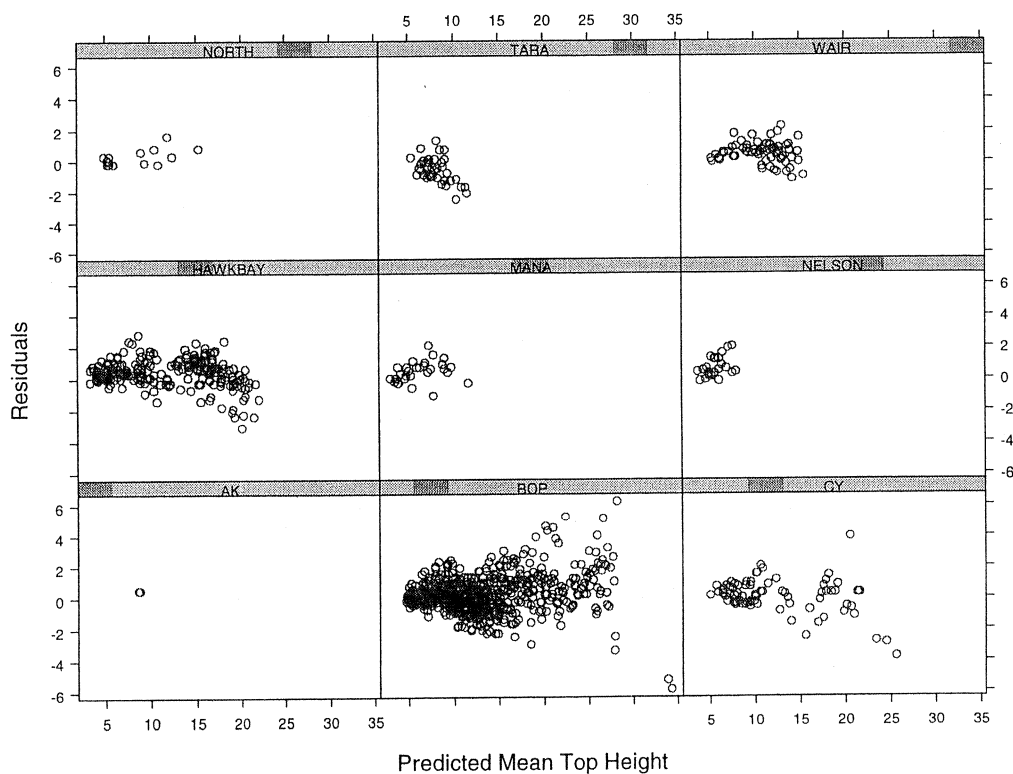


Figure 9-4. Residual plots for all shelterbelt data for each region.

The performance of the Mean Top Height function was also examined by plotting the residuals against factors such as stocking, age of the stand, and time interval of prediction (figures 9-5, 9-6 and 9-7).

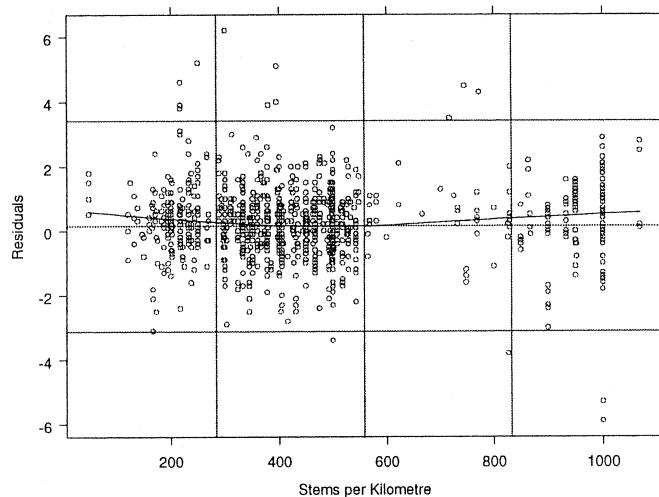


Figure 9-5. Residual plot with a local regression model for all shelterbelt data by Stems per Kilometre.

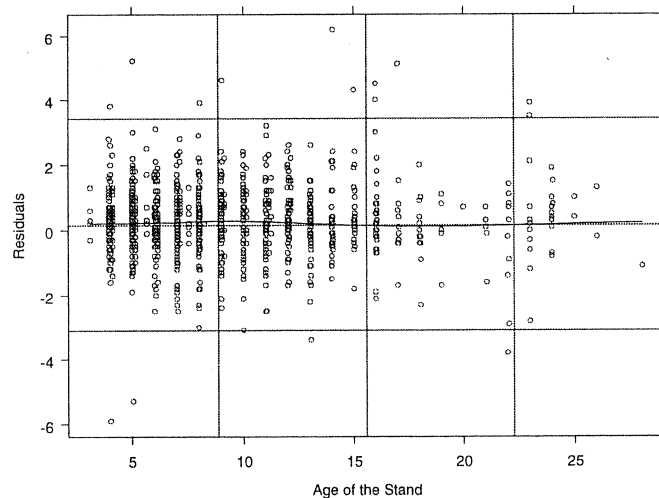


Figure 9-6. Residual plot with a local regression model for all shelterbelt data by Age of the stand.

9-9 also indicated this. Despite the apparent normality, the distribution was not strictly normal (which was determined by the Kolmogorov-Smirnov test) however the residuals were considered to be approximately normal.

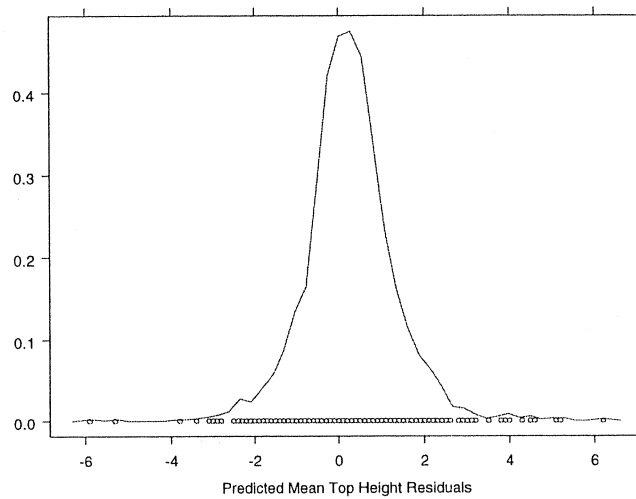


Figure 9-8. Density plot of the Residuals

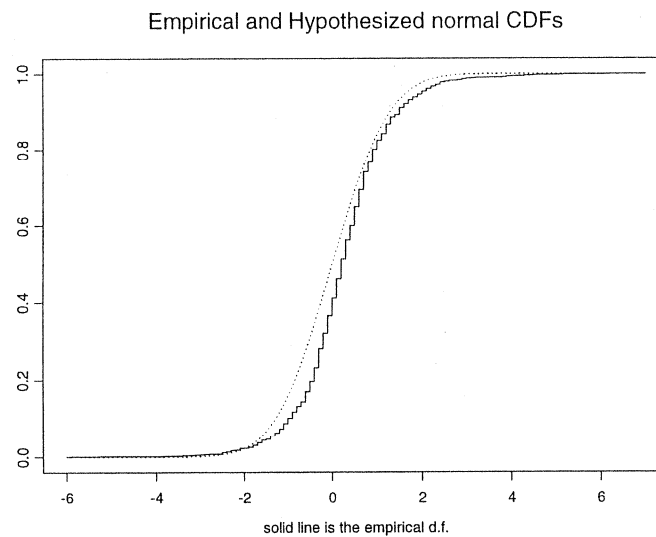
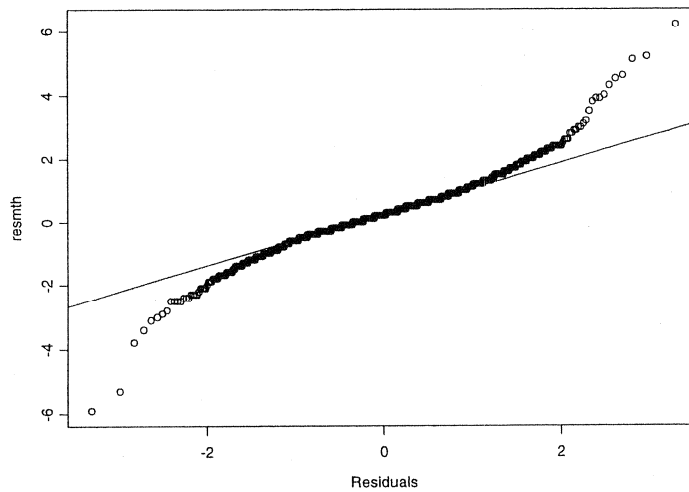


Figure 9-9. Cumulative distribution plot for the normal distribution (solid line) and the distribution of the residuals (dotted line).



9-10. QQ-plot for Mean Top Height residuals

9.4 Validation for the Basal Area Increment Functions:

Fit and residual plots were produced for two Basal Area increment functions (Model 5b and Model 3) using either regional or estimated plot asymptote values and a satisfactory fit was found for both functions (see figures 9-11, 9-12, 9-13, and 9-14).

9.4.1 Fit and Residual plots for the Basal Area Increment Model 5b and Model 3 using regional asymptote estimates:

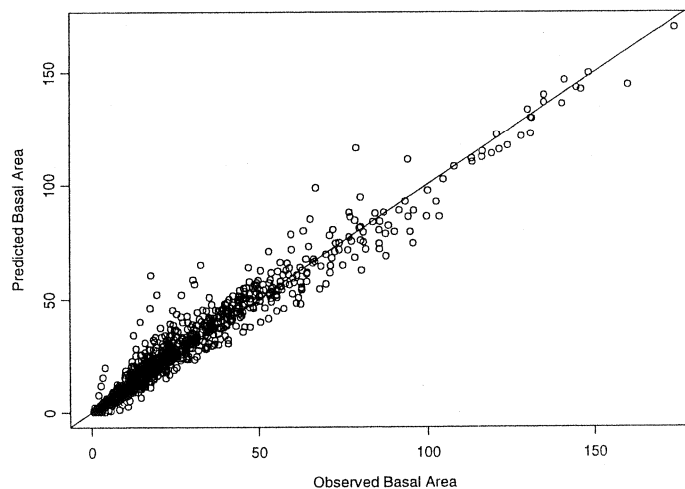


Figure 9-11. Observed versus Predicted shelterbelt Basal Area Values (m^2 / km) for Regional Mixed model 5b.

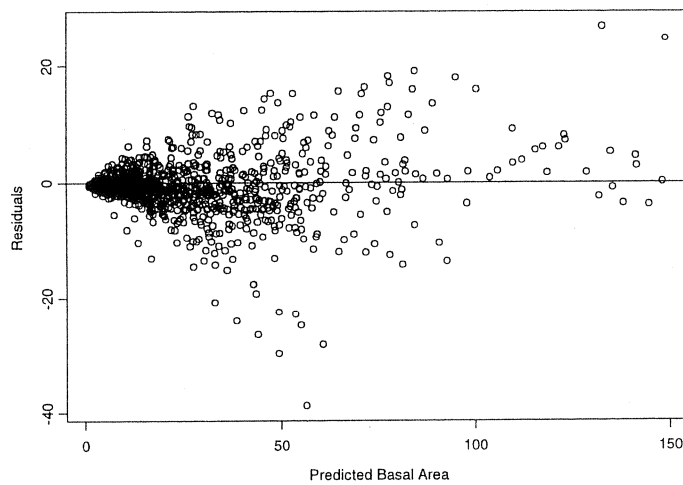


Figure 9-14. Residual plot for the basal area increment Regional Mixed model 3.

**9.4.2 Fit and Residual plots for Basal Area increment Model 5b and Model 3 using
Plot asymptote estimates:**

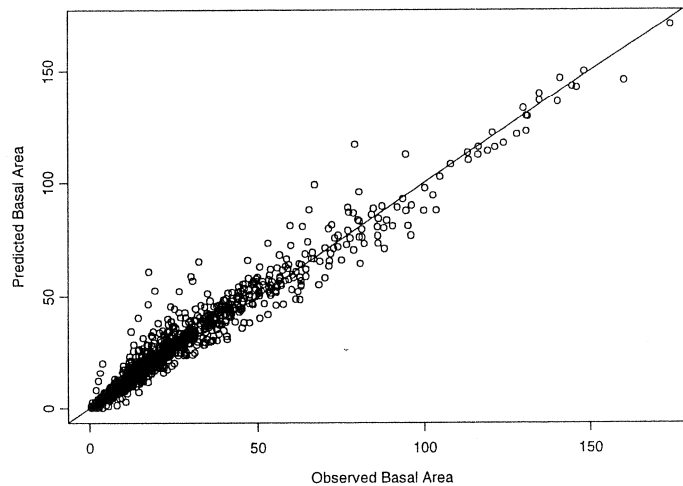


Figure 9-15. Observed versus Predicted shelterbelt Basal Area Values (m^2 / km) for Plot estimate Mixed model 5b.

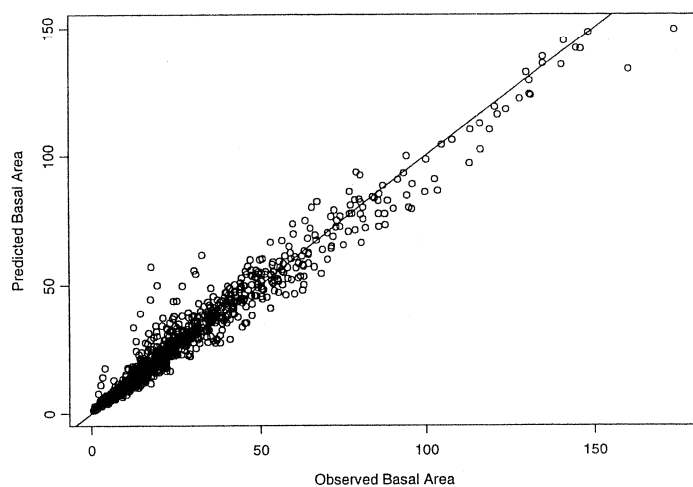


Figure 9-16. Observed versus Predicted shelterbelt Basal Area Values (m^2 / km) for Plot estimate Mixed model 3.

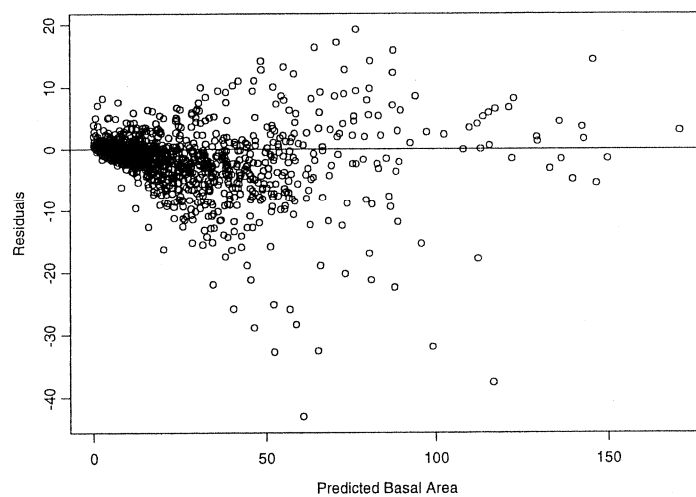


Figure 9-17. Residual plot for basal area increment plot estimate Mixed model 5b.

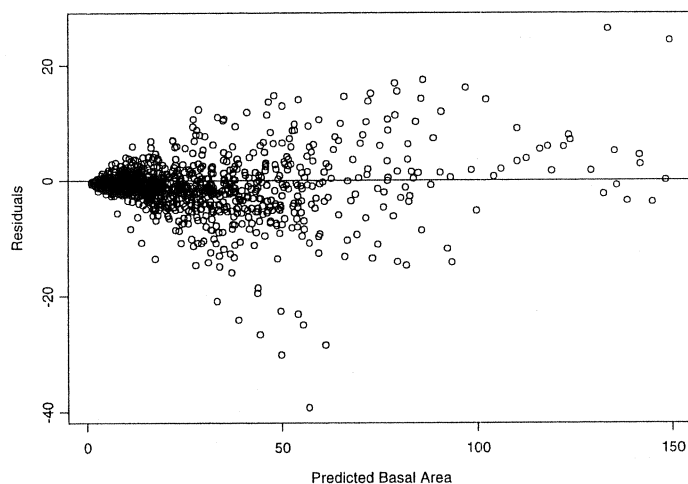


Figure 9-18. Residual plot for Plot estimate Mixed model 3.

Performance of the basal area functions was also checked at the regional level, however due to the similar performance of the basal area functions using either regional or plot asymptote estimates only the regional results for either model 5b or model 3 are shown below:

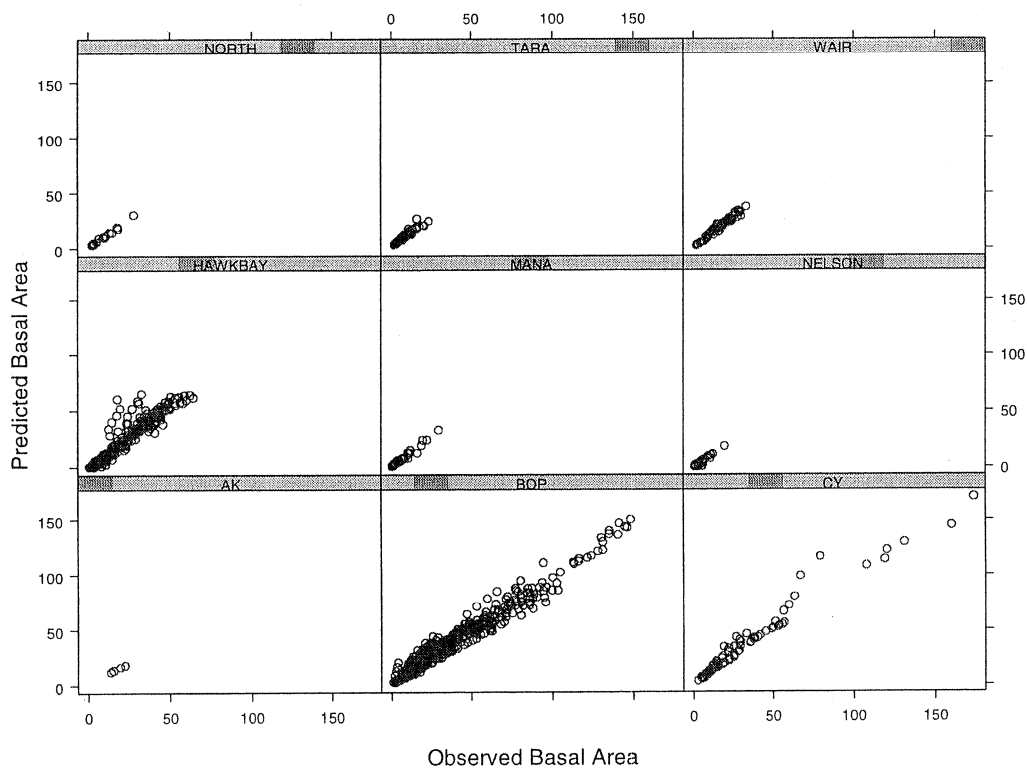


Figure 9-19. Plot of Fitted versus Predicted Values from Mixed model 5b within region.

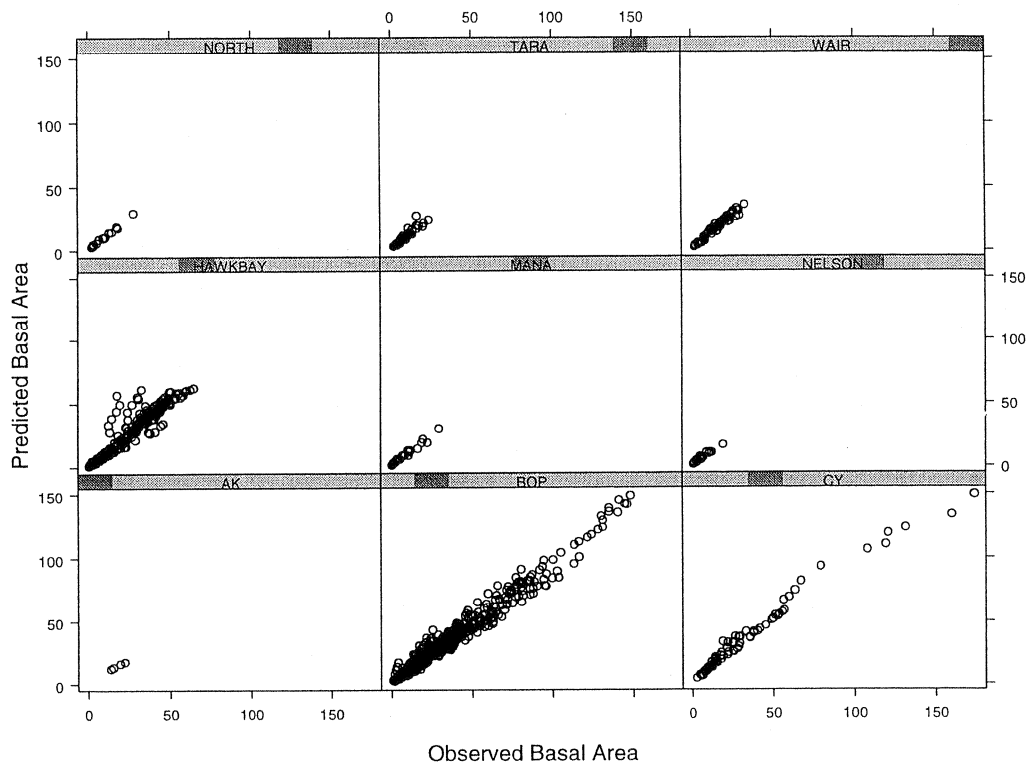


Figure 9-20. Plot of Fitted versus Predicted Values from Mixed model 3 within region.

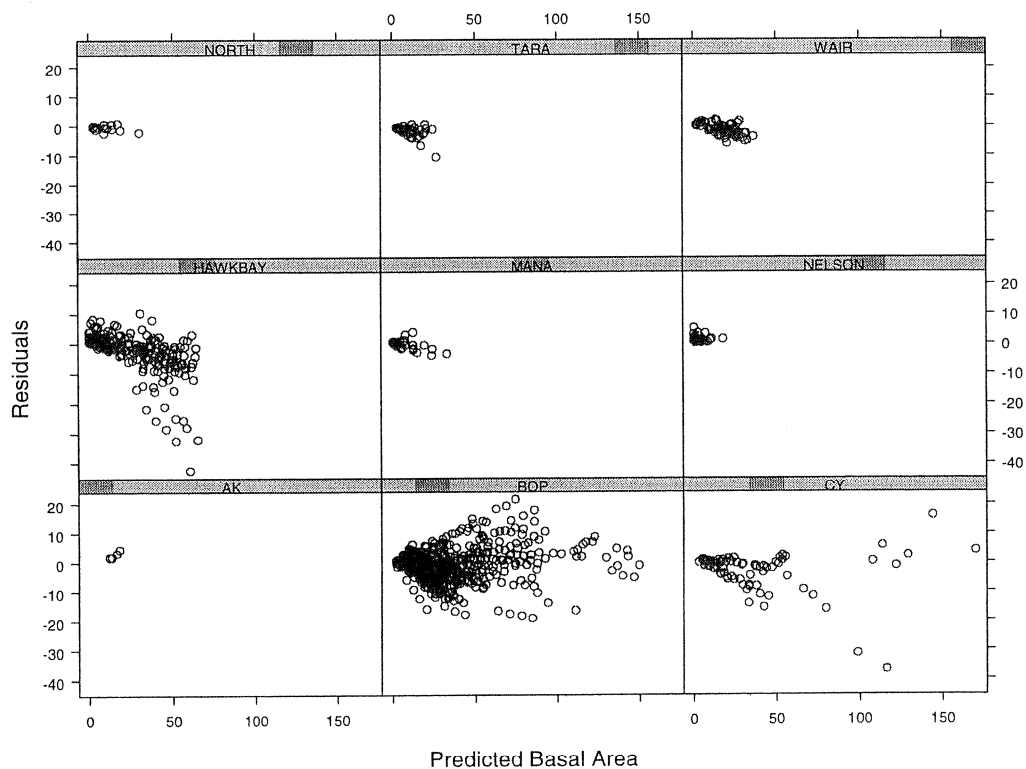


Figure 9-21. Residual plot for Mixed model 5b for all shelterbelt data within region.

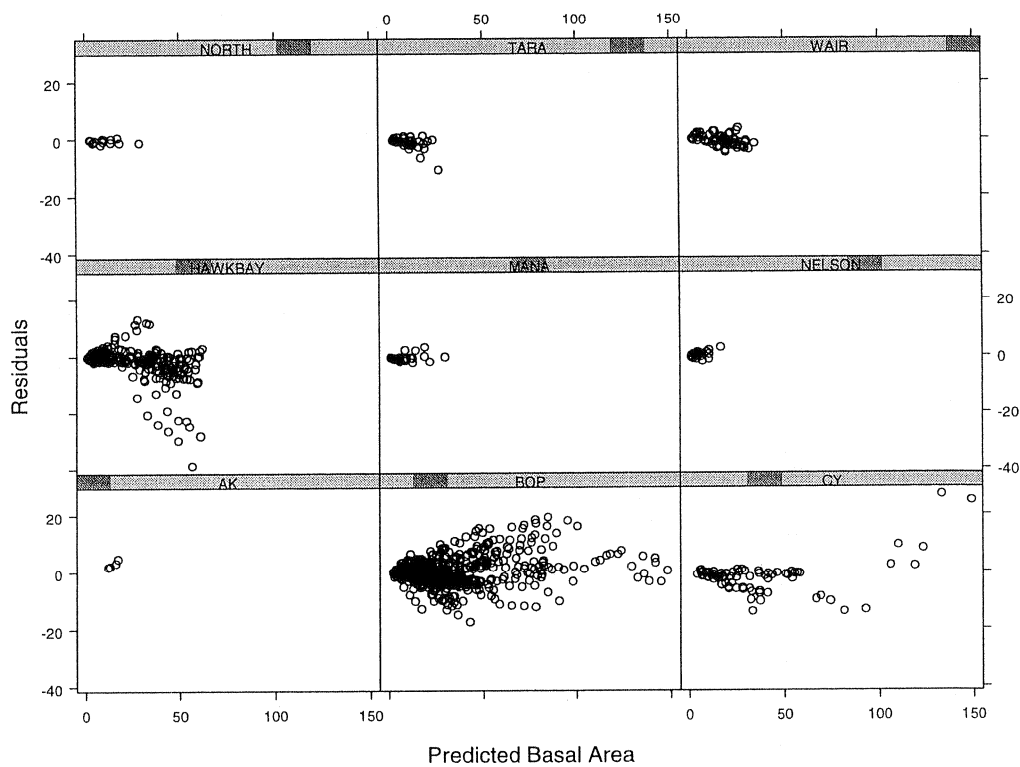


Figure 9-22. Residual plot for Mixed model 3 for all shelterbelt data within region.

9.4.3 Residual Summary Statistics

	Mean	SD	Min	Max	N
Residuals					
All data (5)	-1.449	5.490	-42.680	21.460	1140
	<i>-1.770</i>	<i>5.495</i>	<i>-43.020</i>	<i>19.200</i>	
(3)	-0.678	4.990	-38.800	26.960	
	<i>1.074</i>	<i>4.957</i>	<i>-39.3100</i>	<i>26.510</i>	
Northland (5)	-0.686	0.947	-2.440	0.730	15
	<i>-0.791</i>	<i>0.965</i>	<i>-2.650</i>	<i>0.600</i>	
(3)	-0.629	0.688	-1.890	0.510	
	<i>-0.790</i>	<i>0.718</i>	<i>-2.020</i>	<i>0.260</i>	
Auckland (5)	2.768	1.205	1.760	4.320	4
	<i>2.653</i>	<i>1.161</i>	<i>1.690</i>	<i>4.150</i>	
(3)	3.003	1.356	1.800	4.800	
	<i>2.878</i>	<i>1.320</i>	<i>1.710</i>	<i>4.640</i>	
Bay of Plenty (5)	-0.920	5.164	-19.320	21.460	654
	<i>-1.318</i>	<i>5.159</i>	<i>-22.460</i>	<i>19.200</i>	
(3)	0.083	4.740	-17.560	19.080	
	<i>-0.407</i>	<i>4.696</i>	<i>-18.590</i>	<i>17.410</i>	
Hawkes Bay (5)	-2.637	7.079	-42.680	10.250	232
	<i>-2.862</i>	<i>7.100</i>	<i>-43.020</i>	<i>9.980</i>	
(3)	-2.160	6.308	-38.800	13.050	
	<i>-2.469</i>	<i>6.314</i>	<i>-39.310</i>	<i>12.330</i>	
Manawatu (5)	-0.557	1.641	-4.160	3.710	30
	<i>-0.898</i>	<i>1.537</i>	<i>-4.370</i>	<i>3.500</i>	
(3)	-0.750	1.199	-2.390	3.150	
	<i>-1.084</i>	<i>1.182</i>	<i>-2.900</i>	<i>2.690</i>	
Taranaki (5)	-2.100	1.934	-10.780	0.320	40
	<i>-2.279</i>	<i>1.944</i>	<i>-10.930</i>	<i>0.240</i>	
(3)	-1.570	2.033	-10.970	0.880	
	<i>-1.749</i>	<i>2.071</i>	<i>-11.380</i>	<i>0.550</i>	
Wairarapa (5)	-1.727	1.876	-6.090	1.870	65
	<i>-1.859</i>	<i>1.865</i>	<i>-6.210</i>	<i>1.720</i>	
(3)	-1.200	1.751	-4.960	3.340	
	<i>-1.420</i>	<i>1.702</i>	<i>-5.240</i>	<i>2.650</i>	
Nelson (5)	1.335	1.237	0.000	5.130	24
	<i>1.259</i>	<i>1.224</i>	<i>-0.110</i>	<i>5.000</i>	
(3)	0.043	1.070	-1.880	2.830	
	<i>-0.163</i>	<i>1.030</i>	<i>-2.190</i>	<i>2.350</i>	
Canterbury (5)	-3.395	7.059	-36.940	15.790	76
	<i>-3.711</i>	<i>7.151</i>	<i>-37.590</i>	<i>14.370</i>	
(3)	-2.204	6.203	-14.180	26.960	
	<i>-2.459</i>	<i>6.284</i>	<i>-14.930</i>	<i>26.510</i>	

Table 9-2. Residual Summary Statistics for the two basal area function results. The bold numbers represent the residuals obtained using regional asymptote estimates while the italicised numbers represent those obtained using plot asymptote estimates. Note: (3) = Mixed model 3, (5)=Mixed model 5b.

Either basal area increment model appeared to consistently overpredict on average for seven of the nine regions for model 5b, and six of the regions for model 3. This overprediction was largest for the Taranaki region for both mixed models given the error rates of 18.4% and 14.42% for mixed models 5b and 3 respectively (based on 3.5% of the data). In addition, the results show that either regional basal area increment function underpredicted most on average (for regions with more than four data points) for the Nelson/Marlborough region with error rates of 30% or 0.8% for models 5b and 3 respectively. This level of underprediction may not be surprising given that only 2.1% of the data came from this region. The Bay of Plenty region fitted the best on average, given that the error rates for model 5b and 3 were only 3.04% or 0.28% respectively. These rates are reassuring given that 57.4% of the data came from this region.

The residuals for both basal area functions were also checked with various age, stocking and length of prediction periods for any obvious relationships. See figures 9-23, 9-24 and 9-25.

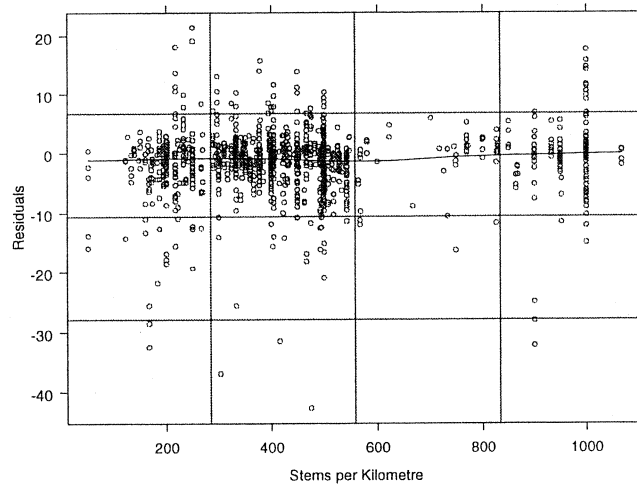


Figure 9-23. Residual plot for model 5b with a local regression for all shelterbelt data by Stems per Kilometre.

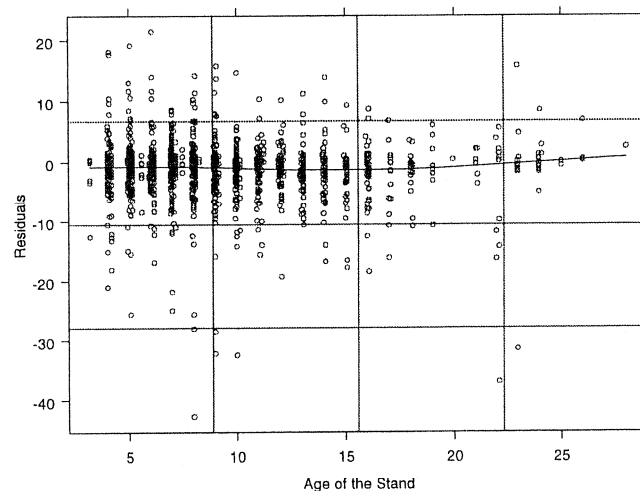


Figure 9-24. Residual plot for model 5b with a local regression for all shelterbelt data by Age.

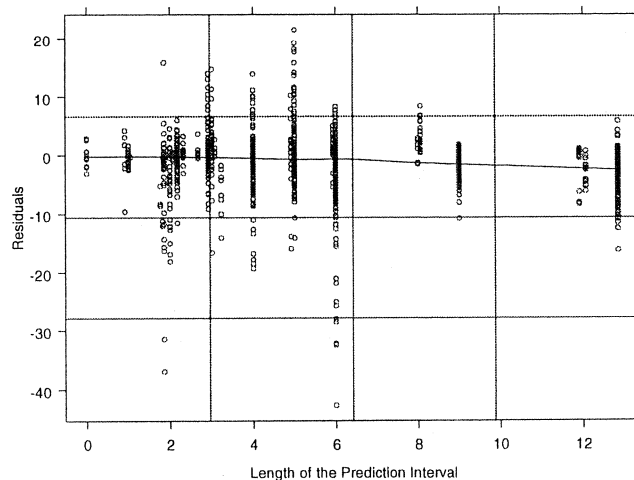


Figure 9-25. Residual plot for model 5b with a local regression model for all shelterbelt data by length of prediction interval (years).

9.4.4 Discussion:

Errors in relation to age:

Both regional basal area models underpredicted on average for seven and overpredicted on average for six age classes (Figures 9-24 and Appendices 4, 11). The largest error rate for model 5b (for age classes with more than one observation) was for the 15-17 age class which reached 10.27%. While the largest error rate for model 3 was for the 25-27 age class which had an error rate of 8.80%. The best fitting age class for both models was the 5-7 age class with error rates of only 0.147% and 0.3% for models 5b and 3 respectively.

Both basal area functions using plot asymptote estimates also underpredicted on average for seven of the thirteen age classes (Appendix 4 and 11). The largest error rate (for age classes with more than one observation) was for the 15-17 age class which was 10.16% for model 5b, while the corresponding value for model 3 was 9.66% for the under five age class. The best predicted class on average resulted in an error rate of 0.25% (for the 21-23 age class) for model 5b while model 3 produced the best error rate of 0.3% for the 23-25 age class.

Errors in relation to stocking:

The regional model 5b overpredicted on average for six and underpredicted on average for five stocking classes (Figure 9-23 and Appendix 4). While model 3 overpredicted for nine of the eleven classes (Appendix 12). The largest overpredicted class (based on more than 5 data points) for either model was the 100-200 stems per kilometre group where the error rates were 24.9% and 15.0% for models 5b and 3 respectively. The best fitted class for both models was the 400-500 stems group where the residual mean was only 1.89% and 0.79% the size of the predicted mean basal area for models 5b and 3 respectively.

The stocking class with the largest error rates (24.9% , 15%) for prediction of basal area was also the stocking class (with more than 5 data points) with the largest error rate for predicting mortality. This suggests that the reason that basal area was being overpredicted was because the number of stems was being overpredicted rather than gross overprediction of basal area increment on each plot.

The basal area model 5b using plot asymptote estimates overpredicted on average for 10 of the eleven stocking classes while model 3 overpredicted for nine of the eleven stocking classes. The largest overpredicted group for either model (based on more than 5 data points) was also the 100-200 stems per kilometre class which saw error rates of 22.95%, and 17.23% for models 5b and 3 respectively. The closest fitting stocking class for model 5b was the 800-900 class with an error rate of only 0.68% while the corresponding value for model 3 was for the 700-800 class with an error rate of 1.86%.

The reason for the ill-fitting 100-200 stems groups is also considered to be highly related to the mortality model.

Errors in relation to length of prediction period:

The regional basal area model 5b underpredicted on average for two of the ten prediction interval classes, however the size of the residuals did not seem to increase considerably as the prediction period increased (Figure 9-25 and Appendix 6). This was also seen for model 3 , however this model underpredicted for three of the ten classes (Appendix 13). The largest error rate (for those groups with greater than six observations) for model 5b was found for the stands that were only predicted one year ahead (14.3%) however this group only contained 3.3% of the data. The largest error rate for model 3 corresponded to the 3-4 year prediction interval (for those classes with greater than six observations) which was 12.26%. The best error rate on average for model 5b was found for the 12-13 prediction period group which contained about 15% of the data and had an error rate of only 1.3%. The corresponding error rate for model 3 was 1.91% which was associated with the 1-2 year prediction interval class.

The plot asymptote basal area model 5b was underpredicted on average for only one of the ten prediction interval classes, while model 3 underpredicted for only three of the ten prediction classes. The size of the residuals did not increase considerably as the prediction period increased for either model (Appendix 6 and 13). The largest error rate for model 5b (for those groups with greater than six observations) was also found for the stands that were only predicted one year apart (15.1%). While model 3 produced the largest error rate (for those groups with greater than six observations) for the 5-6 year prediction class which was 11.5%. The best error rate for model 5b was found for the 12-13 prediction period group which contained about 15% of the data and had an error rate of only 0.4%. The best error rate for model 3 corresponded to the 8-9 interval class with an error rate of 0.14%.

Errors in relation to regional and plot level models:

The basal area functions using either regional or plot asymptote estimates produced very similar results (Appendices 4-6, 11-13). The function using plot asymptote estimates improved on the mean residual in six of the thirteen age classes for both

models and improved on the standard deviations about the mean residual in eight of the thirteen age classes for models 5b and 3. Using plot asymptotes resulted in a reduction in the mean residuals in only two of the eleven stocking classes for both models while the standard deviations were reduced for only four of the eleven stocking classes for model 5b and only three for model 3. The similarity in the residuals for either asymptote model can also be seen in the density plot illustrated below (figure 9-26) while statistical similarity was confirmed using the Kolmogorov-Smirnov test.

Comparison of errors obtained in relation to the mixed model fitted:

The residuals from Mixed model 3 have been shown to be consistently smaller and less variable on average than those obtained from Mixed model 5b. These results are contrary to what was expected given the results obtained when the models were fitted using the modelling data. One possible reason for these contrary results is the formulation of the equations with respect to the stems per kilometre variable. As we have seen in part six and will see in section 9.5, the mortality model is less than adequate given the exposed nature of shelterbelt sites. This has been shown to effect the two basal area models differently. Model 5b (which has shown to fit better than Model 3 in part two) fits a coefficient to the stocking variable which is multiplied by the average taper variable. However, model 3 incorporates the stocking parameter into the rate parameter for the green crown length convex curve. Therefore, the precision of the results from model 5b are more dependent on correct stocking estimates being entered into the model. For example, simply fitting the two models with and without the stocking terms using the modelling data (not allowing for compounding errors) resulted in a 27.7% decrease in the residual mean square (RMS) for Model 3 however, a 40.4% decrease in the RMS was seen when model 5b was fitted with stems per kilometre information. This indicates that the prediction of basal area increment for model 5b relies more heavily on the stems per kilometre variable than model 3.

9.4.5 Distribution of the Residuals:

The distribution of the residuals using either the regional or plot level asymptotes for both models is illustrated in Figures 9-26, 9-27. Skewness is apparent in both distributions (which was confirmed by the QQ-plots and cumulative density plots (figures 9-28 - 9-31)) however the residuals of model 3 was less skewed with skewness statistics of -0.59 and -0.79 for the regional and plot level models compared to statistics of -1.44 and -1.65 for model 5b for the regional and plot level models respectively. Not one of the residual distributions for either model were considered normal when the Kolmogorov-Smirnov test was used.

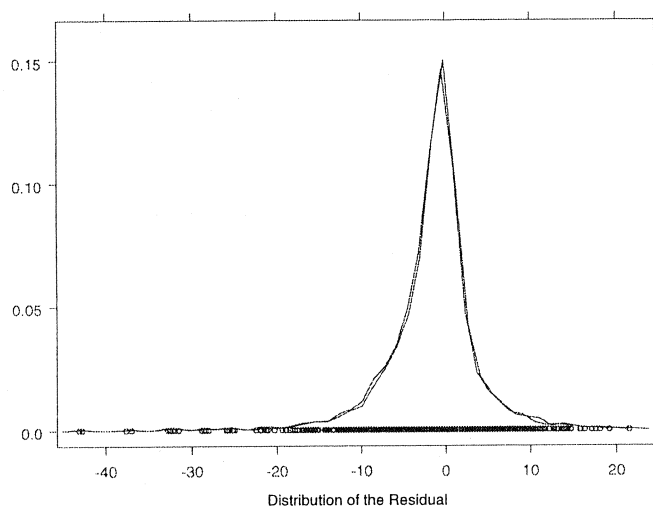


Figure 9-26. Density plot for the residuals for the basal area model 5b. The green line is obtained using the regional estimates while the pink line was obtained using plot asymptote estimates.

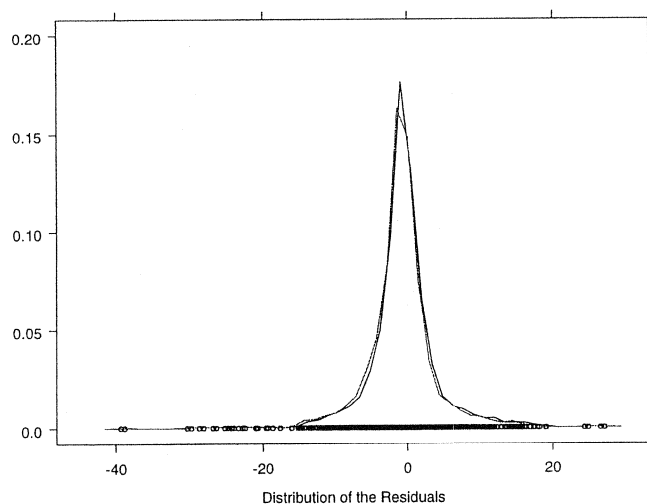


Figure 9-27. Density plot for the residuals for the basal area model 3. The black line is obtained using the regional estimates while the pink line was obtained using plot asymptote estimates.

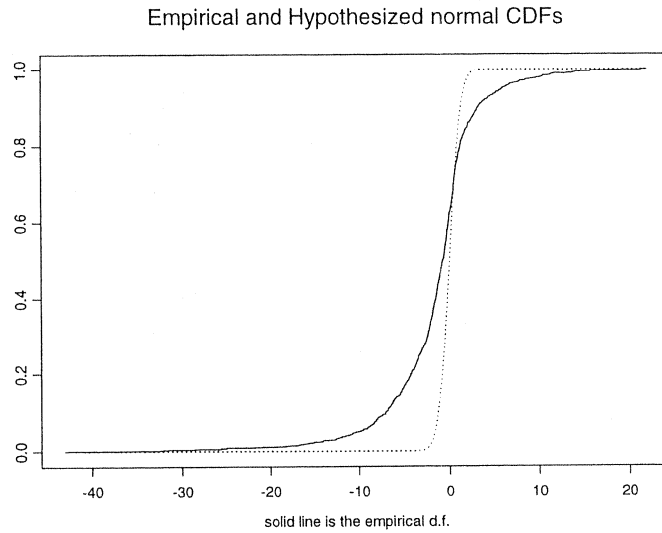


Figure 9-28. Cumulative density plot for residuals from the Regional asymptote basal area model 5b.

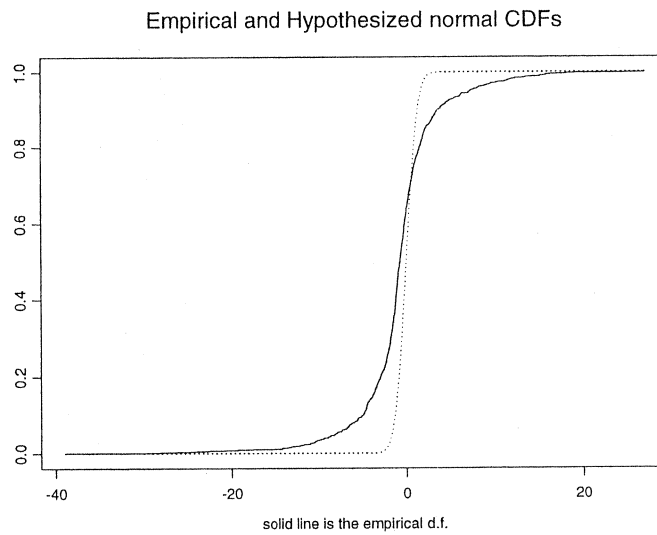


Figure 9-29. Cumulative density plot for residuals from the Regional asymptote basal area model 3.

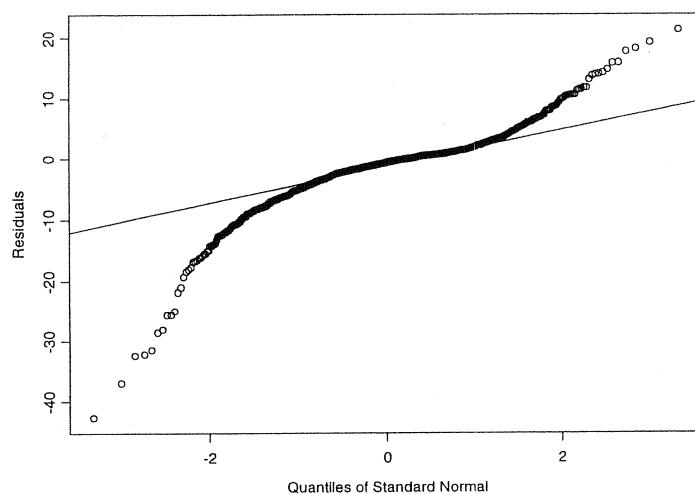


Figure 9-30. QQ-plot for residuals from the Regional asymptote basal area model 5b.

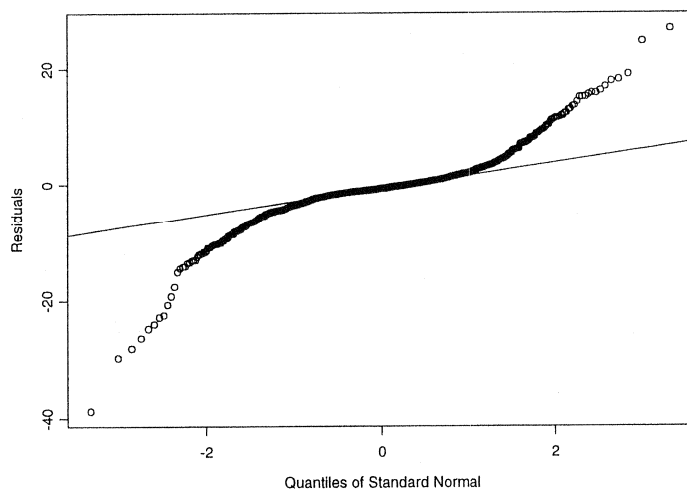


Figure 9-31. QQ-plot for residuals from the Regional asymptote basal area model 3.

9.5 Validation of the Mortality Function:

9.5.1 Fit and Residual Plots:

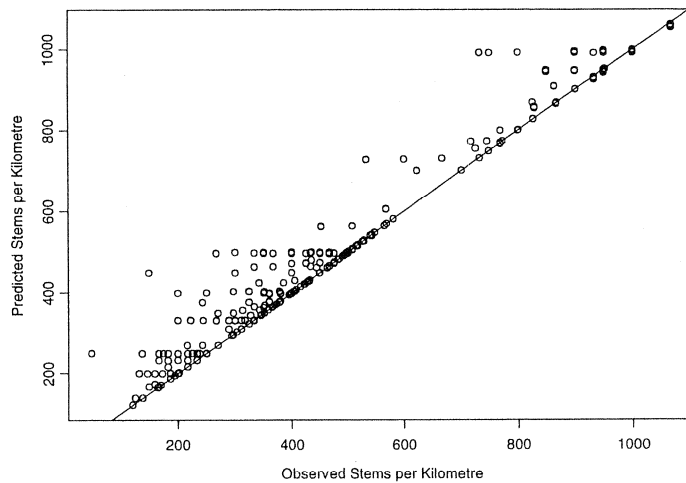


Figure 9-25. Observed versus fitted values for Stems per Kilometre.

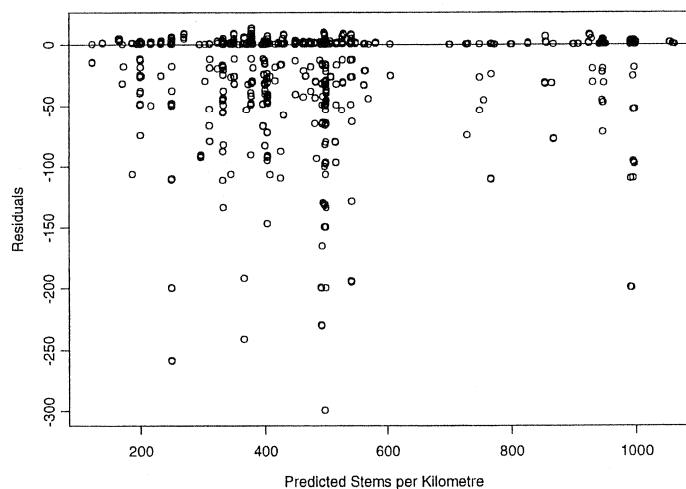


Figure 9-26. Residual plot for Stems per Kilometre.

9.5.2 Residual Summary Statistics:

	Mean	SD	Min	Max	N
All data	-19.532	42.314	-299.300	13.300	1140

Table 9-3. Residual summary statistics for the Mortality function.

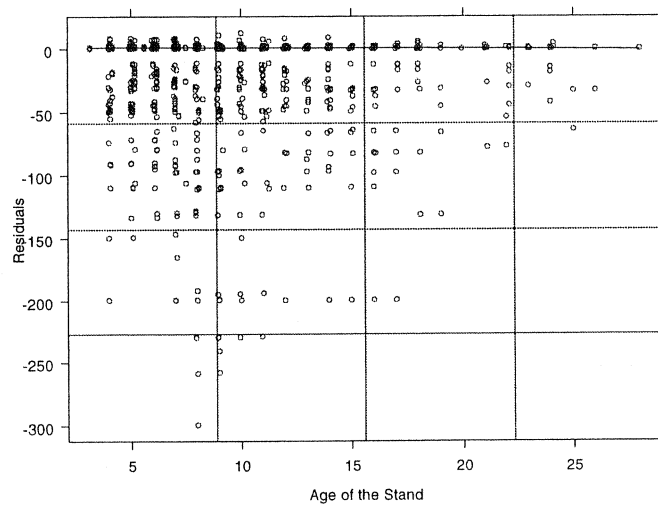


Figure 9-27. Residual plot with a local regression model versus Age of the stand

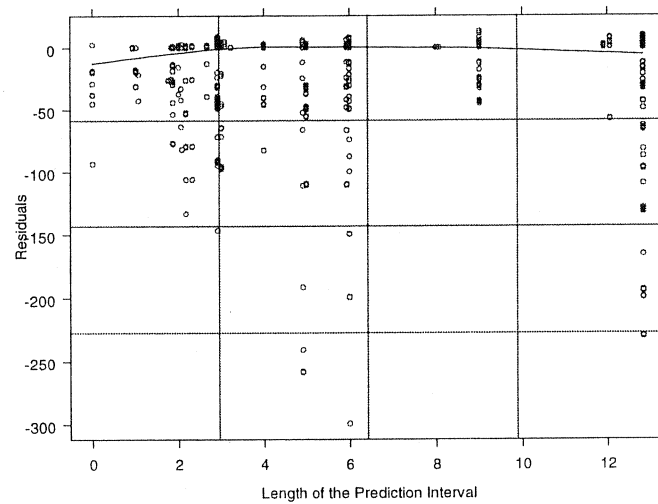


Figure 9-28. Residual plot with a local regression model versus length of prediction interval.

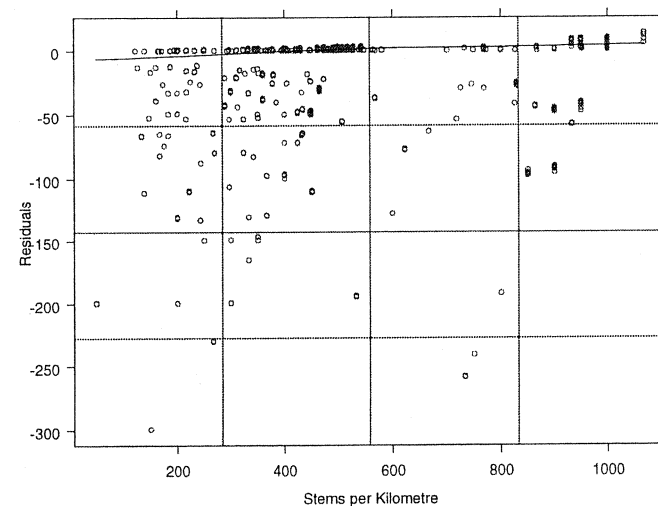


Figure 9-29. Residual plot with a local regression model versus stems per kilometre.

9.5.3 Discussion:

Figures 9-25 and 9-26 show general underprediction of mortality on average, while figures 9-27 and 9-28 show the residuals are not highly related to either age, length of the prediction interval or stocking. Despite the large residuals the error rates did not rise over 6.1% for any of the age classes (Appendix 8), which corresponded to the 7-9 age group. The best error rate (0.2%) was obtained for the 23-25 age group.

The residuals in relation to stocking class showed a large error rate (15.5%) for the 100-200 stocking group (Appendix 9). This was closely followed by an error rate of 12.3% rate that was associated with the 200-300 stocking group. The best predictions on average corresponded to the 900-100 stocking class with an error rate of 0.10%.

The error rates when concerned with the length of the prediction interval (Appendix 10) showed to be smallest for prediction terms of 11-12 years (0.32%) while the worst predictions on average were made for the 5-6 year prediction terms (6.8%). Most importantly the error rate was not found to increase with increasing prediction interval.

9.5.4 Distribution of the Residuals:

As expected the residuals for the mortality function were highly skewed (skewness statistic of -2.89) which further highlights the overprediction of surviving stems per kilometre.

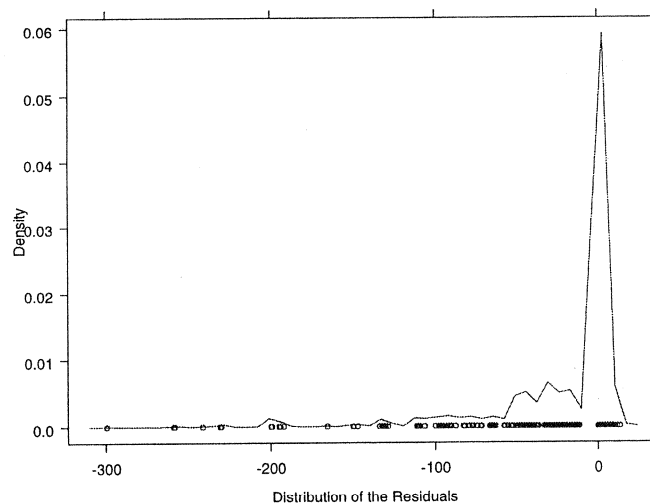


Figure 9-30. Distribution of the residuals for the mortality model.

9.5 Conclusions:

The performance of the complete growth model from single starting values invoked predominantly Mean Top Height, Basal Area increment, Mortality and Average taper functions. These functions require starting estimates of mean diameter, mean height,

age, and mortality. These functions were employed simultaneously and predicted values obtained for each re-measurement of the shelterbelts.

The Mean top Height function produced good results for both the North and South Islands, with a national residual mean average of only 0.25 metres. In addition, no significant biases were seen when the errors were related to stocking, age or the length of the prediction period. The residuals from both the North and South Island models were also examined and were found to be approximately normally distributed.

The selected Basal Area increment functions produced similar results which were not significantly affected by stocking, age or the length of prediction period. However, the validation of Model 3 obtained smaller and less variable residuals on average than Model 5b. This was thought to be predominantly due to the performance of the mortality function given the relative contribution of the stocking term to the two models and thus the relative consequences of incorrect stocking predictions. Therefore, given the degree of difficulty encountered when mortality modelling is undertaken in shelterbelts, the implementation of the mortality model within STANDPAK is often user defined. This feature allows the user to proportionally increase or decrease the amount of mortality predicted and is designed to re-fine the stocking predictions. Given the refined nature of the mortality predictions, Model 5b was chosen as the preferred basal area increment model.

Appendix 1:
Mean Top Height Function: Observed, Predicted and Residual Values for different
age classes:

where MTH= Observed Mean Top Height, P_MTH= Predicted Mean Top Height, RESMTH= Residual

----- AGE=<5 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	152	5.6585526	1.2616842	3.1000000	9.5000000
P_MTH	P_MTH	153	5.4339869	1.1088854	3.1000000	8.5000000
RESMTH		152	0.2184211	0.4392078	-0.6000000	1.7000000

----- AGE=5-7 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	265	8.2086792	1.8518726	4.1000000	14.5000000
P_MTH	P_MTH	273	8.0703297	1.8369669	4.2000000	13.4000000
RESMTH		265	0.1603774	0.7695285	-1.9000000	2.3000000

----- AGE=7-9 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	215	10.8981395	2.0290867	6.6000000	16.8000000
P_MTH	P_MTH	224	10.9160714	2.2251539	6.1000000	16.5000000
RESMTH		215	-0.0046512	0.9374238	-2.5000000	2.4000000

----- AGE=9-11 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	133	13.9631579	2.6930459	7.6000000	20.3000000
P_MTH	P_MTH	146	13.8287671	2.8684240	8.0000000	21.8000000
RESMTH		133	0.1511278	1.0941730	-3.0000000	2.8000000

----- AGE=11-13 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	92	15.8010870	3.2954012	8.4000000	23.9000000
P_MTH	P_MTH	100	15.3470000	3.2352370	8.9000000	23.7000000
RESMTH		92	0.4336957	0.8434984	-1.4000000	3.0000000

----- AGE=13-15 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	91	18.1043956	3.2836265	10.1000000	23.0000000
P_MTH	P_MTH	100	17.4980000	3.2566321	10.1000000	23.1000000
RESMTH		91	0.8054945	1.1504939	-2.5000000	3.9000000

----- AGE=15-17 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	58	19.3879310	3.4513195	11.8000000	26.0000000
P_MTH	P_MTH	63	19.3222222	3.3307678	12.3000000	25.7000000
RESMTH		58	0.1913793	1.4407947	-3.1000000	4.6000000

----- AGE=17-19 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	23	26.5434783	3.1502996	19.7000000	34.4000000
P_MTH	P_MTH	33	25.9000000	3.1227992	20.9000000	34.3000000
RESMTH		23	0.2956522	3.0560638	-5.9000000	6.2000000

----- AGE=19-21 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	4	18.4750000	0.8655441	17.5000000	19.5000000
P_MTH	P_MTH	4	18.7500000	1.4059398	17.3000000	20.1000000
RESMTH		4	-0.2750000	0.6291529	-1.0000000	0.3000000

----- AGE=21-23 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	21	25.9190476	4.0046996	16.1000000	31.7000000
P_MTH	P_MTH	26	24.9653846	3.1842038	17.1000000	28.9000000
RESMTH		21	1.5619048	1.8583531	-1.4000000	5.1000000

----- AGE=23-25 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	12	22.2916667	4.5127009	15.2000000	29.3000000
P_MTH	P_MTH	14	22.5142857	4.3173812	16.9000000	29.7000000
RESMTH		12	0.7666667	1.4208406	-1.7000000	3.5000000

----- AGE=25-27 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	3	20.4000000	1.2288206	19.5000000	21.8000000
P_MTH	P_MTH	3	20.9333333	4.0722639	18.1000000	25.6000000
RESMTH		3	-0.5333333	2.8448784	-3.8000000	1.4000000

----- AGE=27-29 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	1	24.4000000	.	24.4000000	24.4000000
P_MTH	P_MTH	1	20.5000000	.	20.5000000	20.5000000
RESMTH		1	3.9000000	.	3.9000000	3.9000000

Appendix 2:
Mean Top Height Function: Observed, Predicted and Residual Values for different
stocking classes (stems/km) :

where MTH= Observed Mean Top Height, P_MTH= Predicted Mean Top Height, RESMTH= Residual

----- Stems/km =0-100 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	3	8.7666667	0.9609024	7.9000000	9.8000000
P_MTH	P_MTH	5	10.1600000	2.7373345	6.2000000	13.3000000
RESMTH		3	-1.6666667	0.5131601	-2.1000000	-1.1000000

----- STEMS/KM =100-200 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	112	10.8116071	4.9421654	3.6000000	25.5000000
P_MTH	P_MTH	121	10.7793388	4.8295086	3.6000000	21.6000000
RESMTH		112	-0.0062500	1.1446138	-3.1000000	4.5000000

----- STEMS/KM =200-300 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	144	9.3833333	4.6305009	3.1000000	25.8000000
P_MTH	P_MTH	152	9.4651316	4.7671372	3.1000000	25.8000000
RESMTH		144	0.1020833	0.8679587	-2.5000000	2.4000000

----- STEMS/KM =300-400 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	241	12.8746888	5.3455338	3.7000000	28.9000000
P_MTH	P_MTH	258	12.7837209	5.3867801	3.7000000	27.2000000
RESMTH		241	0.5190871	0.9144037	-2.3000000	3.0000000

----- STEMS/KM =400-500 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	319	13.3959248	6.4709571	3.1000000	34.4000000
P_MTH	P_MTH	332	13.2837349	6.3145093	3.3000000	33.9000000
RESMTH		319	0.3589342	1.2039924	-5.3000000	6.2000000

----- STEMS/KM =500-600 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	79	14.0746835	5.5830666	6.8000000	28.4000000
P_MTH	P_MTH	94	15.1287234	6.3322968	5.2000000	34.3000000
RESMTH		79	0.0468354	1.0163832	-5.9000000	2.1000000

----- STEMS/KM =600-700 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	3	12.6000000	4.8135226	7.4000000	16.9000000
P_MTH	P_MTH	3	12.7333333	4.8583262	7.5000000	17.1000000
RESMTH		3	-0.1333333	0.0577350	-0.2000000	-0.1000000

----- STEMS/KM =700-800 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	24	14.3666667	6.8022162	5.9000000	26.7000000
P_MTH	P_MTH	24	14.3791667	6.8980766	5.8000000	27.2000000
RESMTH		24	-0.0125000	1.4099992	-2.9000000	3.9000000

----- STEMS/KM =800-900 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	38	13.0605263	6.4130584	6.7000000	30.3000000
P_MTH	P_MTH	39	13.2794872	6.5755600	5.6000000	28.6000000
RESMTH		38	0.1842105	0.8870011	-1.5000000	2.6000000

----- STEMS/KM =900-1000 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	103	10.2436893	4.1898025	4.4000000	21.8000000
P_MTH	P_MTH	108	10.3768519	4.4145386	3.9000000	25.6000000
RESMTH		103	0.0922330	1.1296095	-3.8000000	2.3000000

----- STEMS/KM =>1000 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	4	8.1750000	2.9375443	4.7000000	11.4000000
P_MTH	P_MTH	4	7.9250000	3.0793668	4.7000000	11.7000000
RESMTH		4	0.2500000	0.4795832	-0.3000000	0.7000000

----- TIME=0-8 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	6	9.0500000	2.8521921	6.6000000	13.4000000
P_MTH	P_MTH	6	8.9666667	2.7645373	6.7000000	12.7000000
RESMTH		6	0.0833333	0.6823977	-0.9000000	1.1000000

----- TIME=0-9 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	130	13.6315385	5.0588633	5.4000000	24.4000000
P_MTH	P_MTH	138	13.4188406	4.8802230	5.1000000	25.6000000
RESMTH		130	0.3361538	1.0625407	-3.8000000	3.9000000

----- TIME=0-12 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	16	14.0812500	5.2636133	7.8000000	25.4000000
P_MTH	P_MTH	16	15.1000000	6.0181946	8.1000000	27.9000000
RESMTH		16	-1.0187500	0.9042262	-3.4000000	0

----- TIME=0-13 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
MTH	MTH	145	16.8262069	6.5277402	6.8000000	34.4000000
P_MTH	P_MTH	166	17.1054217	6.3533743	6.4000000	34.3000000
RESMTH		145	0.4737931	1.3902948	-5.9000000	6.2000000

Appendix 4:
Basal Area Increment Function: Observed, Predicted and Residual Values for
different age classes:

where BALIVE= Live Basal Area, P_BA = Predicted basal area using regional asymptote estimates, P_BA2 = Predicted basal area using plot asymptote estimates, RESBA= Residuals for the Basal area function using regional asymptote estimates and RESBA2 = Residuals for the Basal area function using plot asymptote estimates.

----- AGE=<5 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	153	4.3807843	2.7827520	0.5000000	12.5900000
P_BA	P_BA	153	4.2056863	2.7134959	0.1000000	12.6500000
P_BA2	P_BA	153	4.2624183	2.7225527	0.1000000	12.7400000
RESBA		153	0.1750980	0.8616234	-2.4400000	2.8000000
RESBA2		153	0.1183660	0.8660728	-2.5200000	2.7800000

----- AGE=5-7 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	273	11.2565201	6.5780830	0.9000000	38.0000000
P_BA	P_BA	273	11.2398901	6.2854104	0.1000000	32.0100000
P_BA2	P_BA	273	11.3831136	6.3029354	0.1000000	32.2900000
RESBA		273	0.0166300	2.2469068	-6.8000000	7.8400000
RESBA2		273	-0.1265934	2.2521269	-6.8700000	7.5100000

----- AGE=7-9 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	224	18.8141964	10.4904551	2.7000000	57.6000000
P_BA	P_BA	224	20.0442857	9.7337838	1.1400000	47.6300000
P_BA2	P_BA	224	20.3197768	9.7696401	1.2400000	48.0900000
RESBA		224	-1.2300893	4.2878767	-13.8500000	11.6400000
RESBA2		224	-1.5055804	4.2764062	-14.2700000	11.0600000

----- AGE=9-11 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	146	26.6159589	12.8732592	4.0000000	68.4000000
P_BA	P_BA	146	29.2073288	11.3969093	2.5500000	56.5200000
P_BA2	P_BA	146	29.6178767	11.4559272	2.7200000	56.6500000
RESBA		146	-2.5913699	5.2887761	-16.1800000	14.6900000
RESBA2		146	-3.0019178	5.2715006	-16.4600000	14.1900000

----- AGE=11-13 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	100	31.1490000	12.8391627	10.2200000	66.3400000
P_BA	P_BA	100	33.9677000	11.6948012	11.7900000	65.7100000
P_BA2	P_BA	100	34.3377000	11.8891534	11.9000000	66.6100000
RESBA		100	-2.8187000	4.5323011	-18.1400000	5.9400000
RESBA2		100	-3.1887000	4.5654837	-18.8900000	5.2400000

----- AGE=25-27 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	3	95.5700000	67.7600598	55.9000000	173.8100000
P_BA	P_BA	3	93.1500000	66.5598039	53.8500000	170.0000000
P_BA2	P_BA	3	93.5466667	66.8742382	54.0800000	170.7600000
RESBA		3	2.4200000	1.2468761	1.4000000	3.8100000
RESBA2		3	2.0233333	0.9416121	1.2000000	3.0500000

----- AGE=27-29 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	1	160.1200000	.	160.1200000	160.1200000
P_BA	P_BA	1	144.3300000	.	144.3300000	144.3300000
P_BA2	P_BA	1	145.7500000	.	145.7500000	145.7500000
RESBA		1	15.7900000	.	15.7900000	15.7900000
RESBA2		1	14.3700000	.	14.3700000	14.3700000

Appendix 7:
Basal Area Increment Functions: Observed, Predicted and Residual Values for
different regions:

where BALIVE= Live Basal Area, P_BA = Predicted basal area using regional asymptote estimates, P_BA2 = Predicted basal area using plot asymptote estimates, RESBA= Residuals for the Basal area function using regional asymptote estimates and RESBA2 = Residuals for the Basal area function using plot asymptote estimates.

----- REGION= Auckland-----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	4	17.7975000	3.8502846	13.9200000	22.3200000
P_BA	P_BA	4	15.0300000	2.6698689	12.1600000	18.0000000
P_BA2	P_BA	4	15.1450000	2.7137981	12.2300000	18.1700000
RESBA		4	2.7675000	1.2048617	1.7600000	4.3200000
RESBA2		4	2.6525000	1.1610448	1.6900000	4.1500000

----- REGION= Bay of Plenty-----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	654	29.3453517	28.0128725	0.9000000	148.3600000
P_BA	P_BA	654	30.2655810	26.9345866	1.7600000	149.6700000
P_BA2	P_BA	654	30.6635933	27.1978966	1.8000000	149.8300000
RESBA		654	-0.9202294	5.1642708	-19.3200000	21.4600000
RESBA2		654	-1.3182416	5.1593582	-22.4600000	19.2000000

----- REGION= Canterbury-----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	76	34.0573684	35.3306713	2.8900000	173.8100000
P_BA	P_BA	76	37.4523684	35.5489524	3.0800000	170.0000000
P_BA2	P_BA	76	37.7685526	35.6839645	3.1200000	170.7600000
RESBA		76	-3.3950000	7.0587295	-36.9400000	15.7900000
RESBA2		76	-3.7111842	7.1512067	-37.5900000	14.3700000

----- REGION= Hawkes Bay-----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	232	23.4478017	16.4483656	0.6000000	64.5800000
P_BA	P_BA	232	26.0844397	19.4712207	0.1000000	64.9500000
P_BA2	P_BA	232	26.3097414	19.5225327	0.1000000	65.3200000
RESBA		232	-2.6366379	7.0788243	-42.6800000	10.2500000
RESBA2		232	-2.8619397	7.1000645	-43.0200000	9.9800000

----- REGION=Manawatu -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	30	7.7466667	7.2762596	0.5000000	29.4000000
P_BA	P_BA	30	8.3040000	8.0286206	0.3000000	32.9500000
P_BA2	P_BA	30	8.6446667	8.0848968	0.3200000	33.2500000
RESBA		30	-0.5573333	1.6407272	-4.1600000	3.7100000
RESBA2		30	-0.8980000	1.5370469	-4.3700000	3.5000000

----- AGE=15-17 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	63	480.2063492	202.6369038	150.0000000	1000.00
P_SPH	P_SPH	63	508.1857143	184.9236092	170.8000000	997.8000000
RESSPH		63	-27.9793651	64.7173636	-299.3000000	3.3000000

----- AGE=17-19 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	33	520.6060606	127.6976261	342.0000000	901.0000000
P_SPH	P_SPH	33	541.3060606	119.0713151	424.6000000	900.7000000
RESSPH		33	-20.7000000	35.4658374	-109.4000000	3.5000000

----- AGE=19-21 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	4	476.0000000	374.2735185	139.0000000	800.0000000
P_SPH	P_SPH	4	475.8750000	374.1292413	139.0000000	799.8000000
RESSPH		4	0.1250000	0.1500000	0	0.3000000

----- AGE=21-23 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	26	521.5384615	252.4138238	122.0000000	868.0000000
P_SPH	P_SPH	26	529.6307692	259.9746507	122.0000000	867.8000000
RESSPH		26	-8.0923077	20.4981935	-76.7000000	0.4000000

----- AGE=23-25 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	14	363.1428571	146.6465437	122.0000000	569.0000000
P_SPH	P_SPH	14	363.9500000	145.0198806	122.0000000	568.8000000
RESSPH		14	-0.8071429	3.5111831	-13.0000000	0.5000000

----- AGE=25-27 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	3	630.6666667	234.1203394	494.0000000	901.0000000
P_SPH	P_SPH	3	630.0333333	233.3697995	493.8000000	899.5000000
RESSPH		3	0.6333333	0.7505553	0.2000000	1.5000000

----- AGE=27-29 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	1	800.0000000	.	800.0000000	800.0000000
P_SPH	P_SPH	1	798.9000000	.	798.9000000	798.9000000
RESSPH		1	1.1000000	.	1.1000000	1.1000000

Appendix 9:

Mortality Function: Observed, Predicted and Residual Values for different stocking classes:

where sphlive=stems per kilometre, p_sph = predicted stems per kilometre, and ressph=residual stems per km

----- STEMS/KM =0-100 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	5	50.0000000	0	50.0000000	50.0000000
P_SPH	P_SPH	5	249.3400000	0.2302173	249.1000000	249.7000000
RESSPH		5	-199.3400000	0.2302173	-199.7000000	-199.1000000

----- STEMS/KM =100-200 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	121	180.8099174	22.2487131	122.0000000	200.0000000
P_SPH	P_SPH	121	213.8801653	55.8108899	122.0000000	449.3000000
RESSPH		121	-33.0702479	53.2164599	-299.3000000	1.0000000

----- STEMS/KM =200-300 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	152	251.8750000	31.7294675	202.0000000	300.0000000
P_SPH	P_SPH	152	287.3282895	72.0574940	201.0000000	499.5000000
RESSPH		152	-35.4532895	55.4003968	-230.3000000	1.5000000

----- STEMS/KM =300-400 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	258	360.8604651	29.2458369	303.0000000	400.0000000
P_SPH	P_SPH	258	376.8007752	44.1321750	302.9000000	499.8000000
RESSPH		258	-15.9403101	33.2416386	-165.2000000	2.2000000

----- STEMS/KM =400-500 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	332	466.1807229	32.9582340	403.0000000	500.0000000
P_SPH	P_SPH	332	477.0141566	34.9176469	402.6000000	562.6000000
RESSPH		332	-10.8334337	24.1070109	-110.6000000	3.5000000

----- STEMS/KM =500-600 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	94	533.9468085	20.3223311	502.0000000	600.0000000
P_SPH	P_SPH	94	542.9234043	45.1390740	500.7000000	728.3000000
RESSPH		94	-8.9765957	38.1961756	-194.7000000	3.5000000

----- STEMS/KM =600-700 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	3	637.6666667	25.4034118	623.0000000	667.0000000
P_SPH	P_SPH	3	710.0666667	17.1820643	699.7000000	729.9000000
RESSPH		3	-72.4000000	8.2395388	-77.6000000	-62.9000000

----- STEMS/KM =700-800 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	24	758.8333333	26.7283706	701.0000000	800.0000000
P_SPH	P_SPH	24	803.4833333	88.5742463	700.8000000	991.7000000
RESSPH		24	-44.6500000	89.8778398	-258.7000000	2.6000000

----- STEMS/KM =800-900 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	39	863.6153846	32.4391052	801.0000000	900.0000000
P_SPH	P_SPH	39	914.1564103	60.0677372	800.6000000	995.0000000
RESSPH		39	-50.5410256	35.7984101	-97.7000000	2.8000000

----- STEMS/KM =900-1000 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	108	979.2222222	29.0166995	901.0000000	1000.00
P_SPH	P_SPH	108	980.1925926	27.1493299	899.5000000	999.4000000
RESSPH		108	-0.9703704	15.7063965	-57.8000000	9.8000000

----- STEMS/KM =>1000 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	4	1067.00	0	1067.00	1067.00
P_SPH	P_SPH	4	1057.30	3.4292856	1053.70	1061.50
RESSPH		4	9.7000000	3.4292856	5.5000000	13.3000000

Appendix 10:

Mortality Function: Observed, Predicted and Residual Values for different length of prediction intervals:

where sphlive=stems per kilometre, p_sph = predicted stems per kilometre, and ressph=residual stems per km

----- TIME=0-1 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	38	443.5526316	244.5948329	160.0000000	1000.00
P_SPH	P_SPH	38	446.6631579	243.4521946	172.7000000	997.0000000
RESSPH		38	-3.1105263	10.2772062	-43.4000000	3.2000000

----- TIME=0-2 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	77	449.2077922	213.2605484	122.0000000	1000.00
P_SPH	P_SPH	77	461.9545455	208.1246353	122.0000000	996.7000000
RESSPH		77	-12.7467532	26.6916781	-110.4000000	6.7000000

----- TIME=0-3 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	208	480.4759615	233.5880792	150.0000000	1000.00
P_SPH	P_SPH	208	492.6187500	234.8890205	166.7000000	999.1000000
RESSPH		208	-12.1427885	29.3624978	-133.5000000	5.0000000

----- TIME=0-4 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	120	412.2333333	237.5480043	126.0000000	1067.00
P_SPH	P_SPH	120	435.8266667	223.2379665	139.0000000	1061.50
RESSPH		120	-23.5933333	50.2147769	-299.3000000	13.3000000

----- TIME=0-5 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	171	465.9883041	295.4165949	50.0000000	1000.00
P_SPH	P_SPH	171	495.7035088	285.7838346	170.8000000	999.4000000
RESSPH		171	-29.7152047	62.1999991	-258.7000000	9.0000000

----- TIME=0-6 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	200	498.2300000	281.7510521	167.0000000	1000.00
P_SPH	P_SPH	200	534.6040000	279.9737583	199.1000000	997.1000000
RESSPH		200	-36.3740000	51.0625267	-230.3000000	9.8000000

----- TIME=0-8 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	6	349.0000000	19.6265127	313.0000000	365.0000000
P_SPH	P_SPH	6	360.9500000	3.8365349	357.0000000	364.8000000
RESSPH		6	-11.9500000	17.2383004	-44.0000000	0.4000000

----- TIME=0-9 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	138	368.3188406	156.3013179	195.0000000	901.0000000
P_SPH	P_SPH	138	376.6920290	159.1345572	194.0000000	900.7000000
RESSPH		138	-8.3731884	16.0816509	-77.6000000	2.2000000

----- TIME=0-12 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	16	481.5000000	19.1067178	463.0000000	500.0000000
P_SPH	P_SPH	16	479.9562500	19.0376808	460.3000000	499.6000000
RESSPH		16	1.5437500	0.7571603	0.3000000	3.0000000

----- TIME=0-13 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SPHLIVE	SPHLIVE	166	490.9698795	44.7131497	342.0000000	565.0000000
P_SPH	P_SPH	166	504.5295181	34.7841226	424.6000000	564.6000000
RESSPH		166	-13.5596386	30.8555441	-110.6000000	3.5000000

Appendix 11:

Alternative Basal Area Increment Function: Observed, Predicted and Residual values for different age classes.

where BALIVE= Live Basal Area, P_BA = Predicted basal area using regional asymptote estimates, P_BA2 = Predicted basal area using plot asymptote estimates, RESBA= Residuals for the Basal area function using regional asymptote estimates and RESBA2 = Residuals for the Basal area function using plot asymptote estimates.

----- AGE=<5 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	153	4.3807843	2.7827520	0.5000000	12.5900000
P_BA	P_BA	153	4.7341176	2.4112506	1.0800000	11.5600000
P_BA2	P_BA	153	4.8493464	2.4496194	1.1200000	11.9000000
RESBA		153	-0.3533333	0.9189790	-2.6600000	3.2200000
RESBA2		153	-0.4685621	0.9108451	-2.7800000	3.0600000

----- AGE=5-7 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	273	11.2565201	6.5780830	0.9000000	38.0000000
P_BA	P_BA	273	11.2913553	5.9080211	2.4100000	33.1600000
P_BA2	P_BA	273	11.5241026	5.9739513	2.5300000	33.8100000
RESBA		273	-0.0348352	2.2660096	-10.3400000	8.2800000
RESBA2		273	-0.2675824	2.2790099	-10.8800000	7.7000000

----- AGE=7-9 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	224	18.8141964	10.4904551	2.7000000	57.6000000
P_BA	P_BA	224	19.3840625	9.5721370	5.2700000	48.5100000
P_BA2	P_BA	224	19.8095536	9.7060589	5.4200000	49.5200000
RESBA		224	-0.5698661	4.2071716	-15.1100000	13.0500000
RESBA2		224	-0.9953571	4.2324938	-15.9700000	12.3300000

----- AGE=9-11 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	146	26.6159589	12.8732592	4.0000000	68.4000000
P_BA	P_BA	146	27.6861644	11.0535436	8.3200000	55.9900000
P_BA2	P_BA	146	28.2607534	11.1912764	8.5800000	57.1900000
RESBA		146	-1.0702055	5.0915315	-17.5600000	15.2000000
RESBA2		146	-1.6447945	5.0724958	-18.5900000	14.7100000

----- AGE=11-13 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	100	31.1490000	12.8391627	10.2200000	66.3400000
P_BA	P_BA	100	31.9751000	10.3983349	10.9300000	57.9500000
P_BA2	P_BA	100	32.4304000	10.5906350	11.0000000	58.1900000
RESBA		100	-0.8261000	4.0217479	-14.5000000	11.1500000
RESBA2		100	-1.2814000	3.8299128	-14.6600000	9.4200000

----- AGE=13-15 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	100	39.1607000	13.5014861	12.5000000	75.2000000
P_BA	P_BA	100	42.1753000	10.6355836	16.8400000	64.0900000
P_BA2	P_BA	100	42.6969000	10.8239190	17.0700000	65.2800000
RESBA		100	-3.0146000	5.8081187	-24.0300000	11.1100000
RESBA2		100	-3.5362000	5.6913276	-24.3600000	9.9200000

----- AGE=15-17 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	63	49.7123810	18.1516810	17.6000000	87.8200000
P_BA	P_BA	63	53.8017460	12.5589855	21.9600000	77.3700000
P_BA2	P_BA	63	54.4204762	12.8099508	22.3300000	78.6100000
RESBA		63	-4.0893651	10.2661811	-38.8000000	15.5900000
RESBA2		63	-4.7080952	10.1452968	-39.3100000	14.5300000

----- AGE=17-19 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	33	88.4921212	16.0424614	59.7900000	131.3300000
P_BA	P_BA	33	84.5978788	11.9349528	69.5200000	123.3000000
P_BA2	P_BA	33	85.9893939	11.7052363	70.9100000	123.4200000
RESBA		33	3.8942424	10.2169388	-14.1800000	19.0800000
RESBA2		33	2.5027273	10.1414860	-14.7700000	17.4100000

----- AGE=19-21 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	4	70.4225000	50.0055026	25.7000000	119.0900000
P_BA	P_BA	4	67.5450000	46.7226430	25.1900000	109.9600000
P_BA2	P_BA	4	67.6250000	46.7235055	25.2700000	110.0700000
RESBA		4	2.8775000	4.2661409	-0.1500000	9.1300000
RESBA2		4	2.7975000	4.2495127	-0.2300000	9.0200000

----- AGE=21-23 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	26	101.1907692	35.9378155	27.9500000	148.3600000
P_BA	P_BA	26	100.4465385	35.1544987	27.4900000	148.2100000
P_BA2	P_BA	26	100.5807692	35.1238778	27.6500000	148.3600000
RESBA		26	0.7442308	3.9779062	-10.6000000	6.0300000
RESBA2		26	0.6100000	4.0723536	-11.1400000	5.9100000

----- AGE=23-25 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	14	71.5221429	32.4012759	29.4900000	130.8700000
P_BA	P_BA	14	71.1128571	29.7482388	31.6900000	123.6500000
P_BA2	P_BA	14	71.3078571	29.7331965	32.0100000	123.8200000
RESBA		14	0.4092857	4.0626033	-9.8500000	7.2200000
RESBA2		14	0.2142857	4.1391465	-10.4200000	7.0500000

----- AGE=25-27 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	3	95.5700000	67.7600598	55.9000000	173.8100000
P_BA	P_BA	3	87.8433333	52.9844735	56.6100000	149.0200000
P_BA2	P_BA	3	88.0966667	53.1721594	56.7500000	149.4900000
RESBA		3	7.7266667	14.7775855	-0.9000000	24.7900000
RESBA2		3	7.4733333	14.5899840	-1.0500000	24.3200000

----- AGE=27-29 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	1	160.1200000	.	160.1200000	160.1200000
P_BA	P_BA	1	133.1600000	.	133.1600000	133.1600000
P_BA2	P_BA	1	133.6100000	.	133.6100000	133.6100000
RESBA		1	26.9600000	.	26.9600000	26.9600000
RESBA2		1	26.5100000	.	26.5100000	26.5100000

Appendix 12:

Alternative Basal Area Increment Function: Observed, Predicted and Residual values for different stocking classes.

where BALIVE= Live Basal Area, P_BA = Predicted basal area using regional asymptote estimates, P_BA2 = Predicted basal area using plot asymptote estimates, RESBA= Residuals for the Basal area function using regional asymptote estimates and RESBA2 = Residuals for the Basal area function using plot asymptote estimates.

----- STEMS/KM =0-100 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	5	2.5800000	1.1945711	0.9000000	4.0000000
P_BA	P_BA	5	10.3740000	5.5234482	2.7600000	17.0600000
P_BA2	P_BA	5	10.6720000	5.6965533	2.8200000	17.5700000
RESBA		5	-7.7940000	4.3293510	-13.0600000	-1.8600000
RESBA2		5	-8.0920000	4.5024571	-13.5700000	-1.9200000

----- STEMS/KM =100-200 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	121	11.4442975	9.5355022	0.6000000	53.3400000
P_BA	P_BA	121	13.4560331	11.6218788	1.1900000	56.6000000
P_BA2	P_BA	121	13.8273554	11.7166949	1.2400000	57.1100000
RESBA		121	-2.0117355	5.8381905	-38.8000000	3.3500000
RESBA2		121	-2.3830579	5.8885473	-39.3100000	2.9300000

----- STEMS/KM =200-300 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	152	11.6444737	12.6707902	0.5000000	77.7700000
P_BA	P_BA	152	13.1615132	14.1973036	1.0800000	77.1300000
P_BA2	P_BA	152	13.4926316	14.3361956	1.1200000	77.3100000
RESBA		152	-1.5170395	4.7858099	-28.1700000	7.9800000
RESBA2		152	-1.8481579	4.8018845	-28.6800000	6.4200000

----- STEMS/KM =300-400 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	258	23.0201938	18.3149811	1.1000000	104.8100000
P_BA	P_BA	258	24.2752713	18.7833878	1.8300000	103.9800000
P_BA2	P_BA	258	24.6800388	18.9433743	1.9100000	104.1500000
RESBA		258	-1.2550775	4.0402128	-14.1500000	11.1500000
RESBA2		258	-1.6598450	4.0818701	-14.9300000	9.4200000

----- STEMS/KM =400-500 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	332	29.5341867	24.7852097	1.1000000	130.8700000
P_BA	P_BA	332	29.7659940	23.7091149	2.0900000	123.6500000
P_BA2	P_BA	332	30.1654819	23.9750297	2.1500000	123.8200000
RESBA		332	-0.2318072	4.0214100	-10.5100000	18.1900000
RESBA2		332	-0.6312952	3.9440936	-11.9200000	16.8600000

----- STEMS/KM =500-600 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	94	38.4384043	30.4327297	5.6100000	127.9800000
P_BA	P_BA	94	36.5931915	26.5328355	6.2100000	121.9500000
P_BA2	P_BA	94	37.1113830	26.9016382	6.2800000	122.1100000
RESBA		94	1.8452128	6.1289587	-17.5600000	19.0800000
RESBA2		94	1.3270213	5.8800451	-18.5900000	17.4100000

----- STEMS/KM =600-700 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	3	34.9600000	26.8261309	6.6000000	59.9300000
P_BA	P_BA	3	40.9466667	28.0562298	12.7400000	68.8500000
P_BA2	P_BA	3	41.2666667	28.2001708	13.0100000	69.4100000
RESBA		3	-5.9866667	3.0129277	-8.9200000	-2.9000000
RESBA2		3	-6.3066667	3.2262414	-9.4800000	-3.0300000

----- STEMS/KM =700-800 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	24	58.3254167	49.0245372	3.5000000	160.1200000
P_BA	P_BA	24	59.0854167	44.8437575	6.1600000	141.5600000
P_BA2	P_BA	24	59.4316667	44.7930314	6.2800000	141.7100000
RESBA		24	-0.7600000	9.2723249	-14.1800000	26.9600000
RESBA2		24	-1.1062500	9.3107352	-14.7700000	26.5100000

----- STEMS/KM =800-900 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	39	50.9879487	41.2875128	10.5500000	148.3600000
P_BA	P_BA	39	48.6948718	41.2639270	12.6300000	148.2100000
P_BA2	P_BA	39	49.0789744	41.1968072	12.8000000	148.3600000
RESBA		39	2.2930769	6.5747260	-7.9700000	15.2000000
RESBA2		39	1.9089744	6.4098217	-8.4700000	14.7100000

----- STEMS/KM =900-1000 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	108	29.6691667	24.7316996	3.5000000	173.8100000
P_BA	P_BA	108	30.4343519	23.5530807	4.9700000	149.0200000
P_BA2	P_BA	108	30.8212037	23.6545062	5.0600000	149.4900000
RESBA		108	-0.7651852	4.6443449	-10.9700000	24.7900000
RESBA2		108	-1.1520370	4.7144889	-11.3800000	24.3200000

----- STEMS/KM =>1000 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	4	16.1500000	10.7413531	4.7000000	29.4000000
P_BA	P_BA	4	17.2050000	10.6172988	5.8400000	29.7400000
P_BA2	P_BA	4	17.7450000	10.9786292	6.0000000	30.7100000
RESBA		4	-1.0550000	0.7583095	-2.0800000	-0.3400000
RESBA2		4	-1.5950000	0.7964295	-2.7700000	-1.0000000

Appendix 13:

Alternative Basal Area Increment Function: Observed, Predicted and Residual values for different length of prediction periods:

where BALIVE= Live Basal Area, P_BA = Predicted basal area using regional asymptote estimates, P_BA2 = Predicted basal area using plot asymptote estimates, RESBA= Residuals for the Basal area function using regional asymptote estimates and RESBA2 = Residuals for the Basal area function using plot asymptote estimates.

----- TIME=0-1 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	38	14.3444737	13.6775724	2.5000000	67.1900000
P_BA	P_BA	38	15.7700000	15.6277636	2.9100000	81.3700000
P_BA2	P_BA	38	15.9305263	15.7097412	3.0000000	81.9600000
RESBA		38	-1.4255263	3.1089119	-14.1800000	4.8000000
RESBA2		38	-1.5860526	3.2103360	-14.7700000	4.6400000

----- TIME=0-2 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	77	51.8561039	49.0005478	0.6600000	148.3600000
P_BA	P_BA	77	50.8819481	48.1056211	1.3100000	148.2100000
P_BA2	P_BA	77	51.0497403	48.0750000	1.3800000	148.3600000
RESBA		77	0.9741558	2.3094693	-3.7500000	7.2200000
RESBA2		77	0.8063636	2.3260826	-3.8900000	7.0500000

----- TIME=0-3 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	208	16.8165865	13.9633593	1.9300000	63.1000000
P_BA	P_BA	208	17.4577885	13.4941720	2.3300000	60.5600000
P_BA2	P_BA	208	17.6815385	13.5409883	2.3900000	61.0600000
RESBA		208	-0.6412019	3.1629652	-10.9700000	15.2000000
RESBA2		208	-0.8649519	3.1420974	-11.3800000	14.7100000

----- TIME=0-4 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	120	20.3097500	15.0214582	0.5000000	55.8000000
P_BA	P_BA	120	23.1479167	17.1498204	1.1400000	58.4600000
P_BA2	P_BA	120	23.4685000	17.2178187	1.1900000	58.9700000
RESBA		120	-2.8381667	6.5256478	-38.8000000	9.4400000
RESBA2		120	-3.1587500	6.5420981	-39.3100000	8.7500000

----- TIME=0-5 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	171	20.3673099	15.6108953	0.6000000	64.5800000
P_BA	P_BA	171	22.0042690	16.8002622	1.0800000	61.8300000
P_BA2	P_BA	171	22.3252632	16.8635826	1.1200000	62.2500000
RESBA		171	-1.6369591	5.1588593	-28.1700000	13.0500000
RESBA2		171	-1.9579532	5.1632036	-28.6800000	12.3300000

----- TIME=0-6 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	200	17.5216000	12.5520961	1.6000000	68.4000000
P_BA	P_BA	200	19.3333000	12.8234153	2.1200000	55.9900000
P_BA2	P_BA	200	19.8031000	13.0861055	2.1900000	57.1900000
RESBA		200	-1.8117000	4.4027275	-17.5600000	15.1400000
RESBA2		200	-2.2815000	4.4680259	-18.5900000	13.9900000

----- TIME=0-8 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	6	11.0516667	7.6145319	4.9300000	22.2600000
P_BA	P_BA	6	16.4066667	12.8076753	6.7400000	33.2900000
P_BA2	P_BA	6	16.7300000	13.1586793	6.8000000	34.0700000
RESBA		6	-5.3550000	5.4214417	-14.1500000	-1.6100000
RESBA2		6	-5.6783333	5.7637919	-14.9300000	-1.6700000

----- TIME=0-9 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	138	28.6789855	28.2615529	1.9800000	173.8100000
P_BA	P_BA	138	28.0813768	26.3184623	2.2900000	149.0200000
P_BA2	P_BA	138	28.6406522	26.4252054	2.3600000	149.4900000
RESBA		138	0.5976087	4.7151781	-13.5400000	26.9600000
RESBA2		138	0.0383333	4.6951591	-14.2400000	26.5100000

----- TIME=0-12 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	16	27.5431250	22.4435264	8.8400000	84.7400000
P_BA	P_BA	16	29.9487500	21.8053443	9.0800000	81.6200000
P_BA2	P_BA	16	30.4412500	22.2981778	9.1300000	83.2700000
RESBA		16	-2.4056250	2.2899955	-5.0700000	3.1200000
RESBA2		16	-2.8981250	2.2198205	-5.7400000	1.4700000

----- TIME=0-13 -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	166	44.3531325	27.7980535	8.2100000	116.2900000
P_BA	P_BA	166	42.4881325	25.2133490	8.2500000	100.4300000
P_BA2	P_BA	166	43.1557831	25.7334925	8.2800000	102.2400000
RESBA		166	1.8650000	6.0959673	-12.4700000	19.0800000
RESBA2		166	1.1973494	5.9583776	-14.1400000	17.4100000

Appendix 14:

Alternative Basal Area Increment Function: Observed, Predicted and Residual values for different regions:

where BALIVE= Live Basal Area, P_BA = Predicted basal area using regional asymptote estimates, P_BA2 = Predicted basal area using plot asymptote estimates, RESBA= Residuals for the Basal area function using regional asymptote estimates and RESBA2 = Residuals for the Basal area function using plot asymptote estimates.

----- REGION=Auckland-----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	4	17.7975000	3.8502846	13.9200000	22.3200000
P_BA	P_BA	4	14.7950000	2.5300395	12.1200000	17.5200000
P_BA2	P_BA	4	14.9200000	2.5689816	12.2100000	17.6800000
RESBA		4	3.0025000	1.3558853	1.8000000	4.8000000
RESBA2		4	2.8775000	1.3203125	1.7100000	4.6400000

----- REGION=Bay of Plenty-----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	654	29.3453517	28.0128725	0.9000000	148.3600000
P_BA	P_BA	654	29.2621713	26.4944829	2.1200000	148.2100000
P_BA2	P_BA	654	29.7520336	26.6998483	2.1900000	148.3600000
RESBA		654	0.0831804	4.7398702	-17.5600000	19.0800000
RESBA2		654	-0.4066820	4.6959211	-18.5900000	17.4100000

----- REGION=Canterbury-----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	76	34.0573684	35.3306713	2.8900000	173.8100000
P_BA	P_BA	76	36.2617105	32.3022099	4.1000000	149.0200000
P_BA2	P_BA	76	36.5161842	32.3472302	4.1800000	149.4900000
RESBA		76	-2.2043421	6.2034312	-14.1800000	26.9600000
RESBA2		76	-2.4588158	6.2835512	-14.9300000	26.5100000

----- REGION=Hawkes Bay-----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	232	23.4478017	16.4483656	0.6000000	64.5800000
P_BA	P_BA	232	25.6079310	18.1360532	1.0800000	61.8300000
P_BA2	P_BA	232	25.9165948	18.2018458	1.1200000	62.2500000
RESBA		232	-2.1601293	6.3084273	-38.8000000	13.0500000
RESBA2		232	-2.4687931	6.3138464	-39.3100000	12.3300000

----- REGION=Manawatu-----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	30	7.7466667	7.2762596	0.5000000	29.4000000
P_BA	P_BA	30	8.4963333	6.8541523	1.1400000	29.7400000
P_BA2	P_BA	30	8.8303333	7.0644830	1.1900000	30.7100000
RESBA		30	-0.7496667	1.1992972	-2.3900000	3.1500000
RESBA2		30	-1.0836667	1.1817914	-2.9000000	2.6900000

----- REGION=Nelson/Marlborough -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	24	5.6712500	4.4488800	0.6600000	19.6000000
P_BA	P_BA	24	5.6279167	3.7692705	1.3100000	16.7700000
P_BA2	P_BA	24	5.8337500	3.8737767	1.3800000	17.2500000
RESBA		24	0.0433333	1.0703135	-1.8800000	2.8300000
RESBA2		24	-0.1625000	1.0302648	-2.1900000	2.3500000

----- REGION=Northland -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	15	10.0953333	7.3239323	2.6200000	28.1100000
P_BA	P_BA	15	10.7246667	7.3728071	2.9100000	29.4200000
P_BA2	P_BA	15	10.8853333	7.4734233	3.0000000	29.9100000
RESBA		15	-0.6293333	0.6877029	-1.8900000	0.5100000
RESBA2		15	-0.7900000	0.7182021	-2.0200000	0.2600000

----- REGION=Taranaki -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	40	9.3150000	5.2753170	2.0300000	23.3100000
P_BA	P_BA	40	10.8845000	6.0033123	2.8300000	26.9500000
P_BA2	P_BA	40	11.0642500	6.0786768	2.8900000	27.3600000
RESBA		40	-1.5695000	2.0326778	-10.9700000	0.8800000
RESBA2		40	-1.7492500	2.0709497	-11.3800000	0.5500000

----- REGION=Wairarapa -----

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
BALIVE	BALIVE	65	16.9790769	7.6471577	1.9300000	33.2400000
P_BA	P_BA	65	18.1789231	8.1408737	2.3300000	35.3500000
P_BA2	P_BA	65	18.3993846	8.1942365	2.3900000	35.6100000
RESBA		65	-1.1998462	1.7511144	-4.9600000	3.3400000
RESBA2		65	-1.4203077	1.7015627	-5.2400000	2.6500000