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.

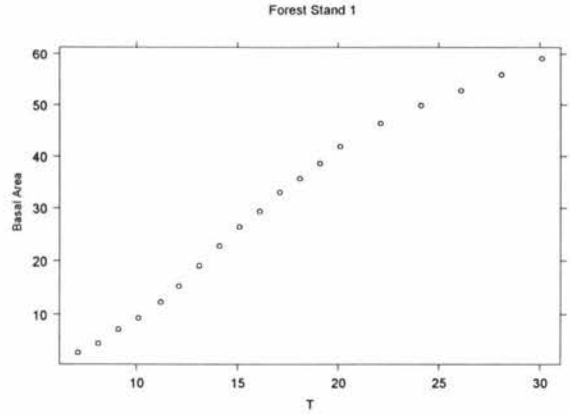

Figure 1-1: Mean basal area (m^2/hectare) against age for a pinus radiata stand from Kaingaroa Forest, New Zealand.

[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
$h(T) = \dfrac{aT^c}{ab + T^c}$	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.