The Predictive Ability of Seven Sigmoid Curves used in Modelling Forestry Growth

By

Anthony Lee

A thesis submitted to the
Institute of Information Sciences and Technology
in partial fulfillment of the requirements for the degree of

Master of Applied Statistics

at

MASSEY UNIVERSITY

February, 2000

Thesis Supervisor: G. R. Wood, Professor of Statistics, Massey University

ABSTRACT

In this thesis we study seven sigmoid growth curve families to determine which best fit pinus radiata basal area against age data. We fit the sigmoid models to the data of distinct plots rather than the pooled data of sets of plots. The seven growth models are the three-parameter Chapman-Richards, Hossfeld, Schumacher, Weibull and Gompertz models and the four-parameter Levakovic and Sloboda models. This investigation was inspired by Dr. Richard Woollons' observation that sigmoid curves vary consistently in their estimation of the asymptote, and those functions giving bigger asymptotes have better goodness-of-fit properties.

It is shown that models with better goodness-of-fit properties are better predictors of basal area. This indicates that the three-parameter Chapman-Richards model and the four-parameter Levakovic and Sloboda models are superior to the Hossfeld, Schumacher, Weibull and Gompertz models. We do however recommend caution in the use of the Levakovic model where convergence of the nonlinear least squares algorithm is often difficult. Also, parameter-effects curvature of the four-parameter models was on occasion seen to be unacceptably large and we recommend careful examination of curvature in the selection of a candidate function.

We demonstrate that the Schumacher model predicts larger asymptotes than the other models but do not conclude that this model has better goodness-of-fit properties, contrary to Dr. Woollons' observation. We do however conclude that models with better goodness-of-fit do have better predictive power.

The study of the growth curves is divided into six parts. Firstly, we investigate fundamental properties of the growth curves with particular attention paid to the point of inflection and the asymptote.

Secondly, the models are fitted to pinus radiata data and goodness-of-fit properties are investigated. Additionally, we discuss practical considerations required when fitting these models using nonlinear least squares; in particular the location of starting

values and reparameterization of the growth models. We note the ease of fitting the Chapman-Richards model and the relative difficulty in fitting the Levakovic and Sloboda models.

Thirdly, we empirically demonstrate that the Schumacher model has the largest asymptote.

Fourthly, we investigate the robustness of the models in predicting basal area using both pinus radiata data and simulated data.

Next, Padé rational approximations of the growth curves are investigated to begin a theoretical study of the observed behaviour of the fitted growth curves in an attempt to explain the goodness-of-fit and predictive power of the curves by providing a common basis of comparison. A possible additional area of research is indicated by this analysis - the estimation of properties at the point of inflection and the fitting of the associated rational function to the data.

Finally, the results in preceding chapters are summarised to rank the growth curves according to their effectiveness in modelling pinus radiata growth data. We do not conclude that one model is optimal in all respects but do give a hierarchy of suitability, with the Chapman-Richards model at the top of this hierarchy.

CERTIFICATION

I, Anthony Lee, certify that this thesis represents original work, except where acknowledged.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor, Professor Graham Wood for his patient support over the last year. He was a constant source of motivation and his enthusiasm and knowledge have added greatly to my own. I am indebted to Dr. Richard Woollons who provided the proposal for this thesis and supplied the data. I am grateful to both Graham and Richard for their willingness to listen and respond to my many questions.

I gratefully acknowledge the financial support by the Teachers' Study Awards programme which allowed me some time from full-time employment to undertake this degree. My thanks also to Mr. Peter Gibson and the management team at Lytton High School for their commitment to the professional development of staff.

I would also like to acknowledge Mr. John Thompson, a colleague and friend over many years, who initiated my interest in continuing university study.

Finally, special thanks to my family for their understanding, encouragement and constant support.

Contents

Part I

Preliminaries

Chapter 1

Introduction

This thesis describes a search for the most appropriate model for predicting mean basal area (m^2/ha) of pinus radiata forestry stands[1] at ages between 30 and 40 years. The basal area (g) of a stem as defined in the *1986 Forestry Handbook* (New Zealand Institute of Foresters) is the cross-sectional area at breast height (1.4 metres), in units of square metres assuming the stem has a circular cross-section. That is, $g = \pi d^2/4000$ where d is the diameter, measured in centimetres. A typical plot of basal-area against age for a cultivated stand of pinus radiata forest is shown in Figure 1-1.

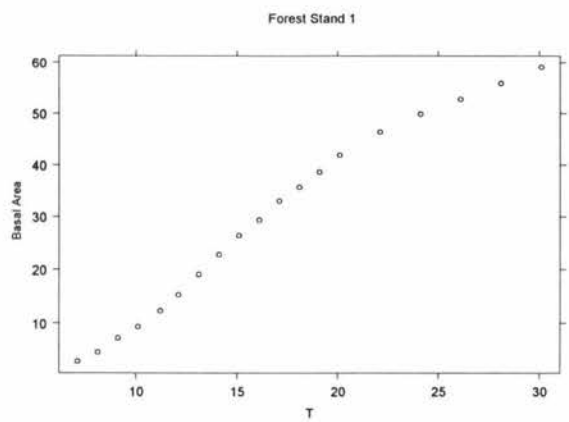


Figure 1-1: Mean basal area (m^2/hectare) against age for a pinus radiata stand from Kaingaroa Forest, New Zealand.

[1] Correctly, these are "plots" but in this thesis we will refer to forestry "stands" to avoid any possible confusion with the many references to graphics plots.

It is clear in this case that basal area will tend towards an asymptote but is not close to it at 28 years (the largest observation plotted). Note also that there is a change in concavity near $T = 14$ giving the data a sigmoid shape. Listed in Table 1-1are seven sigmoid growth curves commonly appearing in forestry modelling. These are the growth curves to be considered in this thesis.

Table 1-1: The seven growth curves to be examined in this thesis.

three-parameter models	
$c(T) = a(1 - e^{bT})^c$	Chapman-Richards/Bertalanffy
$g(T) = ae^{-be^{-cT}}$	Gompertz
$h(T) = \frac{aT^c}{ab + T^c}$	Hossfeld
$s(T) = e^{a - \frac{b}{T^c}}$	Schumacher/log-reciprocal/Lundquist/Korf
$w(T) = a(1 - e^{bT^c})$	Weibull
four-parameter models	
$L(T) = a\left(\frac{T^d}{b+T^d}\right)^c$	Levakovic
$S(T) = ae^{-be^{-cT^d}}$	Sloboda

Kiviste [9] gives examples of seventy five candidate functions but in general these are variations of the seven sigmoid functions listed in Table 1-1. There are similarities among the models with the Levakovic function, with $c = 1$, reducing to a form closely associated with the Hossfeld function. The Weibull function is very similar to the Chapman-Richards function; in the former the power c is associated with T rather than with the factor $(1 - e^{bT})$ in the latter.

There appears to be some confusion about the functional form of the Chapman-Richards model in the literature. It is important to note that the Chapman-Richards model used in this thesis is not the four-parameter Richards model ($y = a(1 - \exp(b - cT))^{1/d}$) that Ratkowsky [2] found to have quite unacceptable intrinsic curvature. It is noted that Zeide [6] refers to this unacceptable intrinsic curvature (page 600) when in fact the model under discussion in his

paper is the Chapman-Richards.

With $d = 1$, the 4-parameter Sloboda function reduces to the Gompertz function. As explained by Zeide [6], these apparent similarities of form can be deceptive and in Section 2.5 this is discussed further in terms of the differential forms of the models.

Additional to this introduction, in Part I properties of the curves are investigated with particular attention given to the relationship between the parameters, the point of inflection and the asymptote. We are particularly interested in which growth curve has the largest asymptote and investigate the ratio of the asymptote level to the y-coordinate of the point of inflection to gain a preliminary understanding of the ordering of asymptote levels among models.

In Part II, stable parameterizations of the curves are developed and fitted to basal area data (courtesy of Dr. R. C. Woollons, Canterbury University) from 22[2] stands of radiata pine. Goodness-of-fit of the growth curves to the data, estimates of the point of inflection and asymptote estimates are compared.

In Part III, the dependence of asymptote level on the slope at the point of inflection is discussed and a hierarchy of asymptote level among models is empirically determined.

In Part IV, the models are refitted to a subset of the data and the accuracy of predicting known basal area is compared between models. Additional data (again courtesy of Dr. R. C. Woollons), with basal area available for ages greater than 40 years in some cases, are also examined with a discussion of the models' comparative predictive precision the main focus.

Also, simulation studies are carried out to empirically determine which curves are superior in predicting basal area at post-data ages. For these simulations we fix properties at the point of inflection and/or the asymptote and use each model to generate data with these properties. Each model is then fitted to the data and predictions of basal area are made at post-data ages. The prediction errors of the fitted models are analysed and compared.

In an attempt to examine all models with a standard basis of comparison and determine features of the curves that will explain their goodness-of-fit and predictive power, Padé rational approximations are developed and analysed in Part V.

[2]There were data from 28 forestry stands in the original set. Two stands had periods of negative growth rate due to mortality (wind damage). Four stands had basal area measured at different times from the remaining plots and these data were omitted for simplicity of analysis and standardised presentation of graphs and tables.

In Part VI , a summary of results is given and an index is provided for the appropriateness of the sigmoid models for estimating basal area at forest ages between 30 and 40 years.

Chapter 2

Growth Curve Properties

2.1 Introduction

Important characteristics of each growth curve are the point of inflection, the slope of the curve at the point of inflection and the asymptote. Each characteristic is now expressed using the growth curve parameters. Illustrative functions for each model are compared in Section 2.3 and the mathematical results are summarised graphically and in tables in Section 2.4.

2.2 Characteristics of the growth curves

The Chapman-Richards/Bertalanffy function

The Chapman-Richards function $c(T) = a(1 - e^{bT})^c$ requires $b < 0$ for an asymptote to exist. If $b = 0$ the function is constant and for $b > 0$ we find $c(T)$ will be undefined for most values of the parameter c. The asymptote is determined by

$$\lim_{T\to\infty} \ln c\,(T) = \ln a\ (b < 0), \quad \text{whence } \lim_{T\to\infty} c\,(T) = a$$

The first and second derivatives are calculated as

$$\begin{aligned} c'\,(T) &= -abc\left(1 - e^{bT}\right)^{c-1} e^{bT} \\ c''\,(T) &= ab^2c\left(1 - e^{bT}\right)^{c-2} e^{bT}\left(ce^{bT} - 1\right) \end{aligned}$$

and the point of inflection is found as the solution to

$$ab^2c\left(1-e^{bT}\right)^{c-2}e^{bT}\left[ce^{bT}-1\right]=0, \quad \text{or } T=\frac{-\ln c}{b}$$

Clearly, for the inflection point to exist at $T>0$ it is required that $c>1$. As for the Schumacher function, the inflection point is independent of a. The slope at the point of inflection is

$$c'\left(\frac{-\ln c}{b}\right)=-ab\left(\frac{c-1}{c}\right)^{c-1}$$

The Gompertz function

Firstly it is noted that the Gompertz function $g(T)=ae^{-be^{-cT}}$ has $g(0)=ae^{-b}$. In the case that $c<0$ and $b>0$, $\lim_{T\to\infty}g(T)=0$; for $c<0$ and $b<0$, $\lim_{T\to\infty}g(T)=\infty$. Also, when $c>0$ and $b<0$ the function is decreasing. Therefore, to ensure growth curve properties it is clear that we must have $c>0$ and $b>0$. Thus

$$\begin{aligned}\lim_{T\to\infty}\ln g(T) &= \lim_{T\to\infty}\ln a+\lim_{T\to\infty}\left[-be^{-cT}\right]=\ln a\\ \text{whence } \lim_{T\to\infty}g(T) &= a\end{aligned}$$

The first and second derivatives are calculated as

$$\begin{aligned}g'(T) &= abce^{-cT-be^{-cT}}\\ g''(T) &= abc^2\left(-1+be^{-cT}\right)e^{-cT-be^{-cT}}\end{aligned}$$

and the point of inflection is found as a solution to

$$-1+be^{-cT}=0, \quad \text{or } T=\frac{\ln b}{c}$$

Clearly, if the inflection point is located at $T>0$ then $b>1$. It is interesting to note that the y-coordinate of the point of inflection is determined by the asymptote level. That is, both the

asymptote and inflection point cannot be specified independently, as

$$g(\frac{\ln b}{c}) = \frac{a}{e}$$

Both b and c determine the x-coordinate of the inflection point but its y-coordinate depends only on a (the asymptote).

The slope of the curve at the point of inflection is calculated as

$$g'\left[\frac{\ln b}{c}\right] = abce^{-\ln b - be^{-\ln b}} = \frac{ac}{e}$$

The Hossfeld function

The Hossfeld function $h(T) = \dfrac{aT^c}{ab + T^c}$ is increasing only for $c > 0$, so this condition is always assumed. The asymptote is determined by

$$\lim_{T\to\infty} h(T) = \lim_{T\to\infty} \frac{aT^c}{ab + T^c} = \lim_{T\to\infty} \frac{a}{\frac{ab}{T^c} + 1} = a, \text{ since } c > 0$$

The first and second derivatives are calculated as

$$\begin{aligned} h'(T) &= \frac{a^2 bcT^{c-1}}{(ab + T^c)^2} \\ h''(T) &= \frac{a^2 bcT^{c-2}\left[ab\left(c-1\right) - T^c\left(c+1\right)\right]}{(ab + T^c)^3} \end{aligned}$$

and the point of inflection is found as the solution to

$$ab\left(c-1\right) - T^c\left(c+1\right) = 0, \quad \text{or } T = \left(\frac{ab\left(c-1\right)}{c+1}\right)^{\frac{1}{c}}$$

For an inflection point located at $T > 0$ we require either that $b > 0$ and $c > 1$ or that $0 < c < 1$ and $b < 0$. It is of some interest that the x-coordinate of the point of inflection depends on a. The slope of the curve at the point of inflection is

$$h'\left[\left(\frac{ab\left(c-1\right)}{c+1}\right)^{\frac{1}{c}}\right] = \frac{\left[a\left(c-1\right)\right]^{\frac{c-1}{c}}\left(c+1\right)^{\frac{c+1}{c}} b^{-\frac{1}{c}}}{4c}$$

The Schumacher/log-reciprocal/Lundquist/Korf function

The Schumacher function $s(T) = e^{a-\frac{b}{T^c}}$ is increasing for $c > 0$, so again this condition is always assumed An obvious anomaly is that this function is undefined at $T = 0$ for $c > 0$. The asymptote is e^a since $\lim_{T\to\infty} \frac{b}{T^c} = 0$.

The first and second derivatives are determined as

$$\begin{aligned} s'(T) &= bT^{-c-1}ce^{a-bT^{-c}} \\ s''(T) &= bcT^{-c-2}e^{a-bT^{-c}}\left(bcT^{-c} - c - 1\right) \end{aligned}$$

and the point of inflection is found as the solution to

$$bcT^{-c} - c - 1 = 0, \quad \text{or } T = \left(\frac{bc}{c+1}\right)^{\frac{1}{c}}$$

For an inflection point located at $T > 0$ we require either that $b > 0$. The point of inflection is here independent of a. The slope of the growth curve at the point of inflection is given by

$$s'\left(\left(\frac{bc}{c+1}\right)^{\frac{1}{c}}\right) = \left(\frac{c+1}{bc}\right)^{\frac{1}{c}}(c+1)\,e^{\frac{ac-c-1}{c}}$$

The Weibull function

The Weibull function $w(T) = a\left(1 - e^{bT^c}\right)$ requires $c > 0$ and $b < 0$ for a growth curve to be sensibly defined, since it is required that $e^{bT^c} \to 0$ as $T \to \infty$. The asymptote is a, since $e^{bT^c} \to 0$ as $T \to \infty$.

The first and second derivatives are calculated as

$$\begin{aligned} w'(T) &= -abcT^{c-1}e^{bT^c} \\ w''(T) &= -abcT^{c-2}e^{bT^c}\left[c - 1 + bcT^c\right] \end{aligned}$$

and the point of inflection is found as the solution to

$$c-1+bcT^c=0, \quad \text{or } T=\left(\frac{1-c}{bc}\right)^{\frac{1}{c}}$$

For an inflection point located at $T>0$ we require $b<0$ and $c>1$. The slope at the point of inflection is

$$w'\left[\left(\frac{1-c}{bc}\right)^{\frac{1}{c}}\right]=-a\,(bc)^{\frac{1}{c}}\,(1-c)^{\frac{c-1}{c}}\,e^{\frac{1-c}{c}}$$

The Levakovic function

The Levakovic function $L(T)=a\left(\frac{T^d}{b+T^d}\right)^c$ is increasing for $d>0$ and its asymptote is determined in the following way

$$\begin{aligned}\lim_{T\to\infty}\ln L(T) &= \ln a+c\ln\left[\lim_{T\to\infty}\frac{1}{\frac{b}{T^d}+1}\right]=\ln a\\ \text{whence } \lim_{T\to\infty}L(T) &= a,\end{aligned}$$

or the asymptote is a.

The first and second derivatives are calculated as

$$\begin{aligned}L'(T) &= \frac{abcdT^{dc-1}}{(b+T^d)^{c+1}}\\ L''(T) &= \frac{abcdT^{dc-2}}{(b+T^d)^{c+2}}\left[bcd-b-T^d-dT^d\right]\end{aligned}$$

and the point of inflection is found as a solution to

$$bcd-b-T^d-dT^d=0, \quad \text{or } T=\left(\frac{b\,(cd-1)}{1+d}\right)^{\frac{1}{d}}$$

For an inflection point located at $T>0$ we require $b>0$, $c>0$ and $cd>1$. The slope of the curve at this point is calculated as

$$L'\left[\left(\frac{b\,(cd-1)}{1+d}\right)^{\frac{1}{d}}\right]=\frac{ac\left[(cd-1)\right]^{\frac{1}{d}(dc-1)}}{d^c\,(c+1)^{c+1}}\cdot(1+d)^{\frac{1+d}{d}}\cdot b^{-\frac{1}{d}}$$

The Sloboda function

The Sloboda function $S(T) = ae^{-be^{-cT^d}}$ is a generalization of the Gompertz function with the addition of d as the power of T. For an asymptote to exist there are two possibilities. Firstly, if $d > 0$ and $c > 0$ then

$$\begin{aligned} \lim_{T\to\infty} \ln S(T) &= \ln a - b \lim_{T\to\infty} e^{-cT^d} = \ln a \\ \text{whence } \lim_{T\to\infty} S(T) &= a \end{aligned}$$

and secondly, if $d < 0$ then

$$\begin{aligned} \lim_{T\to\infty} \ln S(T) &= \ln a - b \lim_{T\to\infty} e^{-cT^d} = \ln a - b \ (d < 0) \\ \text{whence } \lim_{T\to\infty} S(T) &= ae^{-b} \end{aligned}$$

Thus the asymptote can be generated in two ways depending on the sign of the parameter d. If $d > 0$ it is required that $c > 0$ for the asymptote to be defined and then it has level a. If $d < 0$ the asymptote is ae^{-b} and the function is increasing if $b > 0$ and $c < 0$ or if $b < 0$ and $c > 0$. We will see that the Sloboda model has $d < 0$ when this equation is fitted to pinus radiata data. Note that in the case $d < 0$, $S(T)$ is undefined at $T = 0$.

The first and second derivatives are calculated as

$$\begin{aligned} S'(T) &= abcdT^{d-1}e^{-cT^d - be^{-cT^d}} \\ S''(T) &= abcdT^{d-2}e^{-cT^d - be^{-cT^d}} \left[d - 1 - cdT^d + bcdT^d e^{-cT^d}\right] \end{aligned}$$

and the point of inflection is found as the solution to $d - 1 - cdT^d + bcdT^d e^{-cT^d} = 0$. This cannot be solved explicitly for T, but a reasonable approximation may be made by noting that e^{-cT^d} is close to one when T is reasonably large with $d < 0$ and $c > 0$. Therefore, an approximation for the inflection point can be made as

$$T = \left(\frac{d-1}{cd(1-b)}\right)^{\frac{1}{d}}$$

and the slope of the curve at the point of inflection is given approximately by

$$S'\left[\left(\frac{d-1}{cd(1-b)}\right)^{\frac{1}{d}}\right] = abcd\left(\frac{d-1}{cd(1-b)}\right)^{\frac{1}{d}(d-1)} e^{\frac{d-1}{d(-1+b)}} - be^{\frac{d-1}{d(-1+b)}}$$

We will use these approximations in Chapter 4, where the Sloboda model's goodness-of-fit is examined, to estimate the location of the point of inflection and the slope of the curve at the point of inflection.

2.3 An illustrated comparison of the growth curves

Using the properties of the growth curves developed in the preceding section, we determine a formula for each model with the property that the slope of the curve is 2 at the point of inflection $(5, 5)$. This information is not sufficient to determine all parameters of the four-parameter models. For the Levakovic model d is fixed as two and for the Sloboda model c is fixed as two. The formulae, asymptote levels and ratios of the asymptote level (y_{asy}) to the y-coordinate of the point of inflection (y_{inf}) are summarised in Table 2-1 and the curves graphed in Figure 2-1 where the key is presented in descending order of asymptote level. Of interest is the asymptote level of the Schumacher model which is almost double the next largest (the Levakovic model) and is more than three times that of the Weibull model. This supports Dr. Woollons' observation that the Schumacher model consistently gives a larger asymptote than the other models.

Table 2-1: Growth curves with slope 2 at the inflection point $(5, 5)$.

Growth Curve	Formula	Asymptote	$\frac{y_{\text{asy}}}{y_{\text{inf}}}$
Chapman	$15.29\left(1 - e^{-.3187T}\right)^{4.9216}$	15.3	3.1
Gompertz	$13.59e^{-7.4e^{-.4T}}$	13.6	2.7
Hossfeld	$\frac{15T^3}{250 + T^3}$	15.0	3.0
Schumacher	$36.945e^{-\frac{10}{T}}$	36.9	7.3
Weibull	$10.53\left(1 - e^{-.007\,07T^{2.8074}}\right)$	10.5	2.1
Levakovic	$20\left(\frac{T^2}{25+T^2}\right)^2$	20.0	4.0
Sloboda	$.001\,536e^{9.45342e^{-2T^{-1.5850}}}$	19.6	3.9

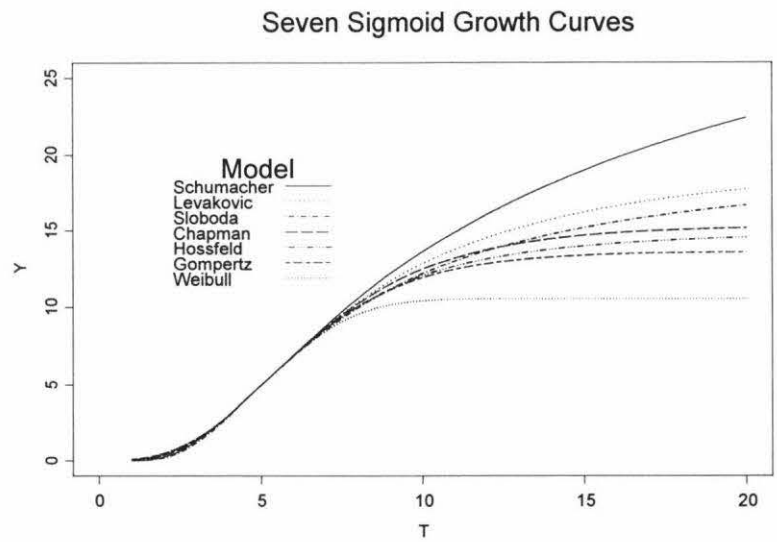


Figure 2-1: A comparison of the growth curves with slope 2 at a point of inflection of $(5, 5)$

The ratio of the asymptote and the y-coordinate of the point of inflection is also given in Table 2-1. This suggests that for a fixed point of inflection with fixed slope, the Weibull function will produce an asymptote markedly smaller than the other sigmoid functions, whereas the Schumacher function will have an asymptote much larger than even the next largest (the Levakovic function). These observations will be explored in greater depth in the next section.

2.4 Summary of inflection point and asymptote properties

Table 2-2 summarises the inflection point and asymptote properties of each sigmoid growth function. In particular, note should be made of the ratio of the level of the asymptote to the y-coordinate of the inflection point. For the Gompertz function this is constant and for the other 3-parameter functions the ratio is only dependent on the parameter c.

Table 2-2: The asymptote, the inflection point, and the ratio of asymptote to y-coordinate of inflection point for each sigmoid growth curve.

Function	General Formula	Asy.	Restr.	Inflection (T)	Inflection (Y)	$\frac{y_{asy}}{y_{inf}}$
Chapman	$c(T) = a\left(1 - e^{bT}\right)^c$	a	$b < 0$	$\frac{-\ln c}{b}$	$a\left(c-1\right)^c c^{-c}$	$\frac{c^c}{(c-1)^c}$
Gompertz	$g(T) = ae^{-be^{-cT}}$	a	$b > 0$ $c > 0$	$\frac{\ln b}{c}$	ae^{-1}	e
Hossfeld	$h(T) = \frac{aT^c}{ab+T^c}$	a	$c > 0$	$\left(\frac{ab(c-1)}{c+1}\right)^{\frac{1}{c}}$	$\frac{1}{2}\left(c-1\right)\frac{a}{c}$	$\frac{2c}{c-1}$
Schumacher	$s(T) = e^{a-\frac{b}{T^c}}$	e^a	$c > 0$	$\left(\frac{bc}{c+1}\right)^{\frac{1}{c}}$	$e^{\frac{ac-c-1}{c}}$	$e^{\frac{1}{c}(c+1)}$
Weibull	$w(T) = a\left(1 - e^{bT^c}\right)$	a	$c > 0$ $b < 0$	$\left(\frac{1-c}{bc}\right)^{\frac{1}{c}}$	$a\left(1 - e^{-\frac{c-1}{c}}\right)$	$\frac{1}{1-e^{-\frac{c-1}{c}}}$
Levakovic	$L(T) = a\left(\frac{T^d}{b+T^d}\right)^c$	a	$d > 0$	$\left(\frac{b(cd-1)}{1+d}\right)^{\frac{1}{d}}$	$\frac{a(cd-1)^c d^{-c}}{(c+1)^c}$	$\frac{d^c(c+1)^c}{(cd-1)^c}$
Sloboda	$S(T) = ae^{-be^{-cT^d}}$	ae^{-b}	$d < 0$	$\left(\frac{d-1}{cd(1-b)}\right)^{\frac{1}{d}}$	$ae^{-be^{\frac{d-1}{d(-1+b)}}}$	$e^{-b+be^{\frac{d-1}{d(-1+b)}}}$

For the 3-parameter models, a plot of the ratio of the asymptote to the y-coordinate of the point of inflection (as functions of c) is shown in Figure 2-2. It is of interest that for c greater than about 1.8, the Schumacher function gives the largest ratio. The Weibull function gives the smallest ratio in this regard. However, when their c parameter is small the Weibull, Hossfeld and Chapman-Richards functions will all have large ratio of asymptote to inflection ordinate.

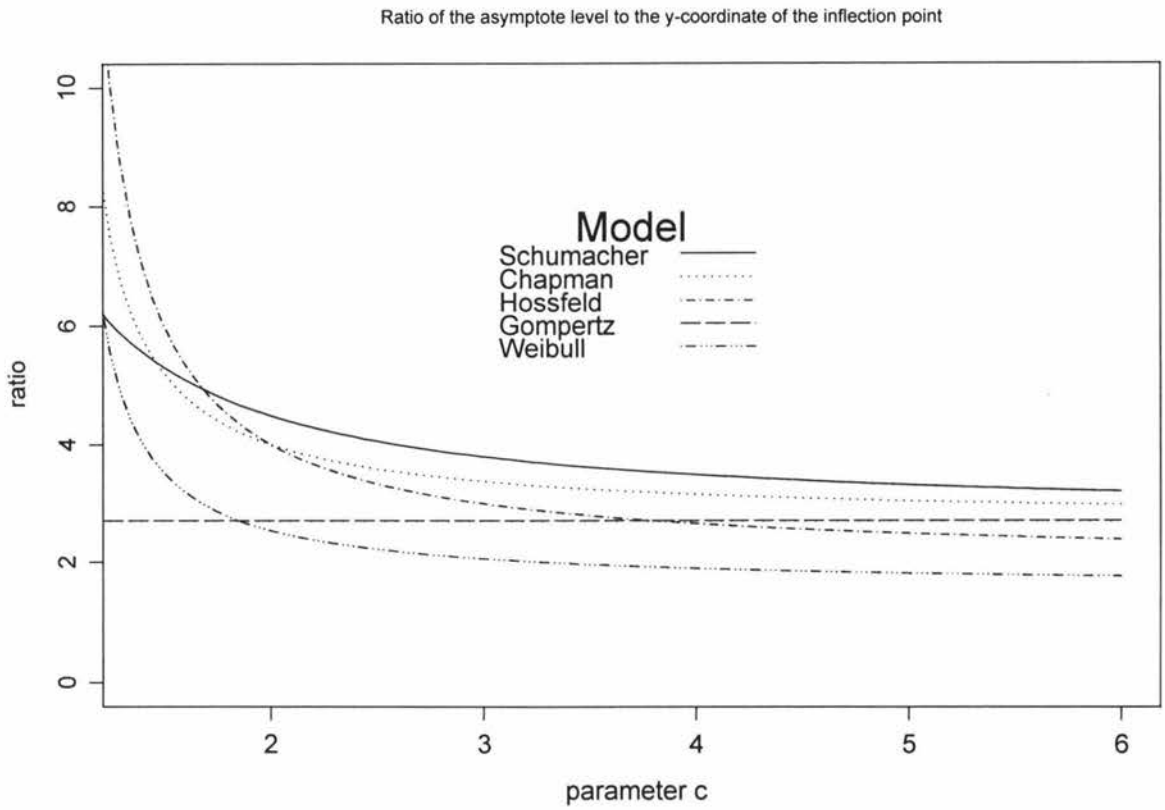


Figure 2-2: The ratio of asymptote to y-coordinate of inflection point plotted against the parameter c for the five 3-parameter growth curves.

A discussion of the fitting of the sigmoid models to pinus radiata forestry data is presented in Chapter 4. However, it will be of interest to preempt this discussion by considering the range of values obtained for the parameter c when the 3-parameter models are fitted to 22 stands of basal area data. The values estimated for parameter c are (refer to Appendix B for a full listing of estimated parameters):

c	Minimum	Maximum
Schumacher	1.1	1.6
Hossfeld	2.7	3.3
Chapman	4.1	6.4
Weibull	2.2	2.7

We use these values of c, to calculate upper and lower limits for $\frac{Y_{asy}}{Y_{\text{inf}}}$ and graph these in Figure 2-3. Clearly, the Schumacher function always has the largest ratio although it is more variable than the other models. The Hossfeld and Chapman-Richards ratios overlap and are next largest. The Gompertz ratio, which is independent of the parameters and equal to e, is always less than those of the Hossfeld and the Chapman-Richards models. The Weibull model clearly has the smallest ratio. This gives the following ordering of $\frac{Y_{asy}}{Y_{\text{inf}}}$ in from largest to smallest.

Schumacher
Chapman-Richards and Hossfeld
Gompertz
Weibull

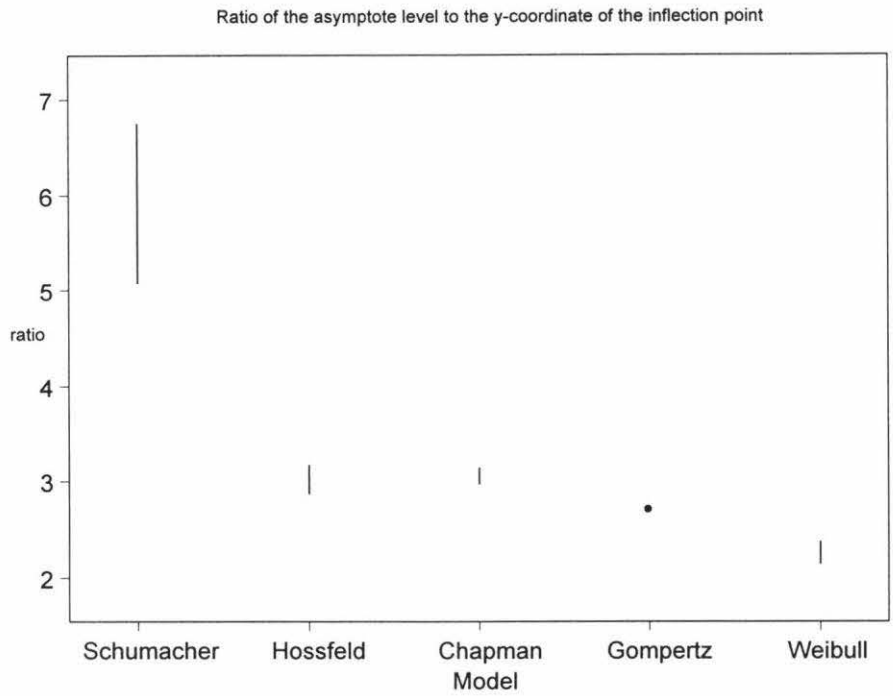


Figure 2-3: Ratio of the asymptote level to the y-coordinate of the inflection point for the 3-parameter growth curves using values of the parameter c that are observed in modelling forestry data.

For the two 4-parameter models, the ratio of the asymptote to the y-coordinate of the point of inflection is shown below in Figures 2-4 and 2-5.

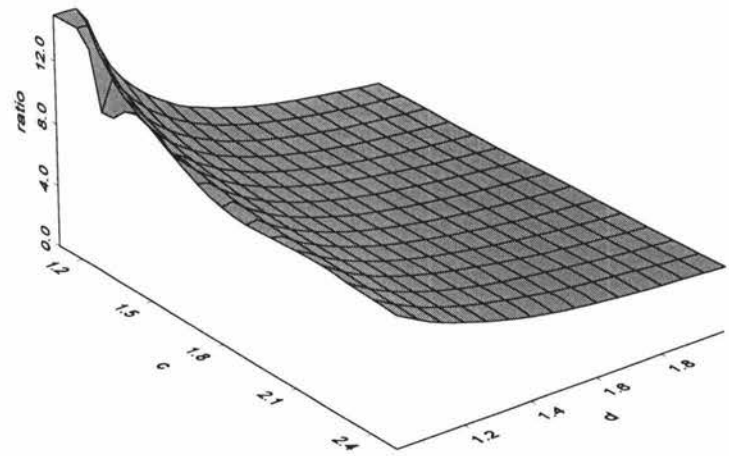


Figure 2-4: The ratio of y_{asy} to y_{inf} for the Levakovic function

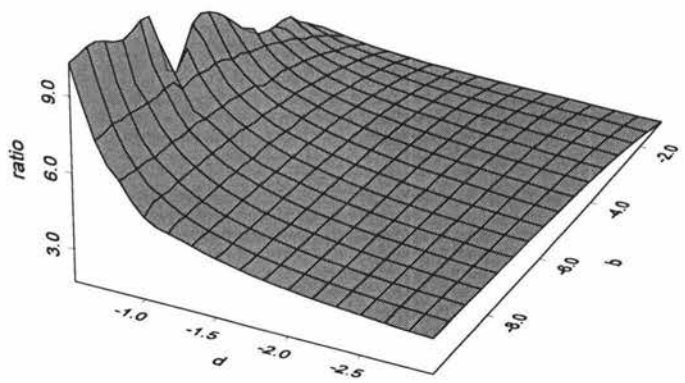


Figure 2-5: The ratio of y_{asy} to y_{inf} for the Sloboda function

It is quite apparent that a reasonably large ratio of asymptote to inflection point is only achieved for a relatively small range of values of the parameters. In Chapter 5, the asymptote levels of the growth curves are investigated in more detail. In particular, the dependence of asymptote level on slope at the inflection point is examined by simulating data using parameters that commonly occur in modelling forestry data.

2.5 Differential equations for the growth curves

Introduction

This section is included for completeness and summarises aspects of the paper by Professor Boris Zeide [6]. He discusses the process of growth and the structure of the sigmoid curves by considering them in differential form. In this form, the instantaneous growth rate, y', is related to the age of the tree (or stand). Zeide describes how the differential equation can be decomposed into components of growth and decline through both subtraction and division. The derivation of the differential forms and decomposition into growth and decline components is shown in the following sections.

The Chapman-Richards function

The Chapman-Richards function $y = c(T) = a\left(1 - e^{bT}\right)^c$ has first derivative $y' = -a\left(1 - e^{bT}\right)^{c-1} cbe^{bT}$. Noting that $1 - e^{bT} = \left(\frac{y}{a}\right)^{\frac{1}{c}}$ and substituting this into the differential equation gives

$$y' = -a\left(\left(\frac{y}{a}\right)^{\frac{1}{c}}\right)^{c-1} cb\left(1 - \left(\frac{y}{a}\right)^{\frac{1}{c}}\right) = bcy - a^{\frac{1}{c}}bcy^{\frac{1}{c}(c-1)}$$

With $a > 0$, $b < 0$ and $c > 0$, $bcy - a^{\frac{1}{c}}bcy^{\frac{1}{c}(c-1)}$ consists of two components with the first driving a decrease in y and the second driving an increase in y. Zeide refers to these as the subtractive components of growth and decline. The division components are available from the differential equation. Using $1 - e^{bT} = \left(\frac{y}{a}\right)^{\frac{1}{c}}$ we obtain

$$y' = -a\left(1 - e^{bT}\right)^{c-1} cbe^{bT} = \frac{-a^{\frac{1}{c}}bcy^{\frac{1}{c}(c-1)}}{e^{-bT}}$$

which has growth and decline components $-a^{\frac{1}{c}}bcy^{\frac{1}{c}(c-1)}$ and e^{-bT}.

The Gompertz function

The Gompertz function $y = g(T) = ae^{-be^{-cT}}$ has first derivative $y' = abce^{-cT}e^{-be^{-cT}}$. Noting that $e^{-cT} = \frac{\ln\left(\frac{y}{a}\right)}{-b}$ and $e^{-be^{-cT}} = \frac{y}{a}$ and substituting these into the differential equation gives

$$y' = bc\frac{\ln\left(\frac{y}{a}\right)}{-b}y = c\left(\ln a\right)y - cy\ln y$$

With $a > 0$, $b > 0$ and $c > 0$ the growth component is $c\left(\ln a\right)y$ and the decline component is $cy\ln y$. The division components can be obtained by substituting $e^{-be^{-cT}} = \frac{y}{a}$ into the differential equation which gives

$$y' = abce^{-cT}\frac{y}{a} = \frac{bcy}{e^{cT}}$$

showing growth and decline components of bcy and e^{cT}.

The Hossfeld Function

The first derivative of the Hossfeld function $y = h(T) = \frac{aT^c}{ab+T^c}$ is determined as $y' = \frac{a^2bcT^{c-1}}{(ab+T^c)^2}$ and simplified by noting that $T^c = \frac{aby}{a-y}$ which gives

$$y' = \frac{a^2bc\frac{aby}{a-y}}{T\left(ab + \frac{aby}{a-y}\right)^2} = \frac{c}{T}y - \frac{c}{aT}y^2$$

This contains two components: $\frac{c}{T}y$ which is the growth component since $c > 0$ and $-\frac{c}{aT}y^2$ which is the decline component since $a > 0$. The division components are found by recognising that $ab + T^c = \frac{aT^c}{y}$ and we substitute this into the differential equation giving

$$y' = \frac{a^2bcT^{c-1}}{\left(\frac{aT^c}{y}\right)^2} = \frac{bcy^2}{T^{c+1}}$$

which has growth and decline components bcy^2 and T^{c+1}.

The Schumacher function

The Schumacher function $y = s(T) = e^{a-\frac{b}{T^c}}$ has first derivative $y' = bcT^{-c-1}e^{a-bT^{-c}} = bcT^{-c-1}y$ and this is simplified by noting that $T^c = \frac{b}{a-\ln y}$. We substitute this into the differential equation giving

$$y' = \frac{bc}{\frac{bT}{a-\ln y}}y = \frac{ac}{T}y - \frac{c}{T}y\ln y$$

Since $a > 0$ and $c > 0$ the subtraction components of growth and decline are $\frac{ac}{T}y$ and $\frac{c}{T}y\ln y$. The division components are immediately available from the differential equation

$$y' = bcT^{-c-1}y = \frac{bcy}{T^{c+1}}$$

which has growth and decline components bcy and T^{c+1}.

The Weibull function

The Weibull function $y = w(T) = a\left(1 - e^{bT^c}\right)$ has first derivative $y' = -abT^{c-1}ce^{bT^c}$. Noting that $e^{bT^c} = \left(1 - \frac{y}{a}\right)$and substituting this into the differential form gives

$$y' = -abT^{c-1}c\left(1 - \frac{y}{a}\right) = bcT^{c-1}y - abcT^{c-1}$$

Since $a > 0, b < 0$ and $c > 0$ the subtraction components of growth and decline are $abcT^{c-1}$ and $bcT^{c-1}y$. The division components are immediately available from the differential equation

$$y' = \frac{-abcT^{c-1}}{e^{-bT^c}}$$

which has growth and decline components $-abcT^{c-1}$ and e^{-bT^c}.

The Levakovic function

The Levakovic function $y = L(T) = a\left(\frac{T^d}{b+T^d}\right)^c$has first derivative $y' = abcd\left(\frac{T^d}{b+T^d}\right)^{c-1}\frac{T^{d-1}}{\left(b+T^d\right)^2}$ and noting that $\frac{T^d}{b+T^d} = \left(\frac{y}{a}\right)^{\frac{1}{c}}$ the differential form is simplified to $y' = \frac{1}{T}\frac{1}{b+T^d}bcdy$. Now using

$T^d = \frac{b\left(\frac{y}{a}\right)^{\frac{1}{c}}}{1-\left(\frac{y}{a}\right)^{\frac{1}{c}}}$ and substituting this into the differential form gives

$$y' = bcdy\frac{1}{T}\frac{1}{b+\frac{b\left(\frac{y}{a}\right)^{\frac{1}{c}}}{1-\left(\frac{y}{a}\right)^{\frac{1}{c}}}} = \frac{cd}{T}y - \frac{cd}{T}y\left(\frac{y}{a}\right)^{\frac{1}{c}}$$

With $a > 0$, $c > 0$ and $d > 0$ the growth component is $\frac{cd}{T}y$ and the decline component is $\frac{cd}{T}y\left(\frac{y}{a}\right)^{\frac{1}{c}}$. The division components are obtained by substituting $b + T^d = T^d\left(\frac{y}{a}\right)^{\frac{-1}{c}}$ into the differential equation giving

$$y' = \frac{1}{T}\frac{1}{T^d\left(\frac{y}{a}\right)^{\frac{-1}{c}}}bcdy = \frac{a^{\frac{-1}{c}}bcdy^{\frac{c+1}{c}}}{T^{d+1}}$$

which has growth and decline components $a^{\frac{-1}{c}}bcdy^{\frac{c+1}{c}}$ and T^{d+1}.

The Sloboda function

The Sloboda function $y = S(T) = ae^{-be^{-cT^d}}$ has first derivative $y' = bcdT^{d-1}e^{-cT^d}y$. Noting that $\frac{\ln\frac{y}{a}}{-b} = e^{-cT^d}$ and substituting this into the differential form gives

$$y' = bcdT^{d-1}\frac{\ln\frac{y}{a}}{-b}y = cdT^{d-1}\left(\ln a\right)y - cdT^{d-1}y\ln y$$

With $c > 0$ and $d < 0$ the growth component is $cdT^{d-1}y\ln y$ and decline component is $cdT^{d-1}\left(\ln a\right)y$. The division components are obtained immediately from the differential form

$$y' = bcdT^{d-1}e^{-cT^d}y = \frac{bcdy}{T^{1-d}e^{cT^d}}$$

which has growth and decline components $bcdy$ and $T^{1-d}e^{cT^d}$.

Discussion

The growth and decline components are summarised in Tables 2-3 and 2-4.

Table 2-3: Integral and differential forms of seven growth curves.

Growth Curve	Integral Form	Differential Equation
Chapman	$y = a\left(1 - e^{bT}\right)^c$	$y' = -a\left(1 - e^{bT}\right)^{c-1} cbe^{bT}$
Gompertz	$y = ae^{-be^{-cT}}$	$y' = abce^{-cT}e^{-be^{-cT}}$
Hossfeld	$y = \frac{aT^c}{ab+T^c}$	$y' = \frac{a^2bcT^{c-1}}{(ab+T^c)^2}$
Schumacher	$y = e^{a-\frac{b}{T^c}}$	$y' = bcT^{-c-1}e^{a-bT^{-c}}$
Weibull	$y = a\left(1 - e^{bT^c}\right)$	$y' = -abT^{c-1}ce^{bT^c}$
Levakovic	$y = a\left(\frac{T^d}{b+T^d}\right)^c$	$y' = \frac{1}{T}\frac{1}{b+T^d}bcdy$
Sloboda	$y = ae^{-be^{-cT^d}}$	$y' = bcdT^{d-1}e^{-cT^d}y$

Table 2-4: Subtraction and division components of growth and decline for seven growth curves.

	Subtraction Components		Division Components	
Growth Curve	Growth	Decline	Growth	Decline
Chapman	$a^{\frac{1}{c}}bcy^{\frac{1}{c}(c-1)}$	bcy	$-a^{\frac{1}{c}}bcy^{\frac{1}{c}(c-1)}$	e^{-bT}
Gompertz	$c(\ln a)y$	$cy\ln y$	bcy	e^{cT}
Hossfeld	$\frac{c}{T}y$	$\frac{c}{aT}y^2$	bcy^2	T^{c+1}
Schumacher	$\frac{ac}{T}y$	$\frac{c}{T}y\ln y$	bcy	T^{c+1}
Weibull	$abcT^{c-1}$	$bcT^{c-1}y$	$-abcT^{c-1}$	e^{-bT^c}
Levakovic	$\frac{cd}{T}y$	$\frac{cd}{T}y\left(\frac{y}{a}\right)^{\frac{1}{c}}$	$a^{\frac{-1}{c}}bcdy^{\frac{c+1}{c}}$	T^{d+1}
Sloboda	$cdT^{d-1}y\ln y$	$cdT^{d-1}(\ln a)y$	$bcdy$	$T^{1-d}e^{cT^d}$

Inspection of the subtractive growth and decline components indicates the similarity between the Hossfeld and Levakovic models but Zeide shows that classification can be achieved more descriptively by considering the logarithm of the division form of the differential equations as shown in Table 2-5.

Table 2-5: The logarithm of the growth curve differential equations (in division form) with a general form for each.

Growth Curve	Division Form	Logarithmic Form	General Form
Chapman	$\frac{-a^{\frac{1}{c}}bcy^{\frac{1}{c}(c-1)}}{e^{-bT}}$	$\ln\left(-a^{\frac{1}{c}}bc\right)+\frac{1}{c}(c-1)\ln y+bT$	$k+p\ln y+qT$
Gompertz	$\frac{bcy}{e^{cT}}$	$\ln bc+\ln y-cT$	$k+\ln y+qT$
Hossfeld	$\frac{bcy^2}{T^{c+1}}$	$\ln bc+2\ln y-(c+1)\ln T$	$k+2\ln y+q\ln T$
Schumacher	$\frac{bcy}{T^{c+1}}$	$\ln bc+\ln y-(c+1)\ln T$	$k+\ln y+q\ln T$
Weibull	$\frac{-abcT^{c-1}}{e^{-bT^c}}$	$\ln(-abc)+(c-1)\ln T-bT^c$	$k+p\ln T+qT^{p+1}$
Levakovic	$\frac{a^{\frac{-1}{c}}bcdy^{\frac{c+1}{c}}}{T^{d+1}}$	$\ln\left(a^{\frac{-1}{c}}bcd\right)+\frac{c+1}{c}\ln y-(d+1)\ln T$	$k+p\ln y+q\ln T$
Sloboda	$\frac{bcdy}{T^{1-d}e^{cT^d}}$	$\ln(bcd)+\ln y-(1-d)\ln T-cT^d$	$k+\ln y+(d-1)\ln T+qT^d$

The general forms indicate two main types. The first is

$$\ln y' = k + p\ln y + q\ln T$$

with the Hossfeld, Schumacher, and Levakovic functions of this type. The other form is

$$\ln y' = k + p\ln y + qT$$

and the Chapman-Richards and Gompertz functions are of this type.

The apparent similarity between the Levakovic and Hossfeld functions in integral (y) form is confirmed in this differential equation classification. It is notable that the Schumacher function, so different to the Levakovic and Hossfeld functions in integral form, has the same differential equation general form. The Weibull equation does not fit into either classification. Zeide notes the difference between the general forms of the Weibull and Chapman-Richards functions and the apparent similarity of these two models in their integral form. Further, the Sloboda function seems to be a combination of the two main forms.

It will be of interest in later sections whether the goodness-of-fit and predictive power of the growth curve models can be related to this differential equation classification.

Part II

Fitting the Growth Curve Models to Basal Area

Chapter 3

Preliminaries to Fitting the Models

3.1 Introduction

In this chapter we discuss the preliminary work required prior to the fitting of the 7 sigmoid model functions to basal area from 22 forestry stands of pinus radiata from Kaingaroa Forest, New Zealand. The data of a typical stand is shown in Figure 3-1.

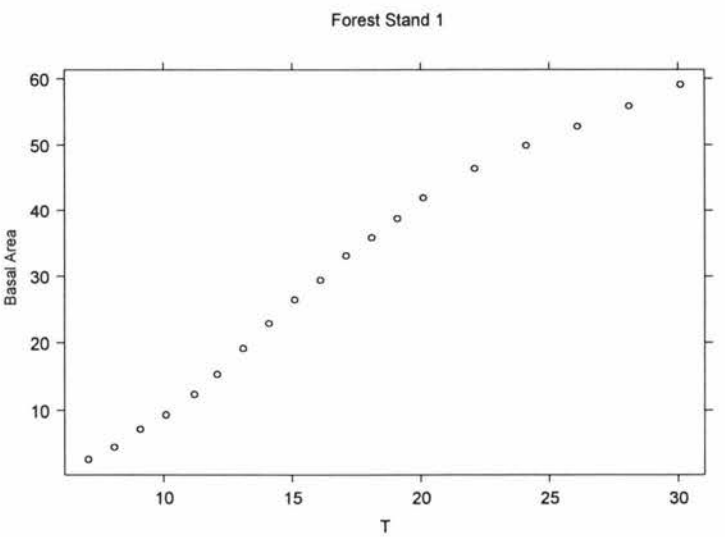


Figure 3-1: Mean basal area against age for a pinus radiata stand from Kaingaroa Forest, New Zealand.

Prior to fitting the models there are two important issues that need to be addressed. The

first is the choice of reasonable starting values to ensure convergence of the algorithm to fit the nonlinear models. The second is curvature which impacts on the precision of estimation from the fitted models. We will see that these two issues overlap and can both be addressed by the choice of a suitable reparameterization of the model function.

3.2 Curvature and the need for reparameterization

We will give a practical guide to the notion of curvature[1] and the need for reparameterization. Detailed, theoretical descriptions can be found in Beale [13], Ross [14], and Bates & Watts [4]. A concise, simple summary of the practical issues is taken from Bates & Watts:

> "...transforming the parameters can have the beneficial effect of making approximate inference regions better and speeding convergence to the least squares value.
>
> An important assumption used in the development of these methods is the assumption that the expectation surface is flat, so the tangent plane provides an accurate approximation."

The general form of a non-linear model is

$$y_n = f(x_n, \theta) + \epsilon$$

where f is the expectation function and x_n is the independent variable. In our study these are seven growth curve models and time in years respectively. The argument θ is a vector of unknown parameters and ϵ is a $N(0, \sigma^2)$ error term. The variance σ^2 is estimated as s^2, the residual mean square, as in linear regression. As θ is varied $f(x, \theta)$ traces a p-dimensional surface where p is the number of unknown parameters. This surface is referred to as the expectation surface.

[1] It is noted that two forms of curvature are defined: parametric-effects curvature and intrinsic curvature. Intrinsic curvature is inherent in the model and we have no control over it. Parametric-effects curvature can be controlled by the choice of parameterization. When the term curvature is used in this thesis it will refer to parametric-effects curvature.

Venables & Ripley [1] use a profile plot (generated from a fitted nonlinear model) to graphically assess the flatness of the expectation surface and an example plot is shown in Figure 3-2.

The profile plot for a parameter a is determined as follows. Suppose that the nonlinear model has residual sum of squares S and the estimate of parameter a is 58.4. This is indicated in Figure 3-2 with a cross at coordinates $(58.4, 0)$. We now choose values of $a^{'}$ in the range from 54 to 64 and these are the x-coordinates in the profile plot for a. For each value of $a^{'}$ the model is refitted and optimized with respect to the other parameters b and c and the residual sum of squares calculated as $S^{'}$. We determine $sign\left(a - a^{'}\right)\sqrt{S - S^{'}}/s$ as the y-coordinate τ of the profile plot. Therefore, tau is a scaled measure of the difference in goodness-of-fit between the optimum model (with $a = 58.4$) and a model with a' fixed at a different value.

If the profiles of each parameter are very nearly linear then the expectation surface will be flat and the precision of estimates will be good. It is clear that there is curvature mainly associated with parameters a and b in the profile plot of Figure 3-2.

Other than the use of the descriptive profile plot to detect curvature, a statistic denoted c^{θ} was suggested by Bates & Watts and implemented in S-Plus. If F is the $F_{p,n-p}$ critical value then it was suggested that a value of $c^{\theta}\sqrt{F} > .3$ indicates unacceptably high curvature. A more complete discussion of curvature, with a detailed example including calculation of $c^{\theta}\sqrt{F}$, is contained in Appendix F.

Reparameterization can reduce the curvature in the fitted model. It involves the replacement of one or more parameters with expected value parameters. This is best described by example. Suppose we fit a nonlinear model f to the data shown in Figure 3-1 and the profile plots show curvature associated with say, parameters a and b as in the profile plot of the model displayed in Figure 3-2. We choose two points that are on the curve in Figure 3-1 (reasonably close by eye is adequate) and reasonably far apart. For our data the two points $(10, 10)$ and $(20, 40)$ will be used.

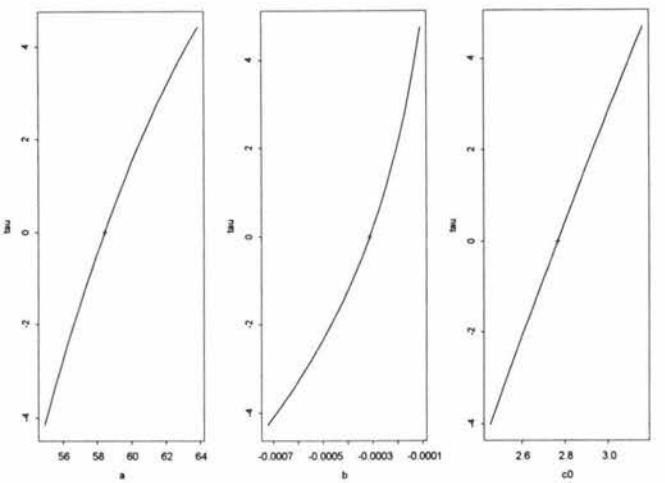


Figure 3-2: A profile plot for a nonlinear model with three parameters a, b and c. It appears that unacceptably high nonlinearity is associated with parameter b.

We now define two expected value parameters

$$\theta_0 = f(10) \text{ and } \theta_1 = f(20)$$

and solve these equations for a and b. That is,

$$a = g_1(\theta_0, \theta_1) \text{ and } b = g_2(\theta_0, \theta_1)$$

and we replace a and b in the original model f with g_1 and g_2. In addition to reducing the curvature associated with a and b we immediately have starting values for two parameters. Namely,

$$\theta_0 = 10 \text{ and } \theta_1 = 40$$

In summary, reparameterization has these advantages:

1. It can speed convergence.

2. Reasonable starting values can be more easily chosen.

3. The model will have better estimation properties. That is, the S-Plus output from **nls()** will show parameter estimates with small standard errors and large t-statistics.

For 4-parameter models it is very difficult to find a reparameterization that works well in all respects. A transformation of two parameters is usually possible but with any more the algebra gets very messy or intractable. Parameterizations that are developed in subsequent sections are in general chosen to reduce the nonlinear behaviour of one or more parameters. This general issue is discussed in some depth by Ratkowsky [3] but the full range of sigmoid curves studied in this thesis are not described in his work. The reparameterizations developed in Section 3-4 have in general been determined independently of Ratkowsky's work but citations have been included where appropriate.

3.3 Using S-Plus to fit nonlinear models

The S-Plus algorithm for nonlinear least squares can be implemented using both first and second derivatives. The Gauss-Newton algorithm is used with a step factor to ensure that the sum of squares decreases at each iteration. The derivatives are generated automatically in S-Plus through the use of the **deriv3()** function by David Smith and supplied as part of the Venables and Ripley *MASS S-Plus Library*. The second derivatives are required to implement the function **rms.curv()** (a function available in *MASS*) to assess curvature. A typical session (illustrated using the Gompertz function) to fit a nonlinear model using S-plus is

- The model is set up with first and second derivatives calculated:

```
gomp <- deriv3(~a*exp(-b*exp(-c0*x)), c(''a'',''b'',''c0''),
 function(a,b,c0,x) NULL)
```

- The model is fitted to the data for Stand 1:

```
gp1.nls <- nls(basal.area[plot==1]~gomp(a,b,c0,age[plot==1]),
 start=list(a=67,b=7,c0=.13), data=forest, trace=T)
```

- Curvature is determined:

```
gp1.curv <- rms.curv(gp1.nls)
```

- A profile plot is produced to assess the curvature:

```
plot(profile(gp1.curv)
```

3.4 Reparameterizations and starting values

Introduction

In the sections that follow initial values will be determined for fitting each model by nonlinear least squares. Determining starting values from properties of the data or transformation followed by linear least squares is often used. Curvature will be examined and reparameterizations (summarised in Appendix L) will be developed where necessary. To illustrate each model, the data from forestry stand 1, graphed in Figure 3-1, will be used.

It seems that a suitable first approximation of the asymptote is when basal area is 70, with point of inflection at about $(15, 20)$. These properties will be used to investigate suitable starting values.

The Chapman-Richards function

From Table 2-3, the Chapman-Richards function $y = c(T) = a\left(1 - e^{bT}\right)^c$ has asymptote a and point of inflection with coordinates $\left(\frac{-\ln c}{b}, a\left(\frac{c-1}{c}\right)^c\right)$. Also $\frac{y_{asy}}{y_{inf}} = \left(\frac{c}{c-1}\right)^c$. Estimating $a = 70$ and an inflection point located at $(15, 25)$ we have an approximation of $\frac{y_{asy}}{y_{inf}}$ as 3 which yields an initial estimate of c as $\left(\frac{c}{c-1}\right)^c = 3$, or $c = 5.7477065$. From the x-coordinate of the inflection point we have

$$\frac{-\ln 5.7477065}{b} = 15, \text{ or } b = -.11658673$$

Using $a = 70$, $b = -.12$ and $c = 5.7$ as starting values for fitting the Chapman-Richards function to the Stand 1 data gives the following edited S-plus output.

```
Parameters:
     Value       Std.  Error   t value
 a   65.392300   0.66125400    98.8914
 b   -0.133922   0.00300943   -44.5009
 c0   6.476840   0.24492800    26.4439
```

```
Residual standard error:  0.402701 on 16 degrees of freedom
Parameter effects:  c^theta x sqrt(F) = 0.3099
```

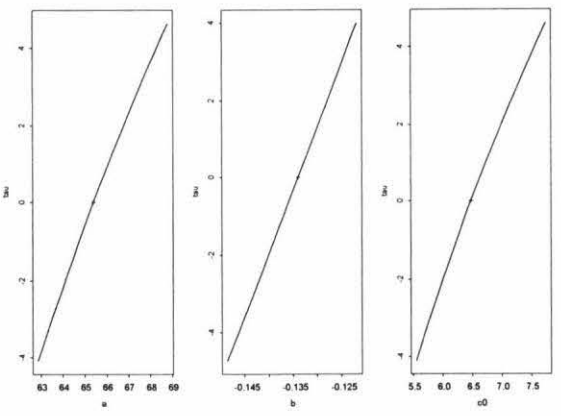


Figure 3-7: Profile plot for Chapman-Richards function fitted to Stand 1 data

Parametric-effects curvature is only just too large using the $c\sqrt{F} < 0.3$ criterion and appears to present mostly through non-linearity in parameters a and c. A stable parameterization is developed using expected values at $T = 10$ and $T = 20$. That is,

$$\theta_0 = a\left(1 - e^{10b}\right)^c \text{ and } \theta_1 = a\left(1 - e^{20b}\right)^c$$

We solve these two equations to give a and c in terms of θ_0, b and θ_1 giving

$$a = \frac{\theta_0}{\left(1 - e^{10b}\right)^{\frac{\ln \frac{\theta_0}{\theta_1}}{\ln \frac{1-e^{10b}}{1-e^{20b}}}}} \text{ and } c = \frac{\ln \frac{\theta_0}{\theta_1}}{\ln \frac{1-e^{10b}}{1-e^{20b}}}$$

and the reparameterized model is

$$y = \frac{\theta_0}{\left(1 - e^{10b}\right)^{\frac{\ln \frac{\theta_0}{\theta_1}}{\ln \frac{1-e^{10b}}{1-e^{20b}}}}} \left(1 - e^{bT}\right)^{\frac{\ln \frac{\theta_0}{\theta_1}}{\ln \frac{1-e^{10b}}{1-e^{20b}}}}$$

which is fitted to the stand 1 data using initial values: $a = 10, b = -0.12$ and $c = 40$. The results are

```
Parameters:
     Value       Std.  Error   t value
 a   9.136080    0.16370100    55.8095
 b   -0.133922   0.00300942    -44.5008
 c0  41.249200   0.15142800    272.4010
Residual standard error:  0.402701 on 16 degrees of freedom
Parameter effects:  c^theta x sqrt(F) = 0.0219
```

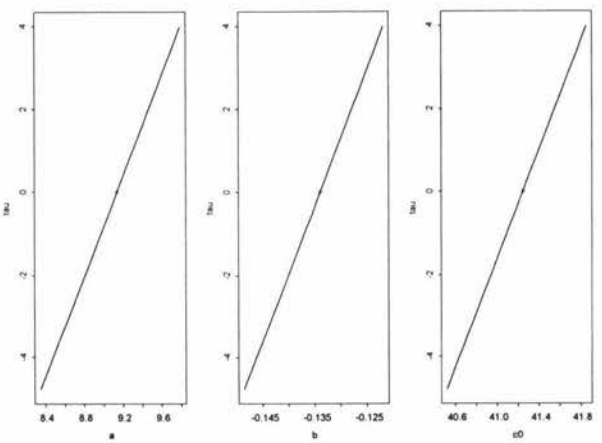


Figure 3-8: Profile plot for reparameterized Chapman-Richards function using expected value parameters for a and c

The curvature is now extremely small and the parameter estimates all have good estimation properties. In fitting the Chapman-Richards function to the data for other forest stands this parameterization should be very suitable. That is,

$$c(T) = \frac{\theta_0}{(1-e^{10b})^{\frac{\ln\frac{\theta_0}{\theta_1}}{\ln\frac{1-e^{10b}}{1-e^{20b}}}}} \left(1-e^{bT}\right)^{\frac{\ln\frac{\theta_0}{\theta_1}}{\ln\frac{1-e^{10b}}{1-e^{20b}}}}$$

and we will see that this reparameterization is very simple to fit. Ratkowsky [3] gives three examples of one-parameter reparameterizations for the Chapman-Richards models but did not give a reparameterization using two expected value parameters.

Zeide [6] studied the Guttenberg data [15] (growth and yield of 103 spruce trees). When we fitted the reparameterized Chapman-Richards model to height for each of the trees the mean intrinsic curvature and the mean parametric-effects curvature of the 103 fitted models were 0.07 and 0.10 respectively. Thus, the fitted models have quite acceptable curvature of either kind.

The Gompertz function

Consider the following transformation of the Gompertz function $y = g(T) = ae^{-be^{-cT}}$

$$\ln\left(-\ln\left(\frac{y}{a}\right)\right) = \ln b - c \cdot T$$

This is a linear model that is fitted, with $a = 70$, to obtain initial estimates of $\ln b$ and c as follows

```
Coefficients:
 Intercept   T
 1.943661    -0.1259171
```

giving $a = 70$, $b = e^{1.943661} = 6.98$ and $c = 0.13$. We use these initial estimates to fit the Gompertz function to the Stand 1 data by nonlinear least squares giving the following output.

```
Parameters:
     Value       Std.  Error   t value
 a   63.607300   0.62784600    101.3100
 b   8.753040    0.30340400    28.8495
 c0  0.150823    0.00301266    50.0629
Residual standard error:  0.478633 on 16 degrees of freedom

Parameter effects:  c^theta x sqrt(F) = 0.2541
```

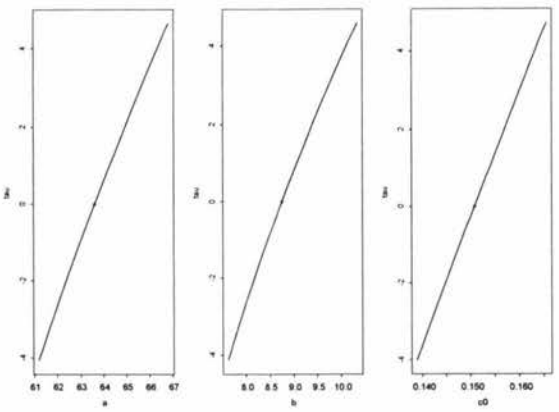


Figure 3-12: Profile plot for the Gompertz function fitted to Stand 1 data

The estimation properties of the parameters are quite good as is the small curvature of the model. A reparameterization is not required for the Gompertz function fitted to the stand 1 forestry data and the original model

$$g(T) = ae^{-be^{-cT}}$$

with starting values $a = 70$, $b = 6.98$ and $c = 0.126$ will be used to fit this model to the other forestry stands.

The Hossfeld function

From Table 2-3, the Hossfeld function $h(T) = \frac{aT^c}{ab+T^c}$ has asymptote of a and an inflection point at $T = \left[\frac{ab(c-1)}{c+1}\right]^{\frac{1}{c}}$, with y-coordinate $\frac{1}{2}(c-1)\frac{a}{c}$. These give the following equations in a, b, c

$$a = 70, \quad \left[\frac{ab(c-1)}{c+1}\right]^{\frac{1}{c}} = 15 \text{ and } \frac{1}{2}(c-1)\frac{a}{c} = 20$$

and solving for a, b, c gives the following estimates of the parameters

$$a = 70, \ b = 336.1 \text{ and } c = 3.5$$

Using S-Plus to fit the Hossfeld function to the data of stand 1 gives the following sequence

of commands and output. The trace option in the call to **nls()** allows the progress of the iterative process to be seen. The first column gives the residual sum of squares at each iteration and the additional columns list parameter estimates at each iteration. Convergence is obtained after four iterations.

```
>hoss <- deriv3(~a*x^c0/(a*b+x^c0),c(''a'',''b'',''c0''),
function(a,b,c0,x) NULL)
>hp1.nls <- nls(basal.area[plot==1]~hoss(a,b,c0,age[plot==1]),
start=list(a=70,b=336,c0=3.5),data=forest,trace=T)
 30.5334 :  70       336      3.5
 19.9403 :  68.7926  242.485  3.40753
 4.29565 :  67.3279  199.836  3.33632
 3.31531 :  67.3057  204.115  3.33771
 3.31505    67.3217  203.637  3.33668
```

The **nls()** function has a summary method that gives parameter estimates, their standard errors and t-statistics. It appears below that the estimate of b is relatively imprecise with its small t-statistic when compared to the other two parameter estimates. That is, the standard error of the parameter estimate of b is rather large compared to those of a and c. Note that the parameter estimates are highly correlated.

```
>summary(hp1.nls)
Formula:  basal.area[plot == 1] ~hoss(a, b, c0, age[plot == 1])
Parameters:
     Value      Std.  Error  t value
 a   67.32170   0.8569600    78.5587
 b   203.63700  32.5241000   6.2611
 c0  3.33668    0.0614782    54.2741
Residual standard error:  0.455182 on 16 degrees of freedom
Correlation of Parameter Estimates:
```

```
     a       b
b    -0.838
c0   -0.868  0.997
```

The curvature of the model is assessed using the **rms.curv()** function and profile plots.

```
>hp1.curv <- rms.curv(hp1.nls)
>hp1.curv
Parameter effects:  c^theta x sqrt(F) = 4.1831
>par(mfrow=c(1,3))
>plot(profile(hp1.nls))
```

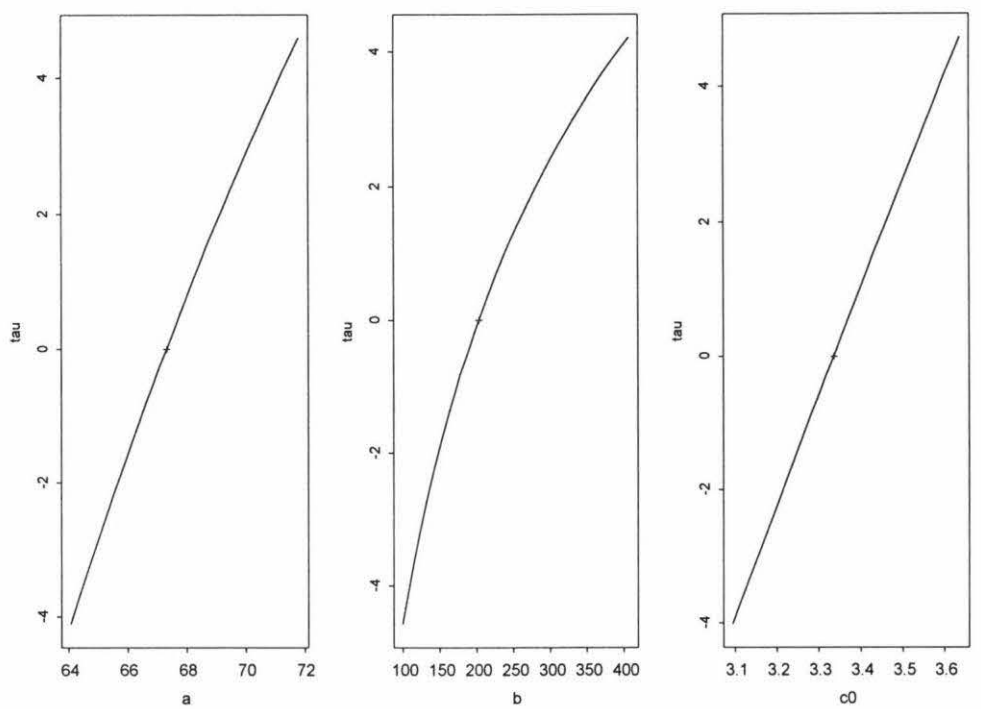


Figure 3-3: Profile plot for the Hossfeld function fitted to Stand 1 data

The curvature is unacceptably high (4.1831) and inspection of the profile plots (Figure 3-3) suggests that a reparameterization using an expected value for b is likely to provide better

estimation properties. A reparameterization is considered using an expected value at $T = 15$. That is, the parameter θ is introduced as

$$\theta = \frac{a \cdot 15^c}{ab + 15^c}$$

We solve for b giving

$$b = 15^c \left(\frac{1}{\theta} - \frac{1}{a} \right)$$

which yields the reparameterized model

$$h(T) = \frac{aT^c}{15^c \left(\frac{a}{\theta} - 1 \right) + T^c}$$

Fitting this model gives the following edited S-Plus output

```
Parameters:
     Value      Std.  Error  t value
 a   67.32080  0.8569160    78.5618
 th  25.57620  0.1529050    167.2680
 c0  3.33675   0.0614791    54.2745
Residual standard error:  0.455182 on 16 degrees of freedom

Correlation of Parameter Estimates:
     a          th
 th  -0.2980
 c0   -0.8680  0.0478
Parameter effects:  c^theta x sqrt(F) = 0.2077
```

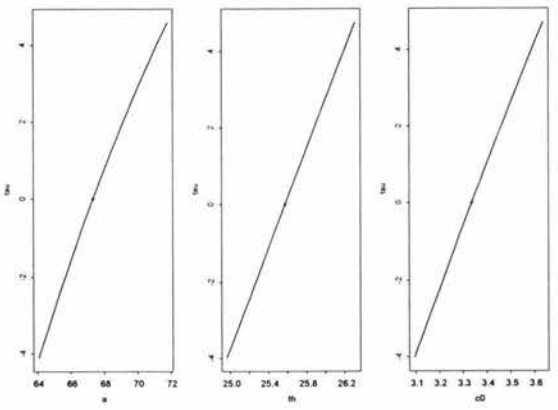


Figure 3-4: Profile plot for reparameterized Hossfeld function using an expected value parameter for b

For the reparameterized model, curvature is within acceptable limits and the profile plot (Figure 3-4) is much more nearly linear for each parameter. The t-statistics of all parameters are large and correlations of parameters a and c with θ are relatively small. Not shown here, but noted, is the slightly faster convergence for the reparameterized model. Therefore, a suitable parameterization that will be used to fit the Hossfeld function to the data for the other forest stands is

$$h(T) = \frac{aT^c}{15^c\left(\frac{a}{\theta} - 1\right) + T^c}$$

The Schumacher function

From Chapter 2, the Schumacher function $y = s(T) = e^{a - \frac{b}{T^c}}$ has an asymptote of e^a and point of inflection with coordinates $\left(\left(\frac{bc}{c+1}\right)^{\frac{1}{c}}, e^{\frac{ac-c-1}{c}}\right)$. Rather than using estimates of the point of inflection and the asymptote to determine starting values consider a transformation of $s(T)$ noting that e^a is estimated as 67 (this being established from the Hossfeld model in the previous section). The transformed model is

$$\ln\left(-\ln\left(\frac{y}{67}\right)\right) = \ln(b) - c\ln(T)$$

which is a linear model that can be fitted to give estimates of $\ln b$ and c. The S-Plus output gives:

```
Coefficients:
 (Intercept)  log(age[plot == 1])
 5.706355     -2.174827
```

and estimates of b and c are calculated as

$$b = e^{5.7064} = 300.78629 \text{ and } c = 2.175$$

The Schumacher function is now fitted using nonlinear least squares with starting values $a = 4.2, b = 301$ and $c = 2.2$. The edited S-Plus output is

```
Parameters:
     Value      Std.  Error    t value
 a   4.51073    0.0368844     122.29400
 b   77.99030   10.6199000    7.34379
 c0  1.52741    0.0614521     24.85530
Residual standard error:  0.609749 on 16 degrees of freedom
Parameter effects:  c^theta x sqrt(F) = 4.6802
```

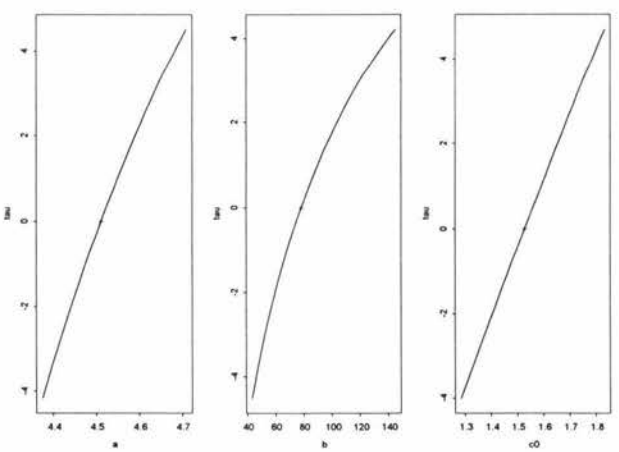


Figure 3-5: Profile plot for the Schumacher function fitted to Stand 1 data

It seems that curvature is associated with parameters a and b and a reparameterization is considered using expected value coordinates at $T = 10$ and $T = 20$. That is,

$$\theta_0 = e^{a - \frac{b}{10^c}} \text{ and } \theta_1 = e^{a - \frac{b}{20^c}}$$

Solving for a and b gives

$$a = \frac{\ln \theta_0 - 2^c \ln \theta_1}{1 - 2^c} \text{ and } b = \frac{20^c}{1 - 2^c} (\ln \theta_0 - \ln \theta_1)$$

and the reparameterized model is

$$y = e^{\frac{\ln \theta_0 - 2^c \ln \theta_1}{1 - 2^c} - \frac{\frac{20^c}{1-2^c} (\ln \theta_0 - \ln \theta_1)}{T^c}}$$

which is fitted using initial values $\theta_0 = 10$, $\theta_1 = 40$ and $c = 2$. The S-Plus output is summarised below and appears much more acceptable in its parameter estimates and curvature. There is also considerable improvement in convergence rate where about half the iterations are required for the reparameterized model compared to that for the original model.

```
Parameters:
      Value     Std.  Error  t value
th0  8.982738  0.2630140     34.1530
th1  40.75260  0.2159450     188.7180
c0   1.52744   0.0614525     24.8556
Residual standard error:  0.609749 on 16 degrees of freedom

Parameter effects:  c^theta x sqrt(F) = 0.0414
```

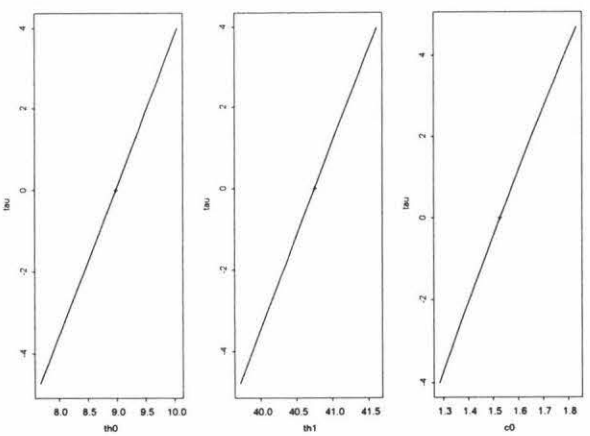


Figure 3-6: Profile plot for reparameterized Schumacher function using expected value parameters for a and b

In summary, the reparameterized model

$$s(T) = e^{\frac{\ln\theta_0 - 2^c \ln\theta_1}{1-2^c} - \frac{\frac{20^c}{1-2^c}(\ln\theta_0 - \ln\theta_1)}{T^c}}$$

seems to be quite adequate for Stand 1 and will be used to fit the data for the other stands.

The Weibull function

Linear regression will be used to obtain starting values for fitting the Weibull function, $y = w(T) = a\left(1 - e^{bT^c}\right)$. Consider the transformation of the model

$$\ln\left[-\ln\left(1 - \frac{y}{a}\right)\right] = \ln(-b) + c\ln T$$

This is a linear model that is fitted (with $a = 70$) to determine initial estimates of b and c as follows

```
Coefficients:
 (Intercept)  log(age[plot == 1])
-8.015476     2.639012
```

This gives initial estimates of parameters b and c as $b = -e^{-8.02} = -.000329$ and $c = 2.64$ which are used as starting values for fitting the Weibull function to the Stand 1 data giving the following edited S-plus output.

```
Parameters:
     Value       Std.  Error    t value
 a   58.445000   0.9423870000   62.01800
 b   -0.000312   0.0000648671   -4.80984
 c0  2.766780    0.0801931000   34.50150
Residual standard error:  0.939986 on 16 degrees of freedom

Parameter effects:  c^theta x sqrt(F) = 4.3421
```

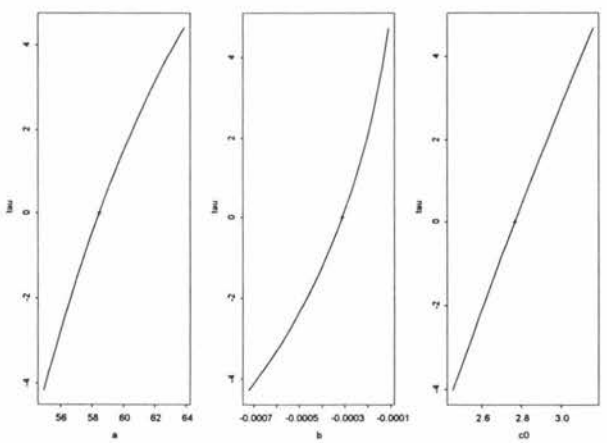


Figure 3-9: Profile plot for Weibull function fitted to Stand 1 data

Large curvature is present with the most badly behaving parameter b. A stable parameterization is developed using the expected value at $T = 10$. That is,

$$\theta_0 = a\left(1 - e^{10^c \cdot b}\right)$$

which is rearranged to give $b = \frac{\ln\left(1-\frac{\theta_0}{a}\right)}{10^c}$.

This gives the reparameterized model suggested by Ratkowsky[3], page 134:

$$y = a\left[1 - e^{\left(\frac{T}{10}\right)^c \cdot \ln\left(1-\frac{\theta_0}{a}\right)}\right], \text{ or } y = a\left[1 - \left(1 - \frac{\theta_0}{a}\right)^{\left(\frac{T}{10}\right)^c}\right]$$

which is fitted to the stand 1 data using initial values: $a = 70, \theta_0 = 10$ and $c = 2.6$. The results are

```
Parameters:
      Value     Std. Error  t value
a     58.44460  0.9423480     62.0202
th0   9.74279   0.3366320     28.9420
c0    2.76682   0.0801936     34.5017
Residual standard error:  0.939986 on 16 degrees of freedom

Parameter effects:  c^theta x sqrt(F) = 0.3116
```

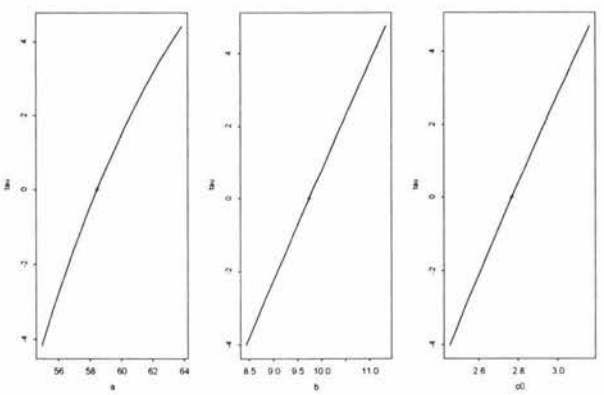


Figure 3-10: Profile plot for reparameterized Weibull function using an expected value parameters for b

The reparameterization has been somewhat successful with very nearly acceptable parametric-effects curvature. However, another reparameterization should be considered since there is still curvature associated with parameter a, as indicated by the profile plot displayed in Figure 3-10.

Ratkowsky suggests replacing a with an expected value parameter θ_0 and also replacing b by $\ln \theta_1$ which gives

$$\theta_0 = a\left(1 - e^{10^c \cdot \ln \theta_1}\right), \text{ or } a = \frac{\theta_0}{1 - \theta_1^{10^c}}$$

and leads to the reparameterized model

$$y = \left(\frac{\theta_0}{1 - \theta_1^{10^c}}\right)\left(1 - e^{T^c \cdot \ln \theta_1}\right) = \theta_0 \frac{1 - \theta_1^{T^c}}{1 - \theta_1^{10^c}}$$

which is fitted to the data for stand 1 using starting values $\theta_0 = 10$, $\theta_1 = e^{-.000312} \approx .9997$ and $c = 2.6$. This fit is summarised below with parameter θ_1 far-from-linear as shown in Figure 3-11.

```
Parameters:
      Value      Std. Error    t value
th0  9.742880  0.3366320000  28.9422
th1  0.999688  0.0000648466  15416.2000
c0   2.766790  0.0801931000  34.5015
Residual standard error:  0.939986 on 16 degrees of freedom

Parameter effects:  c^theta x sqrt(F) = 5.3036
```

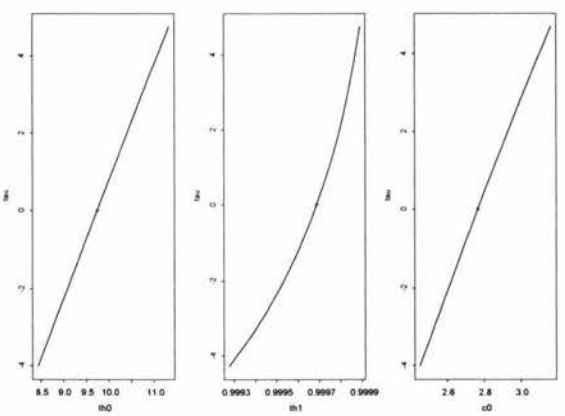


Figure 3-11: Profile plot for the second reparameterization of the Weibull function using an expected value parameter for a

Given the considerable nonlinearity associated with θ_1, there is little advantage in using this formulation of the model. In fitting the Weibull function the reparameterization discussed initially will be used, namely

$$y = a\left[1 - \left(1 - \frac{\theta_0}{a}\right)^{\left(\frac{T}{10}\right)^c}\right]$$

The Levakovic function

From Table 2-3, the Levakovic function $y = L(T) = a\left(\frac{T^d}{b+T^d}\right)^c$ has asymptote a and point of inflection with coordinates $\left(\left(\frac{b(cd-1)}{1+d}\right)^{\frac{1}{d}}, \frac{a(cd-1)^c}{(c+1)^c d^c}\right)$. We estimate $a = 70$ and inflection point $(15, 25)$ giving the following three equations in b, c and d

$$70\left(\frac{15^d}{b+15^d}\right)^c = 25,\ \left(\frac{b\,(cd-1)}{1+d}\right)^{\frac{1}{d}} = 15, \text{ and } \frac{70\,(cd-1)^c}{(c+1)^c\,d^c} = 25$$

which are solved to give

$$a = 70,\ b = 9.98451\,31 \times 10^5,\ c = 2.4 \text{ and } d = 5.3$$

Given the magnitude of b it will be replaced by $b_1 = \ln b$ giving an initial estimate of $b_1 = 13.8$ and the nonlinear model is fitted using these estimates of a, b_1, c and d.

Using S-Plus, the edited output from fitting the Levakovic model is

```
Parameters:
      Value     Std. Error  t value
  a  72.76710  2.329670     31.23490
  b1  6.84594  0.855209      8.00499
  c0  1.70945  0.361573      4.72783
  d   2.59545  0.226152     11.47660
Residual standard error:  0.36373 on 15 degrees of freedom

Parameter effects:  c^theta x sqrt(F) = 16.0794
```

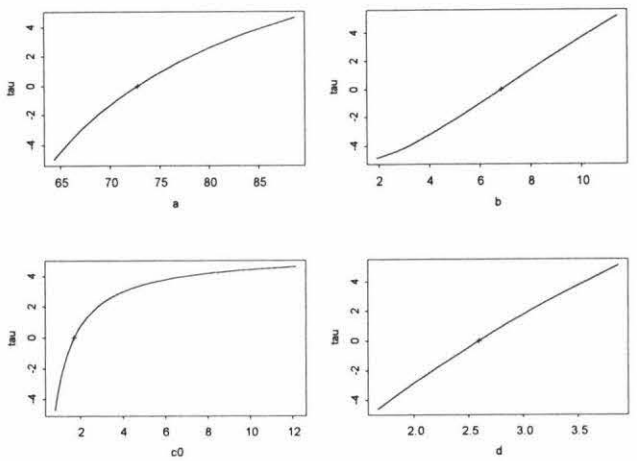


Figure 3-13: Profile plot for the Levakovic function fitted to Stand 1 data

Extremely large parametric-effects curvature is apparent ($c^{\theta}\sqrt{F} = 16.1$) with the parameters a and c exhibiting greatest nonlinearity, as shown in the profile plot of Figure 3-13. A stable parameterization is developed using expected value parameters for a and c at $T = 10$ and $T = 20$. We therefore define θ_0 and θ_1 as

$$\theta_0 = a\left(\frac{10^d}{b_1 + 10^d}\right)^c \text{ and } \theta_1 = a\left(\frac{20^d}{b_1 + 20^d}\right)^c$$

and solving these for c gives

$$\frac{\left(\frac{10^d}{b_1+10^d}\right)^c}{\left(\frac{20^d}{b_1+20^d}\right)^c} = \frac{\theta_0}{\theta_1}, \text{ or } c = \frac{\ln\left(\frac{\theta_0}{\theta_1}\right)}{\ln\left[\frac{\left(\frac{10^d}{b_1+10^d}\right)}{\left(\frac{20^d}{b_1+20^d}\right)}\right]}$$

This expression for c is substituted into $\theta_0 = a\left(\frac{10^d}{b_1+10^d}\right)^c$ and solved for a

$$\theta_0 = a\left(\frac{10^d}{b_1 + 10^d}\right)^{\frac{\ln\left(\frac{\theta_0}{\theta_1}\right)}{\ln\left[\frac{\left(\frac{10^d}{b_1+10^d}\right)}{\left(\frac{20^d}{b_1+20^d}\right)}\right]}}, \text{ or } a = \theta_0\left(\frac{b_1 + 10^d}{10^d}\right)^{\frac{\ln\left(\frac{\theta_0}{\theta_1}\right)}{\ln\left[\frac{\left(\frac{10^d}{b_1+10^d}\right)}{\left(\frac{20^d}{b_1+20^d}\right)}\right]}}$$

giving the reparameterized model

$$y = \theta_0 \left[\frac{T^d \left(b_1 + 10^d\right)}{10^d \left(b_1 + T^d\right)} \right]^{\frac{\ln(\theta_0/\theta_1)}{\ln\left(\frac{10^d(b_1+20^d)}{20^d(b_1+10^d)}\right)}}$$

which is fitted using initial parameter estimates $\theta_0 = 10, b_1 = 6.8, \theta_1 = 40$ and $d = 2.6$ (using the results of the original fit for the estimates of b_1 and d and the expected values of θ_0 and θ_1) with the following results.

```
Parameters:
      Value     Std. Error  t value
th0  9.04520   0.154140     58.68160
b1   6.84557   0.855207      8.00458
th1  41.19930  0.155860     264.33500
d    2.59536   0.226146     11.47640
Residual standard error:  0.36373 on 15 degrees of freedom

Parameter effects:  c^theta x sqrt(F) = 1.629
```

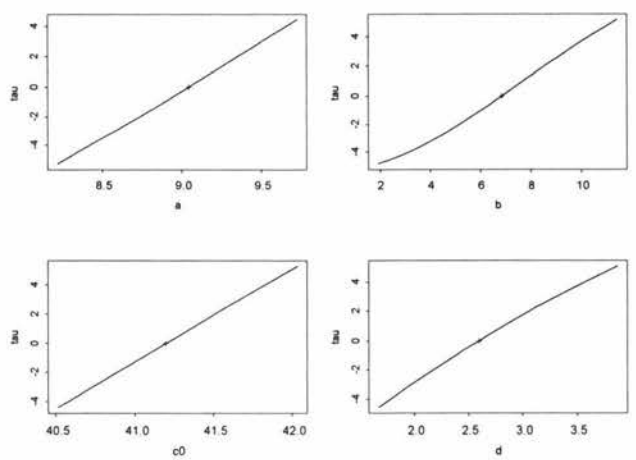


Figure 3-14: Profile plot for reparameterized Levakovic function using expected value parameters for a and c

For estimation purposes, this is much more satisfactory ($c^\theta\sqrt{F} = 1.6$) than the original model although the parametric-effects curvature is larger than desired. However, a 3-parameter

reparameterization is intractable so the following function will be used in fitting the Levakovic model to basal area for other forest stands.

$$y = \theta_0 \left[\frac{T^d \left(b_1 + 10^d\right)}{10^d \left(b_1 + T^d\right)} \right]^{\frac{\ln(\theta_0/\theta_1)}{\ln\left(\frac{10^d(b_1+20^d)}{20^d(b_1+10^d)}\right)}}$$

The Sloboda function

Consider the following form of the Sloboda function, $y = S(T) = ae^{-be^{-cT^d}}$:

$$\ln\left(-\ln\left(\frac{\ln\left(\frac{y}{a}\right)}{-b}\right)\right) = \ln c + d\ln\left(T\right) \tag{3.1}$$

The asymptote is estimated as 80 (this is the asymptote level predicted when the Schumacher model is fitted to the data) and since the asymptote is determined by ae^{-b} we have $b = \ln\frac{a}{80}$ and we substitute for b in equation (3.1) giving

$$\ln\left(-\ln\left(\frac{\ln y - \ln a}{\ln 80 - \ln a}\right)\right) = \ln c + d\ln\left(T\right) \tag{3.2}$$

which is linear in $\ln T$ and can be fitted for a range of values of a. By choosing the particular model that minimizes the residual sum of squares we obtain initial parameter estimates for c and d.

Shown below are the residual sum of squares when the linear model described by equation (3.2) is fitted to Stand 1 data for various values of a. It is seen that $a = 0.6$ provides the best fit for the values considered. Note that it is not necessary to be too precise at this stage; we need a value of a sufficiently accurate to ensure that the estimates of c and d obtained will provide reasonable starting values to ensure the nonlinear least squares algorithm converges. If the nonlinear fit does not converge, we can return to the linear model to find a better choice for a and more accurate estimates of b and c.

a	Residual SS
1	.011228 33
2	.087883 37
.5	.0070747 2
.3	.0091500 8
.6	.0068681 7

Using $a = .6$ to fit the linear model (3.2), the parameter estimates of $\ln c$ and d are obtained as

```
Coefficients:
 Intercept(ln c)   ln T (d)
 4.070445          -1.998474
```

yielding the following initial estimates of the parameters

$$a = .6,\ b = \ln\left(\frac{.6}{80}\right) = -4.9,\ c = e^{4.07} = 58.6 \text{ and } d = -2$$

Using these initial values to fit the Sloboda function to stand 4 data (used rather than stand 1 since the curvature measure proved so extreme to cause the profile plot routine to fail) and nonlinear least squares gives the following results:

```
Parameters:
     Value       Std.  Error   t value
 a   0.910902    0.3902750     2.33400
 b   -4.428640   0.4539990     -9.75473
 c0  58.478400   19.8641000    2.94393
 d   -1.908290   0.0919375     -20.75640
Residual standard error:  0.280609 on 15 degrees of freedom

Parameter effects:  c^theta x sqrt(F) = 67.0879
```

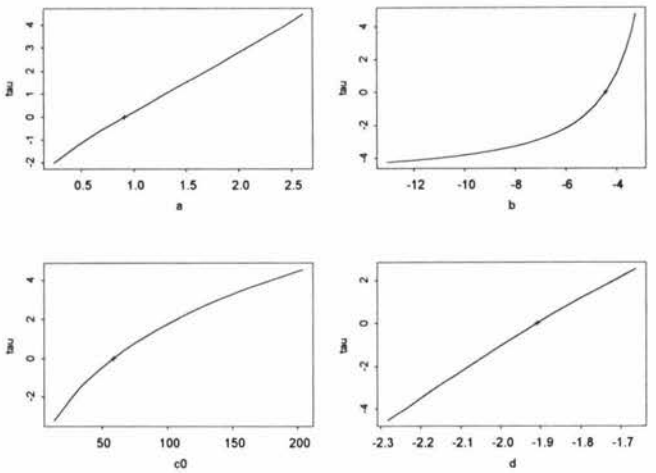


Figure 3-15: Profile plot for the Sloboda function fitted to Stand 4 data

The estimation properties of a and c are not particularly good and there is extremely large curvature mainly through the nonlinearity of b and c. An attempt is made to reduce this by replacing b and c, the most badly behaved parameters, with expected value parameters. A stable parameterization is developed using expected-value parameters θ_0 and θ_1 determined at $T = 20$ and $T = 10$ respectively. That is,

$$\theta_0 = ae^{-be^{-c\cdot 20^d}} \text{ and } \theta_1 = ae^{-be^{-c\cdot 10^d}}$$

We find c by combining $\ln \frac{\theta_0}{a} = -be^{-c\cdot 20^d}$ and $\ln \frac{\theta_1}{a} = -be^{-c\cdot 10^d}$ to give

$$\ln\left[\frac{\ln \frac{\theta_0}{a}}{\ln \frac{\theta_1}{a}}\right] = 10^d c\left(1 - 2^d\right), \text{ or } c = \frac{\ln\left[\frac{\ln \frac{\theta_0}{a}}{\ln \frac{\theta_1}{a}}\right]}{10^d\left(1 - 2^d\right)}$$

and we now solve for b by noting that $\frac{\theta_0}{\theta_1} = \frac{e^{-be^{-c\cdot 20^d}}}{e^{-be^{-c\cdot 10^d}}}$ to give

$$\ln \frac{\theta_0}{\theta_1} = b\left(e^{-c10^d} - e^{-c20^d}\right), \text{ or } b = \frac{\ln \frac{\theta_0}{\theta_1}}{e^{-c10^d} - e^{-c20^d}}$$

The expressions for b and c now give the reparameterized model

$$S(T) = a \exp\left[- \left(\frac{\ln \frac{\theta_0}{\theta_1}}{e^{\frac{\ln\left(\frac{\ln \frac{\theta_0}{a}}{\ln \frac{\theta_1}{a}}\right)}{2^d - 1}} - e^{\frac{2^d \ln\left(\frac{\ln \frac{\theta_0}{a}}{\ln \frac{\theta_1}{a}}\right)}{2^d - 1}}} \right) e^{-\frac{\ln\left[\frac{\ln \frac{\theta_0}{a}}{\ln \frac{\theta_1}{a}}\right]}{10^d(1-2^d)} T^d} \right]$$

and using initial values of $a = 0.9, \theta_0 = 40, \theta_1 = 10$ and $d = -1.9$ we now fit this model to stand 4 data using nonlinear least squares.

```
Parameters:
      Value       Std.  Error   t value
 a    0.910672    0.3902610     2.33349
 th0  35.167600   0.1154510     304.60900
 th1  7.825760    0.1154140     67.80600
 d    -1.908230   0.0919409     -20.75500
Residual standard error:  0.280609 on 15 degrees of freedom

Parameter effects:  c^theta x sqrt(F) = 0.9608
```

There is still quite large parametric-effects curvature present in this parameterization ($c^\theta\sqrt{F} = .96$) and the statistical properties of a do not look promising with its t-value of 2.33. Also, a profile plot is not available with the S-Plus function generating errors through invalid mathematical functions (logarithms of negative numbers).

An alternative parameterization uses expected value parameters for a and b. We start as before with

$$\theta_0 = ae^{-be^{-c\cdot 20^d}} \text{ and } \theta_1 = ae^{-be^{-c\cdot 10^d}}$$

but now solve for a and b rather than a and c as in the first reparameterization. In terms of θ_0 and $\theta_1$this gives a and b as

$$b = \frac{\ln \frac{\theta_0}{\theta_1}}{-e^{-c\cdot 20^d} + e^{-c\cdot 10^d}} \text{ and } a = \theta_0 e^{be^{-c\cdot 20^d}} = \theta_0 e^{\frac{\ln \frac{\theta_0}{\theta_1}}{-e^{-c\cdot 20^d} + e^{-c\cdot 10^d}} e^{-c\cdot 20^d}}$$

The parameterization is then given as

$$S(T) = \theta_0 e^{\frac{\ln\frac{\theta_0}{\theta_1}}{-e^{-c\cdot 20^d}+e^{-c\cdot 10^d}} e^{-c\cdot 20^d}} \; e^{\frac{-\ln\frac{\theta_0}{\theta_1}}{-e^{-c\cdot 20^d}+e^{-c\cdot 10^d}} e^{-c\cdot T^d}}$$

This model is fitted to stand 1 data using initial values, $\theta_0 = 40$, $\theta_1 = 10$, $c = 50$ and $d = -2$ with the following results

```
Parameters:
      Value      Std. Error  t value
th0  8.91964    0.1492390     59.7675
th1  41.16810   0.1447120     284.4820
c0   115.18800  38.7498000    2.9726
d    -2.18753   0.0944351     -23.1644
Residual standard error:  0.355035 on 15 degrees of freedom

Parameter effects:  c^theta x sqrt(F) = 3.9433
```

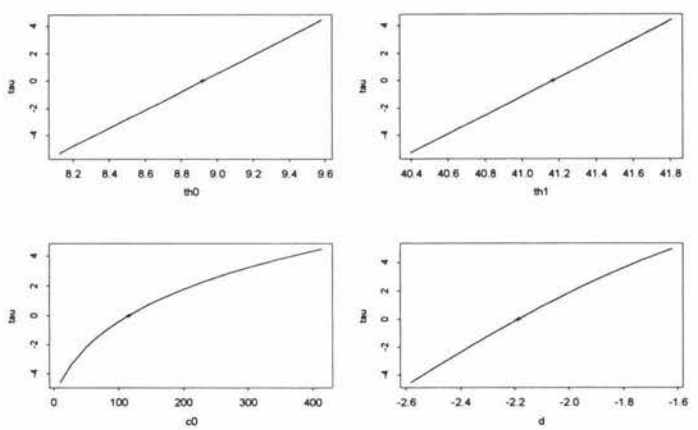


Figure 3-16: Profile plot for the second reparametrization of the Sloboda function using expected value parameters for a and b

The curvature is associated almost exclusively with c. There is little to choose between the two reparameterizations that we have considered. The second formulation of the model

$$S(T) = \theta_1 e^{\frac{\ln\frac{\theta_1}{\theta_0}}{-e^{-c\cdot 20^d}+e^{-c\cdot 10^d}} e^{-c\cdot 20^d}} \; e^{\frac{-\ln\frac{\theta_1}{\theta_0}}{-e^{-c\cdot 20^d}+e^{-c\cdot 10^d}} e^{-c\cdot T^d}}$$

will be used in fitting the Sloboda function to basal area data for the other stands even though its curvature appears less satisfactory than the first. It is however simpler, the profile plots are available and parameter estimates have similar overall properties.

Chapter 4

The Models Fitted to Basal Area

4.1 Introduction

The reparameterizations and starting values developed in Chapter 3 are used to fit the seven growth curve models to basal area data of 22 stands of pinus radiata. The data and their background is detailed in Appendix F. For each model, the following information is presented and discussed

1. Estimates of the asymptote level
2. Estimates of the inflection point
3. Plots of fitted models and raw observations versus age
4. Residuals and residual sum of squares
5. Plots of residuals versus estimated slope of the model function
6. Plots of residuals versus fits
7. Plots of residuals versus age
8. The extent of curvature in the fitted models

The presentation concentrates on the data from stand 1 with similar information from other stands included in the appendices when considered important.

4.2 Estimate of asymptote level

The asymptote levels were calculated for each model fitted to each stand and are displayed in Figure 4-1. For reference, the parameter estimates are tabulated in Appendix B. Note also that the maximum measured basal area from the data is also plotted for each stand.

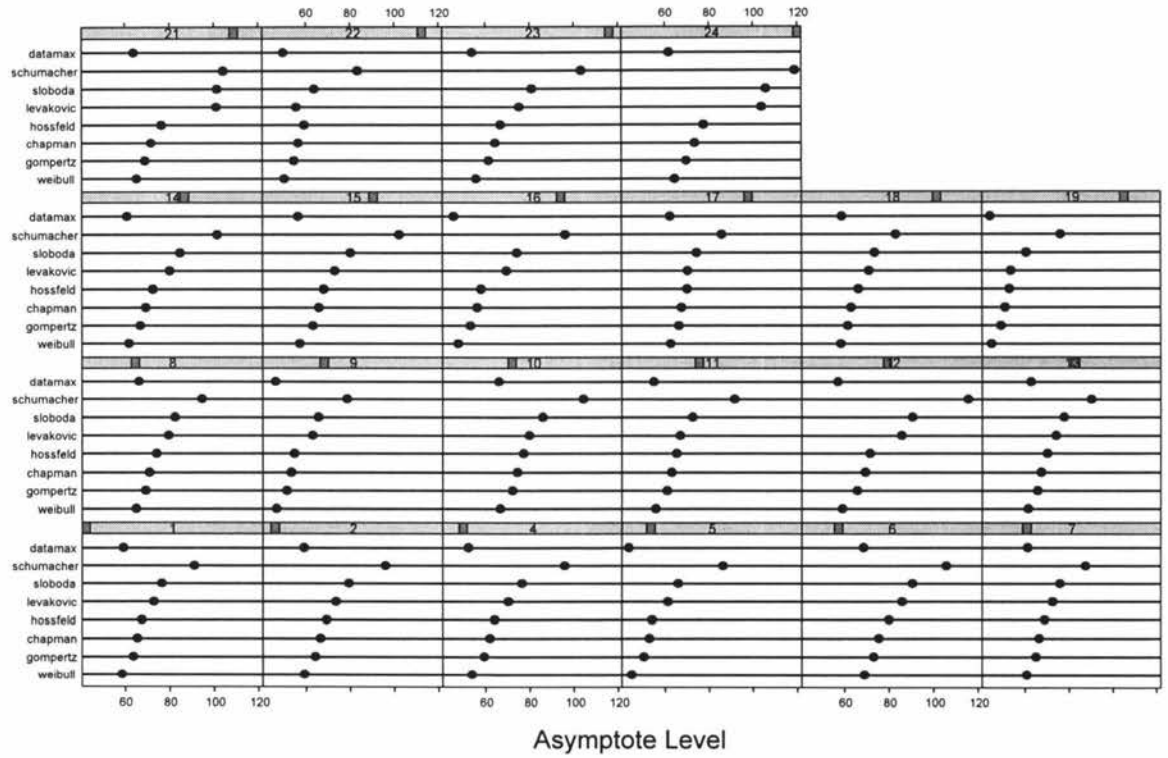


Figure 4-1: The asymptote determined when seven growth curves are fitted to area data of 22 forest stands.

It is clear that the Schumacher model estimates consistently larger asymptotes than the other models. For some stands (e.g. 1, 4, 5, 10, 12, 15, 16 etc.) it is quite noticeably greater than the next largest (generally Sloboda). The Weibull model gives very small asymptotic estimates; in some cases smaller than the maximum area from the data (e.g. stands 1, 8 and

13) and always very similar to the maximum area from the data.

4.3 Estimate of the point of inflection

Using the parameter estimates (listed in Appendix B) and the properties of the functions developed in Chapter 2, estimates of the T-coordinate of the point of inflection are calculated. To illustrate the method of calculation an example using the Chapman-Richards model fitted to Stand 1 data is presented below.

Example Calculation of the T-coordinate of the point of inflection for the Chapman-Richards model for stand 1. The parameter estimates are

$$a = 65.39,\ b = -0.13 \text{ and } c = 6.48$$

Also, recall that the T-coordinate of the inflection point for this model is given by $T = \frac{-\ln c}{b}$ which gives an estimate

$$T = \frac{-\ln 6.18}{-0.13} = 13.95$$

of the T-coordinate of the inflection point.

Figure 4-2 shows the estimated T-coordinate of the inflection point for each model and stand. It is noticeable that the Schumacher model in general fits lower values for this quantity than the other sigmoid models. In general the structure of the relative positions of the T-coordinate is consistent from stand to stand with the exception of stand 21 and stand 22 (discussed in Section 4.6 in relation to goodness-of-fit). The Weibull and Sloboda models estimate larger T-coordinates than the other five models.

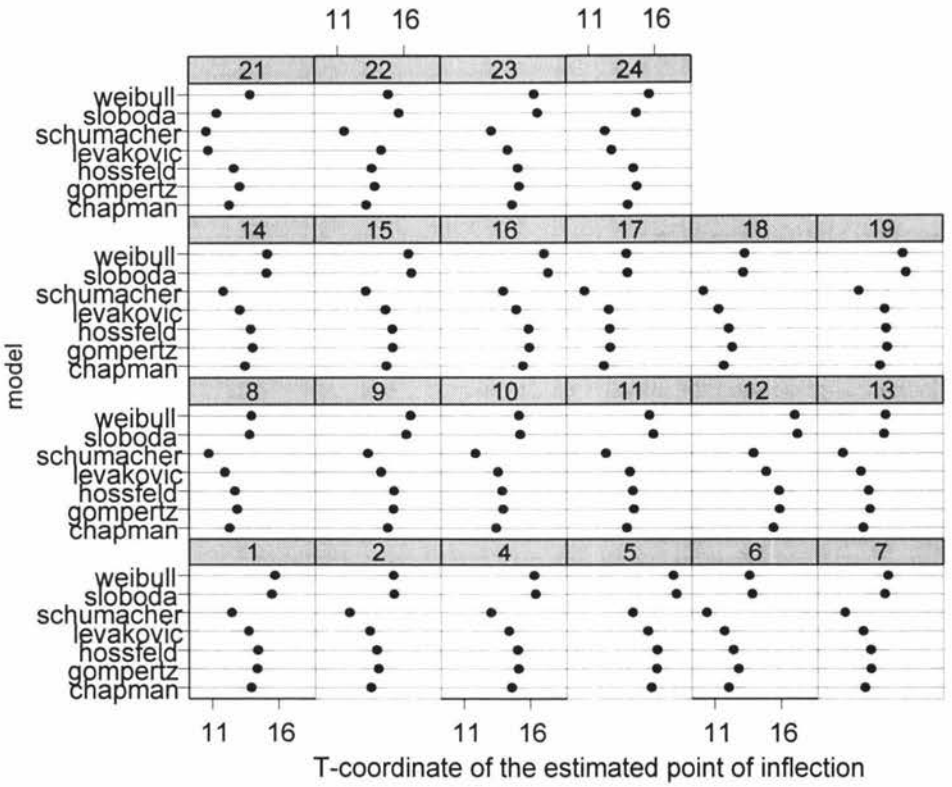


Figure 4-2: T-coordinate of the inflection point for each forestry stand and model

In general lower estimates of the T-coordinate of the point of inflection are associated with higher estimates of asymptote level. This is illustrated in Figure 4-3 but the association is less convincing for the Schumacher model than the other six models.

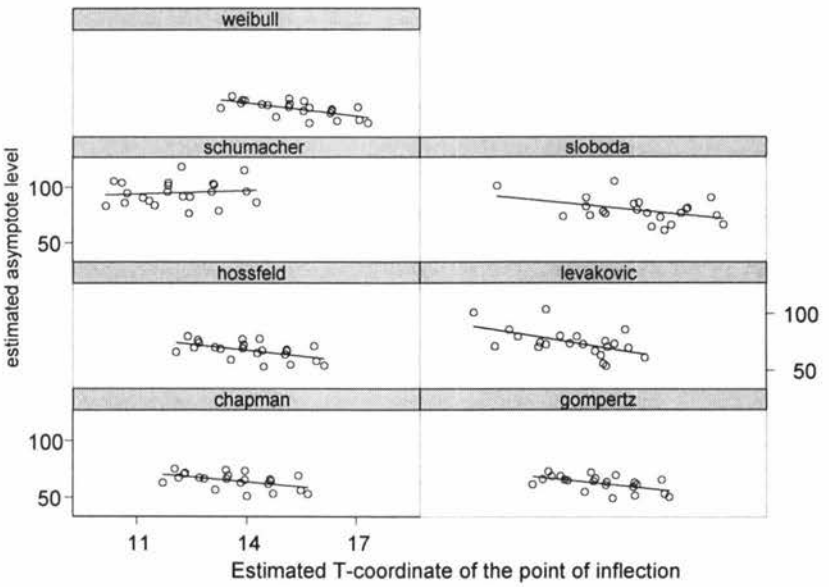


Figure 4-3: Estimated asymptote level plotted against the estimated T-coordinate of the inflection point for each growth curve model.

4.4 Plots of raw data and fitted models

The fit of each model to each stand is displayed graphically. A plot for Stand 1 is presented in Figure 4-4 and Appendix E gives a selection of corresponding plots for Stands 2 to 24.

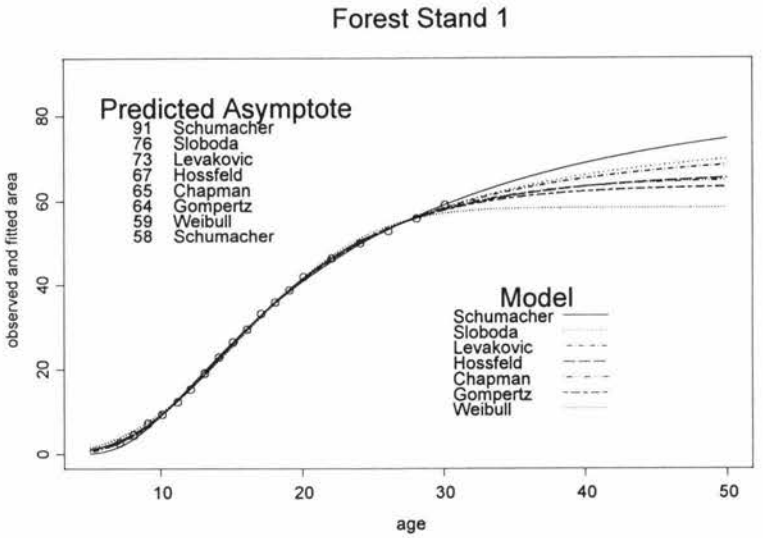


Figure 4-4: The raw basal area data (marked with open circles) and the seven fitted growth curves for Stand 1.

Also shown here (Figures 4-5 to 4-7) are "exploded" portions of the graphs for stands 1 and 24 which will give further information concerning the fit at low ages, ages near the asymptote and high ages for the Hossfeld, Chapman-Richards, Sloboda, Weibull and Schumacher models. These five have been chosen to reflect the varying quality of the models considered. We will see that the Sloboda, a 4-parameter model, appears to be good in all criteria discussed in this chapter as are the 3-parameter Chapman-Richards and Hossfeld models. The Weibull and Schumacher models appear poor: the Weibull model always gives small asymptotes and has quite patterned residuals and poor goodness-of-fit; the Schumacher model also has relatively poor goodness-of-fit, patterned residuals and always gives large asymptotes.

Age 5 years to 12 years The raw data and the fitted models are displayed up to age 12 years in Figure 4-5. Both Schumacher and Weibull models fit quite poorly at low ages, overestimating and underestimating respectively. The fits of all models appear to give closer agreement as the point of inflection (recall that the point of inflection typically is between 10 and 14 years) is neared.

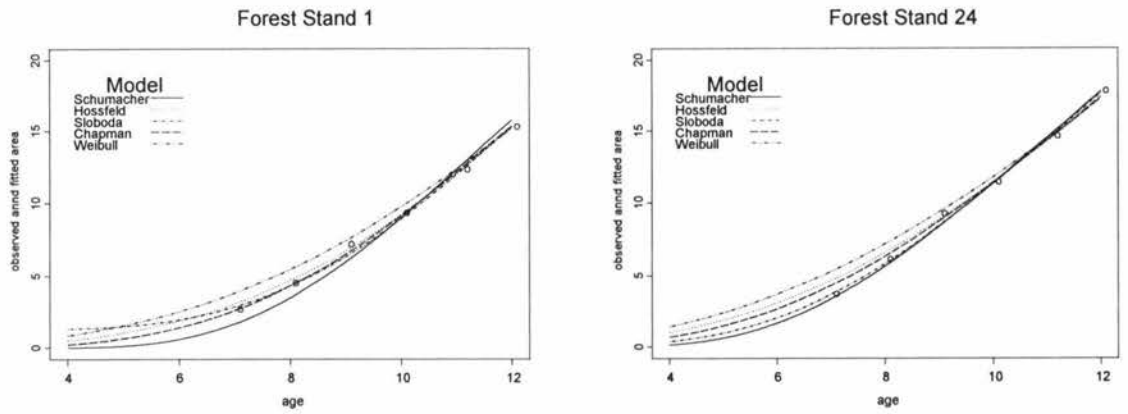


Figure 4-5: Observed and fitted areas for ages up to 12 years for five growth curves.

Age 12 years to 19 years The fitted models and the raw data are displayed for ages between 12 and 19 years in Figure 4-6. These plots show the fitted values near the point of inflection and it is seen that with the exception of the Weibull model (which underestimates here) reasonably close agreement is obtained.

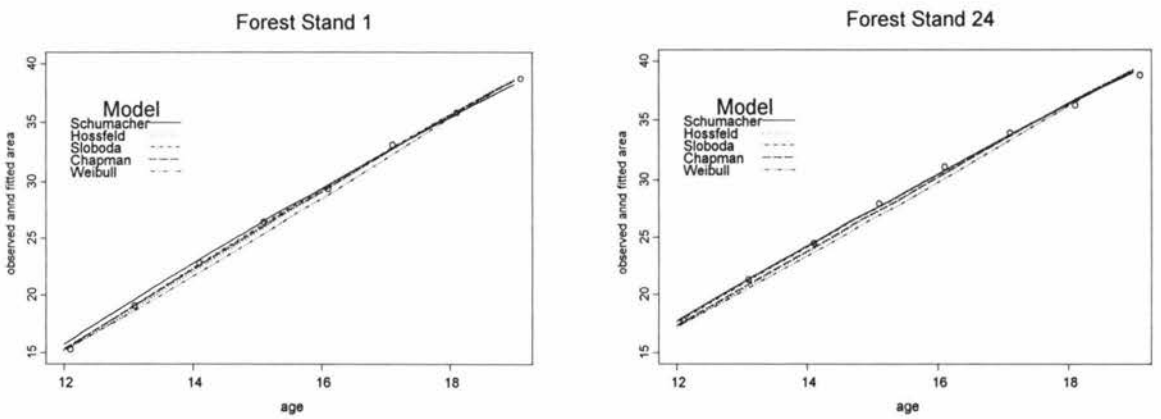


Figure 4-6: Observed and fitted areas for ages from 12 to 19 years for five growth curves.

Age 19 years to 30 years The fitted models and the raw data are displayed for ages between 19 and 30 years in Figure 4-6. The Weibull model overestimates in the age range of about 20 to 26 years and then underestimates from about 28 years. Conversely, the Schumacher model underestimates in the approximate age interval $(20, 26)$ and then fits higher values than all other models from about age 28 years.

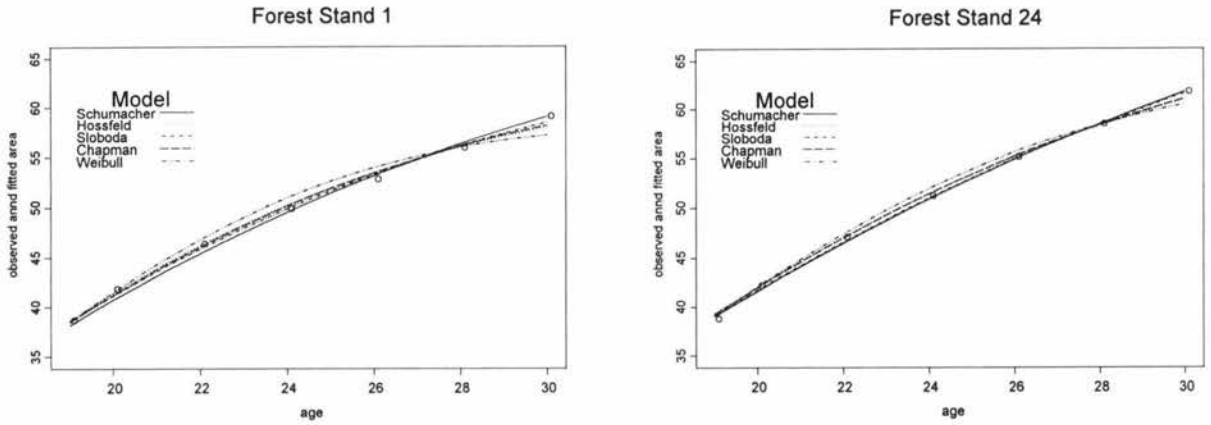


Figure 4-7: Observed and fitted areas for ages from 19 years for five growth curves.

4.5 Residuals

The residuals for each model fitted to forestry Stand 1 data are plotted in Figure 4-8 and residual plots for all models fitted to all stands are contained in Appendix A.

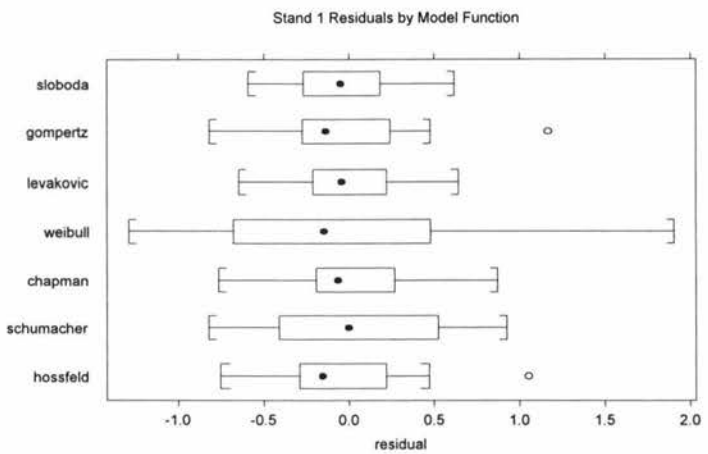


Figure 4-8: The residuals of the seven growth curves fitted to Stand 1 data.

The comments made here concerning Stand 1 are in general applicable to the results of nonlinear modelling for each of the other stands. It is clear that the residuals have a reasonably similar structure for the Hossfeld, Chapman, Levakovic, Gompertz and Sloboda models. The Weibull and Schumacher models often exhibit somewhat greater residual variance with this

more noticeable for the Weibull model. By way of example, we compare the Weibull and the Sloboda models in Figures 4-9 and 4-10 below. It is clear that the better fit is given by the 4-parameter Sloboda model. The Weibull model appears to fit well near the point of inflection but comparatively poorly at both low and high ages. Its poor fit at high ages is particularly worrying since this is where we want to start making predictions.

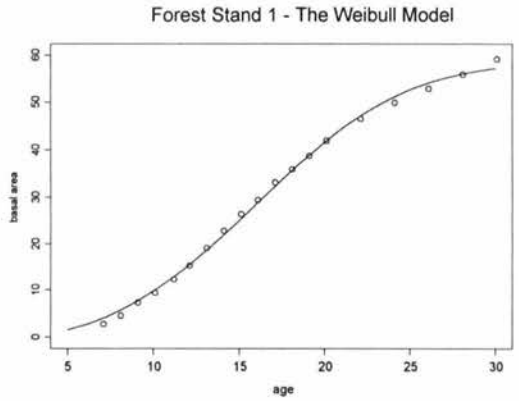


Figure 4-9: The Stand 1 raw data and the Weibull fitted model.

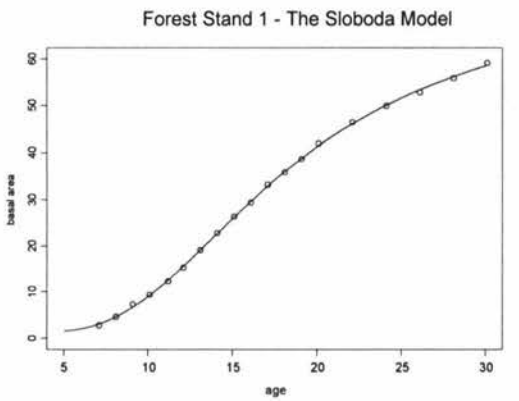


Figure 4-10: The Stand 1 raw data and the Sloboda fitted model.

4.6 Residual sum of squares

The residual sum of squares is tabulated for each model and forestry stand in Table 4-1. Even a cursory inspection of the table indicates the quite poor performance in fitting the Weibull model to the data. An ordering of the seven growth curves in terms of the mean residual sum of squares over the 22 stands is given in Table 4-2.

Table 4-1: The residual sum of squares of seven growth curves fitted to 22 forestry stands.

Stand	Hossfeld	Schumacher	Chapman	Weibull	Levakovic	Gompertz	Sloboda
1	3.32	5.95	2.59	14.14	1.98	3.67	1.89
2	3.47	6.72	2.93	9.71	2.96	4.06	3.47
4	1.63	3.40	0.89	5.88	0.96	1.84	1.18
5	1.07	1.96	0.57	4.56	0.41	1.11	0.34
6	5.04	7.26	4.38	12.93	4.11	7.01	4.44
7	2.70	6.82	1.94	13.87	1.69	3.19	1.93
8	7.33	9.35	6.71	20.07	5.74	9.06	5.81
9	2.15	1.95	1.10	8.21	0.64	2.18	0.56
10	2.93	9.88	2.37	9.63	2.64	3.01	3.38
11	1.83	7.30	1.87	6.21	1.73	1.96	2.00
12	3.18	2.88	1.96	9.35	1.42	3.66	1.34
13	8.06	11.66	7.51	18.56	7.14	8.95	7.42
14	3.55	4.85	2.50	10.88	2.25	4.44	2.42
15	3.13	6.15	2.79	8.54	2.66	3.49	2.95
16	1.77	1.74	1.04	5.22	0.79	2.11	0.77
17	7.26	17.56	6.67	12.95	7.25	6.75	8.51
18	4.06	5.27	3.66	13.45	2.91	5.82	3.09
19	2.15	5.49	2.05	3.80	2.14	2.15	2.95
21	7.51	2.26	6.05	17.55	2.24	11.31	2.24
22	3.22	10.77	3.82	3.16	2.59	2.82	4.53
23	1.89	3.02	1.19	6.80	0.99	2.42	1.04
24	5.18	1.71	3.45	12.75	1.45	7.41	1.45

Table 4-2: Mean residual sum of squares of seven fitted growth curves fitted to 22 forestry stands in increasing order.

Levakovic	Sloboda	Chapman	Hossfeld	Gompertz	Schumacher	Weibull
2.58	2.90	3.09	3.75	4.47	6.09	10.37

Other than the Weibull model and to a lesser extent the Schumacher model (with relatively large mean residual sum of squares) the residual sums of squares are reasonably similar among models. It seems that the Levakovic, Sloboda and Chapman-Richards models have the more

satisfactory fit according to this criterion. An anomaly is the fit to Stand 22 data where the Weibull model outperforms the Hossfeld, Schumacher, Chapman-Richards and Sloboda models. To understand why this is occurring, plots of the Weibull and Sloboda models for Stand 22 follow in Figures 4-11 and 4-12.

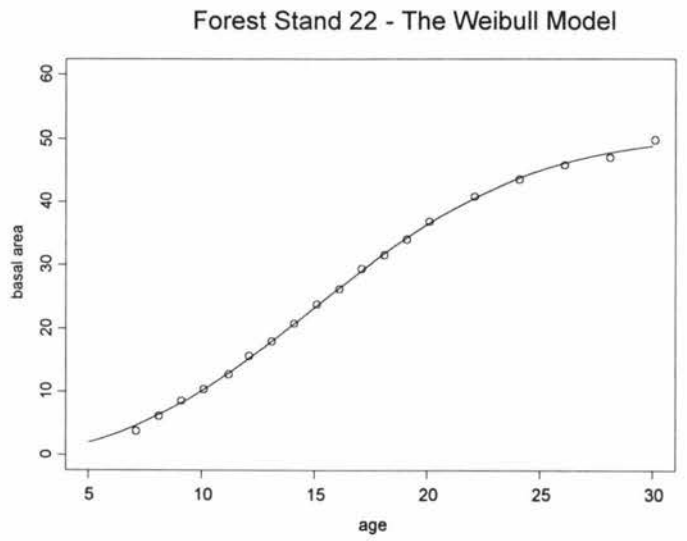


Figure 4-11: The Weibull model fitted to the data of Stand 22.

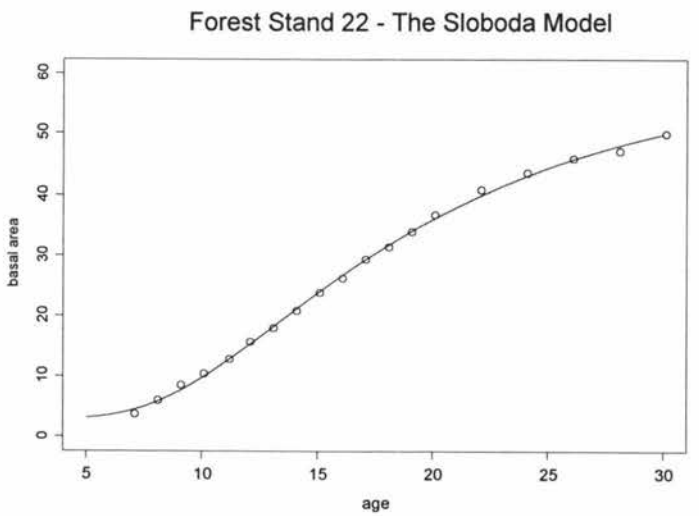


Figure 4-12: The Sloboda model fitted to the data of Stand 22

Compared to the data for Stand 1 (see Figure 4-3), Stand 22 has a smaller slope at the point of inflection (Stand 1 has estimated slope of 3.5 and that of Stand 22 is 2.9) and appears to have a smaller asymptote. The Weibull model describes this behaviour well.

Plots of residual sum of squares versus slope at the point of inflection are shown in Figure 4-13 and provide some interesting information concerning the goodness-of-fit in relation to the slope at the inflection point. It is evident that the Weibull model has poor goodness-of-fit when the slope at the inflection point is relatively large. The Chapman-Richards, Hossfeld, Levakovic and Sloboda models exhibit a very similar structure. For all models, there seems to be some indication of increasing residual sum of squares as the slope at the point of inflection increases but is appears to be much weaker than the relationship for the Weibull model.

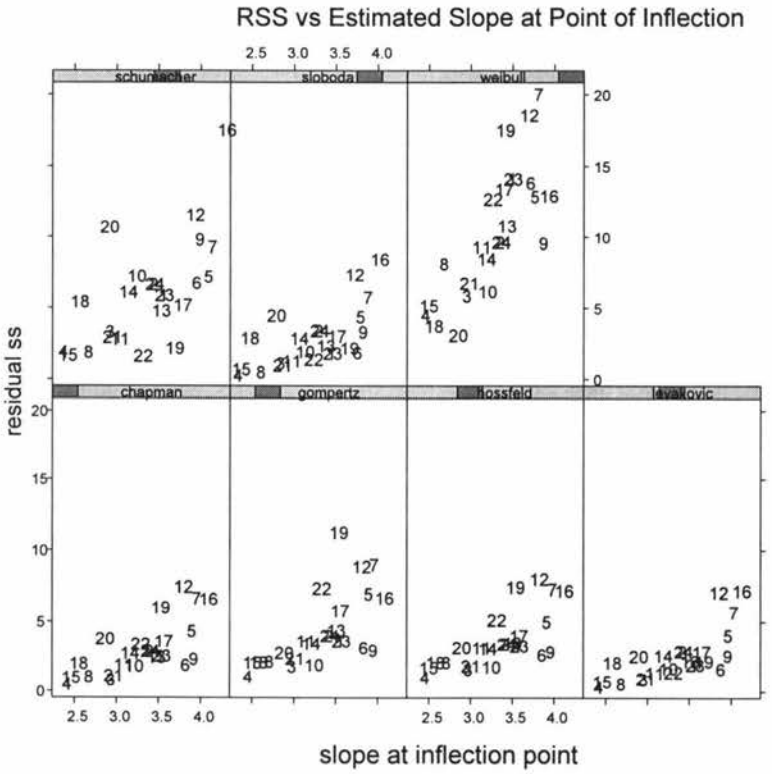


Figure 4-13: Residual sum of squares versus slope at the point of inflection for seven growth curves fitted to 22 area data of 22 forest stands. The labels indicate stand number.

Also of interest is a possible relationship between the asymptote level and residual sum of squares and this information is displayed in Figure 4-14. The Weibull model predicts a relatively

narrow range of asymptotes and a strong relationship with residual sum of squares is evident. There also appears to be an increasing relationship between asymptote level and residual sum of squares for the Chapman-Richards, Gompertz and Hossfeld models although not nearly as strong as for the Weibull model.

The panels for the 4-parameter Sloboda and Levakovic models indicate a wide range of estimated asymptotes while maintaining small residual sums of squares. Of the 3-parameter models, the Hossfeld and Chapman-Richards models are very similar and appear satisfactory although the larger asymptotes predicted by the 4-parameter models are not evident. The Schumacher model is notable for the lack of small asymptotes predicted.

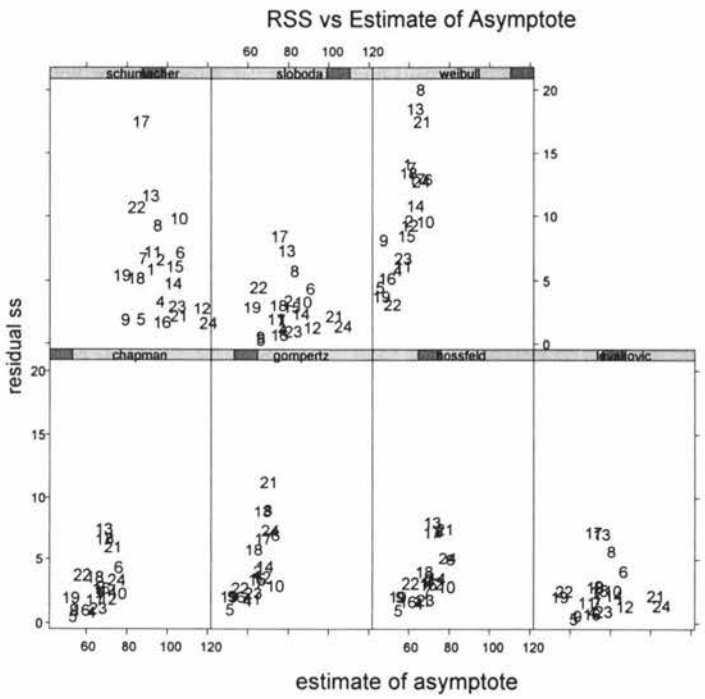


Figure 4-14: Residual sum of squares versus asymptote level for seven growth curves fitted to area data of 22 forest stands. The labels indicate stand number.

4.7 The residual structure

Introduction

The simulations that will follow in later chapters of this thesis require data that is consistent with that given by real forestry plots. In particular the residual structure needs to be examined carefully to determine whether any special patterns are present that should be incorporated into a simulation. Also, it will be instructive to determine whether one or more models display an error structure different to the others. The relationship of the residuals to the slope of the fitted model, to the fitted values and to age will be discussed in this section.

Residuals versus slope

For each model, the residuals obtained from the fitted nonlinear models are plotted (shown in Figure 4-15) against the estimated slope of the fitted growth curve.

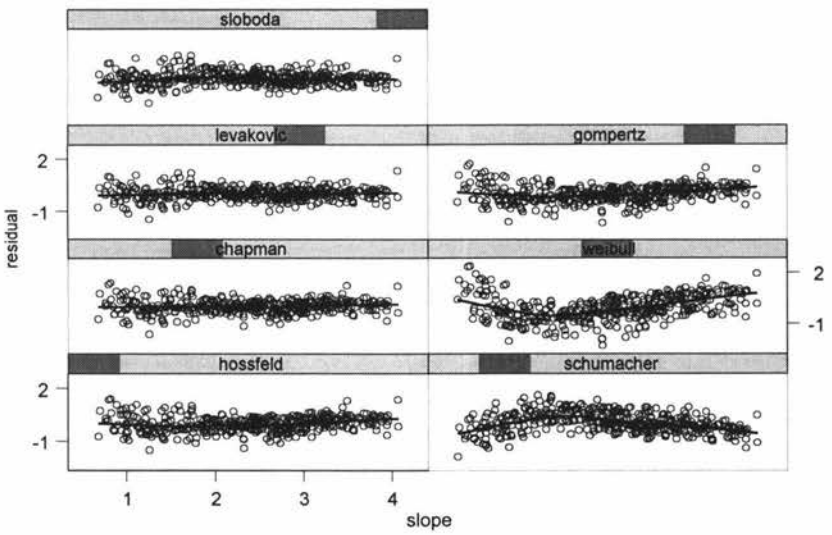


Figure 4-15: Residuals against estimated slope for all models and stands

A local regression model has been fitted in each panel to more clearly show the presence or absence of pattern in the residual plots. It is clear that there is little pattern for the Hossfeld, Chapman-Richards, Levakovic, Sloboda and Gompertz models. Two models appear quite poor:

the Weibull model indicates an error structure that is concave up and the Schumacher model one that is concave down.

It will be instructive to examine the structure of the panels above within each stand for both the Weibull (Figure 4-16) and Schumacher (Figure 4-17) models. For comparison, the Hossfeld model, a model with little curvature, will also be presented in Figure 4-18.

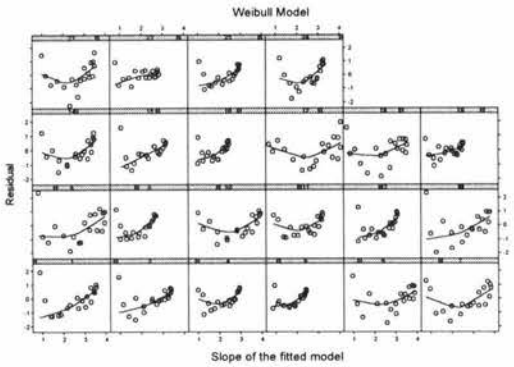


Figure 4-16: Residuals versus slope for the Weibull model fitted to 22 stands.

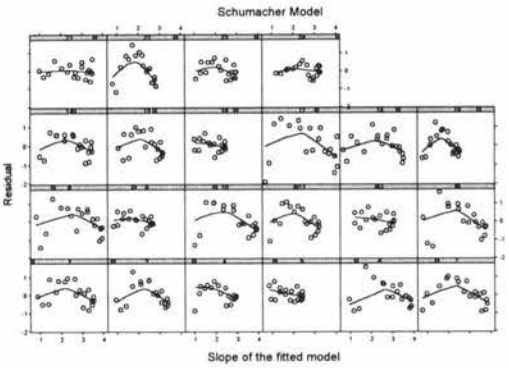


Figure 4-17: Residuals versus slope for the Schumacher model fitted to 22 stands.

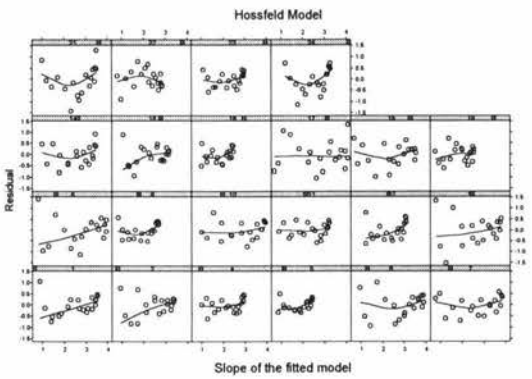


Figure 4-18: Residuals versus slope for the Hossfeld model fitted to 22 stands.

It is clear that the residual plots of the Hossfeld model are much less patterned than those of the Weibull or Schumacher models. The curvature evident for the Weibull model is concave up indicating that the model under-estimates where the slope of the curve is both small and large and over-estimates where the curve has intermediate slopes. For the Schumacher model, the converse occurs. That is, over-estimation where the slope is either small or large and under-estimation between these extremes of slope. While these points will be illustrated further in

the following section where plots of residuals against fitted values will be displayed, the residual structure discussed here can be seen in the plot of the raw data and fitted models for Stand 11 displayed in Figures 4-19 and 4-20.

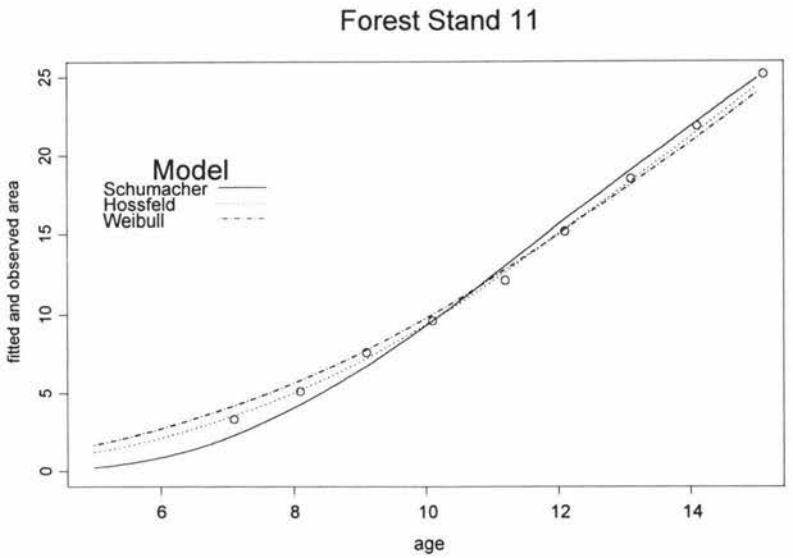


Figure 4-19: The fitted models (Stand 11) together with the raw data displayed for ages from 0 to 15 years.

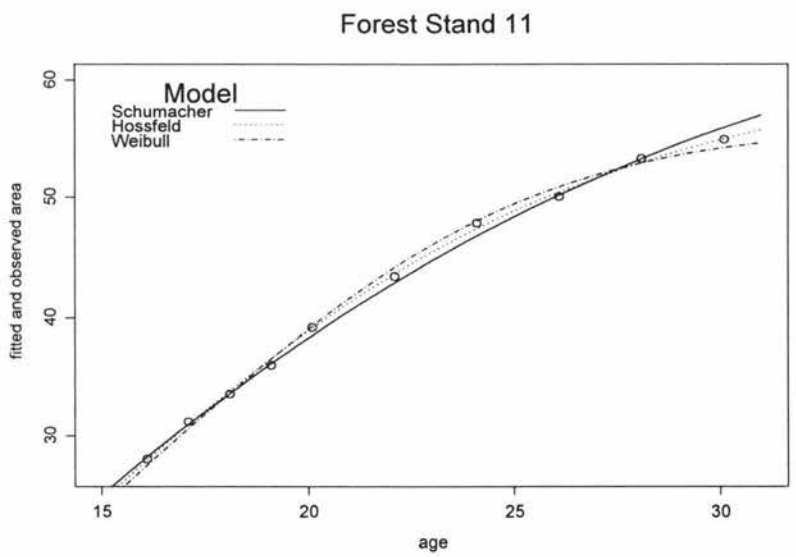


Figure 4-20: The fitted models (Stand 11) together with the raw data displayed for ages from 15 to 30 years.

At ages less than about 10 years, the underestimation of the Schumacher model and overestimation of the Weibull model is clearly seen in Figure 4-18. The roles are reversed near the inflection point (near an age of 12 to 15 years) and then again near an age of 20 years. From ages greater than about 27 years, the Schumacher model continues to run high and the Weibull model runs low.

Plots of residuals versus slope for the other models can be found in Appendix C; they are quite similar in structure to the Hossfeld model and thus display little relationship between residual and slope.

Residuals versus fitted values

Further investigation of structure in the residuals is carried out through examination of possible relationships between the residuals and the fitted values for each growth curve. A plot of these is displayed in Figure 4-21.

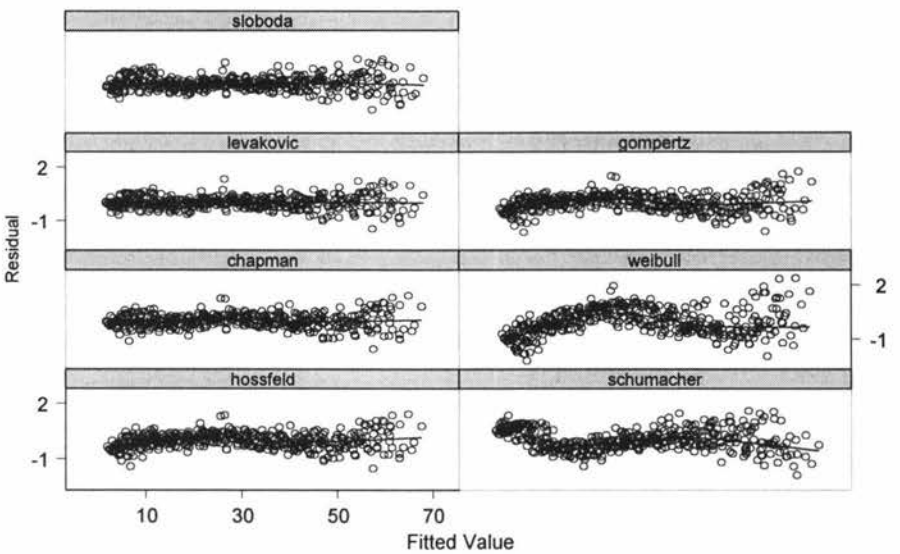


Figure 4-21: Residuals plotted against fitted values for seven growth curves fitted to 22 forestry stands.

The plots of the 4-parameter Levakovic and Sloboda models are remarkably similar to those of the 3-parameter Chapman-Richards and Hossfeld models, all of which display little curvature.

The curvature of the plots for the Gompertz, Weibull and Schumacher models indicate that these are not fitting as well as the others. All models show some "fanning out" of residuals at the high end of the fitted values but this is less severe for the four models mentioned initially. Greater detail can seen by examining individual plots for each stand/method combination. These plots are shown in Appendix D and confirm the 4-parameter models as having the more nearly patternless structure. Of the 3-parameter models the Chapman-Richards looks best with the Hossfeld function less impressive when the individual stands are considered; this plot is displayed in Figure 4-22.

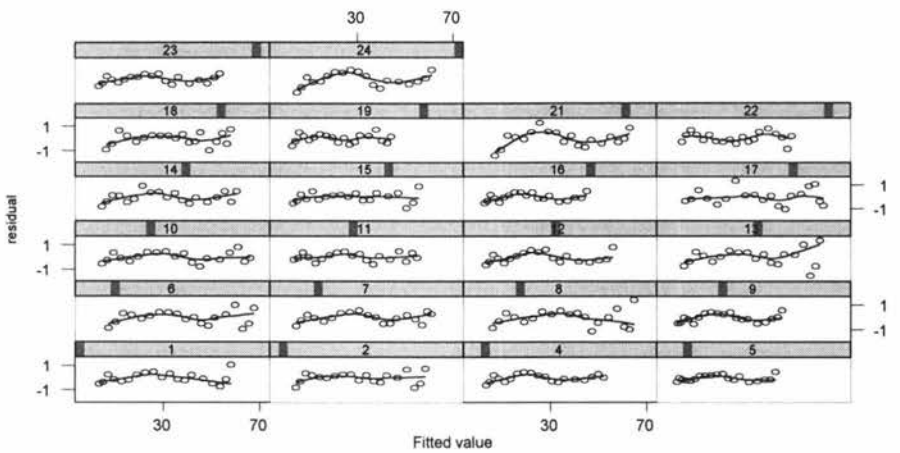


Figure 4-22: Residuals versus fitted values for the Hossfeld model fitted to each of 22 forestry stands.

Residuals versus age

Plots of residuals against age, Figure 4-23, are discussed to provide further understanding of both the fit and the residual structure associated with each sigmoid model.

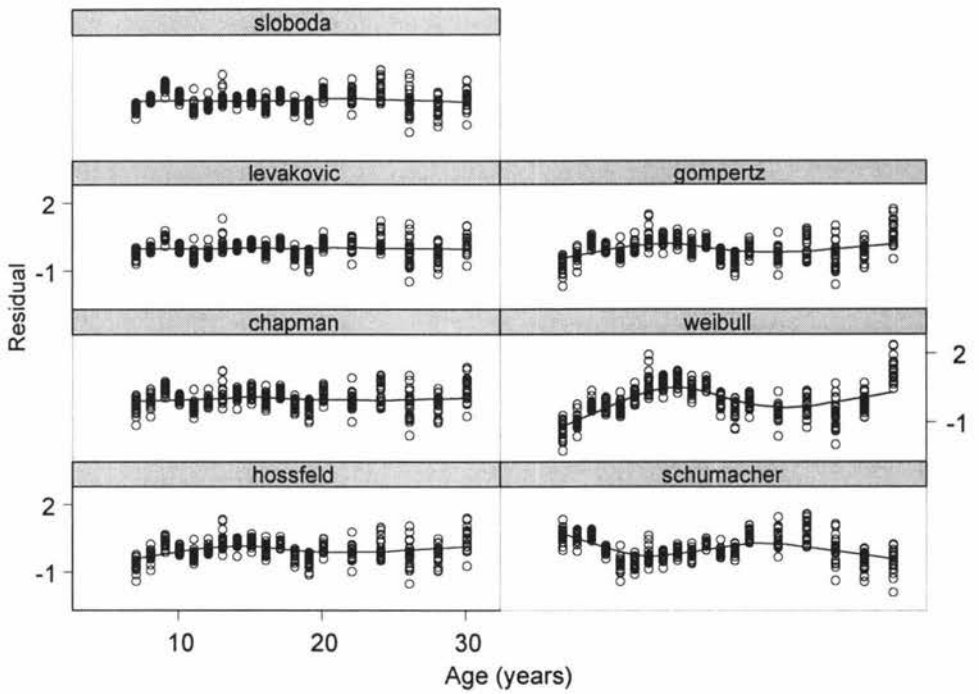


Figure 4-23: Residuals plotted against age for seven growth curves fitted to 22 forestry stands.

As in the previous section, the 4-parameter models (Sloboda and Levakovic) look good and the best of the 3-parameter models is the Chapman-Richards. Each of these models have relatively small residuals at high values of age and a reasonably patternless display is evident. High levels of curvature and/or large residuals at age 30 are apparent for the other models.

Normality of Residuals

Normal plots of the residuals within each fitted model and histograms of the residuals are shown in Figures 4-24 and 4-25.

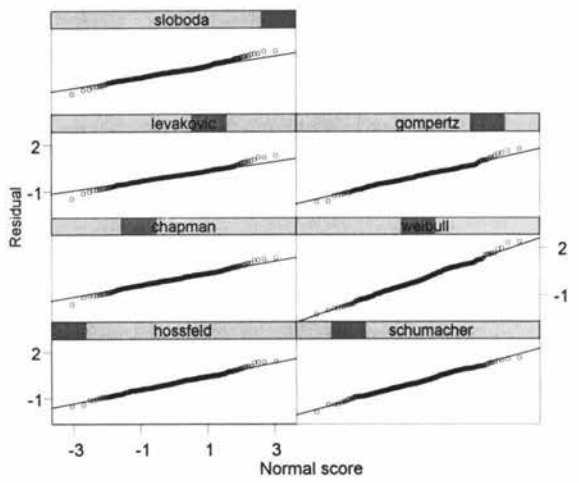


Figure 4-24: Normal probability plots of the residuals of seven growth curves each fitted to basal area data of 22 forestry stands.

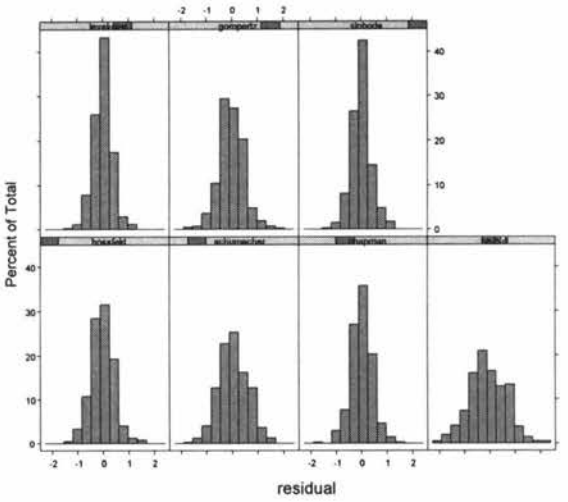


Figure 4-25: Histograms of the residuals of seven growth curves each fitted to basal area data of 22 forestry stands.

The initial impression is that violations of normality, if present, are not too severe. The normal plots look reasonable and the histograms are reasonably symmetric about their means but with shorter tails than normal in a few instances. A more formal test of normality using a Kolmogorov-Smirnov goodness-of-fit test was used to assess the underlying normality of the residuals generated by each model. The p-values are presented in Table 4-3 and indicate some evidence of non-normality for the Schumacher and Levakovic models and very strong evidence of non-normality for the Sloboda model; this is seen to be caused by the tighter than usual tails. The nonlinear model assumes residuals that are $N\left(0,\sigma^2\right)$ and violations of this do not appear serious.

Table 4-3: P-values of Kolmogorov-Smirnov Test[1] applied to the residuals of seven growth curves fitted to 22 forestry stands.

Chapman	Gompertz	Hossfeld	Schumacher	Sloboda	Weibull	Levakovic
0.5[1]	0.5[1]	.5[1]	0.04	0.0005	0.5[1]	0.03

[1] S-plus uses the Dallal-Wilkinson approximation to calculate the p-value in testing composite normality (normality when estimated parameters are not supplied) and is most accurate for p-values ≤ 0.10. When the calculated p-value is greater than 0.1 it is reported as 0.5.

4.8 The ratio of asymptote level to the y-coordinate of the inflection point

We will now examine the ratio $y_{\text{asy}}/y_{\text{inf}}$ (the ratio of the asymptote level to the y-coordinate of the point of inflection) that was discussed extensively in Chapter2. We are particularly interested in how this quantity varies across growth curves. These ratios are calculated using the properties of the sigmoid functions discussed in Chapter 2 and the parameter estimates tabulated in Appendix B. For example, for the Chapman-Richards fitted to stand 1

$$\frac{y_{\text{asym}}}{y_{\text{inf 1}}} = \frac{c^c}{(c-1)^c} = \frac{6.48^{6.48}}{(6.48-1)^{6.48}} = 2.96$$

The values of $y_{\text{asy}}/y_{\text{inf}}$ for each of the seven growth curves fitted to the 22 forestry stands are shown in Figure 4-29. The Schumacher model always gives the largest ratio as we would suspect by recalling that the plots of fitted models always showed the Schumacher model with the highest asymptote level. The Sloboda and Levakovic models provide the next largest ratios and it was evident in Section 4.2 that the second and third largest asymptotes were generated by these models. The Weibull ratio is always smallest. We summarise this order in Table 4-4 and note that the ordering of the 3-parameter models is in agreement with the discussion of Chapter 2.

Table 4-4: The ratio $y_{\text{asy}}/y_{\text{inf}}$ in descending order

Schumacher
Levakovic
Sloboda
Chapman-Richards
Hossfeld
Gompertz
Weibull

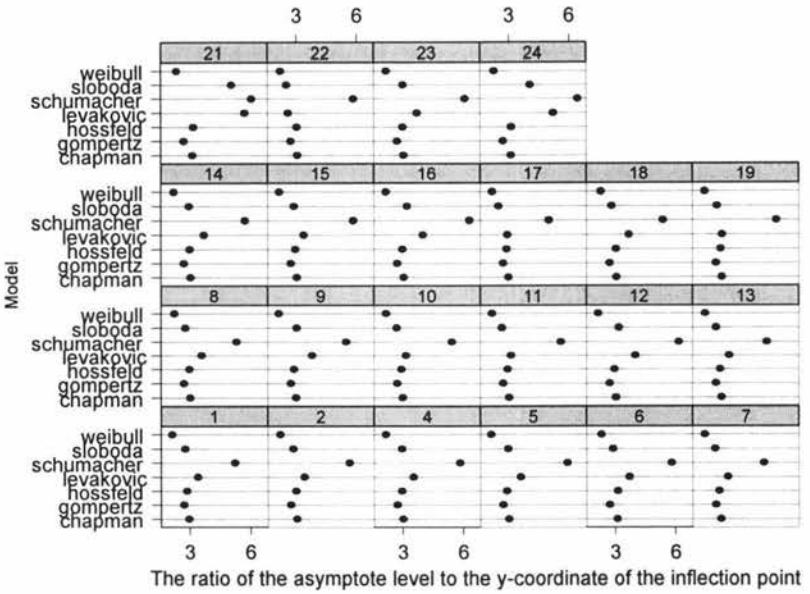


Figure 4-29: Ratio of estimate of asymptote and estimate of the y-coordinate of inflection point

4.9 Parametric-effects curvature

For each model, curvature measures are calculated using Venables' & Ripley's, S-Plus **rms.curv()** function and we recall that the curvature measure $c^{\theta}\sqrt{F}$ is considered to be acceptable if it is less than 0.3. The intrinsic curvature of all models is at acceptable levels and is not discussed here although these are tabulated in Appendix F.3. Parametric-effects curvatures are tabulated in Appendix F.2. As the general pattern is similar for most stands, two plots are presented here, in Figure 4-30 for Stands 1 and 21. Stand 1 is typical and Stand 21 is somewhat unusual.

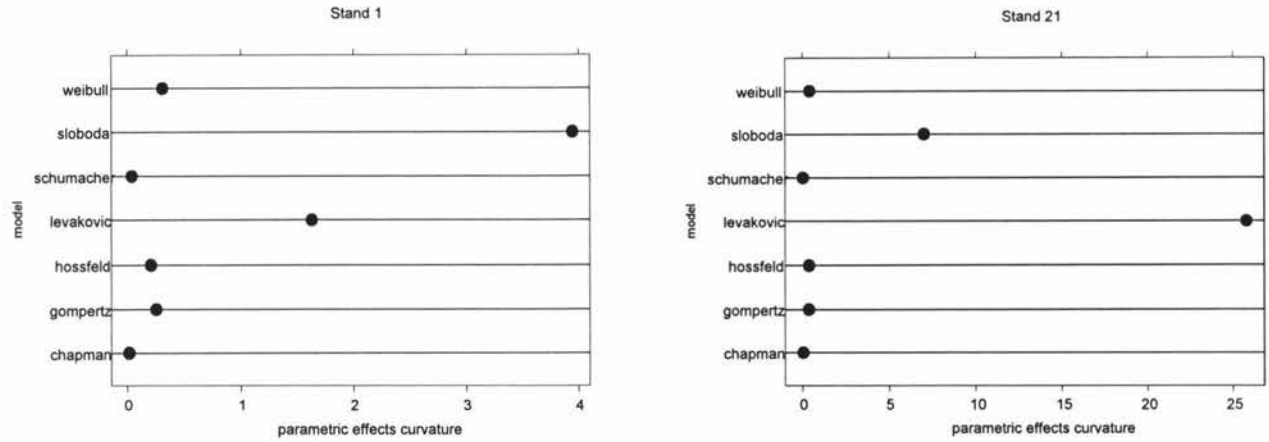


Figure 4-30: Curvature of seven growth curves fitted to basal area data of Stands 1 and 21

Other than for the 4-parameter models, parametric-effects curvature is mostly maintained within the $c\sqrt{F} < 0.3$ bound. The largest curvatures (with one exception being Stand 21) are shown by the Sloboda model and in all cases markedly exceed the 0.3 level for $c\sqrt{F}$. The curvatures of the Levakovic models are also unacceptably high but usually not as extreme as that of the Sloboda model. The Weibull model has $.3 < c\sqrt{F} < .4$ for several stands and both the Gompertz and Hossfeld models have one or two instances of $.3 < c\sqrt{F} < .37$. Both Chapman and Schumacher models are notable for having $c\sqrt{F} < .07$ in all cases indicating very good estimation properties.

The Levakovic and Sloboda models were sometimes difficult to fit and the Levakovic model was particularly sensitive to starting values. On the other hand, the 3-parameter models (reparameterized where necessary) were very simple to fit with little work required for the non-linear least squares algorithm to converge.

4.10 Conclusion

The performance of each model is summarised in Tables 4-4 and 4-5 for each of the qualities discussed in this chapter. On the basis of the information contained in the table it seems clear that the Chapman-Richards and Hossfeld models are reasonably adequate in all criteria and would be adequate candidate functions to model similar data. Of the 4-parameter models, the Levakovic model outperforms the Sloboda model with the only concern being the parametric-

effects curvature with consequent loss of precision in parameter estimates.

Table 4-4: Summary of main results obtained when five 3-parameter sigmoid models are fitted to basal are measurements of 22 forestry stands.

	Chapman	**Hossfeld**	**Gompertz**	**Schumacher**	**Weibull**
Goodness-of-fit	Good	Good	Fair	Poor	Poor
Asymptote level	Variety	Variety	Variety	Always large	Small
Inflection point	Variety	Variety	Variety	Small	Large
Residuals vs slope	Patternless	Slight pattern	Patterned	Patterned	Patterned
Residuals vs fits	Little pattern	Little pattern	Patterned	Some pattern	Patterned
Residuals vs age	Little pattern	Little pattern	Patterned	Patterned	Patterned
Residual normality	Satisfactory	Satisfactory	Satisfactory	Unsatisfactory	Satisfactory
Curvature	Very good	Satisfactory	Satisfactory	Very good	Satisfactory

Table 4-5: Summary of main results obtained when two 4-parameter sigmoid models are fitted to basal area measurements of 22 forestry stands.

	Levakovic	**Sloboda**
Goodness-of-fit	Very good	Very good
Asymptote level	Variety	Variety
Inflection point	Variety	Large
Residuals vs slope	Patternless	Patternless
Residuals vs fits	Little pattern	Little pattern
Residuals vs age	Little pattern	Little pattern
Residual normality	Unsatisfactory	Unsatisfactory
Curvature	Unsatisfactory	Unsatisfactory

It will be useful to present an ordering of the models in terms of their overall suitability in modelling basal area data similar to the 22 stands of pinus radiata data. From best to worst the models are listed as

Chapman-Richards
Levakovic
Sloboda
Hossfeld
Gompertz
Schumacher
Weibull

The reader may question the ordering of the top four models and it is agreed that it is often not as clear-cut as this. The Levakovic model is quite sensitive to starting values and on occasions will not converge easily. There is however little doubt about the superiority of the first four models over the Gompertz, Schumacher and Weibull models. In Part IV we will investigate the accuracy of making predictions of basal area. This ordering may need modification as a result of this work.

Part III

The Growth Model with the Largest Asymptote

Chapter 5

The Asymptote Level of the Seven Growth Curves

5.1 Introduction

In Chapter 2 we investigated the ratio $y_{\text{asy}}/y_{\text{inf}}$ and concluded that of the 3-parameter models the Schumacher model will have the largest value of this ratio. As a result of fitting the growth curves to 22 stands of pinus radiata basal area data we determined that the Schumacher model consistently gave higher estimates of asymptote level than the other six models. In this chapter we will further investigate the asymptote levels determined by the growth curves by examining the dependence of asymptote level on the slope of the growth curve at the point of inflection. We will empirically show that the Schumacher growth curve will always give higher asymptote estimates when fitted to forestry data. A simulation study (the S-Plus script used to run the simulation is listed in Appendix I) is developed as follows:

1. We start with sigmoid curves with slope m at the inflection point. Equations for the asymptote in terms of m and the parameters are determined. For example, the equation for the asymptote of the Hossfeld growth curve can be shown to be

$$a = \left(\frac{4bcm}{(c+1)^2}\right)^{\frac{c}{c-1}} \frac{(c+1)}{b(c-1)}$$

A closed form expression cannot be determined for the Sloboda model. A numerical method (the Newton-Raphson method) is used to first find the t-coordinate of the inflection point and the asymptote is then calculated . The equations of the asymptote are listed in Section 5.2.

2. To simulate values of the asymptote we must choose values of the parameters. We use the parameter estimates found when the sigmoid models were fitted to the Kaingaroa forest pinus radiata data (tabulated in Appendix B). The mean, standard deviation and correlation between parameter estimates, within the 20^{th} and 80^{th} percentiles[1], are calculated and summarised in Tables 5-1 and 5-2. It must be noted that the value of the asymptote level of the Levakovic model is very sensitive to the values of the parameters, and generation of these outside the 20^{th} and 80^{th} percentiles can lead to undefined mathematical operations.

Table 5-1: Mean and standard deviation of the parameter estimates determined when the sigmoid models were fitted to pinus radiata basal area data.

	b		c		d	
	mean	sd	mean	sd	mean	sd
Chapman-Richards	−.123	.00536	5.42	.234		
Gompertz			.144	.00632		
Hossfeld	109	23.1	3.08	.0553		
Schumacher	54.2	7.18	1.38	.0771		
Weibull	−.000531	.000135	2.61	.0519		
Levakovic	6.62	.766	1.60	.218	2.49	.217
Sloboda	−4.08	.279	76.7	15	−2.02	.112

[1] These percentiles were chosen for two reasons: firstly, to ensure that the most commonly occurring parameter values were used and secondly, for most models the parameter estimates were reasonably normally distributed within these percentiles.

Table 5-2: Correlations between the parameter estimates determined when the sigmoid models were fitted to pinus radiata basal area data. The asymptote of the Gompertz model depends only on m and one parameter c so the correlation between parameters is not required.

	Correlation		
	ρ_{bc}	ρ_{bd}	ρ_{cd}
Chapman-Richards	$-.54$		
Hossfeld	.76		
Schumacher	.95		
Weibull	.92		
Levakovic	$-.57$	.99	$-.45$
Sloboda	.24	$-.60$	$-.80$

3. We use the equations (listed in Section 5.2) and parameters developed above and $m \in \{2, 2.5, 3, 3.5, 4, 4.5, 5\}$ to generate 1000 asymptotes for each model at each slope m. The parameters, with the appropriate correlations, were generated using the S-Plus multivariate-normal random number generator, **rmvnorm()** and the S-Plus script used to run the simulation is listed in Appendix I.

5.2 The equations for the asymptote

Routine calculations using the properties of the models discussed in Chapter 2 give the following equations for the asymptote level in terms of the slope of the curve at the point of inflection and the parameters. It is noted that other than the Hossfeld model these are linear in m.

Chapman-Richards	$a = \frac{-m}{b\left(\left(1-\frac{1}{c}\right)^{c-1}\right)}$
Gompertz	$a = \frac{em}{c}$
Hossfeld	$a = \left(\frac{4bcm}{(c+1)^2}\right)^{\frac{c}{c-1}} \frac{c+1}{b(c-1)}$
Schumacher	$e^a = \frac{m}{bc\left(\left(b\frac{c}{1+c}\right)^{-\frac{1+c}{c}}\right)} e^{\frac{1}{c}+1}$
Weibull	$a = -\frac{me^{\frac{c-1}{c}}\left(\frac{1-c}{cb}\right)^{\frac{1-c}{c}}}{bc}$
Levakovic	$a = \frac{mb^{\frac{1}{d}}d^c(c+1)^{c+1}}{c(cd-1)^{c-\frac{1}{d}}(d+1)^{1+\frac{1}{d}}}$
Sloboda	$ae^{-b} = \frac{me^{ct^d}e^{be^{-ct^d}}}{bcdt^{d-1}} e^{-b}$

For the Sloboda model $y = ae^{-be^{-ct^d}}$ the t-coordinate of the inflection point is the solution to $d - 1 - cdt^d + bcdt^d e^{-ct^d} = 0$. A closed form expression for t, in terms of b, c, d cannot be obtained from this equation. We solve $d-1-cdt^d+bcdt^de^{-ct^d} = 0$ numerically to find t for each (b, c, d) and then substitute into the expression for the asymptote. Using Newton's method, the iterative scheme to determine t is

$$t_{n+1} = t_n - \frac{d - 1 - cdt_n^d + bcdt_n^d e^{-ct_n^d}}{-cd^2t_n^{d-1} + bcd^2t_n^{d-1}e^{-ct_n^d}\left(1 - ct_n^d\right)}$$

and with $t_0 = 10$ this converges quickly with the stopping criterion set to $|t_{n+1} - t_n| < 10^{-6}$.

5.3 Simulated aymptote levels

As stated in the introduction, 1000 asymptotes were randomly generated for each slope. The statistics of the asymptotes for slopes two and five at the point of inflection are presented in Tables 5-3 and 5-4 arranged in order of increasing mean asymptote level. These clearly indicate that the Schumacher function will tend to yield a greater asymptote than the other models.

Table 5-3: Basic statistics (ordered by mean asymptote level) for the asymptote level when the slope is two at the inflection point

Model	Mean	Min	LQ	Median	UQ	Max	S.D.
Hossfeld	33.1	20.8	31.6	33.0	34.6	43.4	2.4
Weibull	34.9	30.7	33.7	34.5	35.7	46.8	1.9
Gompertz	37.9	33.3	36.7	37.7	39.0	43.3	1.7
Chapman-Richards	40.1	35.3	39.1	40.1	41.1	45.4	1.6
Levakovic	43.0	39.1	42.2	43.0	43.8	45.8	1.1
Sloboda	47.3	29.3	41.9	46.6	51.4	82.2	7.4
Schumacher	57.6	41.4	53.1	57.5	61.5	79.8	6.3

Table 5-4: Basic statistics (ordered by mean asymptote level) for the asymptote level when the slope is five at the inflection point

Model	Mean	Min	LQ	Median	UQ	Max	S.D.
Weibull	87.1	77.6	84.1	86.3	89.1	126.0	4.7
Gompertz	94.5	82.9	91.6	94.4	97.2	111.0	4.2
Chapman-Richards	100.2	90.0	97.3	100.1	102.8	114.6	4.1
Levakovic	107.6	98.4	105.8	107.7	109.5	114.7	2.8
Sloboda	118.0	78.5	105.3	116.6	128.5	193.1	17.8
Hossfeld	129.2	101.1	122.1	129.1	135.4	163.1	9.6
Schumacher	142.8	95.4	131.9	142.3	153.4	202.1	15.4

The generated asymptotes plotted against slope at the point of inflection for each model are displayed in Figure 5-1.

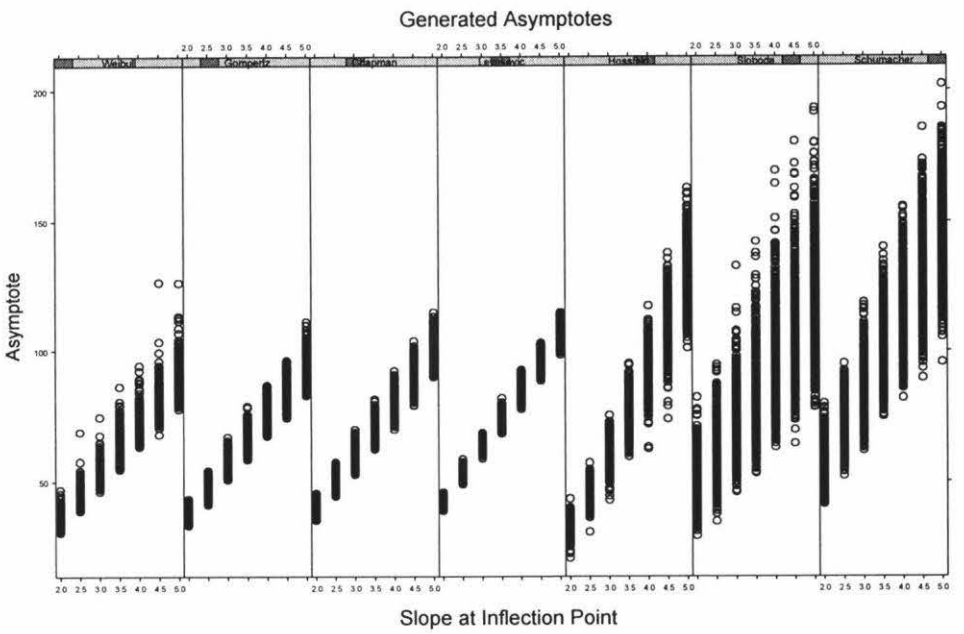


Figure 5-1: One thousand simulated asymptotes for each growth curve with slopes at intervals from 2 to 5 at the point of inflection.

It is clear that Sloboda and Schumacher models tend to generate larger asymptotes than the other four models and it is also evident that there is much greater variation in the size of the asymptote. It is the case for a given data set that a large asymptote estimated by one model tends to mean that a large asymptote will be estimated by other models. It is thus of interest to investigate the relationship between the asymptotes given by the Schumacher model and those of the other models. We regress the asymptote levels of the Schumacher model against those of the other models. A graph of the Schumacher asymptotes plotted against the Chapman-Richards asymptotes is shown in Figure 5-2 and this confirms that relationship is reasonably linear.

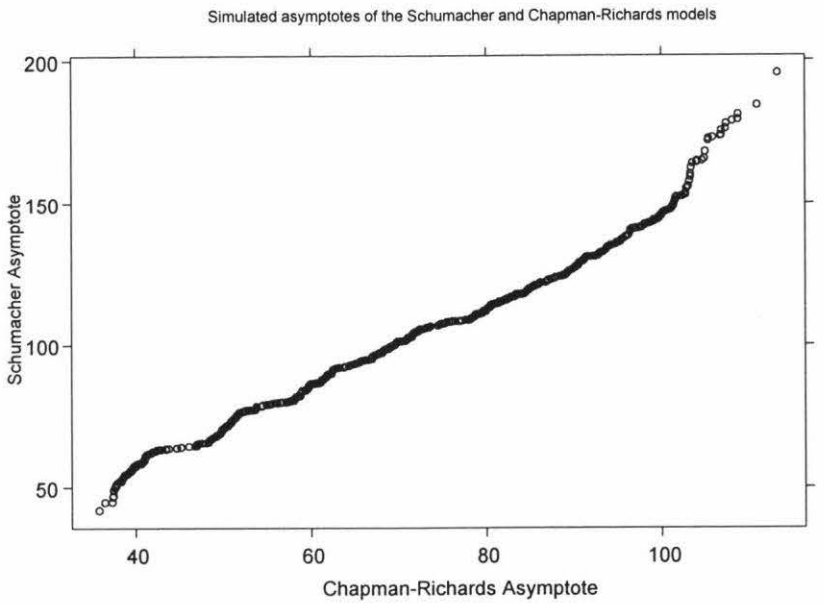


Figure 5-2: The simulated asymptotes of the Schumacher model plotted against those of the chapman-Richards model.

The slopes of the six regression lines are summarised in Table 5-5.

Table 5-5: The slope of the line when the Schumacher asymptote is regressed against the asymptotes of the other models. For example, the Schumacher asymptote is on average 1.5 times greater than the Gompertz asymptote.

Sloboda	Hossfeld	Levakovic	Chapman-Richards	Gompertz	Weibull
1.2	1.2	1.3	1.4	1.5	1.7

This clearly gives an ordering of asymptote level from largest to smallest as Schumacher, Sloboda, Hossfeld, Levakovic, Chapman-Richards, Gompertz and Weibull and confirms the asymptote ordering of the 3-parameter models discussed in Section 2.3.

The ordering of asymptote level when the models were fitted to the pinus radiata data is the same as obtained here, except that the position of the Levakovic and Hossfeld models are interchanged. Although the position of the Levakovic model in the ordering remains in some doubt, it is quite clear that the larger asymptotes will be determined by the Schumacher and Sloboda models and very small asymptotes by the Weibull and Gompertz models.

Part IV

Robustness of the Growth Curve Models

Chapter 6

Predicting Basal Area Using Real Data

6.1 Introduction

We will compare the prediction accuracy of the seven growth models by refitting the models to a subset of the data. For each stand[1], the basal area recorded at the first 17 ages (up to 26.1 years) is used to fit the models and these models employed to predict known basal areas at 28.1 and 30.1 years. The predictions were found using the S-Plus library function **predict.nls()** written by Professor Douglas Bates and included in the Unix version of the Venables & Ripley MASS library[2]. Points of major interest are summarised and discussed here with full results, including standard errors of prediction, tabulated in Appendix H.

6.2 Analysis of prediction errors

The absolute error of prediction

$$\text{prediction error} = |\text{actual basal area} - \text{predicted basal area}|$$

[1] The Levakovic function could not be fitted to the subset of Stand 21 basal area data.
[2] MASS is a library of functions that supplements Modern Applied Statisics with S-Plus.

was determined for each stand-model combination and a rank of 1 to 7 assigned to the prediction error in each stand with rank 1 given to the smallest absolute error of prediction. The results obtained are summarised in Table 6-1 (with mean rank in the last row of the table).

Table 6-1: Seven growth models are used to make predictions at two post-data ages for 22 stands of basal area data. Within each stand and prediction age, ranks of 1 to 7 are assigned to the prediction error with a rank of 1 given to the smallest absolute prediction error. The models are ordered from left to right by decreasing prediction accuracy (increasing mean rank).

Rank	Levakovic	Chapman	Sloboda	Hossfeld	Gompertz	Schumacher	Weibull
1	9	2	14	4	6	5	4
2	11	9	9	7	5	2	1
3	7	18	3	8	4	4	0
4	9	14	1	11	2	5	2
5	6	1	10	14	10	3	0
6	0	0	7	0	17	11	9
7	0	0	0	0	0	14	28
Mean	3	3.1	3.1	3.5	4.3	5	6

From Table 6-1 the important results can be summarised as follows:

1. Of the 3-parameter models the Chapman-Richards model has 66% of its predictions at rank 1, 2 or 3 and the Hossfeld model has 43% at rank 1, 2 or 3. With average ranks of 3.1 and 3.5, the Chapman-Richards and Hossfeld models are the most successful of the 3-parameter models. It is noted that neither of these models give any ranks of 6 or 7.

2. Of the 4-parameter models, the Levakovic has 48% at rank 1 or 2 and the Sloboda model 52% within these ranks. The Levakovic however, has only 14% in ranks 5 to 7 compared with 39% for the Sloboda model.

3. The Weibull, Gompertz and Schumacher models are very poor predictors. The Weibull model has 84% of its predictions at ranks 6 and 7; the Gompertz model has 61% of its predictions at ranks 5 and 6 although it does not yield any rank 7 predictions; the Schumacher model has 57% of its predictions at ranks 6 and 7.

The root mean square prediction error for each growth model is also calculated as

Chapman	Levakovic	Sloboda	Hossfeld	Gompertz	Schumacher	Weibull
1.02	1.05	1.06	1.19	1.40	1.82	2.78

The mean prediction error of each method is presented in the following table. These indicate that the Gompertz, Schumacher and Weibull models give biased predictions. The Levakovic and Chapman-Richards seem to be quite adequate.

Levakovic	Chapman	Sloboda	Hossfeld	Gompertz	Schumacher	Weibull
-0.11	.27	-.52	.57	.93	-1.56	2.43

Standard errors of prediction are tabulated in Appendix H. These are all relatively small with, for example, the Chapman-Richards model giving prediction standard errors typically between 0.2 and 0.5. These can be compared with the Schumacher model where the prediction standard errors are between 0.3 and 0.7.

These results tend to confirm that more accurate and less biased predictions are obtained from the Chapman-Richards, Levakovic and Sloboda models with the Hossfeld model only slightly inferior. A tentative ordering of the growth curves in terms of their accuracy of prediction can now be established and this is compared to the ordering, in terms of goodness-of-fit, given in Section 4.10. There is very good agreement between the rankings with only the Gompertz and Schumacher models swapping positions.

Prediction ordering		Goodness-of-fit ordering	
1	Chapman-Richards	1	Chapman-Richards
2	Levakovic	2	Levakovic
2	Sloboda	3	Sloboda
4	Hossfeld	4	Hossfeld
5	Gompertz	5	Schumacher
6	Schumacher	6	Gompertz
7	Weibull	7	Weibull

It is of interest whether the same goodness-of-fit ordering is preserved when the models are fitted to the early data (up to 26.1 years). The following table gives the mean residual sum of squares, in ascending order, when the models are fitted to this data.

Levakovic	Chapman-Richards	Sloboda	Hossfeld	Gompertz	Schumacher	Weibull
1.9	2.1	2.1	2.6	2.8	3.8	5.4

As can been seen, this simple ordering is reasonably similar to the previous goodness-of-fit with the Chapman-Richards and Levakovic models changing positions and the Schumacher and Gompertz models also interchanging. We see that the prediction ordering is in closer agreement with this goodness-of-fit ordering than previously when all the data was used.

The tentative ordering above may need modification if there are large differences between prediction accuracy at 28.1 years and 30.1 years. Consequently, the above results are further broken down so that results at each age can be discussed separately. Included in the discussion are the four models that appear to perform somewhat satisfactorily: the Chapman-Richards, Levakovic, Sloboda and Hossfeld models and the results are summarised in Tables 6-2 and 6-3.

Table 6-2: Prediction error ranks at age 28.1 years

	Chapman	Hossfeld	Levakovic	Sloboda
1	0	4	6	5
2	7	3	3	3
3	9	3	3	3
4	5	8	5	0
5	1	4	4	6
6	0	0	0	5
7	0	0	0	0
Mean	3	3.2	3.1	3.6

Table 6-3: Prediction error ranks at age 30.1 years

	Chapman	Hossfeld	Levakovic	Sloboda
1	2	0	3	9
2	2	4	8	6
3	9	5	4	0
4	9	3	4	1
5	0	10	2	4
6	0	0	0	2
7	0	0	0	0
Mean	3.1	3.9	2.9	2.6

From Table 6-2 and 6-3 the important results can be summarised as follows:

1. The Sloboda model is comparatively better at 30.1 years than at 28.1 years.
2. The Hossfeld model works very well at 28.1 years but is less effective at 30.1 years.
3. The Levakovic and Chapman-Richards models are reasonably consistent in their performance at 28.1 and 30.1 years.

The root mean square errors are also presented at each prediction age:

28.1 years	Chapman	Hossfeld	Gompertz	Levakovic	Sloboda	Weibull	Schumacher
	.72	.76	.81	.85	1.00	1.65	1.68

30.1 years	Sloboda	Levakovic	Chapman	Hossfeld	Gompertz	Schumacher	Weibull
	1.11	1.22	1.25	1.47	1.81	1.95	3.57

It is clear that the 3-parameter Hossfeld and Chapman-Richards models are more precise than the 4-parameter models at 28.1 years but less effective at 30.1 years. The Gompertz model seems to work reasonably well at 28.1 years but much less effectively at 30.1 years. At 30.1 years there is little difference between the Levakovic and Chapman-Richards models.

6.3 Discussion

These results indicate that the Chapman-Richards and Hossfeld models give reasonably good predictions at both 28.1 and 30.1 years. Compared with the 3-parameter models, the 4-parameter Levakovic and Sloboda models performed very well at 30.1 years and adequately at 28.1 years.

A tentative ordering of the models in terms of accuracy of prediction was given earlier and this can now be confirmed as

Ordering by prediction accuracy	
1	Chapman-Richards
2	Levakovic
2	Sloboda
4	Hossfeld
5	Gompertz
6	Schumacher
7	Weibull

6.4 Predictions using earlier truncations of the data

Following the suggestion of Dr. Richard Woollons, the data were truncated at 17.1, 18.1, 19.1, 20.1, 22.1, 24.1 and again at 26.1 years. The seven models were fitted to each stand within each reduced data set and a prediction made of basal area at 30.1 years. Root mean square prediction errors for each model/data combination are summarised in Table 6-4.

Table 6-4: Root mean square error for prediction of basal area at 30.1 years using models fitted to data up to the truncation age shown.

	Age at which the data was truncated						
	17.1	18.1	19.1	20.1	22.1	24.1	26.1
Chapman	3.6	3.2	3.5	2.5	1.9	1.4	1.3
Gompertz	4.8	5.6	5.6	4.2	3.1	2.1	1.8
Hossfeld	4.7	5.5	5.4	3.8	2.6	1.7	1.5
Schumacher	12.0	7.8	5.5	5.6	4.2	3.2	2.0
Weibull	14.1	14.1	13.0	10.2	7.3	4.7	3.6
Sloboda	6.5	2.7	2.9	3.0	2.2	1.9	1.1
Levakovic	5.6	7.0	6.4	2.7	2.2	1.7	1.2

The Chapman-Richards model provides good predictions irrespective of the age at which the data is truncated. The Sloboda model is quite stable in prediction error for data sets truncated at 18.1 years and beyond, although it is noted that convergence was not possible when fitting the Sloboda model to the data of six stands truncated at 17.1 years. Similarly, the Levakovic model provides good predictions when based on data truncated at 20.1 years and beyond but again it was often difficult to obtain convergence when the model was fitted to the truncated data and in some cases convergence was not possible. The Weibull model is comparatively very poor at all ages.

For each stand, at each truncation age, the models are ranked from 1 to 7 in terms of the absolute error of prediction with 1 assigned to the smallest absolute error. The results are summarised in Tables 6-5 to 6-11 (NA means "convergence not obtained"). In each case the models are ordered in increasing order of median rank.

Table 6-5: Ranks assigned to the absolute prediction error at 30.1 years when the growth curves were fitted to 22 basal area data sets truncated at the 17.1 years.

	Rank								
	1	2	3	4	5	6	7	NA	Median
Chapman	11	4	3	4	0	0	0	0	1.5
Gompertz	4	6	7	2	1	2	0	0	3
Hossfeld	2	8	5	5	2	0	0	0	3
Levakovic	4	2	3	2	5	0	0	6	4.5
Sloboda	0	2	4	3	6	1	0	6	5
Schumacher	1	0	0	4	3	7	7	0	6
Weibull	0	0	0	2	5	6	9	0	6

With fifteen 1 and 2 ranks and an absence of ranks 5, 6, and 7, the Chapman-Richards model clearly provides the best predictions for data truncated at 17.1 years. The Hossfeld model also performs well with seventeen 1, 2 and 3 ranks and an absence of ranks worse than 5. The Gompertz model is reasonably good with seventeen 1, 2 and 3 ranks although it has two ranks of 6. Both the Sloboda and Levakovic models fail to converge for the data of six stands (the same stands in each case) and both the Weibull and Schumacher models are very poor.

Table 6-6: Ranks assigned to the absolute prediction error at 30.1 years when the growth curves were fitted to 22 basal area data sets truncated at the 18.1 years.

	Rank								
	1	2	3	4	5	6	7	NA	Median
Sloboda	11	5	2	1	2	0	0	1	1.5
Chapman	8	10	3	1	0	0	0	0	2
Gompertz	0	1	12	1	4	4	0	0	3
Hossfeld	1	1	2	15	3	0	0	0	4
Levakovic	0	5	3	2	7	4	0	1	5
Schumacher	2	0	0	1	6	13	0	0	6
Weibull	0	0	0	1	0	0	21	0	7

Inspection of Table 6-6 clearly indicates that the Chapman-Richards and Sloboda models are the best when the data is truncated at 18.1 years although convergence was not obtained for one stand for the Sloboda model. The Schumacher and Weibull models continue to provide comparatively poor predictions.

Table 6-7: Ranks assigned to the absolute prediction error at 30.1 years when the growth curves were fitted to 22 basal area data sets truncated at the 19.1 years.

	Rank								
	1	2	3	4	5	6	7	NA	Median
Sloboda	13	4	3	0	1	1	0	0	1
Chapman	3	14	3	2	0	0	0	0	2
Gompertz	1	0	6	6	4	5	0	0	4
Hossfeld	0	2	2	8	9	1	0	0	4
Levakovic	3	1	3	3	4	8	0	0	5
Schumacher	2	1	5	2	4	7	1	0	5
Weibull	0	0	0	1	0	0	21	0	7

As for data truncated at 18.1 years, inspection of Table 6-7 indicates that the Chapman-Richards and Sloboda models continue to well out-perform the other models. There does appear to be some improvement of the prediction quality of the Schumacher model.

Table 6-8: Ranks assigned to the absolute prediction error at 30.1 years when the growth curves were fitted to 22 basal area data sets truncated at the 20.1 years.

	Rank								
	1	2	3	4	5	6	7	NA	Median
Chapman	8	6	6	2	0	0	0	0	2
Sloboda	4	9	3	1	3	1	0	1	2
Levakovic	6	2	6	3	2	0	0	3	3
Hossfeld	1	4	2	9	4	2	0	0	4
Gompertz	1	0	4	5	8	4	0	0	5
Schumacher	2	1	0	2	4	13	0	0	6
Weibull	0	0	1	0	1	1	19	0	7

At a truncation age of 20.1 years the Chapman-Richards and Sloboda models continue to out-perform the other models. There is an improvement in the relative quality of the predictions from the Levakovic model, however the problem of non-convergence is still evident. The Hossfeld model has remained quite consistent over the four truncation ages discussed to this point and for the first time has a median rank better than that of the Gompertz model.

Table 6-9: Ranks assigned to the absolute prediction error at 30.1 years when the growth curves were fitted to 22 basal area data sets truncated at the 22.1 years.

	Rank								
22.1 years	1	2	3	4	5	6	7	NA	Median
Chapman	8	6	5	3	0	0	0	0	2
Sloboda	7	3	4	2	4	2	0	0	3
Levakovic	3	7	8	1	2	0	0	1	3
Hossfeld	2	3	3	6	7	1	0	0	4
Gompertz	0	3	0	5	8	6	0	0	5
Schumacher	1	0	2	4	1	11	3	0	6
Weibull	1	0	0	1	0	2	18	0	7

.Inspection of Table 6-9 indicates that the Levakovic model is beginning to near the Chapman-Richards and Sloboda models in accuracy of prediction.

Table 6-10: Ranks assigned to the absolute prediction error at 30.1 years when the growth curves were fitted to 22 basal area data sets truncated at the 24.1 years.

	Rank								
24.1 years	1	2	3	4	5	6	7	NA	Median
Levakovic	4	7	3	5	2	0	0	1	2.5
Chapman	5	4	7	5	1	0	0	0	3
Sloboda	6	5	0	2	6	3	0	0	3
Hossfeld	3	5	2	5	6	1	0	0	4
Gompertz	3	1	4	2	6	6	0	0	5
Schumacher	1	0	4	2	1	7	7	0	6
Weibull	0	0	2	1	0	5	14	0	7

The evening out in prediction performance among the Chapman-Richards, Levakovic and Sloboda models is apparent from inspection of Table 6-10 and if not for the convergence problems the Levakovic could be out-performing the Chapman-Richards model.

Table 6-11: Ranks assigned to the absolute prediction error at 30.1 years when the growth curves were fitted to 22 basal area data sets truncated at the 26.1 years.

	Rank								
26.1 years	1	2	3	4	5	6	7	NA	Median
Sloboda	9	6	0	1	4	2	0	0	2
Levakovic	3	8	4	4	2	0	0	1	2.5
Chapman	2	2	9	9	0	0	0	0	3
Hossfeld	0	4	5	3	10	0	0	0	4
Schumacher	4	1	3	3	3	4	4	0	4.5
Gompertz	3	1	1		3	13	0	0	6
Weibull	1	0	0	1	0	3	17	0	7

With fifteen ranks of 1 and 2, the Sloboda model is making very good predictions for most stands although it does have six ranks of 5 and 6. The Levakovic model seems to be out-performing the Chapman-Richards model here.

A graphical illustration of the ranks obtained by each modelling function in predicting basal area at age 30.1 years, for each stand, is shown in Figure 6-1.

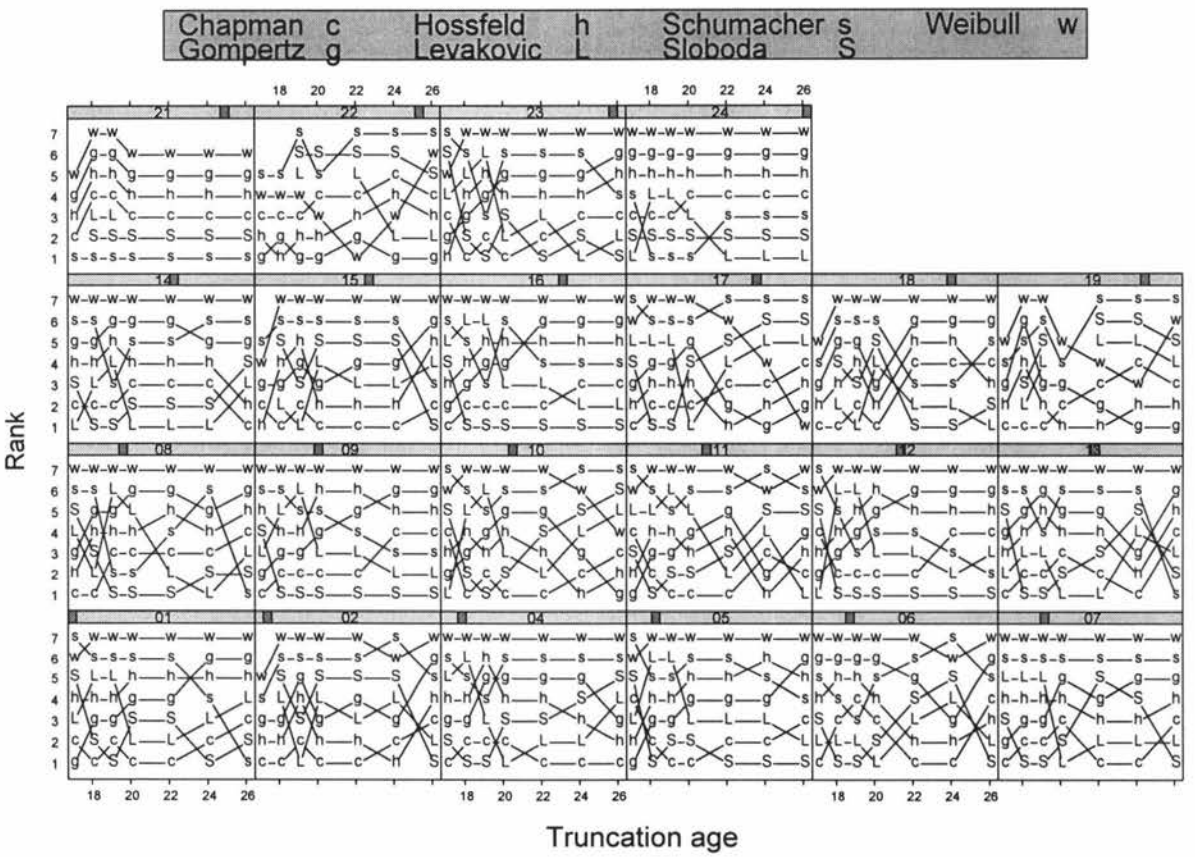


Figure 6-1: Prediction error rankings, by stand, for each model when a prediction is made at 30.1 years. Each prediction is made using a fittted model that has been determined by basal area data up to and including the truncation age.

The panels are somewhat noisy over the first four times where data truncation occurred but settle down somewhat for predictions made based on data up to 22.1, 24.1, and 26.1

years. The plot clearly shows the comparative inadequacy of the Weibull model for prediction although Stand 17 is striking in the improvement of this model as more data is used. Also evident is the deterioration of the comparative performance of the Chapman-Richards model and improvement of the Sloboda model as more data is used and a corresponding improvement in the Levakovic model. In general, the Gompertz model predicts comparatively better when less data is used.

Also of interest is the nature of the prediction error. Figure 6-2 shows the prediction errors obtained at each truncation age. It is of particular note that the Weibull model always predicts very low and in all but a few cases the Schumacher model predicts high. The Sloboda model also predicts high, particularly when less data is used to fit the model. The Hossfeld and Gompertz model tend to predict low and only the Levakovic and Chapman-Richards models seen to give a range of both high and low predictions.

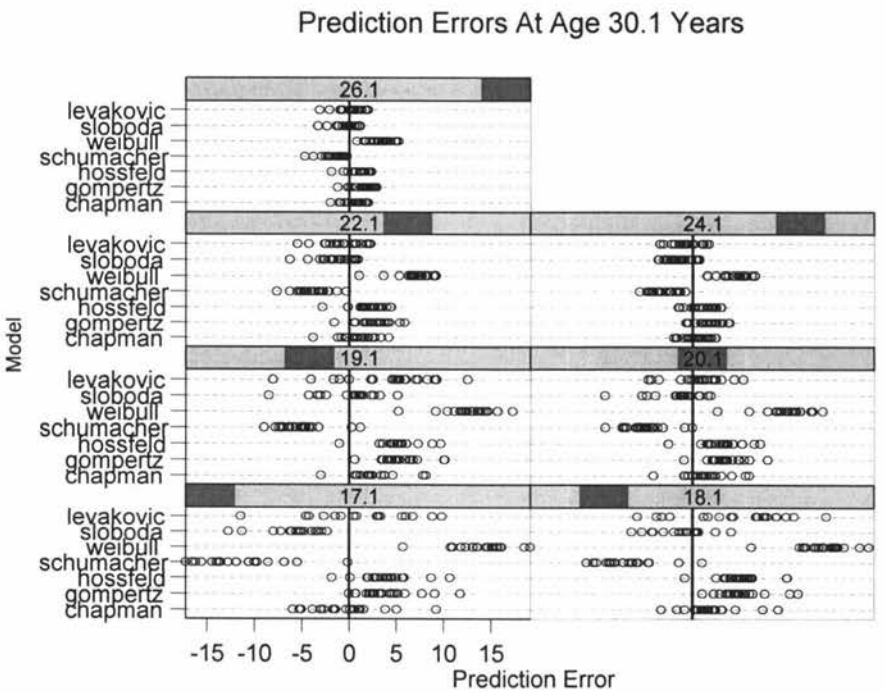


Figure 6-2: Error of prediction at an age of 30.1 years when models are fitted to data truncated at seven ages from 17.1 years to 26.1 years.

6.5 Discussion

It seems that the Chapman-Richards model performs consistently well, whether predictions are made far from the data or close to the data. The Sloboda model provides comparatively good predictions particularly far from the data. The Levakovic model predicted well when the truncation was made not too distant from the age at which the prediction was made. However, it is noted that there were difficulties in fitting the 4-parameter models to the truncated data and in several instances convergence was not obtained. The information gained can now be used to alter the ordering of models in terms of their prediction accuracy from that given in Section 6.3.

Initial ordering by prediction accuracy		Modified ordering by prediction accuracy	
1	Chapman-Richards	1	Chapman-Richards
2	Levakovic	2	Sloboda
3	Sloboda	3	Levakovic
3	Hossfeld	4	Hossfeld
5	Gompertz	5	Gompertz
6	Schumacher	6	Schumacher
7	Weibull	7	Weibull

with the Levakovic model ranked below the Sloboda model mainly because of the difficulties encountered in fitting this model.

6.6 Further predictions of basal area

Data from a further eight stands of pinus radiata were kindly supplied by Dr. Richard Woollons. A valuable feature of this data was that basal area had been determined at large ages; at 41 years in the case of one stand. The data is listed in Appendix K. Table 6-12 details the age where the data was truncated and the greatest age of recorded basal area at which time predictions are made.

Table 6-12: Eight stands of pinus radiata basal area data. The data is truncated at the indicated age, the growth curve models are fitted and a prediction of basal area is made at the maximum recorded age for each stand.

Stand	Age of truncation	Age at which prediction made
1	30.0	34.1
2	34.0	41.0
3	24.8	39.7
4	25.9	37.1
5	25.0	37.0
6	25.0	37.0
7	25.0	37.0
8	32.9	40.1

Prior to investigation the prediction accuracy of the models, goodness-of-fit measures of three models fitted to all the data of the eight stands were calculated. For the Schumacher, Chapman-Richards and Hossfeld models the residual root mean squares were respectively $0.76, 0.93$ and 0.98. If goodness-of-fit is a precursor of prediction accuracy we expect the Schumacher model to work well with these data. This is contrasted with previous data where the Schumacher model had consistently poor goodness-of-fit properties and poor prediction accuracy. Also, curvature measures were calculated for each fitted model. Intrinsic curvatures were found to be quite acceptable in all cases. However, the four-parameter Levakovic and Sloboda models exhibited quite unacceptable parametric-effects curvature as did the three-parameter Weibull model.

In many instances it was very difficult to fit the Levakovic model using non-linear least squares. Convergence was very sensitive to starting values and even very small changes could lead to lack of convergence. There were few difficulties in fitting the 3-parameter models or the Sloboda model. Prediction errors were calculated and the models ranked within each plot in terms of the absolute prediction error (the smallest error ranked as one). The results are summarised in Table 6-13 in decreasing order of the number of ranks of one obtained.

Table 6-13: Ranks assigned to the absolute prediction error for eight stands of forestry data where basal area measurements are available at advanced ages.

	Rank						
Model	1	2	3	4	5	6	7
Sloboda	3	3	1	0	1	0	0
Schumacher	3	3	1	0	0	0	1
Levakovic	1	0	3	0	0	1	3
Chapman	1	0	3	1	3	0	0
Hossfeld	0	1	0	4	2	1	0
Gompertz	0	1	0	2	2	3	0
Weibull	0	0	0	1	0	3	4

It is clear that both the Schumacher and Sloboda models are the very good predictors of basal area when predicting at large ages reasonably far from the data for these eight stands. The absolute errors of prediction are shown in Figure 6-3. Interestingly, the Gompertz model seems to be not much worse than either the Chapman-Richards or Hossfeld models and for Stand 6 it is better. The performance of the Schumacher model for these eight stands is in stark contrast to its predictive accuracy when the original 22 stands of pinus radiata data were used. This improved performance is particularly notable for Stands 5, 6 and 7. These stands are noteworthy for the absence of data between about 15 and 22 years (data near the point of inflection) and the almost linear increasing relationship between 23 and 34 years, shown for Stand 5 in Figure 6-4.

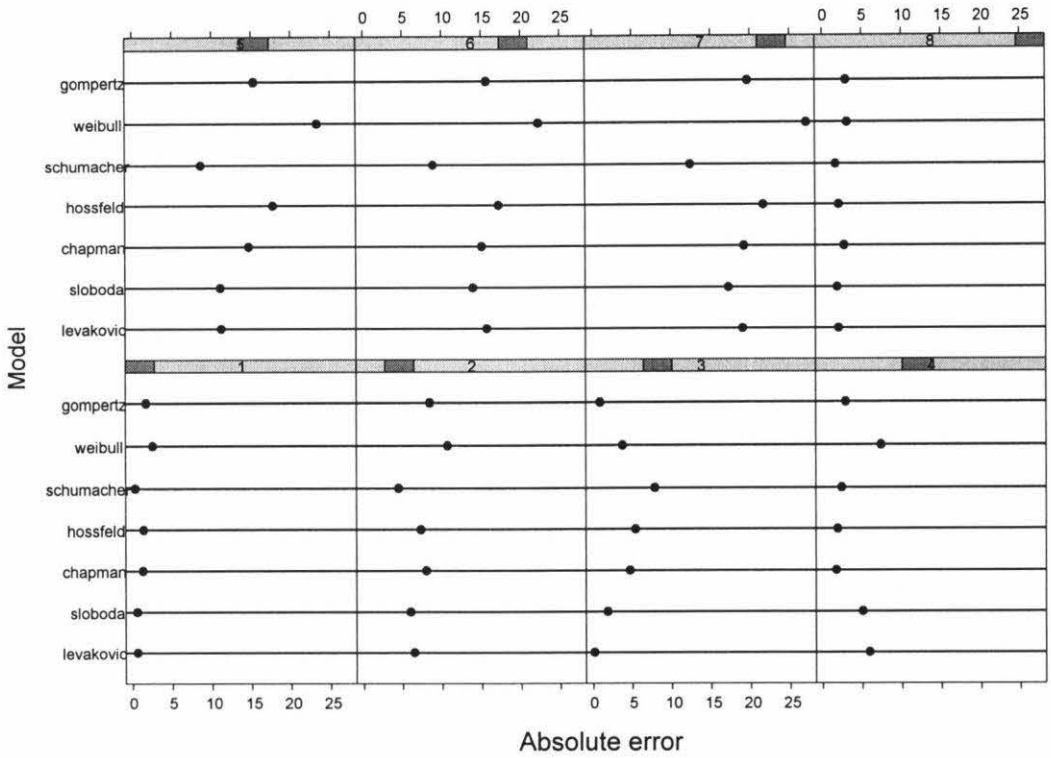


Figure 6-3: Absolute error of prediction for seven growth curves fitted to eight stands of pinus radiata data.

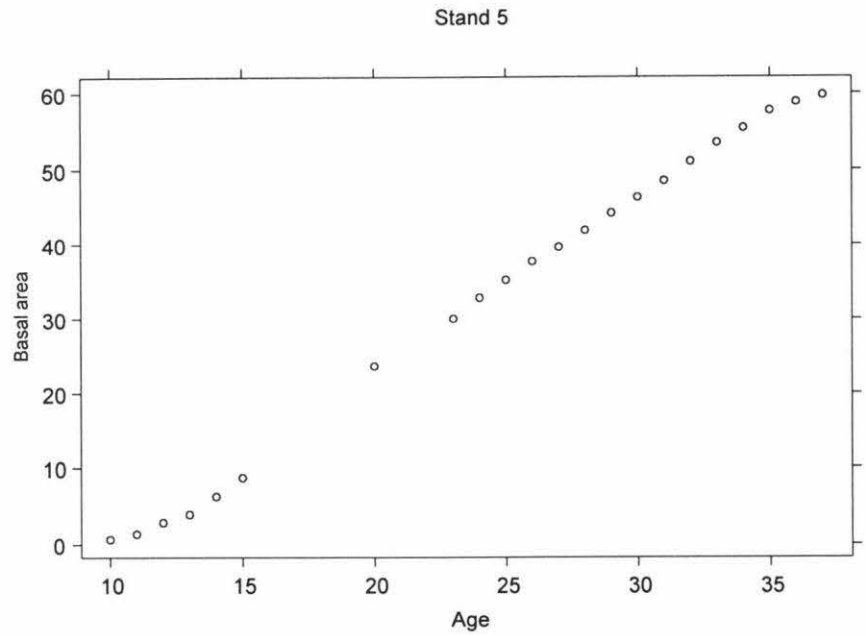


Figure 6-4: Basal area versus age for Stand 5. Note the absence of data near the inflection point and the nearly linear relationship between 22 years and 34 years.

The root mean square prediction errors over the eight stands are shown in Table 6-14. It must first be noted that the sizes of these errors are very large compared to those of Section 6.2 where values close to 1 were typical. The Schumacher model does provide the best predictions with the Sloboda model somewhat more precise than the Levakovic and Chapman-Richards models. If the three unusual data sets (Stands 5, 6, and 7) are eliminated from the calculation of the root mean square prediction error a quite different pattern emerges as shown in Table 6-15.

Table 6-14: Root mean square prediction error for seven growth curves fitted to eight stands of basal area data.

Schumacher	Sloboda	Levakovic	Chapman-Richards	Gompertz	Hossfeld	Weibull
7.1	9.2	10.1	10.7	11.0	12.1	15.7

Table 6-15: Root mean square prediction error for seven growth curves fitted to five stands of basal area data. Stands 5, 6 and 7 have been removed from the calculations of Table 6-14.

Sloboda	Levakovic	Schumacher	Gompertz	Hossfeld	Chapman-Richards	Weibull
3.7	4.0	4.2	4.3	4.3	4.4	6.3

The difference in prediction precision between the Sloboda and Levakovic model is quite small. The Schumacher model has slipped to third place and there is very little difference between the five models with root mean square prediction errors between 4.0 and 4.4. We examine the raw data and fitted models for Stands 5, 6 and 7 to gain an understanding of why the Schumacher model predicts comparatively well for these stands.

When the data were truncated at 25 years none of the models was particularly good in prediction basal area at 37 years. Even when the data is truncated at 32 years and the models fitted to this reduced data set the predictive power at 37 years is not very impressive as shown in Figure 6-5 for the Schumacher, Chapman-Richards and Hossfeld models fitted to Stand 5 data. The Schumacher model will of course give smaller prediction errors for data of this type since it always runs higher than the other models at large ages.

The Schumacher model predicts an asymptote level of 98.4 which looks unreasonable on inspection of the raw data in Figure 6-5. The Sloboda model predicts an asymptote level of 79.8 which is also probably too large. Although the fit of the Chapman-Richards and Hossfeld models is not as good as that of the Schumacher model at large ages, their asymptote levels of 65.2 and 63 do look more believable.

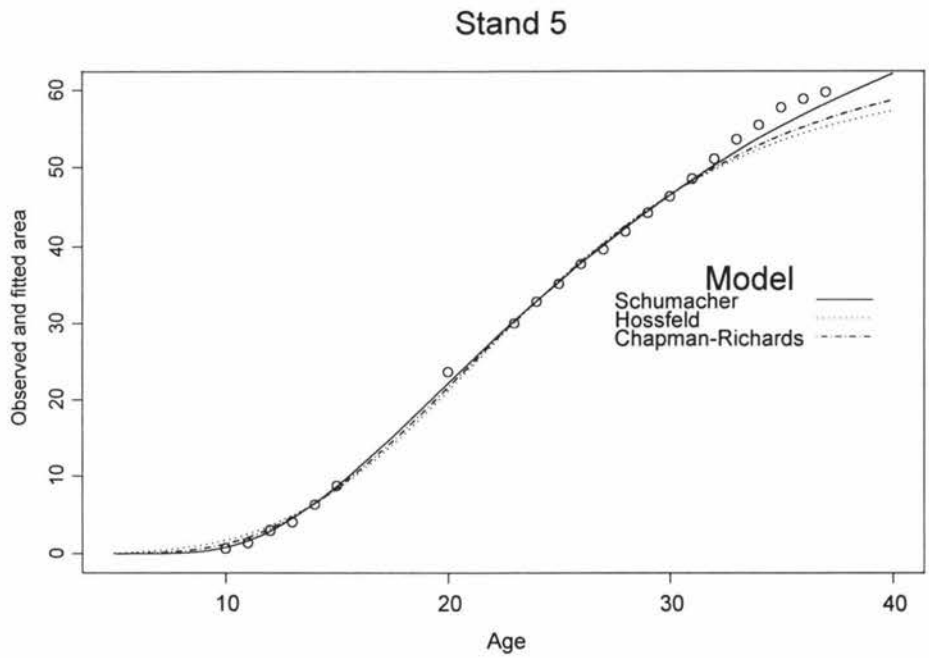


Figure 6-5: Observed basal area and fitted models for Stand 5 data. The fitted models are based on data up to 32 years.

The information gained in this section is based on a very small number of stands and three of these are quite unusual. A possible ordering based on this data set is

1	Schumacher
2	Chapman-Richards
3	Sloboda
4	Levakovic
5	Gompertz
6	Hossfeld
7	Weibull

with the four parameter models ranked below the Chapman-Richards model because of their unacceptably high parametric-effects curvature.

Chapter 7

Predicting Basal Area Using Simulated Data

7.1 Introduction

In this chapter we further investigate the robustness (goodness-of-fit and prediction accuracy) of the growth models using simulated forestry stand data. The aim is to determine whether one or more models have superior goodness-of-fit and provide accurate predictions when data is generated according to one model and all models are tested on this data.

We are interested in three properties of the growth curves: the location of the point of inflection (PoI), the slope of the curve at the PoI and the asymptote level. Two of these properties will be fixed and the third will be free to vary. We will consider four cases and these are specified in Table 7-1 where the properties have been chosen as similar to those of forestry stand data investigated in previous chapters.

Table 7-1: Properties of simulated forestry data for four simulations. PoI is the point of inflection.

Case I	Cases II and III		Case IV
Inflection point at $(10, 20)$	Inflection point location varies		Inflection point at $(10, 20)$
Asymptote 90	Asymptote 90		Asymptote varies
Slope at the PoI varies	Slope at PoI is 3	Slope at PoI is 5	Slope PoI is 3

The simulations will proceed as follows:

1. For each model[1] and case, parameters are calculated to give growth curve equations detailed in Table 7-1. For the Levakovic model c is fixed as 2 and for the Sloboda model d is fixed as -2.

2. For each model and each case, data is generated in the range $T \in (4, 26)$ with a random error e added at each point $(e \backsim Normal(\mu = 0, \sigma = 0.4))$[2]. Thirty data sets are generated for each model giving 180 simulated basal area data sets for cases I, II and III and 210 data sets for case IV.

3. Each model is fitted to each of the data sets and goodness-of-fit is compared.

4. Predictions of basal area are made at $T = 28$, $T = 30$ and $T = 32$ and the prediction errors compared between models.

Example 1 *We will illustrate the process outlined above by examining the fit of the Hossfeld model to a single stand of Chapman-Richards data for case I. The Chapman-Richards function with point of inflection at* $(10, 20)$ *and asymptote level of* 90 *is*

$$c(T) = 90\left(1 - e^{-.0533T}\right)^{1.71}$$

[1] As seen in Section 2.2, the Gompertz model has an inflection point with coordinates of $\left(\frac{\ln b}{c}, \frac{a}{e}\right)$ and slope at the inflection point $\frac{ac}{e}$. This is clearly incompatible with the properties specified for cases I, II and III and the Gompertz model is not used to generate data in these cases,

[2] In Chapter 3 it was shown that the residual sum of squares for a typical stand was approximately 3, giving $s = \sqrt{\frac{3}{18}} = .4$. We also showed that the residuals were reasonably patternless against slope, fitted value, and age for the four models that seem to be promising in predicting post data areas. That is, the Chapman-Richards, Hossfeld, Levakovic and Sloboda models.

We generate data in the range $T \in (4, 26)$ *and add an error at each point to give the data shown in Figure 7-1. The Hossfeld model is then fitted to this data; the residual sum of squares is calculated as* 1.62.

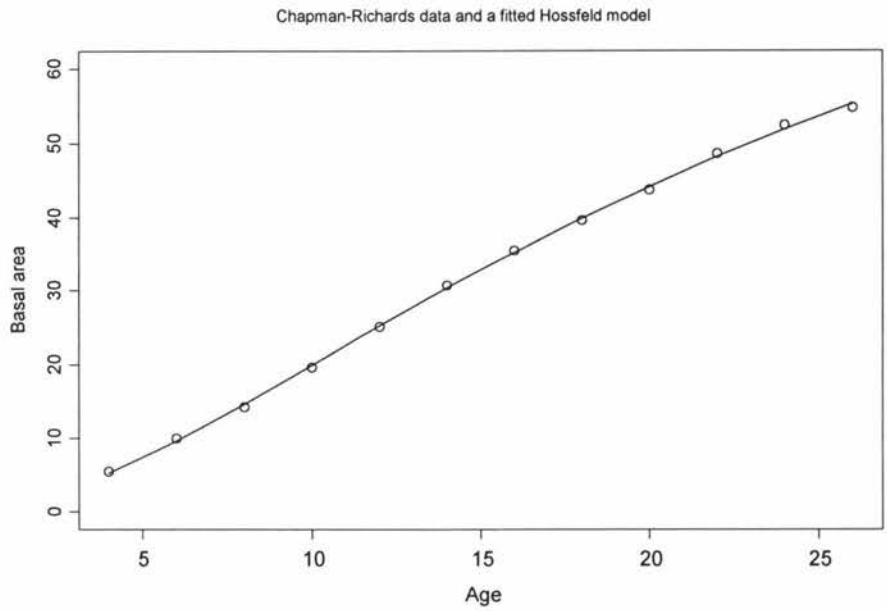


Figure 7-1: Case I data simulated using the Chapman-Richards model with a Hossfeld model fitted.

In general the reparameterizations developed in Chapter 3 will be used in fitting the models by nonlinear least squares. However, in Section 3.4 a one-parameter reparameterization of the Hossfeld model was used. Here we will use a two-parameter reparameterization to allow starting values to be determined more easily This parameterization is developed as follows. Expected values at $T = 10$ and $T = 20$ give

$$\theta_0 = \frac{a10^c}{ab + 10^c} \text{ and } \theta_1 = \frac{a20^c}{ab + 20^c}$$

We solve these equations for a and b giving

$$a = \theta_0\theta_1 \frac{2^c - 1}{2^c\theta_0 - \theta_1} \text{ and } b = \frac{20^c\,(\theta_1 - \theta_0)}{\theta_0\theta_1\,(2^c - 1)}$$

and so the reparameterized model is

$$h(T) = \frac{\theta_0\theta_1(2^c - 1)T^c}{20^c(\theta_1 - \theta_0) + T^c(2^c\theta_0 - \theta_1)}$$

This will be used to fit the Hossfeld model to the simulated forestry data.

7.2 The growth curves

Using the growth curve properties summarised in Table 2-3 and the properties required (specified in Table 7-1) we can routinely determine equations for each curve.

	I	II	III	IV
$c(T)$	$90\left(1 - e^{-.0533T}\right)^{1.71}$	$90\left(1 - e^{-.0621T}\right)^{1.71}$	$90\left(1 - e^{-.10364T}\right)^{1.71}$	$72.96\left(1 - e^{-.087T}\right)^{2.40}$
$g(T)$				$54.366e^{-4.4817e^{-.15T}}$
$h(T)$	$\frac{90T^{1.8}}{220.84+T^{1.8}}$	$\frac{90T^{1.8}}{195.05+T^{1.8}}$	$\frac{90T^{1.8}}{77.769+T^{1.8}}$	$\frac{80T^2}{300+T^2}$
$s(T)$	$90e^{\frac{-144.90804}{T^{1.98382}}}$	$90e^{\frac{-567.01}{T^{1.9838}}}$	$90e^{\frac{-205.81}{T^{1.9838}}}$	$401.8e^{\frac{-9.4868}{T^{.5}}}$
$w(T)$	$90\left(1 - e^{-.012T^{1.336}}\right)$	$90\left(1 - e^{-.016T^{1.336}}\right)$	$90\left(1 - e^{-.032T^{1.336}}\right)$	$52.39\left(1 - e^{-.0057T^{1.93}}\right)$
$L(T)$	$90\left(\frac{T^2}{773.93+T^2}\right)^{.694}$	$90\left(\frac{T^2}{519.62+T^2}\right)^{.694}$	$90\left(\frac{T^2}{187.03+T^2}\right)^{.69381}$	$58.603\left(\frac{T^3}{5000+T^3}\right)^{.6}$
$S(T)$	$.2716e^{5.803e^{-30T^{-2}}}$	$2716e^{5.803e^{-84.116T^{-2}}}$	$.2716e^{5.803e^{-30.284T^{-2}}}$	$12.076e^{2.231e^{-148.65T^{-2}}}$

7.3 The results

I. Data where the point of inflection and the asymptote are fixed

Each of the six growth curve models are fitted to each of the 180 simulated forest stand data sets. The Sloboda and Levakovic functions could not be fitted to the tenth stand simulated using the Chapman-Richards function. Also only 23 of the 30 simulated Schumacher stands could be fitted using the Levakovic model. Table 7-2 summarises the total residual sum of squares in fitting the six sigmoid functions to thirty stands (with the exceptions above) simulated for each model.

Table 7-2: Residual sum of squares for each model/data combination. For example, the bold entry **46.6** is the residual sum of squares when the Chapman-Richards model was fitted to the 30 stands of Chapman-Richards simulated data. The entry below (43.9) is the residual sum of squares when the Chapman-Richards model was fitted to the 30 stands of Hossfeld simulated data.

	Modelling Function - All ages					
Data Model	Chapman	Hossfeld	Schumacher	Weibull	Levakovic	Sloboda
Chapman	**46.6**	46.5	67.2	47.9	37.7[3]	39.2[3]
Hossfeld	43.9	**43.7**	81.5	52.6	37.0	41.3
Schumacher	251.1	417.4	**44.3**	1239.7	27.6[4]	39.9
Weibull	42.0	42.0	54.6	**41.2**	36.7	38.5
Levakovic	51.1	51.0	79.2	47.5	**42.8**	45.6
Sloboda	105.0	168.4	96.2	619.9	42.8	**42.4**

From Table 7-2, it is clear that the Sloboda and Levakovic models are quite satisfactory for all data types with the goodness-of-fit consistently better than when a function models its own data. The Schumacher and Sloboda data are modelled poorly by the 3-parameter Chapman-Richards, Hossfeld and Weibull models. This is unsurprising given that the Sloboda and Schumacher functions with asymptote 90 and inflection point $(10, 20)$ have slopes of 5 and 6 respectively at the point of inflection whereas the other 3-parameter models have slopes close to 2.5 at the point of inflection.

For each stand - data model - fitted model combination, predictions of basal area were made at 28 years, 30 years and 32 years with mean absolute error at 28 years summarised in Table 7-3.

[3] Based on 29 stands
[4] Based on 23 stands

Table 7-3: Mean absolute prediction error at 28 years for each model/data combination.

	Modelling Function - Age 28					
Data Model	Chapman	Hossfeld	Schumacher	Weibull	Levakovic	Sloboda
Chapman	**.30**	.31	.88	.32	.43	.40
Hossfeld	.32	**.30**	.86	.71	.48	.39
Schumacher	2.11	2.14	**.27**	4.19	.39	.32
Weibull	.39	.39	.65	**.44**	.36	.37
Levakovic	.40	.41	1.00	.38	**.56**	.48
Sloboda	1.35	1.40	.60	3.32	.35	**.30**

Inspection of Table 7-3 reveals that the Sloboda model is quite impressive in predicting observations from all the data models considered. The Levakovic model also seems quite robust but it is noted that fitting the Levakovic model to Schumacher data is particularly sensitive to starting values and convergence is often not possible. Also, the Levakovic model is less accurate in predicting its own data than any of the other models except the Schumacher model. If the Schumacher and Sloboda data models are ignored, both Chapman and Hossfeld models give both small and similar prediction errors. It is interesting that the Weibull model does not appear to predict its own data as well as the Chapman, Hossfeld, Levakovic or Sloboda models although it does predict Levakovic data better than any of the other models including the Levakovic model itself.

Prediction errors at ages 30 and 32 years are greater, of course, further from the data but the general observations made at age 28 years are applicable.

Discussion

An important question to be addressed is "Why the slope at the point of inflection is so important to the robustness of the growth curves?" Consider a simple example where the Hossfeld model has slope 6 at the point of inflection $(10, 20)$. These properties yield an Hossfeld function $h(T) = \frac{50T^5}{150000+T^5}$ with asymptote of 50. In order to adequately model the Schumacher data (with inflection point at $(10, 20)$ and slope 6) with an Hossfeld function, the asymptote will be greatly underestimated and consequently predictions will be poor as the asymptote is

approached. A plot of the Schumacher data and the fitted Hossfeld model confirming this is displayed in Figure 7-2.

It is clear that the Hossfeld model fits the Schumacher data poorly at both small and large ages so the asymptote behaviour of the Schumacher data is described poorly. The 3-parameter Hossfeld model lacks the flexibility to describe both the Schumacher inflection point and asymptote properties adequately.

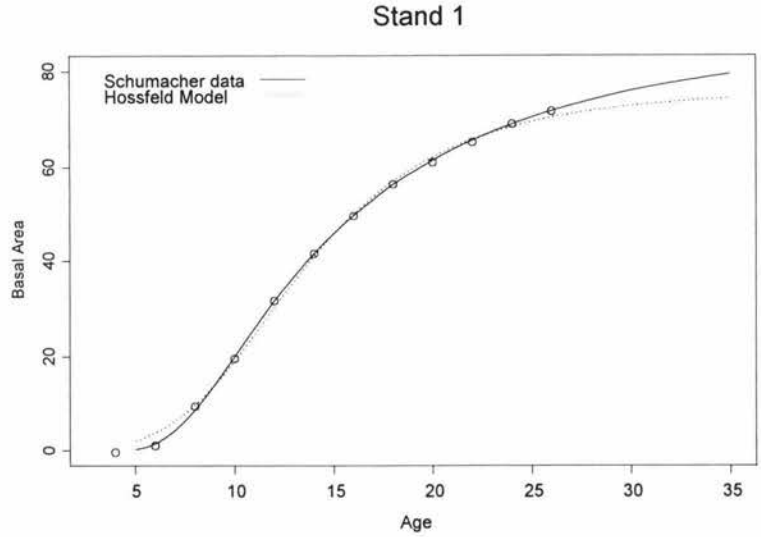


Figure 7-2: The Schumacher data (inflection point at $(10, 20)$ with slope 6 and asymptote 90) and the fitted Hossfeld Model.

As discussed in Section 2.4, for the 3-parameter models, the ratio of the asymptote to the y-coordinate of the inflection point is dependent on only one parameter. To set the inflection point position and the asymptote however, requires the specification of all three parameters with the slope at the inflection point then fixed by these parameters. With only three parameters there is insufficient flexibility to model both the asymptote and the inflection point parameters up to the first derivative.

This is not a problem when the Weibull, Chapman and Hossfeld functions are fitted to one another's data since each has a similar slope at the point of inflection for a given position of the inflection point and asymptote. The Schumacher function though, has much greater inflection point slope than the other 3-parameter models.

Conclusion

If the slope at the point of inflection is small (near 2.5, as it is for the Chapman-Richards, Hossfeld, Weibull and Levakovic data models) then the simpler 3-parameter models, Chapman-Richards and Hossfeld, have very acceptable goodness-of-fit and small prediction errors. If however, the curve has large slope (near 6, as it is for the Schumacher and Sloboda data models) at the inflection point then the Sloboda and Schumacher models have good properties. Further, the Sloboda and Levakovic models are sufficiently flexible to work quite well irrespective of the slope at the inflection point. Therefore the order of the models in descending order of prediction accuracy is

1	Sloboda
2	Levakovic
3 =	Chapman-Richards
3 =	Hossfeld
5	Weibull
6	Schumacher

with the Levakovic model ranked at position two mainly because of the difficulty in obtaining convergence.

II. Data with a fixed asymptote and a slope of three at the inflection point

The Sloboda and Levakovic models do not converge for ten and thirteen stands respectively of the simulated Schumacher data. Table 7-4 summarises the total residual sum of squares in fitting the six sigmoid functions to thirty stands (with the exceptions above) simulated for each model. Shown in Table 7-4 are the residual sums of squares obtained when each model was fitted to each data type.

Table 7-4: Residual sum of squares for each model/data combination. For example, the bold entry **33.8** is the residual sum of squares when the Chapman-Richards model was fitted to the 30 stands of Chapman-Richards simulated data. The entry below (56.8) is the residual sum of squares when the Chapman-Richards model was fitted to the 30 stands of Hossfeld simulated data.

	Modelling Function - All ages					
Data Model	Chapman	Hossfeld	Schumacher	Weibull	Levakovic	Sloboda
Chapman	**33.8**	85.6	1614.1	251.1	121.2	602.1
Hossfeld	56.8	**14.3**	1180.2	718.9	61.0	338.8
Schumacher	1771.1	4597.7	**75.7**	13423.7	162.9[5]	207.9[6]
Weibull	139.8	328.2	2247.5	**30.5**	90.8	824.3
Levakovic	12.2	100.2	1659.4	190.0	**57.6**	468.4
Sloboda	519.6	1482.7	714.2	7574.6	146.4	**106.8**

From Table 7-4, it is striking that only the Levakovic model provides consistently satisfactory goodness-of-fit over all data types. Although the Levakovic model could not be fitted to ten stands of Schumacher data, the residual sum of squares compares favourably with the other models except for the Schumacher itself.

The Chapman-Richards model is quite satisfactory other than when modelling Schumacher and Sloboda data and it models Levakovic data better than the Levakovic model. This small residual sum of squares (12.2) is notably the smallest of any model-data combination.

Summarised below in Table 7-5 are the residual sum of squares in making predictions at ages 30, 35, 40 years for the seven sigmoid functions. These have been averaged over the number of fitted models (generally thirty with the exceptions mentioned above).

[5] Based on 20 stands
[6] Based on 17 stands

Table 7-5: The mean residual sum of squares of prediction (at ages 30, 35 and 40 years combined) when each growth curve model is fitted to each data type. For example, 19.68 is the prediction error when the Chapman-Richards model is used to predict Schumacher data.

Data Model	Chapman	Hossfeld	Schumacher	Weibull	Levakovic	Sloboda
Chapman	**.38**	.95	17.93	2.79	1.35	6.69
Hossfeld	.63	**.16**	13.11	7.99	.68	3.76
Schumacher	19.68	51.09	**.84**	149.15	3.39	3.46
Weibull	1.55	3.65	24.97	**.34**	1.01	9.16
Levakovic	.14	1.11	18.44	2.11	**.64**	5.20
Sloboda	5.77	16.47	7.94	84.16	1.63	**1.19**

The result that stands out is the precision given when the Chapman-Richards model is fitted to the data generated using the Levakovic model. The mean residual sum of squares is only 0.14 compared with 0.64 when the Levakovic function models its own data. The Chapman-Richards model makes good predictions, other than for data generated using the Sloboda and Schumacher functions. The Levakovic model is consistently good as a predictor over the range of data models. Other than for predicting their own data, the Schumacher, Sloboda, Hossfeld and Weibull models are not particularly good especially when compared to the Chapman-Richards and Levakovic models. These results about prediction accuracy closely match the results in terms of goodness-of-fit.

Ranks of one to seven are assigned within each stand to the models in increasing order of absolute prediction error. Table 7-6 summarises the ranks assigned to each model in decreasing order of the number of first ranks obtained.

Table 7-6: Ranks assigned to the growth curve models by absolute prediction error. A rank of one indicates the smallest prediction error within a stand.

	Rank						
Fitted Model	1	2	3	4	5	6	7
Chapman Richards	137	134	182	87	0	0	0
Levakovic	126	187	135	43	7	0	0
Hossfeld	76	79	89	148	37	111	0
Weibull	73	45	39	93	114	45	131
Schumacher	71	17	31	27	33	235	126
Sloboda	54	73	35	96	248	4	0
Gompertz	3	5	29	46	101	121	235

We recall that for each model other than the Gompertz function, 90 predictions are made using a curve's own data. It might be expected that each model would have roughly 90 first ranks. This is clearly not the case with the superiority of the Chapman-Richards and Levakovic models apparent. It is of some note that the predictions gained from fitting the Chapman-Richards function are never worse than fourth of the seven models considered. The Levakovic function has a combined total of 313 first and second ranks compared with 271 for the Chapman-Richards function.

The Sloboda model gives only 54 ranks of one but has a total of 127 ranks of one and two which is fourth best. This model has only four predictions at rank six compared to the Hossfeld model's 111 ranks of six. The Sloboda model performs comparatively poorly here compared to the first simulation. The fit of the Sloboda function to the Weibull data is shown in Figure 7-3. It is seen that the model fits reasonably well but overestimates markedly at 35 and 40 years. This can be compared with the fit of the 3-parameter Chapman-Richards model to the same data shown in Figure 7-4 where prediction errors are much smaller.

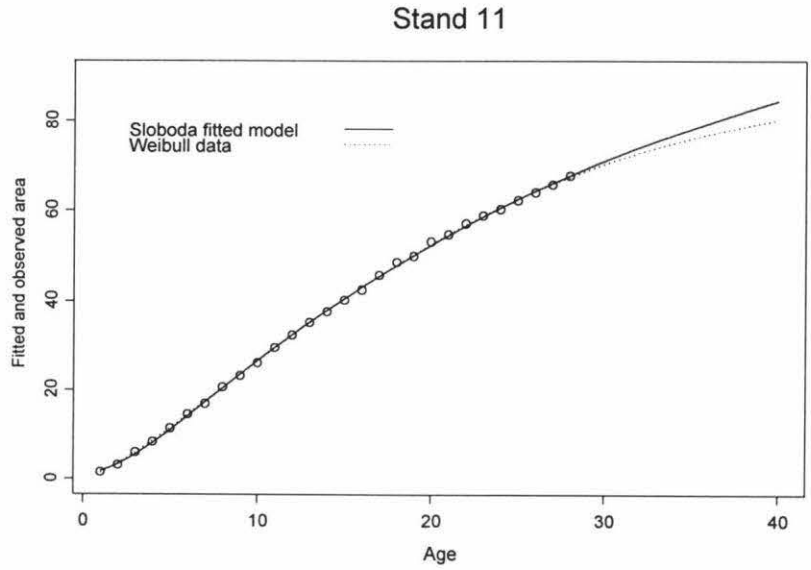


Figure 7-3: A Sloboda model fitted to data of a Weibull function.

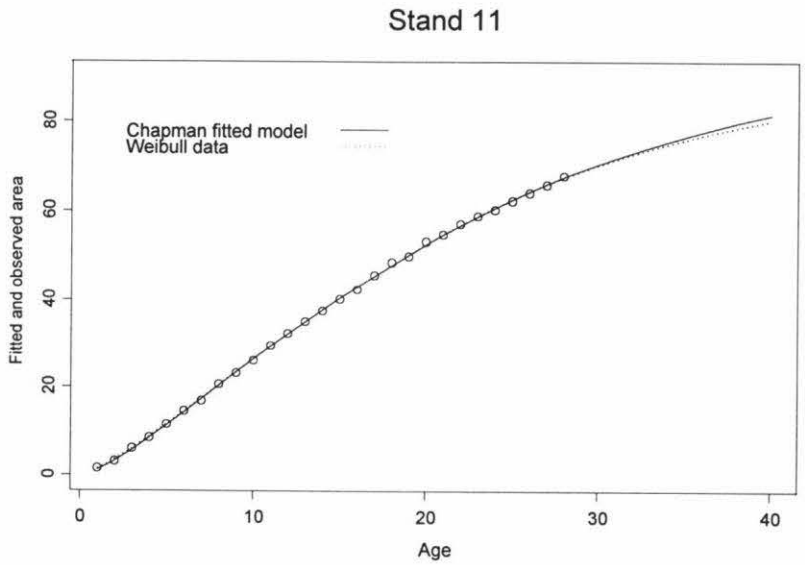


Figure 7-4: A Chapman-Richards model fitted to data of a Weibull function

Conclusion

It seems clear that the Levakovic model performs reasonably well over all the data models. The Chapman-Richards model is extremely effective in predicting from data that is Weibull,

Levakovic and Hossfeld but consistently underestimates when the data is generated using the Sloboda or Schumacher models. Based on the results of this simulation, an ordering of robustness of the models is

1	Chapman-Richards
2	Levakovic
3 =	Sloboda
3 =	Hossfeld
5	Weibull
6	Schumacher

with the Levakovic model failing to gain the best rating because of the difficulty in gaining convergence in several cases.

III. Data with a fixed asymptote and a slope of five at the inflection point

As in the previous simulation, there were difficulties in fitting the Levakovic function to the simulated Schumacher data with convergence not obtained for ten of the thirty simulated stands. Shown in Table 7-7 are the residual sums of squares obtained when each model was fitted to each data type.

Table 7-7: Residual sum of squares for each model/data combination. For example, the bold entry **6.2** is the residual sum of squares when the Chapman-Richards model was fitted to the 30 stands of Chapman-Richards simulated data. The entry below (316.6) is the residual sum of squares when the Chapman-Richards model was fitted to the 30 stands of Hossfeld simulated data.

	Modelling Function - All ages					
Data Model	Chapman	Hossfeld	Schumacher	Weibull	Levakovic	Sloboda
Chapman	**6.2**	381.2	2883.5	238.4	140.6	1120.3
Hossfeld	316.6	**5.2**	964.0	1060.1	11.3	239.3
Schumacher	2019.3	2015.3	**11.9**	7224.5	16.9[7]	33.3
Weibull	167.5	1110.6	4600.0	**9.5**	168.1	1879.0
Levakovic	104.6	102.7	1608.7	502.3	**17.4**	419.9
Sloboda	1177.0	829.5	178.9	4775.4	21.9	**8.1**

Table 7-7 clearly indicates the quite striking superiority of the Levakovic model in modelling data of the other growth curve models. The results are similar to the previous simulation where the slope was three at the point of inflection but here there is a more obvious difference between the Levakovic model and the others.

Table 7-8 summarises the residual sum of squares in making predictions at ages 30, 35, 40 years. These have been averaged over the number of fitted models (generally thirty with the exceptions mentioned above)

[7] Based on 20 stands.

Table 7-8: The mean residual sum of squares of prediction (at ages 30, 35 and 40 years combined) when each growth curve model is fitted to each data type.

	Modelling Function					
Data Model	Chapman	Hossfeld	Schumacher	Weibull	Levakovic	Sloboda
Chapman	**.07**	4.23	32.04	2.65	1.56	12.45
Hossfeld	3.52	**.06**	10.71	11.78	.13	2.66
Schumacher	22.43	22.39	**.13**	80.27	.28	.37
Weibull	1.86	12.34	51.11	**.11**	1.87	20.88
Levakovic	1.16	1.14	17.87	5.58	**.19**	4.67
Sloboda	13.07	9.22	1.99	53.06	.24	**.09**

Again, the goodness-of-fit results of Table 7-7 closely match the prediction accuracy of Table 7-8. The Levakovic function stands out as a consistently good predictor over all the data models. The Chapman-Richards model makes reasonable predictions, other than for Sloboda or Schumacher data. The Gompertz model is consistently very poor. Other than for predicting their own data, the Schumacher, Sloboda, Hossfeld and Weibull models are not particularly good, especially when compared to the Levakovic model. Assigning a rank of one to the model with the smallest absolute error of prediction within each stand and a rank of seven to the greatest, Table 7-9 summarises the ranks obtained by each model.

Table 7-9: Ranks assigned to the growth curve models ordered by absolute prediction error. A rank of one indicates the smallest prediction error within a stand

	Rank						
Fitted Model	1	2	3	4	5	6	7
Levakovic	154	273	80	3	0	0	0
Chapman-Richards	95	76	143	117	109	0	0
Sloboda	84	79	81	80	92	124	0
Weibull	82	28	49	57	88	86	150
Hossfeld	70	57	85	219	109	0	0
Schumacher	55	26	98	0	57	68	236

The Levakovic model is quite impressive in gaining ranks 1, 2 or 3 in all but three cases (remembering that the model could not be fit to 10 stands of Schumacher data). The positioning of the Hossfeld model in Table 7-7 is a little deceptive since it does not yield any ranks worse than 5 and has 127 ranks of 1 and 2. The Sloboda model has 163 ranks of 1 and 2 but also has 124 ranks of 6.

The results at each prediction age show the same pattern evident in the combined results of the three prediction ages and are not presented.

Conclusion

The stability of the Levakovic model over all data types and at all prediction ages is impressive. The other models are not close to Levakovic in consistency of prediction where the data has slope 5 at the point of inflection. Based on the results of this simulation, an ordering of robustness of the models is

1	Levakovic
2	Chapman-Richards
3	Hossfeld
4	Sloboda
5	Weibull
6	Schumacher

IV. Data with a fixed inflection point where the slope is three

As for previous simulations, there were difficulties fitting the Levakovic and Sloboda models to the Schumacher data with sixteen and twelve stands respectively not converging. Shown in Table 7-10 are the residual sums of squares obtained when each model was fitted to each data type.

Table 7-10: Residual sum of squares for each model/data combination. For example, the bold entry **10.9** is the residual sum of squares when the Chapman-Richards model was fitted to the 30 stands of Chapman-Richards simulated data. The entry below (223.3) is the residual sum of squares when the Chapman-Richards model was fitted to the 30 stands of Hossfeld simulated data.

	Modelling Function - All ages						
Data Model	Chapman	Gompertz	Hossfeld	Schumacher	Weibull	Levakovic	Sloboda
Chapman	**10.9**	840.8	42.0	1732.5	639.0	134.7	702.2
Gompertz	223.3	**5.9**	524.6	2730.6	63.5	19.7	907.1
Hossfeld	49.3	1073.2	**19.1**	1328.8	888.9	61.7	389.9
Schumacher	1602.3	8147.1	1466.8	**21.4**	3806.9	31.4[8]	35.6[9]
Weibull	477.1	75.3	900.7	3519.2	**10.4**	133.2	1766.1
Levakovic	46.2	103.1	205.5	1963.2	265.3	**9.7**	742.3
Sloboda	2618.2	9769.3	2506.9	76636.9	1310.6	749.2	**29.6**

Inspection of Table 7-10 shows the striking superiority in goodness-of-fit of the Levakovic model. None of the other models is close in consistency across all data types although the Chapman-Richards model does perform adequately for Hossfeld and Levakovic data in addition to its own.

Table 7-11 summarises the residual sum of squares in making predictions at ages 30, 35 and 40 years for the seven sigmoid functions fitted to thirty stands simulated for each model. These have been averaged over the number of fitted models (generally thirty with the exceptions mentioned above)

[8]Based on 14 stands.
[9]Base on 18 stands.

Table 7-11: The mean residual sum of squares of prediction (at ages 30, 35 and 40 years combined) when each growth curve model is fitted to each data type.

	Modelling Function						
Data Model	Chapman	Gompertz	Hossfeld	Schumacher	Weibull	Levakovic	Sloboda
Chapman	**.12**	9.34	.47	19.25	7.10	1.50	7.80
Gompertz	2.48	**.07**	5.83	30.34	.71	.22	10.08
Hossfeld	.55	11.92	**.21**	14.76	9.88	.69	4.33
Schumacher	17.80	90.52	16.30	**.24**	42.30	.75	.63
Weibull	5.30	.84	10.01	39.10	**.12**	1.48	9.62
Levakovic	.51	1.15	2.28	21.81	2.94	**.11**	8.25
Sloboda	29.09	108.55	27.85	851.52	14.56	8.32	**.33**

The Chapman-Richards model performs well other than for the data generated using the Schumacher and Sloboda models. This is consistent with results in other simulations. The Levakovic model is clearly the best over the range of data models; however, it must be mentioned that convergence is sometimes difficult to achieve. Of some note is the generally appalling performance of the Schumacher model.

As for previous simulations, ranks are calculated and these are summarised in Table 7-12. Again, the clear superiority of the Levakovic model is evident with the next best (the Chapman-Richards model) far behind in prediction performance. The next best is possibly the Hossfeld model since it has no ranks less than 5. This is followed by the Weibull and Sloboda models and finally the very poorly performing Gompertz and Schumacher models.

Table 7-12: Ranks assigned to the growth curve models ordered by absolute prediction error. A rank of 1 indicates the smallest prediction error within a stand.

	Rank, All Ages						
Fitted Model	1	2	3	4	5	6	7
Levakovic	138	229	198	14	3	0	0
Chapman-Richards	104	90	104	218	114	0	0
Sloboda	99	43	13	108	41	293	0
Weibull	87	34	167	121	149	67	5
Hossfeld	53	115	81	145	236	0	0
Gompertz	76	112	57	24	63	220	78
Schumacher	73	7	10	0	24	17	499

Conclusion

The robustness of the Levakovic model over all data types is clearly markedly superior to the other models. Based on the results of this simulation, an ordering of robustness of the models is

1	Levakovic
2	Chapman-Richards
3	Hossfeld
4	Weibull
5	Sloboda
6	Gompertz
7	Schumacher

7.4 Summary

We summarise the results of the four simulations by presenting the growth curve orderings determined through the discussion of the results of each simulation:

	I	II	III	IV
1	Sloboda	Chapman-Richards	Levakovic	Levakovic
2	Levakovic	Levakovic	Chapman-Richards	Chapman-Richards
3	Chapman-Richards,	Hossfeld	Hossfeld	Hossfeld
4	Hossfeld	Sloboda	Sloboda	Weibull
5	Weibull	Weibull	Weibull	Sloboda
6	Schumacher	Schumacher	Schumacher	Gompertz
7				Schumacher

It is clear that there are two classes of models. In one group we have the Chapman-Richards, Levakovic, Sloboda and Hossfeld models that seem to predict accurately with the Levakovic and Chapman-Richards models the best. The second group of poorly performing models consists of the Weibull, Schumacher and Gompertz models.

Part V

The Growth Curves on a Common Footing

Chapter 8

Padé Rational Approximations

8.1 Introduction

To further understand the nature of the growth curves, and put them on a similar footing so that direct comparisons can be made, Professor G. R. Wood suggested that each growth curve be approximated as a series expansion. Rather than using a Taylor series expansion, which will not capture the asymptote behaviour of the curves, it was suggested that this behaviour should be better described by approximation as the ratio of two polynomials - a Padé rational approximation [5].

For each model, we used information at the point of inflection to develop rational approximations, with the aim to determine features of the Padé approximations that would explain the robustness properties of the growth curves observed when fitted to basal area data.

We were also interested to see whether the classification of the growth curve differential equations, discussed in Section 2.5, was consistent with a classification of the Padé approximations.

8.2 Derivation of a Padé approximation

A Padé approximation represents f as the ratio of two polynomials. Since an asymptote exists for the growth curves, each polynomial used should be of the same degree n. That is, the Padé

approximation for $f(x)$ at $x = 0$ has the form

$$R_{2n}(x) = \frac{a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n}{1 + b_1 x + b_2 x^2 + \ldots + b_n x^n}$$

The coefficients are generated to ensure that R_{2n} and f agree at $x = 0$ and also to ensure that the first $2n$ derivatives agree at $x = 0$. We start with a MacLaurin series, $P_{2n}(x)$ for $f(x)$.

$$P_{2n}(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_{2n} x^{2n}, \text{ where } c_i = f^{(i)}(0)/i!$$

This expansion will of course ensure agreement at $x = 0$ between f and P_{2n} up to the $2n^{\text{th}}$ derivative. We now consider

$$\begin{aligned} P_{2n}(x) - R_{2n}(x) &= c_0 + c_1 x + c_2 x^2 + \ldots + c_{2n} x^{2n} - \frac{a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n}{1 + b_1 x + b_2 x^2 + \ldots + b_n x^n} \\ &= \frac{\left(c_0 + c_1 x + c_2 x^2 + \ldots + c_{2n} x^{2n}\right)\left(1 + b_1 x + b_2 x^2 + \ldots + b_n x^n\right) - a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n}{1 + b_1 x + b_2 x^2 + \ldots + b_n x^n} \end{aligned}$$

and we calculate the a_i and b_i to ensure $P_{2n}(x) - R_{2n}(x) = 0$ for $i = 0, 1, \ldots, 2n$.

Firstly, we require that $P_{2n}(0) - R_{2n}(0) = 0$ and so $c_0 - a_0 = 0$. That is, a_0 is the constant term of the MacLaurin series expansion of f. Secondly, $P'_{2n}(0) - R'_{2n}(0) = 0$ giving the following equation in the a_i, b_i and c_i

$$c_1 + b_1 c_0 - a_1 - (c_0 - a_0) b_1 = 0, \text{ or } b_1 c_0 + c_1 - a_1 = 0$$

Similarly, for successive derivatives it is also required that $P^{(i)}_{2n}(0) - R^{(i)}_{2n}(0) = 0$ giving the following $2n - 1$ equations:

$$\begin{aligned} b_2 c_0 + b_1 c_1 + c_2 - a_2 &= 0 \\ b_3 c_0 + b_2 c_1 + b_1 c_3 + c_3 - a_3 &= 0 \\ &\vdots \\ b_n c_0 + b_{n-1} c_1 + \ldots + b_1 c_{n-1} + c_n - a_n &= 0 \\ b_n c_1 + b_{n-1} c_2 + \ldots + b_1 c_n + c_{n+1} &= 0 \\ b_n c_2 + b_{n-1} c_3 + \ldots + b_1 c_{n+2} + c_{n+2} &= 0 \end{aligned}$$

$$\vdots$$

$$b_n c_n + b_{n-1} c_{n+1} + \ldots b_1 c_{2n-1} + c_{2n} = 0$$

Solving these $2n+1$ equations gives the coefficients of the Padé rational approximation $R_{2n}(x)$. The expansion is of course around $x = 0$ since the polynomials are derived from MacLaurin series.

Example For the Schumacher function $s(T) = 90 \exp\left(\frac{-144.9}{T^{1.98382}}\right)$ which has an asymptote of 90 and inflection point at $(10, 20)$, a Padé rational approximation $R_6(x)$ around $T = 10$ will be determined. We first map $T \to T + 10$ (that is, a translation of ten units to the left) and define $s_1(T) = 90e^{\frac{-144.9}{(T+10)^{1.98382}}}$. Now the expansion of $s_1(T)$ will be found around $T = 0$.

The MacLaurin series expansion of $s_1(T)$ is

$$\begin{aligned} P_6(T) = {} & 20.002 + 5.9678T - 4.5041 \times 10^{-5} T^2 - 5.8871 \times 10^{-2} T^3 + \\ & 7.3359 \times 10^{-3} T^4 - 2.6743 \times 10^{-4} T^5 - 6.2985 \times 10^{-5} T^6 + O\left(T^7\right) \end{aligned}$$

and

$$\begin{aligned} P_6(T) - R_6(T) = {} & 20.002 + 5.9678T - .000045041T^2 - .058871T^3 + .0073359T^4 \\ & -.00026743T^5 - .000062985T^6 - \frac{a_0 + a_1 T + a_2 T^2 + a_3 T^3}{1 + b_1 T + b_2 T^2 + b_3 T^3} \end{aligned}$$

Now, we require $P_6^{(i)}(0) - R_6^{(i)}(0) = 0$ for $i = 0, 1, 2, 3, 4, 5, 6$ giving the following 7 equations in the coefficients

$$\begin{aligned} 20.002 - a_0 &= 0 \\ 20.002 b_1 + 5.9678 - a_1 &= 0 \\ 20.002 b_2 + 5.9678 b_1 - .000045041 - a_2 &= 0 \\ 20.002 b_3 + 5.9678 b_2 - .000045041 b_1 - .058871 - a_3 &= 0 \\ 5.9678 b_3 - .000045041 b_2 - .058871 b_1 + .0073359 &= 0 \\ -.000045041 b_3 - .058871 b_2 + .0073359 b_1 - .00026743 &= 0 \end{aligned}$$

$$-.058871b_3 + .0073359b_2 - .00026743b_1 - .000062985 = 0$$

Solving the final three equations gives $b_1 = .36354$, $b_2 = .040756$, and $b_3 = .0023573$; substituting these into the second, third and fourth equations gives $a_0 = 20.002$, $a_1 = 13.239$, $a_2 = 2.9847$, and $a_3 = .23149$.

The Padé approximation of $s_1(T)$ is then given as

$$R_6(T) = \frac{20.002 + 13.239T + 2.9847T^2 + .23149T^3}{1 + .36354T + .040756T^2 + .0023573T^3}$$

and mapping $T \rightarrow T - 10$ (that is, a translation of ten units to the right) we have

$$R_6(T) = \frac{-45.408 + 22.992T - 3.96T^2 + .23149T^3}{-.9171 + .25561T - .029963 \times 10^{-2}T^2 + .0023573T^3}$$

as the Padé approximation of $s(T)$.

The calculations have been performed with relatively little precision and there are small errors in the coefficients. *Mathematica* was used to compute the Padé approximation using the function *Pade*[] that is supplied with the standard calculus packages of Version 3. Using this tool, $R_6(T)$ was calculated as

$$R_6(T) = \frac{-45.4276 + 22.9996T - 3.96105T^2 + .231536T^3}{-.917613 + .255715T - .029974T^2 + .00235787T^3}$$

and both $R_6(T)$ and $s(T)$ are displayed in Figure 8-1. It is clear that $s(T)$ is approximated well by the Padé approximation up to about $T = 30$. However, the asymptote level is not described as well as might be hoped. $R_6(T)$ has an asymptote level of $\frac{.231536}{.00235787} = 98.197$ whereas the Schumacher function $s(T)$ has an asymptote level of 90.

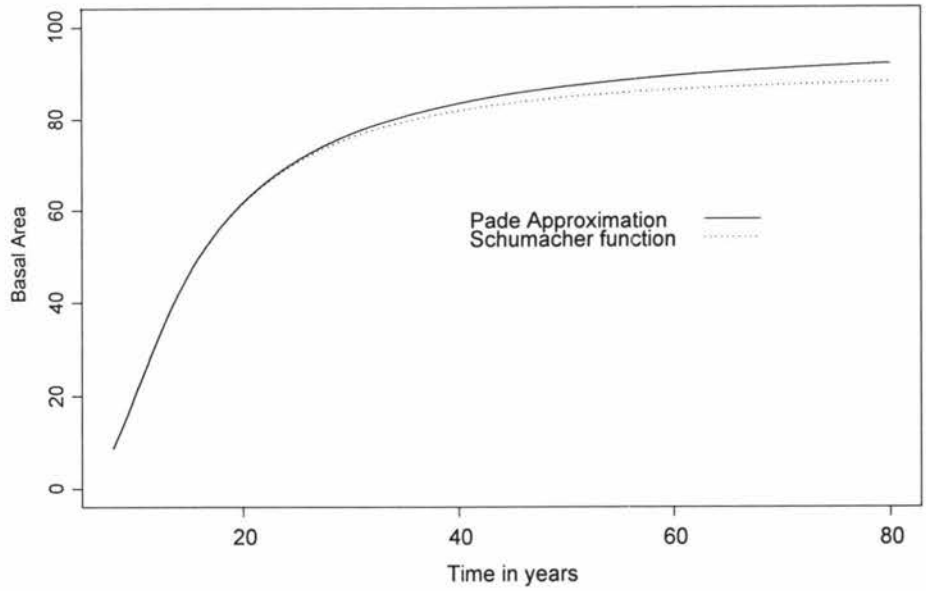


Figure 8-1: The Schumacher function and its Padé approximation $R_6(T)$.

8.3 Rational approximations of the growth curves

Using Mathematica, Padé approximations with polynomials of degree $2, 3, 4, 5, 6, 7, 10$ and 12 in the numerator and denominator were determined for the growth curves with inflection point $(10, 20)$ and asymptote of 90. The rational approximation of the Gompertz function is not calculated; with the inflection point properties specified the Gompertz function will have an asymptote level of $20e = 54.366$ which is clearly incompatible with the requirements.

We require good agreement at the inflection point and this will be ensured by expanding around this point. We also need good approximation up to about $T = 40$ where forestry data occurs and reasonable agreement of asymptote level.

The formulae for the Padé approximations $R_4(T), R_6(T), R_8(T)$ and $R_{10}(T)$ are listed in Appendix J and the asymptote levels determined by each approximation are listed in Table 8-1.

Table 8-1: Asymptote levels given by the Padé approximations.

	Asymptote Level							
Polynomial degree	2	3	4	5	6	7	10	12
Growth Curve	$R_4(T)$	$R_6(T)$	$R_8(T)$	$R_{10}(T)$	$R_{12}(T)$	$R_{14}(T)$	$R_{20}(T)$	$R_{24}(T)$
Schumacher	95.4	98.2	91.1	89.9	90.0	90.0	90.0	90.0
Sloboda	89.4	97.7	91.4	89.3	90.0	90.0	90.0	90.0
Weibull	150.8	79.1	31.1	29.3	−14.7	284.6	50.7	23.9
Chapman-Richards	134.9	77.6	50.1	47.9	258.1	121.5	75.8	73.5
Levakovic	98.3	83.1	79.9	91.9	89.8	89.6	90.0	90.0
Hossfeld	95.0	91.2	90.5	90.2	90.1	90.1	90.0	90.0

The Hossfeld, Schumacher, Sloboda and Levakovic approximations give quite close agreement to the asymptote level of 90. This is contrasted with the poor description of the Chapman-Richards and Weibull models where the asymptote level is unreasonably large or small with no clear relationship with the degree of the approximation. This behaviour is striking when it is recalled that the Chapman-Richards model was among the best in modelling forestry data and the Weibull model consistently the worst.

We will further investigate the most accurate approximation developed - the Padé approximation with degree 12 polynomials in the numerator and denominator. The maximum error of approximation of any of the degree 12 approximations in the range $T \in [5, 26]$ is 7×10^{-6} for the Levakovic approximation with other models typically close to 10^{-9}. The degree 12 approximations can be viewed for $T < 50$ in Figure 8-2.

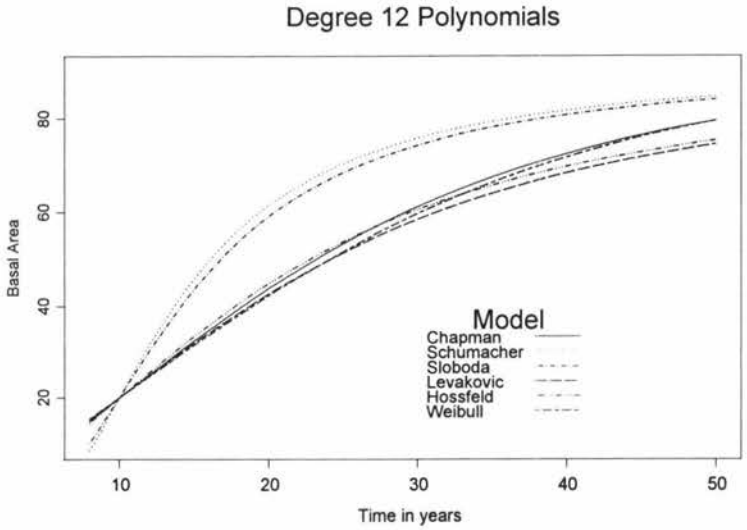


Figure 8-2: The Padé approximations $R_{24}(T)$ displayed for $T \in [8, 50]$.

The clear similarity of the Schumacher and Sloboda approximations is evident.. Over this range of T the four other models behave somewhat similarly. This is contrasted with the behaviour for $T > 50$ seen in Figure 8-3 where three pairings become apparent: firstly the Sloboda and Schumacher models, secondly the Weibull and Chapman-Richards models and thirdly the Levakovic and Hossfeld models.

The approximations of the Sloboda, Schumacher, Hossfeld and Levakovic models display much greater consistency and in general increase monotonically to their asymptote level. There is marked similarity between the Sloboda and Schumacher approximations and these are almost superimposed in Figure 8-3. This is almost the case for the Levakovic and Hossfeld approximations.

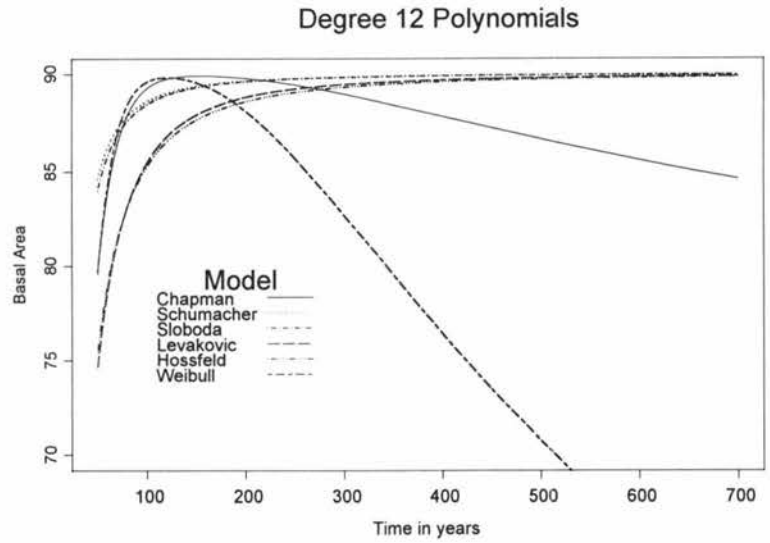


Figure 8-3: The Padé approximations $R_{24}(T)$ displayed for $T \in [50, 700]$.

In summary, the approximations behave similarly in pairs: firstly, the Chapman-Richards and Weibull models, secondly, the Sloboda and Schumacher models and finally the Levakovic and Hossfeld models and we note that each of these pairs have somewhat similar formulae. These groupings are somewhat different to the differential equation classification of Chapter 2 where the Hossfeld, Schumacher and Levakovic models were seen to have the same general form. The Chapman-Richards and Gompertz models were of a second form; the Sloboda model was a combination of the two forms and the Weibull model was of neither form.

The information gained from the plots can be summarised as follows

1. A strong feature is the consistent similarity between the Schumacher and Sloboda approximations. We recall that when predictions were made far from the data in Section 6.6 the Schumacher and Sloboda models gave best prediction accuracy (see Table 6-14 on page 110).

2. The Schumacher and Sloboda give estimates of asymptote level close to 90.

3. The Hossfeld and Levakovic models are similar particularly when viewed after $T = 50$ as shown in Figure 8-3.

4. The asymptote levels of the Hossfeld approximations are close to 90.

5. The Chapman-Richards and Weibull approximations behave similarly but somewhat erratically compared to the other four models.

These points can be contrasted with the robustness of the models when fitted to forestry data

1. The Chapman-Richards model was seen to give accurate predictions close to the data and had quite satisfactory goodness-of-fit properties. The Weibull model was consistently quite poor.

2. The Sloboda model was seen to give accurate predictions and had quite satisfactory goodness-of-fit properties yet the Schumacher model was consistently quite poor although it did have good predictive power when making predictions far from the data.

3. The Hossfeld model had adequate goodness-of-fit and predictive accuracy and the Levakovic model was very good in both respects. It was noted that the Levakovic model was often difficult to fit to forestry data with convergence sometimes not possible.

8.4 The Padé approximations to explain nonlinear least squares fitting

We will show that the nonlinear least squares fit of the growth curve models is largely determined by the properties at the point of inflection. That is, the location of the point of inflection and the first and higher order derivatives at the point of inflection are fundamental to the nonlinear fit.

The nonlinear least squares fit for a particular growth curve model is approximately the true curve and we have noted that the Padé approximation errors are very small so that the Padé approximation is the true curve over the range of interest. We therefore assert that the nonlinear fit is determined by the properties at the point of inflection since the Padé approximation *is* based entirely on these properties.

Suppose we have a data set where we estimate the location of the point of inflection as $(10, 20)$ and the slope at this point as three. We develop degree 12 Padé approximations

(about the point of inflection) of growth curves with these properties and these are displayed for $T \in [8, 26]$ in Figure 8-4 and for $T \in [26, 50]$ in Figure 8-5.

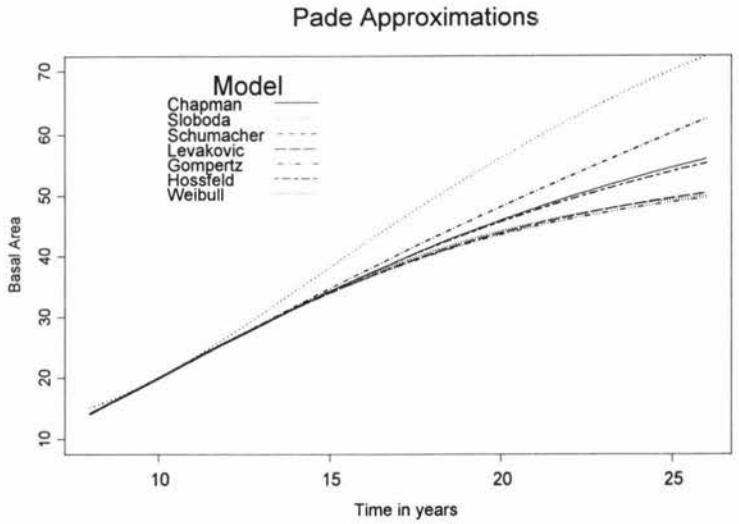


Figure 8-4: The Padé approximations $R_{24}(T)$ displayed for $T \in [8, 26]$. These approximations have been developed from growth curve models with a slope of three at a point of inflection $(10, 20)$.

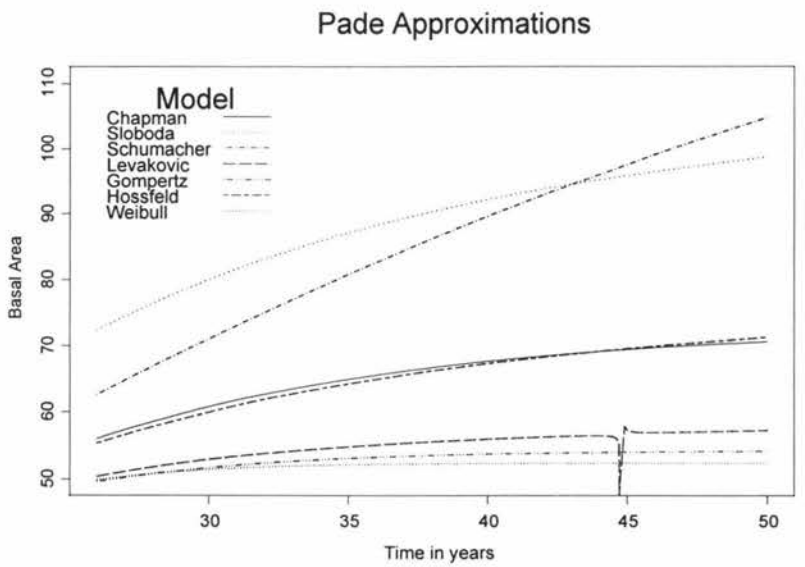


Figure 8-5: The Padé approximations $R_{24}(T)$ displayed for $T \in [26, 50]$

From Figures 8-4 and 8-5 we can note the following points:

1. The Chapman-Richards and Hossfeld models are quite central and this has been exhibited in the pinus radiata data.

2. The Sloboda and Schumacher models run high, particularly the Schumacher model for larger T and this has been noted as a property evident in the eight additional stands of pinus radiata data analysed in Section 6-4.

3. The Weibull and Gompertz models run low and seem unable to capture the properties of pinus radiata data for even moderate T.

The nonlinear least square fit is however not as simple as this with the location of the estimated point of inflection adjusted to allow the models to better fit the data for large T. Fairly typical locations of estimated point of inflection for pinus radiata data, ordered by T-coordinate, are shown in Table 8-2.

Table 8-2: Estimates of the location of the point of inflection when growth curve models are fitted to pinus radiata data. The slope at this point is fairly consistent and close to 2.7.

	Schumacher	Levakovic	Chapman-Richards	Hossfeld	Gompertz	Weibull	Sloboda
T-coordinate	13.2	14.6	14.6	15.1	15.1	16.3	16.6
y-coordinate	15.5	18.6	18.7	19.8	19.9	22.8	24.5

We determine Padé approximations of the growth curve models with properties at the inflection point specified in Table 8-2 and these are displayed for $T \in [8, 26]$ in Figure 8-6 and for $T \in [26, 50]$ in Figure 8-7.

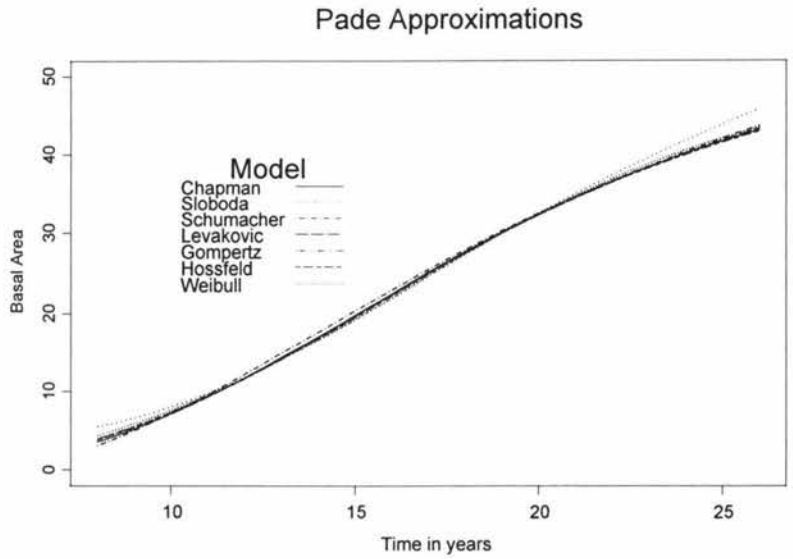


Figure 8-6: The Padé approximations $R_{24}(T)$ displayed for $T \in [8, 26]$. These curves have been developed from growth curve models with properties specified in Table 8-2.

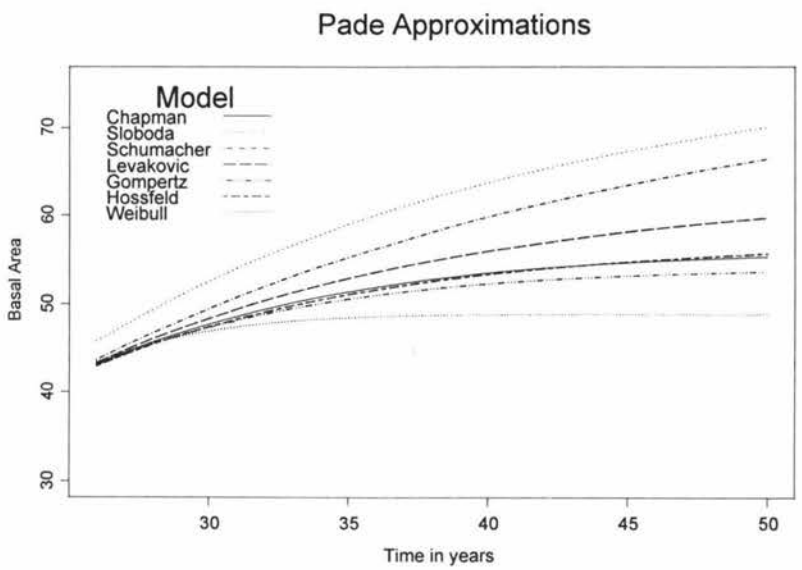


Figure 8-7: The Padé approximations $R_{24}(T)$ displayed for $T \in [26, 50]$. These curves have been developed from growth curve models with properties specified in Table 8-2.

The comparative behaviour of the models is not now as extreme as in Figures 8-4 and 8-5 but the general comments made are applicable here. We do note that the centrality of the

Levakovic model is now evident and this model has good predictive power close to the data as does the Chapman-Richards model. The Weibull model, and to a lesser extent the Gompertz model, always runs low and appears to be quite unsatisfactory even close to the data between 30 and 35 years.

8.5 Conclusion

It is apparent that if predictions are to be made based on early data the information available can in a sense be compressed into information at the point of inflection. For predictions not far from the data the Chapman-Richards model seems central and appropriate. Where predictions are to be made at large T, far from the data, the Schumacher and Sloboda models appear to be more appropriate. we cannot say that any single curve matches the data well but we can eliminate the Weibull model from consideration.

This may provide some direction for further investigation. rational functions that capture the properties of the data at the inflection point could be fitted. That is, we estimate the properties at the point of inflection and fit the associated rational function.

Part VI

Conclusion

Chapter 9

Summary and conclusions

Before providing a detailed summary of the goodness-of-fit and predictive power results for the seven models there are several points arising from the research in this thesis that need to be highlighted.

1. The two-parameter reparameterization of the Chapman-Richards model was very simple to fit with convergence not very sensitive to starting values.

2. We have shown that the Schumacher model consistently estimate the largest asymptote level with the Sloboda model giving the next largest. The smallest asymptotes will be generated by the Gompertz and Weibull models with the Weibull model always smallest.

3. Rational functions that capture inflection point properties may provide an additional class of functions useful in modelling forestry data.

As stated, the Schumacher and Sloboda models will consistently give larger asymptote levels than the other models but when we analyse the goodness-of-fit and predictive power we cannot conclude that models with larger asymptotes are superior candidate functions than those that estimate smaller asymptote levels, particularly when predictions are made at relatively small ages. This will be demonstrated in summary by presenting an ordering of the growth curve in terms of their effectiveness in modelling pinus radiata basal area data and making predictions based on this data. Table 9-1 gives a summary of the orderings of the seven growth curve models, from best to worst, for:

1. The goodness of fit when the models were fitted to 22 stands of basal area data;

2. Prediction robustness - pinus radiata data;

3. Prediction robustness - four types of simulated data;

Table 9-1: Rating in descending order of performance of seven growth curve models studied.

	Pinus radiata data		
	Goodness of Fit Section 4.10	Prediction Robustness (30 years) Section 6.5	Prediction Robustness (35 to 40 years) Section 6.6
1	Chapman-Richards	Chapman-Richards	Schumacher
2	Levakovic	Sloboda	Chapman-Richards
3	Sloboda	Levakovic	Sloboda
4	Hossfeld	Hossfeld	Levakovic
5	Schumacher	Schumacher	Gompertz
6	Gompertz	Gompertz	Hossfeld
7	Weibull	Weibull	Weibull

	Simulated data Robustness, Section 7-4			
	Case I (28 years)	Case II (up to 40 years)	Case III (up to 40 years)	Case IV (up to 40 years)
1	Sloboda	Chapman-Richards	Levakovic	Levakovic
2	Levakovic	Levakovic	Chapman-Richards	Chapman-Richards
3	Chapman-Richards	Sloboda	Hossfeld	Hossfeld
4	Hossfeld	Hossfeld	Sloboda	Weibull
5	Weibull	Weibull	Weibull	Sloboda
6	Schumacher	Schumacher	Schumacher	Gompertz
7				Schumacher

Based on Table 9-1, we will rate the seven growth curves studied in this thesis. In providing this overall rating more weight is given to the results obtained from the pinus radiata data to those obtained from the simulated data. The following table is a final rating from best to worst.

Chapman-Richards
Levakovic
Sloboda
Schumacher
Hossfeld
Gompertz
Weibull

This rating leads us to state that of the seven models considered, the Chapman-Richards, Levakovic and Sloboda models seem reasonable candidate functions in the modelling of pinus radiata basal area data. It is important to note that where the slope at the inflection point is large the Sloboda model will often be more precise than the Chapman-Richards model. Also, the Levakovic model must be used with caution - it is often difficult to fit and is very sensitive to choice of starting values. The Schumacher model must be also considered - it had good predictive power when the eight additional data sets were analysed in Section 6.6. The results may not be reproduced for other species or pinus radiata from other regions, but the index provided above may provide a useful starting point in the modelling of basal area data. Additionally, it is important to assess curvature for the Levakovic and Sloboda models. For the second data set, consisting of eight stands, these two models were seen to have unacceptably high parametric-effects curvature.

These results contrast strongly with those reported by Woollons & Wood [7] where the five 3-parameter models discussed here were investigated. It was determined that the Schumacher and Weibull models were superior candidates than the Gompertz, Hossfeld and Chapman-Richards models. It is noted that this study focussed on goodness of fit rather than prediction accuracy but our goodness of fit results clearly indicate that the Chapman-Richards is superior to both the Schumacher and Weibull models and indeed the Gompertz model is superior to the Weibull model which is consistently poor.

Woollons, Whyte and Liu [8] gave results of an analysis of the Gompertz, Schumacher and Hossfeld models in predicting basal area and reported a standard error of estimate of basal area to be less than 2 m^2/ha for all three and concluded that the Hossfeld model is a viable model that needs to be considered. In Appendix H, we also report quite small standard errors of

prediction (less than 0.7 m^2/ha) for these three models and further conclude that the Hossfeld model is more precise than either the Schumacher or Gompertz model.

Woollons et al., state in the conclusion to their paper [8] as follows:

> "We believe it is unrealistic to expect a unique sigmoid function to consistently perform better than others with forest growth and yield data. A more rational approach is to be aware of the existence of several candidate equations, and to explore their utility with the data in question."

This statement is endorsed and the rating provided above may provide a direction of exploration of possible candidate functions.

Bibliography

[1] Venables, W. N. and Ripley, B. D. (1997) *Modern Applied Statistics with S-Plus.* Springer-Verlag New York, Inc.

[2] Ratkowsky, D. A.. (1983) *Nonlinear Regression Modeling.* Marcel Dekker, Inc.

[3] Ratkowsky, D. A.. (1990) *Handbook of Nonlinear Regression Models.* Marcel Dekker, Inc.

[4] Bates, D. M. and Watts, D. G. (1988) *Nonlinear Regression Analysis & Its Applications.* John Wiley & Sons, Inc.

[5] Gerald, C. F. and Wheatley, P. O. (1994) *Applied Numerical Analysis.* Addison-Wesley Pub. Co.

[6] Zeide, B. (1993) Analysis of growth equations. *Forest Science* **39**, 594-616.

[7] Woollons, R. C. and Wood, G. R. (1992) Utility and performance of five sigmoid yield-age functions, fitted to stand growth data. *IUFRO Conference* 13-17 January, Canberra, Australia.

[8] Woollons, R. C., Whyte, A. G. D. and Liu Xu (1991) The Hossfeld function: an alternative for depicting stand growth and yield. *Japanese Journal of Forestry* **15**, 25-35.

[9] Kiviste, A. K. (1988) Mathematical functions of forest growth. *Estonian Agricultural Academy.* Tartu, 23-33.

[10] Mathsoft, Inc. (1997) *S-Plus 4 Guide to Statistics.*

[11] Wolfram, S. (1991) *Mathematica, a system for doing mathematics by computer.* Addison-Wesley Publishing Co., Inc.

[12] Ryan, T. P. (1997) *Modern Regression Methods.* John Wiley & Sons, Inc.

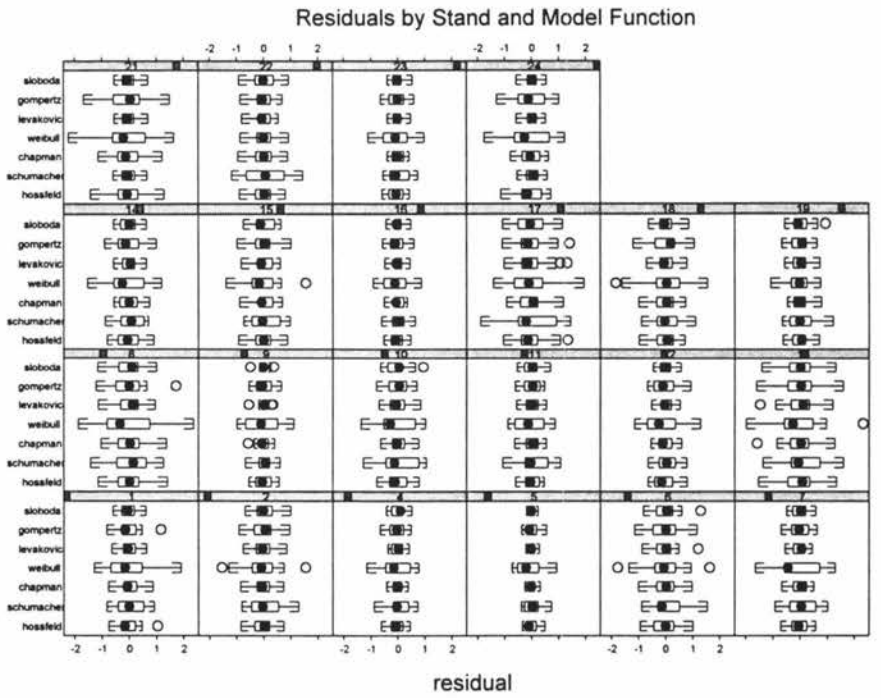
[13] Beale, E. M. L. (1960) Confidence intervals in nonlinear estimation (with discussion). *Journal of the Royal Statistical Society* B **22**, 41-88.

[14] Ross, G. J. S. (1970) The efficient use of function minimization in non-linear maximum likelihood estimation. *Applied Statistics* **40**, 269-282.

[15] Guttenberg, A. R., von. (1915) Growth and yield of spruce in Hochgebirge. Franz Deuticke, Wien.

Appendix A

Residuals of Models Fit to Each Stand



Appendix B

Parameter Estimates

In most cases, reparameterizations of the sigmoid models were used in fitting by nonlinear least squares. The parameter estimates detailed here have been obtained from the transformations used and are estimates relating to the original, simple form of each sigmoid function.

B.1 The Chapman-Richards model

Stand	*a*	*b*	*c*
1	65.39	-0.13	6.48
2	66.39	-0.12	5.35
4	61.78	-0.12	5.46
5	52.83	-0.11	5.81
6	75.37	-0.12	4.43
7	66.43	-0.14	6.27
8	70.80	-0.14	5.38
9	53.31	-0.12	6.23
10	74.22	-0.13	5.74
11	62.80	-0.12	5.63
12	69.27	-0.11	5.29
13	67.38	-0.14	5.80
14	69.27	-0.12	5.15
15	65.92	-0.12	5.55
16	56.22	-0.11	5.12
17	67.35	-0.15	6.20
18	62.89	-0.14	4.96
19	51.01	-0.12	5.48
21	71.75	-0.12	4.17
22	56.64	-0.12	4.96
23	64.39	-0.11	5.11
24	73.52	-0.11	4.40

B.2 The Gompertz model

Stand	*a*	*b*	*c*
1	63.61	8.75	0.15
2	64.21	7.48	0.14
4	59.23	7.77	0.13
5	50.33	8.29	0.13
6	73.00	6.32	0.14
7	65.02	8.36	0.16
8	69.09	7.36	0.15
9	51.44	8.62	0.14
10	72.12	7.88	0.14
11	60.78	7.82	0.14
12	65.62	7.71	0.12
13	65.79	7.86	0.15
14	66.82	7.29	0.14
15	63.22	7.87	0.13
16	53.07	7.54	0.12
17	66.21	8.15	0.16
18	61.41	6.82	0.15
19	49.21	7.69	0.14
21	68.95	6.10	0.13
22	54.77	7.00	0.14
23	61.36	7.41	0.13
24	69.67	6.57	0.12

B.3 The Hossfeld model

Stand	a	b	c
1	67.32	203.67	3.3
2	69.24	88.04	3.0
4	63.95	138.05	3.1
5	53.95	274.14	3.2
6	79.90	28.84	2.7
7	68.86	127.93	3.2
8	74.06	61.51	3.0
9	54.76	267.78	3.2
10	77.12	100.03	3.1
11	65.15	125.09	3.1
12	71.32	135.23	3.0
13	70.14	93.21	3.1
14	72.44	74.20	3.0
15	68.06	140.65	3.1
16	57.97	150.02	3.0
17	69.95	99.30	3.2
18	66.15	45.25	2.9
19	52.97	145.81	3.1
21	76.49	26.42	2.6
22	59.35	75.05	2.9
23	66.84	106.19	3.0
24	77.61	45.31	2.7

B.4 The Schumacher model

Stand	*exp(a)*	*b*	*c*
1	90.99	77.99	1.5
2	95.91	51.69	1.3
4	95.69	50.82	1.3
5	86.11	54.89	1.2
6	105.68	38.55	1.3
7	87.46	78.40	1.6
8	94.42	57.77	1.4
9	78.57	68.85	1.4
10	104.24	60.80	1.4
11	91.42	56.37	1.3
12	115.64	45.59	1.2
13	90.17	65.80	1.5
14	101.35	48.21	1.3
15	102.25	51.80	1.3
16	96.01	42.37	1.1
17	85.66	80.31	1.6
18	82.97	50.45	1.4
19	76.10	52.03	1.3
21	104.19	34.36	1.2
22	83.41	43.27	1.3
23	103.43	43.83	1.2
24	118.82	34.33	1.1

B.5 The Weibull model

Stand	*a*	*b*	*c*
1	58.44	-0.00031	2.7
2	59.25	-0.00057	2.5
4	53.70	-0.00037	2.6
5	44.81	-0.00022	2.7
6	68.80	-0.00138	2.3
7	60.78	-0.00051	2.6
8	65.02	-0.00087	2.4
9	46.72	-0.00026	2.7
10	66.58	-0.00048	2.6
11	55.69	-0.00044	2.6
12	58.90	-0.00032	2.6
13	61.51	-0.00062	2.5
14	61.76	-0.00063	2.5
15	57.19	-0.00035	2.6
16	47.58	-0.00034	2.6
17	62.41	-0.00066	2.6
18	58.22	-0.00126	2.3
19	44.86	-0.00043	2.6
21	65.17	-0.00149	2.2
22	50.41	-0.00070	2.5
23	55.66	-0.00044	2.6
24	64.39	-0.00081	2.4

B.6 The Levakovic model

Stand	*a*	*b*	*c*	*d*
1	72.77	6.85	1.71	2.6
2	73.42	6.87	1.45	2.5
4	70.20	6.57	1.64	2.4
5	61.08	6.53	1.80	2.3
6	85.68	5.65	1.64	2.2
7	72.51	7.00	1.54	2.6
8	79.51	5.91	1.79	2.3
9	62.96	5.70	2.32	2.2
10	79.67	7.78	1.25	2.8
11	66.79	8.15	1.16	2.8
12	85.58	5.26	2.34	2.0
13	74.11	6.72	1.55	2.5
14	80.01	5.80	1.86	2.2
15	73.02	7.18	1.45	2.5
16	69.44	5.31	2.24	2.0
17	70.24	8.64	1.04	3.1
18	70.89	5.56	1.82	2.2
19	53.57	8.57	1.07	2.9
21	101.08	0.82	18.79	1.3
22	55.67	11.07	0.67	3.7
23	75.27	6.01	1.80	2.2
24	103.84	2.66	5.52	1.4

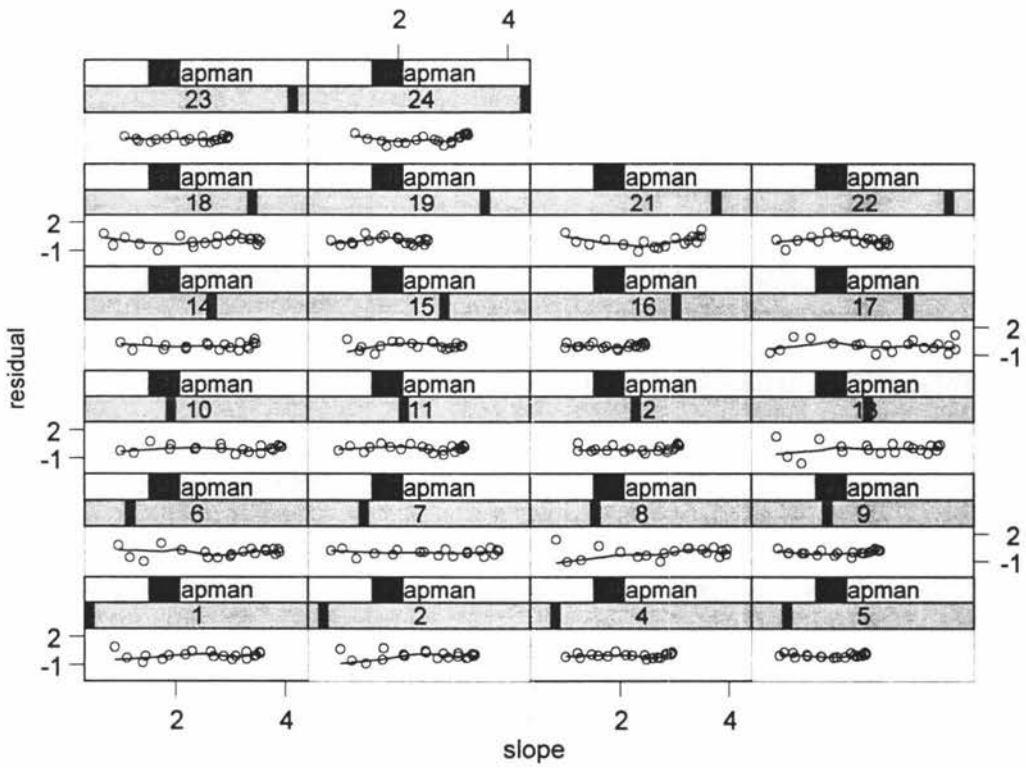
B.7 The Sloboda model

Stand	*a*	*b*	*c*	*d*
1	1.30	-4.08	115.19	-2.1
2	1.40	-4.04	66.84	-1.9
4	0.91	-4.43	58.48	-1.9
5	0.72	-4.52	71.43	-1.9
6	1.74	-3.95	39.62	-1.8
7	1.73	-3.78	120.00	-2.2
8	1.63	-3.92	67.28	-2.0
9	0.52	-4.85	63.81	-1.9
10	2.17	-3.68	111.75	-2.1
11	2.06	-3.56	128.25	-2.1
12	0.57	-5.07	38.01	-1.7
13	1.72	-3.81	92.70	-2.1
14	1.05	-4.39	47.33	-1.8
15	1.25	-4.16	75.07	-1.9
16	0.46	-5.08	34.91	-1.7
17	3.17	-3.16	204.52	-2.4
18	1.38	-3.98	49.68	-1.9
19	1.27	-3.87	87.13	-2.0
21	0.00	-22.44	1.85	-1.3
22	3.03	-3.05	153.55	-2.2
23	0.90	-4.50	47.39	-1.8
24	0.02	-8.40	7.82	-1.4

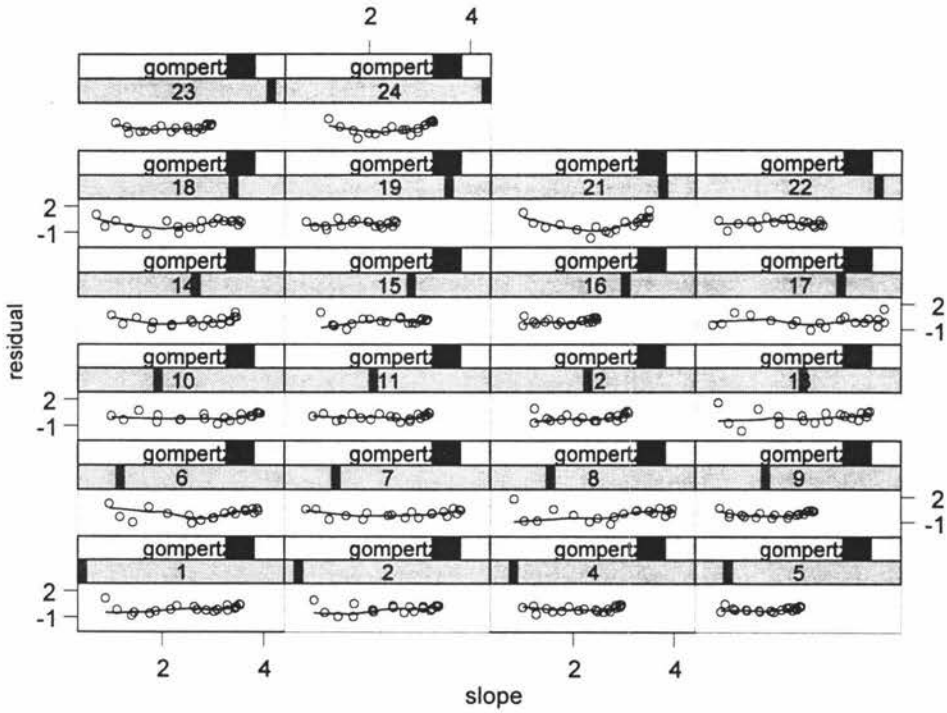
Appendix C

Plots of Residuals against Estimated Slope

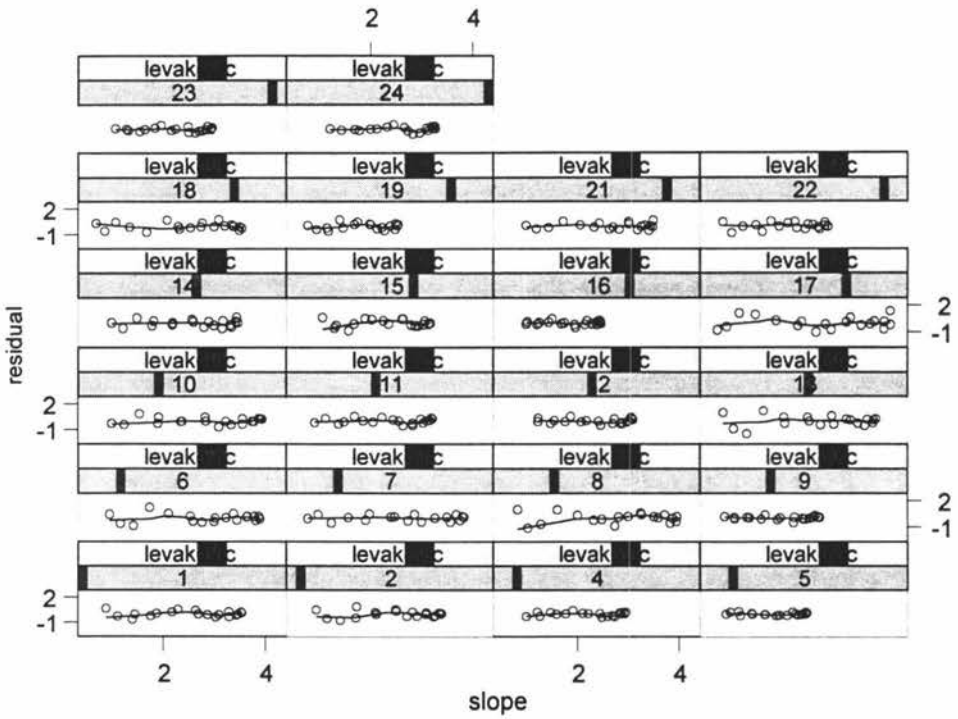
C.1 The Chapman-Richards model



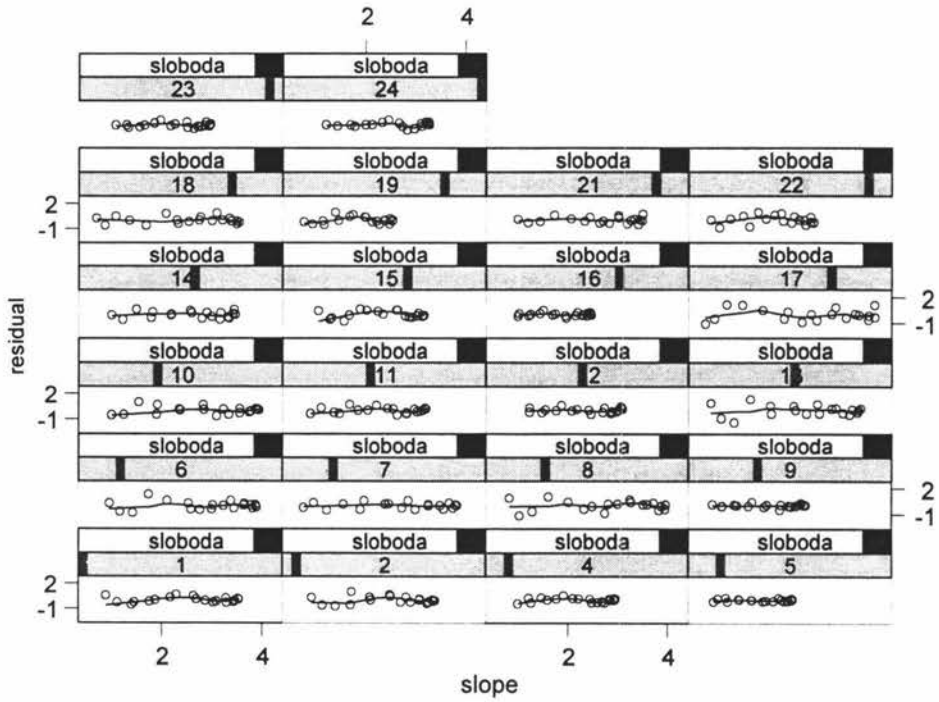
C.2 The Gompertz model



C.3 The Levakovic model



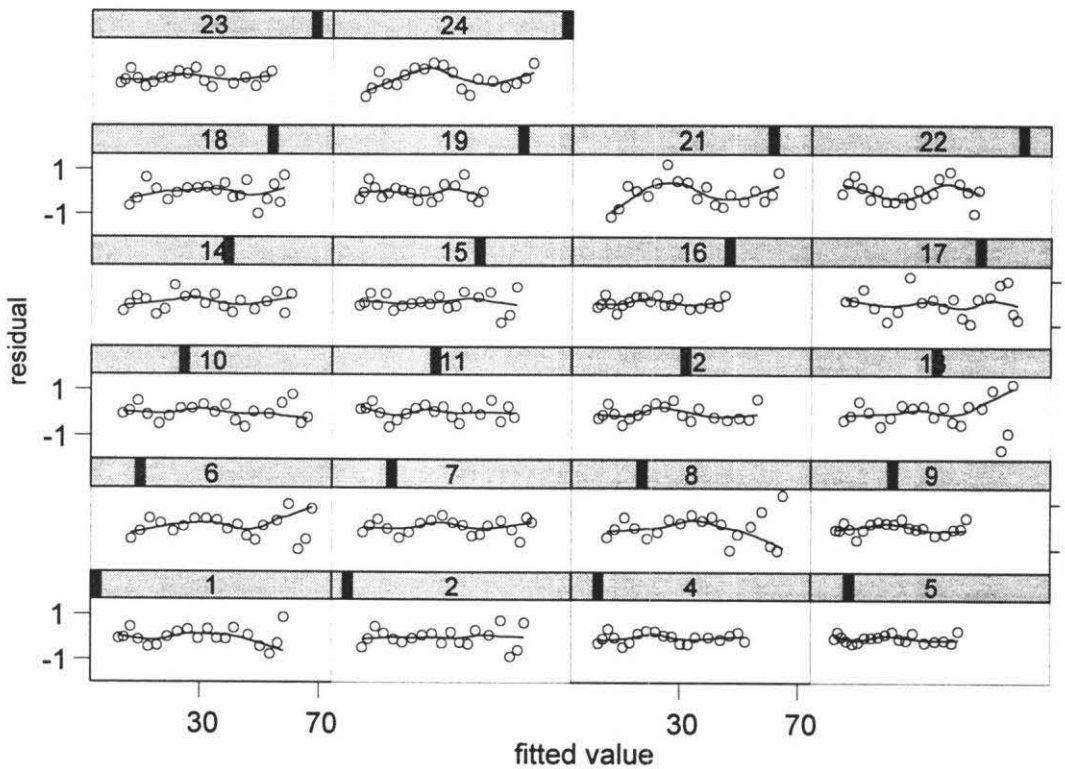
C.4 The Sloboda model



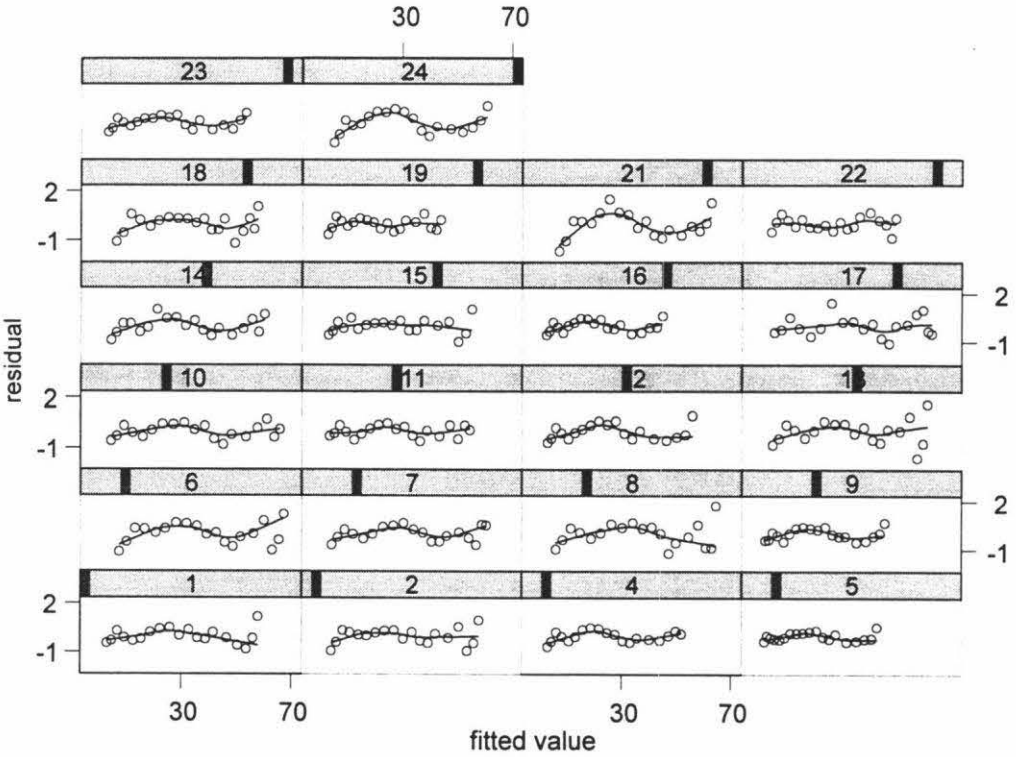
Appendix D

Plots of Residuals Against Fitted Values

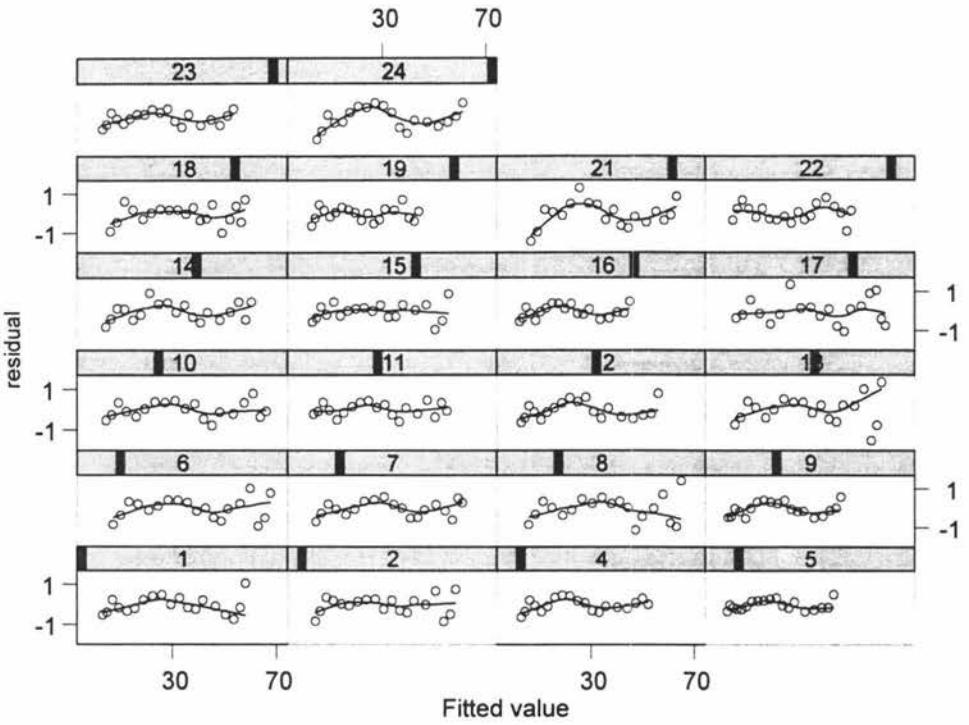
D.1 The Chapman-Richards model



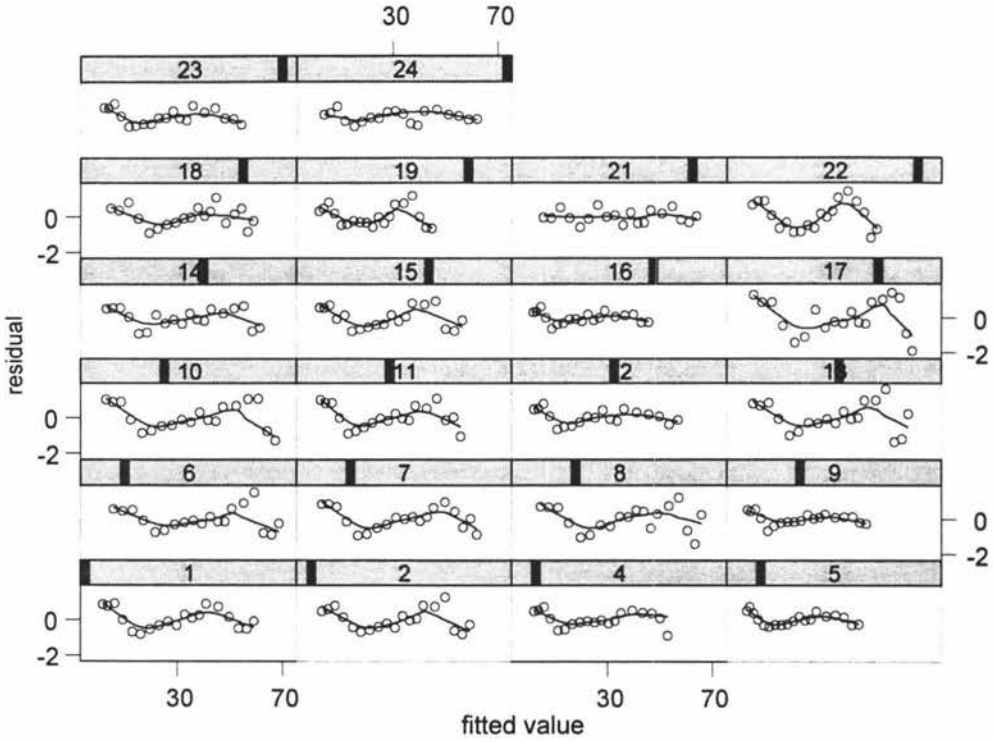
D.2 The Gompertz model



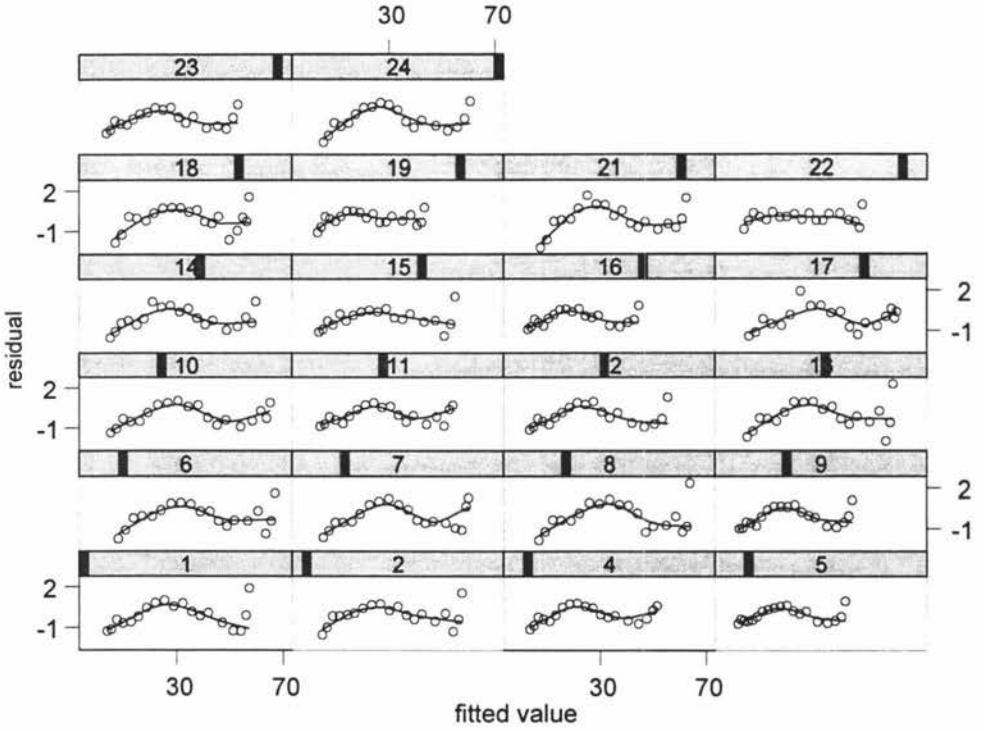
D.3 The Hossfeld model



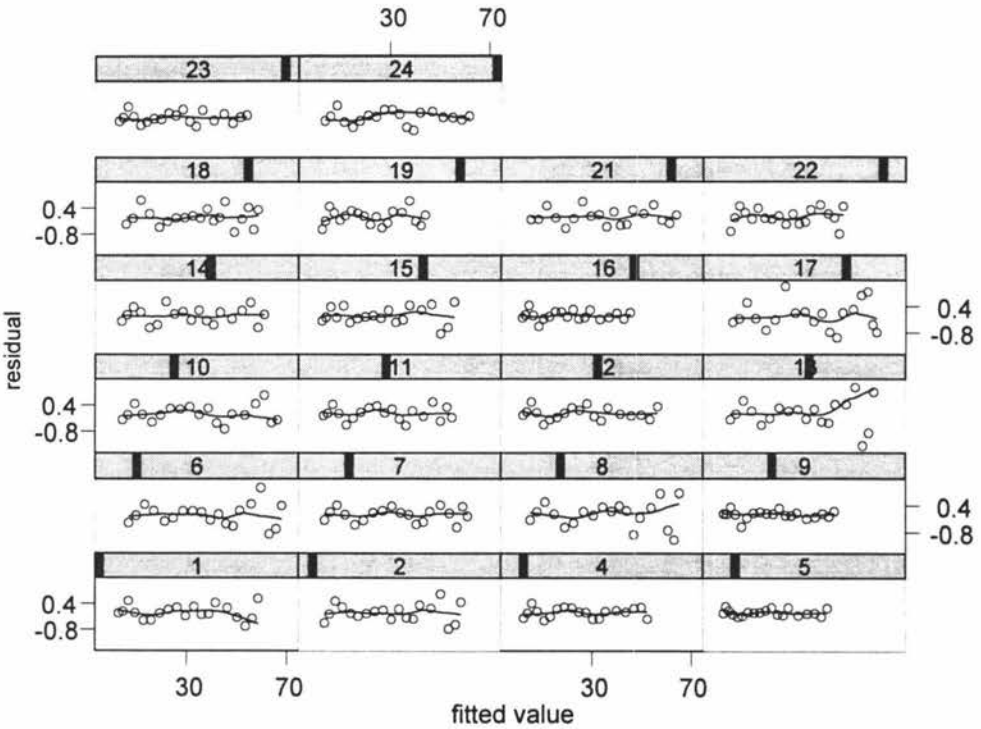
D.4 The Schumacher model



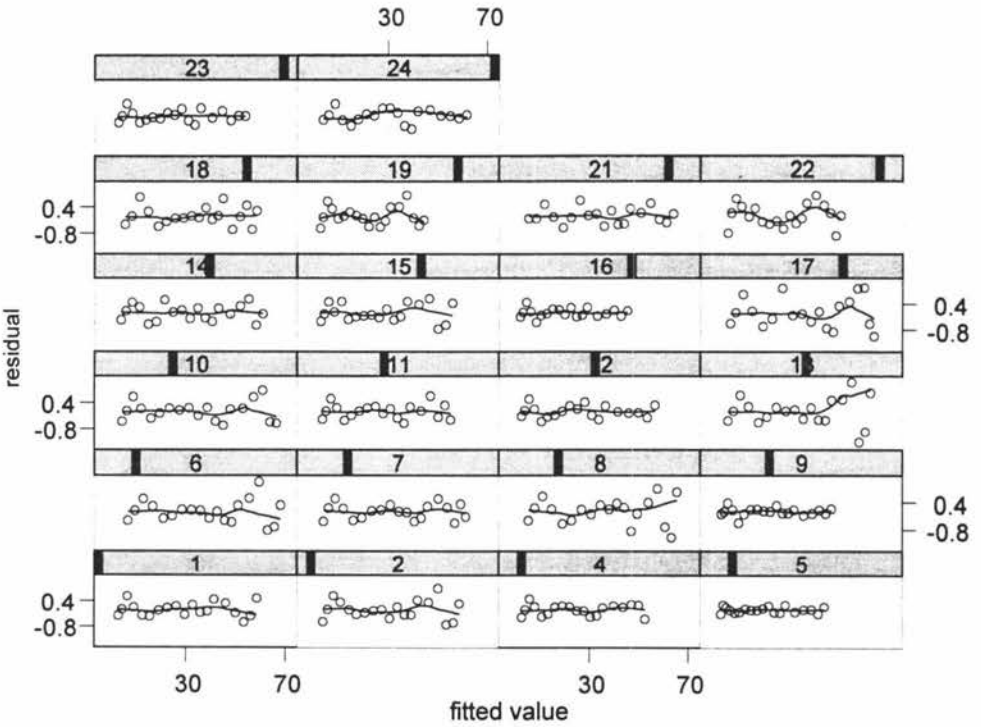
D.5 The Weibull model



D.6 The Levakovic model

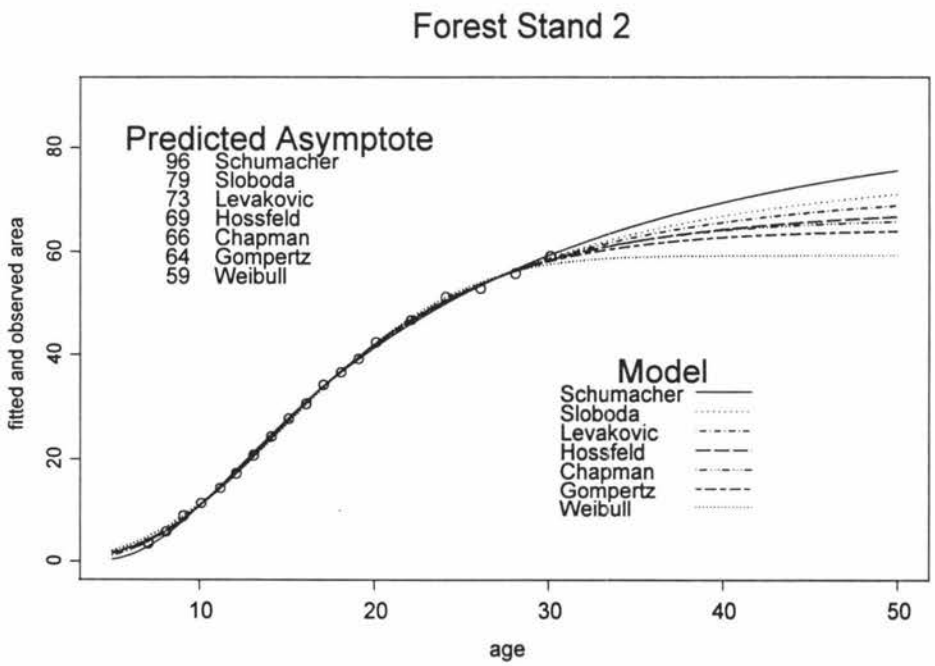


D.7 The Sloboda model

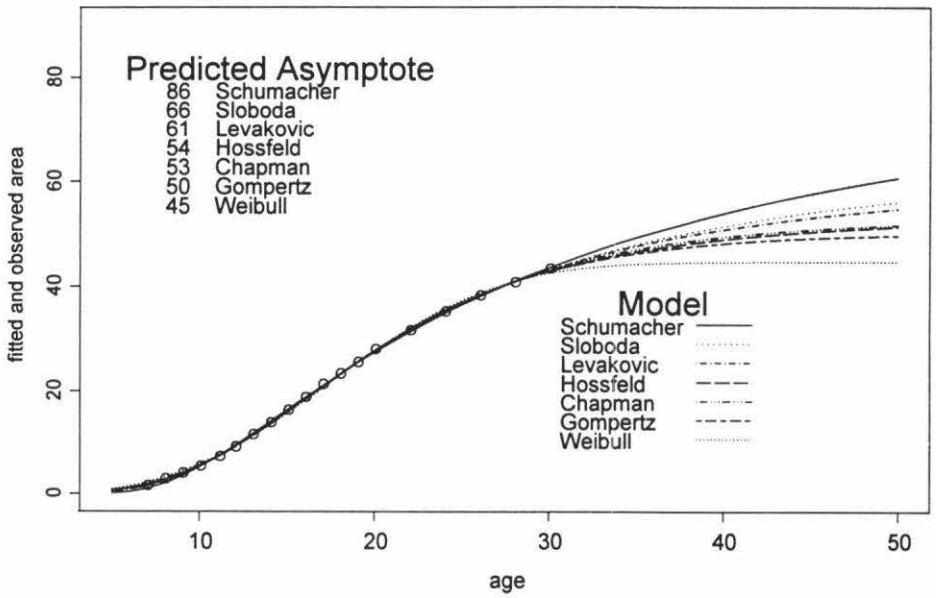


Appendix E

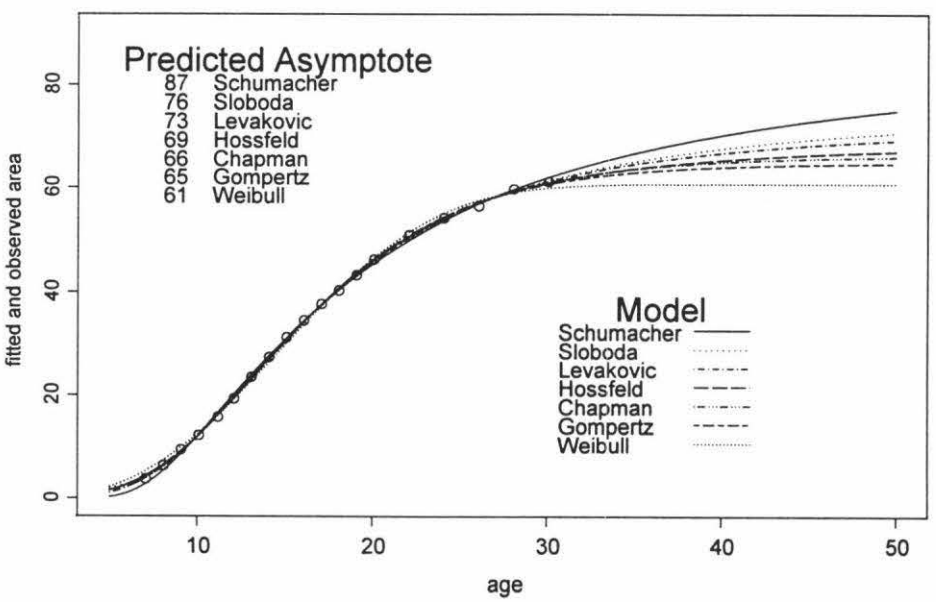
Fitted Models and Raw Data



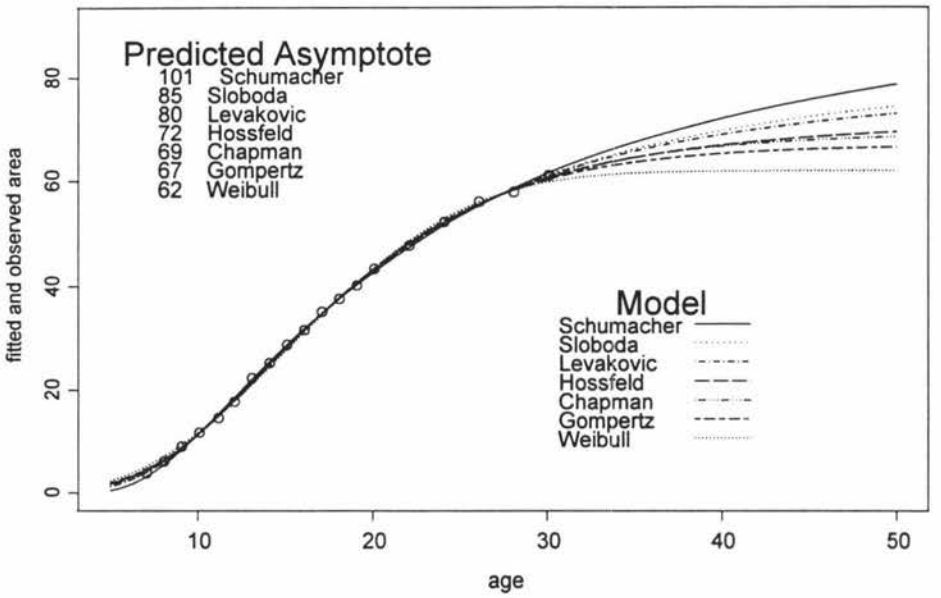
Forest Stand 5



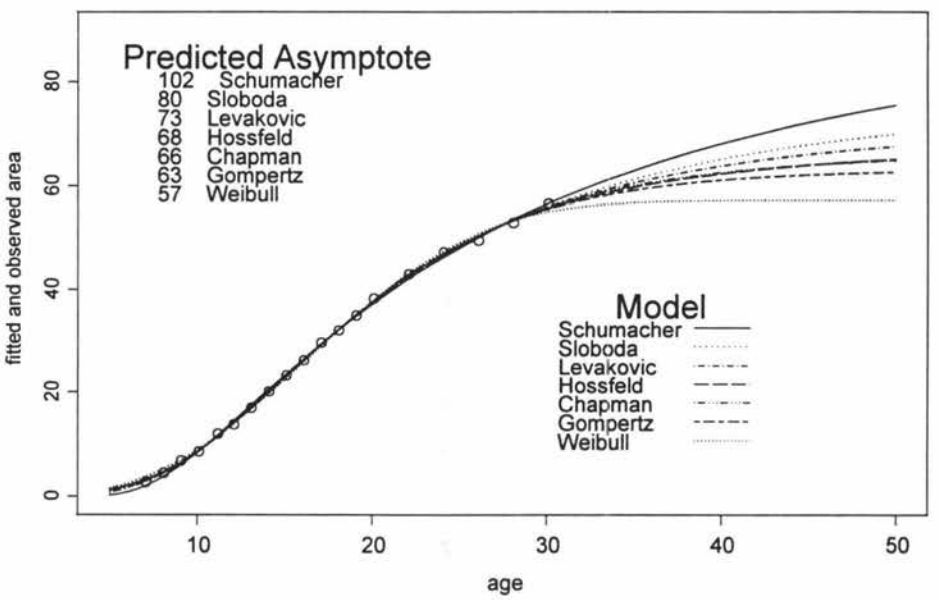
Forest Stand 7



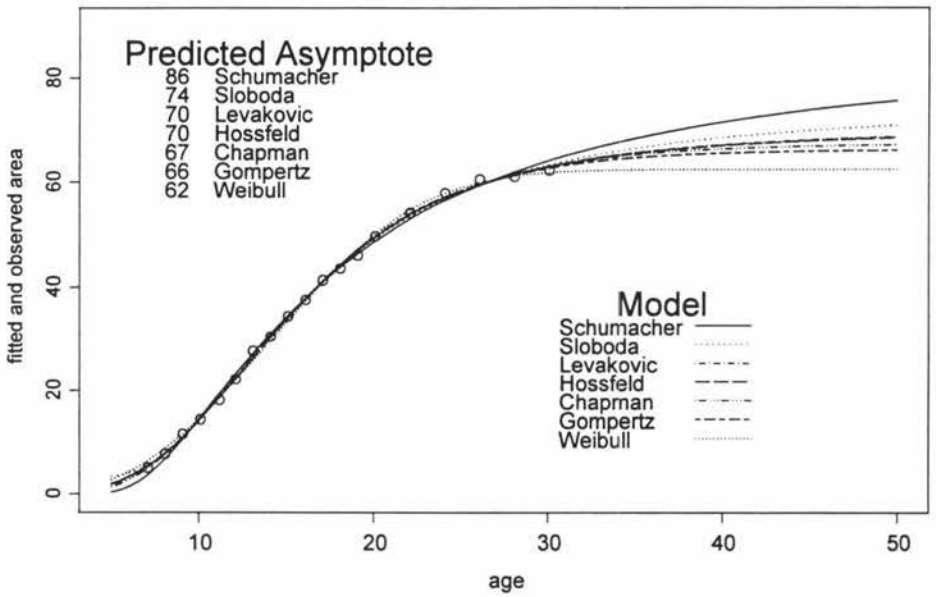
Forest Stand 14



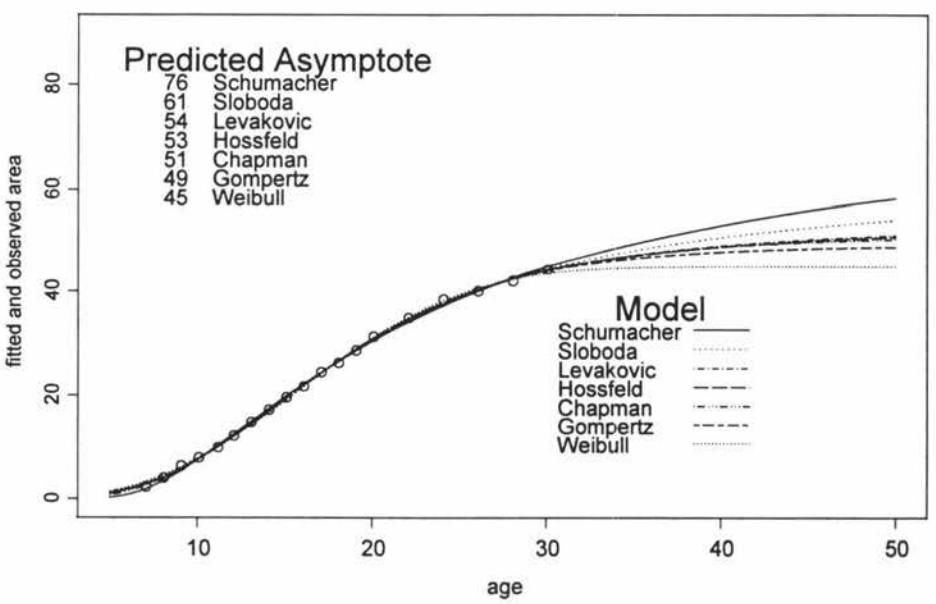
Forest Stand 15



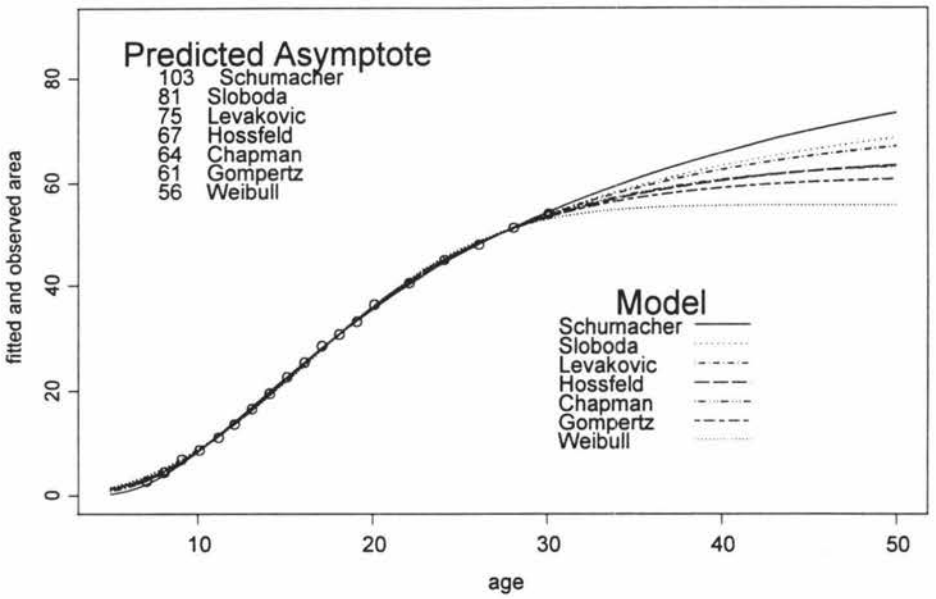
Forest Stand 17



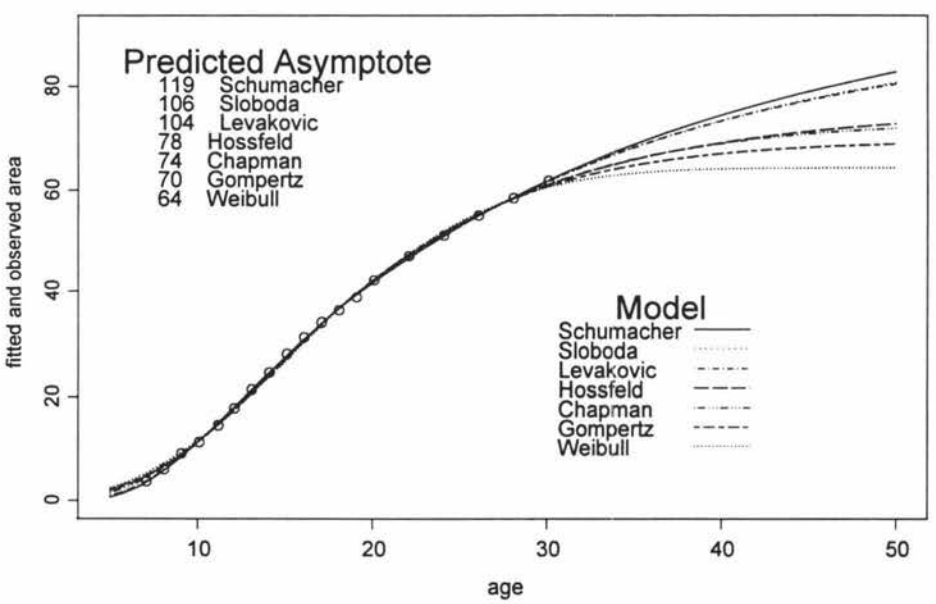
Forest Stand 19



Forest Stand 23



Forest Stand 24



Appendix F

Data for 28 Forestry Stands

Data from 28 forestry stands of pinus radiata with age measured in years and basal area in hectares is listed below. The background of the data is

1. The stands are part of a radiata pine thinning experiment established in 1971 in Kaingaroa Forest, New Zealand;
2. Six thinnings, to 200, 300, 400, 500, 600, and 700/ha, were carried out at age 7 years;
3. Four replicates of each thinning level were laid out in 0.2 ha plots;
4. The site indices (the mean top height in metres at age 20 years) at each thinning were 29.4, 30.6, 31.0, 30.2, 30.6, and 30.8 metres;
5. No late-age fertilization was carried out.

Stand	Age	Basal area	Stand	Age	Basal area	Stand	Age	Basal area	Stand	Age	Basal area
1	7.1	2.7	3	7.1	4.3	5	7.1	1.6	7	7.1	3.7
1	8.1	4.5	3	8.1	7.1	5	8.1	2.9	7	8.1	6.3
1	9.1	7.2	3	9.1	10.6	5	9.1	4.0	7	9.1	9.4
1	10.1	9.3	3	10.1	13.6	5	10.1	5.4	7	10.1	12.2
1	11.2	12.3	3	11.2	17.4	5	11.2	7.3	7	11.2	15.7
1	12.1	15.3	3	12.1	21.0	5	12.1	9.2	7	12.1	19.3
1	13.1	19.1	3	13.1	25.3	5	13.1	11.6	7	13.1	23.5
1	14.1	22.8	3	14.1	29.3	5	14.1	13.9	7	14.1	27.4
1	15.1	26.4	3	15.1	33.0	5	15.1	16.3	7	15.1	31.3
1	16.1	29.4	3	16.1	36.5	5	16.1	18.8	7	16.1	34.5
1	17.1	33.1	3	17.1	40.5	5	17.1	21.3	7	17.1	37.7
1	18.1	35.8	3	18.1	43.0	5	18.1	23.3	7	18.1	40.3
1	19.1	38.7	3	19.1	46.2	5	19.1	25.5	7	19.1	43.2
1	20.1	41.9	3	20.1	50.0	5	20.1	28.0	7	20.1	46.2
1	22.1	46.4	3	22.1	55.0	5	22.1	31.6	7	22.1	50.9
1	24.1	49.9	3	24.1	60.1	5	24.1	35.2	7	24.1	54.1
1	26.1	52.8	3	26.1	57.5	5	26.1	38.3	7	26.1	56.4
1	28.1	55.9	3	28.1	60.4	5	28.1	40.8	7	28.1	59.7
1	30.1	59.1	3	30.1	63.6	5	30.1	43.5	7	30.1	61.2
2	7.1	3.4	4	7.1	2.4	6	7.1	6.4	8	7.1	5.0
2	8.1	5.8	4	8.1	4.1	6	8.1	9.7	8	8.1	8.1
2	9.1	8.8	4	9.1	6.4	6	9.1	13.6	8	9.1	11.8
2	10.1	11.3	4	10.1	8.2	6	10.1	17.0	8	10.1	14.9
2	11.2	14.4	4	11.2	10.5	6	11.2	20.8	8	11.2	18.6
2	12.1	17.2	4	12.1	13.1	6	12.1	24.5	8	12.1	22.4
2	13.1	20.7	4	13.1	16.3	6	13.1	28.7	8	13.1	26.9
2	14.1	24.2	4	14.1	19.3	6	14.1	32.5	8	14.1	30.6
2	15.1	27.6	4	15.1	22.2	6	15.1	36.1	8	15.1	34.7
2	16.1	30.4	4	16.1	24.9	6	16.1	39.2	8	16.1	38.0
2	17.1	34.0	4	17.1	27.7	6	17.1	42.7	8	17.1	41.5
2	18.1	36.5	4	18.1	30.1	6	18.1	45.3	8	18.1	44.3
2	19.1	39.2	4	19.1	32.7	6	19.1	48.0	8	19.1	46.0
2	20.1	42.4	4	20.1	35.5	6	20.1	51.3	8	20.1	49.3
2	22.1	46.8	4	22.1	40.0	6	22.1	56.2	8	22.1	54.2
2	24.1	51.3	4	24.1	43.8	6	24.1	60.8	8	24.1	58.5
2	26.1	52.9	4	26.1	47.3	6	26.1	62.0	8	26.1	59.9
2	28.1	55.8	4	28.1	50.2	6	28.1	65.0	8	28.1	62.0
2	30.1	59.1	4	30.1	52.1	6	30.1	68.4	8	30.1	66.2

Stand	Age	Basal area	Stand	Age	Basal area	Stand	Age	Basal area	Stand	Age	Basal area
9	7.1	1.8	11	7.1	3.3	13	7.1	4.1	15	7.1	2.6
9	8.1	3.0	11	8.1	5.1	13	8.1	6.7	15	8.1	4.3
9	9.1	4.9	11	9.1	7.6	13	9.1	10.2	15	9.1	6.7
9	10.1	6.5	11	10.1	9.6	13	10.1	13.0	15	10.1	8.5
9	11.2	8.4	11	11.2	12.1	13	11.2	16.3	15	11.2	11.9
9	12.1	11.0	11	12.1	15.1	13	12.1	20.0	15	12.1	13.7
9	13.1	13.8	11	13.1	18.5	13	13.1	24.3	15	13.1	16.9
9	14.1	16.5	11	14.1	21.9	13	14.1	27.9	15	14.1	20.1
9	15.1	19.1	11	15.1	25.2	13	15.1	31.6	15	15.1	23.3
9	16.1	21.7	11	16.1	28.0	13	16.1	34.6	15	16.1	26.3
9	17.1	24.5	11	17.1	31.2	13	17.1	38.3	15	17.1	29.7
9	18.1	26.6	11	18.1	33.6	13	18.1	40.7	15	18.1	32.1
9	19.1	28.9	11	19.1	36.0	13	19.1	43.4	15	19.1	35.0
9	20.1	31.2	11	20.1	39.2	13	20.1	46.8	15	20.1	38.3
9	22.1	34.9	11	22.1	43.4	13	22.1	51.2	15	22.1	42.9
9	24.1	38.4	11	24.1	47.8	13	24.1	55.6	15	24.1	47.3
9	26.1	41.5	11	26.1	50.0	13	26.1	55.9	15	26.1	49.5
9	28.1	43.9	11	28.1	53.2	13	28.1	58.9	15	28.1	52.8
9	30.1	46.3	11	30.1	54.8	13	30.1	62.8	15	30.1	56.5
10	7.1	4.0	12	7.1	2.3	14	7.1	3.7	16	7.1	1.9
10	8.1	6.4	12	8.1	3.9	14	8.1	6.1	16	8.1	3.2
10	9.1	9.6	12	9.1	6.2	14	9.1	9.0	16	9.1	5.1
10	10.1	12.2	12	10.1	7.9	14	10.1	11.7	16	10.1	6.4
10	11.2	15.7	12	11.2	10.1	14	11.2	14.5	16	11.2	8.1
10	12.1	19.4	12	12.1	12.8	14	12.1	17.7	16	12.1	10.4
10	13.1	23.6	12	13.1	15.8	14	13.1	22.2	16	13.1	12.8
10	14.1	27.5	12	14.1	19.0	14	14.1	25.1	16	14.1	15.4
10	15.1	31.5	12	15.1	22.3	14	15.1	28.6	16	15.1	17.8
10	16.1	34.9	12	16.1	25.2	14	16.1	31.5	16	16.1	20.0
10	17.1	38.8	12	17.1	28.5	14	17.1	35.1	16	17.1	22.7
10	18.1	41.5	12	18.1	30.8	14	18.1	37.6	16	18.1	24.6
10	19.1	44.4	12	19.1	33.4	14	19.1	40.2	16	19.1	26.9
10	20.1	48.0	12	20.1	36.7	14	20.1	43.4	16	20.1	29.4
10	22.1	53.1	12	22.1	41.4	14	22.1	47.8	16	22.1	33.0
10	24.1	57.9	12	24.1	45.8	14	24.1	52.2	16	24.1	36.7
10	26.1	61.8	12	26.1	49.8	14	26.1	56.0	16	26.1	40.1
10	28.1	63.4	12	28.1	53.1	14	28.1	57.8	16	28.1	42.7
10	30.1	65.9	12	30.1	56.8	14	30.1	60.9	16	30.1	45.5

Stand	Age	Basal area	Stand	Age	Basal area	Stand	Age	Basal area	Stand	Age	Basal area
17	7.1	5.0	19	7.1	2.2	21	7.1	5.3	23	7.1	2.7
17	8.1	7.7	19	8.1	3.9	21	8.1	8.3	23	8.1	4.4
17	9.1	11.5	19	9.1	6.2	21	9.1	12.3	23	9.1	6.8
17	10.1	14.3	19	10.1	7.8	21	10.1	15.3	23	10.1	8.6
17	11.2	18.0	19	11.2	9.8	21	11.2	18.8	23	11.2	11.0
17	12.1	22.1	19	12.1	12.1	21	12.1	22.5	23	12.1	13.6
17	13.1	27.7	19	13.1	14.8	21	13.1	26.8	23	13.1	16.6
17	14.1	30.5	19	14.1	17.2	21	14.1	29.5	23	14.1	19.5
17	15.1	34.4	19	15.1	19.6	21	15.1	32.8	23	15.1	22.7
17	16.1	37.6	19	16.1	21.7	21	16.1	35.3	23	16.1	25.5
17	17.1	41.3	19	17.1	24.5	21	17.1	38.9	23	17.1	28.6
17	18.1	43.5	19	18.1	26.3	21	18.1	41.0	23	18.1	30.8
17	19.1	46.0	19	19.1	28.7	21	19.1	43.6	23	19.1	33.2
17	20.1	49.6	19	20.1	31.3	21	20.1	46.7	23	20.1	36.4
17	22.1	53.9	19	22.1	34.9	21	22.1	50.9	23	22.1	40.5
17	24.1	57.8	19	24.1	38.5	21	24.1	55.2	23	24.1	44.8
17	26.1	60.5	19	26.1	40.1	21	26.1	57.9	23	26.1	47.9
17	28.1	61.0	19	28.1	42.0	21	28.1	60.8	23	28.1	51.2
17	30.1	62.2	19	30.1	44.2	21	30.1	63.9	23	30.1	53.9
18	7.1	5.3	20	7.1	6.7	22	7.1	3.7	24	7.1	3.7
18	8.1	8.3	20	8.1	10.1	22	8.1	6.0	24	8.1	6.1
18	9.1	12.3	20	9.1	14.1	22	9.1	8.5	24	9.1	9.2
18	10.1	15.1	20	10.1	17.5	22	10.1	10.4	24	10.1	11.4
18	11.2	18.4	20	11.2	20.9	22	11.2	12.8	24	11.2	14.6
18	12.1	21.9	20	12.1	24.6	22	12.1	15.7	24	12.1	17.8
18	13.1	25.6	20	13.1	28.6	22	13.1	18.0	24	13.1	21.3
18	14.1	29.0	20	14.1	32.1	22	14.1	20.8	24	14.1	24.5
18	15.1	32.3	20	15.1	35.6	22	15.1	23.8	24	15.1	28.0
18	16.1	35.2	20	16.1	38.3	22	16.1	26.2	24	16.1	31.1
18	17.1	38.4	20	17.1	41.5	22	17.1	29.4	24	17.1	33.9
18	18.1	40.4	20	18.1	43.5	22	18.1	31.5	24	18.1	36.2
18	19.1	42.9	20	19.1	46.0	22	19.1	34.0	24	19.1	38.8
18	20.1	45.8	20	20.1	49.2	22	20.1	36.8	24	20.1	42.2
18	22.1	48.1	20	22.1	53.2	22	22.1	40.9	24	22.1	47.1
18	24.1	51.8	20	24.1	51.7	22	24.1	43.6	24	24.1	51.2
18	26.1	54.9	20	26.1	50.8	22	26.1	45.9	24	26.1	55.1
18	28.1	56.0	20	28.1	51.7	22	28.1	47.1	24	28.1	58.5
18	30.1	58.7	20	30.1	53.9	22	30.1	49.9	24	30.1	61.8

Stand	Age	Basal area	Stand	Age	Basal are
25	8.1	13.0	27	8.1	14.
25	9.1	18.1	27	9.1	19.
25	10.1	23.3	27	10.1	25.
25	11.2	28.4	27	11.2	29.
25	12.1	32.4	27	12.1	33.
25	13.1	36.8	27	13.1	37.
25	14.1	40.6	27	14.1	41.
25	15.1	44.4	27	15.1	44.
25	16.1	46.7	27	16.1	47.
25	17.1	50.3	27	17.1	50.
25	18.1	52.6	27	18.1	52.
25	19.1	55.1	27	19.1	55.
25	20.1	58.7	27	20.1	58.
25	22.1	62.8	27	22.1	61.
25	24.1	65.2	27	24.1	64.
25	26.1	68.1	27	26.1	64.
25	28.1	70.2	27	28.1	62.
25	30.1	73.6	27	30.1	62.
26	8.1	11.6	28	8.1	13.
26	9.1	16.2	28	9.1	17.
26	10.1	21.4	28	10.1	22.
26	11.2	25.8	28	11.2	27.
26	12.1	29.6	28	12.1	30.
26	13.1	33.7	28	13.1	35.
26	14.1	37.7	28	14.1	38.
26	15.1	41.0	28	15.1	41.
26	16.1	44.0	28	16.1	44.
26	17.1	46.9	28	17.1	47.
26	18.1	49.2	28	18.1	49.
26	19.1	51.5	28	19.1	51.
26	20.1	55.0	28	20.1	55.
26	22.1	59.1	28	22.1	58.
26	24.1	60.9	28	24.1	61.
26	26.1	62.6	28	26.1	64.
26	28.1	64.4	28	28.1	63.
26	30.1	67.8	28	30.1	66.

Appendix G

Curvature

G.1 The theory by example

Bates and the S-Plus 4 manual use the Michaelis-Menten expectation function. $f(x) = \frac{\theta_1 x}{\theta_2 + x}$, fitted to Puromycin data which consists of counts per minute of radioactive product (from which the "velocity" of a reaction was calculated) at different concentrations of substrate.

conc	vel
0.02	76
0.02	47
0.06	97
0.06	107
0.11	123
0.11	139
0.22	159
0.22	152
0.56	191
0.56	201
1.10	207
1.10	200

The model is fit using the S-plus nls function

```
puro.nls <-
nls(vel~th1*conc/(th2+conc),data=puro,start=list(th1=200,th2=.1),trace=T)
7964.19 :  200 0.1
1593.16 :  212.024 0.0542874
1201.03 :  211.773 0.0623245
1195.51 :  212.563 0.0639265
1195.45 :  212.672 0.0641023
1195.45 :  212.683 0.0641194
puro.nls
Residual sum of squares :  1195.449
              th1        th2
parameters:
              212.6826   0.06411945
formula:  vel ~(th1 * conc)/(th2 + conc)
12 observations
```

The velocity functions are

$$\begin{aligned}\left[\dot{V}\right]_{\theta_1} &= \frac{x}{\theta_2 + x} \\ \left[\dot{V}\right]_{\theta_2} &= f(x) = \frac{-\theta_1 x}{(\theta_2 + x)^2}\end{aligned}$$

with acceleration functions

$$\begin{aligned}\left[\ddot{V}\right]_{\theta_1\theta_1} &= 0 \\ \left[\ddot{V}\right]_{\theta_1\theta_2} &= \frac{-x}{(\theta_2 + x)^2} \\ \left[\ddot{V}\right]_{\theta_2\theta_2} &= = \frac{2\theta_1 x}{(\theta_2 + x)^3}\end{aligned}$$

These are evaluated at each x for at the parameter pair $(212.7, 0.0641)$ to give the $n \times 2$ velocity matrix $\dot{V}$ and the $n \times 3$ acceleration matrix $\ddot{V}$.

conc	$\left[\dot{V}\right]_{\theta_1}$	$\left[\dot{V}\right]_{\theta_2}$	$\left[\ddot{V}\right]_{\theta_1\theta_1}$	$\left[\ddot{V}\right]_{\theta_1\theta_2}$	$\left[\ddot{V}\right]_{\theta_2\theta_2}$
0.02	.23781213	−601.45826	0	−2.8277304	14303.407
0.02	.23781213	−601.45826	0	−2.8277304	14303.407
0.06	.48348106	−828.65771	0	−3.895899	13354.677
0.06	.48348106	−828.65771	0	−3.895899	13354.677
0.11	.63182079	−771.90283	0	−3.6290683	8867.3501
0.11	.63182079	−771.90283	0	−3.6290683	8867.3501
0.22	.77437522	−579.75927	0	−2.7257136	4081.3747
0.22	.77437522	−579.75927	0	−2.7257136	4081.3747
0.56	.8972921	−305.80681	0	−1.4377377	979.99298
0.56	.8972921	−305.80681	0	−1.4377377	979.99298
1.10	.944936	−172.65517	0	−.81173095	296.63289
1.10	.944936	−172.65517	0	−.81173095	296.63289

The acceleration vectors are resolved into tangential and normal components obtained by QR decomposition of

$$D = \begin{bmatrix} .23781213 & -601.45826 & 0 & -2.8277304 & 14303.407 \\ .23781213 & -601.45826 & 0 & -2.8277304 & 14303.407 \\ .48348106 & -828.65771 & 0 & -3.895899 & 13354.677 \\ .48348106 & -828.65771 & 0 & -3.895899 & 13354.677 \\ .63182079 & -771.90283 & 0 & -3.6290683 & 8867.3501 \\ .63182079 & -771.90283 & 0 & -3.6290683 & 8867.3501 \\ .77437522 & -579.75927 & 0 & -2.7257136 & 4081.3747 \\ .77437522 & -579.75927 & 0 & -2.7257136 & 4081.3747 \\ .8972921 & -305.80681 & 0 & -1.4377377 & 979.99298 \\ .8972921 & -305.80681 & 0 & -1.4377377 & 979.99298 \\ .944936 & -172.65517 & 0 & -.81173095 & 296.63289 \\ .944936 & -172.65517 & 0 & -.81173095 & 296.63289 \end{bmatrix}$$

which gives

$$R_1 = \begin{bmatrix} -2.4442587 & 1569.2905 & 0 & 7.3779527 & -16185.675 \\ 0 & 1320.8873 & 0 & 6.2100956 & -25030.442 \\ 0 & 0 & 0 & 0 & 8369.135 \end{bmatrix}$$

From R_1, a $3 \times 2 \times 2$ acceleration array can be formed

$$\ddot{A} = \begin{bmatrix} 0 & 7.3779527 \\ 7.3779527 & -16185.675 \end{bmatrix} \begin{bmatrix} 0 & 6.2100956 \\ 6.2100956 & -25030.442 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 8369.135 \end{bmatrix}.$$

The elements of $\ddot{A}$ give information on the parameter-effects and intrinsic non-linearity. The first two faces of the array are written as A^θ to denote parameter effects nonlinearity and the last face as A^i to denote intrinsic nonlinearity. The curvature measures are scale dependent; independence is achieved by conversion to relative curvatures

$$C_i = R_{11}^{-T} \ddot{A} R_{11}^{-1} s\sqrt{P}$$

For the Puromycin data $s = \sqrt{\frac{1195.449}{10}} = 10.93, P = 2, R_{11}^{-1} = \begin{bmatrix} -2.4442587 & 1569.2905 \\ 0 & 1320.8873 \end{bmatrix}^{-1} =$

$\begin{bmatrix} -.4091222 & .48606059 \\ 0 & .000757066 86 \end{bmatrix}$ which gives

$$\begin{aligned} C_1 &= \begin{bmatrix} -.4091222 & 0 \\ .48606059 & .000757066 86 \end{bmatrix} \cdot \begin{bmatrix} 0 & 7.3779527 \\ 7.3779527 & -16185.675 \end{bmatrix} \\ &\quad \cdot \begin{bmatrix} -.4091222 & .48606059 \\ 0 & .000757066 86 \end{bmatrix} \cdot 10.93 \cdot \sqrt{2} \\ &= \begin{bmatrix} 0 & -.035323042 \\ -.035323042 & -.059463512 \end{bmatrix} \end{aligned}$$

and similarly for the other two elements of the $\ddot{A}$ array:

$$\begin{aligned} C_2 &= \begin{bmatrix} -.4091222 & 0 \\ .48606059 & .000757066 86 \end{bmatrix} \cdot \begin{bmatrix} 0 & 6.2100956 \\ 6.2100956 & -25030.442 \end{bmatrix} \\ &\quad \cdot \begin{bmatrix} -.4091222 & .48606059 \\ 0 & .000757066 86 \end{bmatrix} \cdot 10.93 \cdot \sqrt{2} \end{aligned}$$

$$= \begin{bmatrix} 0 & -.029731752 \\ -.029731752 & -.15110827 \end{bmatrix}$$

and

$$C_3 = \begin{bmatrix} -.409122 & 0 \\ .486060\,59 & .000\,757\,066\,86 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 \\ 0 & 8369.135 \end{bmatrix} \cdot \begin{bmatrix} -.409122 & .486060\,59 \\ 0 & .000\,757\,066\,86 \end{bmatrix} \cdot 10.93 \cdot \sqrt{2}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & .074145398 \end{bmatrix}$$

C_1 and C_2 give C^θ and C_3 gives C^i. These define a curvature matrix

$$\begin{bmatrix} 0 & -.035323042 & -.059463512 \\ 0 & -.029731752 & -.15110827 \\ 0 & 0 & .074145398 \end{bmatrix}$$

where all entries are small indicating good estimation properties for the model. The curvature measures referred to in the body of this thesis are root mean square (RMS) curvatures which have been developed as a simple overall measure of nonlinearity. The curvature in a direction $\mathbf{u}$ on the tangent plane is

$$c_{\mathbf{u}} = \|\mathbf{u}^T C_n \mathbf{u}\| = \sqrt{\sum_n (\mathbf{u}^T C_n \mathbf{u})^2}$$

where the C_n are defined as above. The mean square curvature is obtained by integrating over all directions $\mathbf{u}$ and dividing by the surface area of a P-dimensional sphere

$$c^2 = \frac{1}{A} \int \sum_n (\mathbf{u}^T C_n \mathbf{u})^2 \, dA$$

which gives (Bates and Watts, 1980)

$$c^2 = \frac{1}{P(P+2)} \sum_n \left[2 \sum_{p=1}^{P} \sum_{q=1}^{P} c_{npq}^2 + \left(\sum_{p=1}^{P} c_{npp} \right)^2 \right]$$

For the Puromycin example

$$\begin{aligned} c^2 &= \frac{1}{2(2+2)} \sum_{n=1}^{3} \left[2 \sum_{p=1}^{2} \sum_{q=1}^{2} c_{npq}^2 + \left(\sum_{p=1}^{2} c_{npp} \right)^2 \right] \\ &= \frac{1}{8} \sum_{n=1}^{3} \left[2 \left(c_{n11}^2 + c_{n12}^2 + c_{n21}^2 + c_{n22}^2 \right) + \left(c_{n11} + c_{n22} \right)^2 \right] \\ &= \frac{1}{8} \sum_{n=1}^{3} \left[3c_{n11}^2 + 2c_{n12}^2 + 2c_{n21}^2 + 2c_{n11}c_{n22} + 3c_{n22}^2 \right] \\ &= \frac{1}{8} \sum_{n=1}^{3} \left[3c_{n11}^2 + 4c_{n12}^2 + 2c_{n11}c_{n22} + 3c_{n22}^2 \right] \end{aligned}$$

and with

$$\begin{aligned} C_1 &= \begin{bmatrix} 0 & -.035323042 \\ -.035323042 & -.05\,9463512 \end{bmatrix} \\ C_2 &= \begin{bmatrix} 0 & -.029731752 \\ -.029731752 & -.15110827 \end{bmatrix} \\ C_3 &= \begin{bmatrix} 0 & 0 \\ 0 & .074145398 \end{bmatrix} \end{aligned}$$

these give

$$\begin{aligned} c^i &= \sqrt{\frac{3}{8} \cdot .074145398^2} \\ &= .0454045\,98 \\ \text{whence } c^i \sqrt{F_{2,10}} &= .0454045\,98 \cdot \sqrt{4.1028} \\ &= .091968692 \end{aligned}$$

and

$$\begin{aligned} c^{\theta} &= \sqrt{\frac{1}{2}(-.035323042^2 + -.029731752^2) + \frac{3}{8}(-.059463512^2 + -.15110827^2)} \\ &= .10466353 \\ \text{whence } c^{\theta}\sqrt{F_{2,10}} &= .10466353 \cdot \sqrt{4.1028} \\ &= .21199985 \end{aligned}$$

Both intrinsic and parameter-effects curvature are within acceptable bounds according to the $c\sqrt{F} < 0.3$ criterion.

G.2 Parametric-effects curvature

Stand	*Chapman*	*Gompertz*	*Hossfeld*	*Levakovic*	*Schumacher*	*Sloboda*	*Weibull*
1	0.02	0.25	0.21	1.63	0.04	3.94	0.31
2	0.02	0.25	0.24	1.85	0.04	5.39	0.28
4	0.02	0.20	0.21	1.62	0.03	4.47	0.30
5	0.01	0.21	0.23	1.62	0.04	3.43	0.37
6	0.02	0.26	0.24	2.18	0.03	4.83	0.27
7	0.02	0.21	0.15	1.07	0.04	3.02	0.24
8	0.03	0.31	0.25	2.35	0.04	5.07	0.29
9	0.02	0.26	0.25	1.99	0.03	3.60	0.37
10	0.02	0.19	0.18	1.20	0.04	4.24	0.23
11	0.02	0.19	0.18	1.19	0.05	4.15	0.24
12	0.02	0.28	0.33	3.27	0.03	5.65	0.42
13	0.03	0.34	0.28	2.33	0.05	6.06	0.29
14	0.02	0.25	0.24	2.15	0.03	4.69	0.30
15	0.02	0.26	0.27	2.16	0.04	6.35	0.33
16	0.02	0.27	0.32	3.03	0.03	5.48	0.41
17	0.03	0.29	0.22	1.28	0.06	5.10	0.20
18	0.02	0.27	0.21	1.92	0.03	4.08	0.27
19	0.03	0.25	0.26	1.57	0.05	6.92	0.25
21	0.03	0.36	0.36	25.79	0.02	7.02	0.40
22	0.03	0.24	0.27	0.89	0.06	6.13	0.18
23	0.02	0.23	0.24	1.94	0.03	4.31	0.34
24	0.02	0.33	0.37	7.54	0.02	6.14	0.44

G.3 Intrinsic curvature

Stand	*Chapman*	*Gompertz*	*Hossfeld*	*Levakovic*	*Schumacher*	*Sloboda*	*Weibull*
1	0.01	0.02	0.02	0.05	0.03	0.02	0.04
2	0.02	0.02	0.02	0.07	0.03	0.03	0.03
4	0.01	0.01	0.01	0.05	0.02	0.02	0.03
5	0.01	0.01	0.01	0.04	0.02	0.02	0.03
6	0.02	0.02	0.02	0.07	0.02	0.03	0.03
7	0.01	0.01	0.01	0.04	0.03	0.02	0.04
8	0.02	0.02	0.02	0.08	0.02	0.03	0.04
9	0.01	0.02	0.02	0.04	0.02	0.02	0.04
10	0.01	0.01	0.01	0.05	0.03	0.02	0.03
11	0.02	0.01	0.01	0.05	0.03	0.02	0.03
12	0.01	0.02	0.02	0.05	0.02	0.03	0.03
13	0.02	0.02	0.02	0.09	0.03	0.04	0.04
14	0.01	0.02	0.02	0.06	0.02	0.03	0.03
15	0.02	0.02	0.02	0.07	0.03	0.03	0.03
16	0.01	0.02	0.02	0.05	0.02	0.03	0.03
17	0.02	0.02	0.02	0.08	0.04	0.03	0.04
18	0.02	0.02	0.02	0.06	0.02	0.03	0.04
19	0.02	0.02	0.02	0.08	0.03	0.04	0.03
21	0.02	0.02	0.02	0.06	0.01	0.04	0.03
22	0.02	0.02	0.02	0.08	0.03	0.03	0.02
23	0.01	0.01	0.01	0.05	0.02	0.02	0.03
24	0.02	0.02	0.02	0.05	0.01	0.03	0.03

Appendix H

Predictions of Basal Area

Age 28.1 Years

		Chapman		Schumacher		Hossfeld		Gompertz		Weibull		Levakovic		Sloboda	
Stand	Measured Area	Pred.	S.E.	Pred.	S.E.	Pred.	S.E.	Pred.	S.E.	Pred.	S.E.	Pred.	S.E.	Pred.	S.E.
1	55.9	55.4	0.3	57.0	0.6	55.0	0.2	55.0	0.2	53.3	0.5	55.2	0.3	55.8	0.3
2	55.8	56.1	0.4	57.6	0.6	55.9	0.4	55.6	0.4	54.4	0.6	56.2	0.5	56.7	0.5
4	50.2	49.9	0.2	51.1	0.3	49.6	0.3	49.4	0.3	48.5	0.6	50.3	0.3	50.5	0.3
5	40.8	40.8	0.2	41.7	0.3	40.4	0.2	40.3	0.2	39.4	0.4	40.8	0.2	41.0	0.2
6	65.0	65.2	0.4	66.8	0.6	65.2	0.5	64.6	0.5	63.7	0.7	65.7	0.6	66.0	0.6
7	59.7	58.8	0.3	60.4	0.6	58.5	0.3	58.3	0.3	56.8	0.6	58.9	0.4	59.3	0.4
8	62.0	62.5	0.4	64.2	0.6	62.4	0.5	62.0	0.5	60.7	0.8	62.9	0.6	63.2	0.5
9	43.9	43.5	0.2	44.6	0.3	43.1	0.3	43.1	0.3	42.0	0.6	43.8	0.3	43.9	0.2
10	63.4	64.6	0.3	66.2	0.4	64.2	0.4	63.9	0.4	62.7	0.8	65.1	0.4	65.3	0.3
11	53.2	52.9	0.4	54.3	0.6	52.6	0.3	52.5	0.3	51.3	0.5	52.8	0.5	53.2	0.4
12	53.1	52.8	0.3	54.1	0.4	52.4	0.4	52.2	0.4	51.2	0.6	53.0	0.4	53.2	0.4
13	58.9	59.2	0.5	60.8	0.7	59.0	0.5	58.7	0.5	57.3	0.7	59.2	0.7	59.7	0.7
14	57.8	58.4	0.4	59.8	0.4	58.1	0.5	57.8	0.5	56.8	0.8	59.1	0.4	59.2	0.4
15	52.8	53.0	0.4	54.4	0.6	52.6	0.4	52.5	0.3	51.2	0.5	52.6	0.5	53.3	0.5
16	42.7	42.5	0.3	43.5	0.3	42.2	0.3	42.1	0.3	41.3	0.5	42.8	0.3	42.9	0.3
17	61.0	62.5	0.5	64.2	0.6	62.4	0.5	62.0	0.5	60.8	0.9	63.2	0.6	63.3	0.6
18	56.0	56.1	0.4	57.6	0.4	56.1	0.4	55.7	0.4	54.7	0.7	56.8	0.5	57.0	0.5
19	42.0	42.9	0.4	43.9	0.4	42.6	0.4	42.5	0.4	41.6	0.5	42.9	0.5	43.3	0.5
21	60.8	60.0	0.5	61.4	0.3	60.0	0.6	59.4	0.6	58.8	0.9			61.4	0.5
22	47.1	48.8	0.4	50.0	0.6	48.7	0.4	48.4	0.3	47.4	0.3	48.2	0.5	49.0	0.5
23	51.2	50.8	0.3	52.0	0.4	50.4	0.3	50.2	0.3	49.2	0.5	50.9	0.4	51.2	0.3
24	58.5	57.7	0.4	59.0	0.3	57.4	0.5	57.0	0.5	56.4	0.8	58.6	0.4	58.7	0.4

Age 30.1 Years

		Chapman		Schumacher		Hossfeld		Gompertz		Weibull		Levakovic		Sloboda	
Stand	Measured Area	Pred.	S.E.	Pred.	S.E.	Pred.	S.E.	Pred.	S.E.	Pred.	S.E.	Pred.	S.E.	Pred.	S.E.
1	59.1	57.3	0.3	59.9	0.7	56.8	0.3	56.6	0.3	53.9	0.6	57.1	0.4	58.0	0.4
2	59.1	58.2	0.5	60.6	0.7	57.8	0.5	57.4	0.5	55.3	0.7	58.3	0.8	59.2	0.7
4	52.1	52.2	0.3	54.3	0.4	51.7	0.4	51.4	0.4	49.6	0.8	52.9	0.5	53.2	0.4
5	43.5	42.8	0.2	44.6	0.4	42.2	0.3	42.2	0.2	40.3	0.5	42.9	0.3	43.3	0.2
6	68.4	67.2	0.5	69.8	0.7	67.2	0.6	66.3	0.6	64.6	0.9	68.1	0.9	68.6	0.8
7	61.2	60.4	0.3	63.0	0.7	60.1	0.4	59.7	0.3	57.3	0.7	60.8	0.5	61.4	0.5
8	66.2	64.2	0.5	66.9	0.7	64.1	0.6	63.4	0.5	61.3	0.9	64.9	0.8	65.4	0.7
9	46.3	45.3	0.3	47.2	0.4	44.8	0.4	44.6	0.3	42.7	0.7	45.9	0.4	46.1	0.3
10	65.9	67.0	0.4	69.7	0.6	66.6	0.5	66.0	0.5	63.8	1.0	68.0	0.6	68.3	0.5
11	54.8	55.0	0.4	57.3	0.7	54.6	0.4	54.2	0.4	52.1	0.7	54.9	0.6	55.5	0.5
12	56.8	55.4	0.4	57.7	0.6	54.8	0.5	54.5	0.4	52.5	0.8	55.9	0.6	56.3	0.5
13	62.8	60.8	0.6	63.5	0.9	60.6	0.6	60.2	0.6	57.8	0.8	60.9	1.0	61.8	0.9
14	60.9	60.6	0.5	63.0	0.5	60.3	0.6	59.7	0.6	57.9	1.0	61.9	0.6	62.0	0.5
15	56.5	55.3	0.5	57.7	0.8	54.8	0.4	54.5	0.4	52.2	0.6	54.8	0.7	56.0	0.7
16	45.5	44.7	0.3	46.5	0.4	44.2	0.4	44.0	0.4	42.4	0.7	45.3	0.5	45.5	0.4
17	62.2	64.2	0.6	66.9	0.7	64.1	0.6	63.4	0.6	61.4	1.1	65.3	0.9	65.5	0.7
18	58.7	57.6	0.5	59.9	0.5	57.6	0.5	56.9	0.5	55.3	0.8	58.6	0.7	58.9	0.6
19	44.2	44.7	0.4	46.5	0.6	44.4	0.5	44.0	0.4	42.5	0.6	44.9	0.7	45.5	0.6
21	63.9	61.9	0.6	64.2	0.4	61.9	0.7	61.0	0.8	59.9	1.1			64.2	0.6
22	49.9	50.8	0.5	52.8	0.8	50.5	0.5	50.0	0.4	48.4	0.4	49.7	0.7	51.2	0.7
23	53.9	53.1	0.3	55.2	0.5	52.6	0.4	52.2	0.4	50.4	0.7	53.4	0.5	53.9	0.5
24	61.8	60.1	0.5	62.3	0.4	59.8	0.6	59.1	0.7	57.7	1.0	61.8	0.6	61.8	0.6

Appendix I

Asymptote Simulation

```
# Function and derivative needed to locate the inflection
   # point of the Sloboda function, given parameters b, c0, and d.
   #
   slob <- function(b,c0,d,t)
    {n1 <- d-1-c0*d*t^d+b*c0*d*t^d*exp(-c0*t^d)}
   der.slob <- function(b,c0,d,t)
    {n1 <- -c0*d^2*t^(d-1)+b*c0*d^2*t^(d-1)*exp(-c0*t^d)*(1-c0*t^d) }
   #
   # n is the number of simulations for each model
   # m is the slope at the point of inflection
   m <- 2
   n <- 1000
   asym <- c(NULL)
   model <- c(NULL)
   asym.b <- c(NULL)
   asym.c0 <- c(NULL)
   asym.d <- c(NULL)
   slope <- c(NULL)
   while (m < 5.5)
   {
```

```
 slope <- c(slope,rep(m,7*n))
#Chapman
temp <- rmvnorm(n, mean=c(-.123,5.42), cov=matrix(c(1,-.54,-.54,1),nrow=2,ncol=2)
sd=c(.00536,.234), d=2)
b <- temp[,1]
c0 <- temp[,2]
asym.chap <- -m/(b*((1-1/c0)^(c0-1)))
asym <- c(asym,asym.chap)
model <- c(model,rep(''Chapman'',n))
asym.b <- c(asym.b,b)
asym.c0 <- c(asym.c0,c0)
asym.d <- c(asym.d,rep(NA,n))
#Schumacher
temp <- rmvnorm(n, mean=c(54.2,1.38), cov=matrix(c(1,.95,.95,1),
nrow=2,ncol=2), sd=c(7.18,.0771), d=2)
b <- temp[,1]
c0 <- temp[,2]
asym.schu <- m*exp(1/c0+1)/(b*c0*((b*c0/(c0+1))^(-1-1/c0)))
asym <- c(asym,asym.schu)
asym.b <- c(asym.b,b)
asym.c0 <- c(asym.c0,c0)
asym.d <- c(asym.d,rep(NA,n))
model <- c(model,rep(''Schumacher'',n))
#Hossfeld
temp <- rmvnorm(n, mean=c(109,3.08), cov=matrix(c(1,.76,.76,1),
nrow=2,ncol=2), sd=c(23.1,.0552), d=2)
b <- temp[,1]
c0 <- temp[,2]
asym.hoss <- (4*b*c0*m/(c0+1)^2)^(c0/(c0-1))*(c0+1)/(b*(c0-1))
asym <- c(asym,asym.hoss)
```

```
asym.b <- c(asym.b,b)
asym.c0 <- c(asym.c0,c0)
asym.d <- c(asym.d,rep(NA,n))
model <- c(model,rep(''Hossfeld'',n))
#Weibull
temp <- rmvnorm(n, mean=c(-.000531,2.61), cov=matrix(c(1,.92,.92,1),
nrow=2,ncol=2), sd=c(.000135,.0519), d=2)
b <- temp[,1]
c0 <- temp[,2]
asym.weib <- -m*exp(1-1/c0)*((1-c0)/(b*c0))^(1/c0-1)/(b*c0)
asym <- c(asym,asym.weib)
asym.b <- c(asym.b,b)
asym.c0 <- c(asym.c0,c0)
asym.d <- c(asym.d,rep(NA,n))
model <- c(model,rep(''Weibull'',n))
#Gompertz
c0 <- rnorm(n,.144,.00632)
asym.gomp <- m*exp(1)/c0
asym <- c(asym,asym.gomp)
asym.b <- c(asym.b,rep(NA,n))
asym.c0 <- c(asym.c0,c0)
asym.d <- c(asym.d,rep(NA,n))
model <- c(model,rep(''Gompertz'',n))
#Levakovic
temp <- rmvnorm(n, mean=c(6.62,1.60,2.49),
 cov=matrix(c(1,-.57,.99,-.57,1,-.45,.99,-.45,1),nrow=3,ncol=3),
sd=c(.766,.218,.217), d=3)
b <- temp[,1]
c0 <- temp[,2]
d <- temp[,3]
```

```
asym.leva <- m*exp(b)^(1/d)*d^c0*(c0+1)^(c0+1)/
(c0*(c0*d-1)^(c0-1/d)*(d+1)^(1/d+1))
asym <- c(asym,asym.leva)
asym.b <- c(asym.b,b)
asym.c0 <- c(asym.c0,c0)
asym.d <- c(asym.d,d)
model <- c(model,rep(''Levakovic'',n))
#Sloboda
temp <- rmvnorm(n, mean=c(-4.08,76.7,-2.02),
cov=matrix(c(1,.24,-.6,.24,1,-.8,-.6,-.8,1),nrow=3,ncol=3),
sd=c(.279,15,.111), d=3)
b <- temp[,1]
c0 <- temp[,2]
d <- temp[,3]
# Find inflection point using Newton's Method
# Initial estimate at t = 10
#
t <- c(rep(10,n))
#
# stopping crireria
#
diff <- rep(10^6,n)
#
while (sum(abs(diff),na.rm=T)>0.000001)
{
diff <- slob(b,c0,d,t)/der.slob(b,c0,d,t)
t <- t - diff
}
#
asym.slob <- (m*exp(c0*t^d+b*exp(-c0*t^d))/(b*c0*d*t^(d-1)))*exp(-b)
```

```
asym <- c(asym,asym.slob)
asym.b <- c(asym.b,b)
asym.c0 <- c(asym.c0,c0)
asym.d <- c(asym.d,d)
model <- c(model,rep(''Sloboda'',n))
m <- m+0.5}
slope.inf <- as.factor(slope)
Asym.Simulate.2 <- data.frame(model,asym,slope,slope.inf,asym.b,asym.c0,asym.d)
Asym.Simulate.2$model <- reorder.factor(Asym.Simulate.2$model,Asym.Simulate.2$asy
function(x)median(x,na.rm = T))
bwplot(model[asym<250]~asym[asym<250],
 data=Asym.Simulate.2,
 main = ''Generated Asymptotes'',
 xlab = ''Asymptote'' )

bwplot(model~asym|slope.inf,
 data=Asym.Simulate.2,
 main = ''Asymptotes Generated for a Given Slope at the Inflection point'',
 xlab = ''Asymptote'')

xyplot(asym~slope|model,
 data=Asym.Simulate.2,
 main = ''Generated Asymptotes'',
 xlab = ''Slope at Inflection Point'',
 ylab = ''Asymptote'',
 layout = c(7,1,1))
```

Appendix J

Padé Rational Approximations

Quadratic and cubic Padé approximations of growth curves with inflection point $(10, 20)$ and asymptote of 90.

Growth Curve	$R_4(T)$	$R_6(T)$
Schumacher	$\frac{29.49-10.36T+.94T^2}{.74-.073T+.0099T^2}$	$\frac{-45.43+23.00T-3.96T^2+0.23T^3}{-0.92+0.26T-0.030T^2+0.0024T^3}$
Sloboda	$\frac{16.72-6.33T+.67T^2}{.83-.058T+.0074T^2}$	$\frac{-12.92+7.32T-1.47T^2+0.11T^3}{-0.24+0.12T-0.010T^2+0.0011T^3}$
Weibull	$\frac{-.55+.93T+.11T^2}{.66+.027T+.00075T^2}$	$\frac{-0.12+0.39T+0.14T^2+0.0019T^3}{0.38+0.054T+0.00063T^2+.000024T^3}$
Chapman-Richards	$\frac{-.54+.50T+.15T^2}{.60+.028T+.0011T^2}$	$\frac{-0.089+0.16T+0.16T^2+0.0020T^3}{0.41+0.047T+0.00095T^2+0.000025T^3}$
Levakovic	$\frac{-.62+.94T+.11T^2}{.75+.014T+.0011T^2}$	$\frac{-0.027+0.16T+0.12T^2+0.0060T^3}{0.23+0.061T+0.00083T^2+0.000072T^3}$
Hossfeld	$\frac{-.27+.25T+.18T^2}{.69+.013T+.0019T^2}$	$\frac{-0.032+0.058T+0.11T^2+0.0085T^3}{0.33+0.043T+0.0014T^2+0.000093T^3}$

Quartic Padé approximations of growth curves with inflection point $(10, 20)$ and asymptote of 90.

Growth Curve	$R_8(T)$
Schumacher	$\frac{-4.09+1.65T+.0043T^2.-.075T^3+.0082T^4}{.25-.050T+.011T^2-.00074T^3+.000090T^4}$
Sloboda	$\frac{.15-.39T+.25T^2-.061T^3+.0060T^4}{.36.-.054T+.011T^2-.00059T^3+.000066T^4}$
Weibull	$\frac{-.017+.10T+.092T^2+.0095T^3+.000024T^4}{.13+.052T+.0031T^2+.000037T^3+.00000077T^4}$
Chapman-Richards	$\frac{-.0052+.022T+.064T^2+.013T^3+000057T^4}{.11+.048T+.0033T^2+.000069T^3+.0000011T^4}$
Levakovic	$\frac{-.065+.27T+.16T^2+.0029T^3-.00018T^4}{.36+.070T-.00094T^2+.000064T^3-.0000023T^4}$
Hossfeld	$\frac{-.0050+.015T+.053T^2+.011T^3+.00035T^4}{.14+.043T+.0026T^2+.00013T^3+.0000039T^4}$

Quintic Padé approximations of growth curves with inflection point $(10, 20)$ and asymptote of 90.

Growth Curve	$R_{10}(T)$
Schumacher	$\frac{3.74-2.91T+.80T^2-.071T^3+.0044T^4+.00081T^5}{.019+.0082T-.0016T^2+.00067T^3-.000050T^4+.0000090T^5}$
Sloboda	$\frac{1.18-1.01T+.35T^2-.052T^3+.0020T^4+.00026T^5}{.17-.015T+.0045T^2+.000042T^3+.000020T^4+.0000029T^5}$
Weibull	$\frac{.011-.067T-.036T^2+.0095T^3+.0014T^4+.0000034T^5}{-.077-.0088T+.0065T^2+.00045T^3+.0000053T^4+.00000012T^5}$
Chapman-Richards	$\frac{-.012+.044T+.12T^2+.017T^3-.00088T^4-.0000040T^5}{.21+.076T+.0021T^2-.00013T^3-.0000033T^4-.000000084T^5}$
Levakovic	$\frac{-.0075+.048T+.059T^2+.0092T^3+.00035T^4+.0000084T^5}{.080+.040T+.0034T^2+.00014T^3+.0000042T^4+.000000092T^5}$
Hossfeld	$\frac{-.00089+.0038T+.022T^2+.0088T^3+.00076T^4+.000013T^5}{.052+.030T+.0037T^2+.00018T^3+.0000086T^4+.00000015T^5}$

Appendix K

Data for Eight Additional Pinus Radiata Stands

Little background is available for this data. They are permanent sample plots of various sizes between 0.1 and 0.4 ha and are located at various sites throughout New Zealand.

plot	age	area	plot	age	area	plot	age	area	plot	age	area
1	14.1	14.5	3	9.6	11.5	5	10.0	0.6	7	10.0	1.5
1	15.1	17.1	3	12.2	18.9	5	11.0	1.3	7	11.0	3.0
1	16.1	19.6	3	14.1	23.8	5	12.0	2.9	7	12.0	5.8
1	16.9	21.8	3	15.3	26.7	5	13.0	3.9	7	13.0	7.7
1	18.1	24.7	3	16.8	29.8	5	14.0	6.2	7	14.0	10.9
1	19.0	27.0	3	19.4	35.2	5	15.0	8.7	7	15.0	14.4
1	19.9	29.2	3	20.2	36.8	5	20.0	23.6	7	20.0	34.4
1	21.9	33.7	3	21.3	38.9	5	23.0	30.0	7	23.0	40.5
1	23.9	38.0	3	21.9	40.1	5	24.0	32.8	7	24.0	43.6
1	25.9	42.0	3	23.1	42.3	5	25.0	35.2	7	25.0	45.9
1	28.1	46.1	3	23.9	43.6	5	26.0	37.7	7	26.0	48.3
1	30.0	48.8	3	24.8	45.3	5	27.0	39.6	7	27.0	50.7
1	31.9	51.8	3	26.0	47.4	5	28.0	41.9	7	28.0	53.1
1	34.1	55.2	3	27.0	49.2	5	29.0	44.3	7	29.0	55.2
2	9.1	11.5	3	28.1	50.7	5	30.0	46.3	7	30.0	57.7
2	12.0	27.6	3	28.8	51.9	5	31.0	48.5	7	31.0	59.7
2	15.0	44.4	3	32.1	56.8	5	32.0	51.0	7	32.0	62.5
2	18.0	59.6	3	33.8	57.9	5	33.0	53.5	7	33.0	65.5
2	21.0	70.8	3	36.9	59.7	5	34.0	55.4	7	34.0	67.5
2	24.0	80.0	3	38.0	59.9	5	35.0	57.7	7	35.0	70.3
2	32.0	89.9	3	39.7	61.4	5	36.0	58.8	7	36.0	71.1
2	33.0	91.6	4	14.2	21.0	5	37.0	59.7	7	37.0	72.9
2	34.0	93.5	4	15.3	23.5	6	10.0	2.0	8	10.8	25.1
2	35.0	95.2	4	15.8	25.5	6	11.0	3.3	8	11.8	28.4
2	36.0	97.5	4	17.1	29.3	6	12.0	5.8	8	14.0	34.2
2	37.0	98.0	4	17.9	32.1	6	13.0	7.3	8	15.4	34.9
2	38.0	98.5	4	18.8	35.6	6	14.0	10.0	8	17.0	37.8
2	39.0	100.6	4	20.0	39.0	6	15.0	13.0	8	17.9	38.6
2	40.0	102.2	4	21.0	42.1	6	20.0	29.2	8	19.1	40.0
2	41.0	102.8	4	22.0	44.6	6	23.0	34.4	8	19.9	41.5
			4	22.9	46.7	6	24.0	36.9	8	20.8	43.2
			4	23.9	49.1	6	25.0	38.9	8	22.0	44.7
			4	25.9	53.8	6	26.0	41.0	8	24.1	45.1
			4	27.9	56.9	6	27.0	43.2	8	24.8	45.3
			4	29.8	59.0	6	28.0	44.9	8	27.0	46.9
			4	32.0	61.0	6	29.0	46.9	8	28.8	48.0
			4	34.0	65.3	6	30.0	48.8	8	31.0	48.5
			4	35.9	69.2	6	31.0	50.9	8	32.9	49.0
			4	37.1	70.8	6	32.0	53.2	8	36.0	49.5
						6	33.0	55.7	8	38.0	51.8
						6	34.0	57.5	8	40.1	53.0
						6	35.0	59.9			
						6	36.0	60.8			
						6	37.0	61.5			

Appendix L

Reparameterizations of the Growth Curves

Other than the first reparameterization of the Hossfeld model and the Weibull model, two parameter reparameterizations are presented using expected values at $T = t_1$ and $T = t_2$.

Three Parameter Models

$c(T)$	$\theta_0 \left(1 - e^{t_1 b}\right)^{\frac{-\ln \frac{\theta_0}{\theta_1}}{\ln \frac{1-e^{t_1 b}}{1-e^{t_2 b}}}} \left(1 - e^{bT}\right)^{\frac{\ln \frac{\theta_0}{\theta_1}}{\ln \frac{1-e^{t_1 b}}{1-e^{t_2 b}}}}$
$h(T)$	$\dfrac{aT^c}{t_1^c \left(\frac{a}{\theta} - 1\right) + T^c}$, and
	$\dfrac{\theta_0 \theta_1 \left(\left(\frac{t_2}{t_1}\right)^c - 1\right) T^c}{t_2^c (\theta_1 - \theta_0) + T^c \left(\left(\frac{t_2}{t_1}\right)^c \theta_0 - \theta_1\right)}$
$s(T)$	$\exp\left(\dfrac{\ln \theta_0 - \left(\frac{t_2}{t_1}\right)^c \ln \theta_1}{1 - \left(\frac{t_2}{t_1}\right)^c} - \dfrac{\frac{t_2^c}{1-\left(\frac{t_2}{t_1}\right)^c} (\ln \theta_0 - \ln \theta_1)}{T^c}\right)$
$w(T)$	$a\left[1 - \left(1 - \frac{\theta_0}{a}\right)^{\left(\frac{T}{t_1}\right)^c}\right]$

$L(T)$ $\quad \theta_0 \left[\frac{T^d\left(b_1+t_1^d\right)}{t_1^d\left(b_1+T^d\right)}\right]^{\frac{\ln(\theta_0/\theta_1)}{\ln\left(\frac{t_1^d\left(b_1+t_2^d\right)}{t_2^d\left(b_1+t_1^d\right)}\right)}}$, $b_1 = \ln b$

$S(T)$ $\quad a \exp\left[-\left(\frac{\ln\frac{\theta_0}{\theta_1}}{\exp\left(\frac{\ln\left(\frac{\ln\frac{\theta_0}{a}}{\ln\frac{\theta_1}{a}}\right)}{\left(\frac{t_1}{t_2}\right)^d-1}\right)-\exp\left(\frac{\left(\frac{t_1}{t_2}\right)^d \ln\left(\frac{\ln\frac{\theta_0}{a}}{\ln\frac{\theta_1}{a}}\right)}{\left(\frac{t_1}{t_2}\right)^d-1}\right)}\right) \exp\left(-\frac{\ln\left[\frac{\ln\frac{\theta_0}{a}}{\ln\frac{\theta_1}{a}}\right]}{t_2^d\left(1-\left(\frac{t_1}{t_2}\right)^d\right)}T^d\right)\right]$, and

$\theta_1 \exp\left(\frac{\ln\frac{\theta_1}{\theta_0}}{\exp\left(-c\cdot t_2^d\right)-\exp\left(-c\cdot t_1^d\right)} - c\cdot t_1^d\right) \exp\left(\frac{-\ln\frac{\theta_1}{\theta_0}}{\exp\left(-c\cdot t_2^d\right)-\exp\left(-c\cdot t_1^d\right)} \exp\left(-c\cdot T^d\right)\right)$