RADIATA PINE YIELD MODELS

by

J.W. Leech

Thesis submitted
for the degree of
Doctor of Philosophy
Australian National University
August 1978

Statement of originality

The Generalized Least Squares technique reported in Appendix 6 of this thesis was developed by Dr. I.S. Ferguson. Apart from this, and where recognized, this thesis is my own work.

J. W. Leech

J.W. Leech

Acknowledgements

This study was undertaken under the supervision of Dr. I.S. Ferguson and Dr. L.T. Carron of the Department of Forestry, Australian National University. Their constructive criticism and support was greatly appreciated. I am especially indebted to Dr. Ferguson for his help in making the Generalized Least Squares analysis possible, and to Mr. J.A. Miles who wrote the program to carry out that analysis.

Assistance has also been received from Professors A. Brown, P.C. Young and C.R. Heathcote, and Dr. S.R. Wilson of the Australian National University, and Dr. A.M.W. Verhagen of the CSIRO Division of Mathematical Statistics.

The continuing encouragement and help of Messrs. N.B. Lewis and A. Keeves of the Woods and Forests Department of South Australia is also gratefully acknowledged.

The study was supported by a Commonwealth Post-graduate Research Award with ANU supplement, and an allowance by the South Australian Government under its tertiary assistance scheme for full time post-graduate education awards.

Permission to use data and to undertake this study was given by the Conservator of the Woods and Forests Department.

Abstract

Data from radiata pine stands in the south-east of South Australia were used to investigate various aspects of stand yield models with a view to establishing a satisfactory predictive model for use in South Australia. In the first phase of the analyses data from unthinned stands were used with Ordinary Least Squares (OLS) techniques to investigate various model structures that have been proposed in the past, to determine whether yield or increment was the better dependent variable, and to investigate conditioning through a known base point, defined as site potential, and taken as yield at age 10. The second phase extended these analyses to include investigation of the effects of thinning variations and soil differences, and also investigated the use of the model both for other forest regions in South Australia and for second rotation stands. Because these analyses were statistically unsatisfactory Generalized Least Squares (GLS) and Bayesian statistical methods were used in the third phase to develop a simple yield prediction model that is statistically sound. This technique offers considerable promise for future work.

The conditioned form of the Mitscherlich or monomolecular model below was the most satisfactory yield prediction model developed for radiata pine stands in South Australia.

$$Y_A = Y_{10}\left\{\frac{1 - \exp(-p(A - 10\exp(-a_1)))}{1 - \exp(-p(10 - 10\exp(-a_1)))}\right\}$$

where

$$p = 0.05271 - 0.006484\ \ln(Y_{10})$$

$$a_1 = -0.003467\ Y_{10}$$

and where

Y_A = yield at age A,

A = age, and,

Y_{10} = site potential.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

I INTRODUCTION

I INTRODUCTION

South Australia has little native forest and of necessity the Government became interested in plantation forestry over a hundred years ago, only 40 years after the state was first settled.

By 1920 the Woods and Forests Department realized that radiata pine, Pinus radiata (D.Don.), had the greatest growth potential of the species tried, and had developed a satisfactory silvicultural system for the management of the species. During the subsequent economic depression plantation establishment increased dramatically. The Department now controls some 76 700 ha of plantation of which 68 900 ha have been planted with radiata pine. This resource is managed on an approximately 50 year rotation and has a comparatively even distribution of age classes as can be seen in Figure I.1.

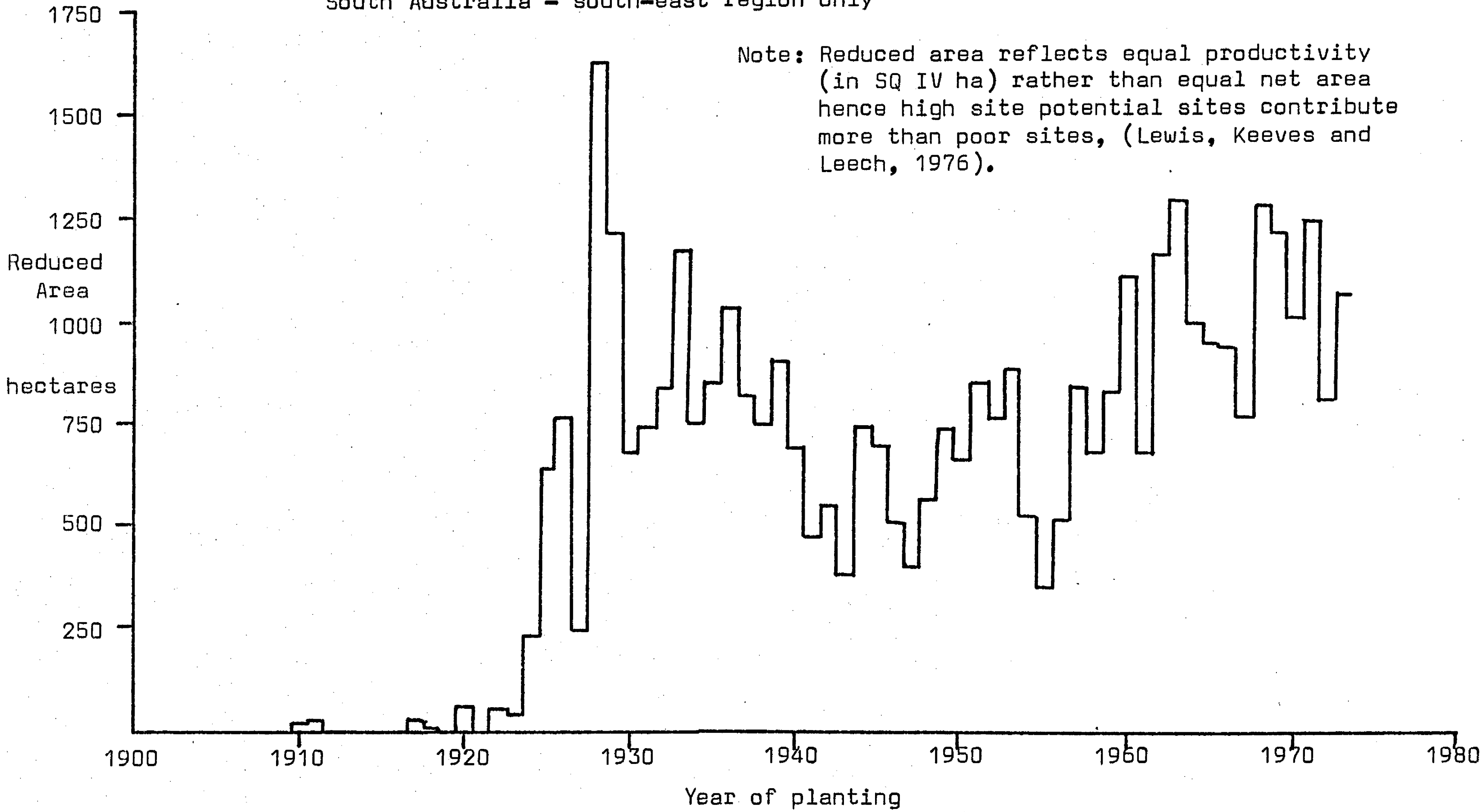
Prediction of future yield in South Australia is especially critical. Current prediction techniques (Lewis, Keeves and Leech, 1976) indicate that the increment of the resource is approximately equal to the commitment to existing industry. The potential for expanding the area of plantations is very limited because of high land prices and the limited area of suitable soils. Moreover Keeves (1966) has shown that the second rotation on any site has, and will have, a lower yield than the first, so that future industry expansion is limited.

The objective of this study was to develop a yield prediction model for the radiata pine plantations in the lower south-east of South Australia so that the Woods and Forests Department can continue to efficiently manage its plantation resource.

However, the study has wider implications. In 1975 Australia had some 565 000 ha of coniferous plantations, amounting to 1.3% of the forest area. Of this area some 394 000 ha or 70% were planted with radiata pine (Australia, Department of Agriculture, 1976). The proportion of

Figure I.1

Pinus radiata age class distribution - Woods and Forests Department of South Australia - south-east region only



unpublished data from Woods and Forests

Australia's wood production obtained from coniferous plantations has increased from 6% in 1950/51 to 18% in 1970/71 (Wilson, 1974) and the FORWOOD Conference (Australian Forestry Council, 1974) predicted that the proportion will be some 57% by the year 2010. On the basis of the existing plantation resource alone radiata pine will become the major commercial forest type in Australia within a relatively few years.

The Mensuration and Management Research Working Group of the Standing Committee of the Australian Forestry Council discussed growth models for radiata pine at a meeting at Caloundra, Queensland, in 1974. The Group acknowledged that differences existed between regions within Australia in the growth of radiata pine, but the Group pointed out that a detailed analysis of the differences in growth and form between regions would lead to a better understanding of the species, and hence lead to better prediction models. The Group concluded that development of a generalized growth model which recognised such differences was both possible and highly desirable.

Because South Australia has a long history of plantation forestry and has probably the best radiata pine growth data available, this study could well provide the basis for managers of radiata pine in other areas to review their long term planning models and provide a basis for a generalized model. Indeed a secondary objective of this study was to investigate the utility of the south-east model in relation to other radiata pine areas in South Australia as a precursor to an investigation of the further transportability of the model.

II THE DATA

Management practice in South Australia

Permanent Sample Plots

VARIABLES

Yield

Age

Site potential

Stand density

Thinning

Soil

Form

Other sources of variation

THE DATA BASE

II THE DATA

South Australia is the driest state in Australia with only 1.2% of the area receiving more than 600 mm of rainfall per year (Bednall, 1957). Intensive forestry is necessarily limited to these higher rainfall areas, the largest area of which is in the south-east of the state. The main plantation resource in the lower south-east has been described by Bednall (1957) and Douglas (1974), and consists of some 100 000 ha of softwood plantations located in a compact unit as shown in Figure II.1. Some 61 000 ha are controlled by the Woods and Forests Department (Woods and Forests, 1976), of which 55 400 ha have been planted with radiata pine.

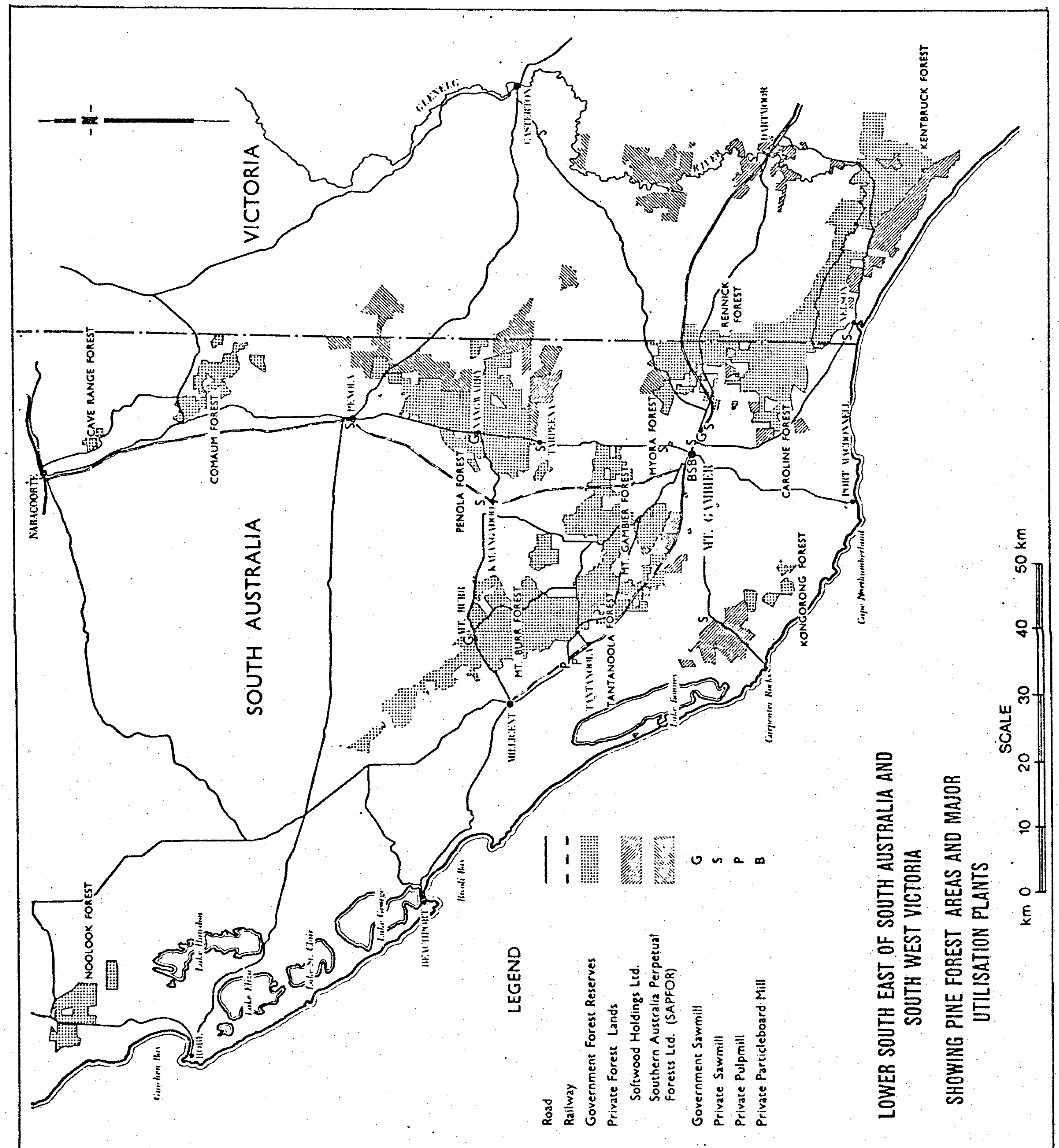
Elsewhere in the state the Woods and Forests Department has some 13 400 ha of radiata pine plantations, predominantly in five forest reserves geographically separated from one another: Bundaleer and Wirrabara Forest Reserves in the Northern region, and Mount Crawford, Kuitpo and Second Valley Forest Reserves in the Adelaide Hills or Central region (Figure II.2). Data from these areas were used to evaluate whether the model developed using data from the lower south-east of the state could be extended to other areas.

Management practice in South Australia

Current management practice in South Australia has recently been described in detail by Lewis, Keeves and Leech (1976). However, some features of current practice need to be reiterated here.

In South Australia radiata pine plantations are stratified into volume productivity classes which are termed site quality classes, volume being considered a more effective basis for stratification than upper stand height (Keeves, 1970). Site quality assessment is based on total volume production to 10 cm top diameter underbark at age $9\frac{1}{2}$ years.

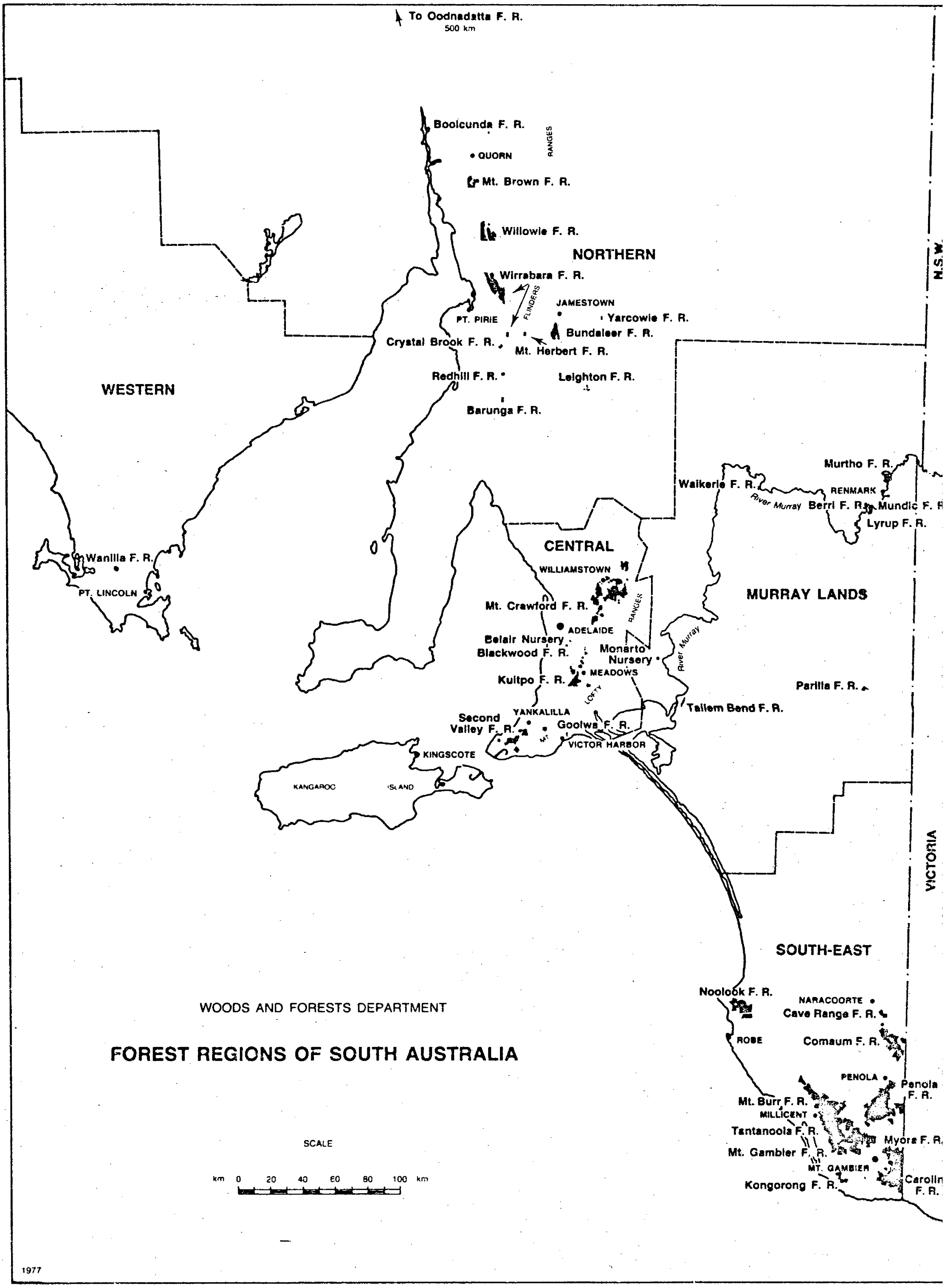
Figure II.1



LOWER SOUTH EAST OF SOUTH AUSTRALIA AND SOUTH WEST VICTORIA

SHOWING PINE FOREST AREAS AND MAJOR UTILISATION PLANTS

Figure II.2


To Oodnadatta F. R.
500 km
Boolcunda F. R.
QUORN
RANGES
Mt. Brown F. R.
Willowie F. R.
NORTHERN
Wirrabara F. R.
FLINDERS
JAMESTOWN
PT. PIRIE
Yarcowie F. R.
Bundaleer F. R.
Crystal Brook F. R.
Mt. Herbert F. R.
Redhill F. R.
Leighton F. R.
WESTERN
Barunga F. R.
N.S.W.
Murtho F. R.
Walkerie F. R.
RENMARK
River Murray
Berri F. R.
Mundic F. R
Lyrup F. R.
CENTRAL
Wanilla F. R.
WILLIAMSTOWN
PT. LINCOLN
MURRAY LANDS
Mt. Crawford F. R.
RANGES
ADELAIDE
Belair Nursery
Blackwood F. R.
Monarto Nursery
River Murray
MEADOWS
Kuitpo F. R.
LOFTY
Parilla F. R.
YANKALILLA
Tailem Bend F. R.
Second Valley F. R.
Goolwa F. R.
MT
VICTOR HARBOR
KINGSCOTE
KANGAROO
ISLAND
VICTORIA
SOUTH-EAST
Noolook F. R.
NARACOORTE
Cave Range F. R.
WOODS AND FORESTS DEPARTMENT
ROBE
Comaum F. R.
FOREST REGIONS OF SOUTH AUSTRALIA
PENOLA
Penola F. R.
Mt. Burr F. R.
MILLICENT
Tantanoola F. R.
Myora F. R.
SCALE
Mt. Gambier F. R.
MT. GAMBIER
km 0 20 40 60 80 100 km
Carolin F. R.
Kongorong F. R.
1977

Inventory is carried out on a five yearly cycle using temporary 0.1 ha plots. The intensity of sampling is such that the average logging unit of some 30 ha has five plots selected at random within site quality strata. In each plot diameters are measured and the trees to be removed during the next five year period are demarcated. Volumes available from thinning are estimated using a tree volume equation, with appropriate adjustment for increment on the thinnings between time of inventory and the scheduled year of thinning (Leech, 1973).

A short term (five year) cutting plan is then produced, delineating where thinning and clear felling should be carried out. The inventory data and the cutting plan are also used to predict yield from the resource some 60 years into the future, using a deterministic simulation model developed by the author. The model developed in this study is intended to replace the yield prediction model currently incorporated in that long term planning model.

Permanent Sample Plots

Following the first forest inventory in South Australia permanent sample plots were established in 1935 and these have been gradually augmented so that there are at present 313 plots in radiata pine plantations in the south-east of the state. These plots have been remeasured at various intervals and provided the data base for this study.

The plots generally cover the range of past silvicultural practice and, although they do not cover recent changes in establishment practice and early maintenance, they are typical of the major part of the forest estate that will contribute to the cut for the next 20 years or so. Models developed from these data can therefore be used in long term planning.

Plots have generally been measured more frequently for basal area and upper stand height than for volume, measurement frequency decreasing with increasing age, but have always been measured for volume at time of thinning. The thinning regime for each plot was prescribed at plot establishment, although some plots have been rescheduled to widen the range of treatments.

Mensuration practice has remained more or less constant since plots were first established and is described in detail elsewhere (Lewis, Keeves and Leech, 1976). However, two aspects of measurement have changed over time:

1 sampling for volume, and,

2 height estimation.

These were considered further to see whether the changes had any serious implications for this study.

Plot volumes have always been estimated from the volume of sample trees, individual trees being estimated by the 3 m or 10 foot sectional method (Jerram, 1939) or the Regional Volume Table (Lewis and McIntyre, 1963; Lewis, McIntyre and Leech, 1973). Thus any variations in form have been taken into account in the estimated plot volumes. Sampling intensity and the method of selection of sample trees have changed gradually over the years. Initially arithmetic mean tree methods (Jerram, 1939) were used, changing later to the use of Jolly's (1950) volume-basal area line and evolving to the current stratified random sampling frame (based on the volume basal area line) developed by Keeves (1961). Nevertheless, analysis of some 32 plots where all trees were measured by the 10 foot sectional method (the data Keeves used) indicated that although the early volume estimates were somewhat less precise than recent methods, they were unbiased (Keeves, pers.comm.).

Initially mean dominant height was estimated. The current definition of 'predominant height' (Lewis, Keeves and Leech, 1976) has only been in use since 1962. This change has more serious consequences than those for volume because the differences between the two measures may be substantial. Estimates of predominant height were available for some plots prior to 1962 but their precision and bias were unknown. The estimates were considered satisfactory for the determination of form estimates but were considered unsatisfactory for the development of height prediction models.

VARIABLES

As in most studies of this kind spanning long periods of time, the data available dictate the variables which can be used in the model.

Yield

The objective of a yield model is to predict the utilisable volume of wood that can be taken from a site. The utilisable volume depends on the volume available and the volume lost or wasted in logging, and this loss or wastage varies considerably depending on the equipment used. However, this study is restricted to estimating the volume available. It is envisaged that separate studies will be carried out from time to time to determine or revise the volume lost or wasted in logging.

The volume available has been defined as including both standing volume and the volume lost due to mortality and thinning, measured underbark in cubic metres per hectare to a 10 cm top diameter.

Age

Age is generally considered to be the most important independent variable in growth and yield studies (Buckman, 1962) and was the only independent variable in many of the earlier models. In South Australia plantations are established in winter using one year old seedlings, however the age of the plantation is taken as the number of years since planting out, ignoring the period in the nursery. All permanent sample plot data were measured in the period between late May and early September, with the measurement program starting in the same locality each year and progressing in the same sequence so that measurements in any plot were generally made in the same month each time. The seasonal fluctuation in growth within each year (Pawsey, 1964) can therefore be ignored.

Site potential

Site quality assessment is carried out in South Australia in the summer when the stand reaches age $9\frac{1}{2}$ years, however, as the plots were all measured in winter the base age for this study was assumed to be 10 years. In this study the total volume yield underbark at age 10 years (Y_{10}) in cubic metres per hectare to a 10 cm top diameter was used as the definition of site potential. The relationship between site potential (Y_{10}) and site quality from the age $9\frac{1}{2}$ assessment is shown in Table II.1.

For plots where there were no measurements at age 10, Y_{10} was estimated by linear interpolation, or if this was not possible, by extrapolation. The extrapolation was based on the average of the first two volume increments available and was confirmed by comparing estimated Y_{10} with extrapolation using Lewis's yield table (Lewis, Keeves and Leech, 1976), and by inspection of the basal area-age trend; many of the plots having been measured for basal area before volume measurement

Table II.1

Relationship between

site quality and site potential

Site quality	Site potential (Y_{10}) m^3/ha to 10cm top diameter underbark
SQ I	273
SQ II	223
SQ III	175
SQ IV	131
SQ V	80
SQ VI	37
SQ VII	7

commenced.

Stand density

Indices of stand density inevitably raise issues concerning definition and measurement (Leech, 1973), however in this study the choice of a variable to use as an index of stand density was restricted. There was no point in using indices based on such variables as standing basal area or upper stand height because such variables change continuously with age and require a separate model to be developed for prediction of future density. Only two of the variables available seemed appropriate, stocking and standing volume.

Stocking, the number of trees standing per hectare, has the advantage of being readily measured in the field, but is not entirely adequate as a index of density (Leech, 1973). Although not so easy to measure, standing volume is measured on all permanent sample plots and can readily be estimated for inventory plots. Standing volume seemed likely to provide a better index than number of trees so both these variables were tried in subsequent analyses.

Thinning

The description of a thinning regime can be separated into three parts (Lewis, 1959; Ford-Robertson, 1971).

1 Thinning type; indicating the categories of trees to be removed in the thinning based on size or crown classification.

2 Thinning grade; indicating the quantity to be removed, expressed in terms of number of trees, basal area or volume.

3 Thinning interval; indicating at what stages in the development of the stand these removals are to be made. This is generally expressed in years although it could be expressed in terms of volume or basal area growth since the last thinning.

The thinning type practised in South Australia has for many years been essentially a thinning from below with all suppressed and sub-dominant trees being removed as well as a proportion of the co-dominants and dominants to help space the trees. Indices of thinning type are generally ratios of either mean tree diameter (Lewis, 1959; Braathe, 1957; Joergensen, 1957) or volume (Lewis, Keeves and Leech, 1976), of the trees removed from the stand to those either before or after thinning. Within the available data the range of thinning type was quite narrow (if thinning type was defined as the ratio of the mean diameter of thinnings to the mean diameter before thinning, the mean thinning type was 0.92 with a standard deviation of 0.04), and thus thinning type was not included as a variable in the subsequent analyses.

The grade of a thinning is a measure of the change in competition level due to that thinning. Grade is commonly specified in terms of either residual basal area (Gentle _et al._, 1962; Robinson, 1968) or residual number of trees (Lewis, 1959, 1963) and in this absolute form is in essence an index of stand density which has already been discussed. Buckman (1962) considered the more logical measure to be the proportion of the forest cut either as volume, basal area or number of trees per unit area. This relative form provides a measure of the shock that the stand has suffered in thinning and should obviously be based on the same variable as stand density.

Thinning interval is defined as the number of years between thinnings. Normal South Australian practice is to thin SQ I and II stands every five years, SQ III every six, SQ IV and V every seven, and SQ VI and VII every eight to ten years depending on the health of the stand. Logging practice generally keeps within one or two years of this ideal. Permanent sample plots include a somewhat wider range of thinning interval than normal plantation practice.

Soil

The soils of the south-east region of South Australia have been described and typed by Stephens (Stephens et al., 1941) who conducted a detailed soil survey of much of the forest area. Three soil profiles have been described on all permanent sample plots and have been allocated to these soil types with three depth phases superimposed. For other regions the soil profiles were allocated to a soil type on the basis of a number of different surveys, but the soil types generally reflect morphological differences on a broader scale than the south-east survey.

Form

When considering the possible effect of form on increment or yield the differences between form factor and taper need to be considered; both being related to different aspects of the concept of form. A number of alternative stand based indices of form were available based on standing basal area (m^2/ha) and a measure of upper stand height, predominant height (m). These indices are crude proxies for the more commonly used tree based indices, but were the best available.

1 Stand form factor, the ratio of standing volume to the product of basal area and predominant height.

2 Stand form factor at age 10, possibly an indicator of differences between soil types or regions because it is unaffected by thinning and is age invariant.

3 Relative stand form factor, the ratio of current stand form factor to stand form factor at age 10.

4 Average stand taper, the ratio of mean tree diameter to predominant height, mean tree diameter being the diameter equivalent to the mean basal area. This was considered more likely to vary between soil types than stand form factor as on some heavier soils higher than average basal area is accompanied by lower than average predominant height.

5 Average stand taper at age 10, possibly a better index than average taper because, as with form factor at age 10, it is age invariant and independent of thinning.

6 Relative stand taper, the ratio of average stand taper to average stand taper at age 10.

Where data at age 10 years were not available, interpolated or extrapolated figures were used.

Other sources of variation

In discussing tree growth in relation to the environment Gaertner (1964) summarised the literature on the effect of nutrition, moisture, temperature and various aspects of light. He and Glock (1955) cited considerable evidence of correlation between rainfall and growth, work supported by Fielding and Millett (1941) for radiata pine. There are few meteorological stations in Australia with records of temperature and hours of sunlight, and even these seldom cover the temporal range of the forest growth data available for this study. Investigation of long term rainfall records revealed many anomalies and discontinuities which made it impossible to develop a useful rainfall index for this study. Moreover, the stations are too sparse to enable rainfall to be estimated for each plot. In view of this lack of suitable data, climatological variables were not included in this study.

Variables reflecting man-made influences such as nutrition and tree breeding were not available in a form suitable for inclusion in these analyses. This is clearly an area which warrants urgent attention in the future, since future yields will be affected by these influences. However, their omission was not critical to this study as the bulk of the present plantation estate was not established from seed orchard stock and has not received intensive treatment with fertilizers.

THE DATA BASE

The data were extracted from manually maintained permanent sample plot registers and files and were coded for punching onto cards. After the data were punched and verified a program developed by the author was used to check the data as rigorously as possible, finally producing appropriately formatted data files and a facsimile of the register. Painstaking reconciliation of this register with the manually maintained register, resolving any remaining sources of difference or ambiguity, ensured that the data were as error free as possible.

The data were largely measured in imperial units, the conversion to metric being made in 1973. The data base thus included metric measurements and metric conversion of imperial measurements, but careful checking reduced potential errors from this source to a minimum.

The data are summarised in a number of tables in Appendix 1. Appendix 1.1 summarises the complete data base, Appendix 1.2 the soil types, and Appendices 1.3 to 1.6 the different data sets used during the analyses.

III STATISTICAL METHODS

III STATISTICAL METHODS

Introduction

The estimation of the relationship between one variable and a number of others is a common problem in forestry and is generally accomplished by the use of multiple linear regression analysis using the 'Ordinary Least Squares' (OLS) technique. Multiple regression requires a model with a linear structure, the linearity referring to the coefficients or parameters of the independent variables.

$$y_i = \sum_{j=1}^{j=k} b_j x_{ij} + e_i \qquad \text{(III.1)}$$

where

y_i is the i'th observation of the dependent variable, $(i=1.....n)$,

x_{ij} is the i'th observation of the j'th independent variable, $(i=1.....n,\ j=1.....k)$,

b_j is the j'th parameter to be estimated, $(j=1.....k)$,

e_i is the error term for the i'th observation,

n is the number of observations, and,

k is the number of parameters.

The linear model can be stated in matrix form as

$$Y = XB + E \qquad \text{(III.2)}$$

where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots x_{1k} \\ x_{21} & x_{22} & \cdots x_{2k} \\ \vdots & \vdots & \cdots \vdots \\ x_{n1} & x_{n2} & \cdots x_{nk} \end{bmatrix}$$

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}$$

$$E = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

OLS is widely used to estimate the parameters of linear models and has been developed and discussed in detail by many authors including Kendall and Stuart (1961), Johnston (1963), Goldberger (1964), Draper and Smith (1966) and Theil (1971). If the analysis of a model violates any of the assumptions that underly OLS analysis, other techniques may be more appropriate. These techniques include two stage least squares (2SLS), generalized least squares (GLS), the use of instrumental variables, lagged variables, weighting and dummy variables. However, OLS generally represents the best starting point for any analysis, providing initial results which can be used to test whether the model conforms with the underlying assumptions.

Linearity is often too restrictive a requirement and models with a nonlinear structure may need to be examined. The general form of a nonlinear model can be represented thus:

$$Y = f(B, X) + E \tag{III.3}$$

where the f operator is used to denote a function nonlinear in the parameters B, and where the notation is as for Equation III.2.

Sometimes a nonlinear model can be transformed (for example by taking logarithms) to obtain a form which is linear in the parameters. These intrinsically linear models, to use Draper and Smith's (1966) terminology, can be estimated in the transformed state using OLS.

In general, however, OLS cannot be used to estimate the parameters of nonlinear models. The normal equations which result from differentiating the objective function are not linear in the unknown parameters, and no exact analytical solution for these equations exists. An iterative approximate solution must be employed. Even so, there is no single algorithm which will unfailingly yield satisfactory estimates of the parameters of nonlinear models.

Nonlinear models which are not intrinsically linear constituted a major interest in this study. The statistical theory relating to parameter estimation for these models is not well developed. However, the results of the linear model theory often seem applicable to them, at least to an acceptable order of approximation (Goldfeld and Quandt, 1972; Box and Tiao, 1973), and hence it seemed appropriate to frame this review around the linear theory results, which are well established and coherent.

PROPERTIES OF ESTIMATORS AND PREDICTORS

The following properties (Kendall and Stuart, 1961; Graybill, 1961; Wonnacott and Wonnacott, 1970) are generally sought for linear estimators:

1 An estimator should be unbiased. An unbiased estimator is one that on average has the same value as the true estimator. For example $\hat{b}_j$ is an unbiased estimator of b_j if

$$E(\hat{b}_j) = b_j \tag{III.4}$$

where

E is the expected value of the parameter.

Bias is defined as

$$B_{b_j} \equiv E(\hat{b}_j) - b_j \tag{III.5}$$

where

B_{b_j} is the bias of the parameter b_j.

2 An estimator should be efficient. When comparing two alternative estimates of b_j, $\hat{b}_j$ and $\hat{\hat{b}}_j$ then the most efficient estimator is the one with the lower variance. The ratio of the variances provides a measure of the relative efficiency if the estimators are unbiased. The relative efficiency of $\hat{\hat{b}}_j$ compared with $\hat{b}_j$ is:

$$R = \frac{\sigma^2_{\hat{b}_j}}{\sigma^2_{\hat{\hat{b}}_j}} \tag{III.6}$$

where

R is the relative efficiency, and,

$\sigma^2_{\hat{b}_j}$ is the variance of the estimator $\hat{b}_j$, $\sigma^2_{\hat{\hat{b}}_j} > \sigma^2_{\hat{b}_j}$.

The most efficient estimator is the minimum variance estimator as by definition there can be no estimator with a greater relative efficiency.

3 An estimator should be consistent. Estimators are said to be consistent if

$$E(\hat{b}_j - b_j)^2 \longrightarrow 0 \tag{III.7}$$

and as

$$E(\hat{b}_j - b_j)^2 = B^2_{\hat{b}_j} + \sigma^2_{b_j} \tag{III.8}$$

an estimator b_j is consistent if

$$\lim_{n \longrightarrow \infty} B_{\hat{b}_j} = 0 \tag{III.9}$$

$$\lim_{n \longrightarrow \infty} \sigma^2_{b_j} = 0 \tag{III.10}$$

If only the bias approaches zero then the estimators are said to be asymptotically unbiased, but not strictly consistent.

4 An estimator should be sufficient. An estimator is said to be sufficient if it contains all the information in the set of observations regarding the parameter to be estimated (Fisher, 1921, 1925; Deutsch, 1965).

OLS estimators possess these properties provided that the data and model conform with the assumptions underlying 'classical normal linear regression', to use Goldberger's (1964) terminology. Under these conditions OLS estimators are also identical to the maximum likelihood estimators. OLS estimators also provide predictors which are best (i.e. minimum variance) linear unbiased predictors (Theil, 1971).

For nonlinear models with independent, normal and identically distributed errors, OLS estimators are likewise identical to the maximum likelihood estimators and are therefore asymptotically efficient, consistent and sufficient estimators (Goldfeld and Quandt, 1972). However, unlike linear models, the small-sample properties of these estimators are not well established. Moreover, in contrast to the exact solutions available for linear models, these properties are further complicated by errors of estimation which may be introduced through the iterative approximate process of solution. Thus the properties of predictors based on small-sample OLS estimators of nonlinear models are not well defined.

MISSPECIFICATION

The assumptions underlying the use of OLS estimation for classical normal linear regression models are as follows (Goldberger, 1964):

1 The variance should be homogeneous over the range of the dependent variable.

2 The error terms should be independent of one another.

3 The error terms should be normally distributed.

4 The rank of the matrix of observations should be equal to the number of parameters to be estimated and less than the number of observations.

5 The variables should be measured without error.

6 The model should have the correct structure and include all the relevant variables, but no others.

If OLS estimation methods are used for a model that is misspecified in terms of these assumptions then the estimates may be biased, inefficient and/or inconsistent depending on the form of the misspecification, as the following sections indicate.

Homogeneity of variance

The variance of the error term is assumed to be independent of the independent variables and homogeneous (Kendall and Stuart, 1961; Johnston, 1963; Wonnacott and Wonnacott, 1970).

$$E(EE') = \sigma^2 I \qquad \text{(III.11)}$$

where

I is the identity matrix of order $n \times n$,

E_i is the error vector of the i^{th} observations, and,

σ^2 is the variance.

If this assumption is violated then the estimators are unbiased but inefficient (Johnston, 1963).

To test the variance for homogeneity the data are generally partitioned and the variance of each cell calculated. There is general agreement (Sokal and Rohlf, 1969; Acton, 1959) that Bartlett's test of homogeneity (1937) is better than either Hartley's (1950) or Cochran's (1941) test although Acton considers that none of these tests is robust, all being sensitive to non-normality in the underlying distribution.

Heterogeneity seemed most likely to arise between different plots, and within any one plot the variance might also increase with increasing age or with decreasing site potential. As insufficient data were available to test for heterogeneity by age within plots, the first test was by plots alone. The data were then pooled and partitioned into age and site potential cells and the cell variances tested for heterogeneity. As a third test the data were ordered on the expected value of the dependent variable and divided into approximately equal cells in a general omnibus test for all other possible sources of heterogeneity.

To overcome heterogeneity, observations are generally weighted by the reciprocal of the square root of the estimated variance (Cunia, 1964; Freese, 1964; Johnston, 1963). If heterogeneity exists and a suitable estimating function for variance cannot be developed then extension to more advanced estimation techniques than OLS may be necessary, but the gain in efficiency must be balanced against the increase in the complexity of the estimation technique.

Serial correlation

The error term in the linear model is assumed to be unbiased, that is

$$E(e_i) = 0 \tag{III.12}$$

and as well the error terms are assumed to be independent of one another (Kendall and Stuart, 1961; Graybill, 1961).

$$E(e_i e_{i'}) = 0 \quad \text{for all } i \neq i' \tag{III.13}$$

If the latter assumption is violated then the OLS estimators will be inefficient (Johnston, 1963).

Serial correlation most commonly occurs in time series due to mis-specification either by omitting variables (Wonnacott and Wonnacott, 1970) or by selecting the wrong model structure (Cochrane and Orcutt, 1949). Errors in the data are another possible source of serial correlation (Cochrane and Orcutt, 1949) but these seemed unlikely to be of importance in this study. As noted earlier the data for this study were all collected at the same time of the year thus eliminating one source of seasonally induced serial correlation.

The most commonly used test for serial correlation is that of Durbin and Watson (1950, 1951).

The statistic is

$$d = \frac{\sum_{i=2}^{i=n} (e_i - e_{i-1})^2}{\sum_{i=1}^{i=n} (e_i)^2} \tag{III.14}$$

where

n is the number of observations,

e_i is the error of the i'th observations, and,

d is the test statistic.

This test in its original form is not an exact test but provides upper and lower bounds to an inconclusive zone for from 15 to 100 observations. Because it is not an exact test a number of authors have developed nominally exact tests of which the one by Theil and Nagar (1961) is probably the most commonly used. In their latest paper Durbin and Watson (1971) concluded that many of these exact tests are "too inaccurate for practical use", and further refined an approximation they had suggested in their earlier papers. However, this also seemed inappropriate for this study.

In this study the number of observations from each plot was limited to 16 or fewer, generally 10–13. Although serial correlation was tested by plots where possible, the results were seldom conclusive because of the low number of observations. The test was also carried out on the pooled data, although there is no adequate test for serial correlation in these circumstances (Heathcote, pers.comm.). The plots in the test data were ordered by site potential, and the observations within each plot ordered by age. The d statistic calculated on this pooled data is biased slightly by the inclusion of the difference between the last observation of one plot and the first of the next and this difference is unlikely to be serially correlated. There were too many observations in the pooled data for the tabulated upper and lower bounds to be used, and the extra calculation necessary for the Durbin Watson approximation of the 'exact' statistic seemed inappropriate in view of the inadequacy of the test for the pooled data. The critical values of the statistic were therefore calculated by the technique of Theil and Nagar (1961), in the absence of better alternatives. For the individual plot data the tabulated Durbin and Watson statistics were extrapolated where necessary.

Normality

When making inferences about model structure using statistical tests of hypotheses the error term is assumed to be normally distributed. Normality could be tested using the Chi-square statistic, but this test is not specific in that it fails to indicate whether skewness or kurtosis is the problem. To overcome this problem Sokal and Rohlf (1969) and Snedecor and Cochran (1967) describe techniques for estimating moment statistics of skewness and kurtosis. These two statistics are then compared with t (two tailed) for infinite degrees of freedom. The Shapiro-Wilk statistic (Shapiro and Wilk, 1965) is more powerful (Shapiro, Wilk and Chen, 1968) than these other statistics but the tests of relative power indicated little gain when the number of observations increased past 50.

Tests of normality could not be carried out by plots because there were too few observations even for the Shapiro-Wilk test. For the pooled data there were more than 100 observations and in these cases the Shapiro-Wilk test is difficult to apply and perhaps even dangerous (D'Agostino, 1971). The moment statistics were therefore selected as the test statistics because they indicated the type of departure from normality and were relatively powerful for the sample sizes used in this study.

Rank

The rank of the matrix of observations must be equal to the number of parameters to be estimated, that is, no exact linear relationship can exist between any of the independent variables. The rank must also be less than the number of observations.

In general, these assumptions are easily met by careful definition and selection of the set of independent variables. However, marked collinearity between any two independent variables gives rise to parameter estimates with very high sampling errors. Thus this condition also needs to be avoided.

Measurement error

OLS assumes that the variables are measured without error (Kendall and Stuart, 1961; Wonnacott and Wonnacott, 1970). If the dependent variable is measured with error, but the error is unbiased, then the variance of the error term for the model is inflated accordingly. This problem is therefore of comparatively little concern, although the increase in the error variance may obscure relationships between the dependent and independent variables.

Errors of measurement in the independent variables may give rise to more serious problems unless;

1 the errors are unbiased, and,

2 the data to be used in subsequent prediction are measured in the same way as those used to estimate the parameters.

Under these circumstances OLS estimators will give unbiased predictions (Wonnacott and Wonnacott, 1970), even though the estimators are biased relative to those appropriate to the independent variables when measured without error.

The measurement practice used in the development of the data base has been described by Lewis, Keeves and Leech (1976) and in part by Leech (1973). Although measurement errors exist, every possible effort has been made to reduce the incidence of these to a minimum by care and by strict adherence to standard procedures. The effect of measurement error was unlikely to be important and was ignored.

Structure

There are three main forms of structural misspecification:

1 choosing the wrong model structure, for example by using the logarithm of the dependent variable where that transformation is inappropriate to the model,

2 omission of an explanatory independent variable, and,

3 inclusion of independent variables that are irrelevant.

The first type of structural misspecification can obviously lead to an inefficient prediction model even though the estimators themselves are efficient.

If relevant independent variables are omitted, either by mistake or because data are not available, then the estimates of the remaining parameters are likely to be biased and inefficient (Wonnacott and Wonnacott, 1970).

If irrelevant independent variables are included then the estimators should be unbiased, but they will be inefficient because there are fewer degrees of freedom in the residuals used to estimate the variance. Because of the 'noisy' parameters the model is likely to be an erratic predictor, especially when used outside the range of the original data.

Although more complex tests of specification have been developed (Ramsey, 1969, 1974), a simple specification test was used in this study. The deviates obtained when the model was fitted to independent test data were regressed against a second order polynomial in each of the independent variables. An analysis of variance was then used to determine whether or not the regressions were significant and the model misspecified.

ESTIMATION

Estimation of the parameters of linear models is relatively simple because an analytical solution exists (Kendall and Stuart, 1961; Theil, 1971; Draper and Smith, 1966). Algorithms which incorporate this technique have been implemented in a number of computer programs for multiple regression analysis and two well-known programs REX (Grosenbaugh, 1967) and SPSS (Nie et al., 1975) were used in this study.

For nonlinear models the objective function cannot be minimized analytically and a number of alternative algorithms (Goldfeld and Quandt, 1972; Sadler, 1975) have been developed to approximate the minimization iteratively.

The algorithms commence from feasible starting values for the parameters and aim to reduce the objective function by successively changing the parameter values until a minimum is reached.

There is no certainty that a particular algorithm will be satisfactory for all models and all data sets. Goldfeld and Quandt (1972) investigated a number of alternative algorithms, testing them against different models and data sets in an effort to compare their effectiveness. No one algorithm was the most efficient for all the examples, but two particular algorithms performed consistently well. Refined versions of these algorithms are implemented in a nonlinear parameter estimation program developed by Bard (1967); the Gauss-Newton method (Eisenpress and Greenstadt, 1966; Carroll, 1961), and the Davidon-Fletcher-Powell method (Fletcher and Powell, 1963; Sadler, 1975).

As the algorithms are iterative it is necessary to use terminating criteria to stop computation at an appropriate end point. Three criteria are appropriate.

1 The change in the parameter estimates between successive iterations should be within a preset tolerance.

$$|\Delta b_i| < d_1 (|b_i| + d_2) \text{ for the i parameters} \qquad \text{(III.15)}$$

In this criterion (Marquardt, 1963) d_1 is the desired tolerance for the parameter and d_2 ensures that if the estimated parameter is close to zero then computation will stop. Setting $d_1 = 0.00001$ and $d_2 = 0.001$ has been found to work well in practice (Marquardt, 1963).

2 The relative change in the objective function between iterations should also be within an arbitrary tolerance, commonly Marquardt's criterion. This is especially useful when the response surface of the objective function is relatively flat for changes in the parameters.

3 The number of iterations should be less than an arbitrary maximum so that if the other criteria fail because the algorithm cannot converge that particular model then computation will cease.

To ensure that the algorithm has converged, that is, a true minimum has been reached (a stationary minimum, but not necessarily a global one), the Hessian matrix (matrix of second order derivatives of the function) should be positive definite (Morrison, 1976). There is no guarantee that the minimum is a global one, but careful specification and testing of the model can reduce any doubt in this regard.

The nonlinear parameter estimation program of Bard (1967) uses Marquardt's (1963) criterion for convergence and provides information sufficient to show whether or not the Hessian matrix is positive definite. The program is flexible and relatively easy to use and was therefore used in this study.

A user supplied subroutine is required to evaluate the function and its partial derivatives. For most growth models the partial derivatives are complex so an additional subroutine was developed to evaluate the partial derivatives numerically. This reduced the complexity of the programming changes necessary between models. The technique adopted is detailed in Appendix 2.

TESTING

Three different types of tests were used in this study:

1 Hypothesis tests to determine whether one model is better than another or to determine whether a model is internally consistent.

2 Tests of the assumptions underlying the model.

3 Tests of the model as a predictor.

Hypothesis testing

In developing a satisfactory model it was necessary to discriminate between models and between alternative forms of each model by testing a null hypothesis against its alternative (Johnston, 1963; Draper and Smith, 1966; Wonnacott and Wonnacott, 1970). Two important considerations in deciding how the alternative hypotheses should be tested were

1 the choice of the test statistic, and,

2 the choice of the significance level.

Analysis of variance based on the F statistic or tests using the t statistic were used to test hypotheses concerning alternative forms of a particular linear model, such as the inclusion or otherwise of an additional parameter. These tests are well known and extensively documented (e.g. Lehmann, 1959).

For nonlinear models the situation is not so clear cut. As noted earlier the small-sample properties of OLS estimators of nonlinear models are not well established. Moreover, estimates of the precision of the parameter estimates are generally based on a linear approximation, which may or may not be sufficiently accurate (Guttman and Meeter, 1965), depending on the particular form of the model and the characteristics of the surface of the likelihood function.

Nevertheless Gallant (1975), who studied models similar in form to those of interest in this study, recommended the use of a test statistic C, analogous to the use of the F statistic in analyses of variance for linear models.

$$C = \sigma_1^2 / \sigma_2^2 \qquad \text{(III.15)}$$

where σ_1^2 and σ_2^2 denote the maximum likelihood estimates of the variance of the respective error terms in the two alternative models, $\sigma_2^2 > \sigma_1^2$.

The statistic is tested against the critical values of the statistic C*.

$$C^* = 1 - i\,F_p / (n - j) \qquad \text{(III.16)}$$

where   where eg $p = 0.05$

F_p = upper 100p% points of an F distribution,

i = number of parameters of interest,

j = the total number of parameters, and,

n = the total number of observations.

Again, following Gallant's (1975) work, the (approximate) t statistic was used to test hypotheses regarding the inclusion or otherwise of individual parameters in a particular model.

Tests of hypotheses involving disparate families of both linear and nonlinear models were achieved by comparing the predictive properties of the models using independent test data. This seemed more appropriate than the complex tests of Cox (1961, 1962).

The level of significance to be used in hypothesis testing must also be carefully considered, bearing in mind the possibility of both Type I errors (rejecting a correct null hypothesis)and Type II errors (accepting a false null hypothesis) (Lehmann, 1959).

Because there was little difference in the effect of each type of error it was desirable to balance the probability of each type of error. As the probability of a Type II error depends on the significance level selected (the probability of a Type I error), the model and the data, it was clearly impossible to set a priori the probability of a Type II error with any confidence.

When the dependent variable was yield and the data were pooled then p=.01 was selected as the appropriate level. When increment was the dependent variable, or when the model was fitted to individual plot data, then the lower level of p=.05 seemed more appropriate. These levels were used to test hypotheses both within and between models. When assumptions in the analysis were tested and when the model was evaluated as a predictor then p=.01 was used to ensure consistency between different developmental lines.

Testing the assumptions underlying the analysis

The models were tested using independent test data to ensure that the assumptions underlying the analysis were not violated. Bartlett's (1937) test was used to test that the variance of the error term was homogeneous. Three different tests were carried out on partitioned data:

1 data partitioned by plots,

2 data partitioned by age/site potential cells, and,

3 data ordered on the estimated value of the dependent variable

and divided into approximately equal sized partitions.

The Durbin-Watson d statistic (Durbin and Watson, 1950, 1951, 1971) was used to test for serial correlation. The test was only applied to the final model selected within each family of models. The data rarely allowed the test to be by plots, so in general the data were pooled, ignoring the slight bias that this may have introduced.

The moment statistics of skewness and kurtosis (Sokal and Rohlf, 1969; Snedecor and Cochran, 1967) were used to test that the errors were normally distributed, an assumption necessary for hypothesis testing.

Testing predictions

One common criterion for a suitable predictor is that it be unbiased over the whole of the regression surface. To test this the independent test data were partitioned and within each partition a t test was used to see whether the mean deviate was significantly different from zero.

Two different ways of partitioning the data were used for these t tests. Data were partitioned by plots in an attempt to discern whether there had been misspecification in relation to plot variables or error characteristics. Further subdivision of the observations within plots into age classes was not possible because of the small number of observations available in each plot. Hence the data were pooled and partitioned into site potential and age classes, in the hope that this would enable problems of misspecification relating to age to be discerned. Site potential was subdivided into three classes, based on boundary values of 200 and 100 m^3/ha. These values corresponded closely to the boundary values of SQ II and III, and SQ IV and V respectively. These data were further subdivided into age classes of 11-16, 17-23, 24-30, 31-39 and 40-50 years, the boundary values being chosen so that the classes spanned roughly equal ranges of yield.

Comparisons between models developed with different dependent variables (yield, periodic increment, transformed yield) were achieved by evaluating each model as a yield predictor on the independent test data. Other things being equal, the best model was selected from those which were unbiased, according to the standard deviation of the deviates.

<u>Summary of testing procedure</u>

The statistical methods adopted can be summarised as follows.

1 For parameter estimation of nonlinear models convergence was confirmed by checking that the Hessian matrix was positive definite.

2 Discrimination between alternative hypotheses was by:

  i Gallant's (1975) test on the variance ratio to test between nonlinear models, or an analysis of variance for linear models, and,

  ii t test on each parameter in turn to test the model for internal consistency.

3 The assumption underlying the analysis was tested on independent test data.

  i Bartlett's test (1937) was used to test for homogeneity of variance:

    a data partitioned by plots,

    b data partitioned by age and site potential, and,

    c data ordered by the estimated value of the dependent variable and partitioned into approximately equal cells.

ii Durbin and Watson's (1950, 1951, 1971) statistic was used to test for serial correlation on data ordered:

a by plots if there were sufficient observations, or

b pooled, ignoring the slight bias this introduces if there were insufficient data to test by plots.

iii Moment statistics of skewness and kurtosis (Sokal and Rohlf, 1969; Snedecor and Cochran, 1967) were used to test for normality.

iv As a further test of misspecification, the deviates were regressed against a second order polynomial in each of the independent variables.

4 The suitability of the model as a predictor was evaluated with independent test data by the following tests:

i the mean deviate for each plot was tested against t to determine whether misspecification had occurred.

ii The mean deviate for each age and site potential cell was tested against t to determine whether the model was biased, especially with respect to age.

5 Alternative prediction models which were otherwise satisfactory were compared by using them to predict yield for the observations in the independent test data. The model with the lowest standard deviation of the deviates was selected.

IV GROWTH AND YIELD MODELS

COMPARISON OF GROWTH AND YIELD MODELS

Graphical models

Polynomial

Grosenbaugh

Bertalanffy

Gompertz-Thomasius

Johnson-Schumacher
Backmann
Hugershoff-Bednarz

Summary

IV GROWTH AND YIELD MODELS

In the biological literature the relative importance of statistical analysis and biological inference in developing models is the subject of much debate. If a model is developed on a purely statistical basis without any deductive reasoning as to the form of the model, then it is likely to be satisfactory only when used in very restricted situations where the data are similar to those used in the estimation of the model. Under these circumstances extrapolation is dangerous and so are inferences at the extremes of the data range. On the other hand if no statistical analysis is used then the model will be of lesser practical value because there will be no indication of accuracy or precision. Kowalski and Guire cautioned (1974):

> "...it must be emphasized that finding a function which makes biological sense has much more to recommend it than searching for a function that will provide only a close mathematical fit. Mere goodness of fit is no justification for adopting a given function since several functions may fit the data equally well."

In principle both biological and statistical inference should be used to develop a model so that it will be useful in a wide range of practical situations. In practice this may be difficult to carry out successfully. Forest growth is the result of the complex interaction between many different and sometimes inter-related processes. Many of these processes have been modelled successfully, but it can be difficult to link them together into one coherent model. It is generally possible to use only relatively simple biological inference and this may tend to limit the formulation of biological hypotheses to very simple approaches.

The pattern of growth can be divided into three phases. In the initial juvenile phase both yield and growth rate are initially low, but both increase until growth rate reaches a maximum. After this phase growth rate declines, but at first mean annual increment still increases, a phase of relatively vigorous growth that changes into a senescent phase after mean annual increment culminates. These phases are shown diagramatically in Figure IV.1.

Figure IV.1 also shows that there are two alternative ways of looking at such a model, the first as a yield model and the second as a growth model. If both growth and yield models are being developed simultaneously for use in practice then they should be compatible, compatibility being formally defined by Clutter (1963) as:

> "when the yield model can be obtained by summation of the predicted growth through the appropriate growth periods or, more precisely, when the algebraic form of the yield model can be derived by mathematical integration of the growth model."

If the growth and yield models are not compatible according to this definition then two different model forms are being used.

COMPARISON OF GROWTH AND YIELD MODELS

Over the years a number of models have been developed for predicting either growth or yield, some simple some complex, and an initial review of these models was necessary to determine which warranted estimation. Of all the variables affecting growth the most important variable is undoubtedly age; indeed, in many of the models it is the only independent variable. This comparison of the various models only considers the effect of age, the other variables are considered later.

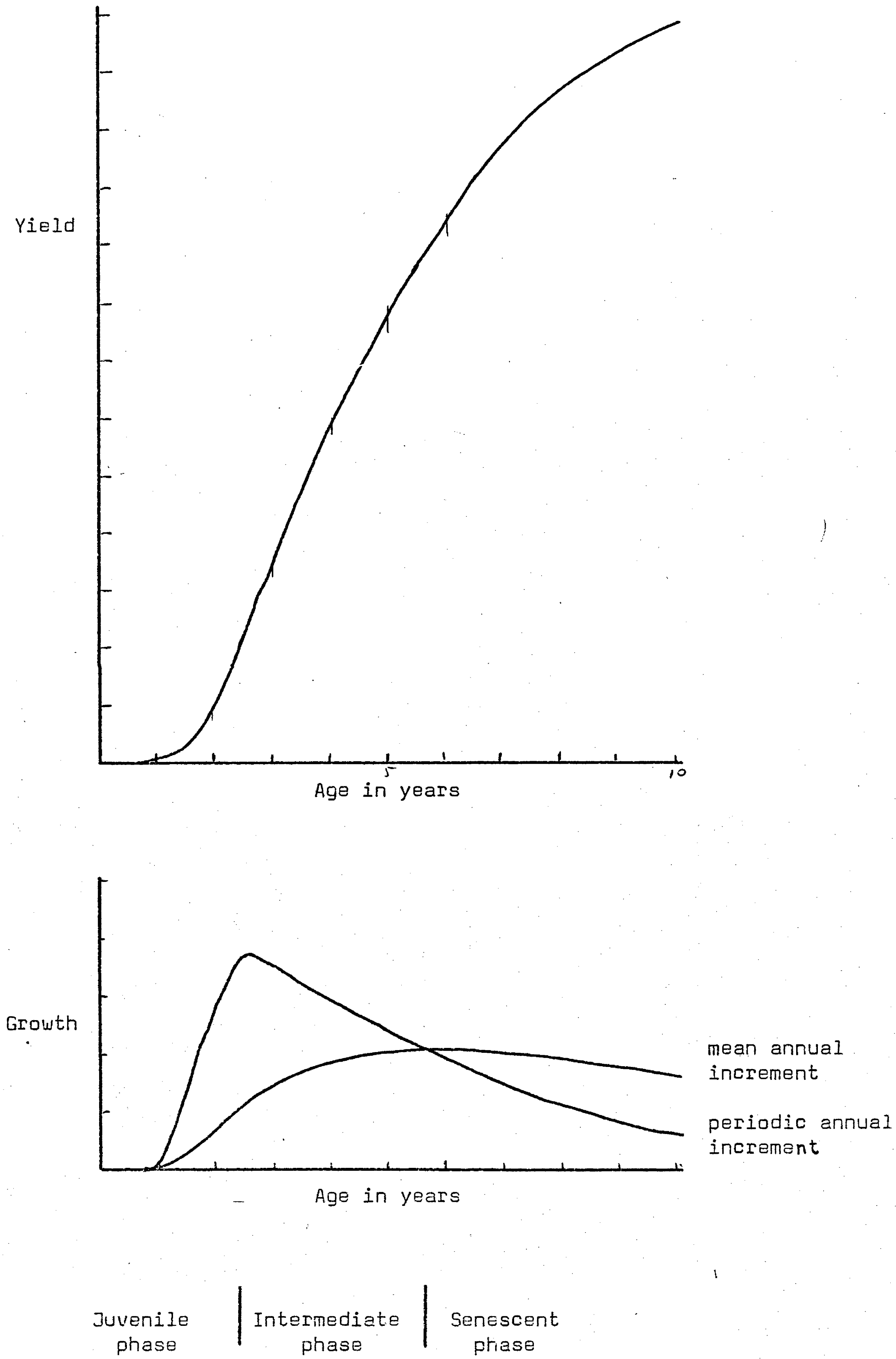
A growth or yield model should in general possess a few simple characteristics.

1 Yield should be zero at age zero, or if yield is to an arbitrary top diameter then yield should be zero at some finite, positive and small age (A_0).

2 Increment after A_0 should always be positive.

Figure IV.1

Relationship between growth and yield



3 Increment should have a single maximum (at age A_i) and after this age it should decrease as age increases.

4 Yield should approach a maximum yield (Y_{max}) asymptotically.

Of course if the data are inadequate or if the model is to be used over a relatively restricted age range then even these requirements may be relaxed.

Table IV.1 summarises many of the models that have been developed in equation form, including both the growth (derivative) and yield (integral) equations to facilitate comparisons. For convenience and consistency some of the models have been reformulated slightly.

Graphical yield models

Graphical yield models have been used in the past in many countries, and in South Australia they have been used for many years to produce radiata pine yield tables. The techniques are flexible and easy to use, but do not allow an estimate to be made of precision, and they are clearly liable to bias.

The first radiata pine yield table in South Australia was produced by Gray in 1931 (Lewis, Keeves and Leech, 1976) using the limiting curve method attributed by Spurr (1952) to Baur in 1877. This method uses single or spot estimates of yield to define upper and lower bounds to yield, these being then divided anamorphically into site potential classes. As more data became available the yield table was revised by Jolly in 1941 and later by Lewis in a series of revisions in 1953, 1957, 1960, 1963, 1968, 1970 and 1973 (Lewis, Keeves and Leech, 1976). As trend data became available the method changed to the directing curve trend method attributed by Spurr (1952) to Heyer in 1846.

These carefully derived graphs have been successfully used by the author in simulation studies (Lewis, Keeves and Leech, 1976). More

Table IV.1

Growth and yield models

Name	dY/dA	Yield	References
polynomial	$b_1 + \sum_{i=1}^{i=n-1} c_i A^i$	$a + \sum_{i=1}^{i=n} b_i A^i$	Marsh (Grut, 1970)
Mitscherlich	$n-pY$	$(n/p)\{1-\exp(-p(A-A_0))\}$	(1910)
logistic	$nY-pY^2$	$(p/n)\{1+\exp(n(A-A_i))\}^{-1}$	Grosenbaugh (1965) Pearl & Reed (1923)
Nelder (1)	$nY-pY^{(1+1/m)}$	$\left\{(p/n)\left[1+\frac{1}{m}\exp(-\frac{n}{m}(A-A_i))\right]\right\}^{1/m}$	(1961)
Nelder (2)	$nY-pY^{(1+m)}$	$\left\{(p/n)\left[1+m\exp(-nm(A-A_i))\right]\right\}^{m}$	(1962), Austin, Nelder & Berry (1964)
Pearl Reed	$q_{n-1}(A)\left[aY-bY^2\right]$	$k\{1+m\exp(q_n(A))\}^{-1}$	(1923)
von Bertalanffy (1)	$nY^{2/3}-pY$	$\left\{\frac{n}{p}\left[1-\exp(-\frac{1}{3}p(A-A_0))\right]\right\}^3$	(1941, 1942)
von Bertalanffy (2) (Chapman-Richards)	nY^m-pY	$\left\{\frac{n}{p}\left[1+c_1\exp(-p(1-m)(A-c_2))\right]\right\}^{\frac{1}{1-m}}$ if $m<1$ and $c_2=A_0$ then $c_1=-1$ if $m>1$ and $c_2=A_i$ then $c_1=(m-1)$	(1941), Richards (1959), Chapman (1961)
von Bertalanffy (3)	nY^m-pY^r	complex	(1941)
Gompertz	$cY\ln(Y/b)$	$a\exp(-\exp(-b(A-A_i)))$	(1825), Winsor (1932)
Thomasius	complex	$a\left\{1-\exp\left[-bA(1-\exp(cA))\right]\right\}^d$	(1964)
Johnson	$-bY/(c+A)^2$	$a\exp(-b/(c+A))$	(1935)
Schumacher	$+bY/A^2$	$\exp(a-b/A)$	(1939), Clutter (1963), Sullivan & Clutter (1972), Ferguson & Leech (1976a)
Backmann	$\exp(a+b(\ln A)+c(\ln^2 A))$	complex	(1943), Prodan (1968), Assmann (1970)
Hugershoff	$a A^2\exp(-cA)$	$\frac{2a}{c^3}\left\{1-\exp(-cA)\left[1+cA+\frac{c^2A^2}{2}\right]\right\}$	Prodan (1968)
Bednarz	$\frac{ab}{c}A^{(b-1)}\exp(-cY)$	$\frac{1}{c}\ln(aA^b+1)$	(1975)
Grosenbaugh	complex	$a+b\{\exp[(n^2-1)U]-nU\}^{nm+1}$ $U=\exp(-b(A-c))$ but U can be any function	(1965)

Where

$q_n(A)$ is a nth degree polynomial in age A, where n is an odd integer commonly 3

Y_r is yield at a base age r

complex indicates that the model is readily fitted in the form specified but less easily in the other form importantly they have been used quite successfully for many years to regulate the total cut in the profitable state afforestation enterprise in South Australia to as near maximum as possible.

However, the possibility of bias and the inability to calculate precision suggests that mathematically formulated models capable of objectively based statistical analysis may be more appropriate. Given adequate statistical precision, mathematical models are easier to use and revise for long term yield regulation calculations, especially with the computer oriented methodology now in use.

Polynomial

The polynomial is the simplest mathematical form for a growth or yield model. Providing that the order of the polynomial is high enough any functional form can be approximated.

$$Y = b_0 + b_1 A + b_2 A^2 + b_3 A^3 + \dots + b_n A^n \qquad \text{(IV.1)}$$

where

Y = yield,

A = age, and,

$b_0, b_1, b_2, b_3, \dots b_n$ are the parameters to be estimated.

Although precise unbiased estimates can be obtained for the parameters of a polynomial it is unlikely to be a satisfactory predictor. The model is likely to behave erratically at the extremities of the data and any extrapolation is extremely dangerous. The polynomial has been used by Marsh (Grut, 1970) and although computationally convenient the absence of any explicit biological structure was sufficient to cast doubt on its utility for this study. It cannot, for example, approach a maximum asymptotically.

Grosenbaugh

In 1965 Grosenbaugh formulated a complex growth model which included many other model forms as special cases. The function is

$$Y = a + b \left\{ \exp\left[(n^2 - 1)U\right] - nU \right\}^{nm+1} \qquad (IV.2)$$

where

$$U = \exp(-d(A-c)) \qquad (IV.3)$$

is often a suitable form for U, but it could be replaced by either a linear or a logarithmic function, and where,

Y = yield,

A = age, and,

a, b, c, d, n and m are parameters to be estimated, or are preset before analysis.

Grosenbaugh tabulated many of the more frequently used models specifying the form of the function U and the particular values of the parameters which each model implies. His objective was to develop a framework within which the various special cases could be compared for a particular data set. His challenge has not yet been taken up, probably because it makes almost impossible demands on the data and on analytical techniques, but the concept of defining a general model that is the starting point of the analysis of a set of data is very appealing.

A computer program to carry out the analyses in the way that Grosenbaugh envisaged was not available and it was considered impractical to develop one in the available time. This study therefore evaluated only a selected set of models.

Bertalanffy

In 1941 Bertalanffy proposed a growth model which seems to have a simple but well founded biological basis. The model was developed over many years (see also Bertalanffy, 1942, 1957, 1969) and encompasses many of the models previously developed. The model was developed from a study of the so-called allometric relationships in organisms (Huxley, 1932), attributed to Snell in 1891. An allometric relationship is said to occur when the relationship between two current attributes (for example volume (x_1), and height (x_2) of an organism can be expressed in the following form;

$$X_1 = a X_2^b \qquad (IV.4)$$

This arises from the assumption that in normal individuals of a population the specific growth rate of one variable has a constant proportional relationship to the specific growth rate of the other.

$$\frac{dX_1}{dA} \propto \frac{dX_2}{dA} \qquad (IV.5)$$

Some objections to this model have however been raised. In particular Haldane pointed out (Laird, 1965; Huxley, 1932) that if each part of an organism is allometrically related to each other part, then the growth of part of the organism is the sum of a number of exponential expressions.

$$Y = \sum_i a_i X_i^{bi} \qquad (IV.6)$$

This sum cannot equal a single allometric expression unless the exponents (the allometric constants) are the same. This is analogous to the pervasive problem of aggregation in econometrics (Theil, 1971). In practice the problem is generally ignored and the growth model is assumed to apply to the aggregate population under study.

Bertalanffy (1941) based his model on the hypothesis that the growth of an organism is the difference between anabolic growth rate (constructive metabolism) and catabolic rate (destructive metabolism), leading to the general form:

$$\frac{dY}{dA} = nY^m - pY^r \qquad \text{(IV.7)}$$

where

Y = yield,

A = age, and,

n, m, p and r are the parameters in the model.

Bertalanffy further noted that if Y was expressed as weight then the catabolic destruction rate could be taken as being proportional to the biomass of the organism itself, thus $r=1$, for many zoological genera. His zoological research suggested three groupings of which the first was the most common:

1 anabolic rate proportional to surface area, $m=2/3$,

2 anabolic rate proportional to weight, $m=1$, and,

3 anabolic rate intermediate between the two, $2/3<m<1$.

In spite of the fact that Bertalanffy recognised three groupings for zoological genera many workers have accepted $m=2/3$ for other biological applications without critically examining the inherent assumptions.

Because the simple model with $m=2/3$ did not perform well in other biological analyses the simple model was 'generalized' to the Chapman-Richards model (Richards, 1959; Chapman, 1961; Pienaar and Turnbull, 1973) although this is still a contraction of the general form that Bertalanffy proposed in 1941.[1]

$$\frac{dY}{dA} = nY^m - pY \qquad \text{(IV.8)}$$

(1) Equation IV.8 is very similar to an equation Verhulst (1844) records but did not pursue, presumably for practical reasons.

$$Y = \left\{\frac{n}{p}\left[1 + c_1 \exp(-p(1-m)(A-c_2))\right]\right\}^{\frac{1}{1-m}} \qquad (IV.9)$$

where

Y = yield,

A = age,

n, p and m are the parameters to be estimated, and, c_1 and c_2 are the constants of integration such that if $c_2=A_0$ the age at which volume growth commences then $c_1=-1$, or, if $c_2=A_i$ the age at which increment culminates then $c_1=(m-1)$ provided $m \neq 1$.

Three variants of Bertalanffy's general model were recognised for this study:

1 m and r allowed to float, the general model,

2 m allowed to float, r=1, the Chapman-Richards model, and,

3 m=2/3, r=1, the simple Bertalanffy model.

It is not possible to integrate the derivative equation for the general form except by numerical methods which were inappropriate for this study (A.Brown, pers.comm.). It could, however, be integrated for certain values of m and r, but the integral often involves exponential and trigonometric terms and generally has age as the dependent variable, which is unsatisfactory. The second level model is the most commonly used form in forestry and can be integrated using Bernoulli's equation (Appendix 3.1).

There are a number of other models that are submodels of the general Bertalanffy model although frequently developed independently, often prior to Bertalanffy's work.

The simplest form is the monomolecular equation (m=0, r=1) also called the "law of physiological dependence" (Assmann, 1970) which was first postulated in a forestry context by Mitscherlich (1910) who suggested that by augmenting a growth factor which is limiting, yield does not increase linearly with the increased factor, but in proportion to the difference between present and maximum yield. Parallel derivation to that used by Bertalanffy suggests that for this restricted case the catabolic rate is proportional to the yield of the organism itself and that the anabolic rate is constant regardless of biomass or time, conclusions that seem appropriate for an established forest crop with a relatively constant source of nutrients in the soil. The equation is:

$$\frac{dY}{dA} = n - pY \tag{IV.10}$$

$$Y = \frac{n}{p}(1 - \exp(-p(A-A_0))) \tag{IV.11}$$

where the variables and parameters are as for Equations IV.8 and IV.9.

The most commonly used growth model is the logistic or autocatalytic which probably originated (Pearl and Reed, 1923; Grosenbaugh, 1965), from the work of Verhulst (1844, 1846), and is a form with m=1, r=2 (or m=2, r=1, n and p negative). This equation is generally formulated in terms of A_i, the age of culmination of increment, and can be stated as:

$$\frac{dY}{dA} = pY - nY^2 \tag{IV.12}$$

$$Y = \frac{p}{n}(1 + \exp(n(A-A_i))) \tag{IV.13}$$

where the parameters and variables are as for Equations IV.8 and IV.9. This equation is symmetric about the point of inflection in the yield equation. For this equation anabolic rate is proportional to yield itself and catabolic rate is proportional to the product of anabolic rate and yield. To overcome some of the restrictions Nelder (1961, 1962) and Austin, Nelder and Berry (1964) proposed more general forms that parallel the second level Bertalanffy equation (see Table IV.1).

Pearl and Reed (1923) used an odd powered polynomial term to define a model of cyclical growth.

$$Y = Y_0 + k \left[1 + m \exp(q(A))\right]^{-1} \qquad \text{(IV.14)}$$

where

Y = yield,

Y_0 = yield at the commencement of the growth cycle and for this analysis is zero,

k = the asymptotic maximum yield for that cycle,

m = a parameter to be estimated, and,

$q(A)$ = an odd powered polynomial function of age A.

They found that a third order polynomial was generally satisfactory, using analytical means to fit the function to few data points. The equation offered no advantage over the second level Bertalanffy and was not considered further.

The allometric constant m in the second level Bertalanffy equation provides an estimate of the fraction of the asymptotic maximum yield that occurs at the culmination of increment (A_i). The fraction is as follows:

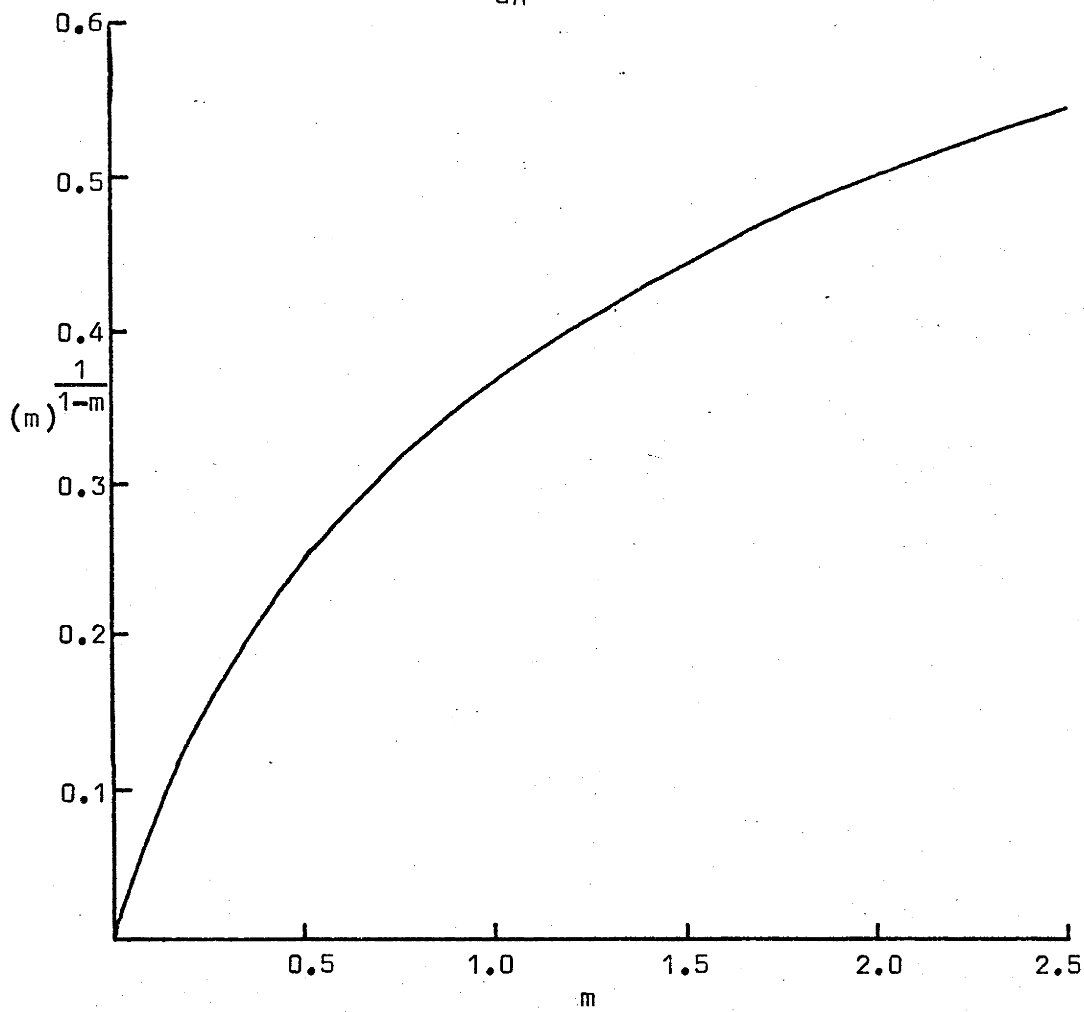
$$(m)^{\frac{1}{1-m}} \qquad \text{(IV.15)}$$

Figure IV.2 shows the way that this fraction changes with m and although for m=1 the fraction is undefined, the limit as $m \rightarrow 1$ is $1/e$. This is the same as for the Gompertz function to be discussed later. Richards (1959) and Pienaar (1966) used a very similar equation to the second level Bertalanffy and claimed that the limiting form as $m \rightarrow 1$ is the Gompertz equation. Pienaar's logic can however be shown to be false if the original Bertalanffy model is used instead of the Chapman-Richards formulation (Appendix 3), but this conclusion is supported by the evaluation of the fraction.

Figure IV.2

Proportion of asymptotic maximum yield that
occurs at age of maximum growth rate
Second level Bertalanffy model

$$\frac{dY}{dA} = nY^{m} - pY$$



m	$(m)^{\frac{1}{1-m}}$	Model
0	0.0	Mitscherlich
2/3	0.296	von Bertalanffy (simple)
1	(0.368)	(by interpolation)
2	0.5	logistic

Gompertz-Thomasius

One of the earliest of all growth models is that of Gompertz (1825) who in a treatise on life expectancy and the calculation of annuities developed a model of growth rate calculation later formulated by Winsor (1932), (see Table IV.1).

$$\frac{dY}{dA} = cY \ln(Y/a) \tag{IV.16}$$

$$Y = a\left\{\exp(-\exp(-b(A-A_i)))\right\} \tag{IV.17}$$

where

Y = yield,

A = age,

A_i = the age of culmination of increment, and,

a, b and c are the parameters to be estimated.

The equation has been widely used, apparently with success, to predict a wide variety of growth responses (Laird, 1965). For this equation increment culminates when yield is 1/e of the asymptotic maximum yield. This seemed unduly restrictive when compared with the more flexible second level Bertalanffy but the model was evaluated because it has been widely used in biological modelling.

Thomasius (1964) combined some of the logic of Mitscherlich (1910) and Gompertz to develop a model for forest growth which is more complicated than either and less well defined (Rawat and Franz, 1974).

$$Y = a\left\{1 - \exp(-b\ A(1 - \exp(cA)))\right\}^d \tag{IV.18}$$

where

Y = yield,

A = age, and,

a, b, c and d are the parameters to be estimated.

This model offered little unless the Gompertz proved to be as satisfactory or better than the other alternative models tried.

Johnson-Schumacher

In 1935 Johnson proposed a simple model for growth to be used after the culmination of periodic annual increment. This was used in simplified form by Schumacher (1939) because it facilitated the formulation of a simple combined model including site potential and stand density as well as age that could be estimated by multiple linear regression analysis.

$$\frac{dY}{dA} = bY / (c+A)^2 \qquad (IV.19)$$

$$Y = a \exp(-b/(c+A)^2) \qquad (IV.20a)$$

or

$$\ln(Y) = \ln(a) - b/(c+A)^2 \qquad (IV.20b)$$

where

Y = yield,

A = age, and,

a, b and c are the parameters to be estimated, and

where for the Schumacher model c=0.

Clutter (1963) used the Schumacher form but Bailey and Clutter (1974) raised age to a power in an effort to define a more flexible model. An inherent assumption of the model is that the age of culmination of mean annual increment is twice the age of culmination of current annual increment. Because of this restriction and because the model has only a limited biological basis it was considered likely that the Johnson-Schumacher model would be inferior to the second level Bertalanffy model. However the model was evaluated because it has been used to predict forest growth satisfactorily.

Backmann

Backmann's formula for forest growth (Prodan, 1968) was based on the premise that the logarithm of growth is proportional to the square of the logarithm of time.

$$\frac{dY}{dA} = \exp(a + b \ln(A) + c (\ln(A))^2) \qquad (IV.21)$$

where

Y = yield,

A = age, and,

a, b and c are the parameters to be estimated.

Although this derivative equation leads to a complex yield equation it can be readily used in practice using arithmetic probability paper. Increment culminates at 15.9% of the asymptotic maximum yield which seemed an unnecessary restriction with little or no biological basis. The equation was considered unlikely to be as satisfactory as the second level Bertalanffy and was not evaluated.

Hugershoff-Bednarz

The Hugershoff equation (Prodan, 1968) assumes that the juvenile phase of growth can be approximated by a quadratic function in age and the senescent phase by an exponential decay model, one phasing into the other in an intermediate stage between the culmination of current and mean annual increment.

$$\frac{dY}{dA} = a A^2 \exp(-cA) \qquad (IV.22)$$

$$Y = \frac{2a}{c^3} \left\{ 1 - \exp(-cA) \left[1 + cA + \frac{c^2 A^2}{2} \right] \right\} \qquad (IV.23)$$

where

Y = yield,

A = age, and,

a, b and c are the parameters to be estimated.

Bednarz(1975) suggested a modification to the two components of the derivative that makes the model more flexible as well as being a mathematically simpler yield form. The model is

$$\frac{dY}{dA} = \frac{ab}{c} A^{(b-1)} \exp(-cY) \qquad (IV.24)$$

$$Y = \frac{1}{c} \ln(aA^{b} + 1) \qquad (IV.25)$$

where

Y = yield,

A = age, and,

a, b and c are the parameters to be estimated, but Bednarz fitted the equation conditioned such that at a base age r, yield was Y_r,

$$Y = Y_r \left\{ \frac{\ln(aA^{b} + 1)}{\ln(ar^{b} + 1)} \right\} \qquad (IV.26)$$

to reduce the number of parameters to be estimated. The model is one of the few models that does not reach an asymptotic maximum yield, yield continuing to increase with increasing age. The Bednarz model was evaluated because it has been previously used to predict radiata pine height growth, in spite of its lack of a coherent biological base.

Summary

Graphical models were considered an inappropriate form for this study and only equation forms were considered. Of these the Bertalanffy model appeared to offer the greatest flexibility, satisfying all the simple biological criteria. Unlike many of the other forms the culmination of increment is not rigidly defined in terms of either a fixed proportion of asymptotic maximum yield or of the age of culmination of mean annual increment. The general form was considered less appropriate than the simpler second level form because it could only be integrated for particular values of the parameters m and r, and not over the complete range.

Graphs of yield against age suggested that increment probably culminates at a relatively young age for most plots, possibly at or even before the first volume measurement at age 8 to 12 (Appendix 1.3b and 1.3c). Models such as the Johnson-Schumacher or Bednarz were thus possibly satisfactory predictors within the range of the data and for that reason were evaluated. The double exponential Gompertz form and the polynomial were evaluated for completeness rather than from any sense of probable utility.

The other models in Table IV.1 were either not evaluated (Backmann, Pearl-Reed and Hugershoff), were evaluated as part of the evaluation of the second level Bertalanffy (Mitscherlich, logistic, simple Bertalanffy and Nelder), or were reconsidered after a simpler form had been evaluated (Thomasius). Grosenbaugh's form was not evaluated because it was impractical.

PART 1

ORDINARY LEAST SQUARES ESTIMATION

V EXPLORATORY ANALYSES OF DATA FROM UNTHINNED STANDS

DATA

SITE POTENTIAL

Conditioning

Model formulation

ANALYSIS

Form of the dependent variable

Single or two stage analysis

STRATEGY

BERTALANFFY MODEL RESULTS

The evaluation of the general model

Periodic annual increment, conditioned

Periodic annual increment, unconditioned

Derivative

Yield, conditioned

Yield, two stage

Summary

JOHNSON-SCHUMACHER

Linear models

Nonlinear models, unconditioned

Nonlinear models, conditioned

Summary

BEDNARZ

OTHER MODEL FORMS

Gompertz

Polynomial

Lewis's yield table

SUMMARY

V EXPLORATORY ANALYSES OF DATA FROM UNTHINNED STANDS

The exploratory analyses were carried out on data from unthinned stands partly to reduce the number of variables to be considered and partly because it seemed more appropriate to first investigate stands that had not been artificially modified by thinning.

DATA

There are a number of plots in the south-east of South Australia which have never been thinned in order to provide a control series against which the more numerous thinned plots can be compared. These plots cover a representative range of age and site potential and therefore provide good data sets on which to evaluate the models of growth and yield for unthinned stands.

The data from unthinned stands were extracted from the data base and divided into two sets.

1 Developmental data; comprising a minimum of nine measurements of volume for each plot over a minimum twenty year growth period.

2 Test data; between five and eight measurements for volume over at least a fifteen year growth period.

The data were divided according to the number of measurements and growth period because it was intended to evaluate the use of individual plot trends. Long trends with at least nine measurements were highly desirable for this type of analysis. There were insufficient of these plots to allow random allocation into development and test sets, so the plots with the shorter trends provided the independent test data.

The two data sets are summarised in Appendix 1.3. The twenty plots in the developmental data included 228 volume measurements with an average growth period of 33 years. The twenty three plots in the independent test data had 157 volume measurements with an average growth period of 20.8 years.

The developmental data cover a narrower range in several respects than the test data. Of the test data plot EP24C was planted at 6x6 feet, EP24E at 9x9 feet, whereas all the other plots in both data sets were planted at 7x7, 8x8 or 9x7 feet. Four plots in the test data, EP24C, EP24E, 433 and 155, are of higher site potential than the developmental data, and three plots, 149, 368 and 369, are poorer. The test data also cover a wider range of forest district, but because of the way the data were allocated there are few measurements at later ages. Because the two data sets were not allocated at random the models developed may be open to question if used to predict the wider range covered by the test data. This problem was not considered critical at this stage because the objective was to use these exploratory analyses to narrow the number of models to be fitted and evaluated, not to arrive at a final prediction model *per se*.

SITE POTENTIAL

Conditioning based on site potential and the effect of site potential on yield are two aspects which warranted careful study.

Conditioning

For a number of growth and yield models it was possible to condition the yield model so that at age 10 yield is the value of site potential (Y_{10}), thus eliminating one or more parameters. For example the second level Bertalanffy model can be conditioned thus:

$$Y = Y_{10}\left\{\frac{1 - \exp(-p(1-m)(A-A_0))}{1 - \exp(-p(1-m)(10-A_0))}\right\}^{\frac{1}{1-m}} \tag{V.1}$$

where

Y = yield,

A = age,

Y_{10} = site potential, yield at age 10, and,

p, m and A_0 are the parameters.

Bailey and Clutter (1974) have argued that conditioning places too much weight on measurement and other errors associated with site potential (or in their case, site index). This seemed unduly pessimistic in this instance. Conditioning is identical to the imposition of an exact constraint on the parameters of the stochastic model. Linear theory (Goldberger, 1964; Theil, 1971) shows that the imposition of exact linear constraints and solving by OLS yields minimum variance, unbiased estimators and thus predictors. It is not clear whether these results hold for nonlinear models. However to the extent that most similar results hold asymptotically for estimators of nonlinear models, it could be expected that these results would also hold asymptotically.

A more powerful technique might be to treat the observations of site potential as unbiased estimators of constraints on the parameters. However this would necessitate precise estimates of the variance associated with the estimate of site potential, and far more complex techniques of parameter estimation. Given that the values of site potential are precise, neglect of these errors seemed unlikely to be of much consequence. A gross check is feasible however. Comparison of the conditioned and unconditioned versions of any one model using the usual tests can eliminate the unlikely possibility that the errors attached to site potential are so large that conditioning results in markedly poorer estimators.

Model formulation

Conditioning relates to the data and parameters for individual plots. It may not serve to take the effects of site potential fully into account, across the entire data set. The remaining parameters of the model may themselves be functions of site potential.

One well known technique for taking these effects into account is to develop an average yield curve and then assume that the other curves are anamorphic, a fixed amount or fraction above or below the average yield curve. Although used by Bednarz (1975) to predict upper stand height growth, the assumption of similar shape is not valid for volume to a top diameter limit, because it takes a varying number of years (A_0) for growth to commence, with growth commencing earlier on the better sites.

In the Bertalanffy model the relationship between culmination of increment and the asymptotic maximum yield is dependent on the allometric constant m (Figure IV.2). As increment culminates for radiata pine at an early age, possibly at or before the first measurement included in the data, (see Appendix 1.3), it seemed unlikely that the effect of site potential on the parameter m could be estimated. Studies by Brickell (1968) and Beck (1971) were also unable to relate m to site potential.

Replacing m by a linear function in site potential would allow the relationship between the age of culmination of current annual increment and the age of culmination of mean annual increment to vary, but it was thought unlikely that satisfactory estimates could be obtained from the available data.

The parameters n and p in the second level Bertalanffy model combine to provide an estimate of the asymptotic maximum yield that a site can achieve. For anamorphic yield curves this asymptotic maximum yield is assumed to be a linear function of site potential.

$$Y_{max} \equiv (n/p)^{\frac{1}{1-m}} = b_0 + b_1 Y_{10} \qquad (V.2)$$

where

Y_{max} = asymptotic maximum yield,

n and p are the parameters of interest that together with the allometric constant m combine to provide the asymptotic maximum yield,

Y_{10} = site potential, and,

b_0 and b_1 are the parameters to be estimated.

Beck (1971) used this linear form to estimate height growth of white pine, although Brickell (1968) added a quadratic term making the yield curves polymorphic. Thus Equation V.2 could be replaced by:

$$n = n_0 + n_1 Y_{10} + n_2 Y_{10}^2 \qquad (V.3)$$

$$p = p_0$$

where n_2 would be equal to zero for anamorphic yield curves, but not equal to zero for polymorphic curves. In terms of Bertalanffy's deductions this implies that the anabolic rate is proportional to site potential but that the catabolic rate is not.

The contradictory hypothesis that anabolic rate is independent of site potential and that catabolic rate decreases with increasing site potential also seemed to be biologically plausible.

$$p = p_0 - p_1 Y_{10} \qquad (V.4)$$

$$n = n_0$$

For a yield model based on Equation V.4 the rate of increase of asymptotic maximum yield increases with increasing site potential, supporting Brickell's model form. This form was also supported by the work of Cilliers and van Wyk (1938). Since both models seemed plausible, both were evaluated.

The age at which volume growth commences (A_0) was incorporated into the models by replacing age A with ($A-A_0$). Site potential influences A_0 so that as site potential increases, A_0 decreases from the base age for the yield curves but never reaches zero. This could be approximated by a simple linear function within the range of the data, or could be fitted by a quadratic. Exponential or reciprocal forms seemed more logical as they allow A_0 to decrease as site potential increases without introducing a turning point.

$$A_0 = a_0 \qquad (V.5a)$$

$$A_0 = a_0 - a_1 Y_{10} \qquad (V.5b)$$

$$A_0 = a_0 + a_1 Y_{10} + a_2 Y_{10}^2 \qquad (V.5c)$$

$$A_0 = a_0 \exp(-a_1 Y_{10}) \qquad (V.5d)$$

$$A_0 = a_0 + a_1 / (a_2 + Y_{10}) \qquad (V.5e)$$

where

A_0 = the age at which volume growth commences,

Y_{10} = site potential, and,

a_0, a_1 and a_2 are the parameters to be estimated.

When site potential is zero then A_0 should be ten. However this was outside the range of the data, and although conditioning to this effect seemed desirable it was not considered essential. After the models were fitted A_0 was tested to determine whether this conditioning was acceptable. If the null hypothesis was accepted then the simpler submodel for A_0 was selected.

Similar derivations were used for parameters in models other than those based on the Bertalanffy form. For most parameters in the equations in Table IV.1 simple polynomials were used to test the effect of site potential.

$$b = b_0 \qquad (V.6a)$$

$$b = b_0 + b_1 Y_{10} \qquad (V.6b)$$

$$b = b_0 + b_1 Y_{10} + b_2 Y_{10}^2 \qquad (V.6c)$$

where

b represents the parameters of the equations in Table IV.1,

Y_{10} = site potential, and,

b_0, b_1 and b_2 are the parameters to be estimated.

For the Johnson-Schumacher model these formulations encompassed the work of Schumacher (1939), Clutter (1963) and Ferguson and Leech (1976a), but as the dependent variable was not yield but the logarithm of yield, two other submodels were also evaluated.

$$b = b_0 + b_1 / Y_{10} \qquad (V.6d)$$

$$b = b_0 + b_1 \ln(Y_{10}) \qquad (V.6e)$$

ANALYSIS

Before the models could be evaluated two further facets of the analytical procedure had to be considered.

1 The form that the dependent variable should take; yield or increment, and if increment, whether instantaneous (the derivative) or periodic.

2 Whether the data should be pooled and the model developed in a single stage process or whether a two stage procedure should be used, the first estimating the parameters for each plot and the second evaluating the effect of site potential.

Form of the dependent variable

There are a number of ways of specifying the dependent variable in a growth or yield model. Three of these were considered: yield, instantaneous increment at a given time, and increment over a given period. The choice of the form of the dependent variable depended in part on the available data and in part on statistical considerations.

If yield is to be used as the dependent variable then serial correlation is likely to be a problem. Yield at any age is largely dependent on yield at earlier ages, especially late in the rotation when increment represents only a small proportion of current volume. This could have been avoided by including an autoregressive disturbance proportional to the earlier measurement as described in Chapter III, but this would have made the analysis more complicated than seemed warranted.

A simpler approach was to assume that first order serial correlation existed between successive measurements of yield and to estimate increment rather than yield. If periodic increment (Pi) is the dependent variable then an autoregressive process is implicitly built into the model. If yield at age A is Y_A and yield i years later is Y_{A+i} then fitting

$$Y_{A+i} = f(A, Y_{10}) \qquad (V.7)$$

assumes that yield is not serially correlated, whereas,

$$Y_{A+i} = s\, Y_A + f(A, Y_{10}) \qquad (V.8)$$

assumes serial correlation. If the coefficient of serial correlation s equals one then periodic increment (Pi) can be estimated:

$$Pi = Y_{A+i} - Y_A = f(A, Y_{10}, i) \qquad (V.9)$$

with the inclusion of the increment period i allowing for the data having variable intervals between measurement. The right hand side of the equation is better formulated in terms of periodic annual increment (Pai) rather than periodic increment (Pi) to make the variance of the dependent variable more homogeneous.

$$Pai = Pi / i = (Y_{A+i} - Y_A) / i \qquad (V.10)$$

If instantaneous increment is used as the dependent variable then the derivative form of the yield model can be evaluated. Because the general form of the Bertalanffy model cannot be integrated to provide a yield equation, the derivative is the only way that the general form can be evaluated. However because the data are estimates of periodic annual increment generally after the age of culmination of increment the derivative models will necessarily have biased estimates of the parameters. Unbiased estimators can be achieved if periodic annual increment is the dependent variable by formulating the function as the difference between two yield equations.

Of the three forms, the difference equation was preferred because it was thought to have fewer problems with respect to serial correlation than yield. The derivative was however evaluated to see whether the contraction of the general Bertalanffy model to the second level form was acceptable, that is, whether r could be taken as 1.0.

Appendix 4.1 shows that for the data sets used the variance of both periodic annual increment and yield could be considered homogeneous and that weighting was unnecessary.

Single or two stage analysis

In a single stage analysis the data are pooled and the final model including age and site potential is estimated from the data. In a two stage analysis a model expressing yield (or periodic annual increment) as a function of age is fitted to each of the plots in turn and then, in a second stage, these estimated parameters are related to site potential and possibly other stand variables.

Single stage analysis had the advantage that the data set was much larger, but the information advantage inherent in the long term trend series was ignored. On the other hand, the two stage analysis takes advantage of the long term trend nature of the data but there are relatively few observations for each of the first stage analyses. Both approaches were evaluated.

STRATEGY

In defining a strategy to be adopted for model evaluation five factors had to be considered.

1 Which functional relationship in terms of age should be used.

2 Which form of the dependent variable should be used: yield, instantaneous increment or periodic annual increment.

3 Whether single or two stage analysis should be used.

4 Whether the model should be conditioned or not.

5 Which functional relationships with respect to site potential should be used.

As it was clearly impractical to carry out an exhaustive analysis of all these facets, the approach adopted was to define *a priori* a relatively simple model and to use this as a base for comparison. Models allowing each factor to vary separately were then evaluated and the results compared with those for the base model. In the event of the base model being inferior a new base model was defined and the analytical procedure repeated. This strategy does not guarantee that the optimal model was selected, but careful analysis, including the fitting of many models that have not been reported in this thesis ensured that most possibilities were tested and compared.

In defining the base model simplicity was important to provide a starting point for comparisons. The base model was as follows.

1 The second level Bertalanffy model was preferred to the other forms in Chapter IV because it appeared to offer the greatest flexibility and satisfied the simple biological criteria.

2 Periodic annual increment was selected as the dependent variable rather than yield to reduce the effects of serial correlation.

3 A single stage analysis was preferred to the two stage analysis because the advantages inherent in the larger number of observations seemed to outweigh the lack of recognition of the trend nature of the data.

4 The conditioned model was preferred to the unconditioned model for the base model because it had fewer parameters to be estimated.

5 The relationship between the parameters and site potential selected were: for p a linear function in Y_{10}, and for A_0 an exponential decay model.

The base model was therefore as follows:

$$Pai = (Y_{A+i} - Y_A) / i \qquad (V.10)$$

$$Y_A = Y_{10}\left\{\frac{1 - \exp(-p(1-m)(A-A_0))}{1 - \exp(-p(1-m)(10-A_0))}\right\}^{\frac{1}{1-m}} \qquad (V.1)$$

$$p = p_0 - p_1 Y_{10} \qquad (V.4)$$

$$A_0 = a_0 \exp(-a_1 Y_{10}) \qquad (V.5d)$$

where

Pai = the dependent variable, periodic annual increment,

Y_A = yield at age A,

A = age,

i = the increment period,

Y_{10} = site potential, yield at age 10, and,

p_0, p_1, a_0, a_1 and m are the parameters to be estimated.

BERTALANFFY MODEL RESULTS

The base model selected was a variant of the second level Bertalanffy model that assumes that the parameter r in the general form (Equation IV.7) can be taken as 1.0. Therefore before the base model was fitted it was necessary to investigate whether the contraction to the second level Bertalanffy form was satisfactory.

The evaluation of the general model

The general model could not be integrated and the only way that the allometric constant r could be evaluated was using the derivative form. Because the data were measurements of periodic increment and not instantaneous increment any parameter estimates were necessarily biased, the extent of the bias depending on the length of the increment period.

This bias could be avoided for the simple Mitscherlich form ($m=0.0$, $r=1.0$) by using a Taylor's series expansion to be described later, but this was inappropriate for the general form. Because the analysis aimed only to evaluate whether r could be taken as 1.0, and did not aim to develop efficient estimators and predictors, it was desirable only that the bias be consistent. The bias did not need to be eliminated. To achieve this the data for this analysis were culled to 103 increment periods of either one or two years.

The general form of the Bertalanffy model was then fitted to the data, together with reduced forms with specific values of m and r. Three secondary models including site potential were evaluated.

$$n = n_0 \qquad p = p_0 + p_1 Y_{10} \qquad (V.11a)$$

$$n = n_0 + n_1 Y_{10} \qquad p = p_0 \qquad (V.11b)$$

$$n = n_0 + n_1 Y_{10} \qquad p = p_0 + p_1 Y_{10} \qquad (V.11c)$$

where

n and p are the parameters in the general Bertalanffy model,

Y_{10} = site potential, and,

n_0, n_1, p_0 and p_1 are the parameters to be estimated.

Of these three submodels Equation V.11b explained slightly more of the variation than Equation V.11a. Although Equation V.11c had an even lower residual sum squares, the estimated parameters had inflated standard errors and were not significantly different from zero. The following results are based on Equation V.11b, but the trend was consistent for all three equations and similar conclusions would have been drawn if they had been used.

The simple Mitscherlich form (m=0, r=1) proved to be a satisfactory form with residual sum squares of 5830.7 compared with 5819.8 if both m and r were allowed to float. The reduction in residual sum squares by allowing either m or r to float was not significant. Models with m=0.0 and r=0.5 and 0.667 respectively were slightly more efficient than the Mitscherlich form but the reduction in residual sum squares was small and was considered insufficient to offset the more complicated form of the yield model. For the second level Bertalanffy (residual sum squares 5823.3) the parameter m was not significantly different from zero. This probably reflects the lack of early age data rather than any inherent structural weakness in the model. However, although the results indicated that the parameter m could be taken as zero the parameter was included in the base model as this inference was of necessity based on a reduced data set. Appendix 4.2 details the residual sum squares for all the models fitted using Equation IV.11b, in both graphical and tabular form.

Periodic annual increment, conditioned

The base model was then fitted to the 208 increment periods of the developmental data (Model 1, Table V.1), together with a number of other models with various structures for the parameters m, p and A_0.

Of the various submodels for A_0 evaluated (Models 2-7, Table V.1) the base model with two parameters (Model 1) was not significantly better than the base model with a_0 set to 10.0 (Model 2). This was sensible as at $Y_{10} = 0$ then A_0 should equal ten. The addition of a constant to the base model (Model 7) was not significant. Neither the polynomial form (Models 3 and 4) nor the reciprocal (Models 5 and 6) were as good as the base model.

For the parameter p, Models 8 and 9 show that the addition of the linear term in Y_{10} was not significant if the two parameter model for A_0 was used, but was significant if the single parameter conditioned model for A_0 was used. Model 2 had a lower residual sum squares than Model 8 for the same number of parameters and was preferred. Models 10 and 11 replacing the linear function by a quadratic were not significantly better than Models 1 and 2.

The allometric constant m was not significantly different from zero for either Model 1 or 2 using a t test, and Gallant's test also showed that Models 12-15 were not significantly better than Models 8, 9, 1 and 2 respectively. Replacing m by a linear function in Y_{10}, Models 16 and 17 also showed no significant improvement.

Table V.1

Bertalanffy; conditioned periodic annual increment

Model				Number of parameters	Residual sum squares	Compared with	
No	Form					Model	Significance
1	m	$p=p_0+p_1Y_{10}$	$A_0=a_0\exp(-a_1Y_{10})$	5	8499.6		
2	m	$p=p_0+p_1Y_{10}$	$A_0=10.0\ \exp(-a_1Y_{10})$	4	8541.6	1	NS
3	m	$p=p_0+p_1Y_{10}$	$A_0=a_0+a_1Y_{10}$	5	8813.7		
4	m	$p=p_0+p_1Y_{10}$	$A_0=a_0+a_1Y_{10}+a_2Y_{10}^2$	6	8689.7	3	NS
5	m	$p=p_0+p_1Y_{10}$	$A_0=a_0/(a_1+Y_{10})$	5	8540.3		
6	m	$p=p_0+p_1Y_{10}$	$A_0=a_0+a_1/(a_2+Y_{10})$	6	8531.7	5	NS
7	m	$p=p_0+p_1Y_{10}$	$A_0=a_0+a_1\ \exp(-a_2Y_{10})$	6	8490.1	1	NS
8	m	p	$A_0=a_0\ \exp(-a_1Y_{10})$	4	8555.1	1	NS
9	m	p	$A_0=10.0\ \exp(-a_1Y_{10})$	3	8756.0	2	Sig
10	m	$p=p_0+p_1Y_{10}+p_2Y_{10}^2$	$A_0=a_0\ \exp(-a_1Y_{10})$	6	8472.7	1	NS
11	m	$p=p_0+p_1Y_{10}+p_2Y_{10}^2$	$A_0=10.0\ \exp(-a_1Y_{10})$	5	8499.1	2	NS
12	m=0.0	p	$A_0=a_0\ \exp(-a_1Y_{10})$	3	8565.7	8	NS
13	m=0.0	p	$A_0=10.0\ \exp(-a_1Y_{10})$	2	8766.2	9	NS
14	m=0.0	$p=p_0+p_1Y_{10}$	$A_0=a_0\ \exp(-a_1Y_{10})$	4	8493.8	1	NS
15	m=0.0	$p=p_0+p_1Y_{10}$	$A_0=10.0\ \exp(-a_1Y_{10})$	3	8562.8	2	NS
16	$m=m_0+m_1Y_{10}$	$p=p_0+p_1Y_{10}$	$A_0=a_0\ \exp(-a_1Y_{10})$	6	8499.1	1	NS
17	$m=m_0+m_1Y_{10}$	$p=p_0+p_1Y_{10}$	$A_0=10.0\ \exp(-a_1Y_{10})$	5	8541.3	2	NS

Note significance shows whether or not the more complicated model was significantly better than the simpler model.

The base model was therefore rejected in favour of the simpler form with two fewer parameters, Equation V.12. For the equation the standard error of each parameter estimate is shown immediately below the estimated parameter.

$$
\begin{aligned}
m &= 0.0 \\
p &= \underset{(0.00108)}{0.02587} - \underset{(0.0503\ 10^{-4})}{0.4627\ 10^{-4}}\, Y_{10} \\
A_0 &= 10.0 \exp(\underset{(0.000052)}{-0.003716}\ Y_{10})
\end{aligned}
\qquad \text{(V.12)}
$$

This model confirmed the conclusion of the exploratory analysis using the derivative that the simple Mitscherlich form with m=0.0 was satisfactory. Also, the simpler submodel for A_0 was biologically sensible, in part validating the statistical analysis.

The underlying assumptions of the analysis were then tested using the 23 plots in the independent test data. Bartlett's test for homogeneity showed no significant heterogeneity when partitioned either by plots, the expected value of the dependent variable, or into age/site potential cells. The deviates were normally distributed, neither the moment statistics of skewness nor kurtosis being significant. However the Durbin-Watson d statistic was 0.998, indicating significant serial correlation. This was surprising as it had been hoped that by using periodic annual increment serial correlation would be avoided.

The Durbin-Watson d statistic for the developmental data was however not significant, indicating that perhaps some difference between the two data sets may be responsible. The developmental data are generally from pre-1940 plantations whereas the test data from post-1940 plantations; the two sets may therefore reflect changes in the pattern of soil type or other geographical variation. The results of the tests are summarised in Table V.2. Regressing the deviates against linear and quadratic models in age and site potential showed the deviates to be

Tests of models

Equation	Model	As a yield predictor			Assumptions of analysis					
		Standard deviation of deviates	Number of cells significantly different from zero		Heterogeneity (X^2)			Normality (t)		Serial correlation (d)
			/23 plots	/13 A/Y_{10} cells	by plots	by A/Y_{10}	by Y	skewness	kurtosis	
V.12	Bertalanffy: pai conditioned	51.51	6	1	NS	NS	NS	NS	NS	0.998
V.13	Bertalanffy: pai unconditioned	57.53	9	3	40.48	NS	NS	NS	NS	0.928
V.19	Bertalanffy: derivative	60.06	7	2	41.82	NS	NS	NS	NS	0.958
V.21	Bertalanffy: yield conditioned	50.50	7	1	47.79	57.42	29.82	NS	NS	0.567
V.22	Bertalanffy: two stage yield conditioned	66.10	8	3						
V.23	Bertalanffy: two stage yield unconditioned	113.72	13	8						
V.27	Johnson-Schumacher: linear unconditioned	61.97	8	4	105.53	87.52	58.21	4.337	8.038	0.469
V.28	Johnson-Schumacher: linear unconditioned	58.10	8	3	146.60	101.96	65.68	NS	10.915	0.644
V.32	Johnson-Schumacher: nonlinear unconditioned	61.62	8	3	64.71	65.30	17.32	-3.460	NS	0.577
V.33	Johnson-Schumacher: nonlinear unconditioned	64.64	11	3	55.95	50.22	NS	NS	NS	0.296
V.37	Johnson-Schumacher: nonlinear conditioned	54.04	8	2	56.02	62.85	30.37	NS	5.640	0.645
V.39	Bednarz: yield conditioned	50.68	6	1	49.08	58.23	29.51	NS	2.812	0.580
V.40	Bednarz: pai conditioned	51.55	7	1	44.94	NS	NS	NS	NS	1.008
V.41	Gompertz: yield conditioned	62.43	8	3	71.12	64.62	35.06	3.696	5.906	0.787
V.42	polynomial: unconditioned	61.72	10	5	67.71	62.43	NS	-5.397	6.254	0.500
	Lewis's Yield Table	52.25	6	2						

Note: Only significant statistics are shown

Blank implies that the test was not applicable independent of both. Replacing the parameters in the linear model by dummy variables for each plot showed that the deviates were related to age for each plot ($F_{45/101}$ =5.78), suggesting misspecification.

The model was then evaluated as a yield predictor using the independent test data. Of the 23 plots 6 had mean deviates significantly different from zero, not unexpected in view of the earlier tests; as did 1 of the 13 age/site potential cells. Excluding from the test data the one plot planted at 6x6 feet overcame the latter problem. This stocking was outside the range of the developmental data and it was concluded that the problem was one of misspecification that could possibly be avoided in later analyses using the wider ranging thinned stand data. The deviates for the test data had a standard deviation of 51.51.

The simplified base model was a satisfactory predictor. Although the estimators were inefficient because the serial correlation assumption had been violated, they were unbiased.

Periodic annual increment, unconditioned

For the evaluation of the unconditioned periodic annual increment models the allometric constant m was set initially to zero and the simple single parameter submodel for A_0 was used. Various linear and quadratic submodels for n and p were evaluated, but the best was Equation V.13 shown below with the standard errors of each parameter estimate immediately below the estimate. Appendix 4.3 summarises some of the models fitted including other submodels for the parameters p and A_0.

$$m = 0.0 \tag{V.13}$$

$$p = \underset{(0.00052)}{0.01895}$$

$$n = \underset{(0.62)}{20.59} + \underset{(0.0200)}{0.1003}\, Y_{10}$$

$$A_0 = 10.0 \exp\,(\underset{(0.000370)}{-0.002084}\, Y_{10})$$

Testing the model against independent test data gave slightly poorer results as can be seen in Table V.2. The deviates were again serially correlated and, unlike the conditioned model, the deviates were heterogeneous by plots. The model was a poorer predictor as the standard deviation of the deviates was 57.53 compared with 51.51, and more cells were significantly different from zero.

There were two possible explanations for the relative inefficiency of the unconditioned model. Firstly, it was possible that the increase in efficiency through the addition of the extra parameter was offset by the decrease in asymptotic efficiency of Bard's program because there were more parameters to be estimated. Secondly, although for linear models the residual sums of squares for an unconditioned model is lower than for a parallel conditioned model (Theil, 1971; Goldberger, 1964), this does not necessarily hold for nonlinear models but depends on the model structure. If the secondary structures for p and A_0 from Equation V.12 are substituted in Equation IV.11 and this equation is reformulated in terms of n rather than Y, then a complex structure relating n to Y_{10} results. The simple linear structure in Equation V.13 is only a crude proxy for this complex structure. If all models had been linear then the structure implied by the unconditioned model would have been a superset of the conditioned model and not a crude proxy.

As the conditioned model was simpler, was a better predictor, and satisfied the assumptions of the analysis as well or better, it was preferred to the unconditioned model.

Derivative

The earlier analysis of the allometric constant r in the general model had shown that r could be taken as 1.0. That analysis and others had shown that the simple Mitscherlich form with m=0.0 was also acceptable. However in the earlier analysis of the derivative the estimates for n and p were biased because the way the data were used assumed that the periodic annual increment and instantaneous increment were the same. Also, the data used were restricted.

Unbiased estimates of the parameters of the simple Mitscherlich form (m=0.0, r=1.0), but not the general form, could be obtained from the full data set by using a Taylor's series expansion of the function (Ferguson and Miles pers.comm.). If yield at age A is Y_A and yield i years later at age (A+i) is Y_{A+i}, and if

$$\frac{dY}{dA} = n - pY \qquad (V.14)$$

the simple Mitscherlich form, then using a Taylor's series expansion

$$Y_{A+i} = Y_A + \frac{i\,dY}{dA} + \frac{i^2}{2}\frac{d^2Y}{dA^2} + \frac{i^3}{6}\frac{d^3Y}{dA^3} + \frac{i^4}{24}\frac{d^4Y}{dA^4} + \dots \qquad (V.15)$$

which can be reformulated in terms of periodic annual increment (Pai)

$$\text{Pai} = \frac{Y_{A+i} - Y_A}{i} = \frac{dY}{dA} + \frac{i}{2}\frac{d^2Y}{dA^2} + \frac{i^2}{6}\frac{d^3Y}{dA^3} + \frac{i^3}{24}\frac{d^4Y}{dA^4} + \dots \qquad (V.16)$$

or

$$\text{Pai} = \frac{dY}{dA}\left\{1 - \frac{i}{2}p + \frac{i^2}{6}p^2 - \frac{i^3}{24}p^3 + \dots\right\} \qquad (V.17)$$

and therefore

$$\frac{\text{Pai}\ \ i\,p}{(1-\exp(-i\,p))} = n - pY \qquad (V.18)$$

Ordinary least squares linear regression theory cannot be used to estimate Equation V.18 even though the right hand side of the equation is linear, because the dependent variable is itself a function of the parameter p. Bard's (1967) program was modified to fit the model iteratively, successive iterations using for the dependent variable the estimate of p from the preceding iteration, each iteration itself iteratively fitting the linear right hand side of the equation. The initial starting values were from the base model. The convergence criteria were those of Marquardt (1963) also used to determine the convergence of each iteration. The technique was in essence an iterative, iterative fit of an apparently linear model.

Equation V.18 was then fitted to the 208 observations using the various secondary model forms suggested earlier, Equations V.2 and V.3. Gallant's test could not be used to compare models as the dependent variable varied with the estimate of p. If Equation V.2 was used then the quadratic term in Y_{10} in the submodel for n was not significantly different from zero using a t test, and the reduced model below with all parameters significantly different from zero was accepted.

$$n = \underset{(1.19)}{21.55} + \underset{(0.00803)}{0.09545}\, Y_{10} \qquad (V.19)$$

$$p = \underset{(0.00153)}{0.01794}$$

If equation V.3 was used then the addition of the quadratic term in the submodel for p was not significant, in that the estimated parameter was not significantly different from zero using a t test.

$$n = \underset{(0.92)}{34.73} \qquad (V.20)$$

$$p = \underset{(0.00312)}{0.03762} - \underset{(0.0000142)}{0.0001354}\, Y_{10}$$

The only way that Equations V.19 and V.20 could be compared was as predictors of the test data because the dependent variables were different. For Equation V.19 the deviates had a standard deviation of 60.06, considerably lower than the 83.76 for Equation V.20 which was then rejected. Table V.2 shows that Equation V.19 was poorer than the conditioned periodic annual increment model even though it satisfied the assumptions of the analysis nearly as well.

The derivative model was therefore rejected and the conditioned periodic annual increment model was preferred.

Yield, conditioned

Various submodels for the parameters m, p and A_0 were then evaluated for the conditioned yield model using the 228 observations. The submodels were variants of the original base model and again reduced in complexity. The allometric constant m was again not significantly different from zero, but in this case the parameter p reduced to a constant rather than a linear function of Y_{10}. The submodel for A_0 remained the two parameter exponential form. Appendix 4.4 summarises the analyses. For Equation V.21 the standard error of each parameter estimate is shown below the estimate.

$$m = 0.0 \qquad \text{(V.21)}$$

$$p = \underset{(0.00154)}{0.01865}$$

$$A_0 = \underset{(0.095)}{9.384} \exp(\underset{(0.000146)}{-0.003334}\ Y_{10})$$

When the estimates of the parameters for this model were compared with the comparable conditioned periodic annual increment model (Model 8, Table V.1), the standard errors of the parameter estimates were consistently larger for the yield model (0.00154 cf. 0.00052, 0.095 cf. 0.081, 0.000146 cf. 0.000085). Periodic annual increment therefore provided more efficient estimators than the equivalent yield model.

As can be seen from Table V.2 the conditioned yield model was as good a predictor, paralleling the conditioned periodic annual increment model. However the assumptions of the analysis were consistently violated, the deviates being heterogeneous, leptokurtic and serially correlated. The heterogeneity was probably associated with the leptokurtosis as Acton (1959) has pointed out that Bartlett's test (like the other tests of homogeneity) is sensitive to nonnormality. This cast doubt on the use of Gallant's test to differentiate between models and Equation V.21 may not in fact be the best conditioned yield model. The Durbin-Watson $\underline{d}$ statistic was significant and considerably lower than for the conditioned periodic annual increment model, and unlike Equation V.12 the $\underline{d}$ statistic for the developmental data was both lower and significant.

The conditioned yield model was not preferred to the conditioned periodic annual increment model because the assumptions of the analysis had been violated consistently and because the estimates of the parameters were less efficient, even though the model was as satisfactory as a predictor.

Yield, two stage

For the two stage analysis the model was fitted to each of the plots in turn and then, in a second stage, the parameters from the first stage were estimated as functions of site potential. The technique has the advantage of making full use of the long term trend data available but there were practical limitations because there were relatively few observations for the first stage analysis. Both conditioned and unconditioned models were evaluated for completeness, but because of the paucity of observations for the first stage analysis the conditioned model with one fewer parameter was thought likely to be superior.

When the conditioned model, Equation V.1, was fitted to the data the standard errors of the parameter estimates were all high and for each plot none of the estimates of p, A_0 or m were significantly different from zero when a t test was used. This indicated a large degree of overfitting that could only be avoided by increasing the data, which was impossible, or by reducing the number of parameters to be estimated. Setting either p or A_0 to zero was biologically unsound and there was no a priori reason for setting them to particular values. The only way that the model could be simplified was to set m to zero, the Mitscherlich form used earlier. When this reduced model was fitted to the data all parameter estimates were significantly different from zero showing that the simpler model was more satisfactory.

When the unconditioned model was fitted to the data (Equation IV.9 with $c_1 = -1$ and $c_2 = A_0$), reduction to the simple Mitscherlich form was again necessary to reduce the standard errors of the parameter estimates so that the parameter estimates were significantly different from zero.

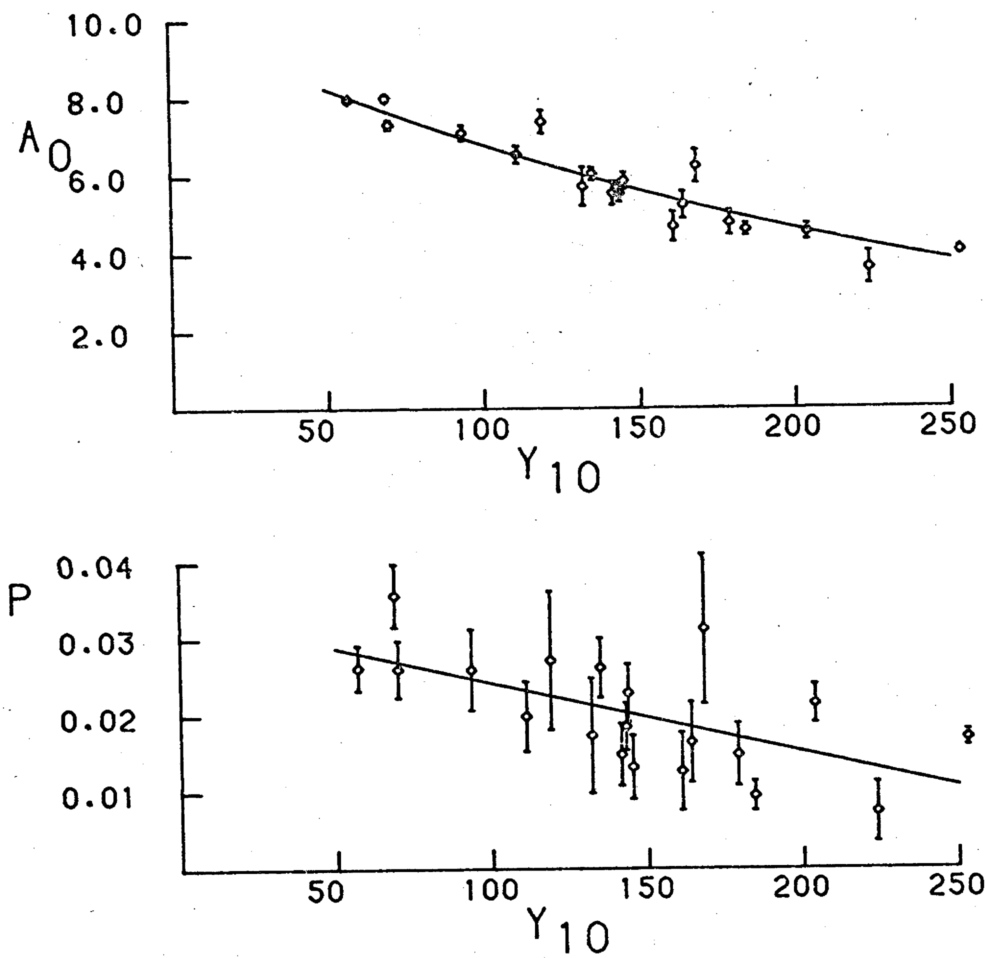
The second stage models were then developed using both linear and nonlinear model structures, the regressions being weighted by the estimated variance of each parameter estimate. For A_0 the conditioned exponential decay model proved to be the superior estimator, consistently better than any of the other forms in Equation V.5. For p the simple linear form was the best for both conditioned and unconditioned models and for n the constant could not be improved upon. Equations V.22 and V.23 were the best models for the conditioned and unconditioned models respectively. For the conditioned model the model and data are graphed in Figure V.1.

$$p = \underset{(0.00410)}{0.03334} - \underset{(0.2677\ 10^{-4})}{0.9070\ 10^{-4}}\ Y_{10} \tag{V.22}$$

$$A_0 = 10.0 \exp(\underset{(0.000327)}{-0.003823}\ Y_{10})$$

and

Bertalanffy two stage analysis
First stage estimates and 95% confidence limits



$$n = \underset{(5.34)}{37.36} \qquad (V.23)$$

$$p = \underset{(0.00431)}{0.03710} - \underset{(0.0281\ 10^{-3})}{0.1042\ 10^{-3}}\ Y_{10}$$

$$A_0 = 10.0\ \exp(\underset{(0.000313)}{-0.003508}\ Y_{10})$$

The two models were then fitted to the independent test data as yield predictors (Table V.2). The conditioned model was inferior to the conditioned periodic annual increment model fitted to the pooled data (standard deviation of the deviates 66.10 cf. 51.51, and more age/site potential cells significantly different from zero). The unconditioned model, Equation V.23, was a very poor model (standard deviation of the deviates of 113.72) parallelling the earlier analysis of the pooled data. The marked reduction in efficiency of even Equation V.22 was attributed to attempting to estimate too many parameters from too few data in the first stage analysis. The two stage OLS analysis was rejected as it was inferior to the conditioned periodic annual increment model developed on the pooled data.

Summary

1 The allometric constant r could be set to 1.0.

2 The allometric constant m could be taken as 0.0.

3 The conditioned model was a superior predictor to the unconditioned model, in part because the gain in asymptotic efficiency of parameter estimation by the simplification in the model structure offset the decrease in efficiency implied by the conditioning, but also because the conditioned structure was not a simple linear contraction of the unconditioned structure.

4 Periodic annual increment was more satisfactory than yield as the dependent variable because it better satisfied the assumptions of the analysis and was as satisfactory as a predictor.

5 Single stage analysis on the pooled data was superior to the two stage analysis using the individual plot trends.

JOHNSON-SCHUMACHER

Linear models

The Johnson-Schumacher model is one of the few models considered that can readily be fitted, albeit in modified form, by multiple linear regression analysis. This form was used by a number of workers including Schumacher (1939) and Clutter (1963). The equation can be formulated as

$$\ln(Y) = b_0 + f(A) + f(Y_{10}) \qquad (V.24)$$

where

$$f(A) = b_1 / A \qquad (V.25)$$

$$f(Y_{10}) = b_2 Y_{10} \qquad (V.26)$$

and where

Y = yield,

Y_{10} = site potential, yield at age 10,

A = age, and,

b_0, b_1 and b_2 are the parameters to be estimated.

Equation V.24 was formulated in this way because the submodels, Equations V.25 and V.26 as used by Schumacher and Clutter, are not wholly satisfactory. For practical reasons the estimated yield at age ten should be within the confidence limits of the estimate of site potential and this is clearly unlikely if Equation V.26 is used. Of the alternatives suggested earlier, Equations V.6a-V.6e, the logarithmic form was logically superior to the polynomial or reciprocal forms with the coefficient of $(\ln(Y_{10}))$ theoretically equal to 1.0; however all were evaluated. Bailey and Clutter (1974) suggested replacing age by a power function which was unsuitable for multiple linear regression analysis. As a simple linear approximation, Equation V.25 was also formulated as a polynomial and as a logarithmic function of age.

Analysis of these linear models showed that the reciprocal of age was considerably superior to the polynomial or logarithmic forms, and that the logarithm of site potential was superior to the other forms, a not unexpected result. The results are summarised in Appendix 4.5a.

$$\ln(Y) = \underset{(0.095)}{4.740} - \underset{(0.30)}{25.64/A} + \underset{(0.0189)}{0.5400} \ln(Y_{10}) \qquad (V.27)$$

The addition of an interaction term was significant:

$$\ln(Y) = \underset{(0.127)}{6.509} - \underset{(3.12)}{61.26/A} + \underset{(0.0349)}{0.1793} \ln(Y_{10}) + \underset{(0.270)}{7.270} \ln(Y_{10})/A \qquad (V.28)$$

but this simple addition to the linear model changed a relatively simple yield model into a considerably more complex one:

$$Y = b_0 \exp(-b_1/A)\,(Y_{10})^{b_2} \exp(b_3 \ln(Y_{10})/A) \qquad (V.29)$$

compared with

$$Y = b_0 \exp(-b_1/A)\,(Y_{10})^{b_2} \qquad (V.30)$$

Table V.2 summarises the results when Equation V.27 and V.28 were fitted to the test data. The models were poor predictors and consistently violated the assumptions of the analysis. This was surprising as Clutter (1963) had claimed that the use of the logarithmic transformation would "generally be more compatible with the statistical assumptions customarily made in regression analysis", although he did not show this to be so in practice.

The linear logarithmic model was therefore rejected and the remaining analyses of the Johnson-Schumacher model were on nonlinear forms where the error terms were considered additive rather than multiplicative.

Nonlinear models, unconditioned

When Equations V.27 and V.28 were fitted as a nonlinear yield model (Equations V.29 and V.30) the sum of the squared residuals was lowered by some 18%, reflecting the avoidance of the logarithmic bias. The estimated parameters in Equation V.31 were all significantly different from those in Equation V.27 using a t test, confirming the belief that the assumption of additivity in the error term had a markedly different effect to the assumption of multiplicativity in the linear logarithmic form.

$$Y = \underset{(0.078)}{\exp(5.200)}\ \underset{(0.43)}{\exp(-27.23/A)}\ Y_{10}^{\underset{(0.0150)}{0.4595}} \qquad (V.31)$$

For all models evaluated yield approaches zero as age approaches zero, that is A_0 approaches zero. This seemed unduly restrictive as the analysis of the Bertalanffy model had indicated that A_0 was site dependent. The addition of the interaction term in Equation V.28 was thought to be a proxy for A_0 so a number of alternative models were evaluated with different nonlinear structures.

Table V.3 shows the results for some of the models fitted, the complete set being summarised in Appendix 4.5b. Model 20 in Table V.3, the nonlinear equivalent to Equation V.28, was surprisingly not significantly better than Model 5, the nonlinear equivalent of Equation V.27, and this could only be attributed to the change in error structure. Model 24, Equation V.32, where age was replaced by a linear function in Y_{10} plus age, was the best model, being considerably better than the form (Model 26) where the addition of parameters resulted in the residual sum squares inflating as misspecification affected the asymptotic efficiency of the program. The quadratic term used by Ferguson and Leech (1976b), Model 36 Table V.3, was also not significantly better. Replacing A_0 by a power function (Model 33, Equation V.33) provided an efficient model, but the model could not be compared with Model 23 because

Table V.3

Johnson-Schumacher; nonlinear.

$Y=b_1 \exp(-b_2/Z_1)\, Y_{10}^{b_3}\, Z_2$			Number of parameters	Residual sum squares	Compared with	
No	Z_1	Z_2			Model	Significance
5	A	1.0	3	491 833		
20	A	$\exp(-b_4 \mathrm{Ln}(Y_{10})/A)$	4	481 892	5	NS
23	$A-a_0$	1.0	4	432 373	5	Sig
24	$A-(a_0+a_1 Y_{10})$	1.0	5	410 256	23	Sig
26	$A-a_0 \exp(-a_1 Y_{10})$	1.0	5	432 808*	23	NS
33	A^{a_1}	1.0	4	428 218	5	Sig
34	$(A-a_0)^{a_1}$	1.0	5	418 181	23	NS
36	A	$\exp(-b_4/A^2)$	4	484 160	5	NS

* Note that misspecification has reduced the efficiency of the estimate.

the extra parameter in Model 23 was not a simple additive increase in structural complexity.

$$Y = \exp(\underset{(0.106)}{5.706}) \exp(-\underset{(1.21)}{35.16}/(A + \underset{(0.2478)}{0.7253} + \underset{(0.00231)}{0.01556}\, Y_{10}))\, Y_{10}^{\underset{(0.0139)}{0.3867}} \qquad (V.32)$$

$$Y = \exp(\underset{(0.152)}{5.795}) \exp(-\underset{(1.12)}{13.33}/A^{\underset{(0.0499)}{0.6392}})\, Y_{10}^{\underset{(0.0141)}{0.4580}} \qquad (V.33)$$

Equation V.32 was not favoured because the estimated parameters are such that estimated value of A_0 was negative, which is biologically unsound. The age scaling implicit in Equation V.33 provided approximately equally efficient estimators with the more satisfactory biological inference that A_0 is zero, and as well the age of culmination of current annual increment is not fixed at half the age of culmination of mean annual increment. However as the model will generally be used only from age 10 both equations could have been satisfactory in practice, so both were tested using the independent test data.

Both equations were less satisfactory than the Bertalanffy yield model as predictors and both violated the assumptions of the analysis, although the nonlinear form was better than the transformed linear form in this regard as can be seen in Table V.2.

<u>Nonlinear models, conditioned</u>

These unconditioned models could hardly be expected to be satisfactory predictors as the estimated yield at age 10 was proportional to a power function in Y_{10}. Within the range of the data the error in estimated yield at age 10 for Equation V.31 varied from 72% to -23%, in a consistent manner, with the error being zero near the mean of the data. This was hardly satisfactory for practical use.

Conditioning the model reduced by two the number of parameters to be estimated; thus equations V.32 and V.33 became

$$Y = Y_{10}\left\{\exp\ b_0\left[1/(A-Z_1)^{Z_2} - 1/(10-Z_1)^{Z_2}\right]\right\} \qquad (V.34)$$

where for Equation V.32

$$Z_1 = b_1 + b_2\ Y_{10} \qquad (V.35)$$

$$Z_2 = 1.0$$

and for Equation V.33

$$Z_1 = 0.0 \qquad (V.36)$$

$$Z_2 = b_1$$

Other alternative structures for Z_1 evaluated included the quadratic and cubic terms. When these models were evaluated the best model was

$$b_0 = \underset{(1.64)}{-37.73} \qquad (V.37)$$

$$Z_1 = \underset{(0.3568)}{0.8826} - \underset{(0.00313)}{0.04339}\ Y_{10} - \underset{(0.0922\ 10^{-4})}{0.6479\ 10^{-4}}\ Y_{10}^{\ 2}$$

$$Z_2 = 1.0$$

Allowing Z_2 to float was not significantly better than fixing the parameter at 1.0, regardless of whether a linear or quadratic structure for Z_1 was used. This model was still unsatisfactory, as Z_1, which in reality is an estimate of A_0, was still negative for the range of the data.

When evaluated as a predictor, Table V.2, Equation V.37 proved to be the best of the Johnson-Schumacher forms, marginally poorer than the Bertalanffy model. The equation violated the assumptions of the analysis and was less satisfactory than the Bertalanffy conditioned yield model.

Summary

The analysis of the Johnson-Schumacher model form was in general unsatisfactory. The advantage that Clutter (1963) claimed for the logarithm of yield over yield as the dependent variable, that it would better satisfy the assumptions of the analysis, was not borne out in practice. Indeed neither dependent variable was satisfactory. The linear model was an unsatisfactory predictor being consistently poorer than the conditioned Bertalanffy yield model. When the model was fitted nonlinearly both the Johnson (1935) form and the power form used by Bailey and Clutter (1974) were equally efficient estimators, but the power form was preferred for biological reasons. This nonlinear model was no better a predictor than the linear model, and was unsatisfactory as the estimated yield at age 10 was in error by more than the likely confidence limits of the estimate of site potential at both high and low site potential levels. When the model was conditioned through Y_{10} at age 10 the exponential power reverted to 1.0 and the best model was the Johnson form. This was the best predictor of the models tested but was still poorer than the conditioned Bertalanffy model.

The Johnson-Schumacher form was rejected for further analysis because it was an inferior predictor, violated the assumptions of the analysis, and, as developed, was biologically untenable at early ages.

BEDNARZ

For the evaluation of the Bednarz model, Equation IV.26 was reformulated to include A_0 and to pass through age 10 values.

$$Y = Y_{10} \left\{ \frac{\ln(a(A-A_0)^b+1)}{\ln(a(10-A_0)^b+1)} \right\} \tag{V.38}$$

where

Y = yield,

Y_{10} = site potential, yield at age 10,

A = age,

A_0 = the age at which volume growth commences, and,

a and b are the two parameters to be estimated.

Various submodels for A_0, a and b were evaluated and the most satisfactory model was Equation V.39 below. The other models are summarised in Appendix 4.6a.

$$a = \underset{(0.00514)}{0.02714} \tag{V.39}$$

$$b = \underset{(0.202)}{1.219}$$

$$A_0 = \underset{(0.433)}{8.938} \exp(-\underset{(0.001028)}{0.004152}\, Y_{10})$$

The equation was as satisfactory a predictor as the conditioned Bertalanffy yield model (Table V.2) as were the estimates.

Because the yield model was an efficient predictor the Bednarz model was also evaluated as a periodic annual increment model using the difference equation (Equation V.10). The results are summarised in Appendix 4.6b where it can be seen that Equation V.40 was the best model.

$$a = \underset{(0.00144)}{0.03007} \qquad (V.40)$$

$$b = \underset{(0.0536)}{1.1436}$$

$$A_0 = \underset{(0.162)}{8.914} \exp(\underset{(0.000246)}{-0.003473}\ Y_{10})$$

This structure was the same as the structure of the yield model. The parameters a and b were independent of site potential confirming Bednarz's results for height. The exponential decay model for A_0 provided the best estimators but the submodel was not conditioned so that at age ten A_0 was 10.0. When the periodic annual increment model was evaluated (Table V.2) the results paralleled the Bertalanffy results very closely.

The analysis showed that the Bednarz model provided a satisfactory prediction model, as good as the Bertalanffy but not better. The model could have been used for subsequent analyses but was not preferred for five relatively minor reasons.

1 To achieve a satisfactory prediction model four parameters were needed compared with three for the Bertalanffy model.

2 The submodel for A_0 was inadequate at extremely low values of site potential as Y_{10} approaches zero.

3 The Bertalanffy model has a more coherent biological basis to its structure and it was felt that extrapolation of the Bertalanffy model might be marginally less hazardous.

4 Although for the Bertalanffy model the allometric constant m was zero, inferring that increment culminates at age A_0, the addition of this parameter in future analyses with better data could approximate the more biologically acceptable sigmoidal form. The Bednarz model cannot readily be extended to allow this sigmoidal form.

5 The Bednarz model does not reach an asymptotic maximum yield.

OTHER MODEL FORMS

Gompertz

The evaluation of the Bertalanffy models indicated that the allometric constant could be taken as zero. This cast doubt on the usefulness of the Gompertz model, which is claimed as the limiting form of the Chapman-Richards equation (similar to the second level Bertalanffy) as m approaches 1.0 (Richards, 1959; Pienaar, 1966).

The Gompertz model was fitted as a conditioned yield model with various polynomial forms for the parameters A_i (the age at which increment culminates) and b, the parameter a in Equation IV.17 being conditioned out of the model. The models fitted are summarised in Appendix 4.7, the best of them being.

$$Y = Y_{10} \left\{ \frac{\exp(\exp(-b(A-A_i)))}{\exp(\exp(-b(10-A_i)))} \right\} \qquad (V.41)$$

where

$$A_i = \underset{(0.30)}{19.47}$$

$$b = \underset{(0.0035)}{0.1161} - \underset{(0.0000257)}{0.0003779}\, Y_{10} + \underset{(0.0705\ 10^{-6})}{0.6817\ 10^{-6}}\, Y_{10}$$

and where

Y = yield,

Y_{10} = site potential, yield at age 10,

A = age,

A_i = age at culmination of volume growth, and with

b were the two parameters estimated.

When this equation was fitted to the independent test data (Table V.2) it was inferior to the Bertalanffy model as a predictor, and the assumptions of the analysis were still violated. Because it was so inferior the Thomasius (1964) model derived from the Gompertz was not evaluated and the line of development was rejected.

Polynomial

The polynomial model form was rejected in Chapter IV (Equation IV.1) because it lacks any biological basis to its structure and because, if used, it was likely to behave erratically as a predictor at the extremities of the data.

The polynomial was evaluated for completeness and to try to provide further quantitative justification for its rejection. Stepwise regression (Draper and Smith, 1966; Efroymson, 1962) was used initially rather than combinatorial screening (Grosenbaugh, 1967), or regression by leaps and bounds (Furnival and Wilson, 1974), because it was simpler and less wasteful of computing resources.

An unconditioned yield model was fitted to a sixth order polynomial in age interacting with a fourth order polynomial in site potential. A second simpler model used a fourth order in age interacting with a second order in site potential. Both reduced to seven parameters but the reduced model with the lower range of powers explained slightly more (0.013%) of the variation than the more complex model, emphasizing the warnings of Draper and Smith (1966) and Grosenbaugh (1967) concerning the disadvantages of stepwise regression. The model with the lower powers was accepted.

$$Y = \underset{}{-212.26} + \underset{(1.74)}{22.44\,A} - \underset{(0.001062)}{0.003895\,A^3} + \underset{(0.00139)}{0.01083\,A\,Y_{10}}$$

$$\underset{(0.000772)}{-0.001914\,A^2\,Y_{10}} + \underset{(0.00001533)}{0.00003694\,A^3\,Y_{10}}$$

$$\underset{(0.1876\ 10^{-7})}{-0.4346\ 10^{-7}\,A^3\,Y_{10}^{\,2}} \qquad (V.42)$$

When fitted to the independent test data (Table V.2) the equation proved to be one of the poorest models evaluated in spite of its having more terms (seven, compared with three for the second level Bertalanffy periodic annual increment model). Equation V.42 has a maximum yield for SQ VII at age 45, outside the range of the data but within the region of interest, other site qualities having maxima soon after age 50. These maxima are biologically unsatisfactory as yield is expected to increase with age, approaching a maximum asymptotically, but never reaching it. Evaluation using either combinatorial screening or leaps and bounds was not carried out because it was unlikely to provide the marked improvement necessary for the polynomial to be useful as a predictor. The polynomial was rejected.

Lewis's yield table

The growth prediction technique currently used in South Australia, embodied in the Woods and Forests Department Yield Regulation system, is the latest version of the graphical yield table developed by Lewis (Lewis, Keeves and Leech, 1976). At the time it was developed it was believed to fit thinned stands of all ages up to age 50, but was believed to overestimate for unthinned stands past age 35. In practice this restriction in its applicability is of little significance as very few stands are left unthinned after age 30. It was of interest to compare the fit of this subjectively defined model with the other models developed in this chapter.

The test data used in this chapter are not truly independent for this analysis as they were included in the data Lewis used, but their effect would have been swamped by the other plots as Lewis used all 313 plots available. Table V.2 shows that the graphical yield table was nearly as good a predictor as the conditioned Bertalanffy periodic annual increment model (standard deviation of the deviates 52.25 compared with 51.51) and better than the Johnson-Schumacher model or the polynomial.

Careful development of a graphical yield table by the directing curve technique as used by Lewis, and as used before computers made the current statistical analyses possible, provides an efficient predictor nearly as good as the best model developed and better than most of the other models. Although open to bias and although confidence limits cannot be calculated, the technique obviously can provide efficient prediction models. Analyses of growth using modern statistical and computational techniques must be very carefully carried out if they are to improve on the techniques used by earlier generations of forest managers.

SUMMARY

The exploratory analyses of the unthinned data confirmed the conclusion of Chapter IV that the second level Bertalanffy model was the best model form. The structure is simple, although nonlinear, and the parameters were more readily interpreted biologically than the other models. The allometric constant m was not significantly different from 0.0, the Mitscherlich form, probably because there were insufficient data at early ages rather than from biological correctness. Despite this limitation the model is satisfactory past age 10, the age that will in practice be used as a minimum. As further data become available perhaps the Mitscherlich form may be replaced by the second level form.

The model was best fitted as a conditioned periodic annual increment model in spite of the theoretical disadvantages induced by the conditioning. This was attributed to the trade off between the statistical disadvantages of the conditioning and the analytical advantages in asymptotic efficiency for the simpler model, and to the nonlinear structure. As foreshadowed, periodic annual increment was better than yield in satisfying the assumptions of the analysis, but even so was unsatisfactory. It was hoped that better model specification including thinning parameters would reduce the serial correlation to an acceptable level.

The Johnson-Schumacher model in linear form was unsatisfactory and, although a nonlinear form was developed that was a relatively satisfactory predictor, it was unsatisfactory as a predictor of early age growth. The Bednarz model was developed easily and was a relatively good predictor, but was rejected because it lacked a coherent biological basis and could not be used at early ages. The Gompertz model was considerably poorer than the Bertalanffy. The polynomial was an inefficient predictor and as it lacked any biological structure it was rejected.

Lewis's yield table was interesting in that it was a marginally poorer predictor than the best models developed, but was better than many of the other models that have been used in previous yield prediction work. Although it is impossible to compute confidence limits for a subjectively defined graphical yield table it clearly demonstrated that careful subjective analysis can provide a good prediction model.

The conditioned Bertalanffy periodic annual increment model was accepted as a suitable model for further analysis on the combined data from thinned and unthinned stands.

VI MODELS BASED ON DATA FROM ALL STANDS

VI MODELS BASED ON DATA FROM ALL STANDS

In Chapter V the conditioned Bertalanffy periodic annual increment model, developed on the pooled data as a difference equation using the yield model, was shown to be the most satisfactory estimator for unthinned stands and one of the best predictors. The objective of this chapter was to extend the model to investigate the effect of thinning and other stand variables such as form and soil type. This model was developed using data from the lower south-east of the state. Its extension to other areas in South Australia and its applicability to second rotation stands were then investigated.

THINNING

In Chapter III three factors which describe a thinning were defined: thinning type, thinning grade and thinning interval. However these variables are not conveniently incorporated in the yield model.

Thinning changes the level of competition in the stand and it was believed that a thinning affects increment some years after that thinning. The use of competition level as a variable does not take into account how the stand reached that level. Buckman (1962) preferred to use a relative measure, the proportion of the forest cut, rather than the absolute level of competition. The proportion is a measure of the effect on the proportion of the site that is occupied immediately following a thinning and therefore represents thinning shock.

Thinning type was ignored in this study as it varied little in the data base, all thinnings being predominantly from below. Thinning interval as such was irrelevant to this study as it is embodied in the change in competition level with age.

Thinning was therefore investigated in terms of two variables:

1 competition level, and,

2 thinning shock.

Competition level

The effect on increment of changing the level of competition was considered by Langsaeter (1941) and Moller (1954a, 1954b) in a qualitative manner. Langsaeter's model, Figure VI.1, can be interpreted in the following way.

Stage I The free growth stage where the individual trees have no influence on their neighbours, volume increment is therefore proportional to the volume of trees standing.

Stage II In this transitional stage the trees are beginning to crowd one another increasingly, but the site is still not fully occupied.

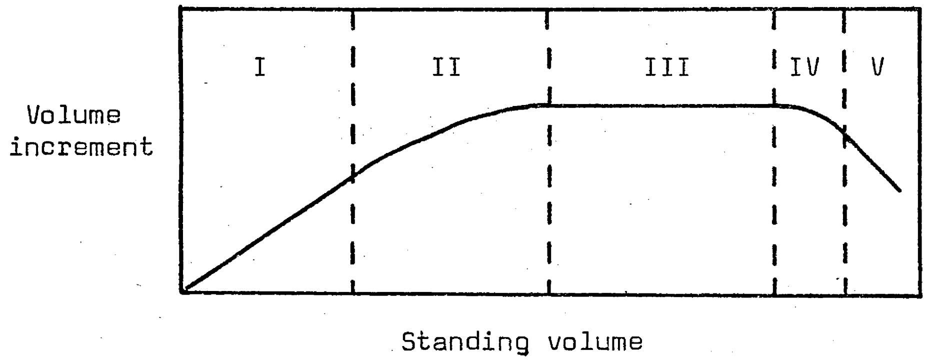
Stage III This broad band denotes the level of full site occupancy in which growth is almost independent of stand density, provided all the trees are healthy.

Stage IV As stand density increases the trees stagnate and growth rate decreases.

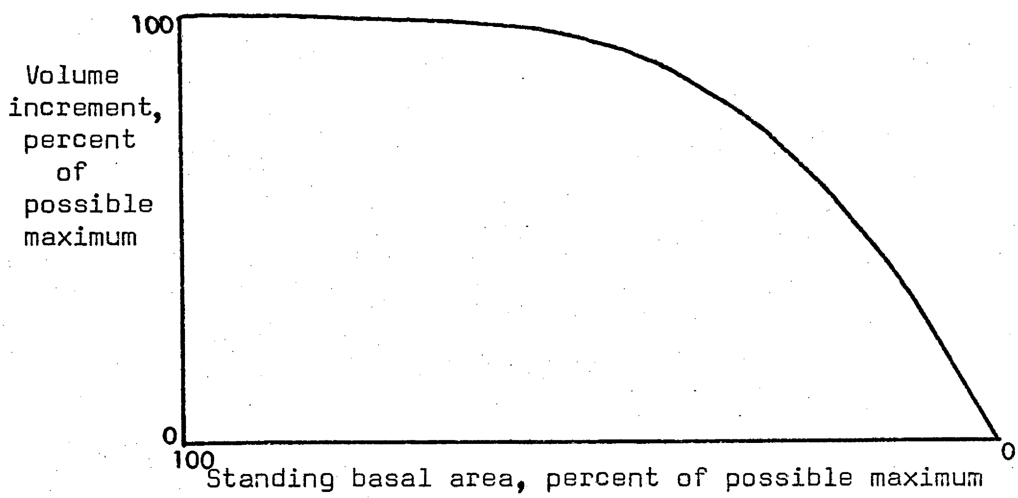
Stage V Competition between trees is so intense that trees become moribund and eventually die.

The most important part of this model for forest management is the shape of stage II and the width of stage III, for South Australian experience suggests that the most economic regime will keep the stand at or about the stage II- stage III boundary. Stage IV is only reached by severely overstocked stands such as those still unthinned at age 30. Moller (1954a, 1954b) used basal area as the competition index and his

Figure VI.1



Relationship between standing volume and volume increment
(Langsaeter, 1941)



Relation between standing basal area and volume increment
(Moller, 1954)

work suggests that Langsaeter's stage III may be quite broad. Moller's work (Figure VI.1) is equivalent to Langsaeter's stages I-III and has been widely cited in the literature. The implication of Moller's model is that increment does not decline as the competition level approaches the maximum the site can sustain but remains near constant. Limited evidence from permanent sample plot trends, notably evidence from adjacent plots X, Y, 304 and 305 reported by Lewis (1962), indicated that Langsaeter's model was more likely to be correct under South Australian conditions.

Two different approaches taking the effect of competition level into account were evaluated. In the first the parameters in the model were reformulated in terms of competition level, and in the second periodic annual increment was corrected by a multiplicative competition level function.

The parameter p in Equation V.1 was the only parameter that could logically be related to the level of competition, for the allometric constant m had already been shown to be close to zero (Chapter V). Moreover A_0 should be independent of changes in the level of competition, because commerical thinning takes place after age A_0. Various models were formulated about the effect of the level of competition on the catabolic destruction rate p.

$$p = p_0 \tag{VI.1a}$$

$$p = p_0 + p_1 D \tag{VI.1b}$$

$$p = p_0 + p_1 D + p_2 D^2 \tag{VI.1c}$$

$$p = p_0 D^{p_1} \tag{VI.1d}$$

$$p = p_0 + p_1 D^{p_2} \tag{VI.1e}$$

$$p = p_0 + \frac{p_1}{1 - p_2 D} \tag{VI.1f}$$

$$p = \frac{p_0}{1 - p_1 D} \tag{VI.1g}$$

where

D = competition level or stand density,

p = the catabolic destruction rate in Equation V.1, and,

p_0, p_1 and p_2 are the parameters to be estimated.

Of the forms the power function (Equation VI.1d) was thought to be biologically the best, but as the others may have been statistically better estimators and predictors they were also evaluated.

Reformulating Equation V.10 as

$$Pai = \left\{(Y_{A+i} - Y_A) / i\right\} Z_1 \qquad \text{(VI.2)}$$

where

Pai = periodic annual increment,

Y_A = yield at age A, and,

i = increment period,

enables Z_1 to be a function of competition level such that it approximates the three middle stages of the Langsaeter model, the likely range of the data. A large number of formulations were possible but consideration of the work of Buckman (1962), Kira <u>et al</u>. (1953), Clutter (1963), Sullivan and Clutter (1972) and Bevege (1972) suggested the following forms would be worthy of evaluation.

$$Z_1 = 1.0 \qquad \text{(VI.3a)}$$

$$Z_1 = b_0 + b_1 D \qquad \text{(VI.3b)}$$

$$Z_1 = b_0 + b_1 D + b_2 D^2 \qquad \text{(VI.3c)}$$

$$Z_1 = b_0 + b_1 / D \qquad \text{(VI.3d)}$$

$$Z_1 = b_0 + b_1 \ln(D) \qquad \text{(VI.3e)}$$

$$Z_1 = b_0 + b_1 \exp(1 / D) \qquad \text{(VI.3f)}$$

$$Z_1 = b_0 + b_1 D^{b_2} \qquad \text{(VI.3g)}$$

where

Z_1 = the function in Equation VI.2,

D = the level of competition or stand density, and,

b_0, b_1 and b_2 are the parameters to be estimated.

The following forms were also evaluated as they exhibit under certain conditions the form of stages II, III and IV in Langsaeter's model.

$$Z_1 = b_0 \; b_1^{(D - b_2)} \qquad (VI.3h)$$

$$Z_1 = b_0 + b_1 / D + b_2 D \qquad (VI.3i)$$

$$Z_1 = b_0 + b_1 D + b_2 D^2 + b_3 D^3 + \ldots + b_n D^n \qquad (VI.3j)$$

If p is replaced by a function in stand density then the maximum asymptotic yield varies with stand density whereas if the function Z_1 is used then the asymptotic maximum yield is independent of stand density. The latter seemed biologically more plausible.

Both number of trees per unit area and standing volume were considered in Chapter II to be satisfactory indices of stand density. Both were evaluated, the models initially using volume because this seemed the better index for this application.

Thinning shock

Immediately after a thinning the site will be less than fully occupied and in this regard less increment will be put onto the standing trees than a stand of the same competition level that was thinned some years earlier.

The effect of thinning shock must be to reduce increment so it was logical to formulate thinning shock models as in Equation VI.2 with Z_1 being a function of thinning shock. The eight measures of thinning shock considered were various combinations of three alternatives:

1 whether the measure should be based on volume or number of trees per unit area,

2 whether the absolute or relative measure should be used, and,

3 whether the effect should be considered to last one year only or whether the effect should last throughout the increment period.

The eight measures of thinning shock were:

$$S = Nt \qquad (VI.4a)$$

$$S = Nt / i \qquad (VI.4b)$$

$$S = Nt / Nb \qquad (VI.4c)$$

$$S = Nt / (Nb\ i) \qquad (VI.4d)$$

$$S = Vt \qquad (VI.4e)$$

$$S = Vt / i \qquad (VI.4f)$$

$$S = Vt / Vb \qquad (VI.4g)$$

$$S = Vt / (Vb\ i) \qquad (VI.4h)$$

where

S = thinning shock,

i = increment period,

Nt = number of trees per hectare removed in thinning,

Nb = number of trees per hectare standing before thinning,

Vt = volume removed in thinning, and,

Vb = standing volume before thinning.

As thinning shock was likely to be masked by the effect of varying competition level it was decided to analyse competition level first. The exact structure to be analysed could then be additive or multiplicative, for example, if competition level was not included in the model then for Equation VI.2.

$$Z_1 = 1 - s_1 S \qquad (VI.5)$$

where

Z_1 = the function in Equation VI.2,

S = thinning shock, and,

s_1 = the parameter to be estimated.

This formulation could also apply if p was a function of competition level. On the other hand if the competition level model was one of the forms in Equation VI.3 such as Equation VI.3c then

$$Z_1 = b_0 + b_1 D + b_2 D^2 + b_3 S \qquad \text{(VI.6a)}$$

$$Z_1 = b_0 + b_1 D + b_2 D^2 + b_3 S + b_4 D S \qquad \text{(VI.6b)}$$

$$Z_1 = b_0 + b_1 D + b_2 D^2 + b_3 S + b_4 D S + b_5 D^2 S \qquad \text{(VI.6c)}$$

were all possible forms, or even

$$Z_1 = (b_0 + b_1 D + b_2 D^2)(1 + b_3 S) \qquad \text{(VI.6d)}$$

where

Z_1 = the function Z_1 in Equation VI.2,

D = the level of competition,

S = thinning shock, and,

$b_0, b_1, b_2, \ldots b_5$ are the parameters to be estimated.

All these were possible models, although the multiplicative model, Equation VI.6d, was thought likely to be poorer than the other structures.

Data

The data used in Chapter V were inappropriate for the development of periodic annual increment models incorporating the effect of thinning. All the available thinned and unthinned data from the lower south-east of South Australia were pooled and randomly allocated by plots such that approximately 60% were used as developmental data, 40% as independent test data. Appendix 1.4 summarises the two data sets and the selection technique.

There were 969 observations in the developmental data and 669 in the independent test data.

Analysis of competition level

The various models incorporating the level of competition were fitted to the data using standing volume as the stand density index. Models 1–14, Appendix 5.1a, summarises the results for Equation VI.1 where the competition level was included in the formulation of p. Models 20–30, Appendix 5.1b, summarises the models for the other developmental line where competition level was included in a correction factor to periodic annual increment.

Of the models where p was reformulated in terms of competition level, Model 4, Table VI.1 and Appendix 5.1a, proved to be the best model. This structure implies that there is no interaction between the level of competition and site potential, for Model 5 with two extra interaction parameters was not significantly better. On the other hand Model 21, Table VI.1 and Appendix 5.1b, proved to be the best of the models where the competition level was included in a correction factor to periodic increment. The implication of this structure is that the qualitative structure of Langsaeter is best fitted by a simple quadratic model, better than Equation VI.3h, Model 27, which was thought to be more logical as it allows curves to closely approximate Langsaeter's stages II, III and IV; very flat at the peak and tapering rapidly at the extremes.

The correction factor approach consistently explained more of the variation than the models where p was reformulated in terms of competition level. The relatively flat response curve was in part a reflection of the paucity of data from extremely heavily thinned stands.

The flat nature of the curve was confirmed by calculating the mean deviate of a subjectively selected subset of the most heavily thinned plots. The mean deviate for these 39 observations was not significantly different from zero. It was inferred that the model was satisfactory within the range of the data.

Table VI.1

Competition models

Model				Number of parameters	Residual sum squares	Compared with	
Number	Equations V1.2, V1.9, m=0, $A_0=10\exp(-a_1Y_{10})$						
	p	Z_1	Density index			Model	Significance
1	$p=p_0+p_1Y_{10}$	$Z_1=1.0$		3	31548.8		
4	$p=p_0+p_1Y_{10}+p_2D+p_3D^2$	$Z_1=1.0$	Volume	5	29779.7	1	Sig
5	$p=p_0+p_1Y_{10}+p_2D+p_3D^2+p_4DY_{10}+p_5D^2Y_{10}$	$Z_1=1.0$	Volume	7	29521.0	4	NS
21	$p=p_0+p_1Y_{10}$	$Z_1=b_0+b_1D+b_2D^2$	Volume	6	28520.1	1	Sig
34	$p=p_0+p_1Y_{10}+p_2D+p_3D^2$	$Z_1=1.0$	Stocking	5	30455.3		
35	$p=p_0+p_1Y_{10}$	$Z_1=b_0+b_1D+b_2D^2$	Stocking	6	28539.7		

The superiority of the correction factor supported the contention that the asymptotic maximum yield is independent of the level of competition.

For completeness the allometric constant m was evaluated again, but Models 15–17, Appendix 5.1a, and Models 31–33, Appendix 5.1b, showed that the Mitscherlich form with m=0.0 was still satisfactory.

Both p and A_0 were reformulated to include stocking at age 10 (N_{10}) so that initial plantation espacement could be evaluated. A number of models were fitted (including Models 18 and 19, Appendix 5.1a, and Models 34 and 35, Appendix 5.1b) but in no case was the addition of N_{10} significant. Although the inspection of the independent test data in Chapter V had suggested that this variable might be significant, the range of the data was probably still too narrow to enable the variable to be included in the model.

If stand density was not standing volume but number of trees per unit area, the other index suggested in Chapter II, then slightly inferior estimators resulted (Table VI.1). Volume was accepted as the best index of the level of competition.

Equations VI.2, VI.7 and VI.8 provided the best model including competition level.

$$\text{Pai} = \left\{(Y_{A+i} + Y_A) \,/\, i\right\} \quad Z_1 \tag{VI.2}$$

$$Y_A = Y_{10} \left\{\frac{1 - \exp(-p(A-A_0))}{1 - \exp(-p(10-A_0))}\right\} \tag{VI.7}$$

$$p = \underset{(0.002694)}{0.009890} + \underset{(0.1232\ 10^{-4})}{0.4837\ 10^{-4}}\ Y_{10} \qquad \text{(VI.8)}$$

$$A_0 = 10.0 \exp(-\underset{(0.000769)}{0.007668}\ Y_{10})$$

$$Z_1 = \underset{(0.109)}{1.405} + \underset{(0.000273)}{0.001168}\ D - \underset{(0.0272\ 10^{-5})}{0.1174\ 10^{-5}}\ D^2$$

and where

Y_A = yield at age A,

A = age,

Y_{10} = site potential,

i = increment period,

Pai = periodic annual increment,

D = the level of competition or stand density (volume),

p and A_0 are the same parameters as in Equation V.10, and,

Z_1 = the correction factor to periodic annual increment.

When Equation VI.8 was fitted to the independent test data the mean deviate was 0.10 and the standard deviation of the deviates was 5.58. The deviates were homogeneous regardless of how they were partitioned and were normally distributed. The Durbin-Watson d statistic was 1.33 and significant, indicating that misspecification was still a problem. Of the 23 plots with more than 10 observations 3 plots had mean deviates significantly different from zero. Two of these plots are on volcanic soils that have always been thought to have different growth trends compared with the other predominantly sandy soils. It was thought that the inclusion of soil variables would possibly overcome this misspecification. Of the 13 age/site potential cells two had mean deviates significantly different from zero, but these two cells were adjacent and had opposite signs for the mean deviate. As the estimated t values were only just greater than the significance level, the results were considered reasonably satisfactory.

Analysis of thinning shock

Thinning shock was then investigated. Because volume proved to be marginally superior to number of trees as the index of competition level, only the four volume based forms of thinning shock, Equations VI.4e–VI.4h, are reported. Some 17 different models were fitted, Appendix 5.2, using both forms of the competition models. If competition level was included in the function for p then, although the addition of thinning shock was significant for any of the forms tried, the reduction in the residual sums of squares was at most 1.5%; considerably lower than the reduction of 3.5% if the competition level was the multiplicative correction factor. The best of the four measures of thinning shock was Equation VI.4h, a relative measure that assumes thinning shock only lasts one year. Surprisingly Equation VI.6d was superior to the additive structures of Equation VI.6a, VI.6b and VI.6c resulting in the multiplicative model, Equation VI.9 below. It could only be inferred that competition level and thinning shock are best considered as acting separately and independently in correcting the periodic annual increment. The structure of Equation VI.9 was clumsy and inelegant but was logical, as well as being statistically the most efficient estimator.

$$p = \underset{(0.002623)}{0.005075} + \underset{(0.1164\ 10^{-4})}{0.5855\ 10^{-4}}\ Y_{10} \qquad \text{(VI.9)}$$

$$A_0 = 10.0 \exp(\underset{(0.000841)}{-0.009172}\ Y_{10})$$

$$Z_1 = (\underset{(0.114)}{1.700} + \underset{(0.2335\ 10^{-3})}{0.4426\ 10^{-3}}\ D - \underset{(0.3050\ 10^{-6})}{0.7380\ 10^{-6}}\ D^2)\ (1.0 - \underset{(0.0731)}{0.4287}\ S)$$

where

p, A_0 and Z_1 are as for Equation VI.7,

Y_{10} = site potential,

D = competition level (volume), and,

S = thinning shock, (Vt/(Vb i)), relative volume, the effect lasting one year only.

Equation VI.9 was then fitted to the independent test data and the results closely resembled those for Equation VI.8. The deviates had a standard deviation of 5.54 about the mean of 0.10, were homogeneous and were normally distributed. The Durbin-Watson d statistic was again significant at 1.33, indicating misspecification. Of the 23 plots 3 had mean deviates significantly different from zero, but only one age/site potential cell was significantly different from zero. When the equation was evaluated as a yield predictor 5 of the 23 plots and one of the 13 age/site potential cells had mean deviates significantly different from zero, again indicating that there was a correlation with soil type. The mean deviate for all observations was not significantly different from zero. The analysis was still unsatisfactory but was the best to date and there was some hope that the incorporation of soil and form variables would provide an even more satisfactory model.

SOIL AND FORM

There was limited evidence from the data that different soil types had different volume-age trends. The shallower terra rossa soils appeared to have a consistently lower increment at later ages than sandy soils of the same site potential, which in turn appeared to have a lower increment than volcanic soils.

As soil type was defined qualitatively rather than quantitatively it was difficult to formulate testable hypotheses concerning the effect on yield. The soil types could only be grouped arbitrarily into a small number of groups and dummy variables (Johnston, 1963; Cunia, 1973) used to determine whether the arbitrary groupings were significantly different or not. This required only that soil variant parameters be recognised, not that any relationship between soil groups be formulated. Hypothesis tests were then used to test different aggregations of soil types.

Of the parameters in the model the parameter p was the one considered most likely to vary between soil types. Equation VI.10 shows the various alternative formulations possible, where Equation VI.10a is the base model and soil invariant.

$$p = p_0 + p_1 Y_{10} \qquad (VI.10a)$$

$$p = d_0 + p_1 Y_{10} \qquad (VI.10b)$$

$$p = p_0 + d_1 Y_{10} \qquad (VI.10c)$$

$$p = d_0 + d_1 Y_{10} \qquad (VI.10d)$$

$$p = d_0 (p_0 + p_1 Y_{10}) \qquad (VI.10e)$$

where

p = the parameter in Equation VI.7,

Y_{10} = site potential,

p_0 and p_1 are the parameters independent of soil types, and

d_0 and d_1 are parameters, dummy variables, one for each soil type.

In Chapter II six form measures were defined based on alternative concepts of form viz. stand form factor and average stand taper. As a form measure could conceivably affect each of the parameters in the model, it was decided to evaluate form by replacing each of the parameters in turn with a linear function of each of the stand form indices in turn.

$$b = b_0 + b_1 F \qquad (VI.11)$$

where

b = the parameter in the model,

F = one of the six alternative stand form indices, and,

b_0 and b_1 are the parameters to be estimated.

Form is a continuous variable, that was thought might vary between soil types. As it was possible that a form index could be a continuous variable proxy for soil type, the two variables were analysed separately and together.

Data

There are a wide range of soils in the south-east of South Australia ranging from deep aeolian sands to shallow heavy soils over limestone. The soil types for each plot are tabulated in Appendix 1.1 and summarised in Appendix 1.2. Because there were 34 different soil profiles recognised, the developmental and test data were combined to give 1638 observations (Appendix 1.4e). Even then each soil type could not be analysed separately as some had too few observations to allow all parameters to be estimated, so the data were subjectively aggregated into twelve morphological groups together with a miscellaneous group with widely divergent, but poorly represented, soil types. These groups are defined in Table VI.2.

The stand form indices used required an estimate of upper stand height at age 10. This was not available for all the 969 observations in the developmental data used in the competition analyses, as many of the older plots were first measured for basal area, volume and mean dominant height, but not predominant height. The 969 observations were culled to 723 that had estimates of predominant height within three years of age 10, thus enabling estimates of stand form factor and average stand taper at age 10 to be used.

For the analyses evaluating form and soil types together, the 1638 observations were culled in the same way to 1271 observations.

Results

When each parameter in Equation VI.8 was allowed to vary between the different groups of soil types, that is the 7x13 parameter model was fitted to the 1638 observations, Bard's (1967) program failed to converge. After considerable effort it was concluded that it was the program and not the use of the program that was at fault and that it could not be used for dummy variables. No suitable alternative program was available.

The procedure finally adopted was a variant of stagewise regression (Draper and Smith, 1966) and although the best technique available it was recognised that the estimators were not true minimum variance estimators and that the parameter estimates may be biased. Standard statistical tests were not necessarily valid and the model evaluation was somewhat arbitrary.

The 1638 observations were divided by soil group and to each of the 13 data subsets three levels of models were fitted based on Equations VI.2, VI.7 and VI.9:

1 Where 6 of the 7 parameters were fixed at the Equation VI.9 estimates, the other being allowed to float.

2 Similar to 1 above, but where p_0 and one other parameter were allowed to float.

3 Where all 7 parameters were allowed to float.

Level 3 was therefore a true minimum variance estimate. The residual sum squares were then aggregated across all 13 subsets.

For level 1 the residual sum squares when p_0 was allowed to float was 44112.7, lower than when p_1 was allowed to float (44227.6) and lower than the other five models (lowest of these 44541.6), compared with allowing no parameters to float (47628.6). For each individual data subset the model allowing p_0 to float was compared with the fit of the combined model and for 6 of the 13 groups the improvement in fit was considerable and would have been significant if Gallant's test was appropriate. The inference was that Equation VI.10b was better than Equation VI.10d and also better than Equation VI.10a; soil seemed to be a significant variable in the model. The best of the level 2 models was when both p_0 and p_1 were allowed to float (residual sum squares 43646.3) but the gain was minimal compared with allowing just p_0 to float. Level 3, allowing all seven parameters to float, had a combined residual sum squares of 42307.3, a marginal reduction over allowing p_0 alone to float, considering it had 91 parameters rather than 19. Fitting Equation VI.10e provided a residual sum squares of 44207.3, poorer than Equation VI.10b. The analysis was unsatisfactory as Gallant's test was possibly inappropriate, but it was concluded that only the parameter p_0 varied between the 13 soil groups.

The 12 estimates of p_0 (excluding the miscellaneous group) were then compared and ordered, as shown in Table VI.2. For a constant site potential, as p_0 decreases the asymptotic maximum yield increases; thus Table VI.2 infers that for a constant site potential (Y_{10}) a volcanic soil will grow at a faster rate towards a higher asymptotic maximum yield than either a terra rossa soil or a brown soil from Comaum.

These twelve groups were then combined into five groups based on the estimates of p_0 and the standard error of these estimates, and also based on morphological and geographical considerations. Because only one parameter was estimated for each soil group it was possible to estimate p_0 for each of the six soil types in the miscellaneous group and to allocate four of these to the other groups so that a total of seven groups were recognised. Further analyses of individual soil types and of other groupings (for example by depth phase and by Forest Reserve) indicated that these seven groups ought not be divided further.

Table VI.2

Initial grouping of soil types

Group Number	Soil Types	Description	Number of Observations	$\hat{P}_o$	Standard error of estimate
1	V-1	Volcanics	83	0.00030	0.00093
2	WS1	Wandilo sands	160	0.00277	0.00083
3	CS1 CS2 CS3	Yellow sands; Myora and Mt. Gambier	174	0.00317	0.00069
4	MM1 MM2 MM3	Mt. Muir sands	62	0.00327	0.00130
5	YS3 K-1 K-2 K-3	White sands	165	0.00424	0.00090
6	NS1 NS2 NS3	Yellow sands; Penola	260	0.00466	0.00062
7	WN1 WN2	Transitionals between groups 2 and 6	32	0.00591	0.00146
8	MB1 MB2 MB3	Yellow sands; Mt. Burr, Tantanoola, Mt. Gambier	382	0.00634	0.00064
9	YM3 KN1 KN2 KN3	White/yellow sand transitionals	75	0.00855	0.00162
10	TF1	Tantanoola flinty sands	107	0.01119	0.00131
11	TR1	Terra rossa's	41	0.01642	0.00267
12	BS1 BS2	Brown soils from Comaum	30	0.01888	0.00310
13	DY3 MS3 RI1 CV1 CV3 RE1	Miscellanea	67		

The groups and the estimated p_0 values are shown in Table VI.3. Because all the data were included in the analysis it was not possible to test the resultant models against independent test data. The residuals were regressed in turn against second order polynomials in age, site potential, competition level and thinning shock, and the regressions were not significant. The residuals were normally distributed and homogeneous but were still serially correlated at d=1.59. For the plots with more than 10 observations for volume the Durbin-Watson d statistic was tested against the extrapolated approximation of the upper and lower bounds. Of the plots 41% were not significant, 26% were inconclusive (5% for positive and 21% for negative serial correlation) and 33% were significant (2% positive, 31% negative). This supported the inference from the pooled estimate that serial correlation was still a problem, but investigation of these plots with significant negative serial correlation gave no insight into the possible cause.

At this stage, form was evaluated using the conditioned periodic annual increment model with competition and thinning shock included but not soil type. This base model was then fitted to the reduced data set of 723 observations. Seven models were fitted for each of the six stand based form indices, replacing each parameter in Equation VI.9 with a linear function in form index. The results are summarised in Appendix 5.3.

Most of the models were not significantly better than the base model and generally the only significant ones were those including sub-parameters of p. The best estimator was the model including p_0 as a function of average stand taper at age 10, although this was only marginally superior to the relative stand taper or the average stand taper at the start of the increment period. Average stand taper at age 10 was preferred because it is age invariant and was thought to

Table VI.3

Final grouping of soil types

Group	Description	Soil types	Number of observations	$\hat{P}_0$	Standard error of estimate
A	Volcanics	V–1	83	0.00030	0.00093
B	Caroline, Wandilo and Myora sands	CS1 CS2 CS3 WS1 MS3	344	0.00302	0.00052
C	Other sands	NS1 NS2 NS3 MB1 MB2 MB3 YM3 KN1 KN2 KN3 YS3 K–1 K–2 K–3 WN1 WN2 MM1 MM2 MM3 RI1 CV1 CV3	999	0.00539	0.00036
D	Tantanoola flinty sands	TF1	107	0.01119	0.00130
E	Brown soils from Comaum and Terra rossas	BS1 BS2 TR1	71	0.01759	0.00201
F	Yellow sands from Comaum	DY3	22	0.01172	0.00225
G	Rendzinas	RE1	12	0.00534	0.00175

better reflect soil differences than the other indices which vary with age. Even though significantly better, the residual sum squares was only marginally below the critical level for Gallant's test, and the reduction in residual sum squares due to the addition of the parameter was only 1%. The estimated parameters were all significantly different from zero using a t test, although three, including the coefficient of form, were only slightly greater than the critical t value.

The reduction in residual sum squares due to the addition of the stand form parameter seemed low (1%) compared with the reduction by the inclusion of dummy variables for soil type (7%) so the combined effect of soil type and form was then investigated.

To the 1271 observations four models were then fitted:

1 A seven parameter model without form or soil variables, paralleling the thinning shock model Equation VI.9 - residual sum squares 36753.3.

2 An eight parameter model including average stand taper at age 10 in the submodel for p_0 - residual sum squares 36404.4.

3 A thirteen parameter model excluding average stand taper at age 10, but allowing p_0 to float for the seven different soil groups developed previously - residual sum squares 33951.2.

4 A fourteen parameter model developed along parallel lines to the thirteen parameter model but including average stand taper at age 10 in the function of p_0 - residual sum squares 33778.3.

Analysis showed that whether average stand taper at age 10 was included or not the addition of soil varying parameters was significant (assuming linear model theory holds for this stagewise analysis). The addition of average stand taper at age 10 was significant if there were no soil parameters but was just below the critical F value if soil parameters were included. The stand based proxies for the tree variables, form factor and taper, were not included in the model.

The thirteen parameter model with p_0 varying for the seven soil groups was the best model, taper at age 10 being but a weak and unsatisfactory proxy for soil type. The estimated value of p_0 for the volcanic soils was not significantly different from zero but was nevertheless included to avoid reestimating the other parameters for the one soil type, and for consistency.

The model based on the 1638 observations was accepted as the best OLS model.

$$Pai = \left\{(Y_{A+i} + Y_A) / i\right\} Z_1 \qquad \text{(VI.2)}$$

$$Y_A = Y_{10} \left(\frac{1 - \exp(-p(A-A_0))}{1 - \exp(-p(10-A_0))} \right) \qquad \text{(VI.7)}$$

where

$$p = p_0 + \underset{(0.1164\ 10^{-4})}{0.5855\ 10^{-4}}\, Y_{10} \qquad \text{(VI.12)}$$

$p_0 = 0.00030$ (0.00093) for volcanic soils

$p_0 = 0.00302$ (0.00052) for Caroline, Wandilo and Myora sands

$p_0 = 0.00539$ (0.00036) for other yellow and white sands

$p_0 = 0.01119$ (0.00131) for Tantanoola flinty sands

$p_0 = 0.01759$ (0.00201) for terra rossa soils and brown soils from Comaum

$p_0 = 0.01172$ (0.00225) for yellow sands from Comaum

$p_0 = 0.00534$ (0.00175) for rendzinas

$p_0 = 0.005075$ (0.002623) for all soil types combined

$$A_0 = 10.0 \exp(\underset{(0.000841)}{-0.009172}\ Y_{10})$$

$$Z_1 = (\underset{(0.114)}{1.700} + \underset{(0.2335\ 10^{-3})}{0.4426\ 10^{-3}}\ D - \underset{(0.3050\ 10^{-6})}{0.7380\ 10^{-6}}\ D^2)\ (1.0 - \underset{(0.0731)}{0.4287}\ S)$$

and where

Pai = periodic annual increment,

Y_A = yield at age A,

A = age,

i = increment period,

p, A_0 and Z_1 are as for Equation VI.7,

Y_{10} = site potential,

D = competition level (volume), and,

S = thinning shock, (Vt/(Vb i)), relative volume, the effect lasting one year only.

EXTENSION TO OTHER REGIONS

Equation VI.12 was developed using data from the lower south-east region of South Australia and is a satisfactory predictor for that region. Of the 68 900 ha of radiata pine plantations administered by the Woods and Forests Department of South Australia, some 16 500 ha are outside that region and data from these areas were used to determine whether the model could be extended to other regions in South Australia and to provide an indication as to whether or not an Australia-wide model is feasible. The available data are relatively sparse (Appendices 1.1 and 1.5) but represent the total available data from these other areas.

The seven parameter model Equation VI.9 was then fitted to each data set in turn. The mean error in periodic annual increment was calculated and a t test used to determine whether this mean deviate was significantly different from zero. The results are summarised in Table VI.4.

Table VI.4

Fit of Equation VI.9 to data from other areas.

Region	Forest	Number of Observations	Mean deviate	Standard error of mean deviate	Significance
Upper south east	Cave Range	12	-5.36	1.51	Sig
	Noolook	37	-6.38	1.05	Sig
Adelaide Hills	Mount Crawford	63	-1.34	0.79	NS
	Kuitpo	92	0.32	0.56	NS
	Second Valley	48	-0.45	0.54	NS
Northern	Wirrabara	26	-6.63	0.69	Sig
	Bundaleer	23	-7.23	1.73	Sig

Upper south-east

For both Noolook and Cave Range Forest Reserves in the upper part of the south-east region of South Australia (Figure I.1), the mean deviate was significantly different from zero, the lower south-east model overestimating considerably.

Of the parameters in the model p_0 was considered the most likely to vary between regions, paralleling the soil type analysis. A number of models were fitted to the data based on the lower south-east model, Equation VI.9:

1 Allowing one parameter to float.

2 Allowing p_0 and one other parameter to float.

3 Allowing periodic annual increment to be corrected by a simple correction factor (C) applied to the lower south-east model.

This analysis provided biased estimators but was the best technique available as the data were too sparse to allow all seven parameters to be estimated. The results are summarised in Table VI.5.

The results were confusing. Allowing p_0 to float was inferior to the simple correction factor, which for Noolook provided the lowest variance estimator. For Cave Range the lowest variance estimator was the model allowing d_0 to float. Allowing d_0 to float was partly a proxy for the correction factor as the terms d_1 and d_2 have relatively little effect on increment. The two models (allowing p_0 to float and the correction factor C) were then fitted to the data and the deviates regressed against age and site potential to determine whether either or both models were satisfactory. A second order polynomial was used for Noolook where there were more data available, a simple linear model for Cave Range.

Table VI.5

Models for other areas

Forest	Number of observations	Total deviates squared								
		Equation VI.11	One parameter allowed to float							
			p_0	p_1	a_1	d_0	d_1	d_2	s	C
Cave Range	12	644.1	399.0	408.1	399.9	293.2	458.0	530.8	644.1	292.6
Noolook	37	3052.3	1783.8	1784.5	1701.3	1176.9	1607.0	2227.2	2921.4	1175.0
Wirrabara	26	1457.3	338.9	717.8	276.3	348.8	521.0	735.1	1299.1	374.4
Bundaleer	23	2717.7	2168.8	2310.6	1372.1	1425.8	1737.0	2111.7	2528.3	1429.6

Forest	Total deviates squared					
	Two parameters allowed to float					
	p_0p_1	p_0a_1	p_0d_1	p_0d_2	p_0d_3	p_0s
Cave Range	398.3	293.2	281.9	378.3	300.8	399.0
Noolook	1715.5	1573.9	1125.7	1584.0	1763.6	1531.1
Wirrabara	271.5	268.4	336.9	303.2	292.1	335.7
Bundaleer	1961.2	256.3	543.9	1589.2	2089.1	2166.3

Forest	Estimated parameters		
	p_0	a_1	C
Cave Range	0.03594		0.771
Noolook	0.03720		0.743
Wirrabara	0.03128	-0.016091	0.690
Bundaleer	0.02656	-0.017776	0.644

The deviates were not significantly related to site potential but were related to age for both alternative models for Noolook, and for the p_0 model for Cave Range. Although the correction factor provided deviates that were not significantly related to age for Cave Range the estimated F value was only just below the critical level. These results were disturbing as it had been hoped that these models would not be significant. Inspection of the data showed that the overestimates were generally associated with the 1967/68 growing season which throughout the state had the lowest rainfall recorded, and these observations had a relatively greater effect on the regressions because of the narrow range of plantation years in the data.

For both forests the data were inadequate and it was concluded that further analysis was not warranted until more data are available. For both forests the correction factor was slightly better than the p_0 model in that the significance levels were slightly lower. Until more data are available the correction factor approach is preferred because it is, and can readily be seen to be, a simple approximation that should be revised as soon as practicable. The correction factor for Noolook Forest Reserve is 0.743 and for Cave Range Forest Reserve 0.771.

Adelaide Hills region

For the three forests in the Adelaide Hills region, Mount Crawford, Kuitpo and Second Valley Forest Reserves, the mean deviates that were obtained when Equation VI.9 was fitted to the data were not significantly different from zero, (Table VI.4). The deviates were regressed against age and site potential (a second order polynomial) and none of the regressions were significant. The lower south-east model is therefore a satisfactory predictor for each of the Forest Reserves in the Adelaide Hills region.

Northern region

For the Forest Reserves of the Northern region, Wirrabara and Bundaleer, there were relatively few data available and for both these forests the mean deviate was significantly different from zero, Table VI.4.

Carrying out analysis similar to that for the other significantly different forests again produced confusing results, Table VI.5. For Wirrabara the lowest total deviates squared was when a_1 was allowed to float, and allowing p_0 to float was superior to the use of the correction factor to increment. For Bundaleer allowing a_1 to float again had the lowest total deviates squared but the correction factor was superior to allowing p_0 to float. Allowing a_1 to float provided an unsatisfactory model for if Y_{10}=100 then A_0=2.0 and 1.7 for the two forests respectively compared with 4.0 for the south-east model. This was biologically unsound as at age A_0 some trees must be 10.5 cm in diameter. It was better to use either the correction factor or p_0.

Fitting these two models to each of the data sets and regressing the deviates against age and site potential showed that the deviates were significantly related to age for Bundaleer but not Wirrabara, and that the deviates were not related to site potential for either forest. Bundaleer provides an even worse example of badly distributed data affected by the 1967/68 drought than Noolook and Cave Range Forest Reserves. The Bundaleer data were from only four plantations, and the youngest of these was the most severely affected by drought with many trees having dead tops and as well there were a number of deaths. As these data were also for a one year increment period the drought effect was accentuated greatly.

The analysis was unsatisfactory because the data were inadequate, but until further data are available it is probably better to use the correction factors of 0.690 for Wirrabara Forest Reserve and 0.644 for

Bundaleer Forest Reserve for two reasons. Firstly the use of the correction factor continually reminds users that it is nominal adjustment, and secondly the improvement in efficiency of the correction factor over p_0 for Bundaleer was greater than the improvement in efficiency of p_0 over the correction factor for Wirrabara.

Summary

The analysis indicated that the lower south-east model can be extended to other areas, although the success of the extension depends on having suitable, sufficient, accurate data available. There was an indication that even when the available data were poor a relatively simple correction factor to the increment may provide a model that can be used until further data become available. No changes were necessary to enable the model to be used for the three Adelaide Hills forests some 400 km from the south-east.

The conclusion was that the same model structure may prove to be satisfactory for other radiata pine areas of Australia. If good data are not available then a relatively simple analysis, estimating p_0 for that area or even estimating the simple correction factor, may provide a satisfactory model until more data are available and all parameters can be re-estimated.

SECOND ROTATION STANDS

There is a considerable area of second rotation plantations of radiata pine in South Australia and these plantations are generally of lower productivity than the first rotation on the same site (Keeves, 1966). It is of considerable practical management importance to know whether, apart from being of different site potential, the yield function is different for second rotation stands compared with first rotation stands. Appendix 1.4f lists the available plots and their plantation year, and although they do not appear to be very well balanced by age, site potential or forest they were the best available and probably represent the current distribution of second rotation stands over age 10 fairly well. There were 33 plots with 157 increment periods.

When Equation VI.9 was fitted to the second rotation data the mean deviate of 0.003, not significantly different from zero, for the mean deviate had a standard error of 0.42. When the deviates were regressed against a second order polynomial in age and site potential the resultant regressions were not significant. Although 73 of the 157 observations were included in the data used to develop Equation VI.11 (7.5% of that data set) these results indicated that second rotation stands have the same yield function as first rotation stands. The pooling implied in the data used to develop the south-east model was justified.

The investigation by Keeves (1966) into the second rotation decline in productivity cites some evidence that yield trends were similar between rotations but that the absolute level was lower. This analysis confirmed that the yield trend is independent of rotation.

SUMMARY

The conditioned Bertalanffy periodic annual increment model was successfully extended for use in thinned stands. Thinning was included in two ways. Firstly, the level of competition was incorporated in a simple quadratic correction to periodic annual increment that approximated Langsaeter's (1941) qualitative model. The very flat response surface was influenced by lack of sufficient heavily thinned plots in the data base and care should be taken not to extrapolate outside the range of the data. Standing volume was better than number of trees as the index of the competition. The second way thinning was incorporated into the model was as thinning shock, the best measure of which was the relative change in volume, with the effect assumed to last only one year.

Investigation of soil type was hampered by the lack of a suitable nonlinear parameter estimation program that could handle dummy variables. The analyses carried out provide biased estimators and although the extent of the bias is unknown it was thought to be relatively small. Form indices were not an adequate continuous variable proxy for soil type. Seven soil groups were recognised, five of practical significance. The resultant model could be used in practice in the south-east if inventory is modified to include the recording of soil type.

Both Equations VI.9 and VI.12 were satisfactory predictors.

Evaluating the model on areas other than the lower south-east of South Australia showed that for the Adelaide Hills region where more data were available, the south-east model (Equation VI.9) was an unbiased predictor. For Wirrabara and Bundaleer Forest Reserves in the Northern region, and Noolook and Cave Range Forest Reserves in the upper part of the south-east, the lower south-east model was biased but the data were inadequate for a detailed analysis. It was concluded that extension of the study to these areas of radiata pine plantation is probably feasible, the level of success depending on the quality and quantity of the data available.

Second rotation stands were shown to be satisfactorily predicted by the lower south-east model. This supported Keeves' (1966) contention that, although the absolute level of site potential changes between rotations, the same yield model can be used for both first and second rotations.

PART 2

GENERALIZED LEAST SQUARES ESTIMATION

VII MODELS FOR UNTHINNED STANDS

GENERALIZED LEAST SQUARES ANALYSIS

MODEL FORMULATION

ESTIMATION OF THE VARIANCE-COVARIANCE MATRIX

MODEL DEVELOPMENT

First stage models

Second stage models

Alternative error structures

Testing the model

Unconditioned model

SUMMARY

VII MODELS FOR UNTHINNED STANDS

GENERALIZED LEAST SQUARES ANALYSIS

In the preceding chapters Ordinary Least Squares (OLS) was used to estimate a yield prediction model. The analysis used to produce the final model did not take explicit advantage of the trend series available within each plot, the data within each plot being treated as if they were random and independent observations. Intuitively it would seem more sensible to utilise the time series nature of the available data by using a two stage approach, analysing each plot in turn and then analysing the coefficients of these plots.

A two stage approach was tried in Chapter V but failed to produce a satisfactory prediction model. In estimating the second stage parameters using OLS it was necessary to treat each of the first stage parameters as if they were independent when in fact they are likely to be highly correlated.

Consideration of this problem led Dr I.S. Ferguson to suggest the possibility of using Generalized Least Squares (GLS) rather than OLS in developing the second stage models, so that the correlation between the parameters in the first stage model would explicitly be taken into account in the development of the second stage models. Dr Ferguson then developed the technique which was programmed in ALGOL by Mr. J.A. Miles of the Department of Forestry, Australian National University. This was then used empirically to determine whether the technique offered significant advantages over OLS in the development of a growth model.

Details of that study (Ferguson and Leech, 1976b) are described in Appendix 6. The study showed that there was a slight improvement in relative efficiency if the structure of the error term was considered to be heterogeneous across plots and a marked increase in relative efficiency

(Appendix 6, Table 5) if the error terms relating to the individual coefficients were assumed to be correlated.

The results of Chapter V, where the difference equation estimating periodic annual increment from the pooled data was found to be superior to the two stage yield model, were reinterpreted in the light of this GLS study. The pooled data approach takes the correlation between the first stage parameters into account but does not utilise the time series to the full. On the other hand, the two stage OLS approach does not take account of the correlation between the parameters but does utilise the trend information inherent in the time series for each plot. GLS enables both to be included in the analysis.

Furthermore, in the OLS analysis, the pooled observations were derived from time series for each plot and were not truly independent. The standard errors of the parameter estimates were therefore under-estimated by an unknown amount that could not be estimated, and it is possible that some parameters were included when they should not have been. GLS analysis avoids this problem.

The GLS study (Appendix 6) dealt only with the Johnson-Schumacher or Clutter model. In the light of the OLS results which showed the second level Bertalanffy model to be superior to that model, further trials of the GLS approach were made using the Bertalanffy model.

MODEL FORMULATION

When the second level Bertalanffy model was developed in a two stage process, the first stage models were clearly nonlinear in the parameters for both the conditioned and the unconditioned model forms.

$$Y = \left\{(n/p)\left[1 - \exp(-p(1-m)(A-A_0))\right]\right\}^{\frac{1}{1-m}} \qquad (VII.1)$$

and

$$Y = Y_{10}\left\{\frac{1 - \exp(-p(1-m)(A-A_0))}{1 - \exp(-p(1-m)(10-A_0))}\right\}^{\frac{1}{1-m}} \qquad (VII.2)$$

The second stage structure used in Chapters V and VI was linear for the parameter p, but nonlinear for the parameter A_0.

$$p = p_0 - p_1\ Y_{10} \qquad (VII.3)$$

$$A_0 = 10\ \exp(-a_1\ Y_{10})$$

where

Y = yield,

A = age,

Y_{10} = site potential, yield at age 10,

A_0 = the age at which volume growth commences,

n and p are the parameters in the second level Bertalanffy model, and,

p_0, p_1 and a_1 are the parameters to be estimated.

This formulation was unsuitable for GLS analysis because the algorithm was designed for linear second stage models only. However by substituting (10 $\exp(-a_1)$) for A_0 in the first stage the model has a more complex first stage, but a simple linear second stage. This enables the GLS program to be used without the development of a nonlinear version.

Apart from Equation VII.3, other formulations including site potential were possible such as those in Equation V.6. Stocking at age ten (N_{10}) is the best measure of initial spacing and was also evaluated. Differences between forests or groups of soil types were investigated by the use of dummy variables (Johnston, 1963; Cunia, 1973).

ESTIMATION OF VARIANCE-COVARIANCE MATRIX

Extending the GLS treatment to embrace a nonlinear first stage using Bard's (1967)program (see Chapter III) assumes that the estimate of the variance-covariance matrix is unbiased and relatively precise.

Although the estimates derived from Bard's program appeared to be sensible estimates it was impossible to confirm analytically that they were unbiased and efficient. Monte Carlo simulation was therefore used to investigate the problem further.

For each of the twenty plots in the developmental data used in Chapter V, Bard's program was used to estimate the parameters n, p and a_1, and the variance-covariance matrix for the following model;

$$Y = (n/p) \left[1 - \exp(-p(A-10 \exp(-a_1)))\right] \qquad \text{(VII.4)}$$

where the parameters and variables are as for Equations VII.1 to VII.3. The unconditioned model was preferred to the conditioned for this analysis because it was more complex and thus more likely to indicate any problems.

For each plot the variance of the residuals for this model was then used, together with a generator of normally distributed random numbers, to define new data sets based on a random disturbance to the original data. For each new data set the parameters were re-estimated and this was repeated until the variance-covariance matrix of these parameter estimates appeared to have stabilised. Commonly this was between 10 000 and 50 000 iterations.

For each plot the variance-covariance matrix estimated by Bard's program was compared with the Monte Carlo estimate using the following test (Morrison, 1976).

$$L = (N-p)\,(\ln |\Sigma_0| - \ln |S| + \mathrm{tr}(S\,\Sigma_0^{-1}) - p) \qquad \text{(VII.5)}$$

where

L = the likelihood, distributed chi-squared with $(p(p+1))/2$ degrees of freedom,

N = the number of observations,

p = the number of parameters,

Σ_0 = the Monte Carlo estimate of the variance-covariance matrix,

Σ_0^{-1} = its inverse,

$|\Sigma_0|$ = its determinant,

S = the estimate of the variance-covariance matrix from Bard's program,

$|S|$ = its determinant, and,

tr = the trace of the product matrix (the sum of the diagonal elements).

For small N, Bartlett (1954) suggested that the statistic should be scaled to yield a new statistic L^* before testing against chi-square.

$$L^* = \left\{1 - \frac{1}{6(N-1)}\ (2p + 1 - 2/(p+1))\right\} L \qquad \text{(VII.6)}$$

This test was particularly rigorous because the Monte Carlo estimate is really a stochastic estimate rather than the true variance-covariance matrix, and a less powerful test would have been more appropriate.

Using a probability level of p=0.05, none of the twenty plots yielded estimated variance-covariance matrices which differed significantly from the Monte Carlo estimate. The estimated variance-covariance matrix from Bard's program was therefore considered acceptable for use in the GLS models.

MODEL DEVELOPMENT

First stage models

The conditioned model, Equation VII.7, was then fitted to the developmental data used in Chapter V.

$$Y = Y_{10} \left\{ \frac{1 - \exp(-p(1-m)(A-10 \exp(-a_1)))}{1 - \exp(-p(1-m)(10-10 \exp(-a_1)))} \right\}^{\frac{1}{1-m}} \qquad \text{(VII.7)}$$

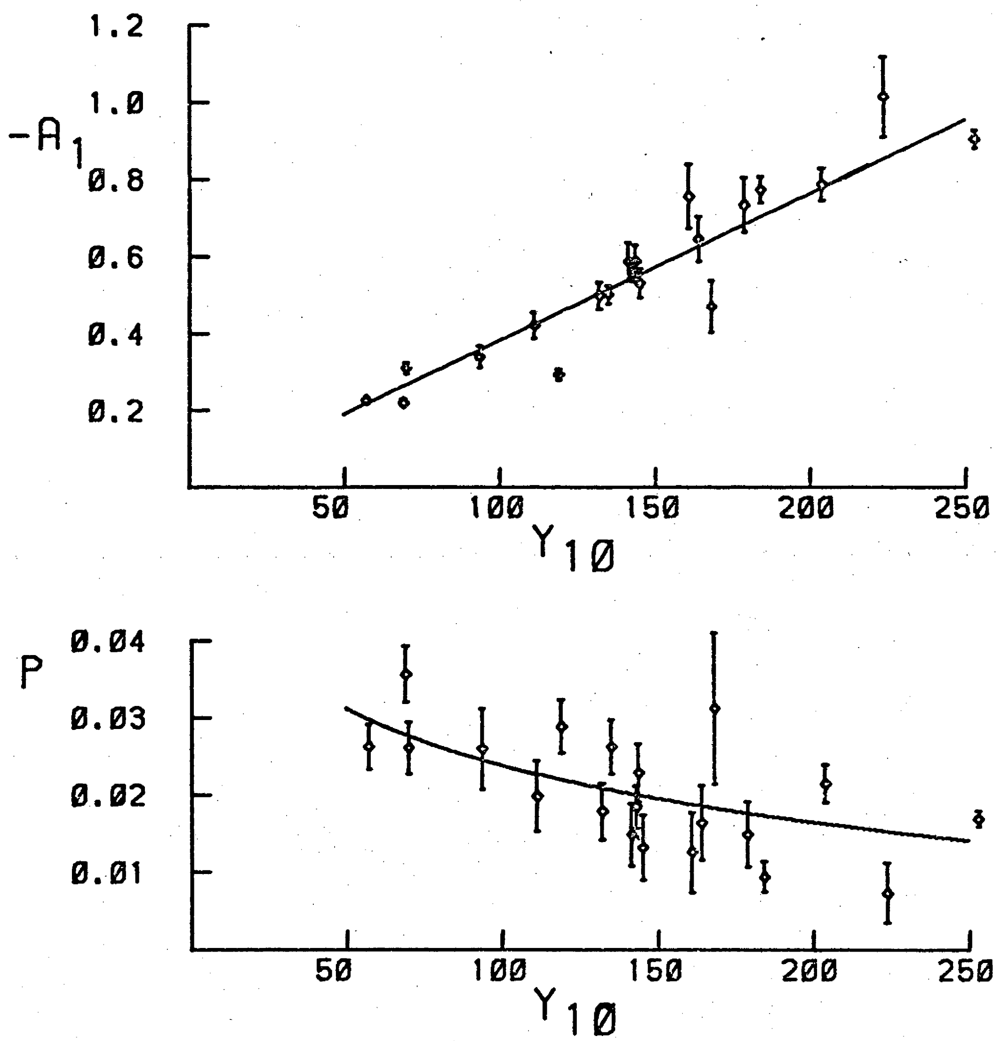
where the parameters and variables are as for Equations VII.1 to VII.3.

The parameter estimates all had large sampling errors and none was significantly different from zero using a t test, paralleling the OLS results. For increased accuracy the analytical partial derivatives were used rather than the approximation described in Appendix 2. The models were refitted with m=0 reducing the number of parameters to be fitted to two for each plot. The parameter estimates for both p and a_1 appeared sensible for all plots, and, although some estimates had large standard errors, the parameters were all significantly different from zero: the results were accepted. The parameter estimates and their 95% confidence limits are summarised in Figure VII.1.

Scatterplots of residuals for each plot gave no indication of heterogeneity or serial correlation within any of the plots. The Durbin-Watson d statistic (Durbin and Watson, 1950, 1951) was calculated even though its value is questionable with so few observations. Published critical bounds only go down to 15 observations and extrapolating these upper and lower bounds to the number of observations for each plot is difficult and unsatisfactory. Recognising the dangers inherent, the statistic was tested against these estimated critical bounds and it was found that none of the plots had values of the d statistic below the lower bound. Of the twenty plots, fourteen fell into the inconclusive zone, six for positive serial correlation and eight for negative. In

Figure VII.1

Conditioned Bertalanffy model
First stage estimates and 95% confidence limits



spite of the large proportion in the inconclusive zone, serial correlation was not considered a problem because it was expected that if serial correlation was a problem then the sign would be consistent for all or most of the plots, and this was patently not so.

Second stage models

Various forms of the second stage models utilising Y_{10} and N_{10} were then evaluated, including one form which permits a direct comparison with the OLS results shown earlier in Equation V.12:

$$p = \underset{(0.00353)}{0.03194} - \underset{(0.2449\ 10^{-4})}{0.7802\ 10^{-4}}\ Y_{10} \qquad \text{(VII.8)}$$

$$a_1 = \underset{(0.000317)}{-0.003803}\ Y_{10}$$

Comparing the standard errors of the parameter estimates of Equations VII.8 and V.12 shows that the GLS estimates are consistently and considerably higher than the OLS estimates (0.00353 cf. 0.00108, 0.2449 10^{-4} cf. 0.0503 10^{-4} and 0.000317 cf. 0.000052). As described earlier, and also on page 261, Appendix 6, this is not an artefact of the GLS method but is caused by the OLS estimates being underestimated because the data from each plot are related and not truly independent.

This suggests that the best models estimated by OLS (Equations VI.9 and VI.12) may contain some unnecessary parameters, with underestimation of the standard errors leading to the possibility of erroneous rejection of the null hypotheses concerning some parameters. However it was impossible to gauge how serious the problem was other than relying on the tests based on independent data. These suggested that Equation VI.9 was satisfactory.

The results for the various second stage models are summarised in Table VII.1. The significance test used to test each parameter is defined in Equation 22, Appendix 6. Equation VII.8, Model 5, was the

Table VII.1

Comparison of alternative unthinned models, conditioned Bertalanffy

Model	First Stage variables												Residual sum squares
	P								a_1				
	Secondary variables								Secondary variables				
	1	Y_{10}	Y_{10}^2	$Ln(Y_{10})$	$Ln(Ln(Y_{10}))$	$^1/Y_{10}$	$exp(-Y_{10})$	N_{10}	1	Y_{10}	Y_{10}^2	N_{10}	
1	*	NS	(NS)					NS	NS	NS	(NS)	NS	17.93
2	*	*						NS	NS	*		(NS)	18.83
3	*	*						NS	(NS)	*			18.98
4	*	*						NS		*			18.99
5	*	*								*			20.23
6	*	*							NS	*			20.22
7	*	*							(NS)	NS	NS		19.51
8	*	*								*	NS		19.88
9	*	*							(NS)	*		NS	19.88
10	*	*								*		NS	20.11
11	*	*						NS		*		NS	18.99
12	*	NS	(NS)							*			18.91
13	*			*						*			19.09
14	NS			NS	NS					*			19.05
15	*			*				NS		*			18.21
16	*			*					NS	*			19.06
17	*			*					(NS)	*		NS	18.86
18	*			*						*		NS	18.97
19	*			*					NS	NS	(NS)		19.01
20	*					*				*			19.42
21	*						NS			*			29.23
22	*									*			30.38

Note
- * parameter significant
- NS parameter not significant
- (NS) parameter not significant and zeta value the lowest of all the estimates.

best of the models including a linear function in Y_{10} (Models 1-12), in that no other parameter was significant when added to this model. Replacing Y_{10} by $(\ln(Y_{10}))$ produced a superior estimator (Model 13, Equation VII.9) that has a lower residual sum squares (19.09 compared with 20.23) and has the same number of parameters. This model could not be improved upon by the addition of other parameters (Models 14-19) or by alternative structures for p (Models 20-22).

$$p = \underset{(0.01524)}{0.07290} - \underset{(0.00317)}{0.01065} \ln(Y_{10}) \qquad \text{(VII.9)}$$

$$a_1 = \underset{(0.000316)}{-0.003813}\, Y_{10}$$

Further analyses were made using Equation VII.9. Each parameter was replaced by dummy variables representing the soil and forest groups defined in Appendix 1.6a. Although the residual sum of squares was smaller (Table VII.2), these more elaborate models were not significantly better than Model 13, Equation VII.9.

Alternative error structures

For Equation VII.9 the correlation between the parameters p and a_1 was 0.76 suggesting that there should be a considerable gain in efficiency through the use of GLS. Bartlett's (1937) test across the twenty plots also showed that the variances were heterogeneous (chi-square 106.3). Table VII.3 shows that the gain in efficiency through the recognition of the correlation between parameters was quite marked, for the relative efficiency assuming heterogeneous independent errors was only 0.182 compared with 1.0 for the assumption of heterogeneous correlated errors. This was a measure of the advantage of the two stage GLS compared with the two stage OLS model. On the other hand the gain in efficiency due to the implicit recognition of heterogeneity was relatively slight, (relative efficiencies of 0.991 compared with 1.0, 0.180 compared with 0.182).

Table VII.2

Comparison of alternative unthinned models conditioned Bertalanffy, by soil type and forest

Model Number	Parameter	Varying with	Number of second stage parameters	Number of groups	Residual sum squares	Compared with	
						Model	Significance
13			3	1	19.09		
23	p_0	Soil type	4	2	18.26	13	NS
24	p_1		4	2	18.25	13	NS
25	a_1		4	2	18.88	13	NS
26	p_0		5	3	18.08	13	NS
27	p_1		5	3	18.03	13	NS
28	a_1		5	3	18.80	13	NS
29	p_0		6	4	17.93	13	NS
30	p_1		6	4	17.76	13	NS
31	a_1		6	4	18.62	13	NS
32	p_0		7	5	17.33	13	NS
33	p_1		7	5	17.23	13	NS
34	a_1		7	5	18.30	13	NS
35	p_0	Forest	5	3	18.50	13	NS
36	p_1		5	3	18.40	13	NS
37	a_1		5	3	18.75	13	NS

Table VII.3

Comparison of alternative error structures for unthinned model

Model	Joint dependent variables			Relative efficiency
	p		a_1	
	Secondary variables			
	1	$Ln(Y_{10})$	Y_{10}	
Heterogeneous correlated	0.07290 ($2.323\ 10^{-4}$)	-0.01065 ($1.006\ 10^{-5}$)	-0.003813 ($9.978\ 10^{-8}$)	1.0
Homogeneous correlated	0.07245 ($2.336\ 10^{-4}$)	-0.01055 ($1.013\ 10^{-5}$)	-0.003818 ($9.940\ 10^{-8}$)	0.991
Heterogeneous independent	0.07416 ($5.010\ 10^{-4}$)	-0.01089 ($2.066\ 10^{-5}$)	-0.003804 ($10.046\ 10^{-8}$)	0.182
Homogeneous independent	0.07387 ($5.075\ 10^{-4}$)	-0.01082 ($2.093\ 10^{-5}$)	-0.003809 ($10.011\ 10^{-8}$)	0.180

Estimated variance shown in brackets below the coefficient

Testing the model

The independent test data used in Chapter V were also used to test Equation VII.9. The standard deviation of the deviates was 46.44, lower than for the OLS models (51.51 for the conditioned periodic annual increment model developed on the pooled data, 66.1 for the two stage OLS). For Equation VII.8, with the same structure as the OLS models, the standard deviation of the deviates was 47.98, indicating that the differences were not solely due to the change in the second stage structure of the parameter p from linear to log-linear. Thus the advantages of GLS over OLS seem to extend to predictors as well as estimators.

Figure VII.2 shows the developmental data and the estimated yield curves for these data. Most of the yield curves seemed to provide a good fit visually, although 73, 89 and 346 are underestimated and 58 and 322 overestimated.

The test data cover a wider range of Y_{10} and N_{10} than the developmental data. The four highest and the four lowest plots with respect to Y_{10} for the independent test data were plotted in Figure VII.3 together with their estimated yield curves. Seven of these eight plots had values of Y_{10} outside the range of the developmental data. For the very low site quality plots Equation VII.9 consistently overestimated yield, especially for plot 369, suggesting structural misspecification. For the very high site quality plots the equation was satisfactory for plots 155 and 433, but EP24C and EP24E appeared to exhibit a different form of yield curve, EP24C being overestimated as well. These latter two plots were originally planted at 6x6 feet and 9x9 feet respectively instead of the more common 7x7, 8x8 or 9x7 feet (2.1x2.1-2.4x2.4 m). These two plots were established on a Tantanoola flinty sand where early rapid growth would normally be expected to be followed by a faster than average

Figure VII.2

Estimated yield functions for
developmental data

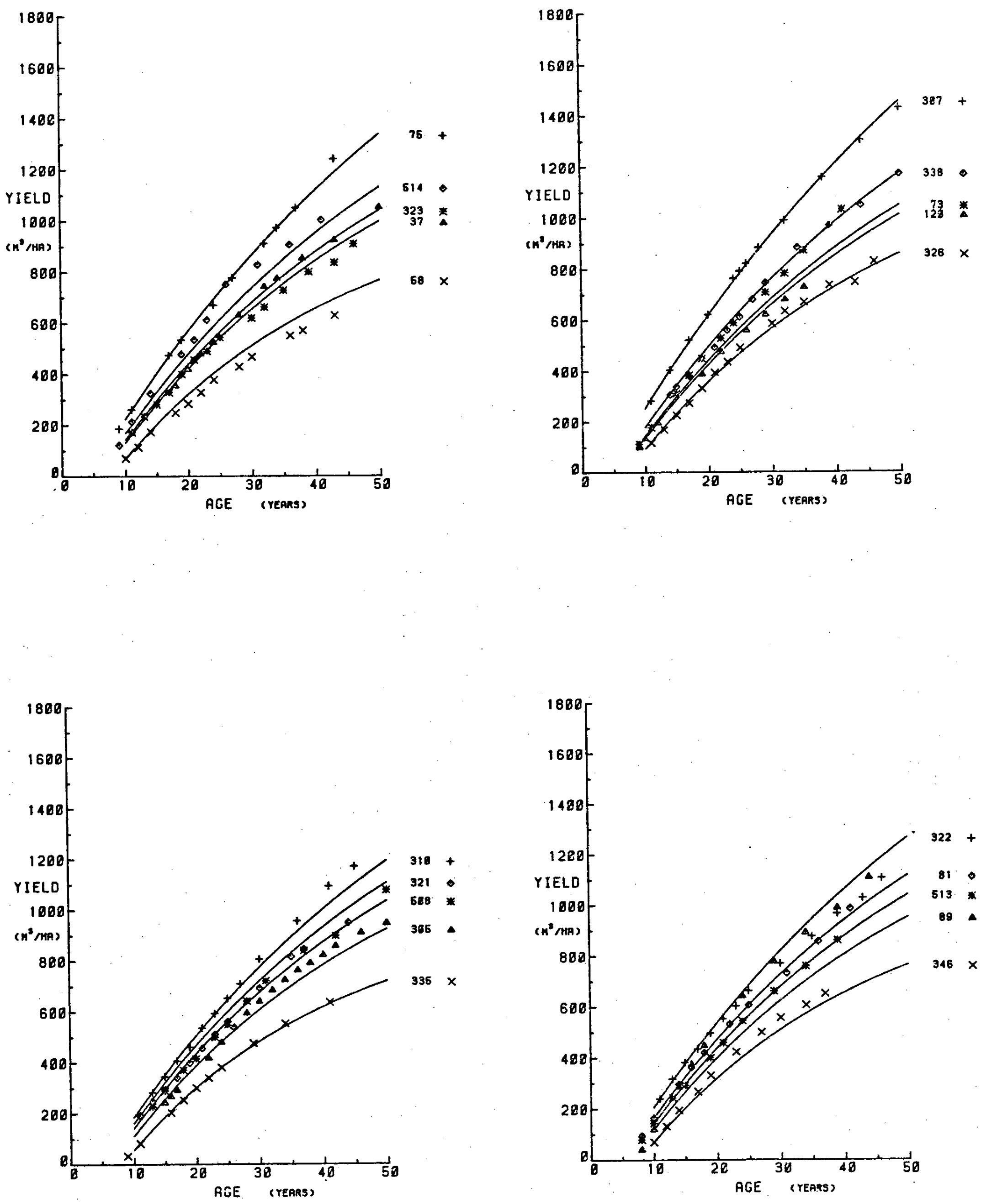
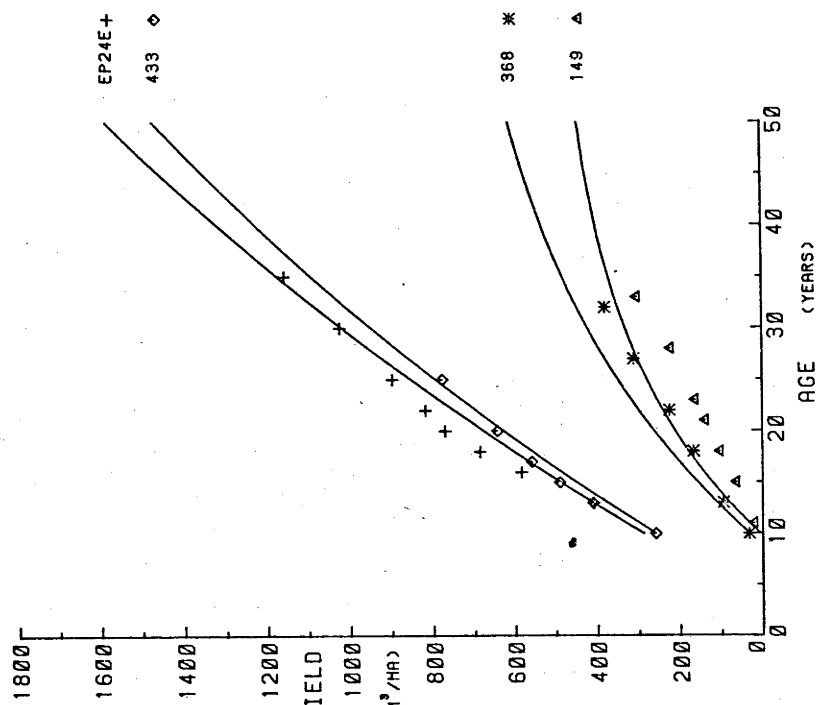
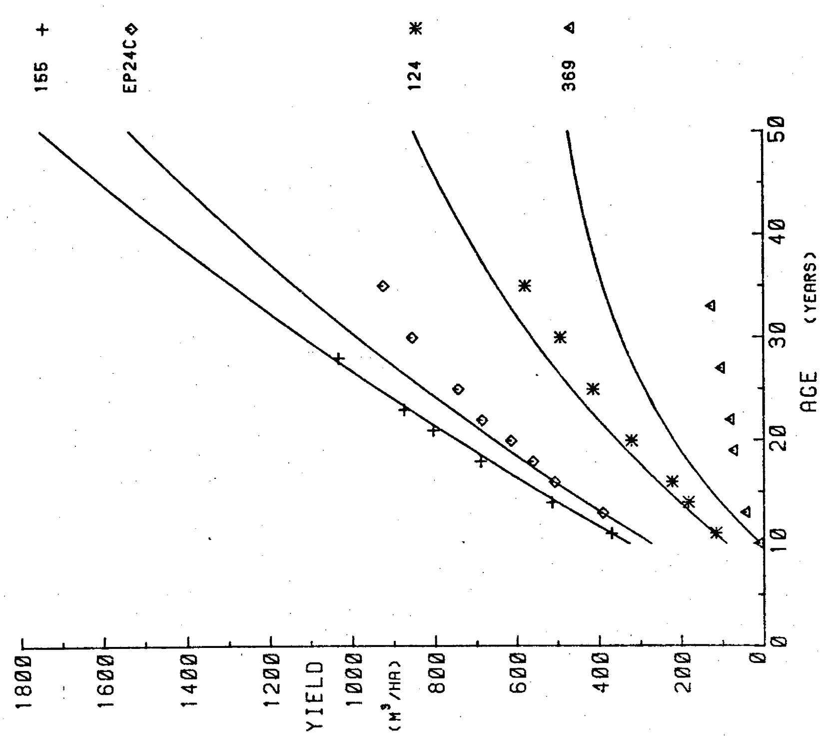


Figure VII.3

Estimated yield functions for selected independent test data





decline in growth (see Chapter VI). These differences in stocking and soil type probably account for the anomalies with respect to these plots, as there were no plots in the developmental data on this soil type or with the initial plantation espacements.

Unconditioned model

For completeness similar analyses were carried out using the unconditioned model, Equation VII.1. The best model was

$$p = \underset{(0.01297)}{0.07575} - \underset{(0.00272)}{0.01109} \ln(Y_{10}) \qquad \text{(VII.10)}$$

$$a_1 = \underset{(0.000376)}{-0.003535}\, Y_{10}$$

$$n = \underset{(0.242)}{7.557} \ln(Y_{10})$$

$$m = 0.0$$

but as with the OLS analysis the unconditioned model was less efficient than the conditioned as a predictor of the test data (standard deviation of the deviates 52.34). Also, the error in using Equation VII.10 to estimate yield at age ten was considerable, ranging from -1% to 15% over the range of the data. The unconditioned model was not considered further.

SUMMARY

The GLS analysis of the unthinned stand data was superior to the OLS analysis in that it satisfied the assumptions of the analysis, was more efficient, yielded accurate estimates of the standard errors of the parameters, and provided a superior predictor.

VIII MODELS FOR THINNED STANDS

FIRST STAGE MODELS

Formulation

Selection

BAYESIAN ESTIMATION

DATA

FIRST STAGE MODEL DEVELOPMENT

Thinning parameters

SECOND STAGE MODEL DEVELOPMENT

Site potential

Other variables

Alternative error structures

EXAMINATION OF THE DEVELOPMENTAL DATA

EXAMINATION OF THE INDEPENDENT TEST DATA

SUMMARY

VIII MODELS FOR THINNED STANDS

FIRST STAGE MODELS

Formulation

The structure of the models developed for thinned stands in Chapter VI using OLS was inappropriate for the GLS analysis because the range of competition levels does not, and cannot, occur in each and every plot but, unlike site potential, competition level does vary with age in a plot. The structure with respect to competition must be simple, for the first stage GLS structure already has two parameters to be estimated from a time series with between 9 and 15 observations.

If Ti_j represents the thinning grade (ratio of volume of trees removed to the volume before thinning) then this is likely to affect the parameter p after the age of the j th thinning, but should not affect the parameter a_1.

A general formulation should therefore include an adjustment to the parameter p for each thinning prior to the current age, but because the number of thinnings varies between plots, and also varies for each observation within a plot, a summation form such as Equation VIII.1 was necessary.

$$p = p_0 + t_1 \sum_{j=1}^{j=k} \left(\left| Ti_j - t_1 \right| \right)^{t_2} \quad \text{for all } A \geqslant A_j \qquad \text{(VIII.1)}$$

where

p = the parameter in the second level Bertalanffy model,

Ti_j = thinning grade of the j th thinning,

k = the number of thinnings,

A = age,

A_j = the age of the j th thinning, and,

t_0, t_1 and t_2 are the parameters to be estimated.

The model parallels the plateau competition effect (Langsaeter, 1941; Moller, 1954a, 1954b) described in Chapter VI, in that for t_0, t_1 and t_2 positive, then p is at minimum if $Ti_j = t_1$ for all j, increasing as Ti_j diverges from t_1, the rate ever increasing. The asymptotic maximum yield, the ratio n/p, would be lower for both unthinned and heavily thinned stands than for "well thinned" stands.

This formulation is also supported by evidence from the series of four plots X, Y, 304 and 305 described by Lewis (1962). Although the plots were comparable at age 17, the total volume production by the control at age 50 was approximately 200 m^3/ha lower than for the two lightly thinned plots, while the more heavily thinned plot was approximately 100 m^3/ha below these two plots.

As it was unlikely that three thinning parameters could be estimated as well as p and a_1, restricted models were formulated at various levels.

$p\ a_1 : t_0 = 0$	No thinning parameters	(VIII.2a)
$p\ a_1\ t_0 : t_1 = 0\ t_2 = 1$	Models with one thinning parameter	(VIII.2b)
$p\ a_1\ t_0 : t_1 = 0\ t_2 = 2$		(VIII.2c)
$p\ a_1\ t_0\ t_1 : t_2 = 1$	Models with two thinning parameters	(VIII.2d)
$p\ a_1\ t_0\ t_1 : t_2 = 2$		(VIII.2e)
$p\ a_1\ t_0\ t_2 : t_1 = 0$		(VIII.2f)
$p\ a_1\ t_0\ t_1\ t_2$	Model with three thinning parameters	(VIII.2g)

Selection

The selection of the best of the seven models is complicated by the variation in results across the various first stage data sets. The technique adopted was to summarise for each plot, and for each of the different numbers of parameters, the model with the lowest residual sum of squares; as well as using F tests to compare different numbers of parameters. The 'best' model then was the model that for the majority of plots had the lowest residual sums of squares and which for the majority of plots had the largest number of significant parameters. If too many noisy parameters were included then this would become apparent in the second stage analysis when no secondary structure would be significant. The level of the best model was, in essence, the most complicated for which second stage models could be determined.

BAYESIAN ESTIMATION

As the data were relatively sparse it was considered unlikely that more than four parameters could be estimated for each plot in the first stage models, and even then, the estimation technique would have to make maximum use of the information. A sequential estimation technique seemed desirable working from the model based on the unthinned data, and then refining these estimates and estimating the thinning parameters using data from thinned stands.

Sequential estimation is possible using Bayesian statistics (Bayes 1763, (1958); Raiffa and Schlaifer, 1961; Lindley, 1972; Box and Tiao, 1973). In Bayesian statistical theory data are used together with a prior distribution from previous experiments to produce a posterior distribution. Equation VII.9 which was developed from unthinned stand data can be used to provide prior estimates of the parameters p and a_1 in an analysis of the thinned stand data. Utilising Bayesian statistics gave a better chance of obtaining satisfactory estimates of the first stage thinning parameters, and made full use of the available data.

Bayesian theory is not without its opponents. However this application can also be treated within the framework of classical statistical theory, by simply regarding the estimates from the unthinned data as prior estimates from a previous experiment. While the Bayesian approach often has considerable appeal for other applications, the debate between Bayesian and Classical schools is largely irrelevant in this case.

Bard's (1967) program can perform such an analysis, offering the option of utilising as the prior estimate the parameter estimate and either:

1 the standard errors of the parameter estimates, or,

2 the variance-covariance matrix of the parameter estimates.

The latter is clearly more appropriate in a mixed Bayesian-GLS analysis of this kind.

DATA

The prior estimate for the parameters p and a_1 was therefore Equation VII.9.

$$p = 0.07290 - 0.01065 \ln(Y_{10}) \qquad (VII.9)$$

$$a_1 = -0.003813\ Y_{10}$$

which has the variance-covariance matrix

$$\sigma^2_{a_1a_1} = 9.9784\ 10^{-8}\ Y_{10}^{\ 2}$$

$$\sigma^2_{pp} = 2.3233\ 10^{-4} - 9.6121\ 10^{-5} \ln(Y_{10}) + 1.0059\ 10^{-5} (\ln(Y_{10}))^2$$

$$\sigma^2_{pa_1} = -1.3038\ 10^{-6}\ Y_{10} + 3.4630\ 10^{-7}\ Y_{10} \ln(Y_{10})$$

OLS estimation had shown that the Bertalanffy model seemed relatively insensitive to the effects of thinning. Hence a prior estimate of zero with an infinite variance and zero covariance was thought likely to be satisfactory for the thinning parameters.

For the posterior model 58 plots were extracted from the data base. These plots all had at least nine volume measurements, the age of first measurement being 13 or less, although 49 of the plots were measured within one year of age ten. The data are summarised in Appendix 1.6b. The average number of observations was 10.5 and the average growth period 24.7 years.

FIRST STAGE MODEL DEVELOPMENT

Using the prior estimate, Equation VIII.2a without any thinning parameters was then estimated. For increased accuracy the analytical derivatives were used instead of the numerical approximation described in Appendix 2. The estimated parameters p and a_1 are shown in Figure VIII.1, together with the 95% confidence limits for each estimate. The estimated parameters show similar trends with site potential to the un-thinned plots which are plotted in Figure VII.1.

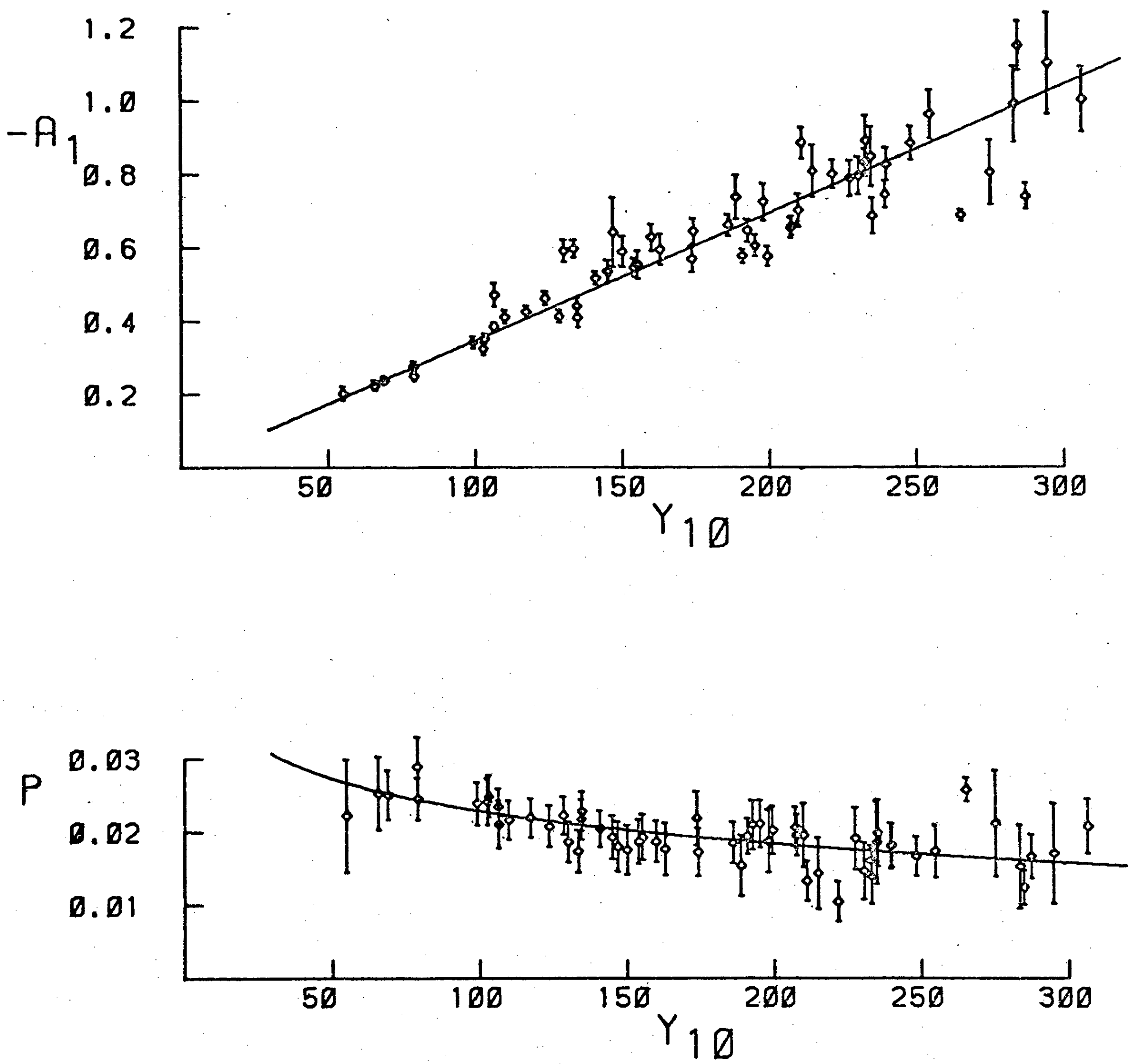
Thinning parameters

The two single thinning parameter models (Equations VIII.2b and VIII.2c) and the three two parameter models (Equations VIII.2d, VIII.2e and VIII.2f) were then fitted to the data.

Comparing the single parameter models with Equation VIII.2a showed Equation VIII.2b to be significantly better for only 9 out of the 58 plots, Equation VIII.2c for only 7. When an extra parameter was added (Equations VIII.2d, VIII.2e and VIII.2f), the results were significantly better in only 4, 3 and 1 plot out of the 58. It was extremely difficult to get parameter estimates for some plots, convergence being extremely slow for Equations VIII.2d, VIII.2e and VIII.2f. This suggests over-parameterisation in the sense that the equations seemed to be statistically incompatible with the data. In view of this, Equation VIII.2g with three thinning parameters was not even fitted to the data.

Figure VIII.1

Conditioned Bertalanffy model
Posterior
First stage estimates and 95% confidence limits



SECOND STAGE MODEL DEVELOPMENT

Site potential

For Equation VIII.2a without any thinning parameters, Bartlett's test (1937) showed that the assumption of homogeneity was invalid (chi-square 404.5) and that the parameter estimates were quite highly correlated, the correlation between p and a_1 being 0.77.

Various forms of the second stage models were then tried for Equation VIII.2a without any thinning parameters, and the initial trials are summarised in Table VIII.1. Using a linear structure in Y_{10} for p (Model 8) was inferior to the logarithmic form (Model 1). A quadratic form in site potential (Model 10) was significantly better than the linear, but variants of this model (Models 11–15) did not show significant gains. The incorporation of the quadratic parameter gave only a slight gain over the logarithmic form. Moreover, the quadratic implied that a minimum value of p was reached at $Y_{10} = 259$, which was within the range of the data, some 7 plots having Y_{10} values in excess of this figure. No tenable argument could be advanced to justify the acceptance of a minimum value within the range of the data, and hence it was rejected.

The posterior model was therefore the simple model Equation VIII.3.

$$p = \underset{(0.00411)}{0.05271} - \underset{(0.000821)}{0.006484} \ln(Y_{10}) \qquad \text{(VIII.3)}$$

$$a_1 = \underset{(0.000151)}{-0.003467}\, Y_{10}$$

Various second stage models (linear functions in Y_{10}) were then fitted to the first stage parameter estimates for Equations VIII.2b and VIII.2c each with a single thinning parameter (three parameters in total). All models had nonsignificant zeta values (see Appendix 6) for the thinning parameter, with the highest zeta value being 0.3, considerably less than the critical value of 1.96. This confirmed the suspicion that 7 or 9

Table VIII.1

Second Stage Models,

Posterior, conditioned Bertalanffy

Model	First stage variables									Residual sum squares
	p					a_1				
	Secondary variables					Secondary variables				
	1	$\ln(Y_{10})$	Y_{10}	Y_{10}^2	N_{10}	1	Y_{10}	Y_{10}^2	N_{10}	
1	*	*					*			64.46
2	*	*			NS		*			64.41
3	*	*				NS	*			64.40
4	*	*				(NS)	NS	NS		63.02
5	*	*					*	NS		64.31
6	*	*					*		NS	64.09
7	*	*				NS	*		(NS)	63.79
8	*		*				*			72.04
9	*		*		NS		*			71.73
10	*		*	*			*			60.10
11	*		*	*	NS		*			59.59
12	*		*	*		NS	*			59.72
13	*		*	*		(NS)	NS	NS		59.53
14	*		*	*		NS	*		NS	59.17
15	*		*	*			*		NS	59.26

Note
- * parameter significant
- NS parameter not significant
- (NS) parameter not significant and zeta value the lowest of all the estimates.

plots with significant thinning parameters out of the 58 plots were too few for satisfactory second stage models to be developed. Equation VIII.3 was accepted as the best model to date and thinning parameters were not included in the model.

Other variables

Stocking at age ten (N_{10}) was evaluated in addition to site potential in the second stage models, but Models 2, 6, 7, 9, 11 and 15, Table VIII.1, show that the variable was not significant.

The data could not be divided into the same seven groups based on soil type used in the OLS analysis because some groups had too few observations. The data were therefore divided into four soil groups based on the groups defined in the OLS estimation.

1 Group C, the main group of sandy soils, 33 plots.

2 Group B, Caroline and Wandilo sands, 17 plots.

3 Group D, Tantanoola flinty sand, 4 plots.

4 Miscellaneous, the other groups (A, E, F and G), 4 plots.

Although the fourth group is heterogeneous in terms of soil type it enabled the other groups to be separated and analysed.

When any one of the three parameters in Equation VIII.3 was replaced by parameters for each soil type and all parameters were re-estimated, increasing the total number of parameters to six, the residual sum squares was reduced by less than 2% (Table VIII.2) and none of the four soil varying parameters differed from any other regardless of which was replaced. Soil type could not be included in the model. The comparison between Equations VII.8 and V.12 in Chapter VII indicated that the standard errors of the parameters in the OLS estimation were underestimated by a factor of between 3.5 and 6. Assuming that this also holds for Equation VI.12 it was evident that that equation may have too many param-

Table VIII.2

Second stage models

Posterior; conditioned Bertalanffy by soil type and forest.

Model Number	Parameter	Varying with	Number of second stage parameters	Residual sum squares	Comparison with	
					Model	Significance
1			3	64.46		
16	p_0	Soil type	6	63.19	1	NS
17	p_1		6	63.22	1	NS
18	a_1		6	63.35	1	NS
19	p_0	Forest	7	61.85	1	NS
20	p_1		7	61.91	1	NS
21	a_1		7	62.85	1	NS

eters, and thus the inability of the GLS analysis to discriminate between groups of soil types seemed quite reasonable.

The forest groupings defined in Appendix 1.6a were evaluated in the same way (Table VIII.2). The yield curves do not vary between forests.

As form variables were not significant in the OLS analysis where the standard errors of the parameters were underestimated they were not even considered in the GLS analysis.

Alternative error structures

Further analysis of this posterior model with alternative error structures (Table VIII.3) showed that the implicit recognition of the correlation between the parameters gave an improvement in relative efficiency: relaxing the assumption of correlation between parameters reduced the relative efficiency to 0.157. However the advantage of explicitly recognizing the variance heterogeneity was minimal and was totally obscured when the assumption of correlation was also relaxed.

EXAMINATION OF THE DEVELOPMENTAL DATA

The posterior model (Equation VIII.3) and the data are plotted in Figures VIII.2a and VIII.2b keeping the data separate by forest area.

Of the thirteen Mount Burr plots, yields were consistently overestimated for two plots, 55 and 123, and were overestimated at later ages, for two other plots, 57 and EP24B. EP24D was estimated satisfactorily although it is adjacent to EP24B. EP24D was planted at 9x9 feet and EP24B at 6x6 feet, so the difference seemed attributable to spacing. However the range of initial spacing in the data was too narrow for this variable to have a significant effect in the GLS analysis.

Table VIII.3

Second stage models

Posterior; conditioned Bertalanffy, comparison of alternative error structures

Model	Joint dependent variables			Relative efficiency
	p		a_1	
	Secondary variables			
	1	$Ln(Y_{10})$	Y_{10}	
Heterogeneous correlated	0.05271 ($1.689\ 10^{-5}$)	−0.006481 ($6.746\ 10^{-7}$)	−0.003467 ($2.293\ 10^{-8}$)	1.0
Homogeneous correlated	0.05059 ($1.556\ 10^{-5}$)	−0.006065 ($6.236\ 10^{-7}$)	−0.003457 ($2.290\ 10^{-8}$)	0.964
Heterogeneous independent	0.04984 ($3.682\ 10^{-5}$)	−0.005930 ($13.982\ 10^{-7}$)	−0.003446 ($2.322\ 10^{-8}$)	0.157
Homogeneous independent	0.04832 ($3.463\ 10^{-5}$)	−0.005633 ($13.128\ 10^{-7}$)	−0.003448 ($2.308\ 10^{-8}$)	0.162

Estimated variances shown in brackets below the coefficient

Figure VIII.2a

Posterior: estimated yield functions for developmental data

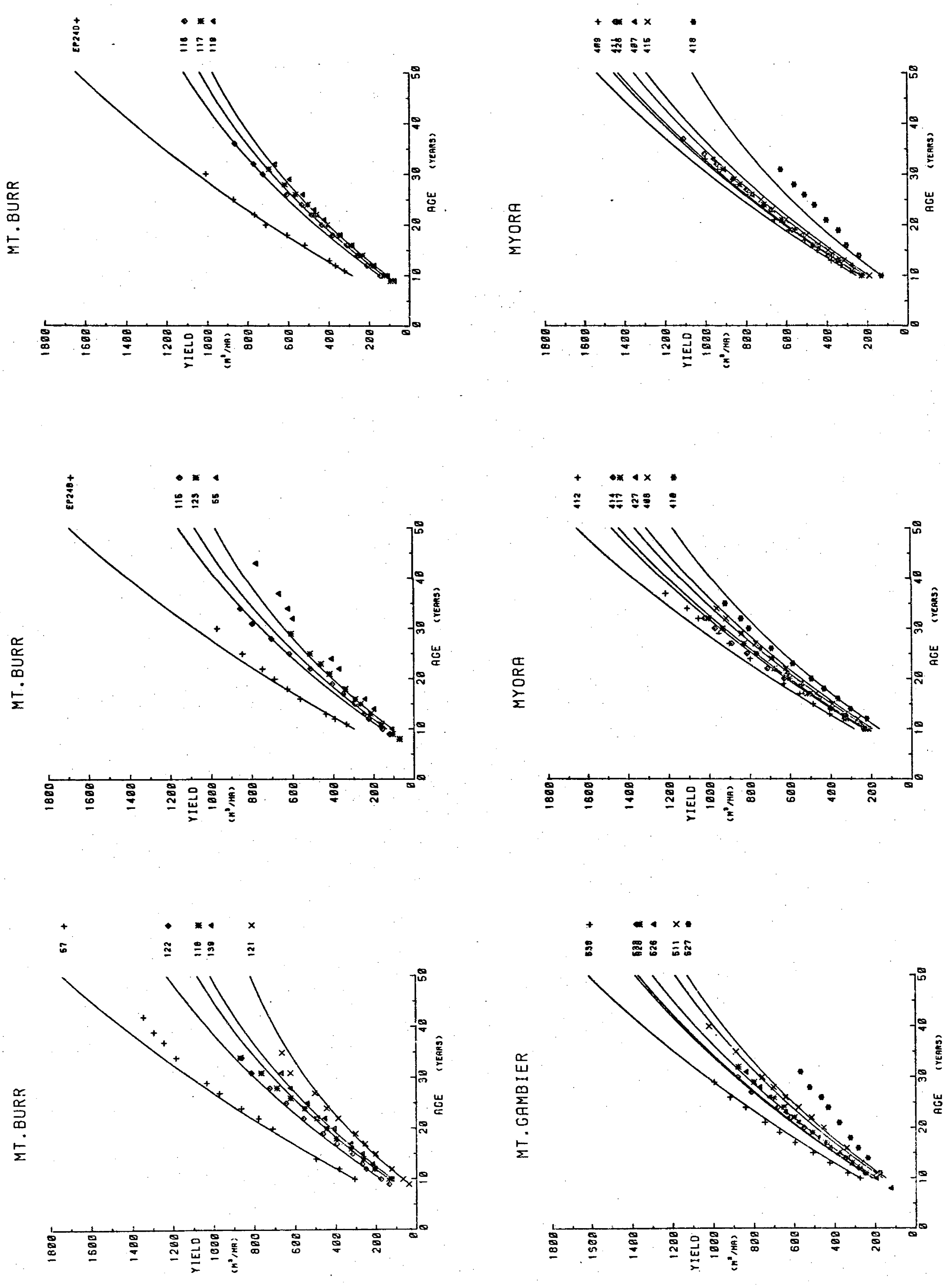
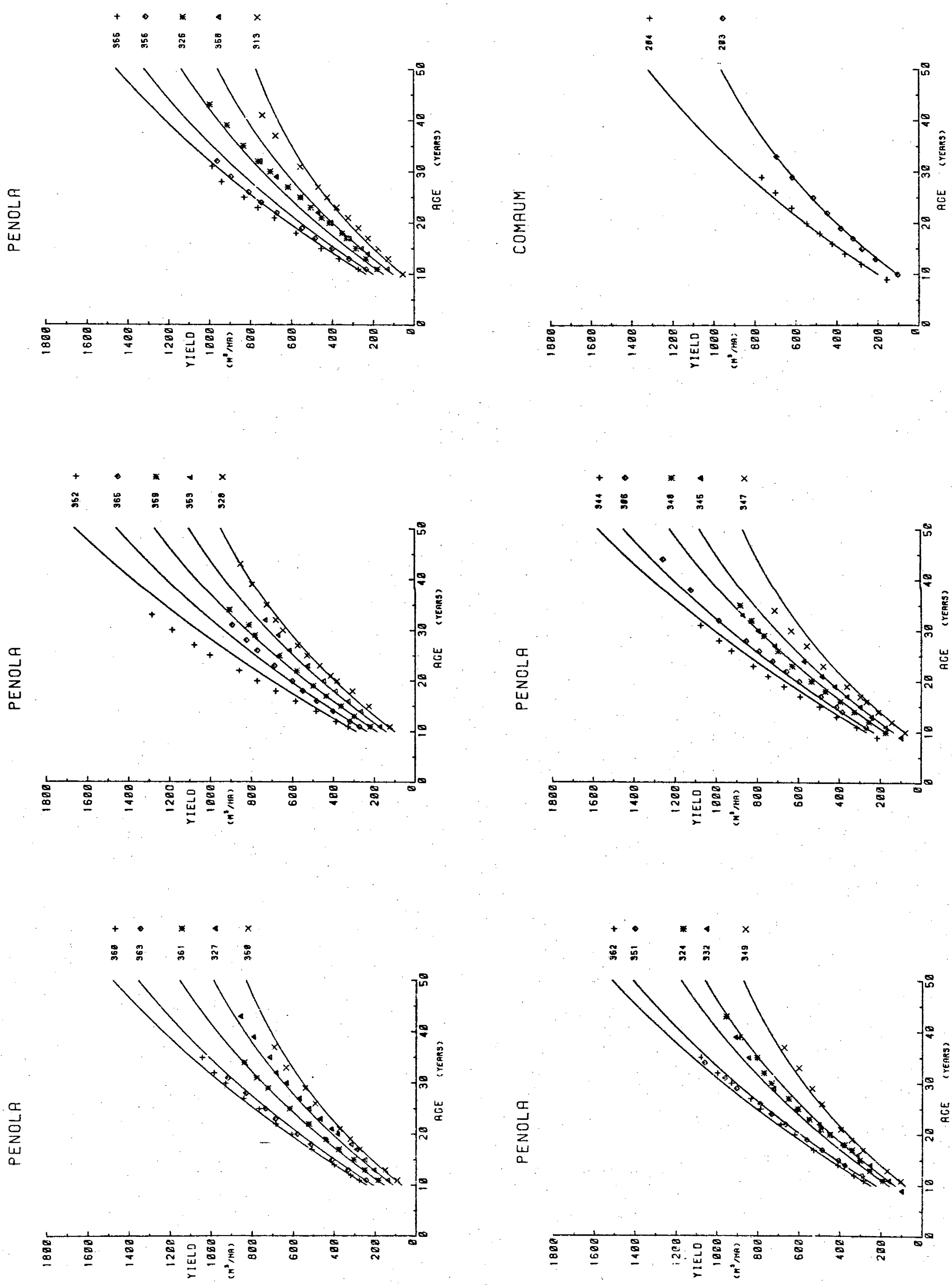


Figure VIII.2b

Posterior: estimated yield functions for developmental data



Five of the six Mount Gambier plots were fitted satisfactorily. In the remaining plot (527) yield was overestimated consistently.

The yield curves overestimated for plots 412 and 418 at Myora, but plot 418 had an inexplicable discontinuity in the actual trend which may be anomalous. Plot 409 at Myora was overestimated at later ages and this plot also had an inexplicable discontinuity.

For the twenty five Penola plots, yields were consistently under-estimated for two plots (352 and 356). The few that were underestimated at later ages (313, 345 and 347) were balanced by others (365 and 324) which were overestimated at later ages. The actual shape of the yield function itself was not really suited to plots such as 325 and 313.

Of the two Comaum plots the yield function overestimates slightly for plot 204 at later ages.

In general the anomalous plots on all forests have similar characteristics. Yields for plots with higher initial stocking or for plots on shallow heavy soils were generally overestimated at later ages. Underestimates seemed to occur where plots have access to a shallow water table, or where plots are located on soils that overly a volcanic base, possibly providing access to more nutrients as the tree roots penetrate the deep sands. The GLS analysis failed to pick up these trends because the soil differences were not well represented in the data.

EXAMINATION OF THE INDEPENDENT TEST DATA

Independent test data with at least 5 measurements for volume over at least a 15 year growth period, and with measurements at or near age 10, were then extracted from the data base. Equation VIII.3 was fitted to these independent test data. The data and the estimated yield function are plotted in Figure VIII.3a and VIII.3b by forest area.

Figure VIII.3a

Posterior: estimated yield functions for test data

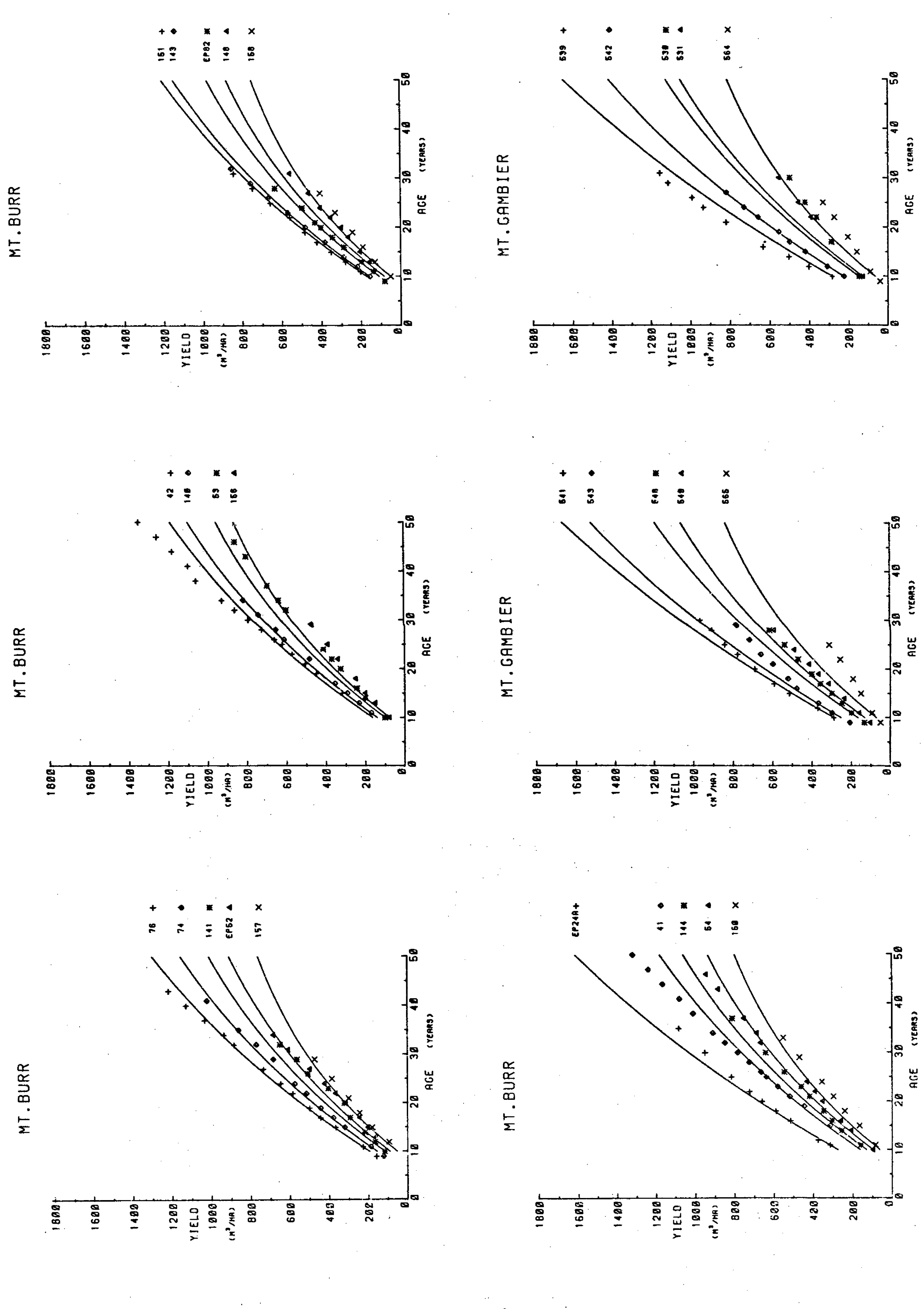
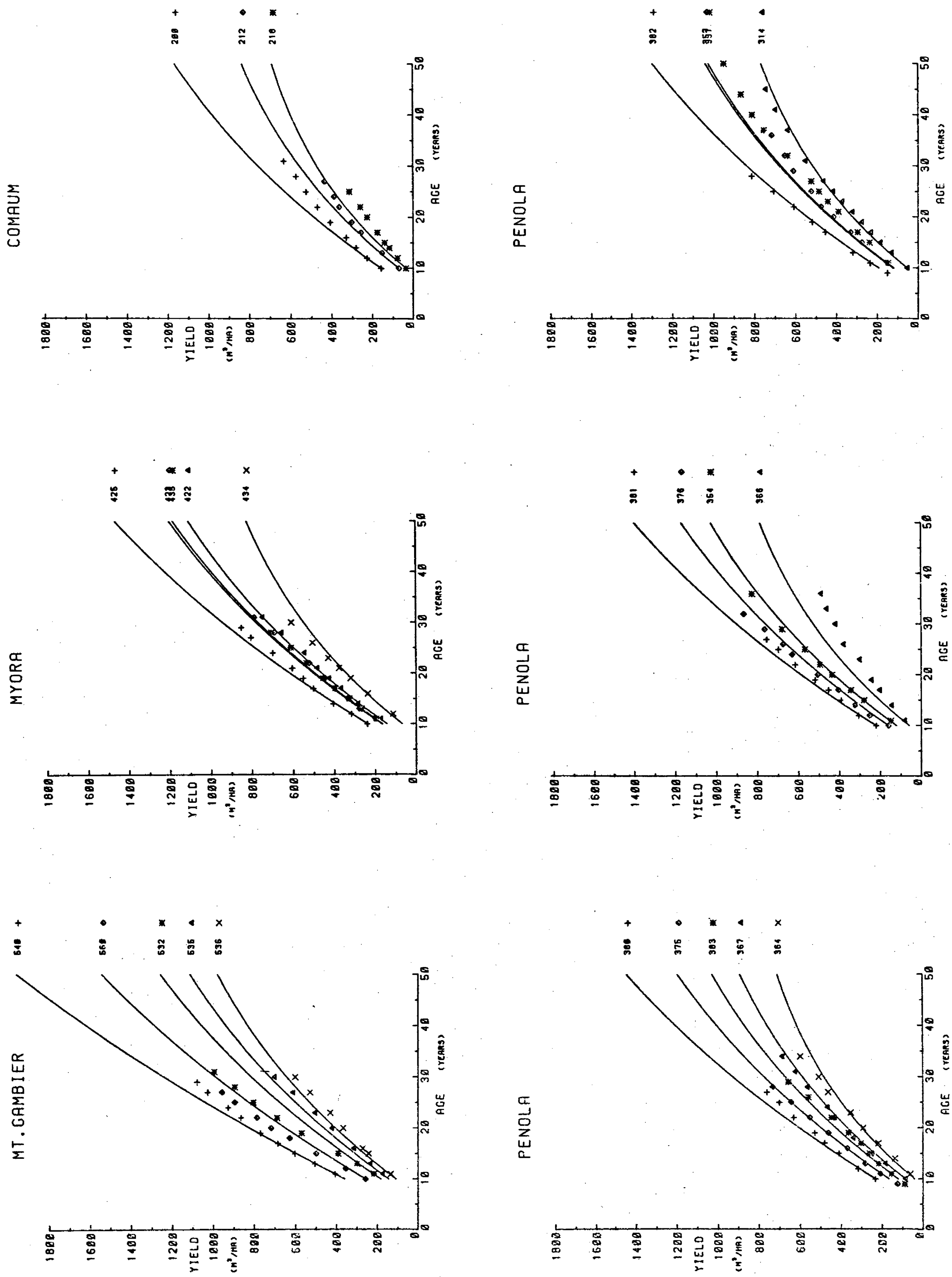


Figure VIII.3b

Posterior: estimated yield functions for test data



When the data were pooled the mean deviate was significantly different from zero, with Equation VIII.3 underestimating yield. For the 378 observations the mean deviate of -7.46 had a standard error of 2.39. This was rather disturbing and seemed too great to be attributable solely to chance. On investigation it was found that the way in which the two data sets were differentiated on number of measurements and on the length of growth period available approximately split the data into two historically different data sets in the pre-1940 plantings and the post-1940 plantings.

Over the years forest practice has changed gradually and this is reflected in two general but important differences between the developmental data and the independent test data. Firstly the tendency has been for the thinning regimes in current sample plot practice to have become heavier over the years and thus the plots in the test data generally have heavier thinning regimes than the plots in the developmental data. The OLS analysis indicated that increment, and hence yield, is slightly lower for heavily thinned stands and although thinning parameters were not significant in the GLS analysis this possibly explains the significantly lower mean deviate for the test data.

Secondly, as the plantation program expanded, the range of soil types planted changed and so did the range in the sample plot series. For example, of the soil groups recognised in the OLS analysis the terra rossa soils with the highest estimated value of the parameter p_0 were represented by only one plot out of 58 in the developmental data, but by six out of 55 in the test data. The change in the distribution of soil type between the two data sets is likely to aggravate the thinning effect, explaining why the mean deviate for the test data was significantly different from zero. This was possible even though neither thinning nor soil type were significant in the GLS analysis because the range of these variables was somewhat reduced.

SUMMARY

The GLS analysis, unlike the OLS analysis, satisfied the statistical assumptions that could be evaluated. Analysis of the first stage models used to develop the prior in Chapter VII indicated that serial correlation was not a problem. Because a Bayesian approach was adopted, tests of serial correlation were irrelevant for the first stage posterior models because these models were influenced by the informative prior. The GLS analysis explicitly took account of the plot induced heterogeneity that the OLS analysis had indicated was the most important of the three sources considered in Chapter III.

The GLS analysis of the Bertalanffy model form produced a simple conditioned yield model (Equation VIII.3). Although superior to the OLS analysis, the behaviour of the independent test data suggests that caution should be exercised in its use in practice, and further it indicates the necessity for the model to be re-developed as more measurements are available for the plots that were used in the independent test data.

IX CONCLUSIONS

IX CONCLUSIONS

The development of this study fell naturally into three separate phases.

In the first phase (Chapter V), Ordinary Least Squares (OLS) techniques were used to estimate and compare a variety of different models using data from unthinned permanent sample plots. This phase centred around a comparison of the Bertalanffy model with other models such as the Bednarz, Johnson-Schumacher (or Clutter), Gompertz and polynomial models. The Bertalanffy model seemed marginally preferable to the Bednarz model, all others being clearly inferior.

A number of different forms of the Bertalanffy model were evaluated including both the Mitscherlich (or monomolecular) and Chapman-Richards forms, and variants of them based on yield, the derivative of yield with respect to age or periodic annual increment as the dependent variables, with or without conditioning through the value of site potential. The results suggested that the allometric constant (r) for the catabolic destruction rate in the Bertalanffy model could be taken as 1.0 while the allometric constant (m) for the anabolic growth rate could be taken as zero.

This Mitscherlich or monomolecular form was preferred, even though it is not sigmoidal in shape. The absence of the point of inflection may reflect the limitations of the data, which did not span very young ages, or it may reflect the actual properties of yield when measured in terms of volume to 10cm top diameter underbark. There are an infinite number of transcendental functions with sigmoidal properties, although few can be fitted with so few parameters, and the sequence: general Bertalanffy, second level Bertalanffy (or Chapman-Richards), Mitscherlich (or monomolecular), provides a logical series of model forms of decreasing complexity.

The conditioned periodic annual increment version of the Mitscherlich form proved to be a superior predictor to its unconditioned analogue and to any of the yield or derivative forms. Unlike linear models, a conditioned nonlinear model is not necessarily an inferior estimator to unconditioned forms with more parameters.

The model currently used for yield prediction, the graphically defined yield table of Lewis, was shown to be a satisfactory predictor but was open to bias and necessarily lacked an objective measure of precision.

In the second phase of the study (Chapter VI), Ordinary Least Squares (OLS) techniques were used to extend the earlier results of the conditioned periodic annual increment model for the unthinned stands to include other stand variables such as those relating to thinning and soil type. Thinning was taken into account in two ways. Firstly, a variable representing the level of competition was incorporated in a manner which approximated the Langsaeter or Moller hypothesis regarding the effects of thinning. Secondly, a variable representing thinning shock was incorporated, essentially as an overriding correction factor to the model. Seven groups of soil types were introduced by appropriate definition of dummy variables and incorporation of these into the model. Form was also investigated but the variables introduced were not found to be useful.

Although some of the estimated parameters were not significantly different from zero, the following model seemed to be the most appropriate of those tested.

$$Pai = \left\{ (Y_{A+i} + Y_A) / i \right\} \quad Z_1 \qquad (VI.2)$$

$$Y_A = Y_{10} \left\{ \frac{1 - \exp(-p(A - A_0))}{1 - \exp(-p(10 - A_0))} \right\} \qquad (VI.7)$$

where

$$p = p_0 + \underset{(0.1164\ 10^{-4})}{0.5855\ 10^{-4}}\ Y_{10} \qquad (VI.12)$$

$p_0 = \underset{(0.00093)}{0.00030}$ for volcanic soils

$p_0 = \underset{(0.00052)}{0.00302}$ for Caroline, Wandilo and Myora sands

$p_0 = \underset{(0.00036)}{0.00539}$ for other yellow and white sands

$p_0 = \underset{(0.00131)}{0.01119}$ for Tantanoola flinty sands

$p_0 = \underset{(0.00201)}{0.01759}$ for terra rossa soils and brown soils from Comaum

$p_0 = \underset{(0.00225)}{0.01172}$ for yellow sands from Comaum

$p_0 = \underset{(0.00175)}{0.00534}$ for rendzinas

$p_0 = \underset{(0.002623)}{0.005075}$ for all soil types combined

$$A_0 = 10.0 \exp(\underset{(0.000841)}{-0.009172}\ Y_{10})$$

$$Z_1 = (\underset{(0.114)}{1.700} + \underset{(0.2335\ 10^{-3})}{0.4426\ 10^{-3}}\ D - \underset{(0.3050\ 10^{-6})}{0.7380\ 10^{-6}}\ D^2)\ (1.0 - \underset{(0.0731)}{0.4287}\ S)$$

and where

Pai = periodic annual increment,

Y_A = yield at age A,

A = age,

i = increment period,

Y_{10} = site potential,

D = competition level (volume), and,

S = thinning shock, $(Vt/(Vb\ i))$, relative volume, the effect lasting one year only.

This model was tested using data from other regions. For the three forests in the Adelaide Hills region where reasonably good data were available, the model was a satisfactory predictor. For the small outlying Forest Reserves of Noolook and Cave Range in the upper part of the south-east region and for Bundaleer and Wirrabara Forest Reserves in the Northern region, the model overestimated, but the data were too sparse to allow detailed investigation. Simple correction factors to increment were developed for use until better data are available.

Importantly to South Australia it was shown that second rotation stands can be considered to have the same yield function as first rotation stands.

In the third phase (Chapters VII and VIII) Generalized Least Squares (GLS) techniques were introduced to overcome statistical defects in the OLS analyses. A yield form of the Mitscherlich model was fitted to each of the plots from unthinned stands in turn. The resulting parameter estimates were then related to differences between the stands, notably in terms of site potential.

The resulting function for unthinned stands was then used as an informative prior in a Bayesian analysis of the data from thinned stands, on a plot by plot basis. The analysis included various models of the effect of thinning but no thinning variables were significant. The parameter estimates were again related to differences between the stands such as site potential and soil type using GLS. The use of the Bayesian analysis enabled a sequential approach to model building to be adopted, in an effort to avoid likely problems from misspecification otherwise introduced by trying to estimate too many parameters from too few data from each plot.

The GLS analysis produced a far simpler model than the OLS analysis because thinning and soil variables were omitted, having failed to yield parameter estimates significantly different from zero.

$$Y_A = Y_{10} \left\{ \frac{1 - \exp(-p(A - 10\exp(-a_1)))}{1 - \exp(-p(10 - 10\exp(-a_1)))} \right\} \qquad \text{(IX.1)}$$

where

$$p = \underset{(0.00411)}{0.05271} - \underset{(0.000821)}{0.006484} \ln(Y_{10}) \qquad \text{(VIII.3)}$$

$$a_1 = \underset{(0.000151)}{-0.003467}\, Y_{10}$$

and where

Y_A = yield at age A,

A = age, and,

Y_{10} = site potential, yield at age 10.

Any comparison of OLS and GLS results must necessarily be somewhat equivocal in the light of the complexity of the models being studied and the inadequacies of the data. Nevertheless, some points need to be stressed in comparing the OLS and GLS results.

Firstly, the OLS technique probably yielded biased estimates of the standard errors of the parameters, underestimating the true values substantially. This casts considerable doubt on the entire sequence of hypothesis testing of a particular model in moving from one model form to the final form accepted. Thus, while the model summarized in Equations VI.2, VI.7 and VI.12 seems appealing, the statistical basis of that model is questionable.

On the other hand, the data used to develop the GLS model did not adequately cover the range of soil types and thinning intensities. The omission of these variables from the GLS model may reflect inadequacies of the data or may be well-founded: only further data and analysis can provide an answer to this.

On balance, the GLS technique seems to offer greater advantages for future work of this kind. The ability to build models sequentially in a Bayesian framework has definite advantages in clarifying the form of the model and reducing the number of alternatives to be tried. The testing of hypotheses is also placed on a sounder basis than that for the OLS analyses as used in this study.

The GLS technique can also be expanded to cater for simultaneous models of other dependent variables beside yield, such as height and basal area. The joint estimation of such models would enable the correlations between these variables to be taken into account and used to improve the efficiency of the estimates. Some preliminary work was carried out along these lines and the results seem promising, but further modifications to the programs are required before the analysis can be completed.

BIBLIOGRAPHY

Acton, F.S. (1959) Analysis of straight line data. Dover 267pp.

Assmann, E. (1970) The principles of forest yield study. Pergamon Press 506pp.

Austin, R.B., Nelder, J.A. and Berry, G. (1964) The use of a mathematical model for the analysis of manurial and weather effects on the growth of carrots. Ann. Bot. 28 (109) p153-162.

? Australia (1977) Woodchips and the environment. Report from the Senate Standing Committee on Science and Environment. AGPS Canberra 454pp.

Australia, Department of Agriculture (1976) Forestry and Timber Bureau Annual Report 1974/75. AGPS Canberra 71pp.

Australian Forestry Council (1974) Report of the Forestry and Wood Based Industries Development Conference. AGPS Canberra 94pp.

Bailey, R.L. and Clutter, J.L. (1974) Base age invariant polymorphic site curves. For Sci. 20 (2) p155-159.

Bard, Y. (1967) Nonlinear parameter estimation and programming. IBM Contributed Program Library 154pp.

Bartlett, M.S. (1937) Propoerties of sufficiency and statistical tests. Proc. Roy. Soc. London Series A 160 p268-282.

Bartlett, M.S. (1954) A note on multiplying factors for various chi-squared approximations. J. Roy. Stat. Soc. Series B 16 p296-298.

Bayes, T. (1763) Solving a problem in the doctrine of chances. Phil. Trans. Roy. Soc. London 53 p370-418, reprinted (1958) Biometrika 45 p293-315.

Beck, D.E., (1971) Polymorphic site index curves for white pine in the southern Appalachians. USFS Res. Paper SO-80 8pp.

Bednall, B.H. (1957) Forestry handbook. Woods and Forests Bull. 6 104pp.

Bednarz, R.W. (1975) A new tree growth function with application to forecasting predominant height. Draft report for APM Forests Pty. Ltd. by Numerical Data Sciences. 15pp.

Bertalanffy, L. von (1941) Untersuchungen uber die gesetzlich beit des wachstums. VII Stoffwechseltypen und wachstumstypen. Biol. Zbl. 61 510-532.

Bertalanffy, L. von (1942) Theoretische biologie. II Stoffwechsel, wachstum. Gebruder Borntraeger, Berlin 362pp.

Bertalanffy, L. von (1957) Quantitative laws in metabolism and growth. Quart. Rev. Biol. 32 (3) p217-232.

Bertalanffy, L. von (1969) General system theory. George Braziller N.Y. 289pp.

Bevege, D.I. (1972) Girth and basal area growth of slash pine (P.elliottii var elliottii Engl.) and their significance to plantation management in Queensland. Qld. Dept. For., Res. Paper 2 69pp.

Box, G.E.P. and Tiao, G.C. (1973) Bayesian inference in statistical analysis. Addison-Wesley 588pp.

Braathe, P. (1957) Thinnings in even aged stands. Fac. For. U. of New Brunswick 92pp.

Brickell, J.E. (1968) A method of constructing site index curves from measurements of tree age and height-its application to inland Douglas Fir. USFS Res. Paper INT-47 23pp.

Buckman, R.E. (1962) Growth and yield of red pine in Minnesota. USFS Tech. Bull. 1272 50pp.

Carroll, C.W. (1961) The created response surface technique for optimizing nonlinear restrained systems. Op. Res. 9(12) p169-184.

Chapman, D.G. (1961) Statistical problems in dynamics of exploited fisheries populations. Proc. 4th Berkeley Symp. on Math. Stats. and Prob. Uni. of Cal. Press IV p153-168.

Cilliers, A.C. and van Wyk, J.H. (1939) A mathematical expression for the growth of trees and their dependence on time and density of stocking. Annals U. of Stellenbosch XVI Series A 2 36pp.

Clutter, J.L. (1963) Compatible growth and yield models. For. Sci. 9 (3) p354-371.

Cochran, W.G. (1941) The distribution of the largest of a set of estimated variances as a function of their total. Ann. Eugenics 11 p45-52.

Cochran, W.G. and Cox, G.M. (1957) Experimental designs. Wiley Int. edition, 2nd edition, 615pp.

Cochrane, D. and Orcutt, G.H. (1949) Application of least squares regression to relationships containing autocorrelated error terms. J. Am. Stat. Assoc. 44 p32-61.

Cox, D.R. (1961) Tests of separate families of hypotheses. Proc. 4th Berkeley Symp. on Math. Stats. and Prob., Uni. of Cal. Press I p105-123.

Cox, D.R. (1962) Further results on tests of separate families of hypotheses. J. Roy. Stat. Soc. Series B (24) p406-424.

Cunia, T. (1964) Weighted least squares method and construction of volume tables. For. Sci. 10 (2) p180-192.

Cunia, T. (1973) Dummy variable and some of their uses in regression analysis. Proc. IUFRO Working Group S4.02 Nancy France 2 p1-146.

D'Agostino, R.B. (1971) An omnibus test of normality for moderate and large sized samples. Biometrika 58 (2) p341-348.

Deutsch, R. (1965) Estimation theory. Prentice Hall 269pp.

Douglas, D.R. (1974) Forestry and associated industry in the Mount Gambier (S.A.) region. Woods and Forests Bull. 15, 3rd edition, 32pp.

Draper, N.R. and Smith, H. (1966) Applied regression analysis. Wiley 407pp.

Durbin, J. and Watson, G.S. (1950) Testing for serial correlation in least-squares regressions. I. Biometrika 37 p409-428.

Durbin, J. and Watson, G.S. (1951) Testing for serial correlation in least-squares regressions. II. Biometrika 38 p159-178.

Durbin, J. and Watson, G.S. (1971) Testing for serial correlation in least-squares regressions. III. Biometrika 58 p1-19.

Efroymson, M.A. (1962) Multiple regression analysis. Article 17 p191-203 in Ralston, A. and Wilf, H.S., eds. Mathematical methods for digital computers. Wiley New York 293pp.

Eisenpress, H. and Greenstadt, J. (1966) The estimation of nonlinear econometric systems. IBM New York Sci. Centre Report 320.2910.

Ferguson, I.S. and Leech, J.W. (1976a) Stand dynamics and density in radiata pine plantations. N.Z. J. Forestry Science 6 (3) p443-454.

Ferguson, I.S. and Leech, J.W. (1976b) Generalized least squares estimation of stand yield functions. Proc. Res. Working Group No.2: Mensuration and Management, Melbourne Victoria 22pp.

Fielding, J.M. and Millett, M.R.O. (1941) Some studies of the growth of Monterey pine. F&TB Bull. 27 31pp.

Fisher, R.A. (1921) The mathematical foundations of theoretical statistics. Roy. Soc. London Phil. Trans. Series A 222 p309-368.

Fisher, R.A. (1925) Theory of statistical estimation. Proc. Camb. Phil. Soc. 22 p700-725.

Fletcher, R. and Powell, M.J.D. (1963) A rapidly converging descent method for minimization. Comp. J. 6(2) p163-168.

Ford-Robertson, F.C. (1971) Terminology of forest science, technology practice and products. English language version. Soc. Am. For. 349pp.

Freese, F. (1964) Linear regression methods for forest research. USFS Res. Paper FPL-17 136pp.

Furnival, G.M. and Wilson, W.W. jnr. (1974) Regression by leaps and bounds. Technometrics 16 (4) p499-511.

Gaertner, E.E. (1964) Tree growth in relation to the environment. Bot. Rev. 30 p393-436.

Gallant, A.R. (1975) Nonlinear regression. Am. Statistician 29 (2) p73-81.

Gentle, S.W., Henry, J.L. and Shepherd, K. (1962) The application of basal area control to thinning of Pinus radiata plantations in New South Wales. 8th Br. Comm. For. Conf. East Africa 16pp.

Glock, W.S. (1955) Tree growth. II: Growth rings and climate. Bot. Rev. 21 p13-188.

Goldberger, A.S. (1964) Econometric theory. Wiley 399pp.

Goldfeld, S.M. and Quandt, R.E. (1972) Nonlinear methods in econometrics. North Holland 280pp.

Gompertz, B. (1825) On the nature of the function expressive of the law of human mortality and on a new mode of determining the value of life contingencies. Phil. Trans. Roy. Soc. London 115 (1) p513-585.

Graybill, F.A. (1961) Introduction to linear statistical models. McGraw Hill 463pp.

Grosenbaugh, L.R. (1965) Generalization and reparameterization of some sigmoid and other nonlinear functions. Biometrics 21 p708-714.

Grosenbaugh, L.R. (1967) REX--Fortran-4 system for combinatorial screening or conventional analysis of multivariate regressions. USFS Res. Paper PSW-44 47pp.

Grut, M. (1970) Pinus radiata-growth and economics. A.A. Balkema Capetown 234pp.

Guttman, I. and Meeter, D.A. (1965) On Beales measure of non-linearity. Technometrics 7 p623-637.

Hartley, H.O. (1950) Maximum F-ratio as a short cut test for heterogeneity of variance. Biometrika 37 p308-312.

Huxley, J. (1932) Problems of relative growth. Methuen London 276pp.

Jerram, M.R.K. (1939) Elementary forest mensuration. Thomas Murby & Co. 124pp.

Joergensen, C. (1957) Thinning experiments. Br. Col. For. Serv. Publ. T45 24pp.

Johnson, N.O. (1935) A trend line for growth series. J. Am. Stat. Assoc. 30 p717.

Johnston, J. (1963) Econometric methods. McGraw Hill 300pp.

Jolly, N.W. (1950) The volume line theory in relation to the measurement of standing volume of a forest (with particular reference to Pinus radiata). Woods and Forests Dept. Misc. Publ. 46pp.

Keeves, A. (1961) A consideration of certain mensuration methods for intensive forest management of coniferous plantations in Australia. Dissertation for Dip. For. U. of Oxford (unpubl).

Keeves, A. (1966) Some evidence of loss of productivity with successive rotations of Pinus radiata in the south-east of South Australia. Aust. For. 30 (1) p51-63.

Keeves, A. (1970) Regulation of yield. Radiata pine symposium, Canberra, 12pp.

Kendall, M.G. and Stuart, A. (1961) The advanced theory of statistics. II Inference and relationship. Griffin 676pp.

Kira, T., Ogawa, H. and Sakazaki, N. (1953) Intraspecific competition amongst higher plants. I Competition-yield-density inter-relationships in regularly dispersed populations. J. Inst. Polytech. Osaka City Univ. 4 (D) p1-16.

Kowalski, C.J. and Guire, K.E. (1974) Longitudinal data analysis. Growth 38 p131-169.

Laird, A.K. (1965) Dynamics of relative growth. Growth 29 p249-263.

Langsaeter, A. (1941) Om tynning i enraldret gran- og furuskog. Meddeleleser f det Norske skogforsokvesen. 8 (1) p131-216.

Leech, J.W. (1973) Tree volume and increment models for radiata pine thinnings. M.Sc. Thesis ANU Canberra (unpubl).

Lehmann, E.L. (1959) Testing statistical hypotheses. Wiley 369pp.

Lewis, N.B. (1959) Modern theory and practice of thinnings in coniferous plantations. Dissertation for Dip. For. U. of Oxford (unpubl).

Lewis, N.B. (1962) Some effects of thinning on value production. Aust. For. 26 (1) p33,34.

Lewis, N.B. (1963) Optimum thinning range of Pinus radiata in South Australia. Aust. For. 27 (2) p113-120.

Lewis, N.B., Keeves, A. and Leech, J.W. (1976) Yield regulation in South Australian Pinus radiata plantations. Woods and Forests Dept. Bull. 23 174pp.

Lewis, N.B. and McIntyre, G. (1963) Regional volume table for Pinus radiata in South Australia. Woods and Forests Dept. Bull. 14 59pp.

Lewis, N.B., McIntyre, G. and Leech, J.W. (1973) Regional volume table for Pinus radiata in South Australia, metric edition. Woods and Forests Dept. Bull. 20 130pp.

Lindley, D.V. (1972) Bayesian statistics, a review. Regional Conference series in Applied Maths. SIAM 2 83pp.

Marquardt, D.W. (1963) An algorithm for least squares estimation of non-linear parameters. J. Soc. Ind. Appl. Math. 11 (2) p431-441.

Mitscherlich, E.A. (1910) Das gesetz des minimums und das gesetz des abnehmenden bodenertrages. Landwirtsch. Jahrbucher 38 p537-552.

Moller, C.M. (1954a) The influence of thinning on volume increment. in Heiberg, S.O. ed. Thinnings, problems and practices in Denmark. State Uni. of New York Coll. For. Tech. Publ. 76 p5-49.

Moller, C.M. (1954b) Gründflachenzuwachs und massenzuwachs mit verschiedenen definitionen. Forstwirtenschaftliches Centralblatt 73 p350-354.

Morrison, D.F. (1976) Multivariate statistical methods. 2nd ed. McGraw Hill 415pp.

Nelder, J.A. (1961) The fitting of a generalization of the logistic curve. Biometrics 17 p89-110.

Nelder, J.A. (1962) An alternative form of a generalized logistic equation. Biometrics 18 p614-616.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K. and Bent, D.H. (1975) SPSS. Statistical package for the social sciences manual. 2nd ed. McGraw Hill 675pp.

Pawsey, C.K. (1964) Height and diameter growth cycles in P. radiata. Aust. For. Res. 1 (1) p3-8.

Pearl, R.P. and Reed, L.J. (1923) On the mathematical theory of population growth. Metron. 3 p6-19.

Pienaar, L.V. (1966) A non linear biomathematical growth model for even aged plantations. Forestry in South Africa 7 p117-124.

Pienaar, L.V. and Turnbull, K.J. (1973) The Chapman-Richards generalization of von Bertalanffy's growth model for basal area growth and yield in even-aged stands. For. Sci. 19 (1) p2-22.

Prodan, M. (1968) Forest biometrics. Pergamon press 447pp.

Raiffa, H. and Schlaifer, R. (1961) Applied statistical decision theory. Harvard Uni. Press 356pp.

Ramsey, J.B. (1969) Tests for specification errors in classical least-squares regression analysis. J. Roy. Stat. Soc. Series B (2) p350-371.

Ramsey, J.B. (1974) Classical model selection through specification error tests. p13-47 in Zarembka, P. (1974) Frontiers of econometrics. Academic press 252pp.

Rawat, A.S. and Franz, F. (1974) Detailed nonlinear asymptotic regression studies on tree and stand growth with particular reference to forest yield research in Bavaria (Federal Republic of Germany) and India. Institutionen för Skogsproduktion, Rapporter och Uppsatser 30 p180-221.

Richards, F.J. (1959) A flexible growth function for empirical use. J. Exp. Bot. 10 (29) p290-300.

Robinson, W.M. (1968) Pruning and thinning practice in Queensland plantations. 9th Comm. For. Conf. 21pp.

Sadler, D.R. (1975) Numerical methods for nonlinear regression. U. of Qld. press 89pp.

Schumacher, F.X. (1939) A new growth curve and its relation to timber yield studies. J. For. 37 p819,820.

Shapiro, S.S. and Wilk, M.B. (1965) An analysis of variance test for normality (complete samples). Biometrika 52 p591-611.

Shapiro, S.S., Wilk, M.B. and Chen, H.J. (1968) A comparative study of various tests for normality. J. Am. Stat. Assoc. 63 p1343-1372.

Snedecor, G.W. and Cochran, W.G. (1967) Statistical methods. Iowa State Uni. press 593pp.

Sokal, R.R. and Rohlf, F.J. (1969) Biometry--the principles and practices of statistics in biological research. Freeman 766pp.

Spurr, S.H. (1952) Forest inventory. Ronald New York 476pp.

Stephens, C.G., Crocker, R.L., Butler, B. and Smith, R. (1941) A soil use survey of the hundreds of Riddoch, Hindmarsh, Grey, Young and Nangwarry, County Grey South Australia. CSIRO Bull. 142 55pp.

Sullivan, A.D. and Clutter, J.L. (1972) A simultaneous growth and yield model for loblolly pine. For. Sci. 18 (1) p76-86.

Theil, H. (1971) Principles of econometrics. Wiley 736pp.

Theil, H. and Nagar, A.L. (1961) Testing the independence of regression disturbances. J. Am. Stat. Assoc. 56 p793-806.

Thomasius, H. (1964) Uber die arbhangigkeit des wachstums der waldbaume von zeit und umwelt. Dungung und melioration in der forstwirtschaft. Institut fur Bodenkunde und Standortslehre, Facultat fur Forstwirtschaft der Technischer Universitat Dresden Tharandt p123-149.

Verhulst, P.F. (1844) Recherches mathematiques sur la loi d'accroissement de la population. Mem. del'Acad. Roy. de Bruxelles T XVIII p1-38.

Verhulst, P.F. (1846) Deuxieme memoire sur la loi d'accroissement de la population. Mem. del'Acad. Roy. de Bruxelles T XX p1-32.

Winsor, C.P. (1932) The Gompertz curve as a growth curve. Proc. Nat. Acad. Sci. 18 p1-8.

Wonnacott, R.J. and Wonnacott, T.H. (1970) Econometrics. Wiley 445pp.

Woods and Forests Department (1976) Annual report 1975/1976. Govt. Printer Adelaide 69pp.

ADDENDA

Bachmann, G., 1943. Wachstun und organische Geit. Bios 15 : 1-195.

Wilson, D., 1974. Australian forest industry trends. For. and Timb. Bur., Canberra, 38pp.

APPENDICES

APPENDIX 1

DATA

Appendix 1.1

SUMMARY OF DATA BASE

Key to abbreviations

Pln	Plantation year
Cpt	Compartment number
BA	Basal area
Vol	Volume
PDH	Predominant height
No	Number of
Min	Minimum
Max	Maximum
Meas	Measurements

SUMMARY OF PLOT MEASUREMENTS

Plot	Pln	Cpt	Soil Type	Site Quality	BA No Meas	BA Min Age	BA Max Age	Vol No Meas	Vol Min Age	Vol Max Age	PDH No Meas	PDH Min Age	PDH Max Age
Mount Burr Forest Reserve													
EP24A	1938	16	TF1	I	11	11	35	9	11	35	10	11	35
EP24B	1938	16	TF1	I	11	11	35	10	11	35	10	11	35
EP24C	1938	16	TF1	I	10	12	35	8	13	35	8	13	35
EP24D	1938	16	TF1	I	11	11	35	10	11	35	10	11	35
EP24E	1938	16	TF1	I	10	12	35	8	13	35	8	13	35
EP52	1940	8	MB3	V	11	13	34	9	13	34	10	13	34
EP82	1944	7	MB3	IV	10	9	28	9	9	28	9	9	28
37	1917	13	V-1	IV	19	18	56	10	18	50	7	18	50
38	1917	13	V-1	III	18	18	56	10	18	50	3	34	50
39	1917	13	V-1	IV+	18	18	56	10	18	50	4	34	50
41	1920	44	V-1	III+	21	15	53	15	15	50	10	15	50
42	1920	34	V-1	III	22	15	53	15	15	50	8	30	50
53	1925	24	MB1	V	16	10	47	11	10	46	5	32	46
54	1925	27	MB1	IV	16	10	47	11	10	46	5	32	46
55	1925	27	MB1	V	16	10	47	10	10	46	9	10	46
57	1925	18	MB2	I	22	10	47	13	10	45	15	10	45
58	1925	14	MB1	V	16	10	47	12	10	43	6	24	43
63	1963	52	MB1	II	0	0	0	0	0	0	7	5	11
73	1929	21	MM1	III	16	6	44	11	9	41	6	22	41
74	1929	21	MM1	III	16	6	44	11	9	41	5	24	41
75	1929	21	MM1	II	17	6	43	10	9	43	11	11	43
76	1929	21	MM1	II	17	6	43	13	9	43	7	24	43
81	1930	1	MB2	III	13	5	41	10	8	41	9	10	41
87	1930	6	MB3	III	14	5	41	8	8	41	5	18	41
89	1930	11	MB3	II/III	14	5	44	10	8	44	8	14	44
115	1936	4	MB1	IV+	13	9	37	13	9	37	10	13	37
116	1936	4	MB2	IV+	15	9	38	13	10	38	12	14	38
117	1936	4	MB3	IV-	13	9	35	13	9	35	10	14	35
118	1936	4	MB3	IV-	13	10	37	13	10	37	11	14	37
119	1936	4	CV1	IV	11	9	35	11	9	35	11	9	35
120	1936	4	MM2	IV	9	9	35	9	9	35	6	19	35
121	1936	2	MB2	V	12	9	38	12	9	38	9	15	38
122	1936	5	V-1	III-	13	9	37	13	9	37	10	13	37
123	1937	3	MB3	IV	13	8	34	12	8	34	10	13	34
124	1937	3	MB2	V	7	11	35	7	11	35	6	14	35
126	1928	2	MB3	IV	14	17	46	12	17	46	11	17	46
127	1928	2	MB3	IV	13	17	46	12	17	46	10	17	46
128	1928	2	MB3	IV-	13	17	46	12	17	46	12	17	46
129	1925	18	MB2	II	0	0	0	0	0	0	9	21	44
139	1938	1	MM3	IV	10	10	34	10	10	34	9	13	34
140	1940	3	CV3	IV+	11	8	34	9	11	34	10	11	34
141	1940	14	MB3	IV	10	8	32	9	10	32	9	10	32
142	1940	15	MB3	IV-	10	8	32	7	11	32	8	11	32
143	1940	14	V-1	III/IV	10	8	32	9	10	32	8	12	32
144	1937	4	MB1	IV	10	11	37	9	11	37	9	14	37
148	1940	2	MB3	V	8	13	31	8	13	31	8	13	31
149	1940	1	YS3	VII	8	8	33	7	11	33	7	11	33
150	1940	8	YM3	V-	9	8	33	7	11	33	7	11	33
151	1940	19	V-1	III	10	8	31	9	11	31	9	11	31
155	1943	3	RI1	I	9	11	30	6	11	28	7	11	28
156	1944	2	MB3	V	9	10	29	7	10	29	8	10	29
157	1942	10	MB3	V/VI	9	12	31	6	12	29	6	12	29
158	1944	4	MB3	V/VI	9	10	29	6	10	27	7	10	27
173	1947	10	MB3	VI/VII	12	15	26	5	15	25	9	15	25
174	1947	10	MB3	VI/VII	12	15	26	5	15	25	8	15	25
175	1946	5	MB2	VI/VII	13	14	26	5	14	24	9	14	24
176A	1955	4	MB1	IV/V	10	10	19	3	10	19	8	10	19
176B	1955	4	MB1	IV/V	10	10	19	3	10	19	8	10	19
176C	1955	4	MB1	IV	10	10	19	3	10	19	8	10	19
176D	1955	4	MB1	IV-	10	10	19	3	10	19	9	10	19
176E	1955	4	MB2	IV/V	10	10	19	3	10	19	8	10	19
176F	1955	4	MB1	IV/V	10	10	19	3	10	19	8	10	19
176G	1955	4	TR1	IV/IV-	10	10	19	3	10	19	9	10	19
176H	1955	4	MB2	IV/V	10	10	19	3	10	19	9	10	19
176J	1955	4	MB1	IV/V	10	10	19	3	10	19	8	10	19
177A	1954	1	MB1	IV/V	10	11	20	2	11	15	7	11	18
177B	1954	1	MB1	V	10	11	20	2	11	15	7	11	18
177C	1954	1	MB1	V	10	11	20	2	11	15	9	11	20
177D	1954	1	MB1	V	10	11	20	2	11	15	7	11	18
177E	1954	1	MB1	V/V+	10	11	20	2	11	15	8	11	20
177F	1954	1	MB1	IV/V	10	11	20	2	11	15	7	11	18
177G	1954	1	MB1	V+	10	11	20	2	11	15	6	11	18
177H	1954	1	MB1	V/V+	10	11	20	2	11	15	9	11	20
177J	1954	1	MB1	V/V+	10	11	20	2	11	15	6	11	18

SUMMARY OF PLOT MEASUREMENTS

Plot	Pln	Cpt	Soil Type	Site Quality	BA No Meas	BA Min Age	BA Max Age	Vol No Meas	Vol Min Age	Vol Max Age	PDH No Meas	PDH Min Age	PDH Max Age
Penola Forest Reserve													
X	1923	8	K-2	IV	17	17	50	15	17	50	13	17	50
Y	1923	8	K-2	IV+	17	17	50	15	17	50	14	17	50
300	1922	4	NS3	II/III	20	13	50	13	16	50	11	23	50
301	1922	4	NS3	III+	19	13	50	13	16	50	12	16	50
304	1923	8	K-3	IV	19	12	50	13	17	50	12	24	50
305	1923	8	K-3	IV	21	12	50	16	15	50	12	22	50
306	1924	1	WN1	II+	15	11	50	13	11	50	8	11	50
307	1924	1	WN1	I-	17	11	50	12	11	50	10	14	50
310	1925	7	NS2	III+	19	10	47	12	13	45	8	21	45
313	1925	1	K-3	V	17	10	47	13	10	45	10	17	45
314	1925	1	K-2	V	17	10	47	13	10	45	8	21	45
315A	1956	8	NS3	VI+	10	9	18	2	11	16	10	9	18
315B	1956	9	NS1	V-	11	9	19	3	11	19	11	9	19
315C	1956	8	NS1	VI	10	9	18	2	11	16	10	9	18
315D	1956	8	NS3	V-	10	9	18	2	11	16	10	9	18
315E	1956	8	NS2	V/V+	10	9	18	2	11	16	10	9	18
321	1927	59	NS3	IV+	16	8	44	12	11	44	8	19	44
322	1927	57	NS3	III	15	8	46	13	11	46	8	21	46
323	1927	55	KN1	IV	17	8	46	14	11	46	9	21	46
324	1927	54	KN2	IV	18	8	46	16	11	46	12	18	46
325	1927	54	K-3	IV	18	8	46	16	11	46	14	15	46
326	1927	57	K-3	IV-	15	8	46	14	11	46	9	21	46
327	1927	56	NS3	IV	17	8	46	16	11	46	11	17	46
328	1927	56	NS3	IV	17	8	46	14	11	46	12	18	46
332	1929	5	NS2	III/IV	18	6	45	13	9	45	8	18	45
335	1929	2	K-1	V/VI	14	6	41	10	9	41	4	24	41
337	1924	6	WS1	IV	18	11	50	12	11	50	9	21	50
338	1924	6	WS1	III	19	11	50	12	14	50	9	21	50
341	1927	59	NS3	IV+	15	17	46	11	17	44	10	18	44
344	1936	34	WS1	I	14	9	36	12	9	36	11	11	36
345	1936	20	WS1	III-	13	9	36	12	9	36	11	11	36
346	1935	22	NS1	V+	11	10	37	10	10	37	9	12	37
347	1935	22	NS1	IV-	12	10	37	11	10	37	10	10	37
348	1935	15	NS1	III+	12	10	38	12	10	38	10	14	38
349	1934	25	NS1	IV-	12	11	40	10	11	40	9	13	40
350	1934	25	NS2	IV/V	12	11	40	10	11	40	9	13	40
351	1936	13	WS1	II	12	12	37	12	12	37	12	12	37
352	1937	3	WS1	I+	12	11	35	12	11	35	12	11	35
353	1937	3	WN2	III-	10	11	35	10	11	35	10	11	35
354	1937	3	K-2	IV	11	11	36	8	11	36	9	11	36
355	1937	5	WS1	II+	11	11	36	10	11	33	10	11	33
356	1937	5	WS1	II	12	11	36	11	11	34	11	11	34
357	1937	6	KN3	IV	11	11	36	9	11	36	10	11	36
358	1937	1	KN1	IV	12	11	36	10	11	36	11	11	36
359	1937	1	KN1	III	11	11	37	11	11	37	11	11	37
360	1937	4	NS1	II+	13	11	37	12	11	37	12	11	37
361	1937	5	NS2	III-	11	11	37	11	11	37	11	11	37
362	1937	6	NS2	I	13	11	37	12	11	37	12	11	37
363	1937	13	NS1	II	12	11	36	10	11	33	11	11	33
364	1937	12	K-1	V	12	11	36	8	11	34	9	11	34
365	1937	8	NS2	II+	12	11	36	11	11	33	11	11	33
366	1937	10	KN1	VI	11	11	36	9	11	36	9	11	36
367	1938	1	NS2	V+	9	10	34	9	10	34	9	10	34
368	1938	10	K-1	VI	7	10	32	6	10	32	7	10	32
369	1938	11	K-1	VII	7	10	33	6	10	33	7	10	33
374	1941	2	WS1	III	8	10	31	8	10	31	8	10	31
375	1944	3	NS2	III	9	7	28	8	9	28	8	9	28
376	1942	2	NS1	III	9	10	32	9	10	32	9	10	32
377	1944	1	WS1	II	7	8	29	6	10	29	6	10	29
379	1941	4	WS1	V+	11	10	32	8	10	29	8	10	29
380	1944	4	WS1	II	10	8	30	8	10	27	8	10	27
381	1944	4	WS1	II	10	8	30	8	10	27	8	10	27
382	1944	3	NS2	II-	9	7	28	8	9	28	8	9	28
383	1944	3	NS1	IV	10	7	29	9	9	29	9	9	29
395	1946	4	KN3	VII	12	16	28	4	16	25	9	16	25
396	1946	4	KN3	VII	13	16	28	4	16	26	10	16	26
397	1949	6	NS3	VI	11	15	25	4	15	22	8	15	24
398	1949	6	NS2	VI	11	15	25	4	15	22	8	15	24
399	1949	6	NS2	VI	11	15	25	4	15	22	8	15	24

SUMMARY OF PLOT MEASUREMENTS

Plot	Pln	Cpt	Soil Type	Site Quality	Type of Measurement BA No Meas	BA Min Age	BA Max Age	Vol No Meas	Vol Min Age	Vol Max Age	PDH No Meas	PDH Min Age	PDH Max Age
Mount Gambier Forest Reserve													
508	1924	10	MB1	IV+	18	11	50	11	13	50	7	25	50
511	1926	9	MB1	III	16	9	44	12	11	44	9	16	44
513	1930	2	MB1	IV+	13	5	39	10	8	39	9	10	39
514	1928	8	WS1	III+	15	7	41	10	9	41	8	14	41
524	1928	28	WS1	II	11	17	45	9	17	45	9	17	45
526	1937	7	TF1	II	14	8	35	12	8	33	10	12	33
527	1937	2	MB2	IV	11	11	35	10	11	35	10	11	35
528	1937	1	MB1	II-	11	11	34	11	11	34	11	11	34
530	1938	21	TR1	IV+	9	10	34	5	10	30	5	10	30
531	1938	18	TR1	IV	10	10	34	5	10	30	5	10	30
532	1940	2	MB1	III+	9	8	31	8	11	31	8	11	31
533	1940	1	MB1	II	11	8	32	10	11	32	10	11	32
535	1940	8	RE1	IV	8	8	30	7	11	30	7	11	30
536	1940	8	RE1	V	8	8	30	7	11	30	7	11	30
538	1938	9	TF1	I	11	10	34	11	10	34	9	11	34
539	1938	5	TF1	I+	11	10	34	9	10	31	8	12	31
540	1940	20	TF1	I+	12	8	32	9	11	29	9	11	29
541	1940	21	TF1	I	10	10	33	9	10	30	9	10	30
542	1944	14	MB1	II	9	8	27	8	10	27	8	10	27
543	1944	16	MB1	II	12	7	29	9	9	29	9	9	29
547	1944	12	MM2	I	7	8	29	5	13	29	6	13	29
548	1944	5	TF1	III	9	9	28	9	9	28	9	9	28
549	1944	1	TR1	IV	8	9	28	8	9	28	8	9	28
550	1944	2	TF1	IV	8	9	28	8	9	28	8	9	28
553	1943	2	TF1	II	9	10	30	8	10	27	8	10	27
555	1926	2	MB3	V	7	27	47	5	27	47	6	27	47
557	1928	1	YS3	V	6	25	45	6	25	45	6	25	45
558	1935	7	MB3	V	8	18	37	8	18	37	8	18	37
559	1935	7	MB3	V	8	18	37	8	18	37	8	18	37
560	1945	4	WS1	I	14	10	27	8	10	27	11	10	27
563	1948	15	TR1	III-	10	9	23	6	9	23	10	9	23
564	1948	15	TR1	V	16	9	25	6	9	25	14	9	25
565	1948	15	TR1	V	16	9	25	6	9	25	14	9	25
571	1946	1	WS1	I	10	11	25	7	11	25	10	11	25
572	1946	1	WS1	I	10	11	25	7	11	25	10	11	25
573	1946	1	WS1	I	10	11	25	7	11	25	10	11	25
576	1948	15	YS3	VII	12	14	26	4	14	24	9	14	24
577	1948	1E	YS3	VII	12	14	26	4	14	24	9	14	24
578	1948	1E	YS3	VI	12	14	26	5	14	23	9	14	24
579	1948	1E	YS3	VI/VI-	12	14	26	5	14	23	8	14	24
Myora and Caroline Forest Reserves													
402	1930	24	CS3	IV	10	18	42	9	18	42	9	18	42
403	1930	22	CS3	IV	7	18	41	7	18	41	6	20	41
404	1930	20	CS2	II	10	18	42	10	18	42	10	18	42
405	1929	8A	CS3	II+	9	19	42	6	19	42	5	21	42
407	1936	83	CS3	II	12	12	38	11	12	38	10	14	38
408	1936	91	CS3	II	12	12	37	12	12	37	11	13	37
409	1936	87	CS3	I-	12	12	36	11	12	36	10	13	36
410	1936	81	MS3	III	11	12	38	11	12	38	10	14	38
411	1935	74	CS1	II+	11	13	39	11	13	39	10	14	39
412	1935	77	CS2	I	11	13	39	11	13	39	10	15	39
413	1937	109	CS2	I	9	11	37	8	11	37	7	13	37
414	1938	112	CS1	II+	11	10	35	11	10	35	10	12	35
415	1938	114	CS3	II-	10	10	33	10	10	33	10	10	33
417	1938	114	CS1	II	11	10	35	11	10	35	10	12	35
418	1938	117	TR1	IV	11	10	35	10	10	35	9	14	35
422	1940	122	CS3	III/IV	10	8	33	8	11	31	8	11	31
423	1940	125	CS3	III-	10	8	31	9	11	31	9	11	31
425	1942	136	CS2	I-	10	9	29	9	10	29	9	10	29
426	1942	136	CS3	II	11	9	31	10	10	31	10	10	31
427	1942	136	CS2	II	11	9	31	10	10	31	10	10	31
433	1944	145	CS3	I	6	10	25	6	10	25	6	10	25
434	1942	134	CS3	V	7	12	30	7	12	30	7	12	30
435	1944	148	CS2	III	8	11	28	8	11	28	8	11	28

SUMMARY OF PLOT MEASUREMENTS

Plot	Pln	Cpt	Soil Type	Site Quality	BA No Meas	BA Min Age	BA Max Age	Vol No Meas	Vol Min Age	Vol Max Age	PDH No Meas	PDH Min Age	PDH Max Age
Comaum Forest Reserve													
200	1938	1	BS1	III-	9	10	31	9	10	31	8	12	31
202	1938	2	BS1	IV+	9	10	31	8	10	28	9	10	31
203	1938	3	BS2	IV	10	10	36	10	10	36	10	10	36
204	1939	7	DY3	II-	10	9	32	10	9	32	10	9	32
209	1942	23	BS1	IV	9	11	29	7	11	29	8	11	29
212	1944	29	DY3	V	9	10	29	7	10	27	8	10	27
218	1948	41	DY3	VI	13	10	25	8	10	25	13	10	25
220A	1956	65	NS3	VI	0	0	0	0	0	0	13	4	18
220D	1956	65	NS3	VI	0	0	0	0	0	0	13	4	18
221A	1956	65	NS3	VI	0	0	0	0	0	0	13	4	18
221D	1956	65	NS3	VI	0	0	0	0	0	0	13	4	18
222A	1958	73	NS3	VII	0	0	0	0	0	0	13	2	16
222D	1958	73	NS3	VII	0	0	0	0	0	0	13	2	16
Cave Range Forest Reserve													
TPC541	1954	11A	CY1	V	3	9	21	3	9	21	3	9	21
TPC542	1954	11A	CY1	VI+	3	9	21	3	9	21	3	9	21
TPC571	1957	14	CY1	V	3	9	18	3	9	18	3	9	18
TPC573	1957	13	CY1	VI	3	9	18	3	9	18	3	9	18
TPC592	1959	17	CY1	IV/V	3	9	16	3	9	16	3	9	16
TPC593	1959	17	CY1	VI	3	9	16	3	9	16	3	9	16
Noolook Forest Reserve													
TPN541	1954	4	NR1	V+	4	9	21	4	9	21	4	9	21
TPN542	1954	5	NR1	IV/V	4	9	21	4	9	21	4	9	21
TPN543	1954	6	NR1	IV/V	4	9	21	4	9	21	4	9	21
TPN544	1954	6	NR1	V	4	9	21	4	9	21	4	9	21
TPN552	1955	8	NT1	V	4	9	20	4	9	20	4	9	20
TPN554	1955	9	NY3	I-	4	9	20	4	9	20	4	9	20
TPN555	1955	7	NT1	VI	4	9	20	4	9	20	4	9	20
TPN556	1955	8	NT1	IV	4	9	20	4	9	20	4	9	20
TPN561	1956	12	NT1	V	3	9	19	3	9	19	3	9	19
TPN562	1956	12	NT1	VI	3	9	19	3	9	19	3	9	19
TPN564	1956	12	NY2	V	3	9	19	3	9	19	3	9	19
TPN571	1957	14	NR1	IV	3	9	18	2	9	18	3	9	18
TPN572	1957	14	NR1	V	3	9	18	2	9	18	3	9	18
TPN585	1958	15	NY2	II/III	3	9	17	2	9	17	3	9	17
TPN591	1959	18	NY3	I	3	9	16	3	9	16	3	9	16
TPN593	1959	18	NY1	III-	3	9	16	3	9	16	3	9	16

SUMMARY OF PLOT MEASUREMENTS

Plot	Pln	Cpt	Soil Type	Site Quality	Type of Measurement BA No Meas	BA Min Age	BA Max Age	Vol No Meas	Vol Min Age	Vol Max Age	PDH No Meas	PDH Min Age	PDH Max Age
Wirrabara Forest Reserve													
908	1949	1	WA1	V	11	14	25	4	14	21	6	14	21
909	1944	11	WA1	VI	11	19	32	4	19	26	6	19	26
910	1944	11	WL1	VI	12	19	32	4	19	26	6	19	26
911	1944	15	WL1	VII	11	19	32	5	19	27	7	19	29
912	1944	4	WR1	VI+	12	19	32	4	19	26	6	19	26
913	1944	4	WR1	V-	12	19	32	4	19	26	6	19	26
914	1945	2	WR1	V	12	18	31	4	18	25	6	18	25
915	1945	2	WA1	V/V-	12	18	31	4	18	25	6	18	25
Bundaleer Forest Reserve													
956	1946	46	BD1	V/V-	10	20	30	4	20	30	8	20	30
957	1946	46	BD1	V-	10	20	30	4	20	30	8	20	30
958	1948	50	BR1	VI+	10	18	28	4	18	28	8	18	28
959	1948	50	BR1	VI	10	18	28	4	18	28	8	18	28
960	1948	50	BR1	VI+	10	18	28	4	18	28	8	18	28
961	1950	57	BR1	VI	6	16	21	3	17	21	5	16	21
962	1950	57	BR1	VI	9	16	26	2	17	18	3	16	19
963	1953	63	BR1	VI+	9	13	23	2	14	15	4	13	16
964	1953	63	BR1	VI	9	13	23	2	14	15	4	13	16
965	1953	63	BR1	VI	9	13	23	2	14	15	4	13	16
966	1953	63	BR1	V/VI	9	13	23	2	14	15	4	13	16
967	1953	63	BR1	V/VI	9	13	23	2	14	15	4	13	16
Mount Crawford Forest Reserve													
701	1929	114A	CA1	V	11	19	45	10	19	45	9	21	45
702	1936	198	CA1	IV	11	12	35	10	12	35	9	14	35
704	1936	202	AS1	II	13	12	37	8	12	37	8	12	37
705	1937	210	CA1	II/III	11	11	36	9	11	33	9	11	33
714	1943	232	AC1	VII+	19	10	31	9	10	31	13	10	31
714A	1943	232	AC1	VII+	15	17	31	5	17	31	9	17	31
715	1944	235	AC1	V	18	9	30	8	9	30	12	9	30
715A	1944	235	AC1	V	15	16	30	5	16	30	9	16	30
718	1940	225	CA1	IV/V	9	13	33	8	13	33	8	13	33
Kuitpo Forest Reserve													
605	1936	160	AL2	III/IV	12	12	38	10	12	38	11	12	38
606	1936	155	ML1	IV/V	12	12	38	11	12	38	11	12	38
607	1936	162	MD1	V-	11	12	38	10	12	38	8	17	38
610	1936	83A	TL1	III	11	12	36	10	12	36	10	12	36
612	1937	149	ML1	IV+	11	11	35	11	11	35	10	14	35
613	1937	149	ML1	V/V-	12	11	37	11	11	35	11	11	35
616	1941	24	AL1	III	8	13	32	8	13	32	8	13	32
617	1942	37	LL1	VI-	9	12	32	7	12	32	9	12	32
617A	1942	37	LL1	VI/VII	7	18	32	5	18	32	7	18	32
619	1943	30	LL1	VI/VII	16	11	31	8	11	28	11	11	30
619A	1943	30	LL1	VI/VII	13	18	31	5	18	28	8	18	30
622	1944	10	ML1	IV+	10	10	30	8	10	28	9	10	30
Second Valley Forest Reserve													
651	1935	50	SA1	IV	12	13	38	11	13	38	11	13	38
658	1944	204	SO1	V	17	10	30	7	10	30	10	10	30
658A	1944	204	SO1	V/VI	15	16	30	5	16	30	8	16	30
658B	1944	204	SO1	V	15	16	30	5	16	30	8	16	30
658C	1944	204	SO1	VI	15	16	30	5	16	30	7	16	30
659	1944	204	SO1	VI	15	10	30	8	10	30	11	10	30
659A	1944	204	SO1	VI-	13	16	30	6	16	30	9	16	30
662A	1949	209	SO1	IV/IV+	12	12	24	5	14	24	9	12	24
662B	1949	209	SO1	IV/IV+	12	12	24	5	14	24	9	12	24

Appendix 1.2

SOIL TYPE

Appendix 1.2a

Key to soil types

SOUTH EASTERN REGION

NS	Nangwarry sand	DY	Deep yellow sand
MB	Mount Burr sand	YS	Young sand
CS	Caroline sand	K-	Kilbride sand
YM	Young/Mount Burr trans.	KN	Kilbride/Nangwarry trans.
MS	Myora sand	WS	Wandilo sand
WN	Wandilo/Nangwarry trans.	RI	Riddoch sand
CV	Coarse valley soil	BS	Brown soil from Comaum
MM	Mount Muir sand	TR	Terra rossa
TF	Tantanoola flinty sand	V-	Volcanic
RE	Rendzina		

CAVE RANGE F.R.

CY	Yellow sand

NOOLOOK F.R.

NR	Red sand	NT	Red/yellow trans.
NY	Yellow sand		

WIRRABARA F.R.

WA	Alluvial loam	WL	Grey brown loam
WR	Red brown earth		

BUNDALEER F.R.

BR	Red brown earth

MOUNT CRAWFORD F.R.

AS	Alluvial sand	AC	Cromer sand
CA	Alluvial sandy clay & clay		

KUITPO F.R.

AL	Alluvial sand	MD	Mid slope loamy sand or sand
TL	Transported loam	ML	Mid-lower slope sandy loam
LL	Laterite ridge loam		

SECOND VALLEY F.R.

SA	Alluvial soils	SO	Laterite soils

Summary by Plots   Appendix 1.2b

SOUTH EASTERN REGION

NS1	NS2	NS3	MB1	MB2	MB3	CS1	CS2	CS3
315B	310	300	53[x]	57	EP52	411	404	402
315C	315E	301	54	81[x]	EP82	414	412	403
346	332	315A	55	116[x]	87[x]	417	413	405
347	350	315D	58[x]	121	89[x]		425	407
348	361	321	63	124	117[x]		427	408
349[x]	362	322	115[x]	129	118[x]		435	409
360	365	327	144	170D	123[x]			415
363	367	328	176A	175	126			422[x]
376[x]	375	341	176B	176E	127			423[x]
383	382	397	176C	176H	128			426
DAV 1	398	220A	176D	527[x]	141[x]			433
DAV 2	399	220D	176F		142			434
		221A	176J		148			
		221D	177A		156			
		222A	177B		157			
		222D	177C		158			
			177D		170A			
			177E		171A			
			177F		171D			
			177G		173			
			177H		174			
			177J		555			
			508		558			
			511		559			
			513					
			528					
			532					
			533					
			542					
			543					

DY3	YM3	KN1	KN2	KN3	YS3	K-1	K-2	K-3
204	150	323	324	357[x]	149	335	X	304
212		358[x]		395	172A	364[x]	Y	305
218		359		396	172D	368	314	313
		366			557	369	354	325
					576			326
					577			
					578			
					579			

x See Appendix 1.2c

MS3	WS1	WN1	WN2	RI1	CV1	CV3	BS1	BS2
410[x]	514	306	353	155	119	140	200	203
	524[x]	307					202	
	560						209	
	571[x]							
	572[x]							
	573[x]							
	337							
	338							
	344							
	345[x]							
	351							
	352							
	355							
	356							
	374[x]							
	377							
	379							
	380							
	381							

TF1	TR1	MM1	MM2	MM3	V-1	RE1		N/A
EP24A	176G	73	120[x]	139[x]	37	535		TP1 B64
EP24B	530	74	547		38	536		TP1 G61
EP24C	531	75			39			TP1 G62
EP24D	549	76			41			TP1 G63
EP24E	563				42			TP1 T61
526	564				122			TP1 T63
538	565				143			TP2 B64
539	418				151			TP2 G61
540								TP2 G62
541								TP2 G63
548								TP3 B64
550								TP3 G63
553								TP4 B64
								TP5 B64

x See Appendix 1.2c

CAVE RANGE FR.

CY1
TPC541
TPC542
TPC571
TPC573
TPC592
TPC593

NOOLOOK F.R.

NR1	NY1	NY2	NY3	NT1
TPN541	TPN593	TPN564	TPN554	TPN552
TPN542		TPN585	TPN591	TPN555
TPN543				TPN556
TPN544				TPN561
TPN571				TPN562
TPN572				

WIRRABARA F.R.

WA1	WR1	WL1
908	912	910
909	913	911
915	914	

BUNDALEER F.R.

BD1	BR1
956	958
957	959
	960

MT.CRAWFORD F.R.

CAI	AS1	AC1
701	704	714
702		714A
705		715
718		715A

KUTPO F.R.

AL1	AL2	TL1	ML1	MD1	LL1
616	605	610	606	607	617
			612		617A
			613		619
			622		619A

SECOND VALLEY F.R.

SA1	SO1
651	658
	658A
	658B
	658C
	659
	659A
	662A
	662B

Appendix 1.2c

Notes on specific plots

The soil type as coded refers to the predominant soil type. These notes refer to possible secondary profiles or variations in a single hole compared with the other holes.

Plot	Variation
58	possibly a Mount Muir transitional
120 139	Mount Burr/volcanic transitional
81 87 89 115 116 117 118 123 141	Mount Burr sand soils over a volcanic base
53	Mount Burr sand but possibly a terra rossa influence
323	Shallow and wet
345	possibly a Nangwarry sand transitional
357	closer to a Kilbride sand
358	closer to a Nangwarry sand
364	transitional sandy swamp soil
374	possibly a Kilbride sand transitional
376	some Wandilo sand influence
410	a mixture of soil types
422 423	Terra rossa influence
524	Mount Burr transitional
527	Mount Muir transitional
571 572 573	possibly a Mount Burr sand transitional

Appendix 1.2d

Occurrence of soil types
by
forest district

Soil type	Soil type (Stephens *et al.*, 1941)	Mount Burr	Tantanoola	Mount Gambier	Myora	Penola	Comaum
NS	Nangwarry sand			x		*	
MB	Mount Burr sand	*	*	*			
DY	Deep yellow sand						*
CS	Caroline sand			x	*		
YS	Young sand	*	*	*		x	
K-	Kilbride sand					*	
KN	Kilbride/Nangwarry					*	
WS	Wandilo sand	x	*	*		*	
WN	Wandilo/Nangwarry					*	
RI	Riddoch sand	x					
BS	Brown sand						*
TF	Tantanoola flinty sand	x	*	*			
CV	Coarse sandy valley soils	x					
MM	Mount Muir	x	x				
TR	Terra rossa	*	*	*	x	x	
V-	Volcanic	x	x	x			
RE	Rendzina			x			

Note: * soil type occurs commonly on this forest district
x soil type occurs as a minor occurrence

Appendix 1.3

UNTHINNED STAND DATA

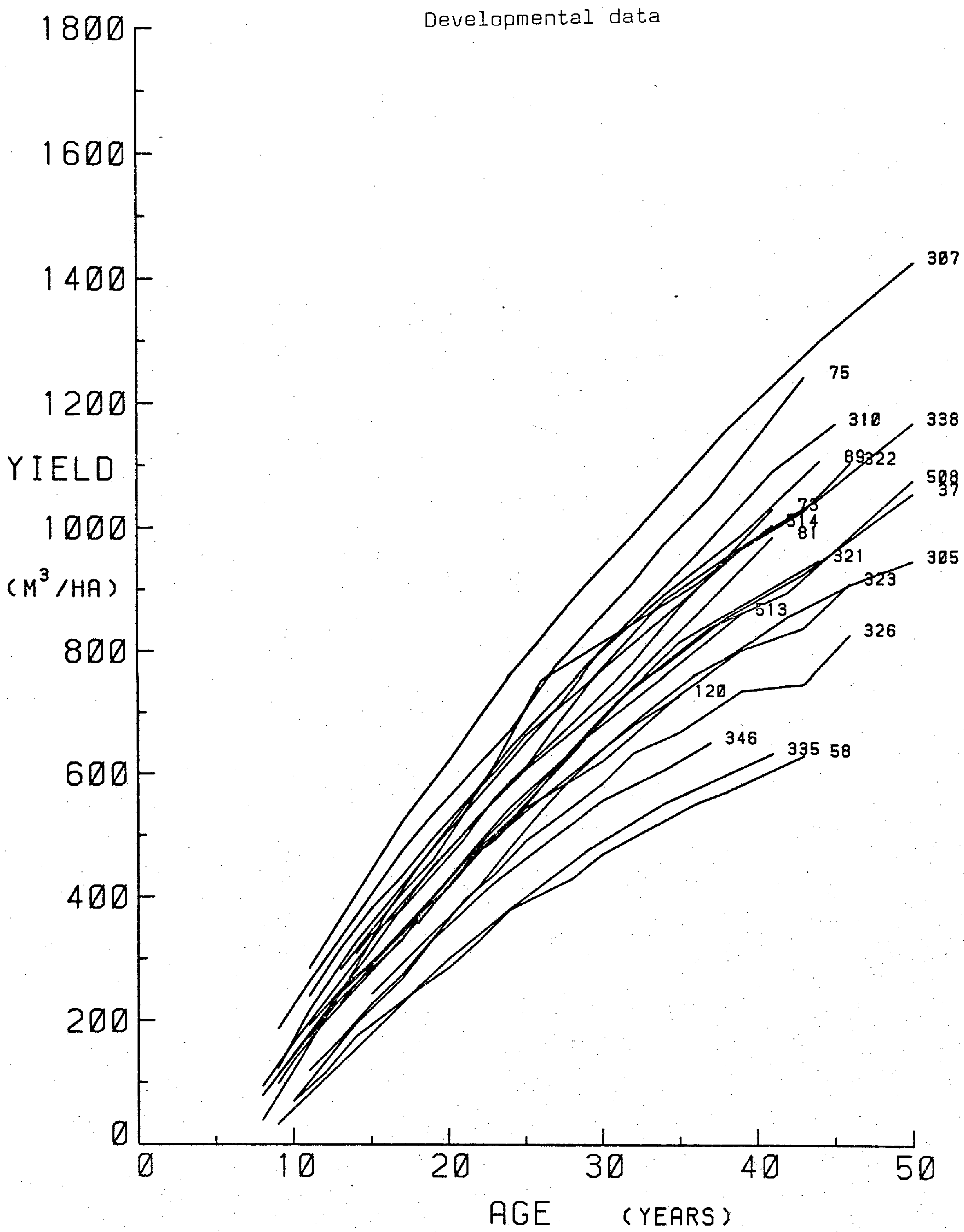
Appendix 1.3a lists the plots that were included in both the developmental and test data sets. These data sets are graphically depicted in Appendices 1.3b and 1.3c.

Appendix 1.3a

Plots in developmental and test data

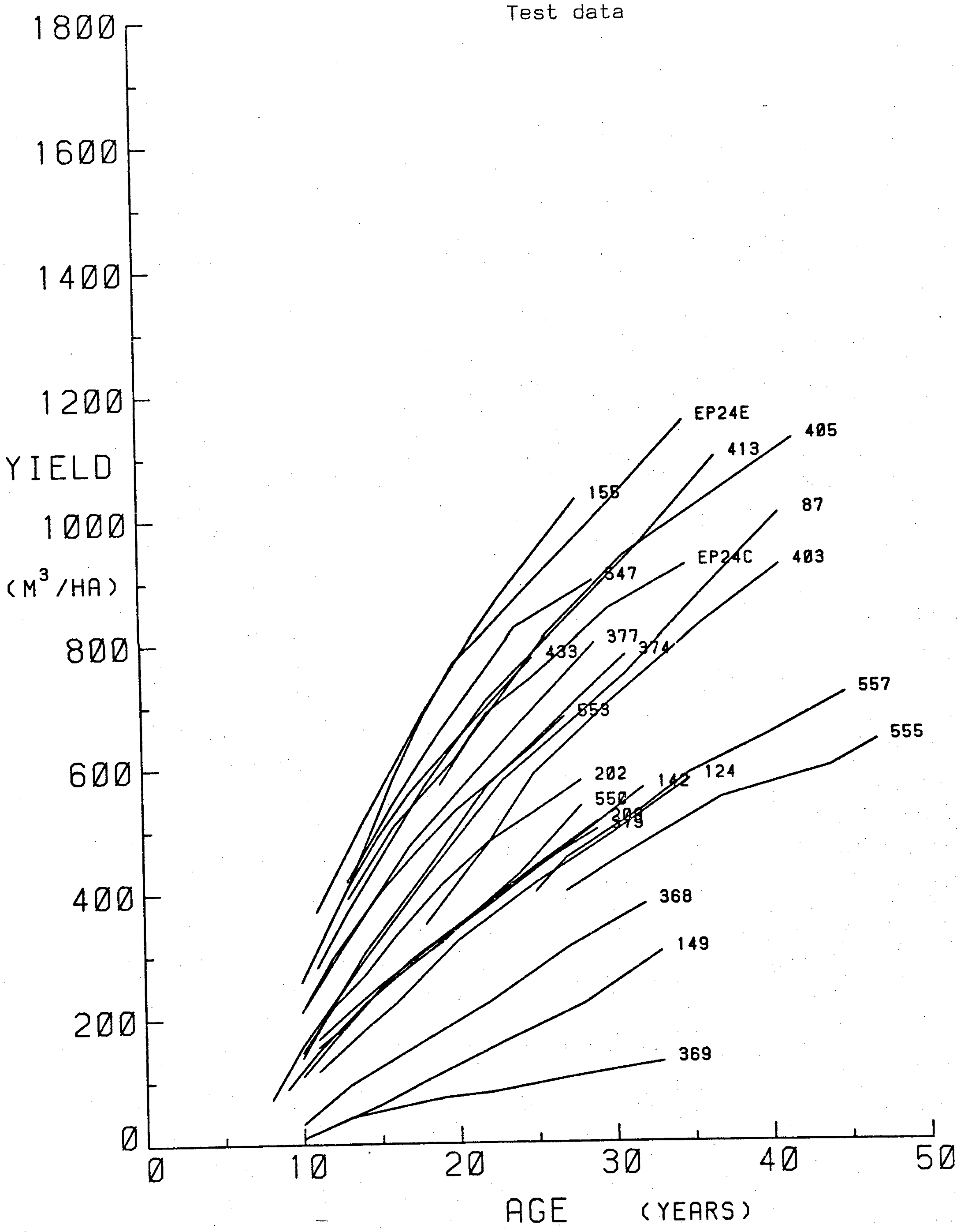
	Developmental data	Test data
Plot numbers	37	EP24C
	58	EP24E
	73	87
	75	124
	81	142
	89	149
	120	155
	305	202
	307	209
	310	368
	321	369
	323	377
	326	379
	335	403
	338	405
	346	413
	508	433
	513	547
	514	550
		553
		555
		557
Number of plots	20	23
Average number of measurements	11.4	6.9
Average growth period	33.0	20.8
Age range	8–50	8–47

Appendix 1.3b



Appendix 1.3c

Test data



DATA FROM THE LOWER SOUTH EAST

Inspection of the data from the lower south east of South Australia (excluding Noolook and Cave Range forest reserves) showed that there were relatively few data at later ages from poorer sites. The data were divided by plots (to ensure the independence of the developmental and test data sets). The developmental data included a random 70% of the low site quality plots and plots measured at later ages, and 50% of the other plots. The objective was to provide a balance by increment periods of approximately 60–40% with the developmental data being better balanced than the test data. The data are summarised by age and site quality in Appendices 1.4a and 1.4b.

The range of periodic annual increment in the developmental data was 7.0 to 62.9 m^3/ha, mean 27.3 with a standard deviation of 8.0. The test data had a mean of 28.6 with a standard deviation of 8.7, somewhat surprising in view of the care taken to provide a better balance in the developmental data, but attributable to the random selection technique.

The data cover a range of thinnings (one plot having received six thinnings and 26 plots five thinnings) with the plots thinned most often being on the better sites. The distribution of the developmental and test data by thinning and site potential is detailed in Appendices 1.4c and 1.4d, where it can be seen that the data cover a relatively wide range of stand conditions.

Appendix 1.2 details the soil types for the data. There are a number of soil types represented by only a few plots so that it was considered impractical to evaluate soil type on independent developmental data testing the models against independent test data. The data were therefore combined for the soil type evaluation. Appendix 1.4e details how the 1638 increment periods are distributed by soil type and site quality. It must be remembered that for example the 20 observations for soil type NS1 and site quality II came from only two plots.

It should be noted that the age of measurement in these tables is the age at the start of the increment period so that although in Appendix 1.4a there appears to be only 3 plots measured after age 45 there were in fact 7 plots measured at age 50.

Appendix 1.4a

Developmental Data by Age and Site Quality

Age of measurement [1]	Site Quality (Y_{10})							Total
	I	II	III	IV	V	VI	VII	
⩽10	1	1	9	5	3	1		19
11–15	27	26	38	52	37	12	4	196
16–20	23	27	48	62	25	16	12	213
21–25	21	30	35	68	26	11	10	201
26–30	16	25	32	46	17	3	1	140
31–35	8	20	26	42	17	3		116
36–40	4	5	15	23	9			56
41–45	1	2	6	11	5			25
46–50			2	1				3
Total	101	136	211	310	138	46	27	969

1 Age at the start of the increment period

Test Data by Age and Site Quality

Age of measurement[1]	Site Quality (Y_{10})							Total
	I	II	III	IV	V	VI	VII	
⩽10		5	2	7	1			15
11–15	27	36	17	41	24	3	2	150
16–20	25	37	21	46	17	5	3	154
21–25	25	29	15	44	18	3	2	136
26–30	15	23	21	35	9	1	1	105
31–35	9	11	8	27	4			59
36–40	1	2	4	16	1			24
41–45	1	1	5	13	1			21
46–50			2	3				5
Total	103	144	95	232	75	12	8	669

1 Age at the start of the increment period

Appendix 1.4c

Developmental Data by Thinning

Operation		Site Quality							Total
		I	II	III	IV	V	VI	VII	
T1	Number	9	13	15	24	20	9	6	96
	Average Age	11.0	13.0	14.4	15.8	12.8	14.9	13.6	
	σ Age	3.0	3.2	4.0	3.2	4.4	4.6	4.0	
	Average No	990	1029	973	919	666	528	491	
	σ No	289	246	239	248	216	235	195	
T2	Number	9	13	15	24	14	4		79
	Average Age	16.1	18.3	21.3	23.2	21.6	21.4		
	σ Age	4.4	3.9	6.4	5.2	7.4	8.4		
	Average No	683	743	632	620	489	386		
	σ No	218	177	184	181	187	254		
T3	Number	9	13	13	19	8	1		63
	Average Age	20.8	23.2	26.9	29.4	33.2	22.3		
	σ Age	5.7	4.9	7.4	5.4	9.3			
	Average No	488	505	455	468	335	132		
	σ No	158	133	148	151	174			
T4	Number	7	13	8	8				36
	Average Age	25.3	28.2	30.4	35.1				
	σ Age	7.8	5.9	8.0	9.7				
	Average No	350	391	357	345				
	σ No	124	119	111	145				
T5+	Number	5	9	1	1				16
	Average Age	30.1	34.4	29.0	28.3				
	σ Age	10.8	9.1						
	Average No	185	283	129	164				
	σ No	135	137						
Total	Number	39	61	52	76	42	14	6	290

Appendix 1.4d

Test Data by Thinning

Operation		Site Quality							Total
		I	II	III	IV	V	VI	VII	
T1	Number	9	13	7	19	17	1	1	67
	Average Age	11.0	11.8	14.6	14.4	13.6			
	σ Age	2.6	2.4	5.7	3.4	4.4			
	Average No	992	955	823	877	768	259	249	
	σ No	303	226	257	256	226			
T2	Number	9	14	8	20	12			63
	Average Age	14.9	17.2	21.1	22.0	20.6			
	σ Age	3.5	3.6	6.9	5.3	6.6			
	Average No	716	658	599	592	516			
	σ No	222	164	194	198	210			
T3	Number	9	14	8	13	2			46
	Average Age	19.1	22.2	26.9	29.7	25.0			
	σ Age	4.6	4.6	8.2	7.0				
	Average No	485	481	470	473	424			
	σ No	161	136	160	185				
T4	Number	7	9	3	6				25
	Average Age	23.5	26.3	33.6	33.6				
	σ Age	6.5	6.6		10.6				
	Average No	383	396	291	320				
	σ No	143	139		131				
T5+	Number	4	3	1	2				10
	Average Age	27.0	29.0	29.7	33.8				
	σ Age	10.4							
	Average No	285	255	152	200				
	σ No	134							
Total	Number	35	53	27	60	31	1	1	211

Appendix 1.4e

Combined Data by Soil Type

Soil Type	Site Quality							Total
	I	II	III	IV	V	VI	VII	
NS1		20	19	27	10	1		77
NS2	11	17	28	13	18	6		93
NS3			36	49	1	4		90
MB1		34	18	53	47			152
MB2	12		9	21	21		4	67
MB3			16	90	39	10	8	163
CS1		30						30
CS2	25	18	7					50
CS3	15	44	8	21	6			94
DY3		9			6	7		22
YM3					6			6
KN1			10	22		8		40
KN2				15				15
KN3				8			6	14
YS3					5	8	12	25
K-1					7	14	5	26
K-2				35	12			47
K-3				55	12			67
MS3			10					10
WS1	47	57	38	11	7			160
WN1	11	12						23
WN2			9					9
RI1	5							5
CV1				10				10
CV3				8				8
BS1			8	13				21
BS2				9				9
TF1	74	18	8	7				107
TR1			5	26	10			41
MM1		21	20					41
MM2	4			8				12
MM3				9				9
V-1			57	26				83
RE1				6	6			12
Total	204	280	306	542	213	58	35	1638

Appendix 1.4f

Second rotation plots

Forest Reserve	Pln	Plot	Site Quality	Age range
Mount Burr	1936	115	IV+	9-37
	1936	116	IV+	10-38
	1936	117	IV-	9-35
	1936	118	IV-	10-37
	1936	119	IV	9-35
	1936	120	IV	9-35
	1937	124	V	11-35
	1940	140	IV+	11-34
	1940	141	IV	10-32
	1940	142	IV-	11-32
	1940	143	III/IV	10-32
	1940	149	VII	11-33
	1940	EP52	V	13-34
	1944	EP82	IV	9-28
Tantanoola	1947	173	VI/VII	15-25
	1947	174	VII+	15-25
	1948	175	VII+	14-24
	1955	177 (1)	V	11-15
Penola	1946	395	VII	16-25
	1946	396	VII	16-26
	1956	315 (2)	V/VI	11-16

Note (1) nine plots

(2) five plots

Appendix 1.5

DATA FROM OTHER REGIONS

Appendix 1.5a

Noolook and Cave Range Data

Forest	Age of measurement	Site Quality							Total	Number of Thinning	Soil Type	
		I	II	III	IV	V	VI	VII				Number of Observations
Noolook											NR1	14
	<20	5	1	2	7	5	5		37		NT1	13
	21–30										NY1	2
	31–40										NY2	3
	41–50										NY3	5
	Total	5	1	2	7	5	5		37	13		37
Cave Range											CY1	12
	<20					6	6		12			
	21–30											
	31–40											
	41–50											
	Total					6	6		12	0		12

Adelaide Hills and Northern Regional Data

Forest	Age of measurement	Site Quality I	II	III	IV	V	VI	VII	Total	Number of Thinnings	Soil Type	Number of Observations
Mt.Crawford	<20		4	4	4	11		6	29		CA1	33
	21-30		2	4	4	11		6	27		AC1	23
	31-40		1		1	4			6		AS1	7
	41-50					1			1			
	Total		7	8	9	27		12	63	26		63
Kuitpo											AL1	7
											AL2	9
	<20			6	15	7	3	6	37		ML1	37
	21-30			8	15	8	3	9	43		MD1	9
	31-40			2	6	4			12		TL1	9
	41-50										LL1	29
	Total			16	36	19	6	15	92	37		92
Second Valley	<20				10	6	10		26		SA1	10
	21-30				6	4	10		20		SO1	38
	31-40				2				2			
	41-50											
	Total				18	10	20		48	20		48
Wirrabara	<20					6	3	1	10		WA1	9
	21-30					6	6	3	15		WL1	7
	31-40										WR1	9
	41-50											
	Total					12	9	4	25	8		25
Bundaleer	<20							3	3		BD1	4
	21-30						4	3	7			
	31-40										BR1	6
	41-50											
	Total						4	6	10	5		10

Appendix 1.6

DATA FOR GLS ANALYSIS

Appendix 1.6a

Plots for prior estimate by Soil Type and Forest

<table>
<tr><th>Soil Type</th><th>Soil type number</th><th colspan="5">Number of plots in alternative groupings</th></tr>
<tr><td>MB1 MB2 MB3</td><td>1</td><td>5</td><td>5</td><td rowspan="2">9</td><td rowspan="4">13</td><td rowspan="4">13</td></tr>
<tr><td>NS1 NS2 NS3</td><td>2</td><td>4</td><td>4</td></tr>
<tr><td>KN1</td><td>3</td><td>1</td><td rowspan="2">4</td><td rowspan="2">4</td></tr>
<tr><td>K-1 K-3</td><td>4</td><td>3</td></tr>
<tr><td>WN1</td><td>5</td><td>1</td><td rowspan="2">3</td><td rowspan="2">3</td><td rowspan="2">3</td><td rowspan="4">7</td></tr>
<tr><td>WS1</td><td>6</td><td>2</td></tr>
<tr><td>MM1 MM2</td><td>7</td><td>3</td><td rowspan="2">4</td><td rowspan="2">4</td><td rowspan="2">4</td></tr>
<tr><td>V-1</td><td>8</td><td>1</td></tr>
</table>

Forest	Number of Plots
Mount Burr	7
Mount Gambier	3
Penola	10

Appendix 1.6b

Plots for posterior estimate

Site potential (Y_{10})							
< 50	50–100	100–150	150–200	200–250	250–300	>300	Total
0	6	15	14	15	7	1	58

Number of observations							
9	10	11	12	13	14	15	Total
18	15	12	9	1	0	3	58

Age of first measurement						
8	9	10	11	12	13	Total
2	9	16	24	5	2	58

Length of period of growth													
20	21	22	23	24	25	26	27	29	31	32	33	34	Total
7	11	10	7	7	6	3	4	1	1	1	5	2	58

APPENDIX 2

EVALUATION OF THE PARTIAL DERIVATIVES

For each evaluation of a partial derivative the function was evaluated at intervals ($h=\pm(n/2)$delta) either side of the estimated parameter value, where delta was a relatively small value and n an integer. Expressing the function as a Taylor's series.

$$f(b+h) = \sum_{r=0}^{r=n} \frac{(h)^r}{r!} f^r(b)$$

where

b = the parameter estimate,

h = $\pm(n/2)$delta, and

$f^r(b)$ is the rth derivative of the function,

and solving the equation for $f^1(b)$ enabled an estimate to be made of the derivative. The accuracy of this approximation depended on the order of the series and on the value of delta.

Evaluation of simple test data using a simple growth model indicated that the fourth order was sufficient giving results within an order of 10^{-4} of the correct figures, however to ensure accuracy a sixth order was used. The value of delta was more difficult to determine. Preliminary evaluation indicated that the optimum value was of the order of 10^{-3} of the estimated parameter value. Smaller values of delta tended to introduce problems with machine noise, and large values of delta reduced the accuracy of the approximation, although changes of order 10^{-1} or 10^{-5} had little effect.

The problem could have been solved by evaluating all the models using a range of delta values, accepting as the best values of delta those which provided the minimum variance estimates. However this would have necessitated many evaluations of a model form, and as there were a large number of models evaluated in the study, it was impractical to carry out this procedure for all models. The delta values were checked for those models for which parameter estimates are reported and for a number of other models within each group.

It is impractical to report these evaluations in this Appendix, but the following example indicates how the analysis was carried out and the sort of results obtained.

The conditioned Bertalanffy periodic increment model

$$Pai = \frac{Y_{10}}{(A_2 - A_1)} \left[\frac{\left\{1 - \exp(-p(A_2 - A_0))\right\} - \left\{1 - \exp(-p(A_1 - A_0))\right\}}{1 - \exp(-p(10 - A_0))} \right]$$

where

$A_0 = 10 \exp(-a_1 Y_{10})$

$p = p_0 + p_1 Y_{10}$

and where

Pai = periodic annual increment between ages A_1 and A_2,

Y_{10} = site potential,

A_0 = the age at which volume growth commences, and,

p_0, p_1 and a_1 are the parameters to be estimated,

was evaluated in Chapter V, Equation V.12, where it was concluded that it was the best of the models developed along that line and also the best model developed by OLS using the unthinned data.

Appendix 2.0a shows the effect of changing the delta values for this model. Over a wide range of delta values the models were all relatively efficient, but within this range the lowest total deviates squared value was at the edge of the range where a slight change in delta made it either impossible to fit the model or provided a markedly less efficient model. The program provides asymptotically efficient estimates rather than true minimum variance estimates and the fluctuations within the bounded region reflect the effect of this on the models. The delta values selected were in the middle of this range.

$\Delta p_0 = .00001$

$\Delta a_1 = .000001$

Here the total deviates squared were 8562.8 although the minimum was 8557.6. These 'middle' values of delta also coincided with the sensible first subjective estimate of 10^{-3} of the estimated value of the parameter.

Conditioned Periodic Increment Model

Residual sum squares for various Δ values

Assuming $\Delta p_1 = \Delta p_0/100$

Δa_1	Δp_0								
	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}
10^{-2}	Δa_i Too large to allow computation to proceed								
10^{-3}	8902.0	8558.2	8558.4	8558.5	8558.5	8557.7	8557.6	8616.6	8904.9
10^{-4}	8902.3	8557.7	8562.9	8562.9	8563.0	8563.2	8561.2	8613.8	8879.3
10^{-5}	8902.3	8557.7	8562.9	8562.9	8563.0	8563.0	8559.4	8623.6	8860.7
10^{-6}	8902.3	8557.7	8562.9	8562.8	8562.8	8563.2	8558.2	8622.1	8831.1
10^{-7}	8902.3	8567.5	8562.9	8563.0	8562.6	8562.8	8562.5	8599.3	8883.8
10^{-8}	8902.3	8557.9	8562.4	8563.3	8559.8	8561.6	8561.5	8632.1	8864.1
10^{-9}	8903.2	8664.8	8609.5	8752.7	8818.4	8728.3	8714.4	8633.2	8859.7
10^{-10}	9929.9	10564.2	9315.0	9150.8	9312.9	9304.8	9307.7	9805.7	10529.6

Other parameters	Δp_1								
	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}	10^{-11}
$\Delta a_1=10^{-6}$ $\Delta p_0=10^{-5}$	8854.3	8558.6	8562.9	8562.9	8563.1	8568.3	8692.4	9337.2	9267.1

APPENDIX 3

SECOND LEVEL BERTALANFFY MODEL

Appendix 3.1

Integration of the second level Bertalanffy equation to give a yield equation.

The second level Bertalanffy equation

$$\frac{dY}{dA} = nY^m - pY$$

can be integrated using Bernoulli's equation.

Dividing by Y^m

$$\frac{dY}{dA}\ Y^{-m} = n - pY^{1-m}$$

and substituting $u = Y^{1-m}$

$$\frac{du}{dA} = \frac{-(m-1)}{Y^m}\ \frac{dY}{dA}$$

$$\frac{du}{dA}\left\{\frac{1}{1-m}\right\} = \frac{dY}{dA}\left\{\frac{1}{Y^m}\right\}$$

$$= n - pu$$

$$\int\frac{du}{pu-n} = \int(m-1)\ dA + c$$

$$\frac{1}{p}\ \ln(pu-n) = (m-1)A + c$$

$$u = \frac{n}{p}\left\{1 + \exp\left\{-p(1-m)A + c\right\}\right\}$$

or

$$Y = \left\{\frac{n}{p}\right\}^{\frac{1}{1-m}}\left\{1 + \exp\left\{-p(1-m)A + c\right\}\right\}^{\frac{1}{1-m}}$$

which can be rewritten as

$$Y = \left\{\frac{n}{p}\right\}^{\frac{1}{1-m}}\left\{1 + c_1\exp\left\{-p(1-m)(A-c_2)\right\}\right\}^{\frac{1}{1-m}}$$

where c_1 and c_2 are constants.

The constants c_1 and c_2 can be defined in a number of ways:

1 $c_2 = A_0$ (the age at which growth commences)

$$Y\Big|_{A=A_0} = 0$$

$$= \left\{\frac{n}{p} \quad (1 + c_1)\right\}^{\frac{1}{1-m}}$$

which can only be defined if $m \neq 1$, and then $c_1 = -1$

2 $c_2 = A_i$ (the age at which growth rate culminates)

$$\frac{dY}{dA} = nY^m - pY$$

$$\frac{d^2Y}{dA^2} = \left\{mnY^{m-1} - p\right\} \frac{dY}{dA} = 0 \text{ at } A=A_i$$

$$\text{ie} \quad Y = \left\{\frac{nm}{p}\right\}^{\frac{1}{1-m}}$$

therefore

$$\left\{\frac{n}{p} \quad (1 + c_1)\right\}^{\frac{1}{1-m}}\Bigg|_{A=A_i} = \left\{\frac{nm}{p}\right\}^{\frac{1}{1-m}}$$

which can only be defined if $m \neq 1$, and then $c_1 = (m-1)$.

For the logistic m=2, ie $c_1 = +1$.

The Gompertz model form (Equation IV.17)

$$Y = a \exp\left\{-\exp\left[-b(A-A_i)\right]\right\}$$

can be reformulated

$$\ln(Y) = \ln(a) - \exp\left\{-b(A-A_i)\right\}$$

therefore

$$\frac{dY}{dA} = Y\, b \exp\left\{-b(A-A_i)\right\}$$

$$= bY \ln(a/Y)$$

which has the same form as the limit form of the second level Bertalanffy equation.

HOWEVER this logic, as used by Pienaar (1966) assumes that a and b are defined at m=1, which is not so if

$$a = \left\{\frac{n}{p}\right\}^{\frac{1}{1-m}} \quad \text{and } b = p(1-m).$$

Therefore the logic is inconclusive if Bertalanffy's model is used, and assumes a and b are defined and not equal to zero as m approaches one.

Appendix 3.2

The limit form of the second level Bertalanffy as m approaches 1.0

The second level Bertalanffy equation (Equation IV.9) with $c_1=-1$ and $c_2=A_0$ can be reformulated into the Chapman-Richards form:

$$Y = a\left\{1 - \exp\left[-b(A-A_0)\right]\right\}^{\frac{1}{1-m}}$$

by substituting

$$a = \left\{\frac{n}{p}\right\}^{\frac{1}{1-m}}$$

$$b = p(1-m)$$

Differentiating

$$\frac{dY}{dA} = nY^m - pY$$

$$= pY\left\{\frac{n/p}{Y^{1-m}} - 1\right\}$$

$$= pY\left\{(a/Y)^{1-m} - 1\right\}$$

$$= bY\left\{\frac{(a/Y)^{1-m} - 1}{1-m}\right\}$$

using L'Hopitals rule

$$\lim_{x \to 0} \frac{a^x - 1}{x} = \ln(a)$$

then for m = 1

$$\frac{dY}{dA} = bY\ \ln(a/Y)$$

APPENDIX 4

EXPLORATORY ANALYSES OF UNTHINNED DATA

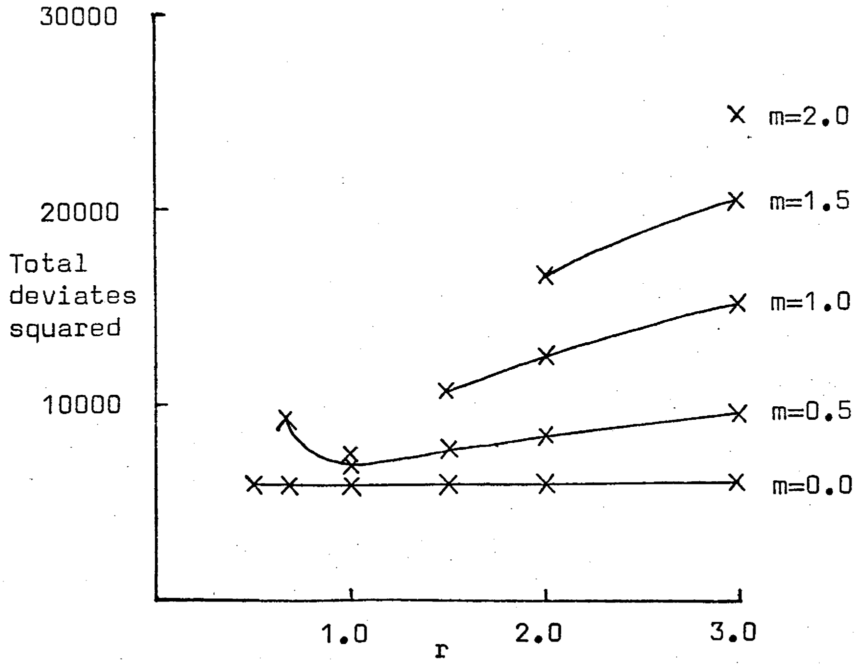
Appendix 4.1

Analysis of trend in variance

To investigate whether yield and periodic annual increment could be used as the dependent variables without the necessity of weighting, the unthinned developmental data were partitioned into age-site potential cells, and periodic annual increment into increment-period cells. The age cells were of eight years, 10-17, 18-25, 26-33, 34-41 and 42-49, the site potential cells having boundaries at Y_{10} = 100 and 200. Increment period was divided into years.

Bartlett's test was applied to the variances for each cell. When yield was investigated the value of Chi-square was 21.1 for 15 cells and when periodic annual increment was investigated the value of Chi-square was 15.5 for the 15 age site potential cells and 12.3 for the seven increment periodic cells. None of these values was significant and it was inferred that the models could be developed unweighted.

Bertalanffy: general model



Total deviates squared

Value of r	Value of m						
	float	0.0	0.5	0.667	1.0	1.5	2.0
float	5819.8[1]	5819.8[2]					
0.5		5820.8					
0.667		5820.0	9436.1				
1.0	5823.3[3]	5830.7	6901.5	7468.6			
1.5		5865.3	7785.1		10861.6		
2.0		5904.3	8592.3		12650.2	16455.3	
3.0		5975.3	9846.1		15408.7	20551.8	24927.3

Note: 1 m=0.002 r=0.682
2 m=0.0 r=0.833
3 m=0.050 r=1.0

Bertalanffy; unconditioned periodic annual increment

Model					Number of parameters	Residual sum squares	Compared with	
No	Form						Model	Significance
1	$m=0.0$	p	$n=n_0+n_1Y_{10}$	$A_0=10.0\exp(-a_1Y_{10})$	4	8461.3		
2	$m=0.0$	$p=p_0+p_1Y_{10}$	n	$A_0=10.0\exp(-a_1Y_{10})$	4	8502.8		
3	$m=0.0$	$p=p_0+p_1Y_{10}$	$n=n_0+n_1Y_{10}$	$A_0=10.0\exp(-a_1Y_{10})$	5	8426.5	1	NS
4	$m=0.0$	$p=p_0+p_1Y_{10}+p_2Y_{10}^2$	n	$A_0=10.0\exp(-a_1Y_{10})$	5	8433.1	2	NS
5	$m=0.0$	p	$n=n_0+n_1Y_{10}+n_2Y_{10}^2$	$A_0=10.0\exp(-a_1Y_{10})$	5	8459.1	1	NS
6	$m=0.0$	p	$n=n_0+n_1Y_{10}$	$A_0=a_0\exp(-a_1Y_{10})$	5	8458.6	1	NS
7	$m=0.0$	$p=p_0+p_1Y_{10}$	n	$A_0=a_0\exp(-a_1Y_{10})$	5	8499.3	2	NS
8	m	p	$n=n_0+n_1Y_{10}$	$A_0=10.0\exp(-a_1Y_{10})$	5	8445.0	1	NS
9	$m=m_0+m_1Y_{10}$	p	$n=n_0+n_1Y_{10}$	$A_0=10.0\exp(-a_1Y_{10})$	6	8442.3	8	NS

Bertalanffy, conditioned yield

Model				Number of parameters	Residual sum squares	Compared with	
No	Form					Model	Significance
1	m=0.0	p	$A_0=10.0\ \exp(-a_1Y_{10})$	2	459 526		
2	m=0.0	p	$A_0=a_0\ \exp(-a_1Y_{10})$	3	395 578	1	Sig
3	m=0.0	$p=p_0+p_1Y_{10}$	$A_0=10.0\ \exp(-a_1Y_{10})$	3	414 282	1	Sig
4	m=0.0	$p=p_0+p_1Y_{10}$	$A_0=a_0\ \exp(-a_1Y_{10})$	4	393 732	3	Sig
4						2	NS
5	m	p	$A_0=10.0\ \exp(-a_1Y_{10})$	3	410 644	1	Sig
6	m	p	$A_0=a_0\ \exp(-a_1Y_{10})$	4	395 542	2	NS
7	m	$p=p_0+p_1Y_{10}$	$A_0=10.0\ \exp(-a_1Y_{10})$	4	403 541	3	Sig
8	m	$p=p_0+p_1Y_{10}$	$A_0=a_0\ \exp(-a_1Y_{10})$	5	393 694	4	NS
9	m=0.0	p	$A_0=a_0$	2	2 894 756		
10	m=0.0	p	$A_0=a_0+a_1Y_{10}$	3	431 897	9	Sig
11	m=0.0	p	$A_0=a_0+a_1Y_{10}+a_2Y_{10}^2$	4	392 675	10	Sig
12	m=0.0	p	$A_0=a_0+a_1/Y_{10}$	3	825 294	9	NS
13	m=0.0	p	$A_0=a_0+a_1/(Y_{10}+a_2)$	4	424 916	10	Sig

Johnson-Schumacher; Linear, conditioned

Model			Number of parameters	Residual sum squares	Compared with	
No	Form				Model	Significance
1	1/A	Y_{10}	3	3.046		
2	1/A	Y_{10} Y_{10}^2	4	2.708	1	Sig
3	1/A	Y_{10} Y_{10} Y_{10}^3	5	2.598	2	Sig
4	1/A	$1/Y_{10}$	3	3.052		
5	1/A	$Ln(Y_{10})$	3	2.629		
6	A	Y_{10}	3	17.772		
7	A	Y_{10} Y_{10}^2	4	17.512	6	NS
8	A	Y_{10} Y_{10}^2 Y_{10}^3	5	17.470	6	NS
9	A A^2	Y_{10}	4	6.076		
10	A A^2	Y_{10} Y_{10}^2	5	5.789	9	Sig
11	A A^2	Y_{10} Y_{10}^2 Y_{10}^3	6	5.692	10	NS
12	A A^2 A^3	Y_{10}	5	3.793	9	Sig
13	A A^2 A^3	Y_{10} Y_{10}^2	6	3.478	12	Sig
14	A A^2 A^3	Y_{10} Y_{10}^2 Y_{10}^3	7	3.405	13	NS
15	Ln(A)	Y_{10}	3	6.480		
16	Ln(A)	Y_{10} Y_{10}^2	4.	6.196	15	Sig
17	Ln(A)	Y_{10} Y_{10}^2 Y_{10}^3	5	6.146	16	NS
18	Ln(A)	$1/Y_{10}$	3	6.677		
19	Ln(A)	$Ln(Y_{10})$	3	6.160		
20	1/A	$Ln(Y_{10})$ $Ln(Y_{10})/A$	4	1.659	5	Sig
21	1/A	Y_{10} Y_{10}^2 Y_{10}^3 Y_{10}/A Y_{10}^2/A Y_{10}^3/A	8	1.580	3	Sig
22	1/A	$1/Y_{10}$ $1/(A\ Y_{10})$	4	2.263	4	Sig

Johnson-Schumacher; nonlinear unconditioned

No	Model: $Y=b_1 \exp(-b_2/Z_1)\ Z_2 \exp(Z_3)$			Number of parameters	Residual sum squares	Compared with: Model	Compared with: Significance
5	$Z_1=A$	$Z_2=Y_{10}^{b_3}$	$Z_3=0.0$	3	491 833		
20	$Z_1=A$	$Z_2=Y_{10}^{b_3}$	$Z_3=b_4\ln(Y_{10})/A$	4	481 892	5	NS
23	$Z_1=A+a_0$	$Z_2=Y_{10}^{b_3}$	$Z_3=0.0$	4	432 373	5	Sig
24	$Z_1=A+a_0+a_1Y_{10}$	$Z_2=Y_{10}^{b_3}$	$Z_3=0.0$	5	410 256	23	Sig
25	$Z_1=A+a_0+a_1Y_{10}+a_2Y_{10}^2$	$Z_2=Y_{10}^{b_3}$	$Z_3=0.0$	6	408 104	24	NS
26	$Z_1=A+a_0\ \exp(-a_1Y_{10})$	$Z_2=Y_{10}^{b_3}$	$Z_3=0.0$	5	432 808*	23	NS
27	$Z_1=A+a_0+a_1/Y_{10}$	$Z_2=Y_{10}^{b_3}$	$Z_3=0.0$	5	423 988	23	NS
28	$Z_1=A+a_0+a_1/(a_2+Y_{10})$	$Z_2=Y_{10}^{b_3}$	$Z_3=0.0$	6	423 988	27	NS
29	$Z_1=A+a_0$	$Z_2=(Y_{10}+b_3)^{b_4}$	$Z_3=0.0$	5	429 667	23	NS
30	$Z_1=A+a_0$	$Z_2=\exp(b_3Y_{10})$	$Z_3=0.0$	4	527 701	23	NS
31	$Z_1=A+a_0$	$Z_2=\exp(b_3+b_4Y_{10})$	$Z_3=0.0$	5	523 403	30	NS
32	$Z_1=A+a_0$	$Z_2=\exp(b_3+b_4Y_{10}+b_5Y_{10}^2)$	$Z_3=0.0$	6	489 442	31	NS
33	$Z_1=A^{a_1}$	$Z_2=Y_{10}^{b_3}$	$Z_3=0.0$	4	442 218	5	Sig
34	$Z_1=(A+a_0)^{a_1}$	$Z_2=Y_{10}^{b_3}$	$Z_3=0.0$	5	418 181	23	NS
35	$Z_1=(A+a_0+a_1Y_{10})^{a_2}$	$Z_2=Y_{10}^{b_3}$	$Z_3=0.0$	6	409 512	24	NS
36	$Z_1=A$	$Z_2=Y_{10}^{b_3}$	$Z_3=b_4/A^2$	4	484 160	5	NS
37	$Z_1=A+a_0$	$Z_2=Y_{10}^{b_3}$	$Z_3=b_4/(A+a_0)^2$	5	431 987	23	NS
38	$Z_1=A+a_0+a_1Y_{10}$	$Z_2=Y_{10}^{b_3}$	$Z_3=b_4/(A+a_0+a_1Y_{10})^2$	6	406 168	24	NS

* Note that misspecification has reduced the efficiency of the estimate.

Johnson-Schumacher; nonlinear, conditioned

Model			Number of parameters	Residual sum squares	Compared with	
No	$Y=Y_{10}\ \exp(b_1(1/(A+Z_1)^{Z_2} - 1/(10+Z_1)^{Z_2}))$				Model	Significance
39	$Z_1 = 0.0$	$Z_2=1.0$	1	3 107 685		
40	$Z_1 = a_0$	$Z_2=1.0$	2	2 887 246	39	Sig
41	$Z_1 = a_0+a_1Y_{10}$	$Z_2=1.0$	3	488 826	40	Sig
42	$Z_1 = a_0+a_1Y_{10}+a_2Y_{10}^2$	$Z_2=1.0$	4	401 963	41	Sig
43	$Z_1 = a_0+a_1Y_{10}+a_2Y_{10}^2+a_3Y_{10}^3$	$Z_2=1.0$	5	392 778	42	NS
44	$Z_1 = 0.0$	$Z_2=a_1$	2	2 891 557	39	Sig
45	$Z_1 = a_0$	$Z_2=a_1$	3	2 887 267*	40	NS
46	$Z_1 = a_0+a_1Y_{10}$	$Z_2=a_2$	4	488 680	41	NS
47	$Z_1 = a_0+a_1Y_{10}+a_2Y_{10}^2$	$Z_2=a_3$	5	401 879	42	NS

* Note that misspecification has reduced the efficiency of the estimate

Bednarz; conditioned yield

Model				Number of parameters	Residual sum squares	Compared with	
No	Form					Model	Significance
1	a	b	A_0	3	2 849 540		
2	a	b	$A_0=a_0+a_1Y_{10}$	4	432 461	1	Sig
3	a	b	$A_0=a_0+a_1Y_{10}+a_2Y_{10}^2$	5	414 686	2	Sig
4	a	b	$A_0=a_0/(a_1+Y_{10})$	4	424 621	1	Sig
5	a	b	$A_0=a_0+a_1/(a_2+Y_{10})$	5	424 011	4	NS
6	a	b	$A_0=a_0 \exp(-a_1Y_{10})$	4	398 897	1	Sig
7	a	b	$A_0=a_0+a_1 \exp(-a_2Y_{10})$	5	397 931	6	NS
8	a	$b=b_0+b_1Y_{10}$	$A_0=a_0 \exp(-a_1Y_{10})$	5	397 876	6	NS
9	$a=a_2+a_3Y_{10}$	b	$A_0=a_0 \exp(-a_1Y_{10})$	5	395 583	6	NS
10	a	$b=b_0+b_1Y_{10}+b_2Y_{10}^2$	$A_0=a_0 \exp(-a_1Y_{10})$	6	397 876	8	NS
11	$a=a_2+a_3Y_{10}+a_4Y_{10}^2$	b	$A_0=a_0 \exp(-a_1Y_{10})$	6	395 401	9	NS

Bednarz; conditioned periodic annual increment

Model				Number of parameters	Residual sum squares	Compared with	
No	Form					Model	Significance
1	a	b	A_0	3	15 209.0		
2	a	b	$A_0=a_0+a_1Y_{10}$	4	8 640.8	1	Sig
3	a	b	$A_0=a_0+a_1Y_{10}+a_2Y_{10}^2$	5	8 631.3	2	NS
4	a	b	$A_0=a_0/(a_1+Y_{10})$	4	8 684.0	1	NS
5	a	b	$A_0=a_0+a_1/(a_2+Y_{10})$	5	8 686.4	4	NS
6	a	b	$A_0=a_0\ \exp(-a_1Y_{10})$	4	8 567.4	1	Sig
7	a	b	$A_0=a_0+a_1\ \exp(-a_2Y_{10})$	5	8 527.4	6	NS
8	a	$b=b_0+b_1Y_{10}$	$A_0=a_0\ \exp(-a_1Y_{10})$	5	8 482.7	6	NS
9	$a=a_2+a_3Y_{10}$	b	$A_0=a_0\ \exp(-a_1Y_{10})$	5	8 505.4	6	NS
10	a	$b=b_0+b_1Y_{10}+b_2Y_{10}^2$	$A_0=a_0\ \exp(-a_1Y_{10})$	6	8 283.8	8	NS
11	$a=a_2+a_3Y_{10}+a_4Y_{10}^2$	b	$A_0=a_0\ \exp(-a_1Y_{10})$	6	8 451.2	9	NS

Gompertz; conditioned yield

Model			Number of parameters	Residual sum squares	Compared with	
No	Form				Model	Significance
1	A_i	$b=b_0+b_1Y_{10}$	3	618 907		
2	A_i	$b=b_0+b_1Y_{10}+b_2Y_{10}^2$	4	429 551	1	Sig
3	$A_i=a_0+a_1Y_{10}$	$b=b_0+b_1Y_{10}$	4	565 706	1	Sig
4	$A_i=a_0+a_1Y_{10}+a_2Y_{10}^2$	$b=b_0+b_1Y_{10}$	5	451 371	3	Sig
5	$A_i=a_0+a_1Y_{10}$	$b=b_0+b_1Y_{10}+b_2Y_{10}^2$	5	423 067	2	NS
6	$A_i=a_0+a_1Y_{10}+a_2Y_{10}^2$	$b=b_0+b_1Y_{10}+b_2Y_{10}^2$	6	419 550	2	NS

APPENDIX 5

ANALYSIS OF DATA FROM ALL STANDS

Competition; reformulation of p

No	Model: Equations VI.2, V.1 $Z_1=1.0$			Number of parameters	Residual sum squares	Compared with: Model	Compared with: Significance
1	$m=0.0$	$p=p_0+p_1Y_{10}$	$A_0=10.0\ \exp(-a_1Y_{10})$	3	31548.8		
2	$m=0.0$	$p=p_0+p_1Y_{10}+p_2D$	$A_0=10.0\ \exp(-a_1Y_{10})$	4	31063.3	1	Sig
3	$m=0.0$	$p=p_0+p_1Y_{10}+p_2D+p_3DY_{10}$	$A_0=10.0\ \exp(-a_1Y_{10})$	5	30987.2	2	NS
4	$m=0.0$	$p=p_0+p_1Y_{10}+p_2D+p_3D^2$	$A_0=10.0\ \exp(-a_1Y_{10})$	5	29779.7	2	Sig
5	$m=0.0$	$p=p_0+p_1Y_{10}+p_2D+p_3DY_{10}+p_4D^2+p_5D^2Y_{10}$	$A_0=10.0\ \exp(-a_1Y_{10})$	7	29521.0	4	NS
6	$m=0.0$	$p=(p_0+p_1Y_{10})(1+p_2D)$	$A_0=10.0\ \exp(-a_1Y_{10})$	4	30855.2	1	Sig
7	$m=0.0$	$p=(p_0+p_1Y_{10})(1+p_2D+p_3D^2)$	$A_0=10.0\ \exp(-a_1Y_{10})$	5	29983.2	6	Sig
8	$m=0.0$	$p=p_0+p_1Y_{10}+p_2/D$	$A_0=10.0\ \exp(-a_1Y_{10})$	4	31298.6	1	Sig
9	$m=0.0$	$p=p_0+p_1Y_{10}+p_2/(1+p_3D)$	$A_0=10.0\ \exp(-a_1Y_{10})$	5	31173.3	2	NS
10	$m=0.0$	$p=(p_0+p_1Y_{10})/(1+p_2D)$	$A_0=10.0\ \exp(-a_1Y_{10})$	4	31039.2	1	Sig
11	$m=0.0$	$p=(p_0+p_1Y_{10})/(1+p_2D+p_3D^2)$	$A_0=10.0\ \exp(-a_1Y_{10})$	5	30862.7	10	NS
12	$m=0.0$	$p=(p_0+p_1Y_{10})\ D^{p_2}$	$A_0=10.0\ \exp(-a_1Y_{10})$	4	31296.7	1	Sig
13	$m=0.0$	$p=(p_0+p_1Y_{10})(p_2+p_3D^{p_4})$	$A_0=10.0\ \exp(-a_1Y_{10})$	6	30783.2	12	Sig
14	$m=0.0$	$p=p_0+p_1Y_{10}+p_2D+p_3D^2+p_4D^3$	$A_0=10.0\ \exp(-a_1Y_{10})$	6	29743.2	2	NS
15	m	$p=p_0+p_1Y_{10}+p_2D+p_3D^2$	$A_0=10.0\ \exp(-a_1Y_{10})$	6	29779.6	4	NS
16	$m=m_0+m_1D$	$p=p_0+p_1Y_{10}+p_2D+p_3D^2$	$A_0=10.0\ \exp(-a_1Y_{10})$	7	29780.0 *	4	NS
17	$m=m_0+m_1Y_{10}$	$p=p_0+p_1Y_{10}+p_2D+p_3D^2$	$A_0=10.0\ \exp(-a_1Y_{10})$	7	29780.1 *	4	NS
18	$m=0.0$	$p=p_0+p_1Y_{10}$	$A_0=10.0\ \exp(-a_1Y_{10})+a_0+a_2N_{10}$	5	31427.2	1	NS
19	$m=0.0$	$p=p_0+p_1Y_{10}$	$A_0=10.0\ \exp(a_0-a_1Y_{10}+a_2N_{10})$	5	31413.6	1	NS

* Note that misspecification has reduced the efficiency of the estimate.

Competition; correction to periodic annual increment

Model				Number of parameters	Residual sum squares	Compared with	
No	Equations VI.2, V.1		$p=p_0+p_1Y_{10}$			Model	Significance
1	$m=0.0$	$Z_1 = 1.0$	$A_0=10.0\ \exp(-a_1Y_{10})$	3	31548.8		
20	$m=0.0$	$Z_1 = b_0+b_1D$	$A_0=10.0\ \exp(-a_1Y_{10})$	5	30003.3	1	Sig
21	$m=0.0$	$Z_1 = b_0+b_1D+b_2D^2$	$A_0=10.0\ \exp(-a_1Y_{10})$	6	28520.1	21	Sig
22	$m=0.0$	$Z_1 = b_0+b_1/D$	$A_0=10.0\ \exp(-a_1Y_{10})$	5	29606.1	1	Sig
23	$m=0.0$	$Z_1 = b_0+b_1\text{Ln}(D)$	$A_0=10.0\ \exp(-a_1Y_{10})$	5	31054.3	1	Sig
24	$m=0.0$	$Z_1 = b_0D^{b_1}$	$A_0=10.0\ \exp(-a_1Y_{10})$	5	29833.6	1	Sig
25	$m=0.0$	$Z_1 = b_0+b_1D^{b_2}$	$A_0=10.0\ \exp(-a_1Y_{10})$	6	29427.7	24	Sig
26	$m=0.0$	$Z_1 = b_0+b_1\exp(-1/D)$	$A_0=10.0\ \exp(-a_1Y_{10})$	5	30151.9	1	Sig
27	$m=0.0$	$Z_1 = b_0+b_1(D-b_2)^2$	$A_0=10.0\ \exp(-a_1Y_{10})$	6	28546.0	1	Sig
28	$m=0.0$	$Z_1 = b_0+b_1/D+b_2D$	$A_0=10.0\ \exp(-a_1Y_{10})$	6	28618.4	20	Sig
29	$m=0.0$	$Z_1 = b_0+b_1D+b_2D^2+b_3D^3$	$A_0=10.0\ \exp(-a_1Y_{10})$	7	28518.6	21	NS
30	$m=0.0$	$Z_1 = b_0+(D-b_1)^{b_2}$	$A_0=10.0\ \exp(-a_1Y_{10})$	6	29669.2	1	Sig
31	m	$Z_1 = b_0+b_1D+b_2D^2$	$A_0=10.0\ \exp(-a_1Y_{10})$	7	28518.3	21	NS
32	$m=m_0+m_1D$	$Z_1 = b_0+b_1D+b_2D^2$	$A_0=10.0\ \exp(-a_1Y_{10})$	8	28520.1 *	21	NS
33	$m=m_0+m_1Y_{10}$	$Z_1 = b_0+b_1D+b_2D^2$	$A_0=10.0\ \exp(-a_1Y_{10})$	8	28519.2 *	21	NS
34	$m=0.0$	$Z_1 = b_0+b_1D+b_2D^2$	$A_0=10.0\ \exp(-a_1Y_{10})+a_0+a_2N_{10}$	8	28489.3	21	NS
35	$m=0.0$	$Z_1 = b_0+b_1D+b_2D^2$	$A_0=10.0\ \exp(-a_1Y_{10}+a_0+a_2N_{10})$	8	28493.4	21	NS

* Note that misspecification has reduced the efficiency of the estimate.

Competition; alternative indices

	Model		Number of parameters	Residual sum squares
No	Equations VI.2 V.1 m=0.0	$A_0=10.0\ \exp(-a_1 Y_{10})$		
	$p=p_0+b_1Y_{10}+b_2D+b_3D^2$	$Z_1=1.0$		
4	D=standing volume		5	29779.7
36	D=stocking		5	30455.3
	$p=p_0+p_1Y_{10}$	$Z_1=d_0+d_1D+d_2D^2$		
21	D=standing volume		6	28520.1
	D=stocking		6	28539.7

Thinning Shock

No	Model: Equations VI.2, V.1 $m=0.0$ $A_0=10.0\exp(-a_1 Y_{10})$		Number of parameters	Residual sum squares	Compared with: Model	Compared with: Significance
1	$p=p_0+p_1 Y_{10}+p_2 D+p_3 D^2$	$Z_1=1.0$	5	29779.7		
2	$p=p_0+p_1 Y_{10}+p_2 D+p_3 D^2$	$Z_1=1.0+s\ Vt$	6	29360.1	1	Sig
3	$p=p_0+p_1 Y_{10}+p_2 D+p_3 D^2$	$Z_1=1.0+s\ Vt/Vb$	6	29371.1	1	Sig
4	$p=p_0+p_1 Y_{10}+p_2 D+p_3 D^2$	$Z_1=1.0+s\ Vt/i$	6	29346.4	1	Sig
5	$p=p_0+p_1 Y_{10}+p_2 D+p_3 D^2$	$Z_1=1.0+s\ Vt/(Vb\ i)$	6	29364.4	1	Sig
6	$p=p_0+p_1 Y_{10}+p_2 D+p_3 D^2+p_4 Vt$	$Z_1=1.0$	6	29437.2	1	Sig
7	$p=p_0+p_1 Y_{10}$	$Z_1=b_0+b_1 D+b_2 D^2$	6	28520.1		
8	$p=p_0+p_1 Y_{10}$	$Z_1=(b_0+b_1 D+b_2 D^2)(1.0+s\ Vt)$	7	27727.4	7	Sig
9	$p=p_0+p_1 Y_{10}$	$Z_1=(b_0+b_1 D+b_2 D^2)(1.0+s\ Vt/Vb)$	7	27517.8	7	Sig
10	$p=p_0+p_1 Y_{10}$	$Z_1=(b_0+b_1 D+b_2 D^2)(1.0+s\ Vt/i)$	7	27587.1	7	Sig
11	$p=p_0+p_1 Y_{10}$	$Z_1=(b_0+b_1 D+b_2 D^2)(1.0+s\ Vt/(Vb\ i))$	7	27253.9	7	Sig
12	$p=p_0+p_1 Y_{10}$	$Z_1=b_0+b_1 D+b_2 D^2+s_1 Vt$	7	27902.1	7	Sig
13	$p=p_0+p_1 Y_{10}$	$Z_1=b_0+b_1 D+b_2 D^2+s_1\ Vt/Vb$	7	27903.2	7	Sig
14	$p=p_0+p_1 Y_{10}$	$Z_1=b_0+b_1 D+b_2 D^2+s_1\ Vt/i$	7	27913.7	7	Sig
15	$p=p_0+p_1 Y_{10}$	$Z_1=b_0+b_1 D+b_2 D^2+s_1\ Vt/(Vb\ i)$	7	27883.1	7	Sig
16	$p=p_0+p_1 Y_{10}$	$Z_1=b_0+b_1 D+b_2 D^2+s_1\ Vt/(Vb\ i)+s_2\ VtD/(Vb\ i)$	8	27882.9	15	NS
17	$p=p_0+p_1 Y_{10}$	$Z_1=b_0+b_1 D+b_2 D^2+s_1 Vt/(Vbi)+s_2 VtD/(Vbi)+s_3 VtD^2/(Vbi)$	9	27884.1 *	15	NS

Note * note that misspecification has reduced the efficiency of the estimate

D the index of competition level was standing volume

Form

Equations VI.2, VI.7 $p=p_0+pY_{10}$ $A_0=10\exp(-a_1Y_{10})$ $Z_1=(d_0+d_1D+d_2D^2)(1+sS)$	Form index											
	Stand Form factor		Stand Form Factor at age 10		Relative Stand Form Factor		Average Stand Taper		Average Stand Taper at age 10		Relative Stand Taper	
	Residual sum squares	Sig	Residual sum squares	Sig	Residual sum squares	Sig	Residual sum squares	Sig	Residual sum squares	Sig	Residual sum squares	Sig
$p_0=p_0+f_1F$	18941.7	NS	18771.3	Sig	18829.4	NS	18759.8	Sig	18759.4	Sig	18759.5	Sig
$p_1=p_1+f_1F$	18864.8	NS	18944.9	NS	18873.6	NS	18777.7	Sig	18781.1	NS	18779.0	NS
$a_1=a_1+f_1F$	18934.1	NS	18815.0	NS	18816.1	NS	18903.9	NS	18870.4	NS	18901.4	NS
$d_0=d_0+f_1F$	18950.8	NS	18813.1	NS	18818.9	NS	18832.9	NS	18785.9	NS	18803.5	NS
$d_1=d_1+f_1F$	18937.0	NS	18916.5	NS	18914.5	NS	18895.1	NS	18798.9	NS	18801.3	NS
$d_2=d_2+f_1F$	18953.6	NS	18947.5	NS	18934.3	NS	18950.3	NS	18770.7	NS	18883.2	NS
$s=s+f_1F$	18879.1	NS	18953.7	NS	18953.9	NS	18954.1	NS	18938.9	NS	18939.3	NS

Compared with model without stand form parameters, Residual sum squares 18954.1

APPENDIX 6

Generalized Least Squares Estimation of Yield Functions

I. S. FERGUSON

J. W. LEECH

ABSTRACT. Data were obtained from 9 measurements of 20 unthinned plots established in Monterey pine plantations in South Australia. A two-stage procedure for estimation of the yield functions was developed, drawing on the theory relating to random coefficients and to seemingly unrelated equations. In the first stage, coefficients relating yield to age for each plot were estimated using ordinary least squares. In the second stage these plot coefficients were then regressed against plot variables such as site index and stocking at age 10. The error terms in the second stage violated the assumptions of ordinary least squares, being heterogeneous across plots and correlated across coefficients. A generalized least squares algorithm was therefore developed and programmed to estimate the final coefficients and other relevant statistics. The algorithm also enabled comparison of the final coefficients based on alternative assumptions about the structure of the error terms. The results showed that the coefficients estimated under the assumption of heterogeneous correlated errors were more efficient than those under other assumptions. Recognition of the correlations between first stage coefficients proved especially important. Comparison of the heterogeneous correlated results with those from ordinary least squares applied to the pooled data from all plots also showed that while the latter estimates of the coefficients seemed robust, their variances were grossly underestimated. Model selection based on ordinary least squares and pooled data may therefore be misleading. Generalized least squares estimators offer substantial advantages in this respect and are consistent and asymptotically efficient. FOREST SCI. 24:27–42.

ADDITIONAL KEY WORDS. Statistical analysis, mathematical models, Monterey pine, *Pinus radiata*.

THIS PAPER deals with the problem of estimating yield functions for plantations of Monterey pine (*Pinus radiata* D. Don) located in the southeast of South Australia. Data used in these analyses were obtained from repeated measurement of permanent plots, often spanning 40 years in time. These data pose a number of problems for efficient estimation of yield functions. A new and more efficient technique has

The authors are, respectively, Senior Lecturer, Department of Forestry, Australian National University, Canberra, Australia; and Assistant Forest Resources Officer, Woods and Forests Department of South Australia, currently undertaking Ph. D. studies at the Australian National University.

The theory, techniques, and initial empirical research for this paper were developed while the senior author was on study leave at the College of Forest Resources, University of Washington. The assistance of Dean J. Bethel and Professor G. F. Schreuder in providing research facilities and computer time is gratefully acknowledged. This paper was first presented to a Workshop organized by Research Working Group 2, Mensuration and Management, of the Standing Committee of the Australian Forestry Council. Thanks are due to a number of participants for their comments. Special acknowledgment is due to Mr. J. Miles of the Department of Forestry, Australian National University, who developed the computer program for generalized least squares estimation; and to the Director of the Woods and Forests Department of South Australia for permission to publish this paper. Manuscript received May 3, 1977.

been developed for this purpose. This paper outlines the approach used and reports the results for a yield function suitable for use in unthinned stands.

Structure of the Growth Model

Nonlinear models such as the Chapman-Richards model or variants of it have been used extensively in recent work in estimating site or yield functions (Pienaar and Turnbull 1973). However, these models pose difficult problems in relation to statistical inference about alternative hypotheses. Moreover some of the problems associated with the data available for this study arise in both linear and nonlinear models. It was therefore simpler to start with linear form for which well-developed techniques of inference were available.

The log/reciprocal model provides a useful starting point:

$$\ln v = b_1 + b_2/a \qquad (1)$$

where ln denotes logarithm to base e,
v denotes volume of the ith observation,
a denotes age of the ith observation,
b_1, b_2 denote the fixed coefficients.

The log/reciprocal model (equation 1) has a number of desirable properties for the present study. Bailey and Clutter (1974) suggested that the simple form in equation (1) could be generalized to provide a polymorphic system of curves by including a further coefficient (c), and by making the slope coefficient plot-specific (b_{2i}):

$$\ln v = b_1 + b_{2i}(1/a)^c \qquad (2)$$

This extension seems unduly restrictive, however, since it must either be fitted by nonlinear regression or by undertaking further transforms of the model to obtain a linear form (Bailey and Clutter 1974). An alternative and more powerful generalization would be:

$$\ln v = b_{1i} + b_{2i}(1/a) + b_{3i}(1/a)^2 + \ldots \qquad (3)$$

where b_{1i} are the coefficients of a polynomial in $(1/a)$.

Equation (3) is linear in the coefficients and can thus be estimated directly using ordinary least squares. Each of the coefficients can be related to site or to other variables which affect differences between the plots. This enables the asymptotic value of volume to vary according to site, while still allowing the point of inflexion to vary with site (cf. Bailey and Clutter 1974).

Since the log-reciprocal transform itself substantially linearizes the relationship, a high-order polynomial is unlikely to be required. Nevertheless the rank of the matrix of independent variables in the polynomial needs to be established before proceeding further. Thus we proceed to discuss briefly the data and the results of the first stage of the estimation process, which involved fitting polynomials to each plot separately.

Data

The data used in this study were derived from successive measurements of a series of 20 permanent plots in Monterey pine plantations in the southeast of South Australia. All the plots had been left unthinned. The first measurement of each plot

TABLE 1. Number of plots showing significant improvement in fit.

Forms of model	Number of plots showing significant improvement
Linear to quadratic	15
Quadratic to cubic	9
Cubic to quartic	5

had been carried out at or near age 10 years. Subsequent measurements were carried out at varying intervals to ages of 40 or 50 years. The standards of measurement were notable for strict adherence to well-established and documented procedures, using highly trained and experienced staff.

The irregular number of observations in the plots posed a problem. Although variable numbers of observations could be handled by the techniques outlined in this paper, the advantages seemed to be outweighed by the additional computational burden.

Thus the data were culled to reduce all plots to nine observations. The first and last measurements were retained in each plot, in order to maintain the maximum period of growth possible. For each plot the surplus observations were culled randomly.

First-Stage Model

A polynomial was fitted to the data[1] from each plot using ordinary least squares. The most appropriate order for the polynomial was not clear, although consideration of the second-stage model suggested that it should be consistent for all plots. Thus linear, quadratic, cubic, and quartic forms of the model were fitted to each plot.

The various forms of the polynomial model were then tested to determine whether the addition of each successive term represented a significant difference over the simpler forms. Inspection of the plot variances of the residuals for any one form of the polynomial model indicated that the plot variances were markedly heterogeneous.

Differences in the pattern of heterogeneity between different forms of the model seemed to eliminate an analysis of variance based on the pooled data. Tests were therefore carried out by plots to establish whether each additional term represented a significant improvement over the previous form. The numbers of the calculated values of the F statistic which exceeded the critical value at the 95 percent probability level are summarized in Table 1.

The results in Table 1 suggested that the quadratic form was probably superior to the linear, but the other comparisons were not so clear. The signs and values of the higher order coefficients in the cubic and quartic forms were notably erratic. Thus quadratic and higher forms were pursued in the second-stage analyses. Since the cubic and quartic proved to be untenable in the second stage, only the results for the quadratic form will be reported in subsequent sections.

The estimated values of the coefficients for the quadratic model are shown in Table 2, together with the values of site index (s_i) and stocking at age 10 years (n_i). In accord with South Australian practice, site index was measured by the estimated volume per unit area the plot would carry at age 10 years.

[1] Age was measured in *tens* of years to provide better-conditioned moment matrices for the first-stage estimates. All subsequent results reflect this scaling of age.

TABLE 2. Estimated coefficients and plot data for quadratic first-stage model.

Plot	Estimated coefficients $\hat{b}_{1i}$	$\hat{b}_{2i}$	$\hat{b}_{3i}$	Site index (m^3/ha) s_i	Stocking (age 10) (Stems/ha) n_i
1	7.902	–3.503	1.131	253.2	1549
2	7.826	–3.479	1.044	223.9	1495
3	7.621	–3.141	0.870	204.1	1690
4	7.956	–4.436	1.873	184.5	1700
5	7.886	–4.532	2.125	179.2	1619
6	7.522	–2.552	0.122	168.5	1673
7	7.525	–2.890	0.402	164.2	1703
8	7.699	–3.983	1.442	161.1	1716
9	7.650	–3.213	0.528	145.2	1549
10	7.434	–3.148	0.717	143.8	1680
11	7.407	–2.716	0.238	143.1	1680
12	7.713	–4.065	1.445	141.5	1737
13	7.275	–2.465	0.062	135.0	1468
14	7.648	–3.637	0.815	132.1	1982
15	7.390	–1.713	–0.991	119.1	1208
16	7.524	–3.374	0.460	111.3	2162
17	7.398	–3.142	0.276	93.7	1834
18	7.093	–2.816	–0.023	70.0	1673
19	7.027	–1.817	–0.954	69.1	1581
20	7.013	–2.153	–0.878	57.2	1609

The data in Table 2 are arranged in descending order of site and provide some visual evidence of a probable correlation between the values of the coefficients and the values of site index.

The estimated values of the elements ($\sigma_{jk}{}^{i}$) of the variance-covariance matrix for these coefficients are summarized in Table 3. Since the matrices for each plot are symmetric, only the diagonal and upper diagonal elements are shown.

Considerable heterogeneity between plots is apparent in the data in Table 3. Bartlett's test of homogeneity was used to examine this problem, using the estimated variances of the residuals for each plot. The calculated value of the test statistic was 65.6. This statistic is approximately distributed as a χ^2 variable with 19 degrees of freedom. The calculated value exceeds the critical value (30.1) of χ^2 at the 95 percent probability level and thus the variances of the residuals are significantly heterogeneous.

Scatter diagrams of the residuals for each plot gave no indication of heterogeneity or of serial correlation within any of the plots. The Durbin-Watson statistic was also calculated for each plot, even though its value is questionable with so few observations. The published critical bounds only go down to 15 observations (Theil 1971). Extrapolating these to 9 observations (and recognizing the dangers inherent), the lower bound is about 0.8 and the upper bound is about 1.5 at the 95 percent probability level. None of the plots had calculated values of the statistic below the lower bound. Six fell in the inconclusive zone (between upper and lower critical bounds) for positive serial correlation and six in the inconclusive zone for negative serial correlation. If serial correlation were present in the first-stage model, one would expect it to be consistent (either positive or negative) for all or most plots. These results suggest that serial correlation was not a serious problem in the data.

TABLE 3. Estimated variances and covariances of first-stage coefficients.

Plot	Estimated variances and covariances ($\times 10^{-4}$)					
	σ_{11}^{i}	σ_{12}^{i}	σ_{22}^{i}	σ_{13}^{i}	σ_{23}^{i}	σ_{33}^{i}
1	9	–36	159	30	–141	129
2	48	–171	652	122	–478	362
3	18	–71	306	61	–272	250
4	8	–34	162	34	–165	171
5	10	–47	242	50	–263	292
6	104	–368	1370	260	–988	731
7	25	–80	269	49	–174	118
8	21	–82	333	68	–283	247
9	23	–77	279	54	–203	154
10	26	–112	516	98	–467	433
11	30	–93	318	59	–210	144
12	24	–105	491	102	–497	519
13	41	–143	515	100	–369	268
14	15	–83	485	107	–639	855
15	74	–230	783	146	–517	356
16	25	–127	655	138	–722	811
17	7	–30	129	25	–113	103
18	10	–39	153	30	–121	99
19	37	–135	522	105	–416	339
20	45	–105	536	104	–386	288

Second-Stage Model

The estimated coefficients for the first-stage models enable yield predictions to be made for any plot in the sample but not for any other plot. The second-stage model is concerned with the development of a more general model, capable of making yield predictions for any plot drawn from the same population as the sample.

Consider the estimated coefficients for the quadratic form of equation (3), shown in Table 2. Each plot can be regarded as a random sample from the population of all plots. Thus these coefficients can be regarded as random variables or "random coefficients," to use the terminology of the literature on growth curves and related work (e.g. Potthoff and Roy 1964, Rao 1965, Grizzle and Allen 1969). Swamy (1971), Rosenberg (1973), and Fearn (1965) have further developed the relevant theory (both classical and Bayesian) regarding random coefficient models, and efficient unbiased estimators of the expected values (and variances) of these coefficients have been developed for various applications. Leak's (1966) pioneering work with repeated measurements in forestry data developed what would now be recognized as large-sample estimators of the expected values and variances for a random coefficients model.

These models need not be limited to random *coefficients*, as Grizzle and Allen (1969) and Rosenberg (1973) have noted. The coefficients of the first-stage model can be postulated to be random *functions* of other exogenous variables. Nevertheless the techniques for estimation of random functions received little attention in the literature on random coefficients models.

The problem of estimating random functions is exactly analogous to that of estimating "seemingly unrelated equations," although this does not seem to have been recognized previously in the literature. The theory and techniques of estimation for the latter problem were first developed by Zellner (1962). Goldberger (1964), Dhrymes (1970), and Theil (1971) also contain useful contributions on the problem.

Before turning to the formal development for estimation of a general model of this type, let us consider the structure of the second-stage model in more detail. The regression coefficients in equation (3) may be postulated to be random functions of site index and any other appropriate variables. As a specific example, let us postulate that the intercept term in equation (3) is a linear function of the logarithm of site index, which seems reasonable because this coefficient determines the asymptotic value of volume as age approaches infinity:

$$b_{1i} = a_{11} + a_{12} \ln s_i + \delta_{1i} \tag{4}$$

The random error (δ_{1i}) may be attributed to the inherently stochastic nature of biological relationships and/or to the large number of potentially important factors which are not taken explicitly into account in the model. For example, the genotype of the planting stock and a multitude of soil and microclimatic factors are known to be potentially important determinants of forest growth, but none of them appear in this model.

The detail of the structure in terms of the variables, and the form in which they are included, may vary between the different regression coefficients. Hence it is desirable to adopt a more general formulation of equation (4):

$$b_{li} = \sum_j a_{lj} z_{ji} + \delta_{li} \tag{5}$$

where the subscript l $(=1, 2, 3)$ is used to denote the lth regression coefficient in equation (3),

z_{ji} denotes the jth $(= 1, 2 \ldots)$ independent variable for the ith plot.

For any one first-stage coefficient it seems reasonable to assume that the errors (δ_{li}) in equation (5) are identically and independently distributed with mean zero and variance denoted Δ_{ll}. However the covariances between the error terms of different first-stage coefficients will not in general be equal to zero (i.e., $\Delta_{lm} \neq 0$), because the coefficients are generally interrelated. For example, other things being equal, a particular soil type is likely to affect all of the first-stage coefficients for a particular plot in some related manner.

The formulation in equation (5) is based on the *true* regression coefficient. Clearly, errors of estimation in the first-stage must also be recognized:

$$\hat{b}_{li} = \sum_{j=1}^{j=k} a_{lj} z_{ji} + \delta_{li} + e_{li} \tag{6}$$

where $\hat{b}_{li}$ denotes the estimated values of the lth coefficient $(l = 1, 2 \ldots)$ for the ith plot,

e_{li} denotes the error of the estimate of the lth coefficient for the ith plot.

In this form, it will be apparent that the second-stage coefficients (a_{lj}) should not be estimated by ordinary least squares applied separately to the data for each coefficient. The combined error term ($\delta_{li} + e_{li}$) does not obey the assumptions underlying ordinary least squares; the variance of one component (e_{li}) being heterogeneous from plot to plot. Moreover more powerful techniques are available which take advantage of the properties of the error term to increase the efficiency of estimation of the second-stage coefficients.

Generalized Least Squares

The model outlined in equation (6) can be developed fully in matrix notation thus:

$$B = ZA + U \tag{7}$$

where $B' = [B_1, B_2, B_3]$

and $B_l' = \left[b_{l,1}, b_{l,2} \ldots b_{l,i} \ldots b_{l,20}\right]$;

$$Z = \begin{bmatrix} Z_1 & 0 & 0 \\ 0 & Z_2 & 0 \\ 0 & 0 & Z_3 \end{bmatrix}$$

and
$$Z_l = \begin{bmatrix} z_{l,1} & \cdot & \cdot & \cdot & \cdot & \cdot & z_{l,20} \\ \cdot & & & & & & \cdot \\ \cdot & & & & & & \cdot \\ \cdot & & & & & & \cdot \\ z_{m,1} & \cdot & \cdot & \cdot & \cdot & \cdot & z_{m,20} \end{bmatrix}$$

$A' = [A_1, A_2, A_3]$

and $A_l' = [a_{l,1}, a_{l,2} \ldots a_{l,m}]$;

$U' = [U_1, U_2, U_3]$

$U_l' = [u_{l,1}, u_{l,2} \ldots u_{l,20}]$

and $u_{l,i} = \delta_{li} + e_{li}$.

Equation (7) is applicable to other first-stage models such as the cubic form by appropriate modification of the dimensions of the components.

Note that the submatrix of independent variables (Z_l) need not be identical for all values of l. Indeed, if all these submatrices were identical, much of the gain in efficiency which accrues from a generalized least squares approach would be lost (Theil 1971).

Using equation (4) as a specific example, the submatrix (Z_l) would contain the vectors derived from $z_{1i} = 1$ and $z_{2i} = \ln s_i$ for $i = 1 \ldots 20$.

Assuming the variance-covariance matrix of errors (U) in equation (6) is known, the theory of generalized least squares (Aitken 1934–1935) may be used to derive best linear unbiased estimates for the second-stage model. These are sometimes referred to as Aitken estimators.

The generalized least squares approach involves a transformation of the model:

$$TB = TZA + TU \tag{8}$$

where T is a square matrix such that:

$$T'T = [E(UU')]^{-1} = W^{-1} \tag{9}$$

where E denotes the expected value operator
and W denotes the known variance-covariance of the error terms in equation (7).

Under these conditions it can be shown (Theil 1971) that

$$E\,[T'U] = 0 \tag{10}$$

$$E\,[T'UU'T] = I \tag{11}$$

where I denotes an identity matrix.

Hence equation (8) fulfills the assumption underlying least squares and can be estimated using ordinary least squares. The resultant Aitken estimators of the coefficients (A) can be shown to be best linear unbiased estimators (Theil 1971).

Of course, the variance-covariance matrix (W) is not known. We therefore follow the usual practice of substituting an estimate for it. Swamy (1971) has developed an unbiased estimator, which will be outlined in the next section, and has shown that so-called *feasible Aitken estimators* (Dhrymes 1970) based on it are consistent and asymptotically efficient. Thus throughout the remainder of this paper we will use W to denote the *estimated* variance-covariance matrix.

The feasible Aitken estimators can be calculated using the following formula:

$$A = [Z'W^{-1}Z]^{-1}\,[Z'W^{-1}B] \tag{12}$$

where $A' = [A_1, A_2, A_3]$

$$[Z'W^{-1}Z]^{-1} = \begin{bmatrix} Z_1'\,W_{11}^{-1}\,Z_1 & Z_1'\,W_{12}^{-1}\,Z_2 & Z_1'\,W_{13}^{-1}\,Z_3 \\ Z_2'\,W_{21}^{-1}\,Z_1 & Z_2'\,W_{22}^{-1}\,Z_2 & Z_2'\,W_{23}^{-1}\,Z_3 \\ Z_3'\,W_{31}^{-1}\,Z_1 & Z_3'\,W_{32}^{-1}\,Z_2 & Z_3'\,W_{33}^{-1}\,Z_3 \end{bmatrix}^{-1}$$

$$[Z'W^{-1}B] = \begin{bmatrix} Z_1'\,W_{11}^{-1}\,B_1 + Z_1'\,W_{12}^{-1}\,B_2 + Z_1'\,W_{13}^{-1}\,B_3 \\ Z_2'\,W_{21}^{-1}\,B_1 + Z_2'\,W_{22}^{-1}\,B_2 + Z_2'\,W_{23}^{-1}\,B_3 \\ Z_3'\,W_{31}^{-1}\,B_1 + Z_3'\,W_{32}^{-1}\,B_2 + Z_3'\,W_{33}^{-1}\,B_3 \end{bmatrix}$$

The partitioned matrices provide a somewhat simpler basis for computation and aid comparisons with other developments later in this paper.

The inverse matrix $[Z'W^{-1}Z]^{-1}$ in equation (12) has been shown (Theil 1971) to provide a consistent estimator of the variance-covariance matrix for the estimated coefficients, i.e.:

$$\mathrm{Var}\,(A) = [Z'\,W^{-1}\,Z]^{-1} \tag{13}$$

where Var is used to denote the variance-covariance operator.

Variance-Covariance Error Matrix

Following Swamy (1971), an unbiased estimator of the variance-covariance error matrix (W) can be derived in two stages. First, the variance-covariance matrix (Γ) for the estimated first-stage coefficients (b_{li}) can be calculated in the usual way:

$$\Gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{bmatrix} \tag{14}$$

where $\gamma_{lm} = (\sum b_{li}\,b_{mi} - \sum b_{li} \sum b_{mi}/n)/(n-1)$.

For the quadratic form of the first-stage model, the following values were obtained:

$$\Gamma = \begin{bmatrix} .6128 & -.6582 & -.1793 \\ -.6582 & .6128 & .2109 \\ -.1793 & .2109 & .0774 \end{bmatrix}$$

This matrix includes the errors of estimation in the coefficients. This component can be estimated separately from the mean (over the 20 plots) of the relevant variance or covariance elements derived in the first-stage fitting by ordinary least squares. Thus the variance-covariance matrix of the δ_{li} terms can be estimated as follows:

$$\Delta = \begin{bmatrix} \Delta_{11} & \Delta_{12} & \Delta_{13} \\ \Delta_{21} & \Delta_{22} & \Delta_{23} \\ \Delta_{31} & \Delta_{32} & \Delta_{33} \end{bmatrix}$$

and

$$\Delta_{lm} = \gamma_{lm} - \sum_i \sigma_{lm}{}^i/n\,. \tag{15}$$

The values in the matrix can be used to estimate the correlation coefficient (l_{lm}) between the error terms (δ_{li}, δ_{mi}) associated with different regression coefficients in the second stage model:

$$l_{12} = -.972, \qquad l_{13} = -.823, \qquad l_{23} = .876\,.$$

The high values obtained suggest that substantial gains in efficiency should accrue from recognition of this feature in the estimating process.

The second stage in estimating the final variance-covariance matrix (W) involves adding the variances and covariances attributable to the first-stage estimation to the Δ matrix, for each plot:

$$W^i = \begin{bmatrix} w_{11}{}^i & w_{12}{}^i & w_{13}{}^i \\ w_{21}{}^i & w_{22}{}^i & w_{23}{}^i \\ w_{31}{}^i & w_{32}{}^i & w_{33}{}^i \end{bmatrix}$$

$$w_{lm}{}^i = \Delta_{lm} + \sigma_{lm}{}^i \tag{16}$$

where the superscript i is used to denote the elements of the W matrix for the ith plot.

The elements of the 20 matrices so derived can be rearranged to form the W matrix:

$$W = \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \end{bmatrix}$$

and

$$W_{lm} = \begin{bmatrix} w_{lm}{}^1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & & w_{lm}{}^2 & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & w_{lm}{}^i & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & w_{lm}{}^{20} \end{bmatrix} \tag{17}$$

Each of the W_{lm} submatrices is a diagonal matrix with zero off-diagonal elements. The diagonal elements are the variance or covariance elements for each plot. It is thus obvious that the variance-covariance matrix W is far removed from the structure embodied in the assumption underlying ordinary least squares, which would imply that W was a diagonal matrix with constant values along the diagonal.

Alternative Forms of Error Matrix

Comparison of the structure of the error matrix with alternative forms seems desirable so that the theoretical gains in efficiency can be examined empirically.

Heterogeneous and Correlated.—The treatment in the preceding sections seems appropriate in view of the evidence to date. Significant heterogeneity existed between the variance-covariance matrices of different plots and high correlations existed between the errors for different regression coefficients. Nevertheless a number of alternative forms can be examined within the same framework. The resulting modifications of equation (12) will be reviewed briefly to provide a basis for subsequent analyses.

Heterogeneous and Independent.—If the combined error term $(\delta_{li} + e_{li})$ is assumed to be heterogeneous across plots but independent for different coefficients, the off-diagonal submatrices in equation (12) equal zero. Thus the components of equation (12) become:

$$[Z' W^{-1} Z]^{-1} = \begin{bmatrix} (Z_1' W_{11}^{-1} Z_1)^{-1} & 0 & 0 \\ 0 & (Z_2' W_{22}^{-1} Z_2)^{-1} & 0 \\ 0 & 0 & (Z_3' W_{33}^{-1} Z_3)^{-1} \end{bmatrix}$$

$$[Z' W^{-1} B] = \begin{bmatrix} Z_1 W_{11}^{-1} B_1 \\ Z_2 W_{22}^{-1} B_2 \\ Z_3 W_{33}^{-1} B_3 \end{bmatrix} \tag{18}$$

Such a model could be estimated by weighted least squares applied separately to each of the three regression coefficients.

Homogeneous and Correlated.—If the error term is assumed to be homogeneous across plots but correlated for different coefficients, the W_{lm}^{-1} submatrices are simply replaced by the scalar value for the inverse of the variance or covariance element concerned:

$$[Z' W^{-1} Z]^{-1} = \begin{bmatrix} \sigma_{11}^{-1} Z_1' Z_1 & \sigma_{12}^{-1} Z_1' Z_2 & \sigma_{13}^{-1} Z_1' Z_3 \\ \sigma_{21}^{-1} Z_2' Z_1 & \sigma_{22}^{-1} Z_2' Z_2 & \sigma_{23}^{-1} Z_2' Z_3 \\ \sigma_{31}^{-1} Z_3' Z_1 & \sigma_{32}^{-1} Z_3' Z_2 & \sigma_{33}^{-1} Z_3' Z_3 \end{bmatrix}^{-1}$$

$$[Z' W^{-1} B] = \begin{bmatrix} \sigma_{11}^{-1} Z_1' B_1 + \sigma_{12}^{-1} Z_1' B_2 + \sigma_{13}^{-1} Z_1' B_3 \\ \sigma_{21}^{-1} Z_2' B_1 + \sigma_{22}^{-1} Z_2' B_2 + \sigma_{23}^{-1} Z_2' B_3 \\ \sigma_{31}^{-1} Z_3' B_1 + \sigma_{32}^{-1} Z_3' B_2 + \sigma_{33}^{-1} Z_3' B_3 \end{bmatrix} \tag{19}$$

where σ_{lm}^{-1} is the inverse of the (constant) variance or covariance elements for all plots.

This formulation does not decompose to separate equations as in the previous section and can only be estimated by joint generalised least squares.

Homogeneous and Independent.—If the error term is assumed to be both homogeneous across plots and independent for different coefficients, the components of equation (12) simplify still further:

$$[Z' W^{-1} Z]^{-1} = \begin{bmatrix} \sigma_{11}(Z_1' Z_1)^{-1} & 0 & 0 \\ 0 & \sigma_{22}(Z_2' Z_2)^{-1} & 0 \\ 0 & 0 & \sigma_{33}(Z_3' Z_3)^{-1} \end{bmatrix}$$

$$[Z' W^{-1} B] = \begin{bmatrix} \sigma_{11}^{-1} Z_1' B_1 \\ \sigma_{22}^{-1} Z_2' B_2 \\ \sigma_{33}^{-1} Z_3' B_3 \end{bmatrix} \tag{20}$$

Clearly this model could be estimated by applying ordinary least squares separately to the data for each regression coefficient, because the σ terms cancel on multiplication.

Homogeneous and Independent Using Pooled Data.—One further possibility warrants comparison. Suppose the assumptions regarding random coefficients or functions are dropped. The first- and second-stage models can then be combined on the assumption that the coefficients are fixed, not random:

$$\ln v_i = \sum_j a_j x_{ij} + \epsilon_i \tag{21}$$

where a_j denotes the jth fixed coefficient,
x_{ji} denotes the jth independent variable for the ith observation,
ϵ_i denotes the random error term.

If the error term in equation (21) is assumed to be homogeneous for all observations, and the error terms for different observations are assumed to be independent, equation (21) can be estimated by applying ordinary least squares to the pooled data covering all plots and measurements..

Analyses

A computer program was developed to compute the second-stage model based on equation (12) (Miles and others 1978).

The first models were based on the assumption that the error terms were heterogeneous and correlated. The cubic form of the first-stage model was eliminated from further consideration at this stage, because the estimated error matrix for the regression coefficients (Δ) had some off-diagonal terms whose square exceeded the value of the product of the corresponding diagonal terms. This indicates (Swamy 1971) that either the assumed model is incorrect or that statistical variability has obscured the underlying relation. In the light of the F tests carried out on the earlier model, reported earlier, the former seemed more likely.

For the quadratic model, different forms of the Z_i submatrices were examined to determine the most appropriate structure for the second-stage models in relation to the independent variables. Site index, stocking at age 10 years, and dummy variables for soil types were included in various forms and combinations.

Tests of significance for an individual independent variable were based on the following test statistic:

$$\zeta = \frac{a_l}{\sqrt{\mathrm{Var}\,(a_l)}} \tag{22}$$

where ζ denotes the test statistic which is asymptotically a $N\ (0, 1)$ variate,
a_l denotes the estimated value of the regression coefficient,
$\mathrm{Var}\,(a_l)$ denotes the estimated value of the variance of the coefficient.

Where joint tests of significance were required to test whether two or more coefficients were jointly significantly different from zero, F tests were carried out based on the alternative models with and without the variables concerned (Theil 1971):

$$F = \frac{(Lp - k_2)}{(k_2 - k_1)} \cdot \frac{SS_1 - SS_2}{SS_2} \tag{23}$$

where L, p denotes the number of equations and plots respectively,
k_1, k_2 denote the numbers of coefficients in models 1 and 2,
SS_1, SS_2 denote the error terms of squares for models 1 and 2 ($SS_1 > SS_2$).

This statistic is distributed as the F statistic with $k_2 - k_1$ on $Lp - k_2$ degrees of freedom.

Having selected the most appropriate form of the second-stage model, comparisons with alternative forms of the error matrix were made by appropriate modification of the data used in the computer program. As noted earlier, the homogeneous independent form using the pooled data could not be estimated with this program and was estimated separately using ordinary least squares.

For alternative forms of the error matrix, the relative efficiency of the model can be gauged by the ratio of the generalized variances, which are determinants of the variance-covariance matrices for the final coefficients:

$$\text{R.E.} = \frac{|\text{Var}_1\,(A)|}{|\text{Var}_2\,(A)|} \tag{24}$$

where R.E. denotes efficiency, measured relative to the heterogeneous correlated form (model 1),

$\text{Var}_1\,(A)$ and $\text{Var}_2\,(A)$ denote the estimated variance-covariance matrices for the regression coefficients in model 1 and model 2 (the alternative).

RESULTS

Heterogeneous and Correlated Error Matrix.—A large number of alternative models, using different forms of the Z_i submatrices, were estimated by generalized least squares, based on the assumption of heterogeneous and correlated error terms. The complete details are too voluminous to report in detail but selected results are summarised in Table 4.

The joint dependent variables shown in Table 4 are the coefficients estimated from the first-stage model and are associated with the intercept term, the reciprocal of age term, and the reciprocal of age squared term respectively (see equation 3). The Z matrix variables in Table 4 comprised a unit vector (1), the logarithm of site index ($\ln s_i$), and the stocking at age 10 years (n_i).

The results for model 1 in Table 4 show that the estimated coefficients for the stocking variable (n_i) were not significantly different from zero in the case of the

TABLE 4. Results of GLS estimation of alternative models.

Model	Z matrix variables	Joint dependent variables $\hat{b}_{1i}$	$\hat{b}_{2i}$	$\hat{b}_{3i}$	Error sums of squares (ss)
1	1	*	*	*	
	$\ln s_i$	*	*	*	30.35
	n_i	n.s.	*	n.s.	
2	1	*	n.s.	*	
	$\ln s_i$	*	*	*	34.66
	n_i		*		
3	1	*		*	
	$\ln s_i$	*	*	*	36.76
	n_i		n.s.		
4	1	*		*	
	$\ln s_i$	*	*	*	40.66
	n_i				

* denotes significant zeta value @ 95 percent probability level;
n.s. denotes not significantly different from zero;
blank denotes the variable was not included.

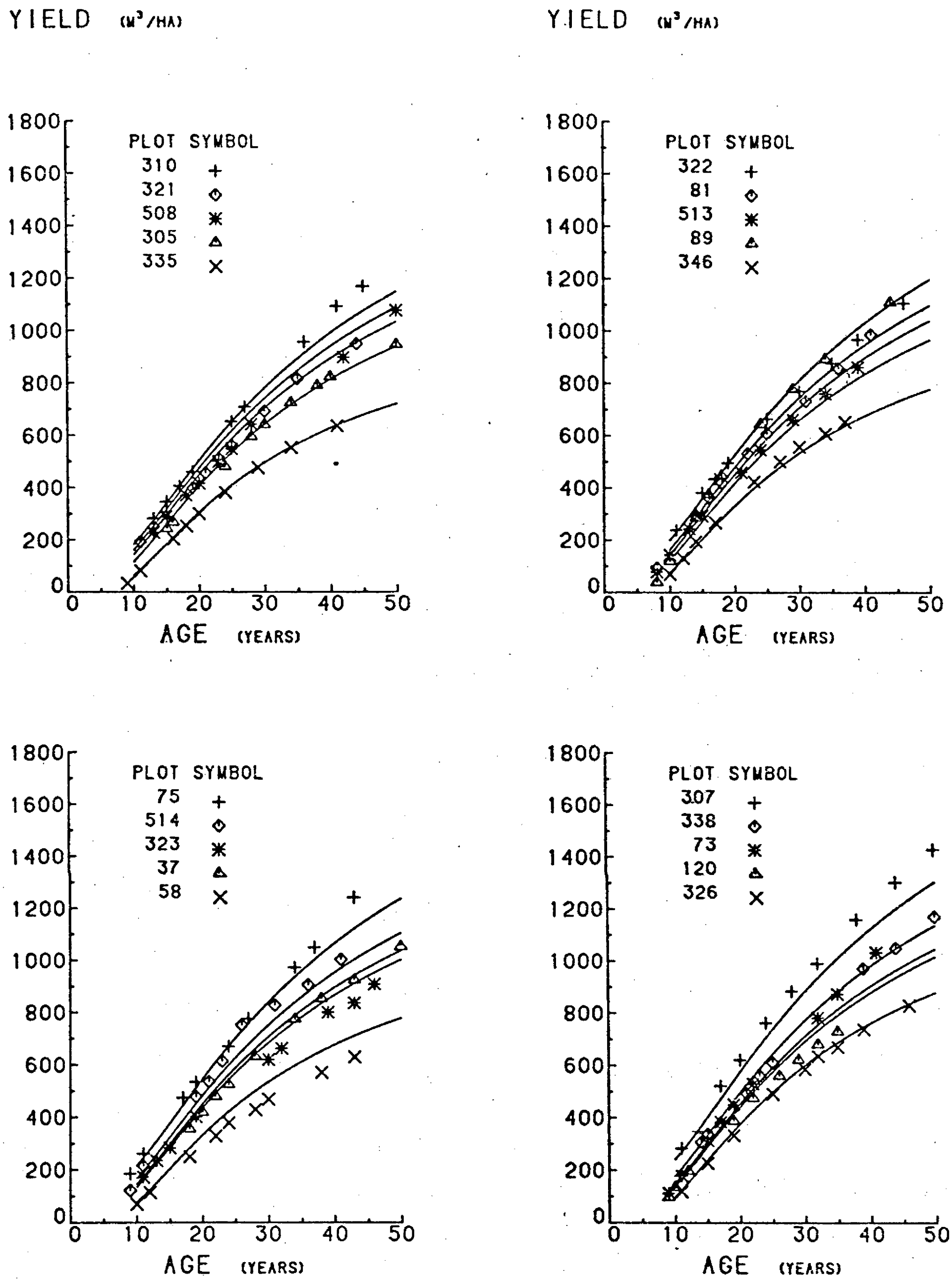


Figure 1. Predicted surfaces and actual values of yield.

first and third dependent variables. Models 2 and 3 show the results for various intermediate deletions and inclusions in the Z matrix, culminating in model 4. All the estimated coefficients were significantly different from zero in model 4, and F tests (see equation 22) showed that the other models were not significantly different from it.

Model 4 was therefore selected as the best model. Expanding both the first and second stages this model can be written:

TABLE 5. Results for alternative forms of error matrix.

Model	Independent variables					Relative efficiency
	1	$\ln s_i$	$\ln s_i/a_i$	$1/a_i^2$	$\ln s_i/a_i^2$	
Heterogeneous correlated	5.155 (.205)[1]	.485 (.009)	–.648 (.0013)	–4.637 (.341)	1.061 (.016)	1.0
Homogeneous correlated	5.161 (.206)	.482 (.009)	–.642 (.0013)	–4.653 (.347)	1.060 (.016)	.92
Heterogeneous independent	4.481 (.638)	.621 (.026)	–.645 (.0013)	–7.127 (6.114)	1.561 (.252)	.00014
Homogeneous independent	4.485 (.639)	.619 (.026)	–.642 (.0013)	–7.136 (6.184)	1.563 (.255)	.00014
Pooled data OLS	5.029 (.012)	.496 (.0005)	–.578 (.0005)	–5.073 (.056)	1.079 (.003)	

[1] Estimated variances are shown in parentheses below the respective coefficient.

$$\ln v_i = \underset{(.45)}{5.155} + \underset{(.09)}{0.485} \ln s_i - \underset{(.04)}{0.648} \ln s_i/a_i - \underset{(.58)}{4.637}/a_i^2 + \underset{(.12)}{1.061} \ln s_i/a_i^2 \quad (25)$$

where the figures in brackets are the standard errors for the coefficient concerned.

Using the results for the model based on heterogeneous and correlated errors (equation 25), graphs of the predicted surfaces and actual values of yield were prepared for the 20 plots (Fig. 1).

The graphs show that the predicted surfaces provided an excellent visual fit for 9 of the plots. The predicted surface tended to deviate from the actual values for three or so of the observations at older ages in 4 of the plots, although the fit was still generally tolerable. In three plots (73, 89, 307), the predicted surfaces consistently underestimated actual yields and in 4 plots (58, 120, 321, 323) they consistently overestimated the actual values, sometimes markedly.

The poor performance of the predicted surfaces in 7 plots is, we believe, related to differences in soil types and/or soil-water regimes. However there were insufficient plots available in the various types to prove this in the present data set. No method of estimation can overcome this type of problem unless additional data or information are available.

Other Forms of the Error Matrix.—A model identical in form to that in equation (25) was re-estimated using alternative assumptions about the nature of the error matrix. The results are shown in Table 5.

The independent variables shown at the top of Table 5 are identical to those in equation (25), unity being used to indicate the intercept term. The estimated values of the coefficients for the heterogeneous correlated model shown in equation (25) are repeated in Table 5, with the estimated values of the respective *variances* shown in brackets immediately below them. This model formed the basis for comparisons with other models (see equation 24) and therefore has a relative efficiency of 1.0.

When the error matrix was assumed to be homogeneous but still correlated, the estimates calculated from equation (18) had a relative efficiency of .92 and the coefficient values differed little from those for the heterogeneous correlated model.

When the error matrix was assumed to be either heterogeneous independent or homogeneous independent, the estimated values of the coefficients (see equations 19 and 20) diverged from those for the heterogeneous correlated model; the differences being well beyond the limits of the confidence intervals in some cases. Relative efficiency also dropped to .00014, indicating the dramatic decrease in precision for the estimates from these two models. The results of the homogeneous independent model correspond with those which would be obtained by applying ordinary least squares to the data for each first-stage coefficient separately. These results highlight the dangers and inefficiency which can arise from this procedure.

The last model in Table 5 was estimated by pooling all data from all plots. The model was then re-estimated using ordinary least squares, assuming the error term to be of homogeneous variance and independent both across and within plots. The estimated values of the coefficients for this model differed little from those of the heterogeneous correlated model, which suggests that the ordinary least squares estimates are reasonably robust under this assumption. However, the estimates of the variances were markedly lower than those for the heterogeneous correlated model and reflect a serious bias in the ordinary least squares results.

This bias stems from the implicit assumption that the pooled observations constitute a random sample whereas the observations within a plot are clearly related. Thus the inherent variation in the pooled data is less than would be the case for a truly random sample, leading to a gross underestimate of the sampling variance attached to the coefficients. As might be expected, scatter plots of the residuals for this model showed a marked serial correlation, in that nearly all the residuals within any one plot had consistent signs.

The underestimation of variances has serious implications. In estimating growth models, hypothesis testing is invariably carried out to choose between alternative models. Underestimation of the variances is likely to result in misleading results from these tests, especially if stepwise regression is used. Under these circumstances the method of model selection may not be nearly as robust as the estimates of the coefficients.

Conclusions

The analysis of yield data from remeasurements of permanent plots spanning a long period of time has a long history in forestry. Prior to the development (in a readily accessible form) of statistical techniques such as multiple regression, the established practice was to fit the data for each plot by eye. A process of "harmonization" was then employed to achieve sensible trends across all plots and sites.

This practice was well founded because, within the limits of the techniques available, it attempted to make the most effective use of the data available. The generalized least squares technique described in this paper represents a rigorous extension of this practice; rigorous in the sense that it is designed to be statistically efficient in the use of the data.

As with the graphical process of harmonization, generalized least squares gives more weight to the plots with the least variable trends. Similarly it recognizes the interrelationships between the parameters of the function. If the plot intercept is to be changed, its impact on other plot coefficients is taken into account and vice versa. Finally it gives due weight to the distinction between observations on plots and those within plots.

The development of computer packages for multiple regression analysis spawned a new wave of estimation of yield functions. Most of this work has focused solely on the deterministic structure of the model; the structure of the error term being largely neglected. The results of this study show that this can be a potentially dangerous course to follow in the case of remeasurements of permanent plots.

Where sufficient remeasurements are available, generalized least squares offers a new and efficient technique for estimation of yield functions.

Literature Cited

AITKEN, A. C. 1934–1935. On least squares and linear combinations of observations. Royal Society of Edinburgh Proc 55:42–48.

BAILEY, R. L., and J. L. CLUTTER. 1974. Base-age invariant polymorphic site curves. Forest Sci 20:155–159.

DHRYMES, P. J. 1970. Econometrics. Harper and Row, New York. 592 p.

FEARN, T. 1975. A Bayesian approach to growth curves. Biometrika 62:89–100.

GRIZZLE, J. E., and D. M. ALLEN. 1969. Analysis of growth and dose response curves. Biometrics 25:357–381.

GOLDBERGER, A. S. 1964. Econometric Theory. Wiley, New York. 399 p.

LEAK, W. 1966. Analysis of multiple systematic remeasurements. Forest Sci 12:69–73.

MILES, J. A., I. S. FERGUSON, and J. W. LEECH. 1978. Program GLS—Generalized least squares for two-stage model development. Forest Sci 24:42.

POTTHOFF, R. F., and S. N. ROY. 1964. A generalized multivariate analysis of variance model useful especially for growth curve problems. Biometrika 51:313–326.

PIENAAR, L. V., and K. J. TURNBULL. 1973. The Chapman-Richards generalization of Von Bertalanffy's growth model for basal area growth and yield in even-aged stands. Forest Sci 19:2–22.

RAO, C. R. 1965. The theory of least squares when parameters are stochastic and its application to the analysis of growth curves. Biometrika 52:447–458.

ROSENBERG, B. 1973. Linear regression with randomly dispersed parameters. Biometrika 60: 65–72.

SWAMY, P. A. V. B. 1971. Statistical inference in random coefficient regression models. Lecture Notes in Operational Research and Mathematical Systems No. 55, edited by M. Beckman, Springer-Verlag.

THEIL, H. 1971. Principles of Econometrics. Wiley, New York. 736 p.

ZELLNER, A. 1962. An efficient method for estimating seemingly unrelated regressions and tests for aggregation bias. J Am Statist Assoc 57:348–368.

Program GLS—Generalized Least Squares for Two-Stage Model Development

Note by J. A. Miles, I. S. Ferguson, and J. W. Leech

PROGRAM GLS estimates the second stage of the two-stage model described by Ferguson and Leech (1978). Estimation of the first stage can be carried out using any standard regression program, provided it can be modified to write out the data in the format required for the second-stage analyses.

The GLS user must supply an external procedure, coded in a suitable language, specifying transformations of the exogenous variables. Each transformed variable is scaled by GLS to improve the conditioning of the weighted moment matrix and thus reduce the possibility of numerical instability in the results. GLS was coded in ALGOL to take advantage of the dynamic run-time storage allocation.

Second-stage models involving 20 plots, 3 first-stage parameters per plot, and up to 9 second-stage parameters, took less than 2 seconds run-time on a Univac 1100/42 under EXEC 8.

Copies of the program, together with user instructions, test data and output, can be obtained from Dr. I. S. Ferguson, Department of Forestry, Australian National University, P. O. Box 4, Canberra, A.C.T. 2601, Australia

Literature Cited

FERGUSON, I. S., and J. W. LEECH. 1978. Generalized least squares estimation of yield functions. Forest Sci 24:27–42.

The authors are with the Department of Forestry, Australian National University, Canberra, A.C.T., Australia. Manuscript received November 17, 1977.