STUDIES ON THE GROWTH AND YIELD OF PLANTATION *EUCALYPTUS GLOBULUS* IN SOUTH WEST WESTERN AUSTRALIA.

BY
GARY INIONS

DECLARATION

This thesis is submitted to the University of Western Australia in fulfilment of the requirement for the degree of Doctor of Philosophy.

Except where acknowledged the work presented in this document is my own.

Gary Brian Inions

February 1992.

ABSTRACT

This thesis details a series of studies on mensurational aspects of plantation grown *E.globulus* Labill. plantations in south west Western Australia. The relationships between extrinsic environmental attributes and site index and the polymorphism evident within the top height development patterns are also examined.

Chapter two gives the results of a study to derive single tree volume equations. Four hundred and thirty sample trees, from 60 plantations, were destructively sampled and their volumes determined. These data were used to parametrize ten functional forms incorporating diameter at breast height (dbh) and total tree height. The equations were validated on an independent data set (n=112). Of the candidate equations presented, the most accurate and precise equations for estimating the merchantable volume under bark of plantation *E.globulus* are the generalized combined variable functional form and the logarithmic functional form. The validation statistics of the logarithmic functional form are significantly improved ($p<0.0001$) by application of Sprugel's (1983) correction factor for logarithmic transformations. If a height measurement is unavailable a second order polynomial equation using only dbh was found to be the most applicable.

The results of a top height development and site index study are detailed in Chapter 3. Stem analysis of 87 site trees, from 57 plots yielded 480 sets of observations of top height (H), site index (S) and age (A) data. The data were used to derive top height development curves via the algebraic difference, the parameter prediction and the Ek-Payandeh methods. Site index equations were also developed. The polymorphism evident in the top height development data was examined via hierarchical agglomerative cluster analysis using the two dimentional profile association metric, described by Fatih *et al.* (1985). This algorithm takes into account the time series relationships of top height data. Plots were clustered such that plots within any cluster display anamorphic top height development patterns while plots of different cluster membership display polymorphism.

The Ek-Payandeh modification of the Chapman-Richards functional form adequately describes the top height growth data while the algebraic difference form of the Chapman-Richards functional form adequately describes the site index data.

The relationships between site index and the polymorphism, encapsulated within the cluster groups, and 74 edaphic physical and chemical, topographic and climatic environmental attributes, sampled from 56 plots, were explored using the multivariate data exploration techniques of semi strong hybrid multidimensional scaling and rotational correlation. Both site index and polymorphism were shown to be related to the ordinate space defined by the environmental variables.

Equations which estimate site index from environmental attributes were constructed using the information derived during data exploration. The derived equation accounted for c.78% of the variation in site index. An equation derived via the application of a stepwise variable selection algorithm accounted for c.80% of the variation in site index. Cluster groups, encapsulating the polymorphism inherent in the data, were separated using environmental variables and discriminant analysis. Groups were also separated using heuristic rules and logistic regression. For maximum separation of groups (c.87%) site index is required along with environmental attributes.

The ability of the equations and allocation techniques described were validated on independent data. All equations recommended validate with acceptable levels of accuracy and precision.

A noteworthy exception was the equation derived via the stepwise variable selection algorithm where no relationship exists between observed and predicted site index. Although this equation yields the most desirable model statistics, it yields the least desirable validation statistics, casting doubt upon the common practise of selecting such attributes by stepwise procedures.

The study presented in the final chapter aimed to develop relationships between yield and attributes of the stand itself, and to assess the influence of the previous landuse on such relationships. Relationships developed, using data from 213 plots gained from the measurement of 7131 trees, showed that yield could be predicted adequately from a function of stand basal area (B) and top height (H). There was no need to separate the data into previous landuse categories (i.e. plots on land that was previously pasture and plots on land that was previously forest). If no measurement of B is available yield is best predicted from functions of stocking (N), S, H and A, if the plot is located on land that was previously pasture, and H, N, and A if the plot is located on land that was previously forest.

All equations recommended in chapter six were validated on independent data and had acceptable levels of accuracy and precision.

ACKNOWLEDGEMENTS

In the preparation of this document numerous people freely gave me assistance and support. I am indebted to the secretarial staff of CALM's Crawley office, particularly Ms Sharon Eccleston, Ms Stella King and Ms Melissa Schifferli, for patiently typing the copious drafts of this document. Secondly, to the staff in general of CALM's Crawley office for providing a very pleasant work environment complimentary to this pursuit.

I am also indebted to Ms D Burton who patiently prepared the frequently altered tables of this thesis, and to Ms C Thomson, for her proof reading endeavours, Ms M Wilke and Ms R Mundy were kind enough to finish these works by desk-top publishing the final document.

The assistance and advice of Mr H Cochane, of the University of Western Australia's Soil Science Department, in the preparation of the particle size analysis methodology was very much appreciated.

The jovial and pleasant manner in which Professor Ian James of Murdoch University, advised me on matters of a statistical nature will be sorely missed.

I am also indebted to the hard-core band of assistants who worked with me during the boring, wet and laborious process of data collection. Particular thanks are due to Tim Duncanson and Gavin Hamilton who, regardless of near miss tree felling events, chainsaw mishaps, frequent rainstorms and hostile farmers, never failed to lose their good nature or humour.

Dr Robert Monserud of the USDA's Intermountain Research Station, Idaho, provided very useful criticism of Chapter Three. More importantly, he influenced the outcome of this work through his unassuming yet very thought provoking views on research and science and by example, through his excellent studies on Douglas fir.

I am also very indebted to my supervisors, Dr David Bell and Dr William Loneragan, for criticisms of drafts and their assistance throughout this project.

Finally, I would like to express my gratitude to the Executive Director of CALM, Dr Syd Shea, for enabling me to undertake this project and for consistently supporting the study throughout its duration. For such an opportunity I will always be grateful.

TABLE OF CONTENTS

Title		*Page*

Title		*Page*

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

CHAPTER ONE

General Introduction

1.1 INTRODUCTION

This thesis presents studies on the mensurational aspects of *E. globulus* Labill. for use in the hardwood afforestation project of the Department of Conservation and Land Management (CALM). It is useful to outline the emergence of this project, prior to the presentation of the studies, so as to place perspective on their possible applications. As the project and this study are based entirely upon plantations of *Eucalyptus globulus* Labill. a brief description of the species is also included.

1.1.1 THE DEVELOPMENT OF CALM'S HARDWOOD AFFORESTATION PROJECTS IN SOUTH WEST WESTERN AUSTRALIA

Two main factors were instrumental in conceiving CALM's hardwood afforestation project. Firstly, the perceived increase in the demand for hardwood fibre within the Pacific Rim (PACRIM) countries. Secondly, the environmental problems attributable to the removal of native vegetation for agricultural purposes had and is, achieving public prominence.

1.1.1.1 HARDWOOD FIBRE TRADE WITHIN THE PACRIM

The Pacific ocean countries trade over 10 000 000 oven dry tonnes (ODT) of wood fibre each year (Hagler 1991). Currently, 60% of this trade is among the northern hemisphere partners, particularly Japan and the United States of America (USA). Japan is the largest importer of hardwood fibre in the world and is the world's second largest producer of paper and paperboards (FAO 1990). However, Japan is almost entirely dependent upon the import of raw material for its paper industry. Only 4% of Japan's printing and writing papers are imported (Streeting and Imber 1991). In recent years both Taiwan and South Korea have emerged as importers of hardwood fibre and together with Japan, account for over 95% of hardwood fibre imports within the PACRIM countries.

Despite the recent downturn in the demand for pulp and paper products and the emergence of the recycled paper industry, the demand for hardwood fibre in the PACRIM is expected to increase. Estimates of the increase in demand vary from 4% to 25% (c.f. FAO 1986; Gibson 1989; Simons 1990; Hegler 1991; Streeting and Imber 1991).

Currently, Australia supplies 42.7% of imported hardwood fibre to the PACRIM countries. The USA supplies 29.1% while Chile and South Africa account for about 5% each (Hegler 1991). Practically all of this supply originates from the native forests of the country of origin.

The ability of these traditional suppliers of hardwood fibre to continue to meet the current demand levels and/or to escalate to meet the future demand, is uncertain (Groome 1989; Hegler 1991). Consequently investments, predominantly from Japanese sources (Anon 1991), are promoting the establishment of hardwood plantations in such countries as Thailand, Vietnam and Chile and are specifically for the supply of hardwood fibre to the PACRIM.

CALM, through the development of its hardwood afforestation project, is attempting to capture a market share of the perceived deficit in supply and attract investment currently injected into other countries of the Pacific Rim.

1.1.1.2 ENVIRONMENTAL PROBLEMS CREATED BY THE REMOVAL OF TREES FROM THE LANDSCAPE

Salinization of waterways and land has emerged as a major environmental and economic problem in the south west of Western Australia. Currently, only 48% of the divertable surface water resources remain fresh (<500 mgL^{-1} total soluble salts) (Western Australian Water Resources Council 1986) and about 440 000 ha of once productive farmland has become salt affected (Schofield and Bari 1991).

Salinization is attributable indirectly to the clearing of native vegetation for agricultural purposes (Wood 1924; Williamson *et al.* 1987). The replacement of deep rooted perennial native vegetation with the shallow rooted annual vegetation used in agriculture, alters the water balances of a site to favour ground water recharge (Peck and Williamson 1987; Schofield *et al.* 1988). The watertable eventually intercepts the surface and discharges the salt naturally stored within the profile (Williamson *et al.* 1987; Ruprecht and Schofield 1991). Other environmental problems indirectly attributed to the removal of native vegetation include, wind and water erosion, waterlogging, deterioration of soil structure and the decline of remnant native vegetation (Fitzpatrick 1983; Conacher 1990).

Methods to restore degraded land are many and each vary widely in their degree of acceptance (Conacher 1990). Most centre on reafforestation (Schofield *et al.* 1989; Schofield and Bari 1991). Consequently, any commercial scheme which seeks to reafforest cleared agricultural land receives community and political support.

It was from the background outlined above, with a perceived world deficit in hardwood fibre on one hand and a demand for trees to be returned to the rural landscape on the other, that CALM launched its hardwood afforestation project in 1988. Basically, the project acquires land from private owners where the landowner enters into an agreement with the State. The land is planted to *E. globulus* by the State and the landowner receives an annuity payment for the use of the land and/or a proportion of the harvest revenue. About 8 000 ha of *E. globulus* plantations have been established up until 1991. The project will escalate its planting program to c. 2 000 ha a^{-1} from 1992 on.

1.1.2 THE SPECIES

Since the first description by Labillardiere in 1800, the taxonomic history of the species *Eucalyptus globulus* Labill. has been controversial. Recently three taxa, that were formerly regarded as species, were reduced to subspecies of *E. globulus* (Kirkpatrick 1974). This latest rearrangement is generally accepted (Chippendale 1976). Using the nomenclature of Pryor and Johnson (1971) the species belongs to the *Eucalyptus* subgenus *symphyomyrtus* section *maidenaria*. The species *E. globulus* is comprised of the following four subspecies;

i *Eucalyptus globulus* Labill. subsp. *globulus*.

ii *Eucalyptus globulus* Labill. subsp. *bicostata* (Maiden *et al.*) Kirkpatr.

iii *Eucalyptus globulus* Labill. subsp. *pseudoglobulus (*Naudin ex Maiden) Kirkpatr.

iv *Eucalyptus globulus* Labill. subsp. *maidenii* (F. Muell) Kirkpatr.

Although Turnbull and Pryor (1984) suggest that there is little difference in the wood characteristics between the four subspecies other authors have shown otherwise. (Turner *et al.* 1983; Dean *et al.* 1990). Only *Eucalyptus globulus* Labill. subsp. *globulus* is used in the hardwood afforestation project discussed above.

The natural distribution of this subspecies is along the east coast of Tasmania, usually within 20 km from the ocean. It also occurs on Flinders and King Islands in Bass Strait and on Wilsons Promontory, Cape Otway and in the Strzelecki Ranges in southern Victoria. The latitudinal range is from 38°30' to 43°30'S while the altitudinal range is from sea level to c. 450 m a.s.l. (Kirkpatrick 1975).

1.1.3 THE USES OF *E. GLOBULUS*

E. globulus is not a species from which high quality sawn timber is derived. Its timber collapses and is prone to surface checking during drying (Campbell and Hartley 1984). However, the timber of this species is used for poles, piles, sleepers, fenceposts and mining timbers (Hall *et al.* 1970; Boland *et al.* 1984; Turnbull and Pryor 1984).

In the developing countries, where it is estimated that over 1000 million people experience fuelwood shortages (CIRC 1982; FAO 1983), *E. globulus* is grown extensively for fuelwood (Jacobs 1981; Gasana 1983; Pohjonen and Pukkala 1990).

The composition of leaf oils of Eucalypts are of commercial value and vary widely between species and subspecies (Sharma and Handa 1982). The different oils have three types of uses in the medicinal, perfumery and industrial markets (Small 1981). *E. globulus* is grown for the extraction of leaf oils. The leaf oils are present in amounts of up to 3% and contain c. 75% cineole. Usually these oils are refined before use in antiseptics, inhalants and embrocants (Baslas and Saxena 1984; Hillis 1984; Dayal and Ayyar 1986). The leaf oils of *E. globulus* form the bases of commercial industries in South Africa, Portugal, Spain, Brazil and China. The small leaf oils industry in Australia does not use *E. globulus* (Small 1981).

The major use of *E. globulus* is for the production of printing and writing papers. It is noteworthy that the first reported pulping of a Eucalypt was in Portugal in 1906 where *E. globulus* was used to produce sulphite pulps. This led to the commercial pulping of the species in 1919 (Watson and Cohen 1969; Algar 1988).

The wood quality attributes important for the pulp and paper industries include basic density, fibre length, extrative content and permeability (Hillis 1972; Higgins 1984). These attributes are expressed very favourably in *E. globulus* and the species is valued for the high quality of pulp and paper yielded (Cromer and Hansen 1972; Farrington *et al.* 1977; Turner *et al.* 1983). Kibblewhite *et al.* (1991) notes that of the hardwood pulps traded throughout the world the pulp of *E. globulus* combines the most important pulp and sheet qualities in a remarkably favourable manner. They give good wet and dry strength properties, good formation due to the short stiff fibres and excellent bulk and optical properties. Likewise, Dean *et al.* (1990) note that *E. globulus* has good surface properties, density, stiffness, opacity and in particular excellent sheet formation.

E. globulus exhibits rapid growth and a wide ecological tolerance to site conditions (Beadle and Inions 1990). This coupled with its ability to produce high quality pulp and paper has contributed to this species being one of the most extensively planted Eucalypts. The species forms the bases of the pulp and paper industries of Portugal, Spain and Thailand while being a major contributor to the industries in Brazil, South Africa, Chile and Argentina (Turnbull and Pryor 1984).

The species is particularly successful in countries with a mediterranean-type climate free of severe frost and prolonged summer dry seasons (Booth *et al.* 1988; Booth and Pryor 1991).

1.1.4 THESIS ORGANISATION.

The major objective of the studies, which collectively comprise this thesis, is to study and quantify mensurational aspects of plantation *E. globulus*. The relationships between some of these aspects and the environment are also pursued. As such the thesis is presented as a series of five related chapters:

Chapter two details a study to derive single tree volume equations. The study tests a number of functional forms for their suitability to the species and examines the effect of some parameter estimation techniques on validation statistics.

Chapter three presents a study which aims to yield top height development and site index curves. A number of methods are employed and compared. A method to account for polymorphism is also pursued.

The use of multivariate data exploration techniques, used to examine the relationships between the top height development pattern and site index and environmental attributes, is presented in Chapter four. Quantitative equations which predict site index and top height development patterns are also developed.

The equations derived and presented in Chapters three and four are validated in Chapter five. Validation of the equations used in isolation and in concert is undertaken on independent data sets.

Chapter six presents a study aimed at developing stand level yield equations. Mensurational aspects detailed in other chapters are employed in Chapter six. The resulting equations are validated on independent data sets.

Each Chapter is preceded by a literature review on the topic pertinent to that chapter.

1.1.5 REFERENCES

Algar, W.H. 1988. Forestry and forests products. *In:* Technology in Australia 1788-1988. Academy of Technological Sciences and Engineering. Parkville.

Anonymous. 1991. Stabilizing the timber market. *Asian Pacific* Forest *Industries*, 2: 6-9.

Baslas, R.K., and S. Saxena. 1984. Chemical examination of essential oil from the fruits of *Eucalyptus globulus* Labill. *Herba Hungarica*, 23: 21-23.

Beadle, C.L., and G.B. Inions. 1990. Limits to growth of Eucalyptus and their biology of production. *In:* Prospects for Australian plantations. J. Dargavel and N. Semple (eds.) ANU. Canberra.

Boland, D.J., M.I.H. Brooker, G.H. Chippendale, N. Hall, B.P.M. Hyland, R.D. Johnson, D.A. Kleining, and J.D. Turner. 1984. Forest trees of Australia. Nelson, CSIRO. Melbourne.

Booth, T.H., and L.D. Pryor. 1991. Climatic requrements of some commercially important eucalypt species. *For. ecol. Manage.*, 43: 47-60.

Booth, T.H., H.A. Nix, M.F. Hutchinson, and T. Jovanovic. 1988. Niche analysis and tree species introductions. *For. Ecol. Manage.*, 23: 47-59.

Campbell G.S., and J. Hartley. 1984. Drying and dried wood. *In:* Eucalyptus for wood production. Hillis, W.E., and A.G. Brown (eds.). CSIRO. Sydney.

Chippendale, G.M. 1976. Eucalyptus nomenclature. *Aust. For. Res.*, 7: 69-107.

CIRC. 1982. Annual Report. Centre for international research cooperation. CSIRO, Melbourne.

Conacher, A.J. 1990. Salt of the earth: Secondary soil salinization in the Australian wheat belt. *Environment*, 32: 5-42.

Cromer, R.N., and N.W. Hansen. 1972. Growth, nutrient uptake and pulping characteristics of young *Eucalyptus globulus*. *Appita*, 26: 187-190.

Dayal, R., and K.S. Ayyar. 1986. Analysis of medicinal oil from *Eucalyptus globulus* subsp. *bicostata* leaves. *Planta Medica*, No. 2.

Dean, G.H., J. French, and W.N. Tibbits. 1990. Variation in pulp and papermaking characteristics in a field trial of *Eucalyptus globulus*. Presented at 44th Appita general conference. Rotorua, New Zealand. 1990.

FAO. 1983. Fuelwood supplies in developing countries. FAO *For. Pap.* 42.

FAO. 1986. Projections of worldwide demand and supply. Implications for capacity and world trade. Rome.

FAO. 1990. Forest Product Yearbook, 1977-88, Rome.

Farrington, A., N.W. Hansen, and P.F. Nelson. 1977. Utilization of young plantation *E. globulus*. *Appita*, 30: 305-309.

Fitzpatrick, E.N. 1983. Land degradation and farming systems in Western Australia: An overview. *In:* Land: The fragile resource. Proceedings of the land management society of Western Australia Conference. Perth, September. 1983.

Gasana, J.K. 1983. Site factors affecting growth of maiden's gum, *Eucalyptus globulus* Labill. subsp. *maidenii* (F.Muell.) Kirkpatr. in Rwana. PhD thesis. Univ. of Idaho.

Gibson, B.A.M. 1989. Wood resources and opportunities for utilization in Australia. *In:* Proceedings of the first international conference on investment and trade opportunities in pulp and paper on the Pacific Rim. Sydney, Australia. 1989.

Groome, J. 1989. Demand and supply of raw materials in East Asia. *In:* Proceedings of the first international conference on investment and trade opportunities in pulp and paper on the Pacific Rim. Sydney, Australia. 1989.

Hagler, R.W. 1991. Wood fibre trade in the Pacific Rim - a shift southward. Paper to the Third International Asian Conference for the Paper Industry. Tokyo. Japan. May 7-9, 1991.

Hall, N., R.D. Johnson, and G.M. Chippendale. 1970. Forest trees of Australia. Aust. Govt. Publ. Serv. Canberra.

Higgins, H.G. 1984. Pulp and paper. *In:* Eucalypts for wood production. Hillis, W.E. and A.G. Brown (eds.). CSIRO. Sydney.

Hillis, W.E. 1972. Properties of eucalypt woods of importance to the pulp and paper industry. *Appita*, 26: 113-122.

Hillis, W.E. 1984. Chemicals. In: Eucalypts for wood production. Hillis, W.E., and A.G. Brown. (eds.) CSIRO. Sydney.

Jacobs, M.R. 1981. Eucalypts for planting. Food and Agricultural Organiation. Rome.

Kibblewhite, R.P., A.D. Bawden, and M.C. Hughes. 1991. Hardwood market kraft fibre and pulp qualities. *Appita*, 44: 325-332.

Kirkpatrick, J.B. 1974. The numerical intraspecific taxonomy of *Eucalyptus globulus* Labill. (Myrtaceae). *Bot. J. Linn. Soc.*, 69: 89-104.

Kirkpatrick. J.B. 1975. Natural distribution of *Eucalyptus globulus* Labill. *Aust. Geogr.*, 13: 22-35.

Peck, A.J., and D.R. Williamson. 1987. Effects of forest clearing on groundwater. *J. Hydorl.*, 94: 47-65.

Pohjonen, V., and T. Pukkala. 1990. *Eucalyptus globulus* in Ethiopian forestry. *For. Ecol. Manage.*, 36: 19-31.

Pryor, L.D., and L.A.S. Johnson. 1971. A classification of the Eucalypts. Australian National University. Canberra.

Ruprecht, J.K., and N.J. Schofield. 1991. Effects of partial deforestation on hydrology and salinity in high salt storage landscapes. I. Extensive block clearing. *J. Hydrol.*, in print.

Schofield, N.J., and M.A. Bari. 1991. Valley reforestation to lower saline groundwater tables: Results from Stene's farm, Western Australia. *Aust. J. Soil Res.*, 29: 635-650.

Schofield, N.J., I.C. Loh, P.R. Scott, J.R. Bartle, P. Ritson, R.W. Bell, H. Borg, B. Anson, and R. Moore. 1989. Vegetation strategies to reduce stream salinities of water resource catchments in south-west Western Australia. Water Authority of Western Australia. *Rep.* WS 33.

Schofield, N.J., J.K. Ruprecht, and I.C. Loh. 1988. The impact of agricultural development on the salinity of surface water resources of south-west Western Australia. Water Authority of Western Australia. *Rep.* WS 27.

Sharma, A., and S.S. Handa. 1982. Chemical grouping of some Eucalyptus species. *J. Tree Sci.*, 1: 105-111.

Simons, H.A. 1990. Competitiveness of Australian forest industries. Report to the Forestry and forest products industry council. March 1990.

Small, B.E. 1981. The Australian Eucalyptus oil industry - an overview. *Aust. For.*, 44: 107-177.

Streeting, M., and D. Imber. 1991. The pricing of Australian woodchip exports. Resource Assessment Commission. *Res. Pap.* No. 4.

Turnbull, C.H., V. Balodis, and G.E. Dean. 1983. Variability in pulping quality of *E. globulus* from Tasmanian provenances. *Appita*, 36: 371-376.

Turner, J.W., and L.D. Pryor. 1984. Choice of species and seed sources. *In:* Eucalypts for wood production. Hillis, W.E., and A.G. Brown, (eds.). CSIRO, Sydney.

Watson, A.J., and W.E. Cohen. 1969. Pulping of Eucalyptus - an historical survey. *Appita*, 22: 17-27.

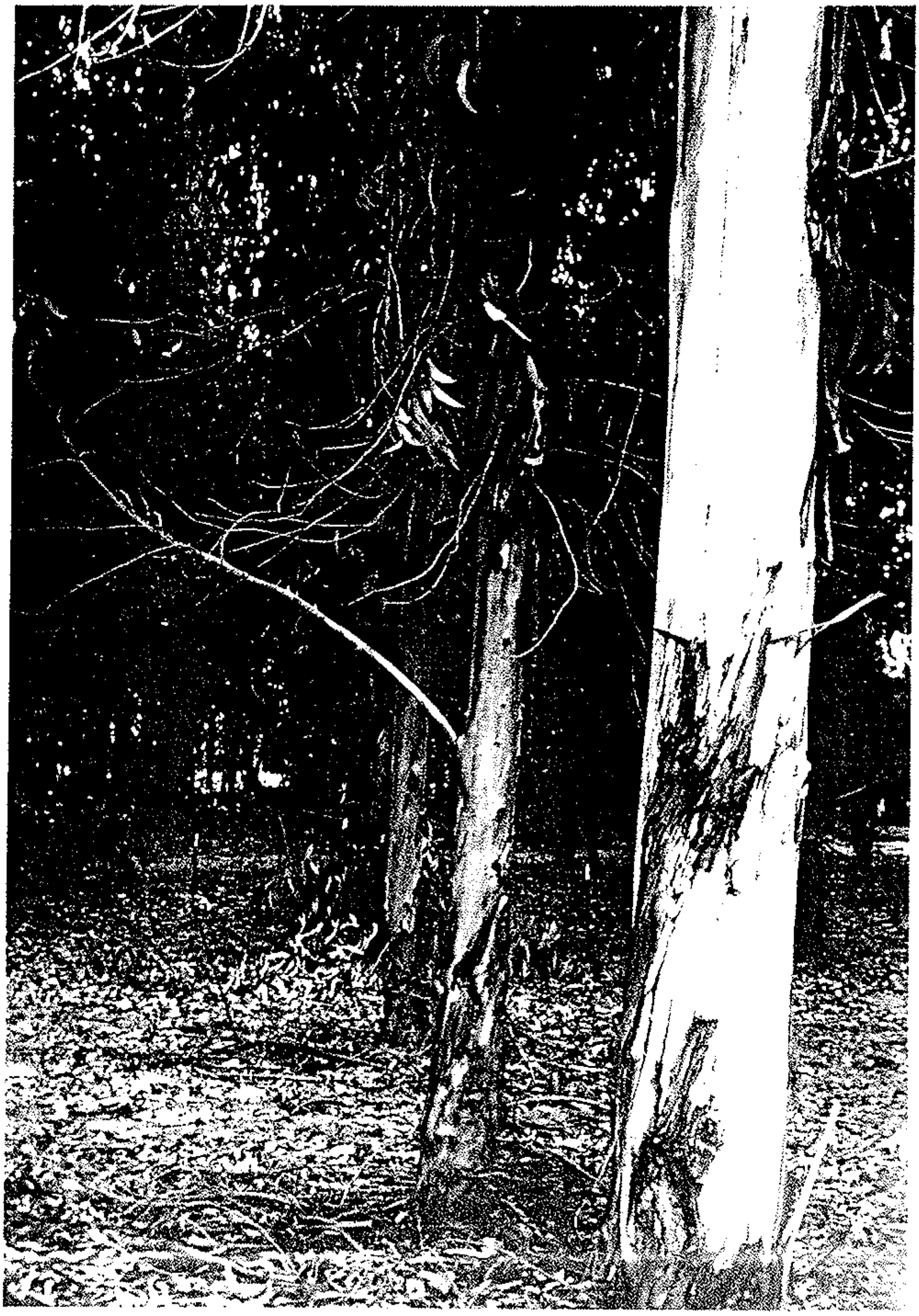
Western Australian Water Resources Council 1986. Water resources perspectives Western Australia: Report No. 2. Water Resources and Water Use, Summary of Data for the 1985 National Survey. *Publ. No. WRC* 7/86.

Williamson, D.R., R.A. Stokes, and J.K. Ruprecht. 1987. Response of input and output of water and chloride to clearing for agriculture. *Hydrol.*, 94: 1-28.

Wood, W.E. 1924. Increase of salt in soil and streams following the destruction of the native vegetation. *J. Roy. Soc. W.A.*, 10: 35-47.

CHAPTER TWO

Estimation of Single Tree Volumes



2.1 LITERATURE REVIEW

2.1.1 INTRODUCTION

Volumes of individual trees can be estimated from functions which are established between some of the more easily measured parameters of a tree, such as the height and diameter at breast height (dbh), and its volume. Such functions are advantageous for forest mensuration and inventory where direct measurement of tree volume is slow and expensive. Single tree volume equations are also useful as primary components of growth models (Munro 1974).

To derive single tree volume equations the volumes of trees in a data set must be known. To derive single tree volumes, stem diameter is measured at regular intervals along its length. The total stem volume is derived by summing the individual volumes of each section, calculated via standard formulae.

2.1.2 CALCULATION OF LOG VOLUMES

Two sources of error contribute to the total error in determining volumes of sectionally measured trees. The first source is introduced when the diameters and lengths of logs are not accurately measured. The second occurs when the mathematical model used to represent the stem or stem segments assumes the shape of the log to be a particular geometrical solid and the form of the log departs from this assumption. For example, if the shape of a log has the form of a frustum of a paraboloid, then the equations of Newton, Huber and Smalian (Husch *et al.* 1972) all provide accurate results. However, if the form of the log departs from parabolic, bias is introduced. Biging (1988) termed this error source model misspecification error.

Studies of the accuracy of model specification compare estimated volumes with "true" volumes, determined by water displacement (Martin 1984) or the sum of the log volumes when the interval between measurements is very small (Goulding 1979). The use of cubic spline functions to portray stem taper and calculate stem volume has also been used (Liu 1980).

Using the water displacement technique to obtain "true" volume, Young *et al.* (1967) examined the accuracy of Smalian's and Huber's formulae when applied to logs of northern hardwoods and softwoods of the U.S.A. They found that for 2.4 m and 4.9 m long logs the average errors associated with Huber's formula (3.5% and -3.7% respectively) were consistently smaller and statistically different from the average errors obtained with Smalian's formula (c. 9.0%). Water displacement was used by Martin (1984) to determine "true" volumes of 243 logs from 75 eastern hardwood trees from West Virginia, U.S.A. Fourteen different equations were then used to estimate these volumes and the results were compared to the true values in both accuracy and precision. Martin found that the formulae of Huber and Newton performed the best followed by that of Smalian. The bias associated with Huber's, Newton's and Smalian's equations were 2.5, 3.9 and 6.9% respectively.

Using the sum of the log volumes when the interval between measurements is very small (usually 0.3 m) to obtain "true volume", Carron and McIntyre (1959) examined the accuracy of Huber's formula when applied to *Pinus radiata* D.Don in the eastern states of Australia. They found errors of -2% for 3 m long logs. The same technique was employed by Goulding (1979) to examine the accuracy of several standard formulae and a spline function when estimating log volumes for *Pinus radiata* in New Zealand. Goulding found that a spline curve had an error that was 60% of the error obtained using Smalian's method. Newton's equation had an error that was 50% of the error associated with Smalian's method. When the distance between measurements was less than 2 m all methods tested had small errors (<2.3%).

Biging (1988) compared the formulae of Smalian, Huber and Newton and a numerical technique using cubic splines. The "true" volume of logs was derived from 2 taper equations for white fir *Abies concolor* (Gord. & Glend.) Lindl. (Lowiana (Gord.)) derived by Biging (1984). Thus the technique facilitated the partitioning of the total error in volume estimation into measurement error and error due to model misspecification. The error due to model misspecification was less than 5% for measurement lengths of 4.9 m for all models tested. Systematic measurement error was estimated at 1 to 4%. Thus total error in volume estimation was less than 9% for all methods tested.

From the literature total errors encountered in estimating volume, assuming a measurement distance of 2.4 to 4.9 m, are approximately 3-9% for Smalian's formula, 3-4% for Huber's formula, 1-4% for Newton's formula and 2-5% for cubic splines. However, as the distance between measurements increases the difference between formulae and the magnitude of the error increases (Brickell 1985).

2.1.3 FUNCTIONAL FORMS OF SINGLE TREE VOLUME EQUATIONS

2.1.3.1 STANDARD FUNCTIONAL FORMS

Single tree volume is usually considered to be a function of tree diameter, usually at breast height, some measure of tree height and an expression of tree form. Less commonly, measures of tree crown or the age of the tree are included. The functional forms of tree volume equations are many and varied. Clutter *et al.* (1983) lists 7 commonly used functional forms as:

i.	Constant form factor	$Y = \beta_1 D^2 H$
ii.	Combined variable	$Y = \beta_o + \beta_1 D^2 H$
iii.	Generalized combined variable	$Y = \beta_o+\beta_1D^2+\beta_2H+\beta_3D^2H$
iv.	Logarithmic	$Y = \beta_1D^{\beta 2}H^{\beta 3}$
v.	Generalized logarithmic	$Y = \beta_o+\beta_1D^{\beta 2}H^{\beta 3}$
vi.	Honer transformed variable	$Y = D^2/(\beta_o+\beta_1H^{-1})$
vii.	Form class	$Y = \beta_o +\beta_1D^2HF$

where; Y = volume

D = diameter at breast height

H = some expression of height

F = an expression of tree form

$\beta_0,\beta_1,\beta_2,\beta_3$ = constants to be estimated.

The constant form factor functional form has been used by Gevorkiantz and Olsen (1955) for prediction of total stem volume of a number of species in the Lake States of the U.S.A. and McNab *et al.* (1985) for prediction of total tree volume of Choctawhatchee sand pine (*Pinus clausa* var. *immuginata* D.B. Ward) in Florida, U.S.A. Green and Strawderman (1986) used this functional form for Stein-rule estimation of the parameters of volume equations for 18 eastern U.S.A. hardwoods. The constant form factor functional form is only suitable where the form of the stem is relatively constant regardless of tree size. As a result this functional form has not been widely used.

The combined variable volume equation has been used to predict merchantable stem volume of slash pine (*Pinus elliottii* Engelm.) in Georgia and Carolina, U.S.A. (Bennett *et al.* 1959) and eastern hardwood trees of Virginia, U.S.A. (Martin 1984). The generalized combined variable functional form has received much attention in the literature and has been used to predict total stem volume of loblolly pine (*Pinus taeda* L.) in Georgia, U.S.A. (Romancier 1961) and trees of the *Cedrus* forests of Morocco (Postaire and M'Hirit 1985).

Because errors associated with tree volume prediction tend to be heteroscedastic (Furnival 1961; Cunia 1964) many equations make use of logarithmic transformations. The logarithmic functional form has been used to predict total stem volume of Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) in British Columbia (Brackett 1973) and loblolly pine in Brazil (McTague and Bailey 1987). Examples of the use of the generalized logarithmic functional form can be found for red pine (*Pinus resinosa* Ait.) in Canada (Newham 1967) and tropical rainforest species in Brazil (Higuchi and Ramm 1985). Horner's transformed variable functional form is an alternative to logarithmic transformation and has been used to predict total stem volume for red pine in Canada (Horner 1965) and forest trees of the Lake States and of Canada (MacDonald and Forslund 1986).

The final class of functional form includes a measure of tree form as an independent variable. Two common measures of tree form often included are the Girard form class and the cylindrical form factor. Girard form class is defined as the diameter (under bark) at the top of the first 4.8 m log, divided by the diameter at breast height outside bark (dbhob). The cylindrical form factor is defined as the ratio of total stem volume to the volume of a cylinder with diameter equal to tree dbh and height equal to the total height of the tree. Form class functional forms are assumed to give greater regional applicability with higher precision (Loetsch *et al.* 1973). On the other hand, Clutter *et al.* (1983) lists the following reasons why functional forms which include measures of tree dbh and height only are preferred:

(i) Measurements of upper-stem diameters are time consuming and expensive;

(ii) Variation in tree form has a much smaller impact on tree volume than height or dbh variation;

(iii) With some species, form is relatively constant regardless of tree size;

(iv) With other species, tree form is often correlated with tree size, so that the dbh and height variables often explain much of the volume variation actually caused by form differences.

An alternative to traditional measures of form is presented by Forslund (1982) who developed a volume equation based on the location of the centre of gravity of the bole and estimated the true volume of individual trees to within 10%, using total height and the diameter at a relative height of 0.3 m. MacDonald and Forslund (1986) modified Forslund's (1982) equation and found it to be more consistently accurate than either the original equation or Horner's transformed variable equation.

A subset of the form class functional form is that of product form, which estimates tree volume from product form and dbh (Smith 1976). Product form is defined as the product of diameter (outside bark) midway along the stem above breast height and total height. However, the optimal height at which a diameter is measured, when calculating product form, varies with species (Roebbelen and Smith 1984).

Whether or not to include a measure of form or product form into volume equations is a matter for conjecture. It has been shown that no practical advantage in volume estimates could be gained from including any measurement of form in addition to the diameter at breast height and the total height (Kozak *et al.* 1969). However, the correctness of this statement will depend on the scale of application of the equation and the nature of the species under study. One approach to this problem is presented by Postaire and M'Hirit (1985) who used cluster analysis to stratify their sample prior to fitting standard functional forms to individual clusters. The clustering criteria was based on measures of tree form.

A variety of functional forms not listed by Clutter *et al.* (1983) have been tabulated by Loetsch *et al.* (1973). The large number of functional forms evident in the literature indicates that there is no functional form which is generally applicable.

2.1.3.2 MISCELLANEOUS FUNCTIONAL FORMS

One deviation from the use of standard functional forms occurs in a study which advocates the prediction of tree volume from a function of tree age and diameter (Sadiq and Smith 1983). The technique circumvents the measurement of tree heights through the inclusion of an age term. The resulting equation predicted the tree volumes of red pine (*Pinus resinosa* Ait.) in Ontario, Canada, more accurately than the standard functional forms tested. A disadvantage of this technique is that the application of the equation is restricted to stands of known age.

Measures of tree crown may also be incorporated into volume equations. Past attempts to include crown variables have met with varying success. While Lohrey (1983) and Farrar (1985) found that the inclusion of crown measures significantly modified the intercept and slope terms of the combined variable functional form, for longleaf pine (*Pinus palustris* Mill.) in the U.S.A., they did not substantially reduce residual variance. A result also supported by Laasasenaho (1982). Crown ratio has been incorporated into four total stem volume equations for Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) as a nonlinear multiplier (Hann *et al.* 1987). One functional form of Hann *et al.* was a "component approach" that divided stem volume into that above and below breast height. The crown ratio term was highly significant in the component functional form, which had the smallest bias and the greatest prediction precision of all equations examined. One disadvantage of this approach is the difficulty and expense involved with obtaining measurements of the crown in the field.

2.1.4 COMPARING FUNCTIONAL FORMS

It is common practice to test a variety of functional forms when fitting a volume equation. A judgement must then be made as to which functional form best fits the aims of the study. Furnival's index (Furnival 1961) is commonly used for comparing equations (Hann *et al.* 1987). However, Green (1983) demonstrated that the ranking of equations according to the index may not be the same as ranking them according to their ability to predict an independent validation data set. Consequently, many studies reserve an independent validation data set to compare individual equations. Predicted values are then compared to known values with the mean of the differences between the actual and predicted values, mean absolute difference and standard deviation of the differences, common comparison statistics (Martin 1984). The mean of the differences is a measure of the accuracy of an estimate while standard deviation of the residuals is a measure of precision. Hann *et al.* (1987) used bias and standard deviation of the residuals as comparison statistics. The mean difference of the residuals was checked with a t-test for departure from zero. The statistics were then combined by taking the square root of the sum of the residual squares to give an overall prediction error. Sadiq and Smith (1983) used Freese's (Freese 1960) Chi- squared accuracy test for comparing the ability of functional forms to predict a validation data set. For a more detailed discussion on validation see Section 5.1.

2.1.5 PARAMETER ESTIMATION

After selection of an appropriate functional form(s), parameter estimation is most commonly accomplished with ordinary least squares. This method assumes homogeneity of variance in the error terms. In most cases this assumption is violated. As previously mentioned logarithmic transformations often alleviates this problem, alternatively many studies choose to weight the equations by $\sqrt{D^2H}$, $1/(D^4H^2)$ or $1/D^4$ (West 1980; Jacobs and Monteith 1981). $1/D^2H$ and

$1/D^2H$ have also been used (Ernst and Hann 1984). Meng and Tsai (1986) define a method for selecting the exponent for the D term for weights using a maximum likelihood function.

Stein-rule estimation is useful for the simultaneous development of multiple volume equations. A stein-rule estimator (Burk and Ek 1982), which shrinks least squares estimates of regression parameters towards their weighted average, was employed to estimate the coefficient in the constant form factor volume equation for 18 species simultaneously (Green and Strawderman 1986). Tests on an independent validation data set revealed that Stein-rule estimates were biased, but predicted better than the corresponding least squares estimates.

2.2 SINGLE TREE VOLUME EQUATIONS FOR PLANTATION *E. GLOBULUS* IN SOUTH WEST WESTERN AUSTRALIA.

2.2.1 INTRODUCTION

Tree volume is usually considered a function of tree diameter and some measure of tree height (McClure *et al.* 1983; Martin 1984; Higuchi and Ramm 1985; Lynch 1988), although crown ratio (Hann *et al.* 1987) and age (Sadiq and Smith 1983) have also been found useful.

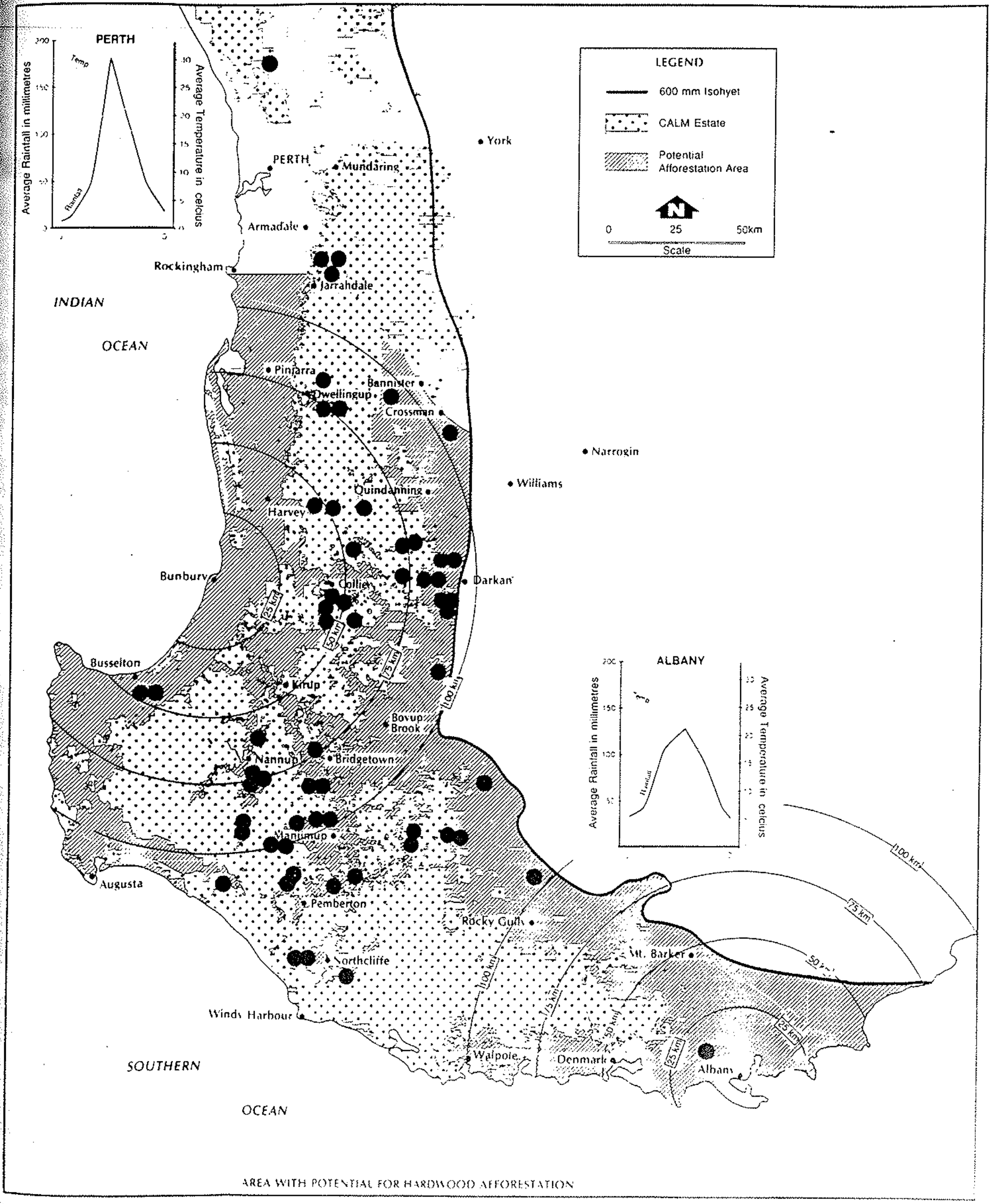
A wide variety of functional forms appear in the literature, some of which are tabulated by Clutter *et al.* (1983) and Loetsch *et al.* (1973). The two lists are not exhaustive and it is evident by the large array of functional forms available that no one functional form is generally applicable. Therefore, it is the aim of this study to, (a) compare a range of the more common functional forms for their ability to estimate the volume of plantation grown *Eucalyptus globulus* Labill. subsp. *globulus* in south west Western Australia, and (b) examine the variation in accuracy and precision caused by estimating the parameters of the equations via ordinary and weighted least squares and, where logarithmic transformations were appropriate, the effect of applying Sprugal's (1983) correction factor.

2.2.2 METHODS

2.2.2.1 DATA

Data from trees felled on plots in *E. globulus* plantations in south west Western Australia, were used in this study. Four hundred and thirty single stem, defect free trees were felled in 60 plantations (Figure 1). From each plantation about 7 stems were subjectively chosen for sampling with an aim of spanning the volume range present in that particular plot. Sample discs were removed from each tree at 0.0 m, 0.5 m, 1.3 m and every metre thereafter along the stem, to an approximate 4 cm top diameter under bark (ub). The diameter (ub) of each disc was recorded as the average of the diameters (cm) of the long and short axis. The diameter under bark was used rather than the diameter over bark, which is the more usual practise. As the bark thickness was observed to vary with geographic locality, this methodology will remove this

Figure 1: Map of south west Western Australia showing the locations of plantations from which the data were drawn.



● location of plantations

Table 1: **Frequency of stems for each height and diameter class. Figures in parenthesis represent stems comprising the validation data set.**

	Height class (m)										
Diameter class (cm)	**0.0 - 3.4**	**3.5 - 6.9**	**7.0 - 10.4**	**10.5 - 13.9**	**14.0 - 17.4**	**17.5 - 20.9**	**21.0 - 24**	**24.5 - 27.9**	**28.0 - 31.4**	**31.5 - 34.9**	**35.0 - 38.5**
0.0 - 3.4	1	4									
3.5 - 6.9		19(6)	25(6)	1							
7.0-10.4		1	42(16)	32(6)	4(1)						
10.5 - 13.9			11(4)	25(15)	23(6)	5(1)	5(1)				
14.0 - 17.4			2	6(3)	21(8)	15(8)	12(5)	3(1)	(1)		
17.5 - 20.9				2	7(2)	9(4)	10(7)	4(1)	2		
21.0 - 24.4					1(1)	4(4)	3(7)	6	1(1)	1	(1)
24.5-27.9					1		2(2)	4	1		(1)
28.0 - 31.5										1	

unwanted source of variance. The total height of each tree was measured directly from the felled tree and recorded to the nearest 0.1 m. The frequencies of stems in each of the diameter and height classes are shown in Table 1.

Tree volumes were determined by summing the volumes of individual sections, determined via Smalian's formula (Clutter *et al.* 1983). Smalian's formula was chosen for its ease of application after ascertaining that the mean of the differences between stem volumes calculated by Smalian's and Huber's (Carron 1968) formulae was not significantly ($p>0.0001$) different from zero.

About 2 trees from each plantation ($n = 112$) were selected at random. These trees were assumed to represent the population and were withheld for equation validation purposes (Table 1).

2.2.2.2 CANDIDATE EQUATIONS

The functional forms of the equations used in this study are listed in Table 2. Where possible the terminology of Clutter *et al.* (1983) is adopted when referring to the functional forms of the candidate equations.

Variables common to the equations are as follows:-

β_0 - β_3	=	Regression parameters particular to specific equations.
V	=	Merchantable volume under bark (m^3) to a top end diameter limit of 4 cm (ub).
D	=	Diameter under bark at 1.3 m above the ground (cm) (dbhub).
H	=	Total height of the tree from base to tip (m).
ln	=	Natural logarithm.

Table 2: Functional forms of candidate equations

Equation	functional form	title	citation
1	$v=\beta_0 D^2 H$	constant form factor	(Green and Strawderman 1986)
2	$v=\beta_0+\beta_1 D^2 H$	combined variable	(Martin 1984)
3	$v=\beta_0+\beta_1 D^2+\beta_2 H+\beta_3 D^2 H$	generalized combined variable	(Postaire and M'Hirit 1985)
4	$v=D^2/(\beta_0+\beta_1 H^{-1})$	Horner's transformed variable	(Horner 1965)
5	$v=D^2H/10^{\beta_0-\beta_1}/(D+\beta_2)^2$	Opie's functional form	(Opie 1976)
6	$ln\,(v)=\beta_0+\beta_1 ln(D)+\beta_2 ln\,(H)$	logarithmic	(Higuchi and Ramm 1985)
7	$ln\,(v)=\beta_0+\beta_1 ln(D^2H)$	logarithmic combined variable	(Higuchi and Ramm 1985)
8	$ln\,(v)=\beta_0+\beta_1 ln(D)$	-	(Loetsch *et al* 1973)
9	$v=\beta_0+\beta_1 D^{\beta_2}$	-	(Avery and Burkhart 1983)
10	$v=\beta_0+\beta_1(D)+\beta_2 D^2$	-	(Loetsch *et al* 1973)

2.2.2.3 NUMERICAL TECHNIQUES

It is well established that the ordinary least squares assumption of homogeneous variance is inappropriate in the case of volume equation development (Schreuder and Anderson 1984; Gregoire and Dyer 1989). A number of approaches towards fitting models with heterogeneous errors recur in the literature. In the first the heterogeneity is defined and its effect nullified by the assignment of weights to the deviations when the equation parameters are estimated (Cunia 1964). Methodologies for the definition of variance heterogeneity are given by McClure *et al.* (1983), Green and Strawderman (1986), Meng and Tsai (1986) and McClure and Czaplewski (1987), while the application of these approaches are common (Gibson and Webb 1968; Knoebel *et al.* 1984; Hann *et al.* 1987)..

In the second, variance heterogeneity is ignored and coefficients are estimated via techniques that are resistant to the problem (Efron and Gong 1983; Schreuder and Anderson 1984; Wu 1986). In an examination of the two approaches, Gregoire and Dyer (1989) conclude that the former methodology offers nonnegligible gains in efficiency, whereas the robust alternatives provide accurate assessment of ordinary least squares parameters even in the presence of heteroscedasticity.

A third approach removes heteroscedasticity by logarithmic transformation of the data prior to the estimation of the parameters, usually by ordinary least squares. The approach is commonly used as an alternative to weighting (Newnham 1967; Higuchi and Ramm 1985; McTague and Bailey 1987). However, a systematic bias is introduced when the unbiased logarithmic estimates are converted back to standard units. In this situation the mean of the normally distributed *ln*(V) for a given *ln*(D) and/or *ln*(H) is replaced by the geometric mean upon conversion with an antilogarithm (Finney 1941) thus introducing a negative bias. The nature of the bias is discussed in detail by Baskerville (1972). To correct for this bias the final result is multiplied by a correction factor calculated from the standard error of the estimate of the regression (Whittaker and Woodwell 1968; Sprugel 1983). With some exceptions (Higuchi and Ramm 1985; McNab *et al.* 1985; Clark *et al.* 1986; Clark and Schroeder 1986) the application of correction factors to volume equations, involving logarithmic transformations, is absent from the literature concerned with single tree volume estimation.

In this study the volume equation parameters for Eqs. [1], [2], [3], [9] and [10] (Table 2) were estimated via ordinary least squares and again via weighted least squares. Statistics derived via weighted least squares are identified by the subscript (WT) with the equation identification number. The variance assumption for weighted least squares was $\sigma^2 \propto \sqrt{D^2H}$ and was found to be a reasonable one during the data exploration phase of this study.

Parameters for Eqs. [6], [7] and [8] were estimated via ordinary least squares after transformation of the variables with a natural logarithm gave homogeneous variance. Predictions of volume were obtained by transforming the variables back to original terms with the exponential function. This exercise was repeated while applying the correction factor (CF)recommended by Sprugel (1983). Hence for Eq. [6], the parameters were estimated when the equation was of the form,

$$ln(V) = \beta_0 + \beta_1\ ln(D) + \beta_2\ ln(H).$$

Volume was predicted when the equation was of the form,

$$V = e^{\beta 0}.D^{\beta 1}.\ H^{\beta 2}. \qquad \text{Eq. [6]}$$

where,

e = base of the natural logarithm

In the final step, volume was estimated when the equation was of the form,

$$V = e^{\beta 0}.\ D^{\beta 1}.\ H^{\beta 2}.\ CF \qquad \text{Eq. } [6]_{(CF)}$$

where,

$$CF = e^{((\sqrt{\Sigma\left(ln Vi - ln \hat{Vi}\right)^2/(n-3)})^2/2)}$$

after Sprugel (1983).

Equations where the correction factor have been applied are identified via the subscript (CF) with the equation identification number. Eqs. [4] and [5] have homogeneous variance and no weighting or logarithmic transformations were applied (Horner 1965; Opie 1976). Where the intercept term is suppressed the subscript (NI) will accompany the equation identification number.

2.2.2.4 VALIDATION CRITERIA

When validating candidate equations two approaches are common. The first uses statistical hypothesis testing to ascertain if the candidate equations are sufficiently accurate to warrant acceptance. The level of accuracy is prescribed. The methodologies for this approach are given by Freese (1960) and Reynolds (1984), both of which are demonstrated and discussed by

Gregoire and Reynolds (1988). In the second, statistical estimates are used for comparative purposes in the absence of any particular standard for accuracy (c.f. Cao *et al.* 1980; Green and Strawman 1986; McTague and Bailey 1987).

The strategy of statistical hypothesis testing is not commensurate with the objectives of this study. Therefore, for the purposes of validation, the volumes of the trees in the independent data set were predicted (V_p) and compared to the assumed actual volumes (V_0). The mean of the deviations ($\overline{D_V}$), the mean of the squared deviations ($\overline{D_V^2}$), the standard deviation of the deviations (D_{VSD}) and the sum of the differences of V_O and V_P (ΣD_V) were used to compare the accuracy and precision of the candidate equations. D_V and ΣD_V are assumed to represent the accuracy of the equations while $\overline{D_V^2}$ and D_{VSD} represent their precision.

In an attempt to identify any pattern occurring in the results, the equations were clustered on the basis of their validation statistics in the following manner. A matrix of associations between equations was calculated via the Gower metric (Gower 1971). The polythetic agglomerative strategy, the unweighted pair-group method using arithmetic averages (UPGMA) was used to impose structure to the association matrix (Gauch and Whittaker 1981). The clustering intensity coefficient *beta* was set at -0.1.

2.2.3 RESULTS

2.2.3.1 CONSTANT FORM FACTOR FUNCTIONAL FORM

The use of the constant form factor functional form assumes that the form of the trees, in the population of interest, is relatively constant regardless of tree size (Clutter *et al.* 1983). To test the validity of this functional form's application to the data set the cylindrical form factor of all trees, with a merchantable stem length greater than 4.3 m, were calculated. Trees with a stem length of 4.3 m and less were excluded from the analysis due to the increasing influence of butt swelling on the value for the cylindrical form factor. Also, trees with a stem length of 4.3 m or less are of little practical importance as they contain very small volumes. The trees remaining ($n = 277^*$) were divided into 10 volume classes of approximately equal numbers. No significant difference was detected ($p>0.3956$) between the means of the cylindrical form factors of each volume class (Figure 2). It is therefore reasonable to assume that the cylindrical form factor functional form is a legitimate candidate model.

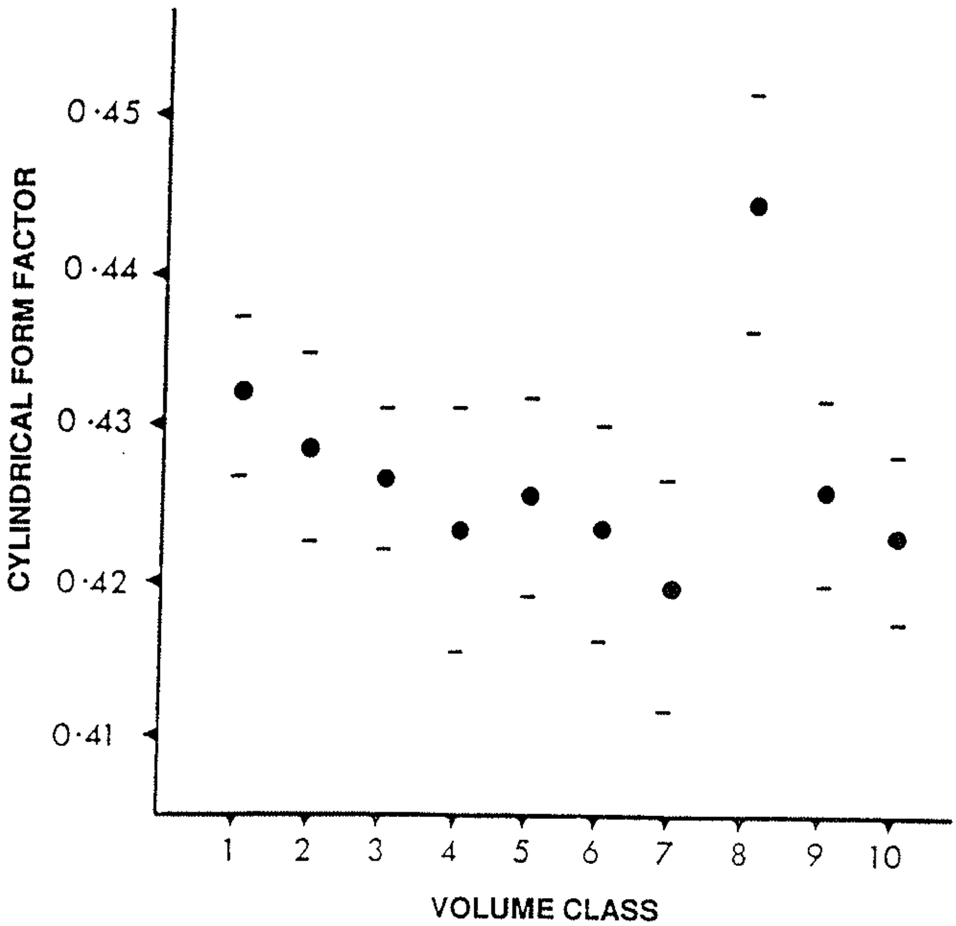
2.2.3.2 EQUATION PARAMETERS AND MODEL STATISTICS

Estimations of the parameters, the squared multiple correlation coefficient (r^2) and the residual mean squared for each equation are given in Table 3. With the exception of Horner's transformed variable functional form (Eq. [4]) all equations have large squared multiple correlation coefficients and small residual mean squares.

Equations containing the D term only (Eqs. [8],[9] and [10]) performed well in terms of r^2 and residual mean square. This would be expected given the relative uniformity of the cylindrical form factor across volume classes.

* does not include the validation data set.

Figure 2: **The mean (S.E.M.) cylindrical form factor for each volume class. ● = mean, - = S.E.M. ANOVA table refers to differences between means.**



Source	Degrees of freedom	Sums of squares	Mean square	F
Model	9	0.011	0.001	1.06
Error	268	0.309	0.001	

Table 3: **Estimations of the parameters, squared multiple correlation coefficient and residual mean squared for each candidate equation.**

Equation	β_0	β_1	β_2	β_3	r^2	RMS
$1_{(NI)}$	-	$3.280x10^{-5}$	-	-	0.9934	0.00036
$1_{(NI)(WT)}$	-	$3.274x10^{-5}$	-	-	0.9941	0.04189
2	0.00156	$3.268x10^{-5}$	-	-	0.9899	0.00036
$2_{(WT)}$	-0.00181	$3.325x10^{-5}$	-	-	0.9912	0.04643
3	-0.00658	$-6.234x10^{-5}$	0.00119	$3.374x10^{-5}$	0.9904	0.00034
$3_{(WT)}$	-0.00890	$-7.814x10^{-5}$	0.00153	$3.396x10^{-5}$	0.9888	0.03981
$3_{(NI)}$	-	$-6.925x10^{-5}$	$6.83x10^{-4}$	$3.440x10^{-5}$	0.9937	0.00034
$3_{(NI)(WT)}$	-	$-8.988x10^{-5}$	$9.90x10^{-4}$	3.468×10^{-5}	0.9945	0.03988
4	64.974	29451.0	-	-	0.8679	$2.55x10^{5}$
5	4.484	$5.522x10^{6}$	$7.591x10^{5}$	-	0.9934	$3.60x10^{-4}$
6	-10.379	1.844	1.175	-	0.9936	0.01433
7	-10.289	0.997	-	-	0.9931	0.01558
8	-9.625	2.767	-	-	0.9780	0.04958
9	-0.00189	$8.015x10^{-5}$	2.714	-	0.9934	0.11340
10	0.0700	-0.0183	0.00144	-	0.9586	0.00184
$10_{(WT)}$	0.119	-0.0243	0.00160	-	0.9606	0.20877

(WT) = parameters estimated via weighted least squares; weighting factor = $\sqrt{D^2H}$
(NI) = no intercept term used.

2.2.3.3 EQUATION VALIDATION

As expected, the equations containing the D term only validated with less accuracy and precision than did equations containing both the D and H term. Of the diameter equations, Eq.$[10]_{(WT)}$ validated with the smallest mean bias while, with the exception of Eq.[8], the precision of these equations was similar (Table 4).

Classification of the equations on the basis of their validation statistics yielded two distinctive groups, namely those equations with both the D and H terms and those with the D term only (Figure 3a). The differences between the two groups were large and masked the subtle groupings within the classification. The classification was therefore repeated with the equations containing the D term only excluded. The second classification was arbitrarily truncated at the four group level (Figure 3b). Group 1 contained Eqs. [1], $[1]_{(WT)}$, $[2]_{(WT)}$ and [5] and is comparably inaccurate although relatively precise. Group 2 contains Eqs. [2],$[3]_{(WT)}$,$[3]_{(WT)}$ $[3]_{(NT)}$, [6], [7] and $[7]_{(CF)}$. This group is second in accuracy to group 3 while being slightly more precise than other groups. Group 3 contains Eqs. $[3]_{(NT)(WT)}$ and $[6]_{(CF)}$ and is the group of highest accuracy. Precision is slightly less than that of the equations in group 2. Group 4 contained Eq [4] only and is less accurate and precise than other groups. Univariate validation statistics for each of the groups is given in Table 5.

Although parameter estimation via weighted least squares is statistically advantageous, its benefits in terms of improvements to validation statistics was varied. For example, Eqs. [1] and [2] gave deviations that were significantly different ($p<0.0001$) than the deviations from the same equations derived via weighted least squares. In this case the deviations were increased after application of weighted least squares. Alternatively, for Eqs.$[3]_{(NT)}$ and [10] the deviations were significantly ($p<0.0007$) decreased by weighted least squares. No significant difference ($p>0.0821$) was detected between the deviations from Eq. [3] for weighted or ordinary least squares. The influence on precision is less evident. With the exception of Eq. [1] where a difference was detected ($p<0.0306$), no significant difference ($p>0.2719$) was found between the squared deviations of the same equations, derived via ordinary or weighted least squares.

Application of Sprugel's (1983) correction to Eqs. [6], [7] and [8] decreased the deviations significantly ($p<0.0001$) in all cases. Although an increase in precision resulted from applying the correction it was not considered to be significant ($p>0.0974$).

Table 4: Validation statistics for candidate equations. $\overline{Dv}$ = mean deviation, $\overline{Dv^2}$ = the mean of the squared deviations, Dv_{SD} = standard deviation of the deviations, ΣDv = sum of the total deviations.

Equations	$\overline{Dv}$	$\overline{D_v^2}$	D_{vSD}	ΣD_v
$1_{(NI)}$	0.00377	2.23×10^{-4}	0.0145	0.4222
$1_{(NI)(WT)}$	0.00403	2.28×10^{-4}	0.0146	0.4521
2	0.00278	2.23×10^{-4}	0.0147	0.3117
$2_{(WT)}$	0.00364	2.09×10^{-4}	0.0140	0.4083
3	0.00238	2.05×10^{-4}	0.0142	0.2795
$3_{(WT)}$	0.00221	2.07×10^{-4}	0.0143	0.2590
$3_{(NI)}$	0.00218	2.13×10^{-4}	0.0145	0.2443
$3_{(NI)(WT)}$	0.00101	2.17×10^{-4}	0.0147	0.1127
4	0.00548	3.12×10^{-4}	0.0168	0.6138
5	0.00376	2.23×10^{-4}	0.0145	0.4211
6	0.00221	2.03×10^{-4}	0.0141	0.2188
$6_{(CF)}$	0.00117	2.01×10^{-4}	0.0142	0.1017
7	0.00298	2.14×10^{-4}	0.0144	0.3308
$7_{(CF)}$	0.00186	2.01×10^{-4}	0.0141	0.2043
8	0.00966	2.30×10^{-3}	0.0473	1.0696
$8_{(CF)}$	0.00621	2.10×10^{-3}	0.0457	0.6823
9	0.00585	2.15×10^{-3}	0.0462	0.6576
10	0.00593	2.11×10^{-3}	0.0458	0.7795
$10_{(WT)}$	0.00279	2.14×10^{-3}	0.0464	0.4531

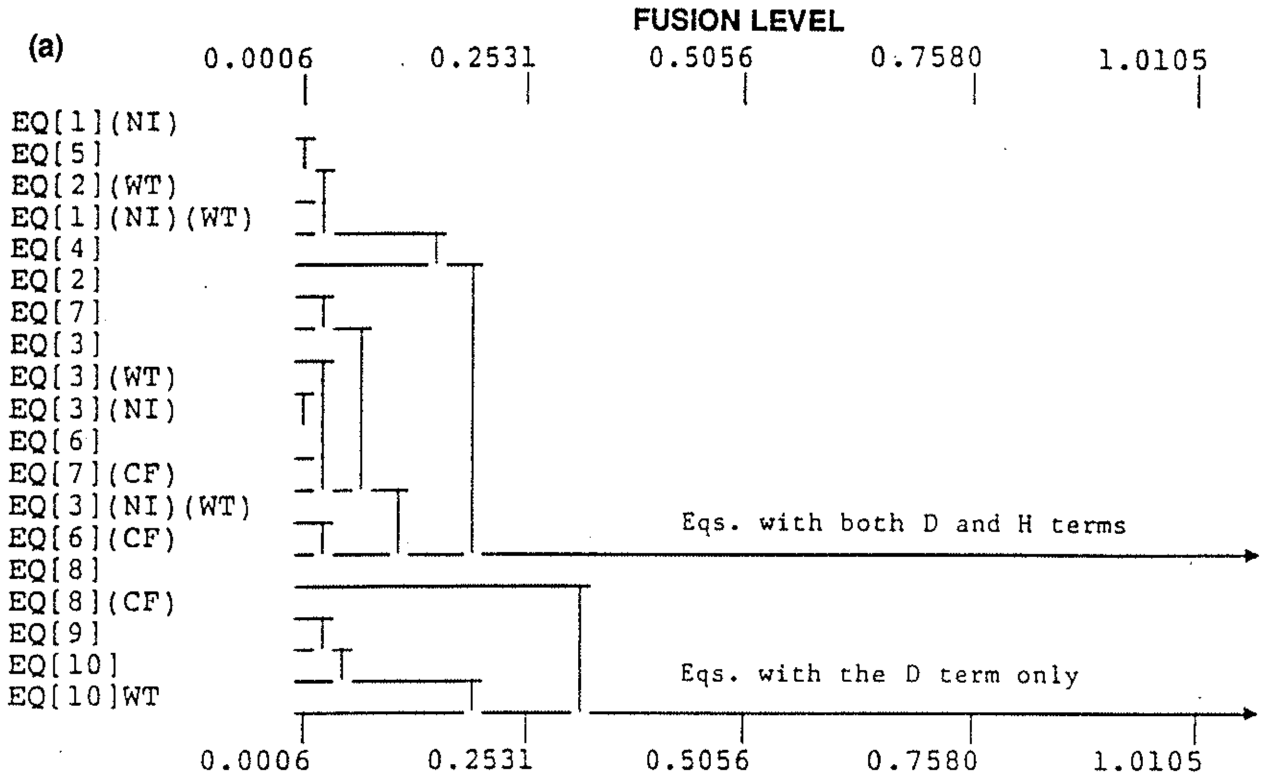
(NI) = no intercept term used.
(WT) = parameters estimated via weighted least squares; weighting factor $\sqrt{D^2H}$.
(CF) = Sprugel's (1983) correction applied.

Figure 3: Dendrograms resulting from the classification of candidate equations based on their validation statistics (metric = Gower; fusion strategy = UPGMA, beta = 0.01).

(a) Dendrogram resulting from the classification of all equations.

(b) Dendrogram resulting from the classification of equations containing the D and H terms.

(a)



(b)

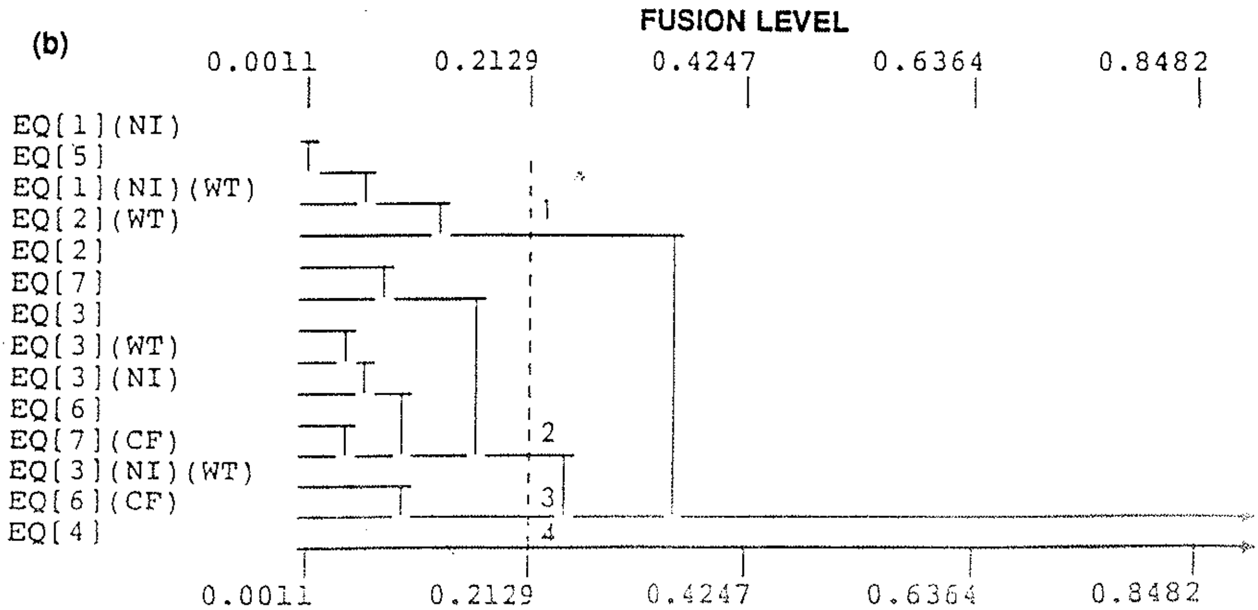


Table 5: Univariate validation statistics for each of the four cluster groups.

Cluster	$\overline{D_v}$			$\overline{D_v^2}$			D_{VSD}			ΣD_V		
	$\bar{x}$	C.V.	N	$\bar{x}$	C.V.	N	$\bar{x}$	C.V.	N	$\bar{x}$	C.V.	N
1	0.0038	4.3	4	2.21×10^{-4}	3.7	4	0.0144	1.9	4	0.4259	4.3	4
2	0.0024	16.2	7	3.09×10^{-4}	3.6	7	0.0143	1.5	7	0.2640	17.6	7
3	0.0011	10.4	2	2.09×10^{-4}	5.4	2	0.0146	2.4	2	0.1072	7.2	2
4	0.0055	-	1	3.12×10^{-4}	-	1	0.0168	-	1	0.6138	-	1

Cluster 1 = Eqs. [1], $[1]_{(WT)}$, $[2]_{(WT)}$, [5]
2 = Eqs. [2], [3], $[3]_{(WT)}$, $[3]_{(NI)}$, [6], [7], $[7]_{(CF)}$
3 = Eqs. $[3]_{(NI)(WT)}$, $[6]_{(CF)}$
4 = Eq. [4]
C.V. = coefficient of variation (%)

2.2.4 DISCUSSION AND CONCLUSION

The relative homogeneity of form for plantation grown *E. globulus* in south west Western Australia is somewhat surprising given the wide range of environmental conditions spanned by the study area (seeChapter 4). Given this uniformity it is not surprising that many of the functional forms fit the data well. The accuracy and precision of most functional forms are within tolerable limits.

Under such circumstances it is reasonable not to examine those functional forms which include a measure of form in the equations (Smith 1976; Roebbelen and Smith 1984). Given the accuracy and precision achieved with the equations examined, any extra effort employed to gather measures of tree form is unnecessary. Likewise, there is no need to stratify the sample on the basis of geometrical shape prior to estimation of the equations' parameters (Postaire and M'Hirit 1985).

The use of weighted least squares to estimate the parameters of the equations, when compared to ordinary least squares, produced varied influences on validation statistics. In the presence of heterogeneity, the use of this technique is critical for constructing confidence limits around the estimates. Therefore, in the absence of evidence that the procedure consistently improves or detracts from the accuracy and precision of the outcome, weighted least squares is the recommended procedure.

The application of Sprugel's (1983) correction factor improved the predictive capabilities of the equations to which it was applied in all cases. Therefore, its application is recommended as standard practice where logarithmic transformations are used as an alternative to using weighted least squares.

It is concluded from this study that either of the following two equations are the most suitable for volume estimation for plantation grown *E. globulus* in south west Western Australia.

$Eq[3]_{(NI)(WT)}$

$$V=-8.9880x10^{-5}.D^2+9.90x10^{-4}.H+3.4687x10^{-5}.D^2.H$$

or

$Eq[6]_{(CF)}$

$$V=(e^{-10.3796}.\ D^{1.8442}.\ H^{1.1746}).\ 1.00719.$$

The relationships between volume and dbhub and height, represented by these equations are given in Figure 4 and Figure 5.

When a height term is unavailable the following equation is recommended.

$Eq[10]_{(WT)}$ $$V=0.11932 - 0.002433.D+0.00160.D^2$$

Figure 4: The relationship between stem volume, diameter at breast height underbark and tree height, represented by Eq. [3] (NI) (WT).

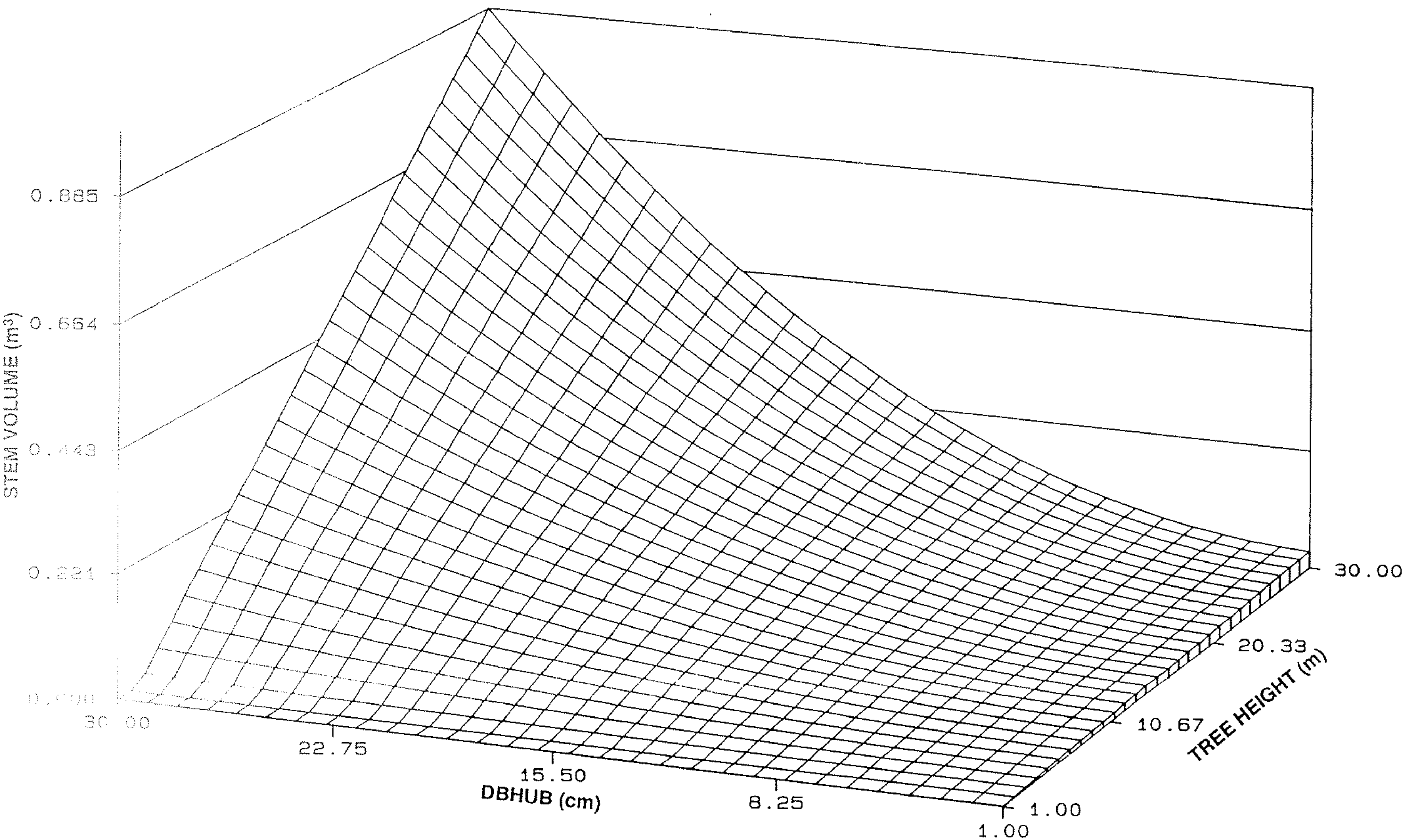
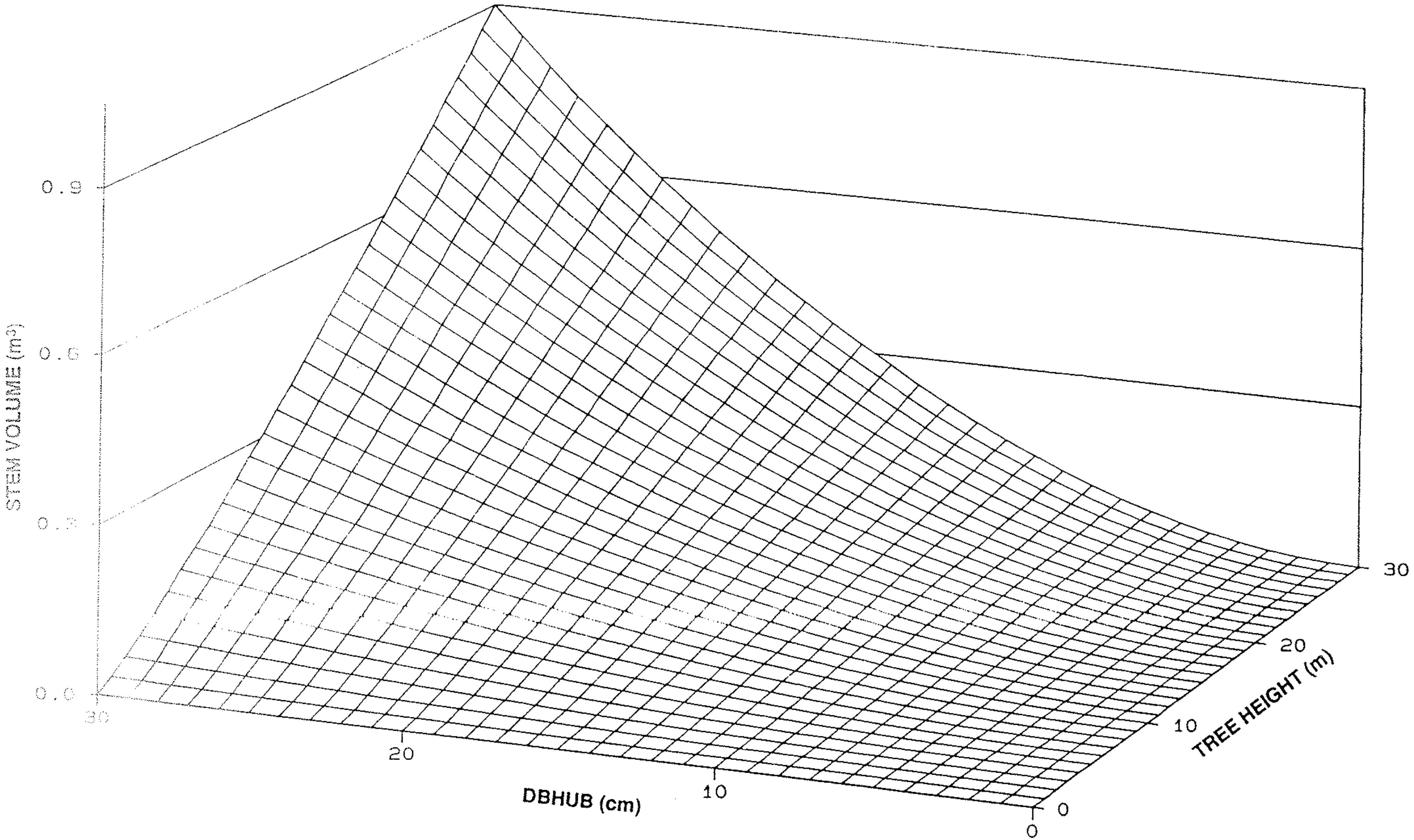


Figure 5: The relationship between stem volume, diameter at breast height underbark and tree height, represented by Eq. [6](CF).



2.2.5 REFERENCES

Avery, T.E., and H.F. Burkhart. 1983. Forest measurement. McGraw-Hill, New York. 331 p.

Baskerville, G.L. 1972. Use of logarithmic regression in the estimation of plant biomass. *Can. J. For. Res.*, 2:49-53.

Bennett, F.A., C.E. McGee, and J.L. Clutter. 1959. Yield of old-field slash pine plantations. *USDA Res. Pap.* SE-107.

Biging, G.S. 1984. Taper equations for second-growth mixed conifers of Northern California. *For. Sci.*, 30:1103-1117.

Biging, G.S. 1988. Estimating the accuracy of volume equations using taper equations of stem profile. *Can. J. For. Res.*, 18:1002-1007.

Brackett, M. 1973. Notes of tarif tree volume compution. *State of Washington, Dept. Nat. Res., Resource Mgt. Report* No 24.

Brickell, J.E. 1985. Stem analysis: A conventional approach to volume determination. *In.* Proceedings - growth and yield and other mensurational tricks: A regional technical conference.*USDA Gen. Tech. Rep.* INT-193.

Burk, T.E., and A.R. Ek. 1982. Application of empirical Bayes/James-Stein procedures to simultaneous estimation problems in forestry. *For.Sci.*, 28:753-771.

Cao, Q.V., H.E. Burkhart, and T.A. Max. 1980. Evaluation of two methods for cubic-volume prediction of loblolly pine to any merchantable limit. *For. Sci.*, 26:71-80.

Carron, L.T., and L. McIntyre. 1959. The taperline (Gray) and sectional methods for merchantable volume of *Pinus radiata* stems. *Aust.For.*, 23: 50-60.

Carron, L.T. 1968. An outline of forest mensuration with special reference to Australia. Australian National University Press, Canberra.

Clark, A., D.R. Phillips, and D.J. Frederick. 1986. Weight, volume, and physical properties of major hardwood species in the Piedmont. *USDA Res. Pap.* SE-255.

Clark. A., and J.G. Schroeder. 1986. Weight, volume, and physical properties of major hardwood species in the Southern Appalachian Mountains. *USDA Res. Pap.* SE-253.

Clutter, J.L., J.C. Fortson, L.V. Pienaar, G.H. Brister, and R.L. Bailey. 1983. Timber management: a quantitative approach. John Wiley and sons. New York 333p.

Cunia, T. 1964. Weighted least squares method and construction of volume tables. *For. Sci.*, 10: 180-191.

Efron, B., and G. Gong. 1983. A leisurely look at the jackknife, the bootstrap, and cross validation. *Am. Statist.*, 37:36-48.

Ernst, S., and D.W. Hann. 1984. Volume and value prediction for young-growth true fir trees. *For. Sci.*, 30:871-882.

Farrar, R.M. 1985. Crown ratio used as a surrogate for form in a volume equation for natural longleaf pine stems. *In*: Proceeding of the 3rd Biennial Southern Silviculture Research Conference. *USDA Gen. Tech. Rep.* S0-54.

Finney, D.L. 1941. On the distribution of a variate whose logarithm is normally distributed. *J. Royal. Stat. Sci., Series B*. 7:155-161.

Forslund, R.R. 1982. A geometrical tree volume model based on the location of the centre of gravity of the bole. *Can. J. For. Res.*, 14:768-774.

Freese, F. 1960. Testing accuracy. *For. Sci.*, 6:139-145.

Furnival, G.M. 1961. An index for comparing equations used in constructing volume tables. *For. Sci.*, 7:337-341.

Gauch, H.G., and R.H. Whittaker. 1981. Hierarchial classification of community data. *J. Ecol.*, 69:537-557.

Gevorkiantz, S.R., and L.P. Olsen. 1955. Composite volume tables for timber and their application in the Lake States. *USDA Tech. Bull.* 1104.

Gibson, D.F., and A.W. Webb. 1968. Regression methods and models for estimating tree volumes of apline ash regrowth. *Aust. For.*, 32:233-242.

Gower, J.C. 1971. A general coefficient of similarity and some of its properties. *Biometrics*, 27:857-871.

Goulding, C.J. 1979. Cubic spline curves and calculation of volume of sectionally measured trees. *N.Z. J. For. Sci.*, 9:89-99.

Green, E.J. 1983. Evaluating the predictive abilities of regressions with PRESS. *For. Sci.*, 29:712-714.

Green, E.J., and W.E. Strawderman. 1986. Stein-rule estimation of coefficients for 18 eastern hardwood cubic volume equations. *Can. J. For. Res.*, 16:249-255.

Gregoire, T.G., and M.E. Dyer. 1989. Model fitting under patterned heterogeneity of variance. *For. Sci.*, 35:105-125.

Gregoire, T.G., and M.R. Reynolds. 1988. Accuracy testing and estimation alternatives. *For. Sci.*, 34:302-320.

Hann, D.W., D.K. Walters, and J.A. Scrivani. 1987. Incorporating crown ratio into prediction equations for Douglas-fir stem volume. *Can. J. For. Res.*, 17:17-22.

Higuchi, N., and W. Ramm. 1985. Developing bole wood volume equations for a group of tree species of central Amazon (Brazil). *Commonw. For. Rev.*, 64:33-41.

Horner, T.G. 1965. A new total cubic foot volume function *For. Chron.*, 45:278-283.

Husch, B., C.I. Miller, and T.W. Beers. 1972. Forest mensuration. Ronald Press. New York. 410 pp.

Jacobs, M.W., and D.B. Monteith. 1981. Feasibility of developing regional weight tables. *J. For.*, 79:676-677.

Knoebel, B.R., H.E. Burkhart, and D.E. Beck. 1984. Stem volume and taper functions for yellow-poplar in the Southern Appalachians. *S.J. Appl. For.*, 8:185-188.

Kozak, A., D.D. Munro, and J.H.G. Smith. 1969. Taper functions and their application in forest inventory *For. Chron.*, 45:278-283.

Laasasenaho, J. 1982. Taper curve and volume functions for pine spruce and birch. *Commun. Inst. For. Fenn.*, 108.

Liu, C.J. 1980. Log volume estimation with spline approximation. *For. Sci.*, 26:361-369.

Loetsch, F., F. Zohrer, and K.E. Haller. 1973. Forest inventory. Vol. II. BLV Verlagsgesellschaft, Munich.

Lohrey, R.E. 1983. Stem volume prediction and crown characteristics of longleaf pine plantations. *In*: Proceeding of the 3rd Biennial Southern Silviculture Reserach Conference. *USDA Gen. Tech. Rep.* SO-54.

Lynch, T.B. 1988. Volume estimation from sample tree counts with certain individual tree volume equations. *For. Sci.*, 34:677-697.

Martin, J.A. 1984. Testing volume equation accuracy with water displacement techniques. *For. Sci.*, 30:41-50.

MacDonald, G.B., and R.R. Forslund. 1986. Application of a geometrical volume equation to species with different bole forms. *Can. J. For. Res.*, 16:311-314.

McClure, J.P., and R.L. Czaplewski. 1987. High order regression models for regional volume equations for loblolly pine and white oak. *USDA Res. Pap.* SE-240.

McClure, J.P., H.T. Schreuder, and R.L. Wilson. 1983. A comparison of several volume equations for loblolly pine and white oak. *USDA Res. Pap.* SE-240.

McNab, W.H., K.W. Outcalt, and R.H. Brendemuehl. 1985. Weight and volume of plantation grown Choctawhatchee sand pine. *USDA Res. Pap.* SE-252.

McTague, J.P., and R.L. Bailey. 1987. Simultaneous total and merchantable volume equations and compatible taper functions for loblolly pine. *Can. J. For. Res.*, 17:87-92.

Meng, C.H., and W.Y. Tsai. 1986. Selection of weights for weighted regression of tree volume. *Can. J. For. Res.*, 16:671-673.

Munro, D.D. 1974. Forest growth models - a prognosis. *In*: Growth models for tree and stand simulation. J.J. Fries (Ed.) *Royal College of Forestry, Res.* Note No. 30.

Newnham, R.M. 1967. A modification to the combined-variable formula for computing tree volume. *J. For.*, 65:719-721.

Opie, J.E. 1976. Volume functions for trees of all sizes. Forests Commission, Victoria, Australia. *For. Tech. Pap.*, 25:27-30.

Postaire, J., and O. M'Hirit. 1985. Application of pattern recognition to volume estimation in forest inventory. *For. Sci.*, 31:53-65.

Reynolds, M.R. 1984. Estimating the error in model predictions. *For. Sci.*, 30:454-469.

Roebbelen, P., and V.G. Smith. 1984. An investigation of product form. *Can. J. For. Res.*, 14:768-774.

Romancier, R.M. 1961. Weight and volume of plantation-grown loblolly pine. *USDA Res. Pap.* No. 161.

Sadiq, R.A., and V.G. Smith. 1983. Estimation of individual tree volume with age and diameter. *Can. J. For. Res.*, 13:32-39.

Schreuder, H.T., and J. Anderson. 1984. Variance estimation for volume when D^2H is the covariable in regression. *Can. J. For. Res.*, 14:818-821.

Smith, V.G. 1976. The use of product form in form-class volume equations. *Can. J. For. Res.*, 6:93-103.

Sprugel, D.G. 1983. Correcting for bias in log-transformed allometric equations. *Ecology*, 64:209-210.

West, P.W. 1980. Use of diameter increment and basal area increment in tree growth studies. *Can. J. For. Res.*, 10:71-77.

Whittaker, R.H., and G.M. Woodwell. 1968. Dimension and production relations of trees and shrubs in the Brookhaven forest, New York. *J. Ecol.*, 56:1-25.

Wu, C.F.J. 1986. Jackknifing, bootstrapping and other resampling methods in regression analysis. *Am. Statist.*, 14:1261-1295.

Young, H.E., W.C. Robbins, and S. Wilson. 1967. Errors in volume determination of primary forestry products. *In*: Proceedings of 14th IUFRO Congress, Munich, Sept. 4-9, 1967. Part VI, Sect.25. pp. 546-562.

CHAPTER THREE

Top Height Development and Site Index Equations for *E. globulus* in South West Western Australia

3.1 LITERATURE REVIEW: LAND EVALUATION FOR FORESTRY PURPOSES

3.1.1 INTRODUCTION

The copious literature relevant to land classification and evaluation is categorised into two conceptual schools. The first is concerned with the identification of land units which are considered homogeneous for the phytological, physiographical or biophysiographical attributes used in their definition. The second classifies land units according to their productive capacity. The two concepts have confused terminology throughout the literature so, for consistency, the terminology of Kilian (1984) will be adopted. *Site classification* defines land units with similar combinations of environmental features of similar ecological effect. *Site evaluation* classifies the site according to its productive capacity based on the ability of the area to produce a certain yield of one or several tree species.

Examples of the site classification approach are given in Smalley (1984), Inions (1990) and Inions *et al.* (1990). Site classification has little relevance to this study, consequently the literature will not be reviewed. Adequate reviews of the topic are given by Daubenmire (1976) and Havel (1980 a,b).

Site evaluation involves the allocation of a productivity measure or class to the site of interest. The productivity measure may be derived by measurement of biomass production (Yarie and Van Cleve 1983) or timber volume (Lewis *et al.* 1976; Mader 1976). Alternatively, indices for productivity, such as the height of the dominant trees of a stand or the height growth over a nominated time period are used (Havel 1968; Economou 1990). By far the most common measure of site productivity is *site index*.

It is the aim of this review to summarise the literature pertaining to site index. No attention will be paid to that literature dealing with other indices for productivity such as height intercepts.

3.1.2 DEFINITION OF SITE INDEX

Site index is defined as the height of the dominant and/or codominant stratum of a stand at a nominal reference age, also referred to as index age (Beck and Trousdell 1973), base age (Bailey and Clutter 1974) or, less commonly, standard age (Carron 1968). The reference age has been recommended to be two-thirds of the rotation length (Carron 1968) or as close to the rotation length as possible (Clutter *et al.* 1983) however, the choice is often arbitrary.

The utility of site index stems from the relationship which exists between site index and the pattern of top height development. By definition site index gives the top height of a stand at one point in time. This implies that the pattern of top height development of the stand will follow a predetermined pattern unique to that site index. As Clutter *et al.* (1983) point out, the fact that the stand of interest has a height of X metres at the reference age is relatively insignificant in comparison. A second utility results from the correlation which exists between site index and stand volume, especially when age and density are accounted for. Finally, it is far easier and less expensive to estimate the site index of a stand than to directly measure standing volume. In a comparison of some site quality indices site index was as well or better correlated to site productivity than some of the more difficult to measure productivity indices (McLeod and Running 1988; Reed and Jones 1989).

3.1.3 DEFINITION OF TOP HEIGHT

Definition of top height, also referred to as dominant or predominant height, is not consistently made in the literature. One attempt, (Rennolds 1978) classifies top height definitions into:

i) Top height as a mean height of a fixed percentage of occurring trees;

ii) Top height as a mean height of a fixed number of largest or tallest trees per unit area, i.e.,

   a) as mean height of a specific number of the largest stems per unit area.

   b) as mean height of a specific number of the tallest stems per unit area.

iii) Top height as maximum height.

Loetsch *et al.* (1973) list a similar classification but include a category where size is determined by crown class.

The definition of top height is inconsistent between regions and subject species. For example, top height in coniferous plantations in Australia is the average of the heights of the 20 to 30 tallest trees per 0.4 ha (50 to 75 trees ha^{-1}). For even-aged regrowth eucalypt forest in Tasmania the mean height of the 12 tallest dominants per 0.4 ha (30 trees ha^{-1}) is used (Cromer and Bowling 1961; West 1982). In Great Britain, top height is defined as the mean of the heights of 40 trees of largest diameter per 0.4 ha (100 trees ha^{-1}) (Johnston and Bradley 1963). This figure corresponds with the IUFRO's recommendation for 100 trees ha^{-1} which is already in use in parts of Europe. In North America the averages of the heights of dominants, or dominants and codominants, is commonly used (Spurr 1952; Husch 1963). For example, Curtis *et al.* (1974b) used the height of one dominant tree on a 0.1 ha plot (10 trees ha^{-1}), while Barrett (1978) defines top height as the height of the tallest tree on a 0.08 ha plot (12.5 trees ha^{-1}). Monserud (1984) defines top height as the heights of the three "best growing" (based on increment cores) dominants on an approximately 0.023 ha plot that was representative of the growing conditions in the stand.

This diversity of definition evident in the literature is of some concern and, as demonstrated by Rennolds (1978), can lead to logical and statistical problems. The problem of inconsistent of definition remains a weakness in the site index literature.

3.1.4 METHODS FOR CONSTRUCTING TOP HEIGHT DEVELOPMENT AND SITE INDEX CURVES.

To date, no adequate system has appeared in the literature which classifies the copious methodologies for deriving top height development and site index systems. One exception is Clutter *et al.* (1983) who classify equations into three types according to the nature of the top height development curves they generate. Namely, (a) anamorphic curves which by definition are proportional, (b) polymorphic disjoint curves which are not proportional, but the curves do not intersect within the age range of interest and, (c) polymorphic nondisjoint curves where there is no constant proportionality and at least some of the curves intersect within the age range of interest. Clutter *et al.* (1983) also classify methodologies for deriving top height development and site index curves as:

i) the guide curve method;

ii) the difference equation method; and

iii) the parameter prediction method.

This classification was used by Grey (1989) in his review of the site index literature, although Smith and Watts (1987) preferred to derive their own classification, grouping the methodologies under:

i) two equation systems;

ii) one equation systems; and

iii) the structural method.

No classification clearly categorises the methodologies and in many cases it is mere semantics as to which category a method is allocated. Consequently, in this review, methods for deriving top height development and site index curves will be discussed under five headings. Three of which are defined by Clutter *et al.* (1983), a fourth termed the Ek-Payandeh method and a fifth for miscellaneous methods which have occasionally appeared in the literature, and are sufficiently different from mainstream methods to warrant mention.

3.1.4.1 THE GUIDE CURVE METHOD

The guide curve method will produce a set of anamorphic top height development curves. Data are usually derived from temporary plots such that a single tree or plot has only one top height/age measurement associated with it. The method is equally applicable to data representing a time series derived via remeasurement of permanent plots or from stem analysis (Curtis 1964).

Seminal works on guide curve construction were graphical, with two procedures common. The first involved drafting two curves based on the upper and lower limits of the data. A family of curves intermediate to these, whose shape was determined from the two limiting curves, was drawn. The method is referred to as *Baur's method*, the *limiting curve method* or the *strip method* (Grey 1945). The second involved plotting top height over a range of sites of various ages and fitting a master curve to these values. A family of curves are located, which are proportional to the master curve and in the same relative position as they are at the reference age. This technique is termed the *guide curve* or *direct curve method* or the *harmonised* or *anamorphic site index technique* (Bruce 1926).

In the more recent literature top height development curves are graphs of functions derived via numerical methods, predominantly regression analysis (Strand 1964). The most common functional form for the guide curve method is that proposed by Schumacher (1939), i.e.,

$$ln(\mathrm{H}) = \beta_o + \beta_1 \mathrm{A}^{-1}$$

where,

H = top height of the stand;

A = age of the stand or tree:

β_0 = the intercept constant which varies for each curve;

β_1 = a constant with the same value for each curve:

ln = natural logarithm.

Brickell (1968) developed top height development curves for Douglas-fir (*Pseudotsuga menziesii* var. *glauca* (Mirb) Franco) using static point data and the Schumacher functional form. Brickell then elaborated on the method by using skewness and kurtosis as well as standard deviations to describe the fractiles of the distributions of heights within age classes when deriving a family of curves.

In a different approach Newberry and Pienaar (1978) use the more complicated Chapman-Richards functional form (Chapman 1961; Richards 1959) i.e.,

$$H = \beta_0\left(1 - e^{-\beta_1 A}\right)^{(1-\beta_2)^{-1}}$$

where,

H, A are previously defined;

β_0, β_1, β_2 = parameters to be estimated;

e = base of the natural logarithm.

When the reference age was substituted into the equation the site index was obtained for the guide curve. Curves for other site index values were obtained from the guide curve equation by holding the shape parameters β_1 and β_2 constant and varying the asymptote parameter β_o as necessary to achieve the required H value when A equals the reference age. The result is a set of anamorphic top height development curves. Other examples of the guide method are found in Bennett *et al.* (1959), Heger (1968) and Smith and Watts (1987).

Two assumptions are inherent in the guide curve methodology. The first, that the guide curve represents the average top height growth on the average site, is false if site index is not constant for all ages. A number of authors have warned of the bias generated by violating this assumption (Heger 1968; Carmean 1972; Beck and Trousdell 1973) while Monserud (1985) and Smith (1984) examine its magnitude. In Monserud's (1985) data, observed site index declined from a height of 20 m at age 50 yrs to a height of 17 m at age 200 yrs, clearly violating the assumption that age and site index are unrelated. This decline resulted in a bias of more than 12 m at age 200 yrs due to the application of the guide curve method. A method to correct such bias is given by Curtis (1964).

The second assumption, that a given site index maintains a constant position in the distribution of heights over time (anamorphism) is also false for most cases. Curve shape is known to vary with site index (Monserud 1985), habitat type (Monserud 1984) and soil group (Carmean and Lenthall 1989). However, the bias introduced by the assumption of anamorphism when polymorphism is applicable is small in comparison to that introduced by applying guide curve methodology when site index varies with stand age (Monserud 1985; Monserud 1988).

Most data used to derive site index systems, which have appeared in the recent literature, result from stem analysis or repeated measurement of permanent plots. These data types gave rise to methodologies which overcame the problems associated with the guide curve method.

3.1.4.2 THE PARAMETER PREDICTION METHOD

The parameter prediction method requires age series data and results in polymorphic disjoint top height development curves. Basically the method involves:

i) fitting a linear or nonlinear top height by age function to each tree or plot in the data set;

ii) using each fitted curve to assign a site index value to each plot; and

iii) expressing the parameter estimates from (i) as functions of stand characteristics, commonly age (King 1966; Barrett 1978; Herman 1978) or site index (Brickell 1968; Trousdell *et al.* 1974).

In a variation from this procedure Smith and Watts (1987) fitted the Chapman-Richards functional form to each plot in their data set. They then expressed the β_2 parameter estimates as a function of site index and the β_0 and β_1 parameters as functions of the β_2 parameter.

Heger (1968) observed that the relationship between top height and site index is basically linear, given that age is held constant, and regressed top height on site index for each age class. These equations were then used to estimate the top heights of plots in each site index class at respective ages. This procedure yields polymorphic top height development curves. Smith (1984), using Clutter's *et al.* (1983) classification of methodologies, described Heger's (1968) method as a parameter prediction method. As no parameters of the final equation are estimated from extrinsic functions, this description is not correct. However, Heger's method often forms the basis from which many top height development curves are developed via the parameter prediction method (Herman *et al.* 1978; Monserud 1984; Alemdag 1988).

An inverse relationship should exist between an equation which predicts top height from site index and age and one which predicts site index from top height and age. However, ordinary least squares regression will not meet this criterion, a point clearly made by Curtis *et al.* (1974 a). As a result, Curtis *et al.* (1974 b) modified Heger's (1968) method into a true parameter prediction method, producing a function to estimate site index and a second function to estimate top height. For example, Curtis' *et al.* (1974a) site index equation could be expressed as,

$$S = \beta_0 + \beta_1 \overline{H}$$

where,

S = estimated site index

$\overline{H}$ = the mean top height from the individual regressions of site index and top height for each age class (after Heger (1968)).

The estimates for β_o and β_1 were then expressed as functions of age.

Separate site index and top height development curves as proposed by Curtis *et al.* (1974a) should (a) coincide at the index age only, and (b) diverge in a constant manner, for all ages (other than index age), as site index diverges from the mean value. However, on application of the Curtis *et al.* (1974a) technique by Curtis *et al.* (1974b), the resulting curves only approximated these conditions.

Dahms (1975) developed a method that derives curves which meet the criteria expected by Curtis *et al.* (1974a). The method begins, as does Heger's, by fitting top height and site index as linear functions of each other while age is held constant. Age is returned to the system by fitting the mean top height, mean site index and the estimates of the parameters of each age class, all as functions of age. This method has been applied successfully (Barrett 1978), although Monserud (1984) found site index curves developed via this method were far better behaved than the top height development curves derived in the same manner. Consequently, Monserud used the logistic functional form (Oliver 1966) to express top height development and went on to build habitat type into both the site index and top height development functions.

Most parameter prediction methodologies follow the procedures already discussed (Heger 1968; Curtis *et al.* 1974b; Dahms 1975; Monserud 1984), some with slight modifications (c.f. Farr 1984; Alemdag 1988). A noteworthy modification is presented by Biging (1985) who argued that the full information available in time series data sets, such as top height development, is not exploited. In Biging's study a modified version of the Chapman-Richards functional form was fitted to a pooled data set. The same functional form was also fitted to each tree in the data set. Parameters of the functions were estimated via ordinary least squares and generalized least squares (see Ferguson and Leech (1978)) with significant differences found between the two sets of estimators. In the example presented, using dominant trees in mixed conifer forest in California, the varying-parameter method proposed by Biging predicted higher asymptotic growth than the pooled data and ordinary least squares methods.

Lappi and Bailey (1988) extend Biging's (1985) argument and propose a statistical model that explicitly described the major random components in the variation in top height development curves. The average height of dominant and codominant trees in the population was expressed as a function of age. Then a model was developed for the variance-covariance structure of deviations from this average height curve due to stands and trees within stands.

The parameter prediction method presents a number of areas where concern is warranted. Firstly, as the parameters for top height development are often expressed as functions of site index, different choices of reference age will give different patterns of development, even if the same data set is used (Heger 1973). The second problem occurs where the predicted top height at the reference age does not equal site index and therefore must be proportionally adjusted (Burkhart and Tennent 1977). Thirdly, it is often difficult or impossible to solve explicitly for site index given top height and age. In this situation site index can only be determined by graphical interpolation or iterative computations (Clutter *et al.* 1983). Finally, where polymorphism is very evident, application of this method requires a flexible, usually nonlinear, functional form to be applied to each tree or plot in the data set. Where parameters are estimated via nonlinear procedures nonconvergence of the iterative procedure is often a problem, particularly for the asymptotic parameter (Biging 1985; Grey 1989).

The problems associated with the parameter prediction method do not exist when the algebraic difference method is used.

3.1.4.3 THE ALGEBRAIC DIFFERENCE METHOD

The algebraic difference method requires time series data and yields either anamorphic or polymorphic top height development and site index curves. This method predicts the top height of a tree or plot from a function including the top height of the tree or plot at age i-p, age i and top height at i-p, where i = time and p = time between remeasurement (Ramirez-Maldonado *et al.* 1988). For example, consider Schumacher's (1939) functional form:

$$ln(H) = \beta_0 + \beta_1 A^{-1}$$

where,

all parameters are previously defined.

Successive measurements of the tree or plot will lay on the same curve thus:

$$\beta_1 = \frac{ln\,(H_i) - ln\,(H_{i-p})}{A_i^{-1} - A_{i-p}^{-1}}$$

$$\therefore \quad ln(H_i) = ln(H_{i-p}) + \beta_1\,(A_i^{-1} - A_{i-p}^{-1})$$

To obtain a site index equation the reference age (A_R) is substituted for A_i such that,

$$ln\,(S) = ln\,(H_{i-p}) + \beta_1\,(A_R^{-1} - A_{i-p}^{-1})$$

Bailey and Clutter (1974) demonstrate that the use of the algebraic difference method yields anamorphic or polymorphic site index and top height development curves which are invariant to the choice of reference age. Another desirable attribute of this method is that top height at the reference age is equivalent to site index. The algebraic difference method using the Schumacher functional form has been used by Smith and Watts (1987), however the method is equally applicable to other functional forms (Clutter *et al.* 1983; Cieszewski and Bella 1989; Rayner 1991).

In an interesting application of the method the algebraic difference form of Schumacher's functional form, which yielded anamorphic top height development curves, was fitted to young (<15 yrs) slash pine (*Pinus elliotti* Engelm). The equation was then joined to an algebraic difference functional form (fitted to trees >15 yrs), based on the Clutter and Jones (1980) polymorphic height increment model, by splining (Borders *et al.* 1984). The resulting system was thus anamorphic for trees younger than 15 years of age and polymorphic for those older than 15 years. It also possessed all the desired attributes of a site index system namely, top height is zero when age is zero, top height at the reference age equals site index, each curve has a separate upper asymptote and the curves are invariant with respect to the choice of reference age.

A related approach models height growth directly by fitting top height increment as a differential function, which on integration over age, yields a model for total top height growth. Inherent in this method is the assumption that instantaneous height growth applies at the average of the two successive ages. An assumption questioned by Borders *et al.* (1984). The differential method has been used to develop polymorphic top height development curves for loblolly pine (*Pinus taeda* L.) by joining the integrated form of the height increment models by segmented regression techniques (Devan and Burkhart 1982). Garcia (1983) used a differential form of Richards functional form (Richards 1959) for predicting top height growth of radiata pine (*Pinus radiata* D. Don) in New Zealand. His model also incorporated a component representative of measurement error.

The differential approach will represent current change in top height in terms of current age. Extending this approach using integro-differential equations, current change is represented in terms of both current age and the sum of all past top heights from the initial age to present (Hamlin and Leary 1987). The advantage gained is the ability to identify trees or plots with similar instantaneous rates of change at a particular age which may have had vastly different patterns of development (Leary and Hamlin 1988).

Although differential equations are considered under the heading of algebraic difference, they are separate entities with conceptual differences (Huseyin 1986). Both methods have received scant attention in the site index literature in spite of their ability to yield top height growth and site index systems with desirable attributes.

3.1.4.4 THE EK-PAYANDEH METHOD

The Ek-Payendeh method yields polymorphic site index curves from time series data. The method involves estimating the parameters of a function from pooled data comprised of site index (S), top height (H) and age (A).

Ek's (1971) expansion of the Chapman-Richards model forms the basis of the methodology, such that:

$$\mathrm{H} = {}_{\beta_0}S^{\beta_1}\left(1 - e^{-\beta_2 A}\right)^{\beta_3 S^{-\beta_4}}$$

In this form the equation may not be solved for S (Payandeh 1974a). To overcome this deficiency Payandeh (1974b) re-arranged the functional form and re-estimated the parameters such that:

$$\mathrm{S} = {}_{\beta_0}H^{\beta_1}\left(1 - e^{-\beta_2 A}\right)^{\beta_3 H^{-\beta_4}}$$

Although the methodology has been used with success (Carmean and Hahn 1981; Carmean and Lenthall 1989), Monserud (1984) found the models lacked the flexibility to track the polymorphism in height growth across the range of site indices evident in his data set.

A problem associated with this methodology is that the predicted height will not equal site index at the reference age. Although this difference is occasionally small enough to be of little practical concern (Newnham 1988) it is nevertheless important. The magnitude of the differences may be decreased by weighted regression (Smith and Watts 1987; Newnham 1987; Carmean and Lenthall 1989).

This technique has received scant attention in the literature with little variation among the functional form used. Modification of other functional forms along the lines described may prove useful.

3.1.4.5 MISCELLANEOUS METHODS

Some methodologies have appeared in the literature which differ from those classes previously discussed. For example, Stout and Shumway (1982) recommend the use of both top height and diameter for developing top height development and site index curves, but fail to explain how their model will accommodate differences in taper caused by silvicultural manipulation and density variations. Zeide (1978) proposed a method whereby two height measurements, at different ages, are required to define a top height development curve. Zeide goes on to claim that the diversity in growth curves of forests throughout the world can be reduced to a few curve types via his method. In both cases no further developments have appeared in the literature.

Smith and Watts (1987) compare a number of methodologies with a method they term the structural method. The structural method takes into account the fact that both top height and site index contain stochastic error. The method compared well to the other methodologies applied to the same data set. Although the structural method performed well for the data set to which it was applied, it is questionable whether equal success would be achieved if a data set of greater age range or pertaining to different species was used (c.f. Smith 1984; Payandeh

1988). The structural method is unlikely to receive much attention because of the disadvantages associated with linear models, which form the basis of the structural method (Payandeh 1983; Smith and Kozak 1984).

3.1.5 POLYMORPHIC NONDISJOINT TOP HEIGHT DEVELOPMENT

Few workers have constructed polymorphic nondisjoint top height development and site index curves. In such cases one or more extrinsic attributes are usually required. For example, Zahner (1962) defined three separate systems for loblolly pine on three separate soil types. To apply this system not only are S, H and A required but also soil type. A similar system was derived by Newberry and Pienaar (1978) involving six soil types. Whether these systems are termed polymorphic nondisjoint or sets of polymorphic disjoint equations pertaining to a particular soil type is mere semantics. However, when a continuous, extrinsic attribute, such as stand density, is included in the equations polymorphic nondisjoint top height development and site index curves are the result (Alexander *et al.* 1967).

3.1.6 FUNCTIONAL FORMS

Functional forms which describe top height development and site index curves are either linear or nonlinear in nature. Linear models are less flexible and may require many parameters to describe the data adequately (Cieszewski and Bella 1989). They may also yield spurious results when used to extrapolate beyond the bounds of the data from which they were developed (c.f. Payandeh 1988; Smith 1988). Nonlinear functional forms are generally more flexible and often have biological bases which provide reasonable estimates upon extrapolation (Pienaar and Turnbull 1973; Smith and Kozak 1984). Although Smith (1984) and Smith and Watts (1987) question the nonlinearity of top height development, most top height development and site index equations in the recent literature are of the nonlinear variety.

Cieszewski and Bella (1989) classify nonlinear functional forms as either fractional forms or modifications of the exponential function. Examples of the fractional class listed by Cieszewski and Bella (1989) include:

$$H = \frac{A^2}{\beta_0 + \beta_1 A + \beta_2 A^2}$$

and;

$$H = \left(\frac{A}{\beta_0 + \beta_1 A^{\beta_2}} \right)^{\beta_3}$$

A fractional functional form which has received little attention is that of Morgan *et al.* (1975): where;

$$H = \frac{\beta_0 \beta_1 + \beta_2 A^{\beta_3}}{\beta_1 + A^{\beta_3}}$$

when $\beta_o = 0$ the model reduces to Hill's (1913) model and when $\beta_o=0$ and $\beta_1=1$ it reduces to the Michaelis-Menten (1913) rectangular hyperbola. The parameter β_o allows the model to have a nonzero intercept.

Exponential functions occur more frequently in the literature and examples include the Gomertz functional form (Yang *et al.* 1978);

$$H = \beta_0 \, e^{-e^{\beta_1 - \beta_2 A}}$$

or;

$$H = e^{\beta_0 - \beta_1 \beta_2 A}$$

the logistic functional form (Monserud 1984);

$$H = \frac{\beta_0}{1 + e^{\beta_0 - \beta_1 A}}$$

the modified Weibull functional form (Yang *et al.* 1978);

$$H = \beta_0 \left(1 - e^{-\beta_1 A^{\beta_2}}\right)$$

the Chapman-Richards functional form (Richards 1959);

$$H = \beta_0 \left(1 - e^{-\beta_1 A}\right)^{(1-\beta_2)^{-1}}$$

and Bailey's (1980) functional form;

$$H = \beta_0 \left(1 - e^{-\beta_1 A^{\beta_2}}\right)^{\beta_3}$$

Many modifications of these functional forms are possible (see Ratkowsky (1983)). For example Ek's (1971) modification of the Chapman-Richards function previously discussed and Monserud's (1984) modification of the logistic functional form, such that:

$$H = \frac{\beta_0 S^{\beta_1}}{1 + e^{\beta_2 - \beta_3 \ln A - f \ln S}}$$

where,

f = a function incorporating forest habitat type.

Although the Chapman-Richards functional form has received prominence in the site index literature no one equation has been found to be universally applicable. As with site index methodology there is a lack of agreement as to which functional form should be used with the choice dependent upon the situation to which it is applied.

3.1.7 PARAMETER ESTIMATION TECHNIQUES

Estimates of parameters for linear functions are obtained by minimising the sums of squares of the errors; the ordinary least squares method. Parameter estimates for nonlinear functional forms are obtained through an iterative procedure until some stopping criterion is met. In each iteration a set of trial coefficients, slightly different from the set in the last iteration, are tested for their fit against the data. The measure of the success of the fit varies with the fitting procedure employed, and iteration continues until no significant improvement can be made to the way the estimates fit the data. Methods for estimation of nonlinear parameters are many (Bard 1974) and some examples appearing in the site index literature include the Gauss-Newton (Ratkowsky 1983) method used by Cieszewski and Bella (1989), Marquardt's (1963) method as used by Borders *et al.* (1984) and the secant method (Ralston and Jennrich 1979) as used by Newnham (1988). Other parameter estimation techniques include maximum likelihood (Garcia 1983) and generalized least squares (Ferguson and Leech 1978). Some of the above techniques are compared by Biging (1985) and Borders *et al.* (1988).

Concern has been expressed about the use of ordinary least squares with repeated measures (Sullivan and Reynolds 1976; Ferguson and Leech 1978). However, Elston and Grizzle (1962) and Sullivan and Clutter (1972) both conclude that ordinary least squares estimators are adequate and it is not necessary to use more elaborate fitting procedures when faced with autocorrelation in a time series. This view has been supported by Gertner (1985), who found that parameter estimates derived via ordinary least squares approximate those derived via generalized least squares for autocorrelated time series data.

In an examination of the effect of autocorrelation on the parameter estimates for top height development and site index curves, Monserud (1984) used a first order autoregressive parameter. The parameter was estimated simultaneously along with the functional parameters. Monserud then forced the autoregressive parameter to equal zero but found the values of the functional parameters remained practically unaltered. Monserud (1984) concluded that the problem of autocorrelation could be ignored without biasing the parameter estimates and found no evidence to question the traditional practise of ignoring the autocorrelation problem when developing top height and site index curves.

Most evidence suggests that the effect of autocorrelation may be ignored. However, some studies suggest the effect will vary according to the functional form and the measurement interval used (Gertner 1985; Borders *et al.* 1988).

Although ordinary least squares estimates of parameters are seemingly unaffected by autocorrelation in a time series, it is well established that the variance estimates can be seriously biased and confidence intervals consequently incorrect (Sullivan and Reynolds 1976). In such cases where the variance changes are large enough to introduce bias, alternative estimation procedures such as autoregressive moving average models (Monserud 1986), maximum likelihood (Garcia 1983) or generalized least squares (Ferguson and Leech 1976) are required.

3.1.8 SITE INDEX AND STOCKING

Site index, as an indicator of the productive capabilities of a site, has gained its popularity partly from the assumption that the measure is independent of stand density. Although some North American conifers appear to be influenced more by density than most other species (Alexander *et al.* 1967; Barrett 1973; Bennett 1975), the general hypothesis will hold true with top height remaining unaffected by density until very low or very high densities are encountered (Carmean 1975; Haggland 1981; Monserud 1984). Even then, it is not universal to have dominant height affected as demonstrated by Pienaar and Shiver (1984), who determined the extent to which the parameter estimates of the Chapman-Richards functional form were affected by differences in planting densities from 370 to 2950 stems ha^{-1}. Results indicate no consistent relationship with any of the estimated parameters and stocking.

Distinction between density and stocking is neither clear cut nor consistently made, despite attempts at definition (Curtis 1970). The estimation of stand density has a variety of purposes, such as to make ecological distinction between species in mixed species stands (Chisman and Schumacher 1940), to develop guidelines for stocking control (Gingrich 1967), to define levels of thinning intensity (Drew and Flewelling 1979) and to predict annual increments (Hall 1979) to name but a few. Sampling rules for the estimation of stand density fall into two general categories (i) plot sampling and (ii) distance sampling (Payandeh and Ek 1986). Distance measures use plots of *n* trees with the plot size defined as a function of the distance from the plot centre to a sample tree or some function of its location. Some estimators based on distance measures are reviewed by Payandeh and Ek (1986). Density estimators from plot samples can be divided into five groups. These are, basal area measures, other diameter based measures, volume based measures, height based measures and crown based measures (Curtis 1970; West 1982). Some of these measures are compared for their suitability to *E. grandis* (Hill) Maiden plantations in South Africa by Bredenkamp and Burkhart (1990).

The hypothesis that top height is unaffected by differences in stand density is poorly tested against the range of stand density estimators available. With some exceptions (Alexander *et al.* 1967; Monserud 1984) the hypothesis is usually tested against stocking (Pienaar and Shiver 1984). There remains a need to examine the universally accepted assumption of the independence of site index from stand density with the true measures of stand density available.

3.2 TOP HEIGHT DEVELOPMENT AND SITE INDEX EQUATIONS FOR *E. GLOBULUS* PLANTATIONS IN SOUTH WEST WESTERN AUSTRALIA

3.2.1 INTRODUCTION

Intensive management of a forest estate requires accurate assessment of site quality. Of the two conceptual schools for site quality assessment, site classification and site evaluation, forest management practices in south west Western Australia have generally taken the site classification approach (Havel 1968, 1975a, 1975b, 1980; Strelein 1988; Wardell-Johnson *et al.* 1989; Inions 1990; Inions *et al.* 1990). Until very recently, the site evaluation approach has not been applied to the Western Australian forest estate (Rayner 1991).

The utility of the site evaluation approach has been discussed elsewhere. The historical development of the approach is described by Tesch (1981) and Monserud (1988) while reviews of the topic are provided by Jones (1969), Carmean (1975), Hagglund (1981) and Grey (1989) and will not be discussed further (see section 3.1).

Uses of land evaluation techniques, in particular site index, are many. In its simplest use site index is merely a label assigned to a tract of land as a symbol of that land's productive potential. Site index more often forms the base for many yield tables (Tesch 1981). For most top height and site index systems a unique top height development curve is generated for each value of site index. The top heights may then be used for the explicit prediction of current and future yield (Sullivan and Clutter 1972) or as input variables in more complex growth functions (e.g; Hahn and Leary 1979; Chang 1988).

It is the aim of this study to develop site index and top height development equations for *E. globulus* in south west Western Australia.

3.2.2 METHODS

3.2.2.1 DATA

Data for deriving top height development patterns and plot site indices were gathered from trees felled in 57 plots established in *E. globulus* plantations in south west Western Australia (Figure 1). Trees considered suitable were the tallest, defect free dominants per plot and were assumed to have been dominants throughout their developments. Height development patterns of c.8 trees per plot, selected from different strata, were compared to that of the site trees. No evidence was found to suggest that site trees were other than dominant through their development.

Eighty seven dominant site trees were felled for stem analysis, procedures for which are detailed in Chapter two (see section 2.2.2.1). Plots were about 0.04ha in area. Top height is defined as the average height of the tallest 40 stems ha^{-1} (i.e., 1 or 2 trees ($\overline{X} = 1.5$ trees) were selected per plot depending upon plot size).

3.2.2.2 PLOT SITE INDEX AND TOP HEIGHT DEVELOPMENT

In order to establish top height growth from stem analysis, a method for estimating the height of the tree at a specified age is required, as the true height at the age corresponding to the ring count at a crosscut will always be located at some distance above that cut (Dyer and Bailey 1987). There are a range of methods for estimating true heights from stem analysis data (Carmean 1972; Lenhart 1972). In a comparison of six such methods, Dyer and Bailey (1987) conclude that the method of Carmean (1972) gave the most accurate results and yielded estimates that were not significantly different from the actual heights.

In this study Carmean's (1972) method was used to estimate the top height of individual trees at consecutive ages. The method assumes that, on average, a crosscut will fall in the middle of a year's height growth. The raw stem analysis data were thus adjusted to estimate the tree height corresponding to the age at each cross-cut.

For each tree, the top height by age co-ordinates where plotted and smoothed by fitting a continuous second derivative cubic spline to the data (Pizer 1975). Each tree was examined for erratic growth. In the few cases where erratic growth occurred, errors were detected in the ring count data and corrected. Site trees from within the same plot displayed similar patterns of top height development. Consequently, the top height development pattern chosen to represent a particular plot was taken to be the average of the top heights, at each age, of the site trees from that plot. Fifty seven such curves were obtained.

In this study the reference age is 5 years. Although a seemingly short space of time it is nonetheless half the rotation length and prevented the need to extrapolate top height growth to obtain site index for all but two of the plots, where an extrapolation of one year was required. Extrapolation was achieved by solving the Chapman-Richards functional form for the reference age. The parameters of the equation were estimated from plots with similar top height development patterns, defined via cluster analysis (see section 3.2.3.3.4).

As the expected rotation length is 10 years all top height development data was truncated at year 12, so as to allow for some extension of the rotation. The data consist of 480 observations of stand top height (H), age (A) and site index (S). Univariate statistics for top height for each year and site index are given in Table 6.

Table 6: Univariate statistics for top height (m) for each year.

Age	Mean Top Height	n	Std. Dev.	Range
1	1.6	57	0.8	0.5-4.5
2	4.5	57	1.7	1.1-8.4
3	7.3	57	2.3	3.3-11.9
4	10.1	57	2.7	5.7-16.0
5 (site index)	12.3	55	2.7	7.3-18.4
6	14.2	43	3.3	8.8-22.5
7	15.6	35	3.8	9.8-27.6
8	16.5	31	3.2	11.4-24.6
9	17.8	26	3.4	13.0-26.4
10	19.0	24	3.4	13.4-28.4
11	20.3	19	4.1	13.7-30.2
12	21.3	19	4.4	14.1-31.6

3.2.2.3 PLOT DENSITY AND STOCKING

Site index has gained its popularity from the assumption that it is independent of stand density. To test the validity of this assumption, plot site index was examined for any correlation with density. At each plot about 40 trees were measured for diameter at breast height over bark (dbhob), bark thickness, total height and crown radius. These data were not applicable for the calculation of density measures from distance sampling procedures, such as those discussed and evaluated by Payandeh and Ek (1986). Of the density measures available from plot sampling (terminology after Payandeh and Ek (1986)) many require additional reference to previously determined relationships (West 1982). Such relationships do not currently exist for *E. globulus* in south west Western Australia, and the data are too limited for their construction.

Consequently, the density measures used in this study are confined to basal area ha^{-1}, stems ha^{-1} and projected crown cover (PCC). PCC is defined as the sum of the crown cross-sectional areas per hectare.

3.2.2.4 NUMERICAL ANALYSIS

3.2.2.4.1 PARAMETER ESTIMATION

Estimation of parameters for linear models was by the standard criterion of least squares. In such cases it is assumed that the errors are independent and identically distributed normal random variables with a mean of zero and a finite variance. The estimates of the unknown parameters are maximum likelihood estimators of the parameters in the model.

Parameter estimates of the nonlinear functional forms were also derived via the least squares criterion. However, estimators for nonlinear equations will not have the properties possessed by those of linear models. The two estimators have properties which are only asymptotically equal. That is, only as the sample size increases to infinity do the properties of the nonlinear estimator approach those of the linear estimator.

To estimate the parameters of the nonlinear functional forms, an iterative process was used where the initial values of the parameters are altered until the error sums of squares is minimized, or at least no significant reduction occurs. In all cases the derivative free secant method of Ralston and Jennrich (1979) was used.

3.2.2.4.2 PATTERN RECOGNITION: CLASSIFICATION OF PLOTS BY TOP HEIGHT DEVELOPMENT PATTERN.

The use of numerical classification procedures has been used often for the analysis of resource data (Webb *et al.* 1984; Carleton *et al.* 1985; Curry and Slater 1986; Callaway and Clebsch 1987; Pojar *et al.* 1987). In most cases the procedure begins by creating a matrix of pairwise associations between plots or sites via some association measure, such as the Gower metric (Gower 1971) or the Bray-Curtis metric (Bray and Curtis 1957). As such each attribute, on which the classification is based, is considered to be an independent entity. This assumption is incorrect when applied to top height development data as the top height of a stand at any one year is not independent of the top height in other years. The application of standard association metrics will not account for the time series inherent in the data. To address this problem a matrix of associations between plots was calculated using the two-dimensional profile algorithm of Faith *et al.* (1985).

The algorithm adds to the usual quantitative dimension of each attribute, a second dimension of order thereby creating what Faith *et al.* (1985) term profile attributes. A profile attribute is then made up of individual characters, in this case top height, that are explicitly ordered along a second dimension, in this case time. The measure is related to the general form for measures of spatial autocorrelation based upon cross products (Hubert *et al.* 1981). Application of the algorithm requires a parameter P to be specified. This parameter affects the neighbourhood of influence (terminology after Faith *et al.* (1985)). As the neighbourhood of influence becomes broader the metric will produce distance measures which approach measures derived via the application of standard association metrics. In this study P was set to equal 0.5, thus defining a narrow neighbourhood of influence.

The hierarchical agglomerative clustering strategy, the unweighted pair-group arithmetic averaging (UPGMA) was used to impose structure to the association matrix (Gauch and Whittaker 1981). Although the method is sometimes prone to minor misclassifications, it has the advantage of taking more than one plot into account at any fusion. The clustering intensity coefficient *beta* was set at 0 (Booth 1978).

For comparative purposes, a second association matrix was derived via the Bray-Curtis metric (Bray-Curtis 1957). As such, the top height at each year is treated as an independent attribute. The UPGMA fusion strategy was used to impose structure to the association matrix. The clustering intensity coefficient *beta* was set at -0.1. Under such conditions the clustering strategy is slightly "space-dilating" and resists the formation of a single large group by favouring the formation of a number of even-sized groups (Booth 1978).

The acceptability of imposed groups was examined by ordinating the sites with principal co-ordinate analysis and examining the position of group members in component space (Gower 1967).

3.2.3 RESULTS

3.2.3.1 SITE INDEX AND DENSITY

Univariate statistics for three measures of stand density are given in Table 7. No significant correlation was detected between plot site index and PCC or basal area. The correlation between site index and stocking was significant ($p<0.01$) (Table 8). This probably results from the practice of planting less stems ha^{-1} in areas of low rainfall, thus low site index and higher stockings in areas of high rainfall thus higher site index. Regressing plot site index on stocking yields an equation which explains only 11.9% of the variance. It is assumed that the relationship between plot site index and stocking is an artifact of management history, rather than a meaningful relationship. Therefore, the assumption of independence between plot site index and stand density is not seriously violated in this study.

Table 7: Univariate statistics for stocking (stems ha^{-1}), basal area (m^2ha^{-1}) and projected crown cover (m^2ha^{-1}).

Variable	Mean	N*	Std. Dev	Range
Stocking	832	53	372	180-2000
Basal area	14.0	53	7.5	2.9-32.7
PCC	3975.2	53	2271.1	1259-11395

* 4 plots have no density measures associated

Table 8: Pearson's correlation coefficients between plot site index and 3 measures of stand density. Significance levels are given in parenthesis.

Variable	Basal area	PCC	Stocking
Site index	0.0184 (0.0114)	0.1796 (0.1982)	0.3449 (0.0114)

3.2.3.2 SELECTION OF FUNCTIONAL FORMS

Polymorphism was evident in the graphs of top height development. Therefore, the functional form selected to explain top height development would need to be flexible in nature and preferably sigmoidal. Five candidate functional forms were tested for their ability to fit the data. These were the Schumacher functional form (Schumacher 1939), the Chapman-Richards functional form (Richards 1959; Chapman 1961), the Weibull functional form (Yang *et al* 1978), the logistic functional form (Monserud 1984) and Hossfald's fractional functional form (Zakrzewski 1986). Parameter estimates and model statistics are given in Table 9.

The model statistics pertaining to the Schumacher functional form are not directly comparable to the other models as they pertain to a different dependent variable, namely *ln* (H). The Schumacher functional form also lacked the ability to track the sigmoidal nature evident in many of the top height development patterns. The fractional functional form of Hossfald did not fit the data well and was not considered further. Of the three remaining models the Chapman-Richards functional form displayed the most desirable model statistics and will form the basis of subsequent model development.

Table 9: Parameter estimates and model statistics for 5 candidate functional forms.

Functional form	β_0	β_1	β_2	A	B	C	D
Schumacher	3.05	-2.83	-	50.82	0.11	0.00	0.33
Chapman-Richards	23.65	0.21	1.51	3794.97	7.96	-0.02	2.81
Weibull	22.76	-0.92	1.31	3806.05	7.98	-0.29	2.82
Logistic	6.14	-0.25	-	8745.05	18.29	0.00	4.27
Hossfald	-18.10	4.71	-0.27	31933.65	76.03	5.81	6.46

A = Residual sums of squares
B = Residual mean square
C = Mean residual
D= Standard deviation of the residuals

3.2.3.3 TOP HEIGHT DEVELOPMENT AND SITE INDEX EQUATIONS

3.2.3.3.1 THE EK-PAYANDEH METHOD

The parameters of Ek's (1971) and Payandeh's (1978) modifications of the Chapman-Richards functional form, such that:

$$H = {}_{\beta_0}S^{\beta_1}\left(1-e^{-\beta_2 A}\right)^{\beta_3 S^{-\beta_4}} \quad [11]$$

$$S = {}_{\beta_0}H^{\beta_1}\left(1-e^{-\beta_2 A}\right)^{\beta_3 H^{-\beta_4}} \quad [12]$$

were estimated. For each equation a number of starting values were tried during the parameter estimation iterative procedure, to ensure the estimates were not local solutions. The iterative procedure used to estimate the parameters of Eq.[12] did not meet the convergence criterion and the parameters presented are those of the equation with the smallest sum of the squared errors. In an attempt to meet the convergence criterion the $H^{-\beta_4}$ term was dropped from Eq.[12] such that:

$$S = \beta_0 H^{\beta_1} \left(1 - e^{-\beta_2 A}\right)^{\beta_3}$$

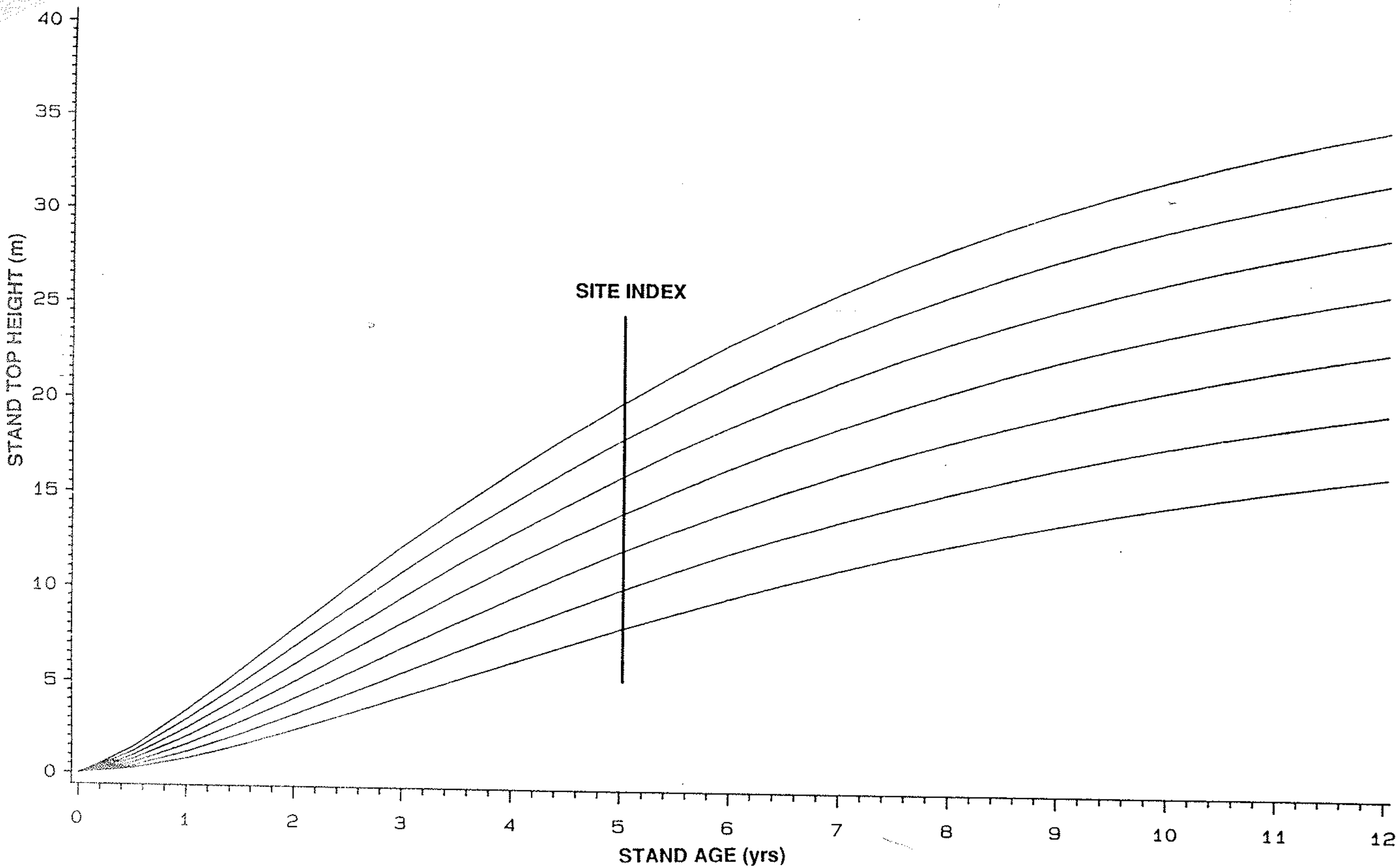
In this case the iterative procedure did meet the convergence criterion. Parameter estimates and model statistics are given in Table 10. Residual analysis showed that no heteroscedasticity was evident. The resulting equations Eqs.[11] and [13] yield doubly asymptotic, polymorphic top height development and site index curves (Figure 6).

Table 10: Parameter estimates and model statistics for the Ek-Payandeh models

Equation	β_0	β_1	β_2	β_3	β_4	A	B	C	D
[11]	4.27	0.76	0.17	3.03	0.27	707.78	1.49	0.00	1.10
[12]	10.27	0.06	47.53	975.61	63.28	3779.67	7.96	0.75	2.60
[13]	1.43	0.59	0.12	-0.86	-	1458.32	3.06	0.00	1.75

A = Residual sums of squares
B = Residual mean squres
C = Residual mean
D = Standard deviation of the residuals

Figure 6: **Top height development curves derived via the Ek-Payandeh method (Eq. [11]).**



3.2.3.3.2 THE PARAMETER PREDICTION METHOD

Two functional forms were considered for this methodology. Firstly, Schumacher's functional form because only two parameters require estimation. Secondly, the Chapman-Richards functional form because it is considered the most appropriate for this data set.

Schumacher's functional form.

The β_0 and β_1 parameters of Schumacher's functional form were estimated for each plot. These parameters were then estimated from functions of S such that:

$$ln(H) = \beta_0 + \beta_1 A^{-1} \qquad [14]$$

where,

$$\beta_0 = 2.4 + 0.05S \qquad (r^2 = 0.46) \qquad [14a]$$

and,

$$\beta_1 = 3.38 - 3.07\beta_0 + 0.21S \qquad (r^2 = 0.72) \qquad [14b]$$

Analysis of the residual errors from predicted heights showed that although the mean error was small ($\overline{X}$ =0.82m) the spread was large (std.dev. = 2.2m). A relationship between plot age and error magnitude was evident. Therefore, an age term was incorporated into Eqs. [14a] and [14b] such that:

$$\beta_0 = 2.39 - 2.39A^{-1} + 0.08S \qquad (r^2 = 0.68) \qquad [14c]$$

and,

$$\beta_1 = 5.80-4.69A^{-1} + 0.31S-3.87\beta_0 \qquad (r^2=0.78) \qquad [14d]$$

Although the incorporation of an age term into the parameter prediction equations improved the r^2 of the models, no improvement in the residual statistics was detected ($\overline{X}$ = 0.38m; std dev. = 4.26m). Use of equations [14c] and [14d] will also have the undesirable property of yielding different values for β_0 and β_1 for different ages. Eqs. [14c] and [14d] are not considered further.

Chapman-Richards functional form

The Chapman-Richards functional form was fitted to each plot. Nonconvergence of the iterative procedure occurred in six plots so parameter estimates from these were excluded from further analysis. The three parameters of the Chapman-Richards functional form were expressed as functions of S such that:

$$H = \beta_0\left(1 - e^{-\beta_1 A}\right)^{(1-\beta_2)^{-1}} \quad [15]$$

where,

$$ß_o = -341.43+15.13S + 8.0 \times 10^{-6} S^5 + 2153.98 S^{-1} \quad [15a]$$
$(r^2 = 0.35)$

$$ß_1 = 0.84 - 0.10 S - 0.005 ß_o + 0.005 S^2 \quad [15b]$$
$(r^2 = 0.58)$

$$ß_2 = 1.34 - 0.036 S + 4.20/ß_1 \quad [15c]$$
$(r^2 = 0.66)$

This set of equations has a negative bias and a large spread of errors ($\overline{X}$ = -2.7 m; std.dev. = 5.17 m). This may stem from the correlation which exists between parameters (Table 11). Thus variation in parameter estimates, which are unrelated to variation in site conditions, may be compensated for by other parameters, reducing the equations ability to explain large amounts of the variation among parameters.

Table 11: Pearson's correlation coefficients between the parameter estimates of the Chapman-Richards functional form fitted to each plot. Significance levels are given in parenthesis.

	β_0	β_1	β_2
β_0	-		
β_1	-0.60 (0.0001)	-	
β_2	-0.41 (0.0025)	0.80 (0.0001)	-

In an attempt to overcome this problem, the β_2 parameter was expressed as a function of the β_1 parameter such that:

$$\beta_2 = 0.94 + 3.97\,\beta_1 \qquad (r^2 = 0.64) \qquad [16]$$

Eq. [16] was substituted back into the Chapman-Richards functional form such that:

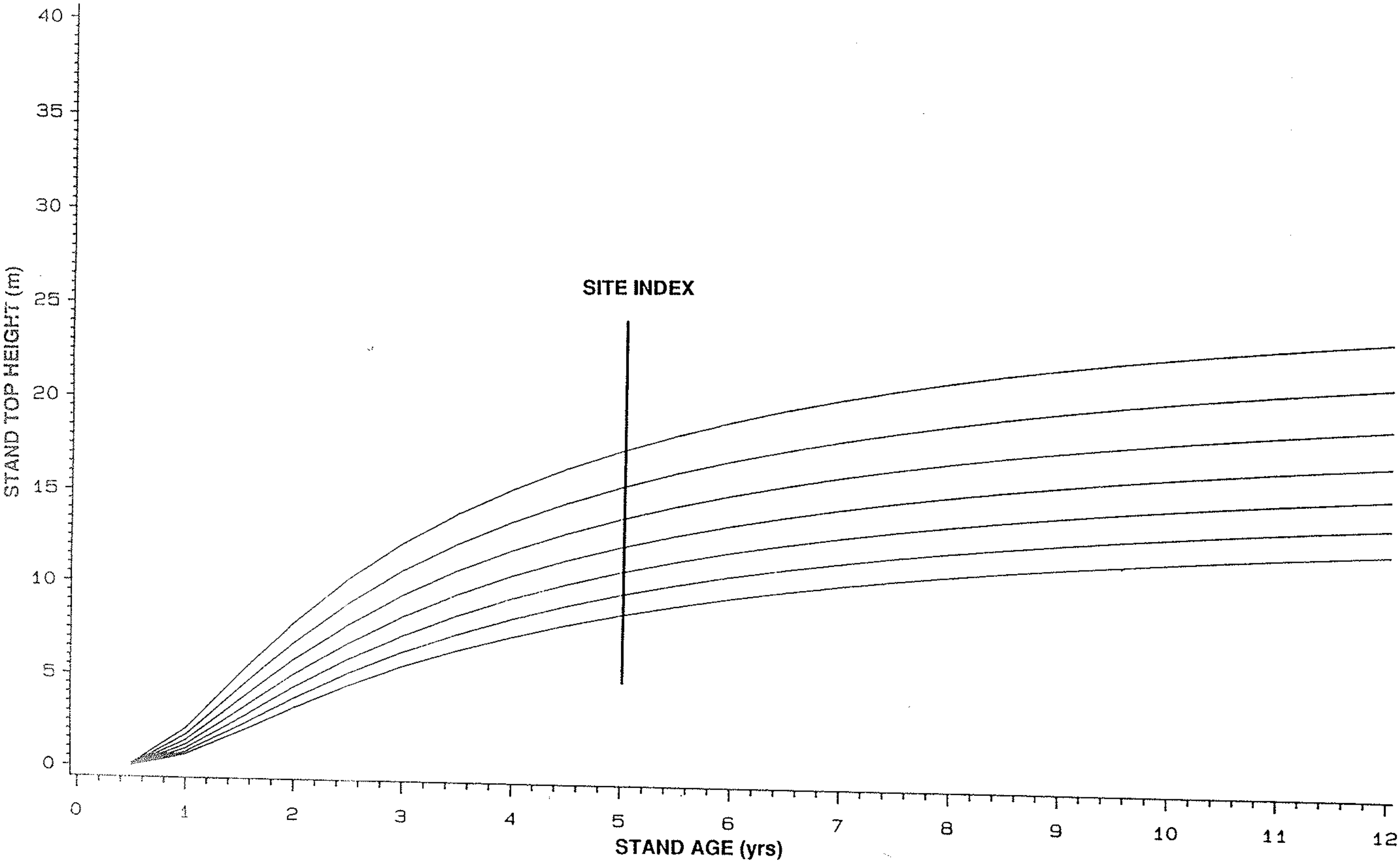
$$H = \beta_0\,(1 - e^{-\beta_1 A})^{[0.94 + 3.97\,\beta_1]} \qquad [17]$$

and the parameters β_0 and β_1 were estimated for each plot. In this case the convergence criterion of the iterative procedure was met for all but one of the plots.

The sum of the individual residual sums of squares resulting from the fit of Eq. [17] to each plot was significantly ($p<0.05$) larger than that yielded by applying the standard three parameter Chapman-Richards functional form to each plot. Only plots which met the convergence criterion in both cases were used in this comparison. Expressing the individual parameter estimates from Eq. [17] as functions of S showed no relationship, with the r^2 values for the relationship between β_0 and β_1 and S being 0.03 and 0.09 respectively.

As the predictions resulting from Eqs. [15], [15a], [15b] and [15c] contain unacceptable errors and yield top height development patterns devoid of reality for most values of S, Eqs. [14] [14a] and [14b] will be used to yield top height development curves via the parameter prediction method. These equations yield polymorphic top height development curves (Figure 7).

Figure 7: **Top height development curves derived via the parameter prediction method (Eq. [14], [14a], [14b]).**



3.2.3.3.3 THE ALGEBRAIC DIFFERENCE METHOD

When estimating the parameters of the algebraic difference equation, based on the Chapman-Richards functional form, non-overlapping increment periods were used. That is, the top height increment from A=1 to A=2 and A=2 to A=3 etc were used, but not A=1 to A=3.

The following equation resulted:

$$H_2 = H_1 \left[\frac{1 - e^{-0.15A_2}}{1 - e^{-0.15A_1}} \right]^{1.43} \qquad [18]$$

where,

H_2 = top height at age A_2

H_1 = top height at age A_1

A_2 = age at measurement 2

A_1 = age at measurement 1

Residual sums of squares	=	333.85 m
Residual mean squares	=	0.79 m
Residual mean	=	0.12 m
Standard deviation of residuals	=	0.88 m
N	=	422

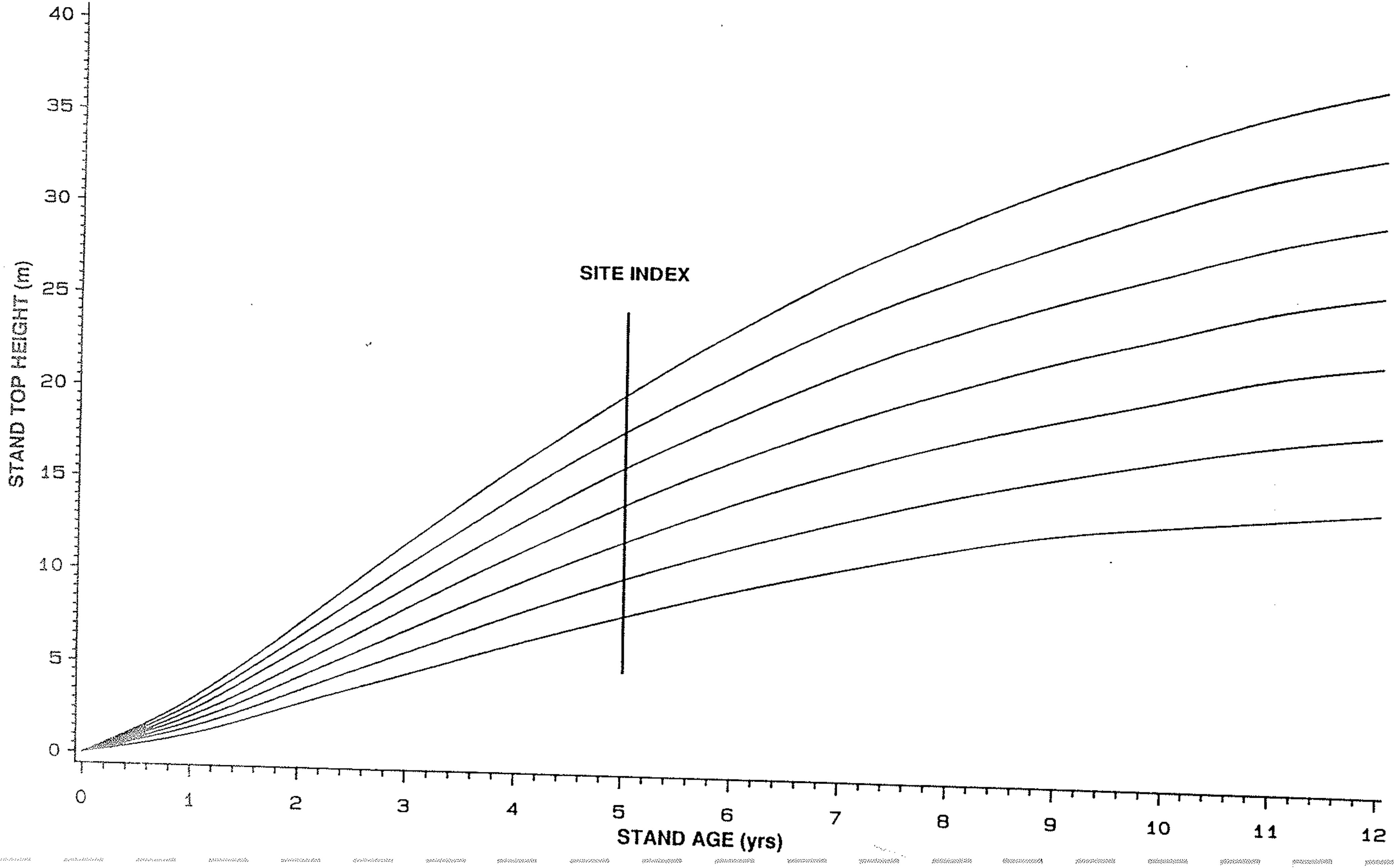
Residuals resulting from Eq. [18] displayed desirable qualities with no evidence of heteroscedasticity.

Substitution of the reference age for A_2 and S for H_2 yields the site index equation:

$$S = H_1 \left[\frac{1 - e^{-0.75}}{1 - e^{-0.15A_1}} \right]^{1.43} \qquad [19]$$

Application of Eqs. [18] and [19] yields doubly asymptotic top height development and site index curves (Figure 8).

Figure 8: Top height development curves derived via the algebratic difference method (Eq. [18]).



3.2.3.3.4 CLASSIFICATION OF PLOTS BY TOP HEIGHT DEVELOPMENT PATTERN

The classification using standard cluster analysis techniques (i.e., metric = Bray-Curtis; fusion strategy = UPGMA, *beta* = -0.1) was arbitrarily truncated at the seven group level (Figure 9). In this case each attribute of the plot (i.e., top height at each year) is considered independent.

Group A: contains 2 plots of low site index ($\overline{X}$ = 9.9 m; Std. dev. = 0.2m). These plots show very slow top height development after age 4.

Group B: contains 11 plots of low site index ($\overline{X}$ = 9.8 m; Std. dev. = 0.5m). Group B does not show the marked slowing of growth as does group A.

Group C: contains 5 plots with very low site index values ($\overline{X}$ = 7.8 m; Std. dev. = 0.5 m).

Group D: contains 8 plots of medium site index ($\overline{X}$ = 13.0 m; Std. dev. = 1.3 m). This group displays the greatest within group variability for site index.

Group E: contains 10 plots of high site index ($\overline{X}$ = 14.9 m; Std.dev. = 0.8 m).

Group F: contains 15 plots of medium site index ($\overline{X}$ = 11.9 m; Std. dev. = 0.6 m).

Group G: contains 6 plots of the highest site index ($\overline{X}$ = 18.4 m; Std. dev. = 0.8 m).

Although this classification minimized the within-group variance while maximizing the between group variance for top height, some within group polymorphism was evident. Consider the top height development curves for the 10 plots which comprise group E (Figure 10). Although all plots in this group have similar values for site index, some polymorphism exists, particularly for ages greater than the reference age.

Figure 9: **Dendrogram resulting from the classification of plots based on top height (metric = Bray-Curtis; fusion strategy = UPGMA, beta = -0.1).**

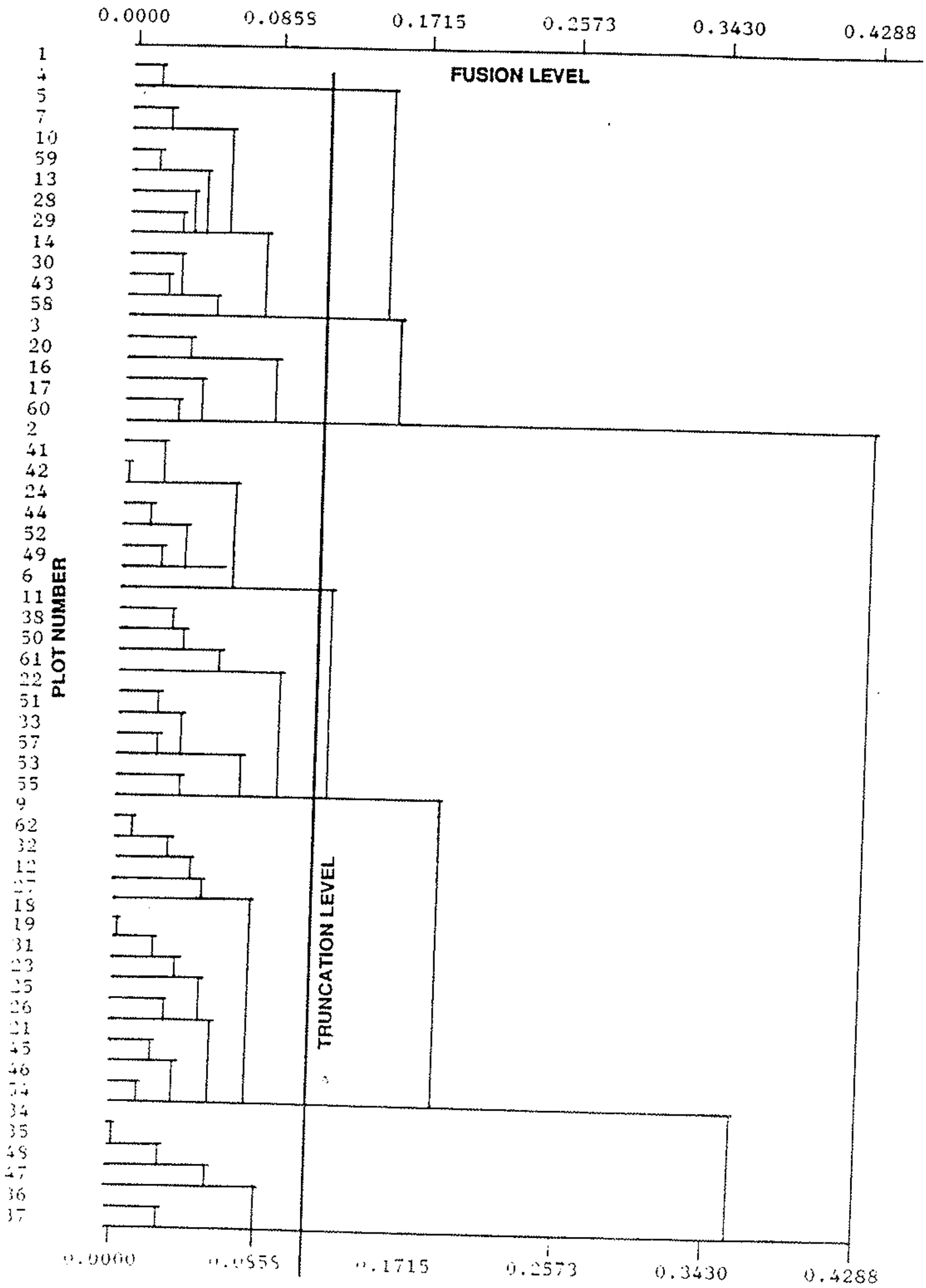
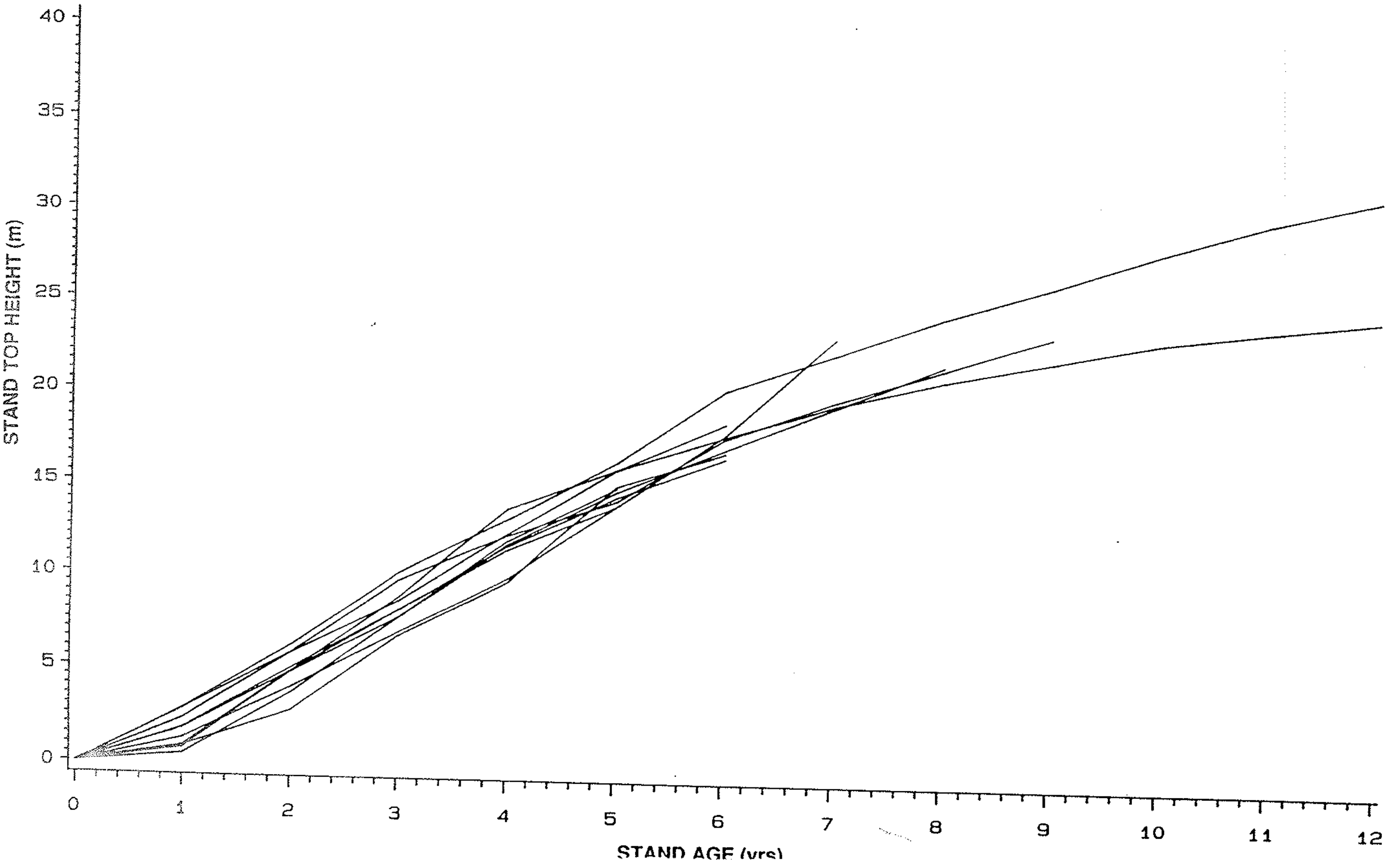


Figure 10: **Top height development curves of plots which comprise cluster group E.**



The classification of plots when the top height development pattern was considered as a two-dimensional profile attribute, after Faith *et al.* (1985), was arbitrarily truncated at the eight group level (Figure 11). Group membership was examined in component space as a guide to the success of the classification and its level of truncation (Figure 12). Good separation of group members was achieved along principal axis 1, which explained 95.5% of the variance accounted for.

Group 1: contains the same 2 plots of group A, defined via standard clustering techniques.

Group 2: contains 13 plots of medium site index ($\overline{X}$ = 13.7 m; Std. dev. = 1.1 m). Plots in this group are characterized by rapid initial top height growth, with a sharp decline in the rate of top height development after the reference age.

Group 3: contains 14 plots of low to medium site index ($\overline{X}$ = 11.9 m; Std. dev. = 0.6 m). Plots in this group are characterized by slow initial growth followed by a minimum of curvature in top height development pattern after age three.

Group 4: contains 5 plots of medium to high site index ($\overline{X}$ = 14.6; Std. dev. = 0.5 m). Plots in this group are characterized by slow initial top height development becoming rapid after the third year.

Group 5: contains 5 plots of very low site index ($\overline{X}$ = 7.8 m; Std. dev. = 0.5 m). Plots in this group are characterized by very slow top height development at young ages and a linear top height development pattern after age 3.

Group 6: contains 11 plots of low site index ($\overline{X}$ = 9.9 m; Std. dev. = 0.6 m). Plots in this group are characterized by slow early top height development with a sharp decline in top height growth after the reference age.

Group 7: contains 6 plots of high site index ($\overline{X}$ = 18.1 m; Std. dev. = 1.0 m). Plots in this group are characterized by rapid top height development throughout the rotation, particularly for ages 1 to 7.

Group 8: contains 1 plot only with a site index of 19.3 m. The top height development pattern for this plot is similar to that of group 7 and will be treated as a member of that group for subsequent analysis and discussion.

Treating top height development as a two dimensional profile attribute allowed the formation of groups of similar top height development pattern. The within group polymorphism evident in Figure 10 does not occur when the methodology was applied. For example, group E, defined via standard clustering methods, is split with five members allocated to other groups. The remaining members form group 4. No evidence of within group polymorphism exists for this group (Figure 13).

Figure 11: Dendrogram resulting from the classification of plots based on top height development pattern (metric = two dimensional profile algorithm; fusion strategy = UPGMA, beta = 0).

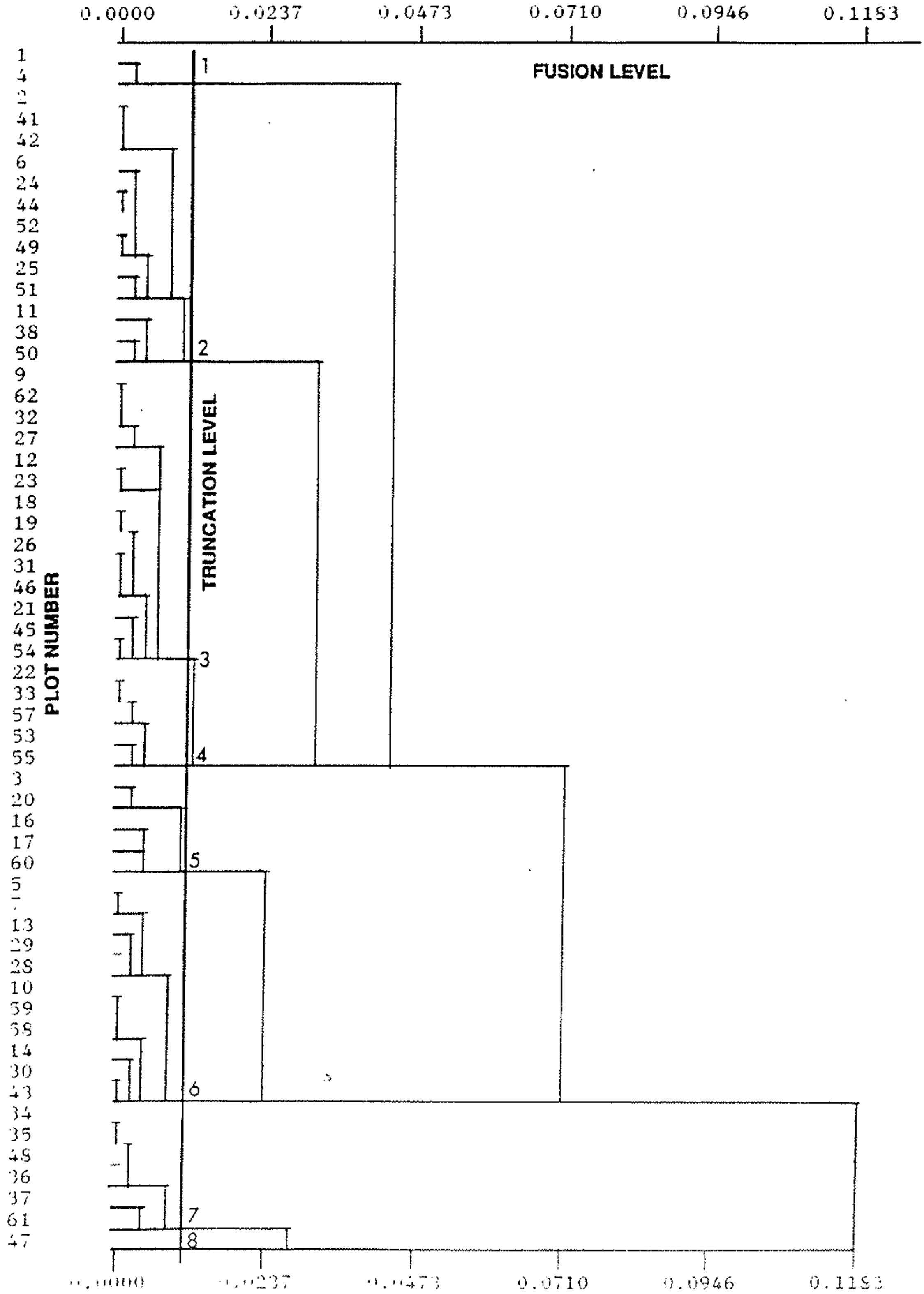


Figure 12: Position in component space of group members, defined via cluster analysis based on top height development patterns.

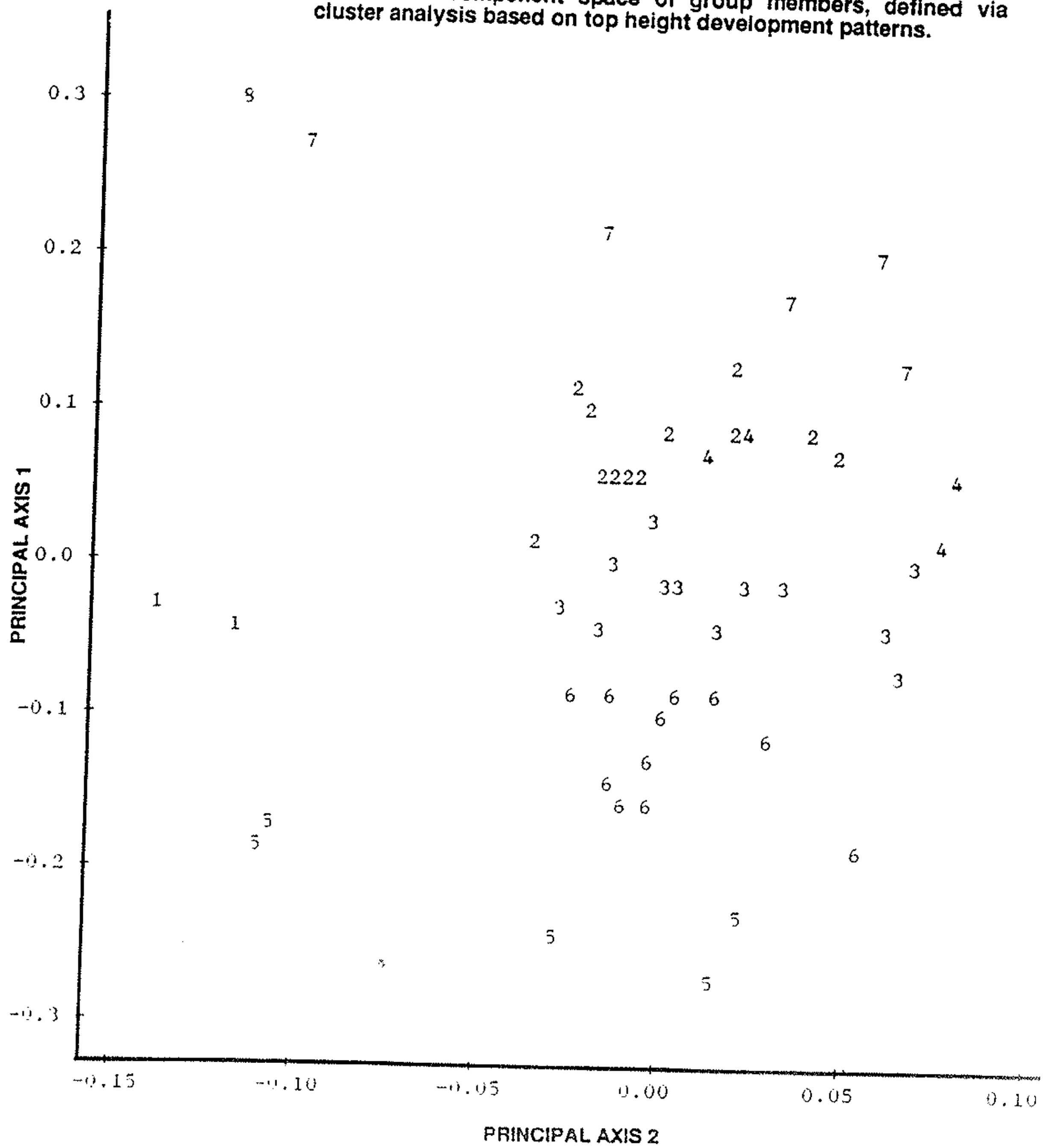
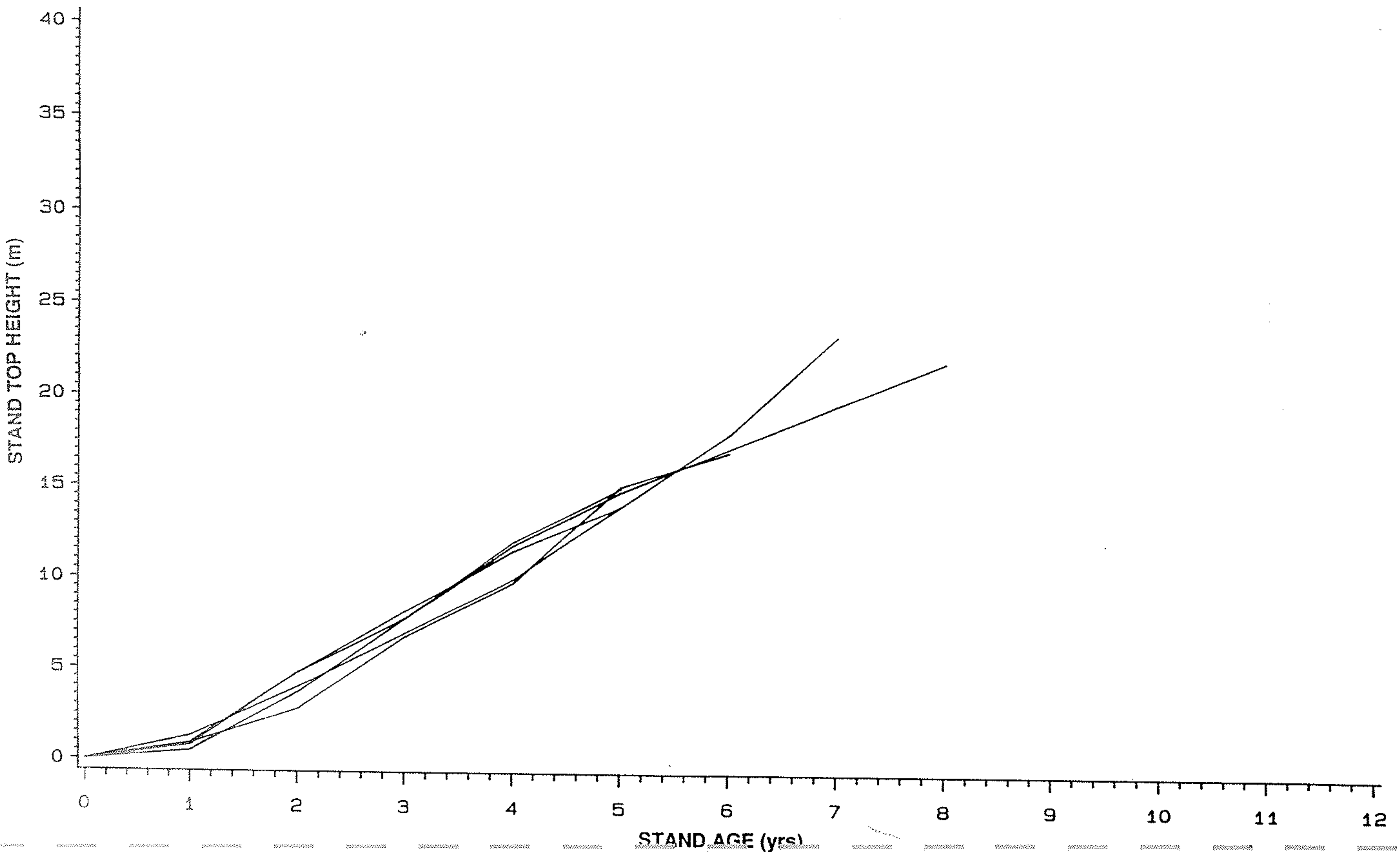


Figure 13: **Top height development curves of plots which comprise cluster group 4.**



In an attempt to track the polymorphism evident in the data, the parameters of the Ek-Payandeh functional form were estimated for each cluster group. In all cases difficulty was experienced meeting the convergence criterion of the iterative procedure for other than local solutions. As much of the polymorphism had now been confined to that which exists between groups, the standard form of the Ek-Payandeh functional form was probably over-parameterised. The functional form was thus modified to:

$$H = {}_{\beta_0}S^{\beta_1} \left(1 - e^{-\beta_2 A}\right)^{\beta_3} \qquad [20]$$

Parameter estimates and model statistics resulting from the application of Eq. [20] to each cluster group are given in Table 12. The resulting set of equations yields doubly asymptotic polymorphic nondisjoint top height development curves (Figure 14). To derive Figure 14, the mean value of S for each cluster was used when applying Eq. [20].

Using the method of Ratkowsky (1983), the hypothesis, that the data for cluster groups could be combined and explained by a single equation, was tested for all pairs. The hypothesis was rejected (p>0.0001) in each case.

Parameters of algebraic difference form of the Chapman-Richards functional form were estimated for each of the cluster groups. The resulting set of equations yields doubly asymptotic polymorphic nondisjoint top height development curves (Figure 15). This system of equations also yield site index equations when the reference age is substituted for the A_2 term and S is substituted for the H_2 term.

The hypothesis that the data for cluster groups could be combined was again tested. The hypothesis was again rejected for all combinations (p>0.0001). The parameter estimates and model statistics resulting from fitting the algebraic difference form of the Chapman-Richards functional form to each cluster group are given in Table 13.

Table 12: Parameter estimates and model statistics derived from fitting the modified Ek-Payandeh functional form (Eq. [20]) to each cluster group.

Cluster group	β_0	β_1	β_2	β_3	A	B	C	D
1	2.88	0.71	0.24	.03	1.40	0.09	-0.01	0.28
2	5.47	0.67	0.14	1.21	68.43	0.76	0.02	0.85
3	2.29	1.02	0.17	1.65	75.58	0.64	-0.01	0.70
4	11.12	0.49	0.17	1.89	26.93	1.00	0.02	0.94
5	7.80	0.71	0.09	1.47	36.64	0.72	-0.03	0.75
6	19.12	0.08	0.20	1.87	129.52	1.17	0.01	0.92
7	1.56	1.15	0.17	1.49	33.61	0.86	0.03	0.91

A = Residual sums of squares
B = Residual mean square
C = Mean residual
D = Standard deviation of the residuals

Table 13: Parameter estimates and model statistics derived from fitting an algebraic difference equation to each cluster group.

Cluster group	β_0	β_1	A	B	C	D
1	0.36	1.37	3.30	0.21	0.08	0.43
2	0.21	1.35	51.12	0.67	0.07	0.81
3	0.20	1.78	40.12	0.38	0.06	0.61
4	0.24	2.18	35.71	1.49	0.15	1.18
5	0.06	1.29	20.25	0.42	0.06	0.64
6	0.16	1.58	37.30	0.36	0.03	0.60
7	0.15	1.40	64.26	1.89	0.18	1.34

A = Residual sums of squares
B = Residual mean square
C = Mean residual
D = Standard deviation of the residuals

Figure 14: **Top height development curves derived via applying Eq. [20] to each cluster group.**

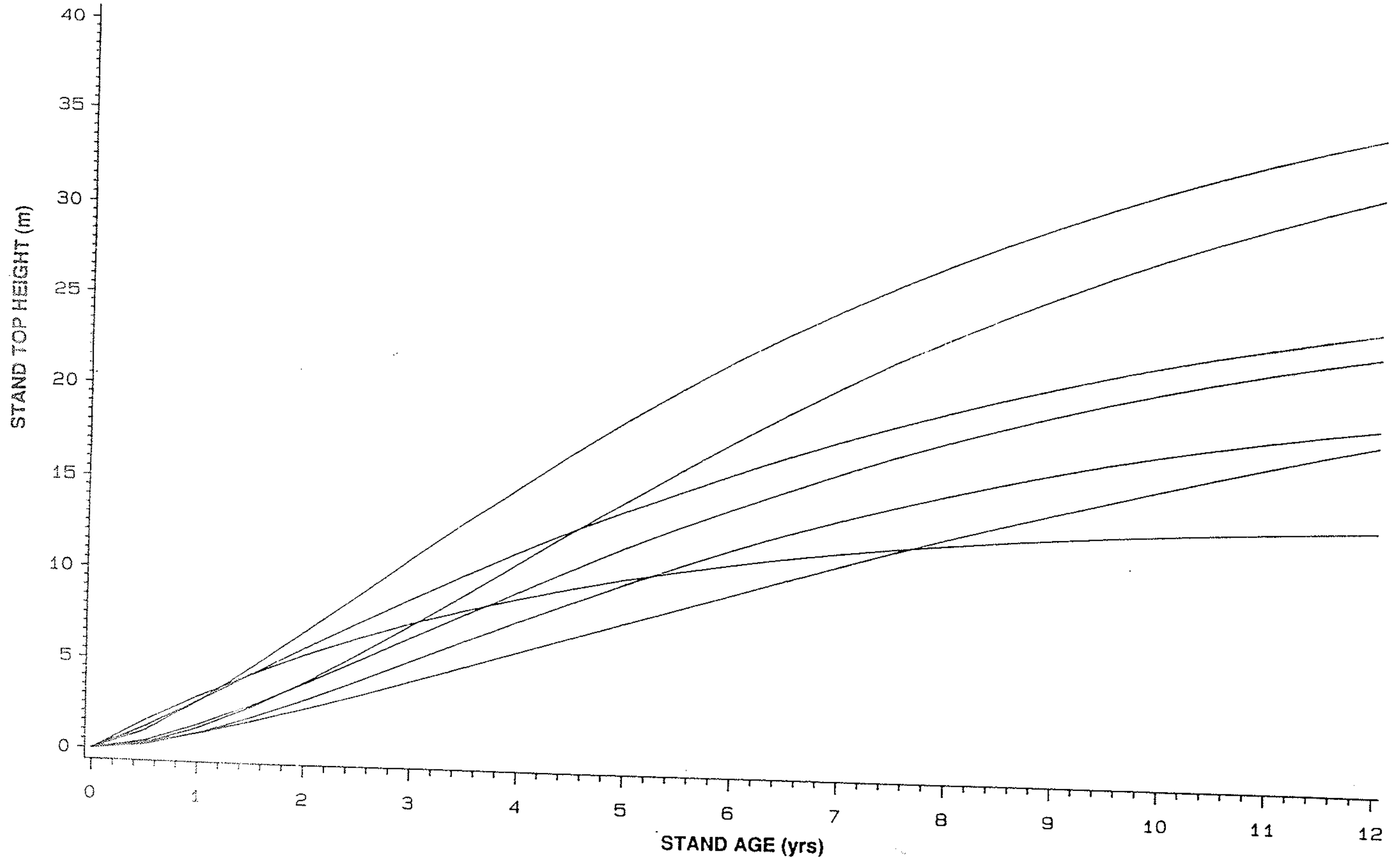
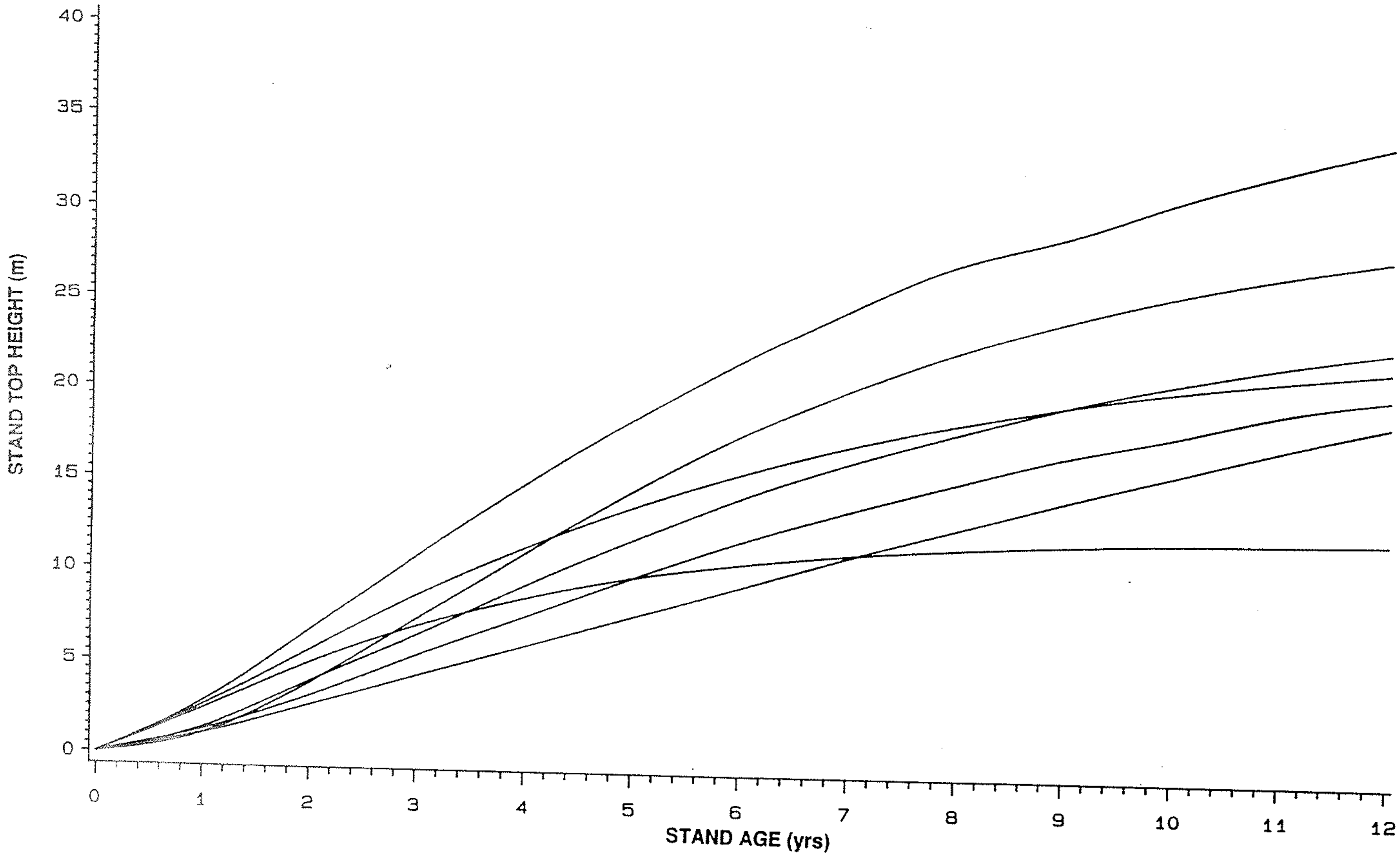


Figure 15: Top height development curves derived via fitting an algebratic difference form of the Chapman-Richards functional form to each cluster group.



3.2.3.4 RESIDUAL ANALYSIS OF CANDIDATE TOP HEIGHT DEVELOPMENT AND SITE INDEX EQUATIONS

3.2.3.4.1 TOP HEIGHT DEVELOPMENT EQUATIONS

The residuals of candidate top height development equations were compared. The two equations derived via the parameter prediction method (Eqs. [14] and [15]) yielded the worst residual statistics. Eqs. [14] and [15] displayed large mean residuals with large variances. The Ek-Payandeh equation (Eq. [11]) yielded the smallest mean residual but suffered from a wide spread of errors. The algebraic difference equation (Eq. [18]) and the modified Ek-Payandeh equation fitted to each cluster group (Eq. [20]) both exhibited small mean residuals and the smallest spread.

It should be pointed out that the residuals generated by application of Eq. [18] will appear more favourable than is warranted. At each prediction it is assumed that A_1, A_2 and H_1 are known where, in reality H_1 will also be a prediction and contain error. Equations [18], [20] and Eq. [18] applied to each cluster group, yield the most favourable residual statistics (Table 14).

Table 14: Univariate statistics for the residuals generated from individual top height development equations.

Equation	Mean	Std. Dev	Range	Skewness	Kurtosis
[11]	0.06	1.22	-5.77 - 4.84	0.01	3.68
[14]	0.66	3.39	-11.61 - 9.91	-1.10	1.67
[15]	-1.69	4.63	-23.57 - 8.11	-1.67	3.31
[18]	0.12	0.88	-2.58 - 3.58	0.49	1.73
[20]	0.15	0.90	-3.50 - 4.48	0.40	2.52
[18]*	0.07	0.79	-2.36 - 3.63	0.72	2.44

* Eq. [18] applied to each cluster group

In the presence of autocorrelation estimates of the parameters of an equation will remain theoretically unbiased, but the estimation procedure will become less efficient (Gertner 1985). However, the estimates of variance will be biased, resulting in confidence intervals that are too narrow (Draper and Smith 1981). As a guide to the magnitude to which the residuals are correlated, a residual correlation matrix was calculated for each equation (Table 15 a,b,c,d,e,f). Residuals from all top height development equations show significant correlation between years. However, the autocorrelation for Eqs. [11], [14] and [15] were particularly prominent. The autocorrelation between residuals for Eqs. [18], [20] and Eq. [18] applied to each cluster group are generally confined to the first order lag intervals (i.e., between years 1 and 2, 2 and 3, 3 and 4 etc) making these two equations more attractive than the other candidates. Of all equations Eq. [18] applied to each cluster group displays less correlation among residuals than other equations.

Table 15: Correlation matrix between residuals derived via the following methods (a) Ek-Payandeh Eq. [11], (b) parameter prediction Eq. [14], Schumacher's functional form, (c) parameter prediction Eq. [15] Chapman-Richards functional form, (d) the difference equation Eq. [18], (e) the modified Ek-Payandeh equation Eq. [20] applied to each cluster and (f) the difference equation Eq. [18] applied to each cluster. The significance levels of the correlation, where significant are given in parenthesis.

(A) Year	1	2	3	4	5	6	7	8	9	10	11	
1	-											
2	0.71 (0.0001)	-										
3	0.52 (0.0001)	0.88 (0.0001)	-									
4	0.29 (0.03)	0.69 (0.0001)	0.83 (0.0001)	-								
5	-0.10 -	-0.13 -	-0.01 (0.04)	0.27 -	-							
6	-0.42 (0.004)	-0.48 (0.001)	-0.33 (0.03)	-0.25 -	-0.09 -	-						
7	-0.53 (0.0009)	-0.63 (0.0001)	-0.43 (0.009)	-0.44 (0.008)	0.01 -	0.90 (0.0001)	-					
8	-0.46 (0.01)	-0.59 (0.0004)	-0.43 (0.01)	-0.39 (0.03)	-0.09 -	0.78 (0.0001)	0.94 (0.0001)	-				
9	-0.24 -	-0.20 -	-0.06 -	-0.20 -	-0.21 -	0.54 (0.004)	0.77 (0.0001)	0.92 (0.0001)	-			
10	-0.24 -	-0.18 -	-0.04 -	-0.21 -	-0.18 -	0.49 (0.01)	0.73 (0.0001)	0.87 (0.0001)	0.98 (0.0001)	-		
11	-0.28 -	-0.25 -	-0.17 -	-0.38 -	-0.22 -	0.47 (0.04)	0.75 (0.0002)	0.84 (0.0001)	0.94 (0.0001)	0.98 (0.0001)	-	
12	-0.30 -	-0.27 -	-0.19 -	-0.39 -	-0.22 -	0.44 -	0.75 (0.0002)	0.82 (0.0001)	0.92 (0.0001)	0.96 (0.0001)	0.99 (0.0001)	-

(B) Year	1	2	3	4	5	6	7	8	9	10	11
1	-										
2	0.29 (0.02)	-									
3	0.06 -	0.87 (0.0001)	-								
4	-0.12 -	0.67 (0.0001)	0.83 (0.0001)	-							
5	-0.49 (0.0002)	0.15 -	0.26 (0.05)	0.55 (0.0001)	-						
6	-0.63 (0.0001)	-0.27 -	-0.10 -	0.12 -	0.55 (0.0001)	-					
7	-0.69 (0.0001)	-0.51 (0.001)	-0.28 -	-0.12 -	0.33 (0.05)	0.87 (0.0001)	-				
8	-0.55 (0.001)	-0.50 (0.004)	-0.31 -	-0.08 -	0.29 -	0.75 (0.0001)	0.94 (0.0001)	-			
9	-0.41 (0.04)	-0.12 -	0.01 -	0.02 -	0.18 -	0.54 (0.005)	0.76 (0.0001)	0.92 (0.0001)	-		
10	-0.27 -	-0.14	-0.04 -	-0.12 -	-0.01 -	0.36 -	0.64 -	0.83 (0.0007)	0.97 (0.0001)	-	
11	-0.24 -	-0.22 -	-0.14 -	-0.24 -	-0.07 -	0.33 -	0.65 (0.0003)	0.79 (0.0001)	0.92 (0.0001	0.97 (0.0001)	-
12	-0.23 -	-0.24 -	-0.17 -	-0.27 -	-0.09 -	0.29 -	0.63 (0.003)	0.76 (0.0001)	0.89 (0.0001)	0.95 (0.0001)	0.99 (0.0001)

(C) Year	1	2	3	4	5	6	7	8	9	10	11
1	-										
2	0.71 (0.0001)	-									
3	0.51 (0.0001)	0.92 (0.0001)	-								
4	0.40 (0.002)	0.85 (0.0001)	0.97 (0.0001)	-							
5	0.45 (0.0006)	0.73 (0.0001)	0.89 (0.0001)	0.98 (0.0001)	-						
6	0.33 (0.03)	0.72 (0.0001)	0.86 (0.0001)	0.96 (0.0001)	0.99 (0.0001)	-					
7	0.19 -	0.55 (0.0006)	0.71 (0.0001)	0.89 (0.0001)	0.96 (0.0001)	0.98 (0.0001)	-				
8	0.16 -	0.49 (0.0005)	0.66 (0.0001)	0.88 (0.0001)	0.96 (0.0001)	0.98 (0.0001)	0.99 (0.0001)	-			
9	0.23 -	0.64 (0.0004)	0.74 (0.0001)	0.90 (0.0001)	0.96 (0.0001)	0.98 (0.0001)	0.99 (0.0001)	0.99 (0.0001)	-		
10	0.24 -	0.64 (0.0008)	0.74 (0.0001)	0.88 (0.0001)	0.94 (0.0001)	0.97 (0.0001)	0.97 (0.0001)	0.99 (0.0001)	0.99 (0.0001)	-	
11	0.19 -	0.68 (0.001)	0.74 (0.0003)	0.86 (0.0001)	0.91 (0.0001)	0.95 (0.0001)	0.96 (0.0001)	0.97 (0.0001)	0.99 (0.0001)	0.99 (0.0001)	-
12	0.20 -	0.67 (0.002)	0.72 (0.0004)	0.84 (0.0001)	0.89 (0.0001)	0.93 (0.0001)	0.95 (0.0001)	0.96 (0.0001)	0.98 (0.0001)	0.99 (0.0001)	0.99 (0.0001)

(D) Year	1	2	3	4	5	6	7	8	9	10	11
1											
2		-									
3		0.18	-								
4		-0.01 -	0.48 (0.0002)	-							
5		-0.34 (0.01)	0.41 (0.002)	0.59 (0.0001)	-						
6		0.06 -	0.37 (0.01)	0.23 -	0.35 (0.02)	-					
7		-0.03 -	0.51 (0.001)	0.25 -	0.43 (0.008)	0.53 (0.001)	-				
8		0.08 -	0.27 -	-0.09 -	0.23 -	0.33 -	0.47 (0.006)	-			
9		0.17 -	-0.04 -	-0.44 (0.02)	-0.04 -	-0.06 -	0.19 -	0.52 (0.005)	-		
10		0.03 -	0.09 -	-0.22 -	0.13 -	0.01 -	0.46 (0.02)	0.43 (0.03)	0.75 (0.0001)	-	
11		-0.12 -	0.24 -	0.24 -	0.49 (0.02)	0.22 -	0.66 (0.001)	0.19 -	0.36 -	0.69 (0.0006)	-
12		0.02 -	0.31 -	0.11 -	0.34 -	0.19 -	0.74 (0.0003)	0.29 -	0.37 -	0.73 (0.0004)	0.83 (0.0001)

(E) Year	1	2	3	4	5	6	7	8	9	10	11
1	-										
2	0.34 (0.008)	-									
3	0.02 -	0.75 (0.0001)	-								
4	-0.18 -	0.44 (0.0007)	0.70 (0.0001)	-							
5	0.02 -	-0.17 -	-0.07 -	0.41 (0.001)	-						
6	-0.12 -	-0.30 (0.04)	-0.18 -	0.18 -	0.46 (0.002)	-					
7	-0.13 -	-0.50 (0.002)	-0.25 -	-0.10 -	0.21 -	0.80 (0.0001)	-				
8	-0.04 -	-0.32 -	-0.15 -	0.01 -	0.29 -	0.74 (0.0001)	0.92 (0.0001)	-			
9	-0.25 -	0.07 -	0.25 -	0.11 -	0.08 -	0.36 (0.06)	0.63 (0.0005)	0.87 (0.0001)	-		
10	-0.07 -	0.13 -	0.23 -	-0.04 -	-0.09 -	0.05 -	0.41 (0.04)	0.70 (0.0001)	0.95 (0.0001)	-	
11	-0.01 -	0.12 -	0.08 -	-0.23 -	-0.15 -	0.01 -	0.41 (0.07)	0.63 (0.003)	0.86 (0.0001)	0.95 (0.0001)	-
12	-0.01 -	0.10 -	0.03 -	-0.26 -	-0.17 -	-0.01 -	0.42 (0.07)	0.61 (0.005)	0.82 (0.0001)	0.92 (0.0001)	0.99 (0.0001)

(F) Year	1	2	3	4	5	6	7	8	9	10	11
1											
2		-									
3		-0.01 -	-								
4		-0.23 -	0.26 (0.04)	-							
5		-0.55 (0.0001)	0.17 -	0.46 (0.0004)	-						
6		-0.01 -	0.20 -	0.02 -	-0.01 -	-					
7		0.05 -	0.44 (0.008)	-0.006 -	0.09 -	0.48 (0.004)	-				
8		0.29 -	-0.13 -	-0.53 (0.002)	-0.34 -	0.19 -	0.16 -	-			
9		0.48 (0.01)	-0.08 -	-0.57 (0.002)	-0.27 -	0.04 -	0.02 -	0.54 (0.004)	-		
10		0.38 -	-0.03 -	-0.34 -	-0.12 -	0.15 -	0.26 -	0.35 -	0.73 (0.0001)	-	
11		0.09 -	0.01 -	0.11 -	0.19 -	0.21 -	0.51 (0.002)	-0.12 -	0.23 -	0.62 (0.004)	-
12		0.31 -	-0.15 -	-0.11 -	-0.16 -	0.11 -	0.57 (0.01)	0.15 -	0.32 -	0.72 (0.0005)	0.68 (0.001)

3.2.3.4.2 SITE INDEX EQUATIONS

To estimate plot site index from A and H, two site index equations are available, the site index equation from the Ek-Payandeh method (Eq. [12]) and the modified difference equation (Eq. [19]). As the parameters of the algebraic difference equation were estimated for each cluster group, a specific site index equation was also available for each cluster group.

When predicting S most of the error incurred stemmed from stands of young age. Predicting S from stands of three years or less produced errors which were unacceptably high. When stands of three or less years were excluded Eqs. [12], [19] and Eq. [19] fitted to each cluster group, yielded acceptable residuals. Equation [19] fitted to each cluster group displayed the most acceptable residuals (Table 16).

Table 16: Univariate statistics for the residuals generated from individual site index equations.

Equation	Mean	Std. Dev.	Range	Skewness	Kurtosis
Eq. [12]	-0.07	1.74	-6.78 - 8.69	0.51	2.66
Eq. [12]	-0.20	1.26	-3.44 - 3.16	0.25	-1.10
Eq. [19]	0.13	2.24	-10.92 - 11.44	0.35	5.12
Eq. [19]^A^	-0.10	1.10	-4.87 - 3.58	-0.11	2.25
Eq. [19]*	0.10	1.77	-10.09 - 8.07	0.18	4.87
Eq. [19]*^A^	-0.08	0.96	-4.87 - 3.17	-0.42	3.45

* Eq. [19] fitted to each cluster group
A only stands older than 3 years considered.

3.2.4 DISCUSSION AND CONCLUSION

3.2.4.1 TOP HEIGHT AND SITE INDEX

The reference age in this study was arbitrarily selected to be 5 years, 50% of the rotation. In most published accounts the reference age varies between 50 and 100 years with 50 years at breast height a common reference age for the slower growing North American species (Carmean 1972; Monserud 1984, 1985; Smith 1984; Biging 1985). For fast growing pine species a reference age of 20 years is common (Garcia 1983; Grey 1989). For the more productive eucalypt species reference ages are younger still. For example, 10 years for *E. globulus* in Portugal (Tome 1988) and seven years for *E. globulus* in Rwana (Gasana and Loewenstein 1984).

In this study the variation in top height for any one year had stabilized by year five. For example, the coefficient of variation for years 1 to 12 was 49.3%, 37.6%, 30.9%, 26.5%, 22.3%, 23.3%, 24.4%, 19.6%, 19.1%, 18.2%, 19.9% and 20.6% respectively. Therefore, shortening the reference age to five years should not detract from the explanatory value of the index when predicting top height development. The logistic advantage gained by being able to assign a measured site index to a stand early in the rotation is also attractive.

3.2.4.2 PLOT SITE INDEX AND STAND DENSITY

The data used in this study were too few to adequately test the hypothesis that top height is unaffected by differences in stand density using the range of stand density estimators available (West 1982). This remains a weakness of this study.

Only stocking showed a relationship with site index and this was considered an artifact of management history. The scarcity of plantations with high stocking levels on poor sites prevented the use of a sampling strategy which may remove this relationship. The correlation between stocking and site index was assumed to be of little consequence given that PCC and basal area show no such relationship. Basal area and density measures based on basal area have been shown to be more valid density estimators by West (1982), although Bredenkamp and Burkhart (1990) found relative spacing useful.

3.2.4.3 TOP HEIGHT DEVELOPMENT EQUATIONS

3.2.4.3.1 THE EK-PAYANDEH METHOD

Although Eq. [11] yields acceptable model statistics, the predicted top height at the reference age did not coincide with the corresponding site index values (Table 17). Such results have been reported in other studies where this functional form has been applied (Kabzen 1971, Hahn and Carmean 1982). Hahn and Carmean (1982) lessened the magnitude of the differences by weighting during the parameter estimation procedure while Newnham (1988) presents a method where the equation is constrained to pass through the appropriate height at the index age. However, Newnham (1988) reported a loss of accuracy compared with the unconstrained method. In this study the magnitude of the differences are small enough to be of little practical significance and therefore remain uncorrected.

A further disadvantage of this methodology stems from the fact that Eq. [11] uses S to modify the Chapman-Richards functional form and therefore the methodology will not be invariant to the choice of reference age (Heger 1973). How the top height development curves generated by Eq. [11] differ with different choices of reference age was not determined in this study.

Table 17: Site index and the estimated top heights at age 5 years.

Site Index (m)	Estimated Top Height At Age 5 (m) Eq. [11]	 Eq. [14]
8	7.9	9.0
10	9.8	10.1
12	11.8	11.4
14	13.8	12.9
16	15.7	14.6
18	17.7	16.5
20	19.5	18.7

3.2.4.3.2 THE PARAMETER PREDICTION METHOD

The failure of the Chapman-Richards functional form to form the base of the parameter prediction method was unexpected. This functional form has performed well in other methodologies and has been used successfully by other authors with this method (Trousdell *et al.* 1974; Biging 1985; Smith and Watts 1987; Kerr and Bowling 1991). Even when the method of Smith and Watts (1987), where only the β_o parameter is predicted from a function of S, was applied to this data set, no success was attained.

The failure of this method is particularly surprising given the success of the Ek-Payandeh functional form. The Ek-Payandeh functional form is merely a modification of the Chapman-Richards functional form where the β_o and β_2 parameters of the Chapman-Richard functional form are allowed to vary with different values of S, i.e.:

$$\beta_0 = \Psi_0 S^{\Psi_1}$$

$$\beta_2 = \Psi_2 S^{\Psi_3}$$

where,

β_0, β_2 = *parameters of the Chapman – Richards functional form*,

$\Psi_0 \Psi_1 \Psi_2 \Psi_3$ = *parameters of the Ek – Payandeh functional form*.

This suggests that the parameters of the Chapman-Richards functional form may be expressed as functions of S using the parameter prediction method. Although this was shown to be true the resulting predictions were unrealistic and the functions of S used to predict the β_o and β_2 parameters of the Chapman-Richards functional form did not resemble what might be reasonably expected.

One possible explanation for the failure of this method is that the functional form lacked the flexibility to track the marked polymorphism evident in the data. Monserud (1984) also considered that the Chapman-Richards functional form lacked the flexibility to track polymorphism.

The correlation between the three parameters of the function made biological interpretation of the parameter values difficult. The individual parameter values derived from each plot, could vary markedly without this variance being associated with shifts in site conditions. Attempts to predict parameter values from site index are hampered by such multicollinearity.

Schumacher's functional form was more successful as a base function to this methodology. However, the resulting top height development equations suffer from a number of disadvantages. Firstly, the top height development curves are unrealistically asymptotic. Secondly, the predicted top height at the reference age does not equal site index (Table 17), with top height over estimated for small values of S and under estimated for larger values. These errors are of higher magnitude than those resulting from the application of Eq. [11] and are large enough to be of practical importance. A procedure for eliminating this problem has been described by Burkhart and Tennent (1977) but was not applied to Eqs. [14], [14a] and [14b].

Thirdly, the equations are not invariant to the choice of reference age, which has been discussed previously. A final disadvantage with this method is that given H and A, S must be estimated implicitly by an iterative procedure which is tedious and logistically unappealing.

Finally, the use of this method may yield equations which are over parameterized. Such would seem to be the case for Eqs. [14] and [15]. The problems resulting from multicollinearity and over parameterization are well documented (Hocking 1976; Verbyla 1986).

3.2.4.3.3 THE ALGEBRAIC DIFFERENCE METHOD

Top height development curves derived via algebraic differences (Eq. [18]) are less asymptotic than other sets of curves. The equation has the desirable properties of (a) the top height at the reference age equals site index, (b) each curve has a separate upper asymptote, (c) the curves are invariant with respect to the choice of reference age and (d) site index may be estimated from Eq. [18].

3.2.4.3.4 THE CLASSIFICATION OF PLOTS BY TOP HEIGHT DEVELOPMENT PATTERN

Numerical classification is a technique rich in forestry applications (Turner 1974) yet is mainly confined to that area of forestry science concerned with the grouping of lands with similar extrinsic attributes (Spies and Barnes 1985; Wardell-Johnson *et al.* 1989; Inions 1990; Inions *et al.* 1990). Although some examples of numerical classification techniques exist in the forest science literature for volume table construction (Postaire and M'Hirit 1985) and deriving productivity indices (Harding *et al.* 1985) most examples are confined to the field of community ecology (Orloci 1988).

Classifying plots on the basis of their top height development patterns may be viewed as a stratification strategy. Stratification yields more efficient, and therefore more precise, estimators of the model's parameters where the variables are homogeneous within a stratum but heterogeneous between strata (Golder and Yeoman 1973). To date stratification for deriving top height development curves has been on the basis of site index class (Carmean 1972) or some extrinsic environmental variable such as soil type (Carmean and Lenthall 1989). No system of top height development and site index equations has been developed where plots have first been stratified via cluster analysis on the basis of their top height development patterns.

Methods of cluster analysis are many (c.f. Booth 1978; Green 1980; Gauch and Whittaker 1981; Faith *et al.* 1987). The choice of association metric and fusion strategy is usually arbitrary and subject to the objective of the classification. In this study describing top height classification which explained more of the polymorphic nature of the data set than did the standard clustering technique. The two-dimensional profile attribute algorithm of Faith *et al.* (1985) has not been applied for forestry purposes, except for demonstration (Faith *et al.* 1985) and in this study. The success of the classification suggests that the algorithm would be applicable for stratifying other data sets involving time series data.

Although the technique was able to confine polymorphism to that which exists between groups, the resulting set of equations suffer from a number of difficulties. Firstly, as with all polymorphic nondisjoint top height development equations, the equation pertaining to the plot of interest must be known before the system can be applied. This renders the system described of little use unless some allocation criterion is available. This criterion will be described and discussed in Chapter four.

Secondly, where Eq. [20] is used, the predicted top height at the reference age may not equal site index. This problem does not occur when the algebraic difference equation is applied to each cluster.

Finally, Eq. [20] is not invariant with reference age. Again this problem is overcome when the algebraic difference equation is applied.

3.2.4.3.5 AUTOCORRELATION AMONG RESIDUALS

The use of ordinary least squares with repeated measures, as was done in this study, resulted in residuals which were significantly autocorrelated in every example. The autocorrelation was of larger magnitude with some equations than for others. Under such situations, hypothesis tests and variance estimates may be biased, as the standard assumption of independence is violated (Sullivan and Reynolds 1976; Monserud 1987). Such effects may be removed with the use of techniques such as fitting appropriate autoregressive moving average (ARMA) models (Monserud 1986; 1987; Yamaguchi 1986; Wigley *et al.* 1987). With the exception of Monserud (1984) these techniques are not being used for constructing top height development and site index equations and were not applied in this study.

Although variance estimation may be biased in the presence of autocorrelation, there is evidence to suggest the problem may be ignored when estimating parameters with ordinary least squares (Elston and Grizzle 1962; Sullivan and Clutter 1972; Monserud 1984). In this study the problem of autocorrelation was ignored, apart from merely quantifying its magnitude. Whether this course of action is justified will be examined during validation.

This study has used a number of methods to derive top height development and site index equations, however, no recommendations as to which are the most appropriate will be made until after model validation, which is the topic of Chapter five.

3.2.5 REFERENCES

Alemdag, I.S. 1988. Site index equations for white spruce in the Northwest Territories, Canada. *For. Ecol. Manage.*, 23:61-71.

Alexander, R.R., D. Tackle, and W.G. Dahns. 1967. Site indexes for lodgepole pine, with corrections for stand density: methodology. *USDA Res. Pap.* RM-29.

Bailey, R.L., 1980. The potential of Weibull-type functions as flexible growth curves: discussion. *Can. J. For. Res.*, 10:117-118.

Bailey, R.L., and J.L. Clutter. 1974. Base-age invariant polymorphic site curves. *For. Sci.*, 20:155-159.

Bard, Y. 1974. Nonlinear parameter estimation. Wiley, New York.

Barrett, J.W. 1973. Latest results from the Pringle Fall ponderosa pine spacing study. *USDA Res. Pap.* PNW-209.

Barrett, J.W. 1978. Height growth and site index curves for managed stands of ponderosa pine in the Pacific Northwest. *USDA Res. Pap.* PNW-232.

Beck, D.E., and K.B. Trousdell. 1973. Site index: accuracy of prediction. *USDA Res. Pap.* SE-108.

Bennett, F.A. 1975. Slash pine: some intriguing growth and yield relationships. *USDA Res. Pap.* SE-133.

Bennett, F.A., C.E. McGee, and J.L. Clutter. 1959. Yield of old-field slash pine plantations. *USDA Res. Pap.* SE-107.

Biging, G.S. 1985. Improved estimates of site index curves using a varing-parameter model. *For. Sci.*, 31:248-259.

Booth, T.H. 1978. Numerical classification techniques applied to forest tree distribution data. 1. A comparison of methods. *Aust. J. Ecol.*, 3:297-306.

Borders, B.E., R.L. Bailey, and M.L. Clutter. 1988. Forest growth models: parameter estimation using real growth series. *In*: Forest growth modelling and prediction. A.R. Ek *et al.* (eds.) *USDA Gen. Tech. Rep.* NC-120.

Borders, B.E., R.L. Bailey, and K.D. Ware. 1984. Slash pine site index from a polymorphic model by joining (splining) nonpolynomial segments with an algebraic difference method. *For. Sci.*, 30:411-423.

Bray, J.R., and J.T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.*, 27:325-349.

Bredenkamp, B.V., and H.E. Burkhart. 1990. An examination of spacing indices for *Eucalyptus grandis*. *Can. J. For. Res.*, 20:1909-1916.

Brickell, J.E. 1968. A method of constructing site index curves from measurements of tree age and height - its application to inland Douglas-fir. *USDA Res. Pap.* INT-47.

Bruce, D. 1926. A method of preparing timber yield tables. *J. Agric. Res.*, 32:543-557.

Burkhart, H.E., and R.B. Tennant. 1977. Site index equations for radiata pine in New Zealand. *N.Z.J. For. Sci.*, 7:408-416.

Callaway R.M., and E.C. Clebsch. 1987. A multivariate analysis of forest communities in the Western Great Smoky Mountains National Park. *Mid. Am. Nat.*, 118:107-120.

Carleton, T.J., R.K. Jones, and G. Pierpoint. 1985. The prediction of understorey vegetation by environmental factors for the purposes of site classification in forestry: an example from northern Ontario using residual ordination analysis. *Can. J. For. Res.*, 15:1099-1108.

Carmean, W.H. 1972. Site index curves for upland oaks in the central states. *For. Sci.*, 18: 109-120.

Carmean, W.H. 1975. Forest site quality evaluation in the United States. *Adv. Agron.*, 27:209-269.

Carmean, W.H., and J.T. Hahn. 1981. Revised site index curves for balsam fir and white spruce in the Lake States. *USDA Res. Note.* NC-269.

Carmean, W.H., and D.J. Lenthall. 1989. Height-growth and site-index curves for jack pine in north central Ontario. *Can. J. For. Res.*, 19:215-224.

Carron, L.T. 1968. An outline of forest mensuration with special reference to Australia. Australian National University Press, Canberra.

Chang, S.J. 1988. Consistent growth and yield prediction through integration of whole stand and diameter distribution models. *In*: Forest growth modelling and prediction. A.R. Ek *et al.* (eds.). Vol 1. *USDA Gen. Tech. Rep.* NC-120.

Chapman, D.G. 1961. Statistical problems in population dynamics. *Proc. Fourth Berkely Symposium on Mathematical Statistics and Probability.* Univ. Calif. Press, Berkely.

Chisman, H.M., and F.X. Schumacher. 1940. On the tree-area ratio and certain of its applications. *J. For.*, 38:311-317.

Cieszewski, C.J., and I.E. Bella. 1989. Polymorphic height and site index curves for lodgepole pine in Alberta. *Can. J. For. Res.*, 19:1151-1160.

Clutter, J.L., J.C. Fortson, L.V. Pienaar, G.H. Brister, and R.L. Bailey. 1983. Timber Management. A quantitive approach. John Wiley and Son, New York.

Clutter, J.L., and E.P. Jones Jr. 1980. Prediction of growth after thinning in old field slash pine plantations. *USDA Res. Pap.* SE-217.

Cromer, D., and P. Bowling. 1961. The development of a yield table for *Eucalyptus obliqua* regrowth in southern Tasmania. Paper presented to 2nd World Eucalypt Conference - Brazil.

Curry, P., and F.M. Slater. 1986. A classification of river corridor vegetation from four catchments in Wales. *J. Biogeography*, 13:119-132.

Curtis, R.O. 1964. A stem-analysis approach to site-index curves. *For. Sci.*, 10:241-156.

Curtis, R.O. 1970. Stand density measures: An interpretation. *For Sci.*, 16:403-414.

Curtis, R.O., D.J. DeMars, and F.R. Herman. 1974a. Which dependent variable in site index-height-age regressions? *For. Sci.*, 20:74-87.

Curtis, R.O., F.R. Herman, and D.J. DeMars. 1974b. Height growth and site index for Douglas-fir in high-elevation forests of the Oregon-Washington Cascades. *For. Sci.*, 20:307-316.

Dahms, W.G. 1975. Gross yield of central Oregon lodgepole pine. *In*: Management of lodgepole pine ecosystems symposium proceedings. Washington State Univ. Co-op.Ext.Serv.,Pullman.

Daubenmire, R. 1976. The use of vegetation in assessing the productivity of forest lands. *Bot. Rev.*, 42:115-143.

Devan, J.S., and H.E. Burkhart. 1982. Polymorphic site index equations for loblolly pine based on a segmented polynomial differential model. *For. Sci.*, 28:544-555.

Draper, N.R., and H. Smith. 1981. Applied regression analysis. 2nd ed. John Wiley and Sons, New York.

Drew, T.J., and J.W. Flewelling. 1979. Stand density management: an alternative approach and its application to Douglas-fir plantations. *For. Sci.*, 25:518-532.

Dyer, M.E., and R.L. Bailey. 1987. A test of six methods for estimating true heights from stem analysis data. *For. Sci.*, 33:3-13.

Economou, A. 1990. Growth intercept as an indicator of site quality for planted and natural stands of *Pinus nigra* var. *pallasiana* in Greece. *For. Ecol. Manage.*, 32:103-115.

Ek, A.R. 1971. A formula for white spruce site index curves. *Univ. Wis. For. Res. Notes* No 161.

Elston, R.C., and J.E. Grizzle. 1962. Estimation of time response curves and their confidence bounds. *Biometrics*, 18:148-159.

Faith, D.P., M.P. Austin, L. Belbin, and C.R. Margules. 1985. Numerical classifications of profile attributes in environmental studies. *J. Environ. Manage.*, 20:73-85.

Faith, D.P., P.R. Minchin, and L. Belbin. 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio*, 69:57-68.

Farr, W.A. 1984. Site index and height growth curves for unmanaged, even-aged stands of western hemlock and sitka spruce in south east Alaska. *USDA Res. Pap.* PNW-326.

Ferguson, I.S., and J.W. Leech. 1978. Generalized least squares estimation of yield functions. *For. Sci.*, 24:27-42.

Garcia, O. 1983. A stochastic differential equation model for the height growth of forest stands. *Biometrics*, 39:1059-1072.

Gasana, J.K., and H. Loewenstein. 1984. Site classification for Maiden's gum, (*Eucalyptus globulus* subsp. *maidenii*), in Rwanda. *For. Ecol. Manage.*, 8:107-116.

Gauch, H.G., and R.H. Whittaker. 1981. Hierarchical classification of community data. *J. Ecol.*, 69:537-557.

Gertner, G.Z. 1985. Efficient nonlinear growth model estimation: Its relationship to measurement interval. *For. Sci.*, 31:821-826.

Gingrich, S.F. 1967. Measuring and evaluating stocking and stand density in upland hardwood forests in the central states. *For. Sci.*, 13:38-53.

Golder, P.A., and K.A. Yoemans. 1973. The use of cluster analysis for stratification. *Applied statistics*, 22:213-219.

Gower, J.C. 1967. Multivariate analysis and multidimensional geometry. *Statist.*, 17:13-28.

Gower, J.C. 1971. A general coefficient of similarity and some of its properties. *Biometrics*, 27:85-871.

Green, R.H. 1980. Multivariate approaches in ecology: The assessment of ecological similarity. *Am. Rev. Ecol. Syst.*, 11:1-14.

Grey, D.C. 1989. Site Index - a review. *S.A. For. J.*, 148:28-32.

Grey, H.R. 1945. Site classification of coniferous plantations. *Aust. For.*, 9:9-17.

Hagglund, B. 1981. Evaluation of forest site productivity. *Forestry Abstracts*, 42:517-527.

Hahn, J.T., and W.H. Carmean. 1982. Lake states site index curves formulated. *USDA Gen. Tech. Rep.* NC-188.

Hahn, J.A., and R.A. Leary. 1979. Potential diameter growth functions. *In*: A generalized forest growth projection system applied to the Lake States region. *USDA Gen. Tech. Rep.* NC-79.

Hall, M.J. 1979. The effect of stocking density, site quality and age on the volume of pulpwood produced per acre. *Appita*, 13:22-29.

Hamlin, D., and R. Leary 1987. Methods for using an integro-differential equation as a model of tree height growth. *Can. J. For Res.*, 17:353-358.

Harding, R.B., D.F. Grigal, and E.H. White. 1985. Site quality evaluation for white spruce plantations using discriminant analysis. *Soil. Sci. Soc. Am. J.*, 49:229-232.

Havel, J.J. 1968. The potential of the Swan Coastal Plain for *Pinus pinaster* Ait. Plantations. *Bull. For. Dept. West. Aust.* No 76.

Havel, J.J. 1975a. Site vegetation mapping in the northern jarrah forest (Darling Range). I. Definition of site-vegetation types. *Bull., For. Dept. West. Aust.* No 86.

Havel, J.J. 1975b. Site-vegetation mapping in the northern jarrah forest (Darling Range). II. Location and mapping of site-vegetation types. *Bull. For. Dept. West. Aust.* No. 87.

Havel, J.J. 1980a. Application of fundamental synecological knowledge to practical problems in forest management. I. Theory and methods. *For. Ecol. Manage.*, 3:1-29.

Havel, J.J. 1980b. Application of fundamental synecological knowledge to practical problems in forest management. II. Application. *For. Ecol. Manage.*, 3:81-111.

Heger, L. 1968. A method of constructing site-index curves from stem analysis. *For. Chron.*, 44:11-15.

Heger, L. 1973. Effect of index age on the precision of site index. *Can. J. For. Res.*, 3:1-6.

Herman, F.R., R.O. Curtis, and D.J. DeMars. 1978. Height growth and site index estimates for noble fir in high-elevation forests of the Oregon-Washington Cascades. *USDA Res. Pap.* PNW-243.

Hill, A.V. 1913. The combinations of haemoglobin with oxygen and carbon monoxide. *Biochem. J.*, 7:471-480.

Hocking, R.R. 1976. The analysis and selection of variables in linear regression. *Biometrics*, 32:1-50.

Hubert, L.J., R.G. Golledge, and C.M. Costanzo. 1981. Generalized procedures for evaluating spatial autocorrelation. *Geogr. Anal.*, 13:224-233.

Husch, B. 1963. Forest mensuration and statistics. Ronald Press Co. New York.

Huseyin, K. 1986. Differential and difference equations through computer experiments. Springer-Verlag. New York.

Inions, G.B. 1990. Classification and evaluation of sites in karri (*Eucalyptus diversicolor* F.Muell) regeneration. I. Edaphic and climatic attributes. *For. Ecol. Manage.*, 32:117-134.

Inions, G.B., Wardell-Johnson, G. and Annels A.A. 1990. Classification and evaluation of sites in karri (*Eucalyptus diversicolor* F.Muell) regeneration II. Floristic attributes. *For. Ecol. Manage.*, 32:135-154.

Jones, J.R. 1969. Review and comparison of site evaluation methods. *USDA Res. Pap.* RM-51.

Johnson, D.R., and R.T. Bradley. 1963. Forest management tables. *Common W. For. Rev.*, 42:217-227.

Kabzems, A. 1971. The growth and yield of well stocked white spruce in the mixed wood section of Saskatchewan. *Sask. Dep. Nat. Resour. For. Branch Tech. Bull. No.* 5.

Kerr, M.F., and C. Bowling, 1991. Polymorphic site index equations for four New Brunswick softwood species. *Can.J.For.Res.*, 21:728-732.

Kilian, W. 1984. Site classification and mapping - principles and trends. Proc. symposium on site and productivity of fast growing plantations. Pretoria, South Africa.

King, J.E. 1966. Site index curves for Douglas-fir in the Pacific Northwest. *Weyerhaeuser Forestry paper*. No 8.

Lappi, L., and R.L. Bailey. 1988. A height prediction model with random stand and tree parameters: an alternative to traditional site index methods. *For. Sci.*, 34:907-927.

Leary, R. and D. Hamlin. 1988. Methods for localizing regional growth models formed as second order differential equations. *In*: Forest growth modelling and prediction, A.R. Ek, *et al.* (eds.). *USDA Gen. Tech. Rep.* NC-120.

Lenhart, D.J. 1972. An alternative procedure for improving height/age data from stem analysis. *For. Sci.*, 18:332.

Lewis, N.B., A. Keeves, and J.W. Leech. 1976. Yield regulation in South Australian *Pinus radiata* plantations. *Woods and Forests Department*, Adelaide. S.A. Bulletin 23.

Loetsch, F., F. Zohrer, and K.E. Haller. 1973. Forest inventory, Vol. II. BLV Verlagsgesellschaft, Munchen.

Mader, D.R. 1976. Soil-site productivity for natural stands of white pine in Massachusetts. *Soil Sci. Soc. Am. J.*, 40:112-115.

Marquardt, D.W. 1963. An algorithm for least squares estimation of nonlinear parameters. *J. Soc. Industrial and Applied Mathematics*, 11:431-441.

McLeod, S.D., and S.W. Running. 1988. Comparing site quality indices and productivity in ponderosa stands of western Montana. *Can. J. For. Res.*, 18:346-352.

Michaelis, L., and M.L. Menten. 1913. Die kinetik der Invertinwirkung. *Biochem. J.*, 49:333-369.

Monserud, R.A. 1984. Height growth and site index curves of inland Douglas-fir based on stem analysis data and forest habitat type. *For. Sci.*, 30:943-965.

Monserud, R.A. 1985. Comparison of Douglas-fir site index and height growth curves in the Pacific Northwest. *Can. J. For. Res.*, 15:673-679.

Monserud, R.A. 1986. Time-series analysis of tree-ring chronologies. *For. Sci.*, 32:349-371.

Monserud, R.A. 1987. Time series analysis applied to tree-ring series. Proc. Int. Sym. on ecological aspects of tree-ring analysis. G.C. Jacoby *et al.* (eds.) 1986, August; Tarrytown, NY.

Monserud, R.A. 1988. Variations on a theme of site index. *In*: Forest growth modelling and prediction. A.R. Ek, *et al.* (eds.). *USDA Gen. Tech. Rep.* NC-120.

Morgan, P.H., L.P., Mercer., and N.W. Flodin. 1975. General model for nutritional responses of higher organisms. *Proc. Nat. Acad. Sci. USA.*, 72:4327-4331.

Newberry, J.D., and L.V. Pienaar. 1978. Dominant height growth models and site index curves for site-prepared slash pine plantations in the lower coastal plain of Georgia and North Florida. *Univ. of Ga. Plantation Mgt. Res. Coop. Res. Pap.* No 4.

Newnham, R.M. 1988. A modification of the Ek-Payendeh nonlinear regression model for site index curves. *Can. J. For. Res.*, 18:115-120.

Oliver, F.R. 1966. Aspects of maximum likelihood estimation of the logistic growth function. *J. Am. Statist. Assoc.*, 61:697-705.

Orloci, L. 1988. Community organization: recent advances in numerical methods. *Can. J. Bot.*, 66:2626-2633.

Payandeh, B. 1974a. Formulated site index curves for major timber species in Ontario. *For. Sci.*, 20:143-144.

Payandeh, B. 1974b. Nonlinear site index equations for several major Canadian timber species. *For. Chron.*, 50:194-196.

Payandeh, B. 1983. Some applications of nonlinear regression models in forestry research. *For. Chron.*, 59:244-248.

Payandeh, B. 1988. An assessment of the structural method of deriving a black spruce site equation: discussion. *Can. J. For. Res.*, 18:1351.

Payandeh, B., and A.R. Ek. 1986. Distance methods and density estimators. *Can. J. For. Res.*, 16:918-924.

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. *For. Sci.*, 19:2-22.

Pienaar, L.V., and B.D. Shiver. 1984. The effect of planting density on dominant height in unthinned slash pine plantations. *For. Sci.*, 30:1059-1066.

Pizer, S.M. 1975. Numerical computing and mathematical analysis. Chicago: Science Research Associates, Inc., Chapter 4.

Pojar, J., K. Klinka, and D.V. Meidinger. 1987. Bioclimatic ecosystem classification in British Columbia. *For. Ecol. Manage.*, 22:119-154.

Postaire, J., and O. M'Hirit. 1985. Application of pattern recognition to volume estimation in forest inventory. *For. Sci.*, 31:53-65.

Ralston, M.L., and R.I. Jennrich. 1979. DUD a derivative-free algorithm for nonlinear least squares. *Technometrics*, 1:7-14.

Ramirez-Maldonado, H., R.L. Bailey, and B.E. Borders. 1988. Some implications of the algebraic difference approach for developing growth models. *In*: Forest growth modelling and prediction. A.R. Ek, *et al.* (eds.) *USDA Gen. Tech. Rep.* NC-120.

Ratkowsky, D.A. 1983. Nonlinear regression modelling; A unified practical approach. Marcel Dekker, inc. New York.

Rayner, M.E. 1991. Site and dominant height growth curves for regrowth karri (*Eucalyptus diversicolor* F Muel.) in south west Western Australia. *For. Ecol. Manage.*, 44: 261-285.

Reed, D.D., and E.A. Jones. 1989. Assessing multiple estimates of plantation productivity. *Can. J. For. Res.*, 19:948-954.

Rennolds, K. 1978. "Top height"; its definition and estimation. *Commonw. For. Rev.*, 57:215-219.

Richards, F.J. 1959. A flexible growth function for empirical use. *J. Exp. Bot.*, 10:290-300.

Schumacher, F.X. 1939. A new growth curve and its application to timber yield studies. *J. For.*, 37:819-820.

Smalley, G.W. 1984. Classification and evaluation of forest sites in Cumberland Mountains. *USDA Gen. Tech. Rep.* SO-50.

Smith, J.H.G., and A. Kozak. 1984. New non-linear models can improve estimates of growth and yield. *Commonw. For. Rev.*, 63:41-45.

Smith, V.G. 1984. Asymptotic site-index curves, fact or artifact? *For. Chron.*, 60:150-156.

Smith, V.G. 1988. An assessment of the structural method of deriving a black spruce site equation: Reply. *Can. J. For. Res.*, 18:1351-1352.

Smith, V.G., and Watts, M. 1987. An assessment of the structural method of deriving a black spruce site equation. *Can. J. For. Res.*, 17:1181-1189.

Spies, T.A., and B.V. Barnes. 1985. A multifactor ecological classification of the northern hardwood and conifer ecosystems of Sylvania Recreation Area, Upper Peninsula, Michigan. *Can. J. For. Res.*, 15:949-960.

Spurr, S.H. 1952. Forest inventory. Ronald Press. Co. New York.

Stout, B.B., and D.L. Shumway. 1982. Site quality estimation using height and diameter. *For. Sci.*, 28:639-645.

Strand, L. 1964. Numerical constructions of site-index curves. *For. Sci.*, 10:410-414.

Strelein, G.J. 1988. Site classification in the southern jarrah forest of Western Australia. *Dept. Cons. Land Manage. Res. Bull.* No. 2.

Sullivan, A.D., and J.L. Clutter. 1972. A simultaneous growth and yield model for loblolly pine. *For. Sci.*, 18:76-86.

Sullivan, A.D., and M.R. Reynolds. 1976. Regression problems from repeated measurements. *For. Sci.*, 22:382-385.

Tesch, S.D. 1981. The evolution of forest yield determination and site classification. *For. Ecol. Manage.*, 3:169-182.

Tome, M. 1988. A growth and yield model for Eucalypt plantations in Portugal. *In:* Forest growth modelling and prediction, A.R. Ek, *et al.* (eds.). *USDA Gen. Tech. Rep.* NC-120.

Trousdell, K.B., D.E. Beck, and F.T. Lloyd. 1974. Site index for loblolly pine in the Atlantic coastal plain of the Carolinas and Virginia. *USDA Res.Pap.* SE-115.

Turner, B.J. 1974. Applications of cluster analysis in natural resource research. *For. Sci.*, 20:343-349.

Verbyla, D.A. 1986. Potential prediction bias in regression and discriminant analysis. *Can. J. For. Res.*, 16:1255-1257.

Wardell-Johnson, G., G.B. Inions, and A. Annels. 1989. A floristic classification of the Walpole-Nornalup National Park, Western Australia. *For. Ecol. Manage.*, 28:259-279.

Webb, L.J., J.G. Tracey, and W.T. Williams. 1984. A floristic framework of Australian rainforests. *Aust. J. Ecol.*, 9:169-198.

West, P.W. 1982. Comparison of stand density measures in even-aged regrowth eucalypt forest of southern Tasmania. *Can. J. For. Res.*, 13:22-31.

Wigley, T.M.L., D.D. Jones, and K.R. Briffa. 1987. Cross-dating methods in dendrochronology. *J. Arch. Sci.*, 14:51-64.

Yarnagucki, D.A. 1986. Interpretation of cross correlation between tree-ring series. *Tree-Ring Bulletin*, 46:47-54.

Yang, R.C., A. Kozak, and J.H.G. Smith. 1978. The potential of Weibull-type functions as flexible growth curves. *Can. J. For. Res.*, 8:424-431.

Yarie, J., and K. Van Cleve. 1983. Biomass and productivity of white spruce stands in interior Alaska. *Can. J. For. Res.*, 13:767-772.

Zahner, R. 1962. Loblolly pine site index curves by soil groups. *For. Sci.*, 8:104-110.

Zakrzewski, W. 1986. Empiric model of height increase of spruce-pine stands in north eastern Poland. *Ann. Warsaw Agric. Univ. SGGW-AR For. Wood Technol.* No. 34.

Zeide, B. 1978. Standardization of growth curves. *J. For.*, 76:289-295.

CHAPTER FOUR

Relationships between Site Productivity and Environmental Attributes for *E. globulus* in South West Western Australia.

4.1 LITERATURE REVIEW: EVALUATION OF SITE QUALITY WITH ENVIRONMENTAL ATTRIBUTES.

4.1.1 INTRODUCTION

It is the aim of this review to summarise the literature concerned with the relationships between environmental attributes and some measure of stand productivity. Scant attention is paid to the body of literature which covers the classification of land, based on soil, climatic, topographical or floristic attributes, which may or may not be related to some measure of the productive potential of the classification unit (Inions 1990; Inions *et al.* 1990). The literature for this review is derived primarily from the last decade and has a bias toward the Anglo-American and Western European sources.

Direct measurement of site quality is usually obtained from site index. Where a stand is degraded, is too young or has been heavily cut, few suitable trees will be available for the required height and age measurements. Consequently, many studies have sought to predict site quality from functions of environmental attributes.

Most attempts to relate environmental attributes to forest productivity have relied on expressing site index as a linear function of soil and topographic variables (Carmean 1975). Other attributes less commonly used are those pertaining to the climate of the study area. Climatic variables are usually used in combination with soil and/or topographic variables (Blyth and MacLeod 1981; White 1982a; Hunter and Gibson 1984) and less commonly in isolation (Christie and Lines 1979; Farr and Harris 1979). With some exceptions topographic variables are also rarely used in isolation (Stage 1976; Verbyla and Fisher 1989).

Floristic attributes may be used in functions to express productivity (McLean and Bolsinger 1973; Corns and Pluth 1984). However, the common practice is to identify sites of similar floristic composition, indicative of a comparatively homogeneous physical environment (Daubenmire and Daubenmire 1968; Layser 1974; Havel 1975; Pfister and Arno 1980; Inions *et al.* 1990). The derived floristic communities may then be used directly for productivity ratings (Havel 1968; Inions *et al.* 1990). Literature pertaining to synecological theory, methodology and application to forest management has been reviewed by Carmean (1975), Daubenmire (1976a) and Havel (1981a,b) and will not be dealt with further in this review.

4.1.2 CLASSES OF ENVIRONMENTAL VARIABLES USED FOR SITE QUALITY EVALUATION

4.1.2.1 EDAPHIC VARIABLES

Edaphic variables may be classified as chemical, generally expressing the soil nutrient status of a site or physical, generally expressing the moisture holding status of a site. The two classes of edaphic variable are usually used in the same function. For example, Brown and Loewenstein (1978) developed a model to predict site index of mixed conifer stands in northern Idaho, U.S.A., and found that 70% of the variation in site index was explained by a function incorporating extractable Ca, exchange acidity, cation exchange capacity, organic matter, total N, soil to rock ratio and clay content of the soils. Likewise, Munn and Vimmerstedt (1980) found that A horizon thickness, depth to B2 horizon, depth to restriction, pH, $kgha^{-1}$ of available P and Mn, $kgha^{-1}$ of exchangeable Ca, Mg and K, % base saturation and soil organic matter were significant variables in a function to predict the site index of yellow-poplar (*Liriodendron tulipifera* L.) in Ohio, U.S.A. Similar approaches have been reported for black spruce (*Picea mariana* Mill. B.S.P.) and balsam fir (*Abies basamea* L. Mill.) in Newfoundland, Canada (Page 1976), white pine (*Pinus strobus* L.) in Massachusetts, U.S.A. (Mader 1976), jack pine (*Pinus banksiana* Lamb.) in central Ontario, Canada (Schmit and Carmean 1988), blue gum (*Eucalyptus globulus* Labill.) in Rwanda (Gasana and Loewenstein 1984), forest trees of the Black Forest, West Germany (Stahr 1979) and *Eucalyptus camaldulensis* Dehnh. plantations in Nigeria (Buckley 1988).

Soil chemistry is difficult to include in standard forest inventory and consequently, many studies attempt to exclude variables requiring laboratory analysis. Fralish and Loucks (1975) found that the precision of an equation to predict the site index of aspen (*Populus tremuloides* Michx.) in Wisconsin, U.S.A., originally comprised of variables measurable in the field, was only slightly increased by adding soil chemical variables such as Mg, Ca, K, and P. On the other hand Mader (1976) found that equations comprised of easily determined soil physical factors gave fair results but the best results required the inclusion of more complicated physical and chemical parameters. For *Pinus radiata* D.Don stands in New South Wales, Australia, 88% of the variation in site index was explained by a function comprised of total soil P and exchangeable Ca only (Truman *et al.* 1983). In contrast, Turvey *et al.* (1986) concludes that soil physical parameters were predominant in discriminating between volume production classes while soil chemical parameters were predominant in discriminating between geological groups for *P. radiata* also in New South Wales, Australia. The geological soil groups were then used to separate growth patterns of *P. radiata* (Ryan 1986). On the other hand Grey (1979) found that edaphic variables have very poor predictive qualities in functions where the site index of *Pinus patula* Schlecht and Deppe in South Africa was the dependent variable.

The depth at which a soil sample is taken has considerable influence on the precision of derived relationships between forest productivity and soil, as the concentration of nutrients varies vertically through the profile. Powers (1980) found the concentration of mineralizable N decreased exponentially with soil depth, under planted stands of *Pinus ponderosa* Laws. in California and Oregon, USA. Similar results occur in Australian forest soils (Charley 1981; Inions 1990). Less certain is how the variance of a nutrient concentrations differs with depth through a profile. While Powers (1980) and Inions (1990) found that the variance of nutrient concentrations decreases with depth, Shumway and Atkinson (1978) found the coefficient of variation for mineralizable soil N increased with depth and therefore recommended that sampling be confined to the top 0 to 15cm of the profile.

A number of studies have examined the spatial distribution of root systems. The results are consistent and show that root concentrations decline sharply with soil depth (Moir and Bachelard 1969; Roberts 1976; Squires *et al.* 1978; Powers 1980). The roots are most concentrated in the surface 25cm presumably where nutrient uptake is also high. Therefore the vertical position from which a soil sample is taken is of considerable importance.

Soil samples are usually taken from a specified horizon (Munn and Vimmerstedt 1980; Hunter and Gibson 1984) or depth (Saunders *et al.* 1984; Turvey *et al.* 1986). Likewise, multiple samples from the one site may be from specified horizons or depths. For example, Mader (1976), Brown and Leowenstein (1978), and Schmidt and Carmean (1988) sampled two, three and four horizons, respectively. Page (1976) sampled from three set depths confined to the upper 30cm of the soil profile. Powers (1980) took five evenly-spaced samples to a depth of 95cm. Truman *et al.* (1983) took two samples from set depths from the upper 37.5cm of the profile, while White (1982a) took two samples from the upper 25cm.

Rarely are explanations offered for selection of horizons or depths as the sampling criteria. One exception is Page (1976) who explains that soil samples were collected from set depths above 30 cm, as below that point profile development was usually minimal and conifer roots were scarce. Also, fixed depth sampling was used to ensure an equal number of observations from each plot for statistical expediency. In an attempt to address the problem Monserud *et al* . (1990) analysed their soil properties by classifying each profile five ways: (i) by A, B and C horizons; (ii) by position of the horizon in the profile (ie., uppermost, second, third); (iii) by four soil depth classes; (iv) by plot coverage weighted by horizon thickness; and (v) by plot average above a potentially limiting depth. The limiting depth was the depth to the horizon with a bulk density of 1.7g cm $^{-3}$ and higher.

Most studies, where edaphic attributes are derived from multiple samples from the same profile, treat the samples as separate independent variables when deriving relationships with forest productivity. For example, Schmidt and Carmean (1988) sampled from four horizons to find only the pH of the BC horizon was significantly related to the site index of jack pine on glacial lacustrine soil in Canada. Broadfoot (1969) found that a number of different sampling depths were influential in equations predicting the site index of a number of different species.

Few studies were found where the multiple samples were treated dependently in forms such as ratios between depths or horizons and this remains a gap in the current knowledge. One notable exception is Fralish and Loucks (1975) who summed nutrient contents to depths of 30 cm, 60 cm, 90 cm and 150 cm. Another is Monserud *et al*. (1990) who use plot averages derived by weighting values by horizon thickness.

Many edaphic attributes are temporally dynamic, a fact rarely addressed in the forest productivity - soil literature. Most authors imply the variables are invariant with time by not addressing the unwanted source of variance. However, a wide range of temporal variability among nutrients, in soils of a variety of forest ecosystems, has been reported (Usher 1970; Weaver and Forcella 1979; Haines and Cleveland 1981; Vance and Henderson 1984). Temporal variability may be plant induced (Auten 1945a; Munn and Vimmerstedt 1980), occur with seasons (Powers 1980), flood induced (Peterson and Rolfe 1985), or occur as a result of logging (Albert and Barnes 1987). Broadfoot (1969) considered that the dynamic nature of the interrelationships between tree and soil and the possible effects of a tree induced change in soil characteristics, as possible reasons that equations he developed did not predict site index of new populations with sufficient precision. On the other hand, Malcolm (1970) found no cyclical change with age in the depth of litter and fermentation horizons in a study of site factors affecting Sitka spruce (*Picea sitchensis* (Bong.) Carr.) growth in Scotland. Will and Ballard (1976) concluded that over a range of soil types, changes in soil characteristics under *Pinus radiata* are inconsistent in magnitude and direction. A view supported by Hunter and Gibson (1984).

Where considered, the approaches to dealing with temporal variation of chemical attributes differ widely. For example, White (1982a) studying Scots pine (*Pinus sylvestris* L.) in Great Britain, simply states that all variables used in his study are assumed to be invariant with time. Munn and Vimmerstedt (1980) studying yellow-poplar (*Liriodendron tulipifera* L.) in Ohio, U.S.A. state that because of temporal variation of independent variables the models they developed are apt to underestimate sites not currently supporting yellow poplar. As such they warn that their models are limited in application. Auten (1945a) identified those variables subject to temporal variation before developing predictive models (Auten 1945b). Page (1976) adjusted edaphic attributes to correct for the effect of cover type and of stands at different stages of development, on soil properties. Adjustments were based on defined trends in soil properties through time, under stands of different types, available from previous studies (Page 1974). The magnitude of temporal variation will differ with each edaphic variable under consideration. For example, samples collected from identical locations at five sites in the autumn of back-to-back drought years, showed average differences of 89% for KCl- extractable N and about 29% for mineralizable N (Munn and Vimmerstedt 1980).

As with temporal variation spatial heterogenity of soils provides an unexplained source of variance unless it is accounted for when sampling. Many studies have examined the heterogenity of soil chemical properties (Usher 1970; Blyth and MacLeod 1978; Keeney 1980) which is known to occur laterally and vertically (Charley 1981). For example, Usher (1970) took 1152 samples of soil under Scots pine forest in Scotland to find that total N was randomly distributed both horizontally and vertically. However, few soil site studies take quantitative account of this unwanted source of variation.

The magnitude of within-site variation differs with the attribute being sampled for. Inions (1990) expressed the within-site variance as a proportion of the between-site variance for 15 soil chemical attributes sampled from karri (*Eucalyptus diversicolor* F. Muell) forest soils in Western Australia. The proportion was highest for total N and Olsen extractable P (Watanable and Olsen 1965) and least for total P and K. For forestry plots of 0.01 ha in Scotland, the number of samples necessary to obtain mean values to within 10% at $p>0.05$ was found to be six samples for total N, nine for total P and 29 for acetic-acid extractable P (Blyth and MacLeod 1978). Gessel *et al.* (1973) found that evaluation of the amount of N in forest soils required more intensive sampling than for other properties in the Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) forests they studied: 25 cores were required to obtain <10% difference from the true mean. Even for the same element, the magnitude of the within-site variance will differ with the way the element is extracted from the soil. For example, the number of samples needed to estimate the population mean for mineralisable N was much higher than the number required to estimate total N (McNabb *et al.* 1986).

Spatial heterogenity may be correlated with microsite differences such as the position of understory plants (Charley 1981) but this is not universal (McNabb *et al.* 1986). The magnitude of spatial heterogenity also varies with geographic area. For example, McNabb *et al.* (1986) found the variability of N and C in surface soils was not uniform between six forest types in the Oregon Cascades, USA.

Few studies have identified the number of samples required to accurately predict soil chemical values prior to deriving functions which predict site productivity. One notable exception is Blyth and MacLeod (1981) who use the sampling intensity recommended by Blyth and MacLeod (1978) prior to predicting the yield of Sitka spruce (*Picea sitchensis* (Bong.) Carr.) in Scotland from environmental attributes. The large number of samples required per plot limits the practical applicability of this approach.

Many studies choose to avoid the use of soil chemical variables because of the spatial heterogeneity (Shoulders and Tiarks 1980) while others are prepared to accept the increased variance and collect one sample only from each site (Fralish and Loucks 1975; Page 1976). However, the most common strategy is to make a composite sample from two or more sub samples. For example, Hamilton and Krause (1985) pool 12 soil samples to form one composite sample per plot, Harding *et al.* (1985) dug one soil pit per site but formed a composite sample by pooling soil from the four sides of the pit while Schmoldt *et al.* (1985) took a single sample from each of four pits per site to form their composite sample. The precision of estimation of soil chemical variables is increased using composite samples, however no indication of the variance is obtained.

A less common strategy is to use the mean value, from a number of samples within a plot. For example, Auchmoody and Clay-Smith (1979) use the mean value of three soil samples per plot. Kabzems and Klinka (1987a) collected 15 samples per site to create three composite samples, each of 5 samples, and used the mean value of the composites to represent the site value. The use of means is of limited practical applicability because of the restrictive fiscal and temporal cost, however some indication of the variability of the parameters under investigation is obtained.

Analytical procedures normally employed to derive estimates of the nutrient status of soils, differ in their effectiveness in measuring nutrient quantities available to the plant. For example, mineralizable soil N has been expressed seven ways in a study to examine the utility of various indices of mineralizable N for predicting the response to N fertilisation by loblolly pine (*Pinus taeda* L.) in the south eastern United States (Lea and Ballard 1982). All differed in their effectiveness. The different expressions for soil N are reviewed by Keeney (1980) and will not be discussed further.

Many studies express one element in a variety of ways and treat each expression as independent variables in subsequent analysis (Blyth and MacLeod 1981a; White 1982a; Saunders *et al.* 1984). Expressions may be in standard analytical nomenclature such as parts per million or percent (Charley 1981) or as a weight per area or volume of soil (Munn and Vimmerstedt 1980; Monserud *et al.* 1990). Whether a particular expression is significantly related to a growth parameter is dependent on the geographic locality for any one species (Hunter and Gibson 1984; Saunders *et al.* 1984; Turvey *et al.* 1986; Schmidt and Carmean 1988) or the tree species for any one geographic locality (Broadfoot 1969; Carmean 1975).

4.1.2.2 CLIMATIC VARIABLES

Climate and its processes form a continuum in space and time. Within the wide range of atmospheric conditions that occur, an infinite variety and combinations of attributes can be derived to characterise the climate, or the components thereof, for a geographic locality. Many recent studies classify large geographic areas into zones of comparable climate or homoclimes. These classifications are based on an array of many climatic attributes (Miller and Auclair 1974; Farr and Hard 1987; Inions 1990). Homoclimes may then be related to the distribution of forest communities (Newnham 1968; van Groenewoud 1984) or the productive capability of a forest community (Inions 1990).

The major disadvantage of the classification approach is the difficulty in obtaining the climatic attributes required. Also, the availability of such data rarely corresponds to forest inventory plots from which productivity data are derived, forcing the necessity for extrapolation if yield and climate are to be compared. Although models (Running *et al.* 1987) and methodologies (White 1979; Hutchinson and Bischof 1983) have been developed to enhance extrapolation, they are usually beyond the practical bounds of forest management. One common approach to overcoming such problems is to restrict studies to one geographic area or homoclime. This strategy has been employed by a number of studies which concentrate on edaphic and topographic variables (Page 1976; Buckley 1988; Schmidt and Carmean 1988). Other studies which seek to relate environment and forest quality across two or more homoclimes often incorporate simple expressions for climate into predictive equations.

One common way of describing climate is through geographic locality. Variables such as latitude, longitude and elevation have proven effective for predicting forest quality (Evans 1974; Farr and Harris 1979; Grey 1979). Including such variables into equations comprised of edaphic and topographical variables, often improves the precision of the prediction (Cook *et al.* 1977), although this is not a universal result (Fourt *et al.* 1971).

4.1.2.2.1 PRECIPITATION

Mean annual rainfall is a variable commonly used for predicting forest quality (Cook *et al.* 1977; Schiller 1982; Hunter and Gibson 1984). Although not always significantly related to forest quality (Cook *et al.* 1977) it is easily obtained from maps showing isohytes. Rainfall distribution is also used, for example, in a dendroclimatological study of loblolly pine in south eastern U.S.A., Chang and Aguilar (1980) examined the relationship between monthly, seasonal and annual precipitation and annual radial growth. Growth was positively related to the total rainfall during the previous summer. Similarly, the heights of the dominant and codominant strata of 20 year old loblolly, slash (*Pinus elliottii* Engelm. var. *elliottii*), longleaf (*Pinus palustris* Mill) and shortleaf pines (*Pinus echinata* Mill.) in the Gulf Coastal Plain of the U.S.A. were related to the rainfall during the warm (April-September) and cool (October-March) seasons (Shoulders and Tiarks 1980). Floristic communities within the karri forests of south west Western Australia, which differ in productive potential, have been separated along rainfall gradients, particularly precipitation during the summer months (December - February) (Inions *et al.* 1990). Likewise, Churchill (1968) found a close relationship between the prehistorical distribution of karri and the rainfall of the wettest and driest months of the year. Similar results have been obtained in Rwanda, where the rainfall in December was a major discriminant between site index classes of *E. globulus*. (Gasana and Loewenstein 1984).

Rainfall distribution is usually expressed as the total for a set of months, usually corresponding to a particular season. Rarely is the distribution throughout a year taken into account. Jackson and Gifford (1974) used a range of polynomial coefficients to provide differential weighting for the rainfall in each month to form a composite variable reflecting annual rainfall distributions. The analysis found that both mean annual precipitation and some of the composite seasonal rainfall distribution variables were significantly related to the periodic volume increment of *Pinus radiata* in New Zealand.

The number of days on which precipitation occurs has been used as an expression of climate. Chang and Aguilar (1980) accumulate the number of days on which precipitation exceeded 0.25mm on a monthly, seasonal and annual basis. Inions (1990) and Inions *et al.* (1990) use the number of days on which precipitation occurred in each quarter, when studying the productive potential of karri forest. This variable is used less frequently than total annual rainfall or variables which express rainfall distribution.

Other forms of precipitation are rarely used in site quality studies. One exception is White (1982a) who used snow depth and rainfall on a quarterly basis in a study of Scots pine in Great Britain.

4.1.2.2.2 TEMPERATURE

Ambient temperature has been used in site quality studies. It may be expressed on an annual, seasonal or monthly interval and as a maximum, minimum or mean value (Chang and Aguilar 1980; White 1982a).

Departures from an optima may also be used as an expression for temperature. For example, Jackson and Gifford (1974) use seasonal departures of ambient temperatures from a postulated optima of 5°C at night and 20°C during the day for their study of *Pinus radiata* in New Zealand. These optima were derived from the studies of Hellmers and Rook (1973). Hunter and Gibson (1984), also studying *Pinus radiata* in New Zealand, derive their own optima and use deviations from 12°C as an independent variable.

Temperature may also be expressed as the number of days during which a threshold temperature was reached. For example, Chang and Aguilar (1980) use the number of days with a maximum temperature equal to or greater than 32.2°C and the number of days with minimum temperature equal to or less than 0°C.

Total annual growing degree days, where one degree day is accumulated for each 1°C rise in temperature above the daily mean when the daily mean temperature is above the minimum thresholds of 0.0°C, 5.0°C, 15.6°C and 21.1°C, was highly correlated with the site index of Sitka spruce along the North Pacific coast (Farr and Harris 1979). Accumulated temperature has been used when comparing forest productivity in Britain and Europe (Christie and Lines 1979). Soil temperature at 15 and 50cm depths have also been used (Chang and Boyer 1977; Chang and Aguilar 1980). Although this variable may be considered as edaphic it is classified as climatic for the purposes of this review.

4.1.2.2.3 RADIATION

The use of variables which describe radiation are more often restricted to mechanistic physiological studies rather than used in empirical models which predict forest productivity. Where used however, radiation has been found useful for predicting forest quality. The three monthly means of total incoming solar-radiation at ground level was a major discriminant between productivity levels of Scots pine (*Pinus sylvestris* L.) across Great Britain (White 1982a). An empirical association between the absorption of radiation by the canopy as a whole and its utilisation in dry matter productivity was developed by Specht (1981) for Australian vegetation. Doley (1982) describes a method which uses only latitude and cloud cover for estimating the daily integral of global radiation at a site. Utilization of the radiation, in the production of dry matter, is also predicted. The single parameter value is recommended as a measure of the productive potential of the site.

A study in the Appalachian mountains, West Virginia, U.S.A., developed a relationship between the radiative index of dryness (i.e. the ratio of yearly sums of net radiation to those of the latent heat of precipitation) and forest biomass (Tajchman 1984). Similarly, in a comparative study of the Appalachian mountains site and a site in the Brindabella mountains of the Australian Capital Territory, this radiative index of dryness was related to forest productivity with optimum growth potential occurring at a value of 0.75 (Tajchman and Lacey 1986).

Although variables derived from measures of radiation have shown a close association with forest productivity in studies of both broad (White 1982a) and narrow (Tajchman 1984) geographic ranges, the difficulty inherent in obtaining these measurements prevents their more frequent use. The study by Doley (1982) provides a method to overcome this problem but has received little attention in the literature. Similarly, annual total solar radiation may be calculated from aspect, gradient and latitude from an algorithm developed by Swift (1976).

4.1.2.2.4 MISCELLANEOUS CLIMATIC VARIABLES

A range of miscellaneous climatic variables have been used in the following forest productivity studies, Christie and Lines (1979), Chang and Aguilar (1980) and White (1982b). They include the number of days without frost on a monthly, seasonal and annual basis, frost free season. day length (hr), growing season (days), potential and annual evapotranspiration, annual runoff (mm), visibility (coded), wind direction (degrees) and wind speed (ms^{-1}). Of these variables only growing season length, frost free season and day length were considered important in the study by Christie and Lines (1979).

4.1.2.3 TOPOGRAPHICAL VARIABLES

With the exception of edaphic attributes, topographic variables are the most common class of variable used in studies which relate environmental attributes to forest productivity. Most studies use topographic variables in functions also comprised of edaphic and/or climatic attributes (Mader 1976; Cook *et al*. 1977; Grey 1979; Blyth and MacLeod 1981; Meeuwig and Cooper 1981), although some studies have used topographic variables exclusively (Stage 1976; McNab 1985; Verbyla and Fisher 1989). Some studies have found little relationship between topographic variables and forest productivity (Mader 1976; Page 1976) while others have found the variables discriminatory (Meeuwig and Cooper 1981; Schmoldt *et al.* 1985).

A common expression of topography is slope, which is usually expressed as a percentage. The variable has been used in studies of site quality for northern hardwoods in Wisconsin and upper Michigan, U.S.A. (Schmoldt *et al.* 1985), yellow poplar in Ohio, U.S.A. (Munn and Vimmerstedt 1980), lodgepole pine (*Pinus contorta* Dougl. var. *latifolia* Engelm.) and white spruce (*Picea glauca* (Moench) Voss) in Alberta, Canada (Corns and Pluth 1984) and white pine in Massachusetts, U.S.A. (Mader 1976). The variable has been significantly related to the diameter growth of *Quercus macrocarpa* Michx. in Kansas, U.S.A. (Abrams 1985), basal area growth of *Pinus monophylla* Torr. & Frem. and *Juniperus osteosperma* (Torr.) Little, in Nevada, U.S.A. (Meeuwig and Cooper 1981), site index and mean annual increment of *Pinus patula* schlecht et. Cham. in Transkei (Grey 1979) and the site index of jack pine in Ontario, Canada (Schmidt and Carmean 1988). White (1982a) expressed slope in degrees from horizontal, when relating environmental variables to the site quality of Scots pine in Great Britain and found the variable significant. The shape of the slope is a variable which is difficult to obtain in a quantitative measure. Where used, the shape of the slope is usually coded in a nominal scale (Grey 1979; Munn and Vimmerstedt 1980).

Topographic position, defined as the position of the stand in relation to the total length of the slope, is often used in site quality assessment. Cook *et al.* (1977) coded topographic position into six nominal classes when studying the site quality of Scots pine in north east Scotland. The variable has also been expressed as a percentage of distance from a ridgetop (White 1982a; McNab 1985). The distance of a plot from a ridge top, expressed as a ratio with the total slope length, has been found to bear a significant relationship with the height of yellow poplar in Ohio, U.S.A. (Munn and Vimmerstedt 1980). Blyth and MacLeod (1981) express topographic position as the ratio of the difference in elevation between the sample plot and the ridge crest to the difference in elevation between the sample plot and the valley floor. The variable was significantly related to the local yield classes of Sitka spruce in north east Scotland.

Degrees azimuth, as a measure of aspect, is another common topographical variable (Cook *et al.* 1977; Grey 1979; Munn and Vimmerstedt 1980; Meeuwig and Cooper 1981). Many studies assume that degrees azimuth may be coded as a cosine function with the minimum in a predetermined quadrat. This assumption was first proposed by Gaiser (1951). The use of higher order sine or cosine functions to represent asymmetries was introduced by Carmean (1967). Significant relationships have been found when predicting the site index of northern hardwoods in the U.S.A. by Schmoldt *et al.* (1985), who use the transformation of Lloyd and Lemmon (1970). Similarly, aspect expressed as 1 + sine (azimuth) was significantly related to the site index and mean annual increment of lodge pole pine and white spruce in Alberta, Canada (Corns and Pluth 1984).

Stage (1976) argues that expressions for the effect of aspect should always be considered as terms involving an interaction with slope. Stage illustrates a method using the relationship between the site index of western white pine (*Pinus monticola* Dougl.) and aspect, slope and habitat type. The methodology has been used successfully by McNab (1985).

Topographic shelter, defined as the sum of the skyline angles in the eight principal compass directions, was significantly related to the local yield class of Scots pine in Scotland (Cook *et al.* 1977). Similar results were recorded for Sitka spruce (Blyth and MacLeod 1981). Microtopex, defined as the sum of the angles formed by the plot centre and points 30 m away on the eight principal compass directions, has not been found useful. Exposure, the amount of skyline at or below 0°, given as a percentage, was significantly related to the height of young *Eucalyptus delegatensis* R.T. Baker in Tasmania, Australia (Keenan and Candy 1983).

Drainage class, coded on a nominal scale, is related to the site index of lodgepole pine and white spruce in Alberta, Canada (Corns and Pluth 1984) but not to the site index of white pine in Massachusetts U.S.A. (Mader 1976).

Geomorphic shape of the land surface is a topographic variable that accounts for the concentration or dilution of surface water and nutrient, but has seldom been included in studies of site quality. Surface shape was included in a site quality study of yellow poplar in Ohio, U.S.A. but was not found to be significant (Munn and Vimmerstedt 1980). However, McNab (1985) studying yellow poplar in Georgia and Central Virginia U.S.A., describes a method to rate the surface shape of landforms, and found the variable significantly related to tree height at age 50 years.

Other topographic variables used less commonly are distance to a bog margin. This variable was used in a site quality study of black spruce in northern Minnesota, U.S.A. by Heinselman (1963) and Watt and Heinselman (1965). Distance to the sea was included in a study of the site quality of balsam fir and black spruce in western Newfoundland Canada, but was not found to be significant (Page 1976).

The advantage of topographic variables is the ease with which the data are derived. Many variables can be obtained from maps, while others are obtained *in situ* with simple measurement. One study was able to successfully discriminate between productivity classes with topographical data derived from satellite imagery (Fox *et al.* 1985).

4.1.2.4 MISCELLANEOUS VARIABLES

A range of variables, used in studies of site quality, do not fit neatly into the edaphic, climatic or topographical classes of attributes. These variables are presented under the collective heading of miscellaneous attributes.

Foliar nutrient values have been successfully related to site quality in a number of studies. In the Douglas-fir ecosystems on Vancouver Island Canada, foliar properties were highly correlated with soil properties and site index. Increases in soil nutrient availability were correlated with increased foliar N concentrations of the current year foliage. A consistent correlation was shown to exist between increased soil nutrient availability, particularly for N, Mg and Ca and decreased foliar Mn and Al and site index was significantly greater on sites with greater quantities of nutrient (Kabzems and Klinka 1987b). Growth intercepts were positively correlated with foliar concentrations of N, P and K in jack pine stands in New Brunswick, Canada (Hamilton and Krause 1985). The site index of loblolly pine was also significantly correlated to foliar N concentrations (Lea and Ballard 1982). A significant relationship was found to exist between foliar P and Ca and site index for *Pinus radiata* at Mullions Range State Forest, Australia. Although site index could be predicted via foliar P and Ca, direct calculation from soil parameters was preferred (Truman *et al.* 1983). However, Saunders *et al.* (1984), also working on the site quality of *Pinus radiata* in Australia, conclude that the individual measure most closely associated with site index was seedling foliar P.

White (1982a) used the amount of monoterpenes in the terminal shoot of Scots pine when studying site quality in Great Britain. This approach has received no further attention.

Shrub competition has been shown to significantly affect stand growth (Brand and Janas 1988) but has received scant attention in the literature concerned with the prediction of site productivity. This is probably due to the fact that shrub competition is not an inherent site factor and is capable of being manipulated with silvicultural practice. However, White (1982a) does incorporate a visual estimate of the mean height of ground flora, as a measure of interspecies competition in a site quality study in Great Britain. During a similar study White (1982b) included in a model to predict site quality from environmental attributes, variables which express the degree of intraspecific competition. The severity of grazing by large herbivores was also included in the same study.

Depth to an impenetrable layer has been used by Schmoldt *et al.* (1985) but was found to have little relationship with site quality. Blyth and MacLeod (1981) use a similar variable titled, effective depth, which is calculated as total rootable depth X 0.01 X (100-stone content), where stone content is expressed on a percentage of volume basis.

The site index of longleaf pine was related to the thickness of the sand layer and the buried topography of a compacted formation (Oliver 1978). In Oliver's study the subsurface compacted formation was mapped on a 1214 ha watershed in South Carolina, U.S.A. The depth to water table is a variable used often in site quality studies in east European countries. It has an inverse relationship with the height growth of plantations in south east European Russia (Malan'in 1985). Similarly, models are developed to predict the height of plantation grown species at 100 y in relation to the water table in west Siberian lowland (Grigor'ev 1986).

4.1.3 MEASURES OF SITE QUALITY

The measure chosen to represent site quality is of considerable importance. If the measure is more an artifact of silvicultural history than site quality, equations comprised of environmental attributes used to predict the parameter, are unlikely to be successful. Another consideration is the practical application of the measure. For example, site index requires the measurement of stand top height and stand age and is usually unaffected by stocking and intermediate cutting (except thinning from above). On the other hand volume production requires the calculation of stand volume (see Clutter *et al.* (1983) for methodologies) and age and is affected by stand density and intermediate cutting. Consequently, site index is applicable where stands vary in stocking density and silvicultural history, while volume production is applicable where stands have equivalent stocking density and silvicultural history and where the extra effort required to measure stand volume is logistically and fiscally feasible.

4.1.3.1 SITE INDEX

The most common measure of site quality used as the dependent variable, in equations comprised of environmental variables is, site index. Successful correlations between environmental variables and site index have been obtained for *Pinus radiata* in New Zealand (Hunter and Gibson 1984), aspen in Wisconsin, U.S.A. (Fralish and Loucks 1975), white pine in Massachusetts, U.S.A. (Mader 1976), mixed coniferous stands in Idaho, U.S.A. (Brown and Loewenstein 1978), loblolly, slash, longleaf and shortleaf pine in the Gulf Coast Plain of the

U.S.A. (Shoulders and Tiarks 1980), balsam fir and black spruce in Newfoundland, Canada (Page 1976), lodgepole pine and white spruce in Alberta, Canada (Corns and Pluth 1984), jack pine in Ontario, Canada (Schmidt and Carmean 1988), Sitka spruce along the Pacific coast of North America (Farr and Harris 1979) and *Eucalyptus globulus* in Rwana (Gasana and Loewenstein 1984). A more comprehensive list of earlier work is given in Carmean (1975).

The success of using site index as the dependent variable is not universal. For example, Schmoldt *et al.* (1985) found that after Chapman-Richards functions were fitted to yield data, for northern hardwoods in the U.S.A., significant correlations were derived between the upper asymptote and maximum growth rate of each fitted curve and the concentrations of extractable Ca, Mg and aspect. No significant correlations were detected between site index and these site factors.

4.1.3.2 AVERAGE STAND HEIGHT

The average height of a stand has been used as the dependent variable in a study to predict site quality of mixed conifer stands in Idaho, U.S.A. from soil and topographic variables (Brown and Loewenstein 1978). Soil and topographic variables explained 70% of the variation in site index and 94% in the average height of the stand. Although a seemingly high proportion of the variation in average height was explained, a major portion of the explanation (71%) is attributable to the inclusion of stand age into the equation. Similar results are reported by Munn and Vimmerstedt (1980).

The average height of a stand has also been used in a site quality study of *Eucalyptus delegatensis* R.T. Baker in Tasmania, Australia. The average heights were standardised for age prior to correlation with environmental attributes (Keenan and Candy 1983).

Average height is not a common measure of site quality. Although successfully used in the small number of studies mentioned the variable is noticeably absent from any major review on the topic of forest site quality (Carmean 1975; Hagglund 1981). Because the variable is more difficult to quantify than site index and is prone to variation bought about by silvicultural manipulation, I do not consider it a useful parameter if used in isolation.

4.1.3.3 HEIGHT GROWTH INTERCEPTS

Height growth intercepts are a common measure of site quality (Thrower 1987). This method was developed for conifer species that have distinct internodes marking annual height growth. Height growth intercepts use information on height growth for some relatively short period as an index to site productivity. Few studies use the measure as the dependent variable when relating environmental variables to site quality. One exception is Hamilton and Krause (1985) who found a significant relationship between ericaceous plant cover, drainage class, extractable P, exchangeable K and Ae horizon development and the two, three and four year height growth intercepts of jack pine in New Brunswick, Canada.

4.1.3.4 VOLUME

An alternative to obtaining site quality information from height measures is to estimate site quality from volume-age relationships. However, the volume attained by a stand at any given age can be greatly affected by factors other than site quality. Unless these factors are controlled or adjustments are made to reflect their influence, volume production differences between stands will have little relationship with site quality. As a result total volume production is rarely used as the dependent variable in studies which relate site quality to the environment.

Some exceptions include Brown and Loėwenstein (1978) who found that 86% of the variation in total volume production of mixed conifer stands in Idaho, U.S.A. was explained by soil and topographical properties and age. Age alone accounted for 42% of the variation. Likewise, Mader (1976) found that total board and cubic volume of white pine in Massachusetts, U.S.A. were related to age and soil factors. However, the standard errors associated with the derived equations were large and the use of equations predicting site index rather than volume was recommended.

Mean annual volume increment (MAI) has been used as the dependent variable by Grey (1979) while predicting the site quality of *Pinus patula* in South Africa. Although a significant relationship was derived, the use of site index as the dependent variable, gave higher r^2 values. Similar results are reported by Corns and Pluth (1984), who predicted the site quality of mixed wood forests in Alberta, Canada.

Mean annual increment is a time dependent statistic and its success as a measure of site quality may well depend on the age at which MAI is calculated. An interesting approach to this problem was contrived by Buckley (1988) who used the maximum MAI increment of *Eucalyptus camaldulensis* in Nigeria as the dependent variable. Significant equations were derived explaining between 47.8 and 71.4% of the variation in maximum MAI increment, depending on the independent variables selected.

4.1.3.5 BASAL AREA AND TREE DIAMETER

Measures of basal area or diameters are rarely used as statistics for site quality. The most successful study of this type is that of Meeuwig and Cooper (1981) who used potential basal area growth as a site quality index for pinyon-juniper (*Pinus monophylla* Torr. & Frem. and *Juniperus osteosperma* (Torr.) Little) woodlands in Nevada. Potential basal area growth was successfully determined by a function comprised of slope gradient, exposure, landform and parent material.

The relationship between soil, topographic variables, the age and diameter distributions of *Quercus macrocarpa* Michx. and *Quercus muehlenbergii* in Kansas U.S.A. has been studied by Abrams (1985). High growth rates were correlated with low topographic slope, low available NH_4 and NO_3 and high K values.

Studies in which basal area or diameter are successfully related to environmental variables are the exception rather than the rule. Other studies have shown that other measures, such as site index or MAI provide better estimates of site quality (Mader 1976; Schmoldt *et al.* 1985).

4.1.3.6 BIOMASS

Site quality is occasionally measured by biomass production. The measure is difficult to obtain and therefore, is rarely used as a practical measure of forest site quality. However, it has been used as the dependent variable in a study in the Appalachian mountains, U.S.A. (Tajchman 1984) and the Brindabella mountains, Australia (Tajchman and Lacey 1986). In both studies biomass was successfully predicted by the radiative index of dryness (see Section 4.1.2.2.3). The effort required to obtain measures for both the dependent and independent variables of this function severely limits the practical application of such measures and functions.

4.1.3.7 MISCELLANEOUS PRODUCTIVITY MEASURES

In a study of the site quality of white spruce in Minnesota, U.S.A., Harding *et al.* (1985) used cluster analysis to define "growth groups" based on combinations of site index, MAI and basal area. Environmental variables were then used in a discriminant analysis to allocate independent plots to a "growth group". Except for the effort involved with the collection of MAI, basal area and site index for each plot, this approach has wide practical application but has received scant attention in the literature.

Four indices of site quality were compared with the volume growth of ponderosa pine (*Pinus ponderosa* Laws.) stands in Western Montana, U.S.A. (McLeod and Running 1988). Indices based on quantifying the biophysical factors or physiological processes that control productivity (available moisture index; the sum of annual precipitation and soil available water capacity and a relative index of seasonal photosynthesis from computer simulations (Running *et al.* 1987)) worked as well as those based on tree or stand measurements (i.e., site index and leaf area index). However, site index was by far the simplest measure to obtain.

4.1.4 CONCLUSION

The literature pertaining to evaluating site productivity from environmental attributes is vast. By far the most common measure of site productivity used in such studies is site index.

The environmental variables used to predict site productivity are many and varied. The most common are those edaphic and topographic variables which are easily measured. There are no universally accepted set of variables which have proven predictive value for all studies. Whether an environmental variable will be related to site productivity will depend on;

(i) the measure of site productivity used;

(ii) the tree species involved;

(iii) the geographic locality of the study; and

(iv) the scale of the study (i.e. does the study area cover 100 ha or 100 km^2).

4.2 RELATIONSHIPS BETWEEN SITE PRODUCTIVITY AND ENVIRONMENTAL ATTRIBUTES FOR *E.GLOBULUS* IN SOUTH WEST WESTERN AUSTRALIA. I. DATA EXPLORATION, A MULTIVARIATE APPROACH.

4.2.1 INTRODUCTION

Knowledge of the productive capabilities of the land comprising a plantation estate, yields economic advantages through the avoidance of unprofitable sites. Knowledge of productive capability also provides an improved basis for managerial decision making and prediction of wood volume over the estate, stratified on a productivity basis. The most frequently used measure of productivity is site index. However, under certain criteria site index is unattainable.

The situation exists in south west Western Australia where the productive capabilities of land must be assessed at year zero of the rotation. This is because the land base for the *E. globulus* plantation estate is acquired from private landowners who may receive an annuity beginning at year zero of the rotation. Secondly, under this scheme, land is being acquired in areas where there is limited knowledge of plantation forestry as a land use. Thus, a method which is capable of assessing the productive capabilities of a tract of land at year zero of the rotation is required.

The usual approach to such a situation is to develop a linear regression equation to predict a measure for productivity such as site index, from environmental attributes. Although some successes have been attained with this approach many studies are limited in (i) their success in accounting for a significant proportion of the productivity variation (Corns and Pluth 1984; Monserud *et al.*1990), (ii) their applicability being restricted by the need to derive variables which require laboratory analysis (Daubenmire 1976), and (iii) the synergistic and nonlinear nature of many environmental attributes are often not accounted for by linear combinations of single environmental attributes, thus the predictive capabilities of the equation are reduced (Mc Quilkin 1976). One approach used nonlinear regression to overcome some of these problems (Czarnowski *et al.* 1971) while some others do include interactive terms in their equations (Jackson and Gifford 1974; Corns 1983).

The use of multivariate techniques to explore patterns within data, prior to formal hypothesis testing and equation construction may be one way some of these deficiencies may be overcome but the practice is absent from the literature concerned with predicting productivity from environmental attributes. Where multivariate techniques are used, their purpose is usually to reduce the number of independent variables (Page 1976; White 1982a) or to summarise the data set into a few dimensions which are subsequently used as the independent attributes (Keenan and Candy 1983; Green *et al.* 1989). Other uses of multivariate techniques include the use of cluster analysis, used by Harding *et al.* (1985) to group sites into "growth response groups" based on site index, basal area and mean annual biomass increment. A discriminant function was then developed to allocate plots to a growth response group on the basis of environmental attributes. A study by Kabzems *et al.* (1987a) used principal components and discriminant analysis to examine the relationship between subjectively defined groups of land and quantitative environmental attributes. Kabzems *et al.* (1987b) went on to use detrended correspondence analysis to summarise floristic data, principal components analysis to summarise soil data and canonical correlation analysis to examine the relationship between soils and floristics, all of which were then related to site index.

It is the aim of this study to;

(i) identify the gradients in the environmental data which account for most of the information contained within the data set; and

(ii) identify the nature and direction of these environmental gradients through multidimensional space, derived from top height development data.

4.2.2 METHODS

4.2.2.1 PLOT SELECTION

Seventy four environmental attributes were obtained from 56 plots in *E. globulus* plantations in south west Western Australia. These plots also yielded data used to derive the site index and top height development equations detailed in Chapter 3. None of the plots had been fertilised except at the time of planting. Half of the plots were placed in stands established on agricultural pasture, while half were placed in stands established on unimproved land (Figure 1).

Plots were located with an aim of obtaining the widest possible distribution along the ecological gradients of the study area. Placement of plots within stands avoided soil boundaries or any other factor which may have yielded heterogeneous growth patterns within the plot.

4.2.2.2 PRODUCTIVITY CRITERIA

The stand density, previous land use and the magnitude of the competition exerted by the understory (where present) varied between plots. As these factors affect stand productivity (Cromer 1973; Skinner and Attiwill 1981) the use of many productivity measures, such as total stand volume, basal area or mean annual increment, was not considered reasonable.

The productivity criteria used in this study will be (i) site index, as it is considered to be less affected by the above influences than other criteria, and (ii) cluster groups, defined in Chapter 3 (Section 3.2.2.4.2). Relating environmental attributes to the cluster groups will indicate possible causes for the polymorphic nature of the top height development patterns observed.

4.2.2.3 ENVIRONMENTAL VARIABLES

4.2.2.3.1 CLIMATIC VARIABLES

The climatic data were obtained from the biological prediction system (BPS). This program interpolates between meteorological stations using the methodologies of Wahba and Wendelberger (1980) and Hutchinson (1987). The program requires latitude, longitude and elevation as input and yields values for a range of climatic variables, as demonstrated by Busby (1986), Nix (1986) and Booth *et al.* (1988).

The twenty one climatic variables obtained for each plot are described in Table 18. Univariate statistics for each climatic attribute are given in Table 19.

4.2.2.3.2 EDAPHIC VARIABLES

Field Procedures

At the centre of each plot a soil pit was excavated to 2 m or to a restrictive layer. Excavations were dug with a mini excavator (Plate 1). Each soil pit was considered representative of that plot and the soil profile was described (Northcote 1971) and photographed for future reference (Plate 2).

Fifteen soil samples from a depth of 10 cm were collected from within the plot boundaries. These samples were combined to form three composites of five samples each. The three composite samples were used to derive edaphic chemical values and to obtain some indication of the magnitude of the heterogeneity of derived variables. One composite sample, of five samples, was collected from within the soil pit at a depth of 30 cm. This sample was also used to derive edaphic chemical variables.

To determine bulk density, soil cores of 206.2 cm^3 were extracted at 10 cm, 30 cm, 50 cm, 100 cm and 150 cm or until a restrictive zone was encounted. Coring followed the procedure of Blake and Hartge (1986).

Plate1: The machine used to excavate soil pits.

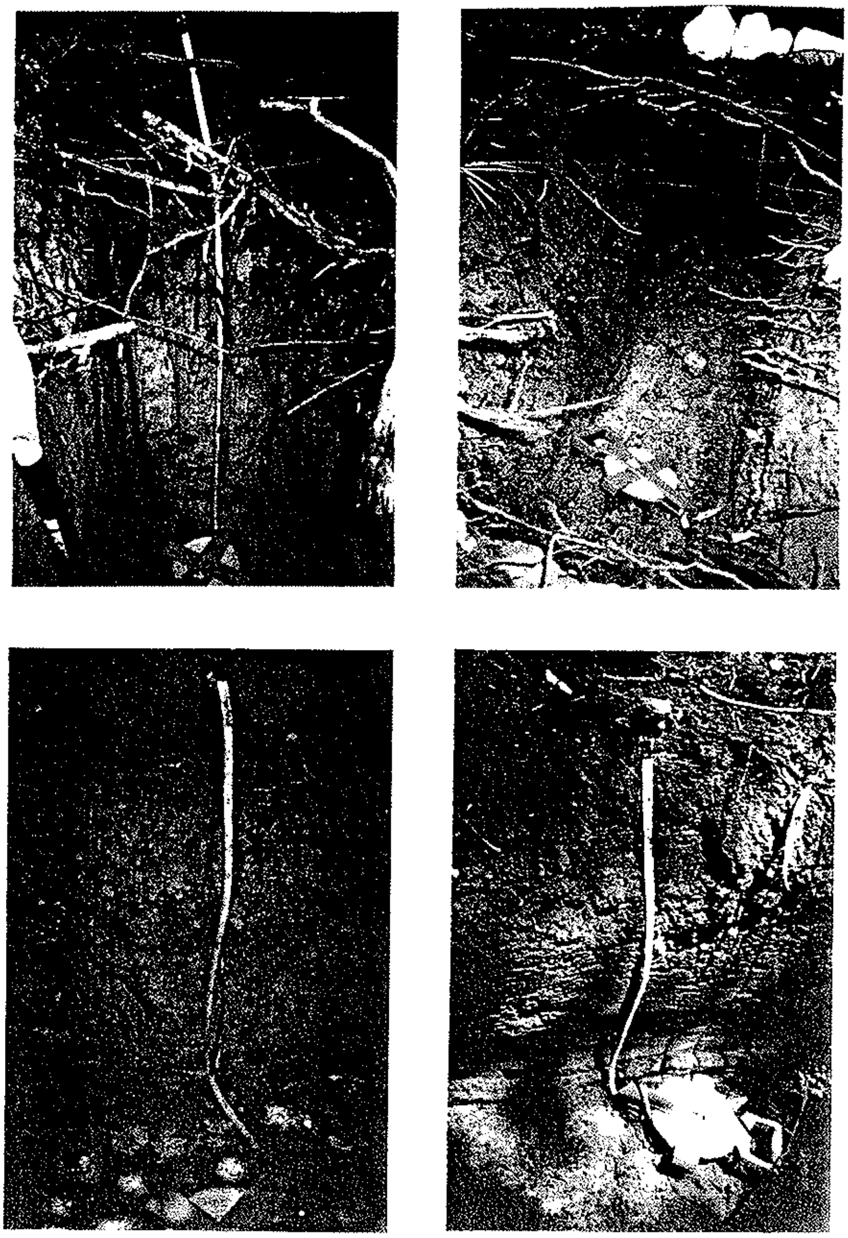


Plate 2: A range of soil types encountered in this study.
(A) a uniform profile of organically stained sand, (B) a gradational profile of gravelly sand, (C) a duplex profile of gravelly sand over clay; and, (D) a duplex profile of loamy sand over clay

Laboratory Procedures

Composite soil samples were air-dried and analysed via the following methodologies. Phosphorous content was determined colorimetrically using the method of Murphy and Riley (1962) at 882 nm. The potassium content was determined using the same solution with atomic absorption at 766.5 nm. For determining pH and conductivity a solution of soil and water with a 1:5 soil/water ratio was mixed. pH was measured with a pH meter while the conductivity was measured using a conductivity meter calibrated against a 0.01 M KCl solution.

The organic carbon content was determined by the Walkley-Black method (Walkley and Black 1934). The soil was treated with 8% dichromate and concentrated sulphuric acid and the resultant chromous ions were measured colorimetrically at 600 nm. Nitrogen, expressed as ammonium ions was determined by extracting the soil with 1 M potassium chloride solution. The ammonium content of the solution was determined colorimetrically at 630 nm by the indophenol blue reaction. Nitrogen, expressed as nitrate, was determined after the soil was extracted with 0.02 M aluminium sulphate solution. The nitrate content of the solution was determined by a specific ion electrode. Reactive iron was determined after the soil was extracted with Tamm's reagent (Ammonium oxalate-oxalic acid at pH 3.25). The iron content of the solution was determined by atomic absorption at 248.3 nm. Edaphic chemical variables derived in this manner are described in Table 18, while their univariate statistics are given in Table 19.

Bulk densities of the soil samples were determined after the cored soil samples were dried at 105°C to a constant weight. Bulk density is the oven-dried mass of the sample divided by its volume. Soil particles greater than 2 mm diameter were separated physically by sieving. After weighing, the amounts of such particles were expressed as a bulk density.

Percentages by weight of sand, silt and clay of the fine fraction were determined by the hydrometer method (Gee and Bauder 1986). To enhance separation and dispersion of aggregates, soils were pretreated with hydrogen peroxide to remove organic matter. A solution of water, hydrogen peroxide and soil was heated at 90°C until reaction ceased. Fifty grams of treated soil was transferred to a beaker and mixed with 250 ml of distilled water and 100 ml of sodium-hexametaphosphate solution (50 g/L). The sodium-hexametaphosphate creates particle repulsion as a result of elevation of the particles zeta potential. This process is accomplished by saturating the exchange complex with sodium. The resulting suspension was mixed for five minutes with an electric mixer then transferred to a sedimentation cylinder and mixed with distilled water to bring the volume up to 1 L. It should be noted that many chemical and physical dispersion techniques exist (c.f. Edwards and Bremner 1967; Kubota 1972; Mikhail and Brimer 1978), however, the technique outlined above was found to be most appropriate given the logistic constraints of this study.

Once in the sedimentation cylinder the suspension was mixed by end-over-end shaking for one minute and hydrometer readings taken at 30 s, 60 s, 1.5 h and 24 h. Readings were also taken from a "blank" solution of 900 mL of distilled water and 100 mL of sodium-hexatetaphosphate solution. Ambient temperatures at the time of each measurement were also recorded. A standard hydrometer, ASTM no. 152 H, with Bouyoucos scale in g/L was used. Percentages by weights of sand, silt and clay were derived from these measurements using the formulae of Gee and Bauder (1986).

Random samples were selected and the sediment and suspension from the sedimentation cylinder were washed through a 53 μm sieve. The separated sand particles were transferred into weighing dishes and dried at 105°C to a constant weight. The percentage sand derived in this manner was used as a cross check of that derived by the hydrometer method. In all cases the two figures did not vary enough to be of concern and the hydrometer derived values were accepted.

Particle sizes were expressed as percentages by weights at each depth. Variables derived in this manner are described in Table 18 while their univariate statistics are given in Table 19.

4.2.2.3.3 OTHER VARIABLES

Rooting depth was determined by the depth to restricting materials, such as bedrock or hardpan, or from observations of the rooting pattern in those soils lacking restricting materials. Slope was expressed as a percentage, while topographic position is defined as the percentage of distance from a ridge top in relation to the total slope length. Aspect was transformed by 1 + sine (degrees azimuth) and 1 + cosine (degrees azimuth). These variables and their univariate statistics are given in Tables 18 and 19.

Other environmental attributes were measured from each plot, but after initial screening, were not found to possess any relationship with productivity. They also have the undesirable property of containing some subjective elements. As such, the following variables were not included in subsequent analysis, soil colour, understorey vegetation height and cover, topographic shelter and drainage class.

Table 18: Variable codes and descriptions of environmental attributes.

VARIABLE CODE	DESCRIPTION	UNITS
RAD	Mean annual radiation	$MJM^{-2}day^{-1}$
RAD_HM	Highest monthly radiation	$MJM^{-2}day^{-1}$
RAD_LM	Lowest monthly radiation	$MJM^{-2}day^{-1}$
RAD_R	Radiation range (RAD_HM-RAD_LM)	-
RAD_S	Radiation seasonality	
RAD_WQ	Radiation in the wettest quarter	$MJM^{-2}day^{-1}$
RAD_DQ	Radiation in the driest quarter	$MJM^{-2}day^{-1}$
TEMP	Mean annual temperature	°C
TEMP_HM	Highest monthly temperature	°C
TEMP_LM	Lowest monthly temperature	°C
TEMP_R	Temperature range (Temp_HM-Temp_LM)	-
TEMP_S	Temperature seasonality	
TEMP_WQ	Temperature in the wettest quarter	°C
TEMP_DQ	Temperature in the driest quarter	°C
RAIN	Total annual rainfall	mm
RAIN_HM	Rainfall in the wettest month	mm
RAIN_LM	Rainfall in the driest month	mm
RAIN_R	Rainfall range (RAIN_HM-RAIN_LM)	-
RAIN_S	Rainfall seasonality	
RAIN_WQ	Rainfall in the wettest quarter	mm
RAIN_DQ	Rainfall in the driest quarter	mm
P10	Phosphous at 10 cm	ppm
NIT10	Nitrate at 10 cm	ppm
AMM10	Ammonium at 10 cm	ppm
K10	Potassium at 10 cm	ppm
ORG_C10	Organic carbon at 10 cm	%
FE10	Reactive iron at 10 cm	ppm
CON_D10	Conductivity at 10 cm	dSm^{-1}
PH10	pH at 10 cm	
C_NIT10	ORG_C10÷NIT10	-
C_AMM10	ORG_C10 ÷AMM10	
P_NIT10	P10 x NIT10	
P_AMM10	P10 x AMM10	-
P30	Phosphous at 30 cm	ppm
NIT30	Nitrate at 30 cm	ppm
AMM30	Ammonium at 30 cm	ppm

(Cont.)

Table 18 (Cont).

VARIABLE CODE	DESCRIPTION	UNITS
K30	Potassium at 30 cm	ppm
ORG_C30	Organic carbon at 30 cm	%
FE30	Reactive iron at 30 cm	ppm
CON_D30	Conductivity at 30 cm	dSm^{-1}
PH30	pH at 30 cm	
C_NIT30	ORG_C30 ÷ NIT30	-
C_AMM30	ORG_C30 ÷ AMM30	-
P_NIT30	P30 x NIT30	-
P_AMM30	P30 x AMM30	-
BD10	Bulk density at 10 cm	$g\ cm^{-3}$
CLAY10	Percentage clay content at 10 cm	% (wt)
SILT10	Percentage silt content at 10 cm	% (wt)
SAND10	Percentage sand content at 10 cm	% (wt)
CBD10	CLAY10 x (BD10/100)	-
SABD10	SAND10 x (BD10/100)	-
SIBD10	SILT10 x (BD30/100)	-
GRBD10	% weight of the coarse fraction at 10cm x (BD10/100)	-
BD30	Bulk density at 30 cm	$g\ cm^{-3}$
CLAY30	Percentage clay content at 30 cm	% (wt)
SILT30	Percentage silt content at 30 cm	% (wt)
SAND30	Percentage sand content at 30 cm	% (wt)
CBD30	CLAY30 x (BD30/100)	-
SABD30	SAND30 x (BD30/100)	-
SIBD30	SILT30 x (BD10/100)	-
GRBD30	% weight of the coarse fraction at 30cm x (BD30/100)	-
BD50	Bulk density at 50 cm	$g\ cm^{-3}$
CLAY50	Percentage clay content at 50 cm	% (wt)
SILT50	Percentage silt content at 50 cm	% (wt)
SAND50	Percentage sand content at 50 cm	% (wt)
CBD50	CLAY50 x (BD50/100)	-
SABD50	SAND50 x (BD50/100)	-
SIBD50	SILT50 x (BD50/100)	-
GRBD50	% weight of the coarse fraction at 50cm x (BD50/100)	-
DEPTH_IP	Rooting depth	cm
SLOPE_PC	Slope percent	%
TOP_POS	Topographic position	%
SHAPE_C	Aspect transformed with a cosine function	-
SHAPE_S	Aspect transformed with a sine function	-

Table 19: Univariate statistics for each environmental attribute.

VARIABLE CODE	MEAN	N	STD. DEV.	SKEWNESS	KURTOSIS
RAD	17.18	56	0.72	0.00	-0.82
RAD_HM	27.64	56	0.86	-0.04	-0.83
RAD_LM	7.58	56	0.51	0.66	-0.39
RAD_R	20.06	56	0.53	-0.81	1.31
RAD_S	1.17	56	0.03	-0.44	-0.37
RAD_WQ	8.99	56	0.48	0.41	-0.57
RAD_DQ	26.45	56	0.84	-0.37	-0.97
TEMP	15.61	56	0.50	2.44	10.56
TEMP_HM	28.90	56	1.10	-0.35	-1.00
TEMP_LM	5.99	56	0.98	1.80	3.52
TEMP_R	22.91	56	1.85	-0.68	-0.43
TEMP_S	1.47	56	0.12	-1.43	1.65
TEMP_WQ	11.00	56	0.72	2.07	5.40
TEMP_DQ	20.94	56	0.80	0.18	-1.06
RAIN	1041.79	56	232.23	-0.13	-0.70
RAIN_HM	199.94	56	45.92	-0.09	-0.85
RAIN_LM	14.86	56	4.11	0.46	-0.95
RAIN_R	185.08	56	44.91	-0.05	-0.78
RAIN_S	2.13	56	0.22	-0.40	0.34
RAIN_WQ	540.38	56	122.20	-0.17	-0.88
RAIN_DQ	55.23	56	55.23	0.69	-0.66
P10	11.78	161	20.37	3.92	17.06
NIT10	2.41	161	2.38	3.22	12.49
AMM10	6.74	161	3.23	1.49	2.68
K10	86.63	161	72.51	1.91	4.10
ORG_C10	2.58	161	0.79	-0.56	-0.62
FE10	900.90	161	509.90	1.02	1.61
CON_D10	0.07	161	0.24	9.08	89.05
PH10	5.92	161	0.41	-0.96	2.08
C_NIT10	1.57	161	0.89	0.65	-0.43
C_AMM10	0.43	161	0.18	1.16	2.38
P_NIT10	56.84	161	191.70	5.74	35.57
P_AMM10	89.39	161	186.78	4.85	27.72
P30	4.85	56	7.67	5.23	30.50
NIT30	1.73	56	1.18	3.16	12.78
AMM30	4.12	56	2.03	1.86	4.86
K30	60.45	56	66.19	2.65	8.32

(Cont.)

Table 19(Cont).

VARIABLE CODE	MEAN	N	STD. DEV.	SKEWNESS	KURTOSIS
ORG_C30	1.24	56	0.74	1.26	0.89
FE30	597.21	56	425.47	1.13	0.89
CON_D30	0.07	56	0.36	7.11	49.05
PH30	6.03	56	0.51	-1.46	2.94
C_NIT30	0.88	56	0.68	1.92	3.48
C_AMM30	0.33	56	0.20	1.69	3.70
P_NIT30	11.41	56	30.42	5.96	37.40
P_AMM30	91.06	56	175.75	3.95	16.86
BD10	1.20	53	0.31	0.59	-0.43
CLAY10	7.63	55	5.58	3.00	14.49
SILT10	13.54	55	8.38	0.56	-0.28
SAND10	78.83	55	10.47	-0.58	-0.17
CBD10	0.06	53	0.04	2.45	9.63
SABD10	0.58	53	0.23	0.72	0.22
SIBD10	0.09	53	0.06	0.33	-0.99
GRBD10	0.47	53	0.46	0.68	-1.10
BD30	1.45	55	0.35	0.40	-0.77
CLAY30	11.37	55	7.49	0.84	0.40
SILT30	12.08	55	7.86	0.82	0.27
SAND30	76.55	55	13.27	-0.71	-0.19
CBD30	0.09	54	0.06	0.76	0.06
SABD30	0.66	54	0.30	0.81	0.21
SIBD30	0.09	54	0.06	0.86	0.27
GRBD30	0.60	55	0.54	0.50	-1.18
BD50	1.40	53	0.29	0.74	-0.12
CLAY50	22.34	53	20.27	1.09	0.20
SILT50	12.19	53	8.49	0.44	-0.87
SAND50	65.47	53	24.95	-0.75	-0.62
CBD50	0.22	52	0.24	1.17	0.01
SABD50	0.58	52	0.32	0.82	-0.32
SIBD50	0.11	52	0.07	0.83	-0.50
GRBD50	0.49	53	0.52	0.77	-0.79
DEPTH_IP	76.70	56	31.67	0.56	-0.28
SLOPE_PC	6.32	56	4.00	0.66	-0.11
TOP_POS	53.09	56	35.24	-0.04	-1.54
SHAPE_C	0.00	56	0.55	0.41	1.05
SHAPE_S	0.00	56	0.52	0.24	-0.24

4.2.2.4 NUMERICAL ANALYSIS

4.2.2.4.1 CLUSTERING TECHNIQUES

Clustering techniques were used to examine the magnitude of the heterogeneity of the edaphic chemical variables. Each composite sample, from 10 cm depth, was considered independent during the clustering procedures. If composite samples from the same plot cluster together, one would have no need for concern that the sample was not a reasonable representation of that plot.

Each observation was range - standardised by subtracting the minimum value and dividing by the range, so that each attribute attains a maximum value of one. Standardisation of attributes to equal maxima prevents the domination of the classification by those attributes with large scales. An association matrix between composite samples was calculated using the Bray-Curtis metric (Bray and Curtis 1957). The polythetic agglomerative clustering strategy, the unweighted pair-group method using arithmetic averages (UPGMA), was used to impose structure to the association matrix (Gauch and Whittaker 1981). The clustering intensity coefficient *beta* was set at -0.2 (Booth 1978).

4.2.2.4.2 ORDINATION TECHNIQUES

The objective of ordination techniques is to summarise continuous variation in complex data sets (Orloci 1988). The summaries serve two purposes, firstly, to reduce data to dimensions which are manageable. Secondly, for hypothesis generation concerning the relationship which may exist between the subject of the ordination, such as vegetation composition, and extrinsic influences (Austin and Greig-Smith 1968; Clymo 1980; Austin 1985; Bowman and Minchin 1987; Minchin 1987a).

Since the 1950's when ordination techniques were first applied to ecological data (Goodall 1954; Bray and Curtis 1957) a copious literature has developed on the topic. Austin (1985) grouped ordination techniques into three classes. Firstly, those he termed "early ordination methods" which includes polar ordination (Bray and Curtis 1957), principal components analysis (PCoA) (Gower 1966), canonical variate analysis (CVA) and canonical correlation analysis (CCA) (Gauch and Wentworth 1976). Secondly, those he termed "Cornell techniques" which includes reciprocal averaging (RA) (Hill 1973) and detrended correspondence analysis (DCA) (Hill and Gauch 1980). Finally, those he termed "multidimensional scaling" (MDS) (Prentice 1977).

Some comparative studies of ordination techniques suggest RA should be preferred to PCoA, as RA is less sensitive to curvilinear distortions (Kendell 1970) in reduced dimensions (Austin 1976a, b; Fasham 1977; Gauch *et al.* 1977; Gauch *et al.* 1981; Oksanen 1983). Gauch *et al.* (1981) suggest that DCA should be preferred to RA and MDS on the grounds of computational efficiency. However, Austin (1985) questions these comparisons and argues that such comparisons, which use simulated data sets (Minchin 1987b), should examine a reasonable number of data sets and the distribution of samples should be varied rather than using a single type. Austin (1985) states that neither DCA or MDS can be recommended without reservation from the above studies, as alteration in the number of replications of data sets and/or distribution of sample points may lead to different conclusions (Gauch *et al.* 1977; Minchin 1987a, b). A more recent set of comparative studies, conclude that MDS is more robust to changes in sampling distribution, data set replication, response curve shape and noise level of the data, than other ordination techniques (Faith *et al.* 1987; Minchin 1987a; Belbin in press).

Fundamental to any ordination technique is a measure to describe the association of sites to other sites on the basis of their attributes. Because response patterns in ecological space do not conform with any particular model (Austin 1976b) the choice of association measure is paramount. The importance of this choice has been well emphasized (Austin and Noy-Meir 1971; Gauch 1973; Beals 1984), however, little agreement has been reached as to which measure is most suitable. As argued by Faith *et al.* (1987) the choice of dissimilarity for ordination must consider the robustness of the measures relationship with ecological distance over the range of response models which may be encountered. Faith *et al.* go on to state "... that the development and evaluation of dissimilarity - based ordination methods must be guided by only those aspects of the relationship between dissimilarity and ecological distance which are robust". They also suggest that the poor performance of many ordination techniques results from those techniques having in-built measures of dissimilarity which are inappropriate to the underlying response model.

Robustness is defined as the ability to recover an Euclidean ordination space in the presence of highly skewed and noisy unimodal responses, uneven representation of sites in the underlying space and different degrees of response in different parts of the space. Once the robustness of the dissimilarity measure is taken into account comparative studies of ordination techniques have favoured MDS (Michin 1987a, b; Belbin *in press*).

MDS yields an ordinate space where the dimensionality is user defined. The procedure starts with a symmetric dissimilarity matrix of sites. Faith *et al.* (1987) recommend the use of the Kulczynski (Hajdu 1981), Bray-Curtis (Bray and Curtis 1957) or the relativized Manhattan measures (Sokal and Michener 1957) when calculating the dissimilarity matrix due to the robust nature of these metrics. The ordination space is derived such that the distances between sites in the reduced dimensions, match the distances represented by the dissimilarity matrix as much as possible. The success or otherwise of the match is judged by a *stress* value. The lower the stress value the better the match. If the dissimilarity measure is assumed to have a linear relationship with the Euclidean distances in the reduced space, the technique is termed metric multidimensional scaling (MMDS) (Torgenson 1952). In contrast, nonmetric multidimensional scaling (NMDS) assumes only monotonicity where the configuration of sample pairs are in rank order with the dissimilarities (Kruskal 1964). It is argued that monotonicity should be preferred in view of our lack of knowledge concerning species response models (Prentice 1980), however, the risk of losing information through simplification is apparent (Shepard 1974).

Faith *et al.* (1987) developed a multidimensional scaling technique which combines both metric and nonmetric MDS. This hybrid multidimensional scaling (HMDS) procedure is more robust than either MMDS or NMDS (Faith and Norris 1989).

Belbin (in press) modified Faith *et al.* (1987) HMDS procedures by applying ordinal fitting procedures to values greater than a user defined threshold and using Guttman rank imaging for the monotone regression (Guttman 1968). Belbin termed this technique semi-strong hybrid multidimensional scaling (SSH). Belbin compared SSH with other ordination techniques on 3240 simulated data sets which varied in their sampling design, response curves and noise and found that SSH gave significant improvements over HMDS which itself out performed MMDS, NMDS, DCA, CCA, RA, and PCoA.

SSH was used during the data exploration phase of this study. The technique was used to ordinate the plots on the basis of the environmental attributes (environmental ordination). The aim here was to reduce the dimensionality of the data set, with minimum loss of information, and identify those variables most influential in summarizing the reduced space. The resultant axes are then used as construct variables and their relationship with productivity explored.

Plots were also ordinated on the basis of the top height at each year (vegetation ordination). These axes, reflecting growth patterns, were used to explore the relationship between the reduced space and environmental variables.

To commence the vegetation ordination a matrix of dissimilarities between plots was calculated via the Bray-Curtis metric (Bray and Curtis 1957) after the recommendations of Faith *et al.* (1987). As top height is recorded in metres no standardization was required. To commence the environmental ordination the dissimilarity matrix was calculated via the Gower metric (Gower 1971).

MDS procedures sometimes produce local optima in the algorithms search for the minimum stress. In this study 10 randomly assigned starting configurations were used as inputs to the procedure to overcome this problem.

The choice of the number of dimensions to use is subjective. The pattern of increasing stress with decreasing dimensionality can be used to indicate a dimensionality after which there is little gain in explanatory value. This choice should not be taken lightly as high correlations with environmental variables may be lost if too few dimensions are used.

Axes resulting from the ordinations were interpretated in terms of their relationship with site index and cluster groups and individual environmental attributes. The terminology of Faith and Norris (1989) is adopted where the axes of the ordinations will be referred to collectively as dimensions. Directions in the ordination space that produce correlations with environmental or productivity variables will be referred to as vectors and correlates with ordination axes will be referred to as gradients.

The technique of rotational correlation (Dargie 1984) was used to derive vectors. The technique finds the vector in ordination space where the projection of sites onto the vector are maximally correlated with the values of the extrinsic attribute of concern. The direction of the fitted vector corresponds to the direction of maximum slope of the hyperplane fitted by multiple regression.

Rotational correlation is a linear procedure. As such nonlinear relationships in ordinate space will be underrepresented. To identify such relationships all extrinsic attributes were plotted against the ordination axis and examined for nonlinear behaviour.

4.2.3 RESULTS

4.2.3.1 SOIL HETEROGENEITY

The classification of samples was arbitrarily truncated at the 16 group level (Figure 16). Most of the samples collected from the same plot were classified within the same cluster group. As such, the assumption that a composite will yield a reasonable approximation of the edaphic chemical variables of a plot is not unreasonable. As some heterogeneity is obvious the mean value of the three composites will be used in further numerical analysis.

The value of most edaphic variables will differ between sampling depths. Edaphic chemical variables have values which are generally significantly less, when sampled at 30 cm rather than

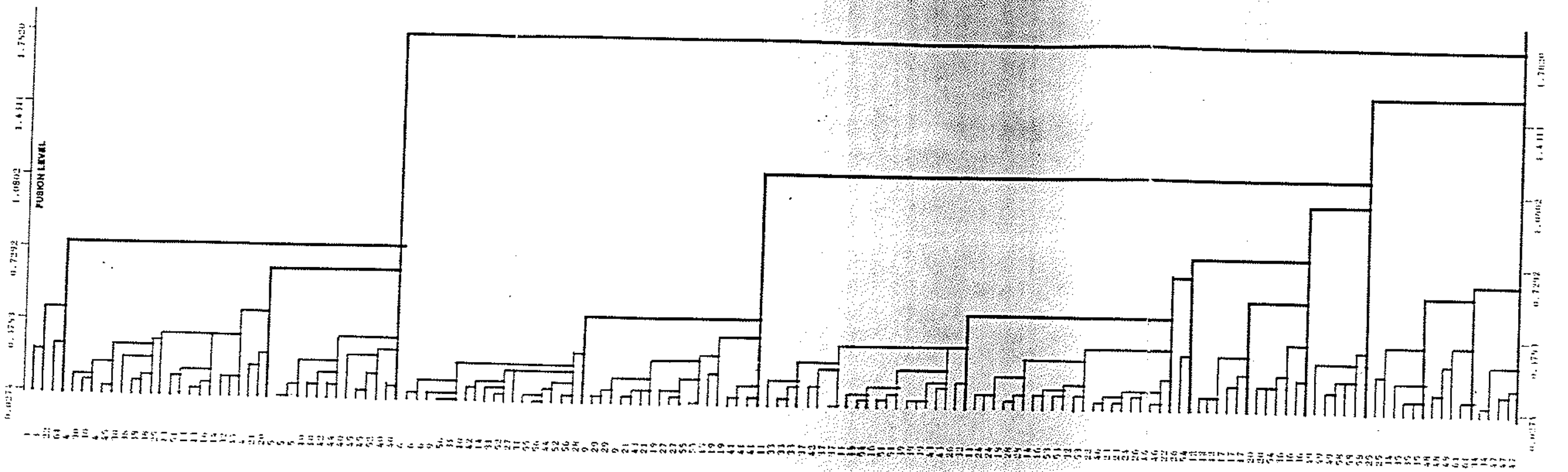
10 cm depth (Table 20). Edaphic physical attributes representing clay content increased significantly with increasing sampling depth, while those representing sand content significantly decreased. Attributes representing silt content did not differ significantly with sampling depth (Table 20).

As the depth at which a sample was taken affects the value of edaphic attributes, edaphic attributes derived from different depths will be treated as independent in subsequent analysis.

Table 20: Mean values of edaphic variables at different sampling depths. Significance levels pertain to differences between means.

VARIABLE	DEPTH (CM)				SIGNIFICANCE LEVEL
	10	30	50	100	P>F
P	11.53	4.85			0.0156
NIT	2.23	0.73			0.0389
AMM	6.62	4.12			0.0001
K	87.04	60.45			0.0365
ORG_C	2.63	1.24			0.0001
FE	892.16	597.21			0.0012
CON_D	0.07	0.07			0.9396
PH	5.89	6.03			0.0976
C_NIT	1.59	0.88			0.0001
C_AMM	0.44	0.33			0.0025
P_NIT	56.92	11.41			0.0531
BD	1.20	1.45	1.40	1.43	0.0003
CLAY	7.63	11.37	22.34	24.71	0.0001
SILT	13.54	12.08	12.19	12.43	0.7932
SAND	78.83	76.55	65.47	62.80	0.0001
CBD	0.06	0.09	0.22	0.25	0.0001
SABD	0.58	0.66	0.58	0.57	0.4221
SIBD	0.09	0.09	0.11	0.13	0.3382
GRBD	0.47	0.60	0.49	0.48	0.5431

6: Dendrogram resulting from the classification of replicate composite soil samples based on soil chemical values (metric = Bray-Curtis; fusion strategy = UPGMA, beta = -0.1).



4.2.3.2 ENVIRONMENTAL ORDINATION

Choosing the number of dimensions required to adequately reflect the information contained within the data was approached in two ways. Firstly, six separate SSH ordinations, in one to six dimensions, were performed and the stress value corresponding to each was examined graphically. The stress value decreased rapidly with increased dimensionality with minimal reduction occurring after the fourth dimension (Figure 17). Secondly, vector correlation coefficients, determined by rotational correlation, were examined in each dimension. If too few dimensions were chosen, important environmental vectors may become obscured. All vectors with large correlation coefficients which existed in six dimensions were also prominent in four dimensions. As such an SSH ordination of four dimensions was accepted. Individual axes resulting from this ordination will be referred to as SSHE1, SSHE2, SSHE3 and SSHE4 respectively.

A maximum of 50 iterations were used to determine the final configuration of plots in reduced dimensional space. Ten randomly selected configurations were submitted as starting values for the iterative procedure, to guard against local optimia solutions. The dissimilarity value, used as a cut off point for hybrid multidimensional scaling, was 0.2859 which is also the value of the third quartile of all dissimilarity values. The final stress value was 0.097 which is in accordance with Kruskal's (1964) recommendations. The directions of vectors along which the projection of plots have the greatest correlation with individual environmental variables were determined for each variable. Vectors with high correlation coefficients ($r > 0.5$) are given in Table 21. The most prominent vectors are those of climatic attributes, particularly those expressed as the wettest and driest quarters or months. Vectors of edaphic physical attributes yield the next highest correlation coefficients. Vectors of edaphic chemical variables did not display as high correlation coefficients as the edaphic physical or climatic variables. Vectors of TOP_POS and DEPTH_IP have comparatively low correlation coefficients (Table 21).

Figure 17: Plot of stress by the number of dimensions for the environmental ordination (●) and the vegetation ordination (■).

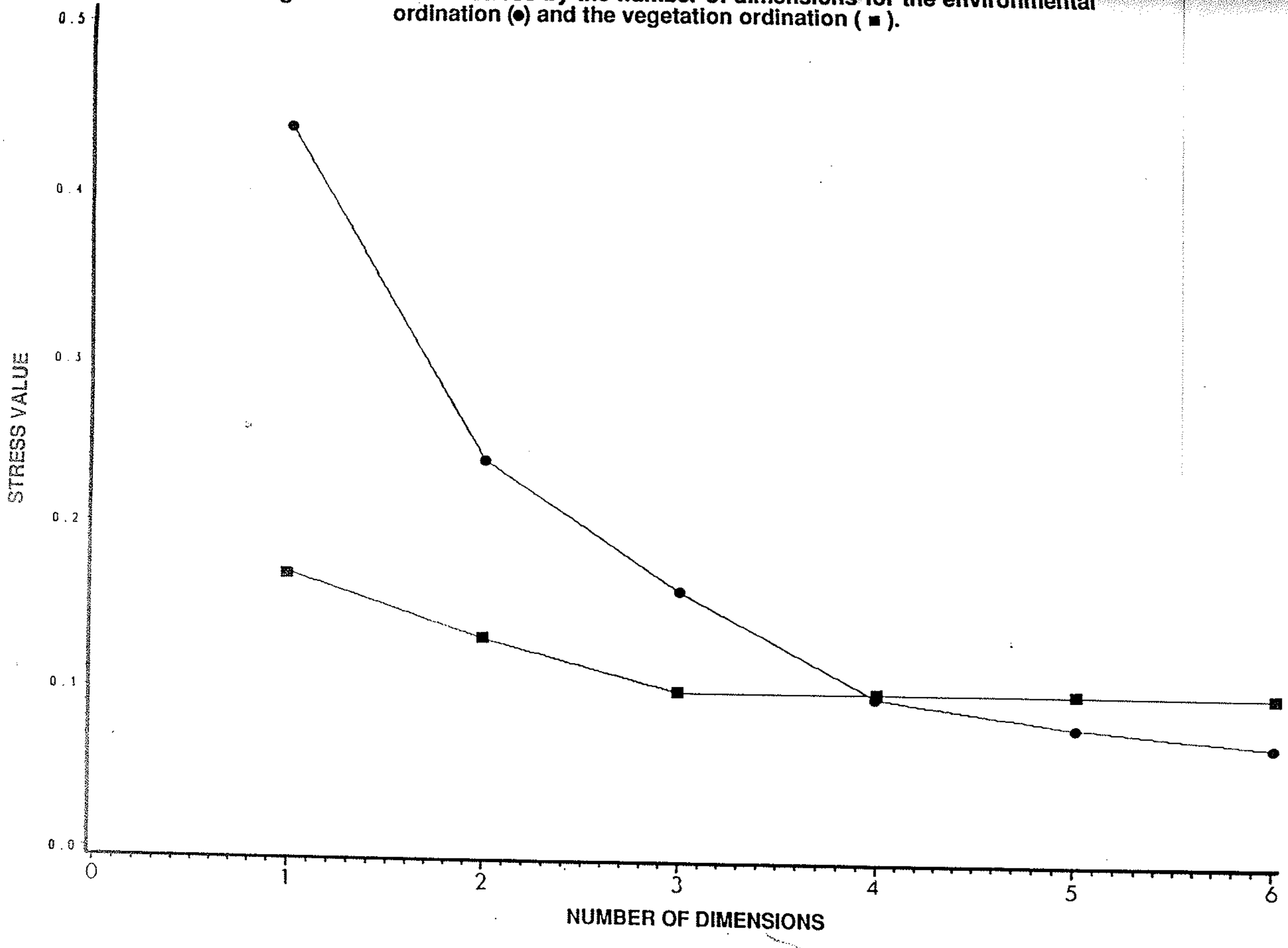


Table 21: Correlations for major environmental vectors in four dimensional space defined by environmental ordination.

VARIABLE	CORRELATION	VARIABLE	CORRELATION
NIT10	0.6652	RAD_R	0.9464
AMM10	0.5057	RAD_S	0.9237
K10	0.6267	RAD_DQ	0.9629
FE10	0.7971	RAD_DQ	0.9782
PH10	0.5795	TEMP	0.6699
C_NIT10	0.6146	TEMP_HM	0.9643
NIT30	0.7590	TEMP_LM	0.7423
K30	0.6255	TEMP_R	0.9014
FE30	0.7806	TEMP_S	0.7886
PH30	0.6159	TEMP_WQ	0.6746
SAND10	0.5139	TEMP_DQ	0.9457
SILT10	0.8225	RAIN	0.9413
CLAY30	0.7712	RAIN_HM	0.9072
SAND30	0.7806	RAIN_LM	0.9384
SILT30	0.8518	RAIN_R	0.8990
CLAY50	0.7259	RAIN_S	0.8932
SAND50	0.8161	RAIN_WQ	0.9203
SILT50	0.8450	RAIN_DQ	0.9505
RAD	0.9780	DEPTH_IP	0.5777
RAD_HM	0.9718	TOP_POS	0.5807
RAD_LM	0.9372	SITE INDEX	0.6380

4.2.3.2.1 THE RELATIONSHIP BETWEEN ENVIRONMENTAL ORDINATE SPACE AND PRODUCTIVITY

It should be noted that there is no requirement that productivity measures are related to the environmental ordinate space as this space merely represents the relationship between plots, based on environmental similarity, that are arbitrary references for plotting points in multidimensional space. However, productivity may be considered a function of the environment, and as such, one would expect productivity to have some relationship to this ordinate space.

The correlation between site index and its vector was 0.638. However, if site index is correlated with each individual axis of the ordinate space only SSHE2 and SSHE3 yield significant results (Table 22). Plotting site index against individual axes showed that where a relationship was evident it was linear in form.

Table 22: Pearson's correlation coefficients and their significance between the four axis of the environmental ordination and site index.

AXIS	PEARSON'S CORRELATION COEFFICIENT	SIGNIFICANCE P>X
SSHE1	0.2235	0.0978
SSHE2	-0.4829	0.0002
SSHE3	-0.5217	0.0001
SSHE4	-0.1109	0.4157

The position of plots in two dimensions (SSHE2, SSHE3) and the direction of some major vectors is given in Figure 18. Site index is seen to increase in value towards the lower left hand section of Figure 18. This is also the general direction of the SILT and CLAY vectors, some edaphic chemical vectors such as NIT and summer rainfall vectors such as RAIN_DQ. The TEMP_DQ, RAD and RAD_DQ vectors increase toward the top right hand section of the figure in an opposite direction to that of the site index.

Axes SSHE2 and SSHE3 were correlated with individual environmental variables to identify significant linear gradients (Table 23). The strongest gradients associated with SSHE2 are those of edaphic physical variables. This is particularly so for those attributes that are products of bulk densities and whose vectors gave poor correlations with the total ordination space. The strongest gradients associated with SSHE3 are those of climatic variables. Some edaphic chemical variables are correlated with both axes, although the strongest gradients for this class of variable exists for NIT10, NIT30 and FE30 along SSHE3.

Cluster groups defined in Chapter 3, were not clearly separated in four dimensional space. Given that the final configuration of plots in four dimensions is based upon environmental attributes rather than top height development patterns, this result is not surprising. However, if the hypothesis that the environment is the cause of the polymorphic top height development patterns is accepted, some separation could be expected. To pursue the hypothesis further a discriminant function was developed where plots were allocated to cluster groups on the basis of their scores on the four SSHE axes. The classification summary table for this analysis (Table 24) shows that only 46% of the plots were correctly allocated by this function.

Table 23: Pearson's correlation coefficient for the relationship between environmental ordination axis SSHE2 and SSHE3 and environmental attributes.

VARIABLE	SSHE2	(P>X)	SSHE3	(P>X)
NIT10	-0.3766	0.004	-0.5223	0.0001
AMM10	-0.4625	0.0003	-	-
K10	-0.4357	0.0008	-	-
ORG_C10	-0.4055	0.0019	-	-
PH10	-0.3148	0.0181	-	-
NIT30	-0.3378	0.0109	-0.6686	0.0001
K30	-0.4640	0.0003	-	-
ORG_C30	-	-	-0.3620	0.0061
FE30	-	-	-0.5933	0.0001
PH30	-0.3949	0.0026	-	-
C_NIT30	0.3867	0.0032	-	-
BD10	0.4399	0.001	0.4790	0.0003
SAND10	0.6988	0.0001	0.3872	0.0035
SILT10	-0.7212	0.0001	-0.4306	0.001
SIBD10	-0.6604	0.0001	-0.4621	0.0005
BD30	0.3854	0.0037	0.4479	0.0006
CLAY30	-0.5670	0.0001	-0.3088	0.0218
SAND30	0.7282	0.0001	0.5173	0.0001
SILT30	-0.6889	0.0001	-0.5790	0.0001
CBD30	-0.4398	0.0009	-	-
SABD30	0.4368	0.001	-	-
SIBD30	-0.6227	0.0001	-0.6025	0.0001
BD50	0.4567	0.006	0.4649	0.0005
CLAY50	-0.7453	0.0001	-	-
SAND50	0.8666	0.0001	-	-
SILT50	-0.7673	0.0001	-0.4828	0.0003
CBD50	-0.7094	0.0001	-	-
SABD50	0.5829	0.0001	-	-
SIBD50	-0.8414	0.0001	-0.4525	0.0008
GRBD50	0.3734	0.0059	-	-

Cont.

Table 23 (Cont.)

VARIABLE	SSHE2	(P>X)	SSHE3	(P>X)
RAD	-	-	0.8220	0.0001
RAD_HM	-	-	0.7989	0.0001
RAD_LM	-	-	0.8462	0.0001
RAD_R	0.4245	0.0011	0.4629	0.0003
RAD_S	-	-	-0.8444	0.0001
RAD_WQ	-	-	0.8746	0.0001
RAD_DQ	-	-	0.7607	0.0001
TEMP	0.4457	0.0006	-	-
TEMP_HM	-	-	0.5691	0.0001
TEMP_LM	-	-	-0.3499	0.0082
TEMP_R	-	-	0.5239	0.0001
TEMP_S	-	-	0.4071	0.0018
TEMP_DQ	0.3741	0.004	0.5880	0.0001
RAIN	-	-	-0.8276	0.0001
RAIN_HM	0.3367	0.0112	-0.6678	0.0001
RAIN_LM	-0.3658	0.0056	-0.7663	0.0001
RAIN_R	0.3777	0.0041	-0.6126	0.0001
RAIN_S	0.5177	0.0001	0.3199	0.0162
RAIN_WQ	0.3217	0.0156	-0.6961	0.0001
RAIN_DQ	-0.3131	0.018	-0.7908	0.0001
DEPTH_IP	0.3516	0.0079	-	-

Figure 18. Position of plots in two dimensions defined by SSHE2 and SSHE3 and the direction of some prominent environmental vectors. Circle size is a function of site index; increasing circle size represents increasing site index values. Vectors are of equal and arbitrary length. Vectors which appear shorter are a result of directional differences.

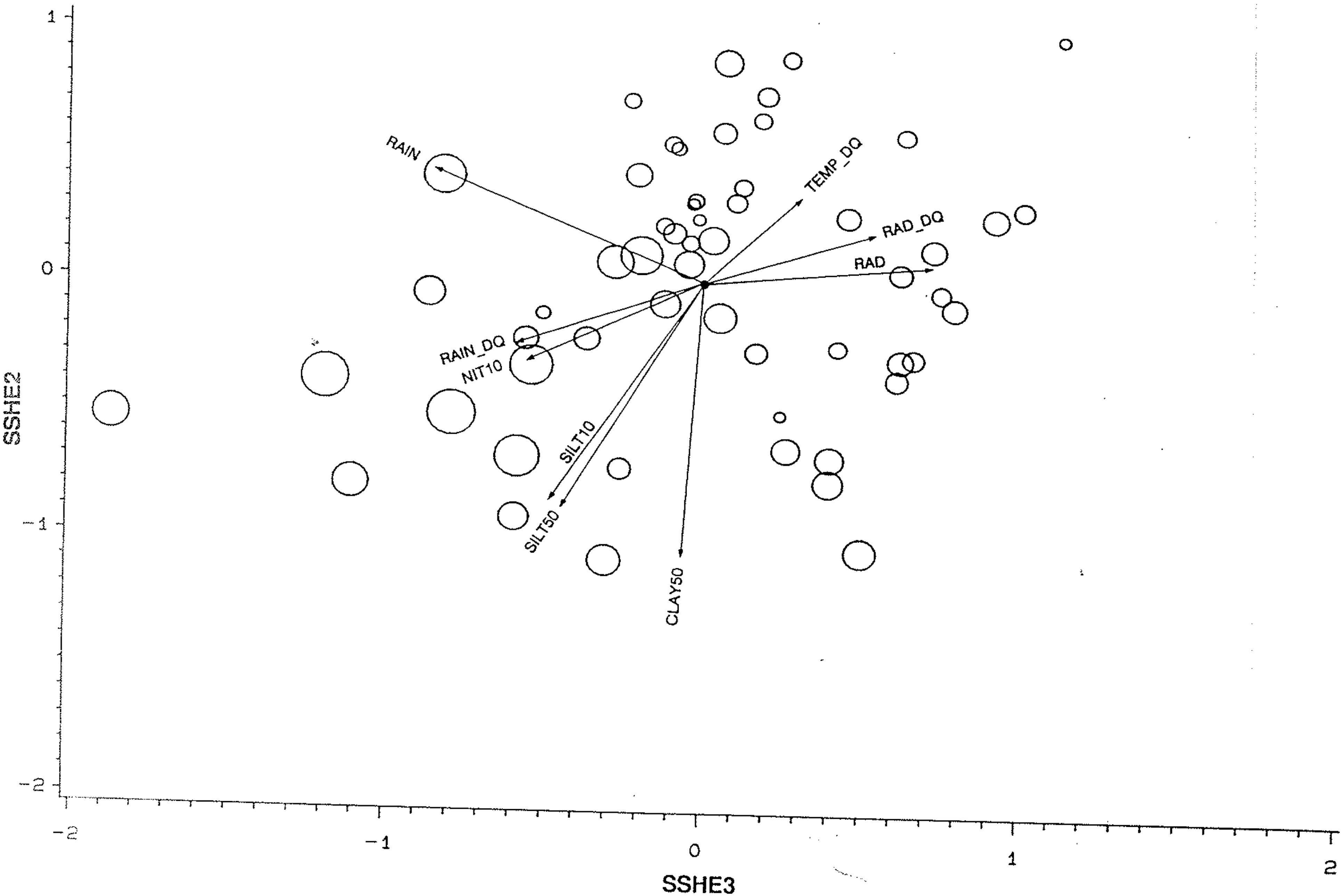


Table 24: Classification summary table from an discriminant analysis. Plots were allocated to a cluster group by discriminant function comprised of SSH environmental ordination axis.

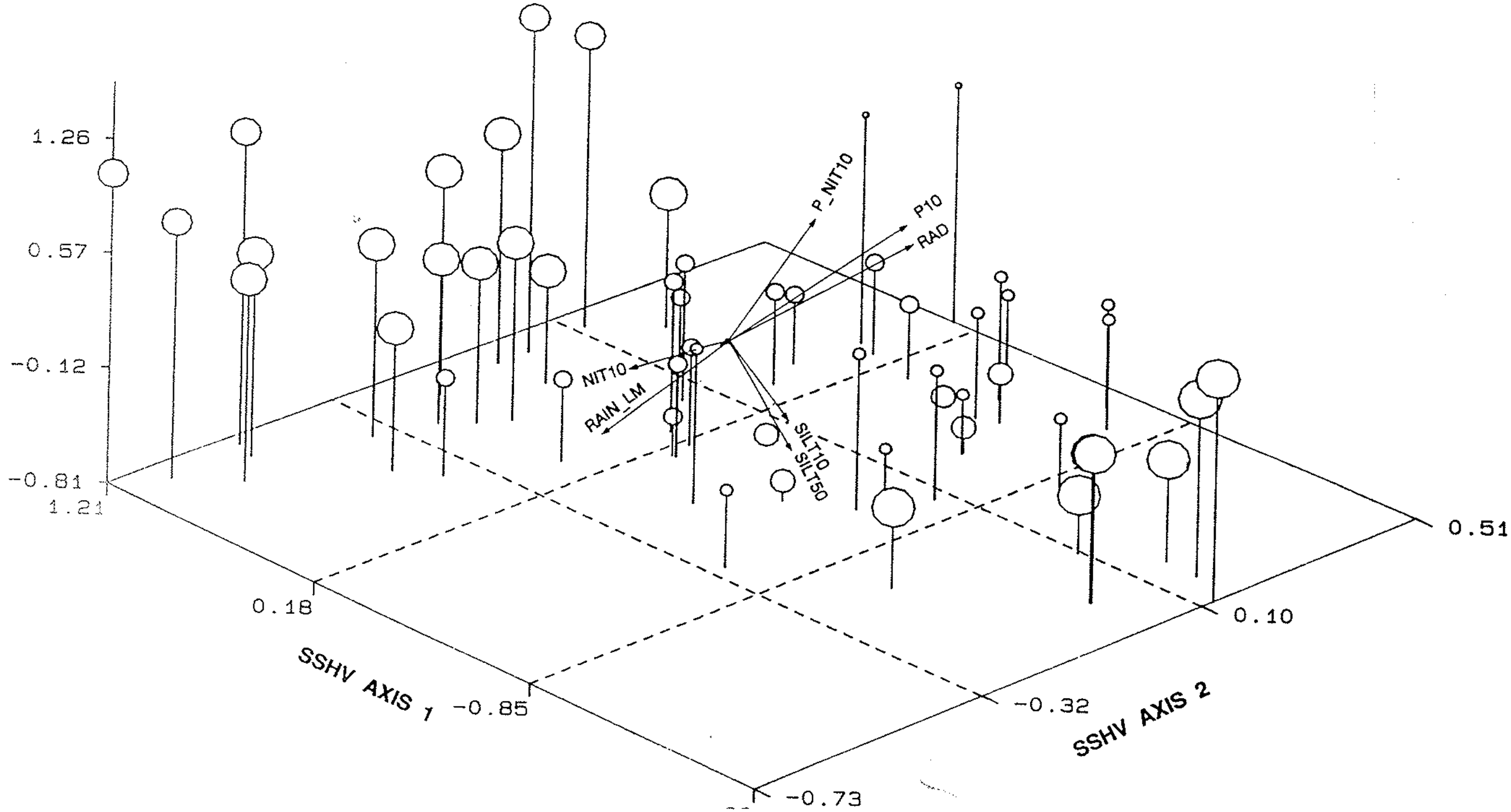
Predicted cluster group	CLUSTER GROUP 1	2	3	4	5	6	7
1	2 (100)	-	-	-	-	-	-
2	1 (7.69)	4 (30.77)	-	3 (7.69)	1 (7.69)	2 (15.38)	2 (15.38)
3	-	-	10 (79.62)	1 (7.69)	-	2 (15.38)	-
4	-	2 (40.00)	-	3 (60.00)	-	-	-
5	2 (40.00)	1 (20.00)	1 (20.00)	-	-	-	1 (20.00)
6	2 (18.18)	-	3 (27.27)	-	3 (27.27)	3 (27.27)	-
7	-	1 (14.29)	-	1 (14.29)	1 (14.29)	-	4 (12.50)

4.2.3.3 VEGETATION ORDINATION

As for the environmental ordination, the stress value and the strength of the vector correlations were assessed in each of six dimensions. On the basis of these assessments three dimensions were considered optimal for representing the information contained within the top height data (Figure 19). A maximum of 50 iterations were used to determine the final configuration of plots. Again ten randomly selected configurations were submitted as start values to guard against local solutions. The cut off dissimilarity value used for hybrid multidimensional scaling was 0.25. The final stress value of the ordination was 0.099 which is also within Kruskal's (1964) recommendations (Figure 17). Individual axes resulting from this ordination will be labelled SSHV1, SSHV2 and SSHV3 respectively.

The position of plots in three dimensional space and the directions of some of the more prominent vectors are given in Figure 19. The vector with the highest correlation is that of site index (r = 0.9736). Other prominent vectors are listed in Table 25.

Figure 19: Position of plots in three dimensions, defined by SSHV1, SSHV2 and SSHV3 and the directions of some prominent environmental vectors. Circle size is a function of site index; increasing circle size represents increasing site index values. Vectors are of equal and arbitrary length. Vectors which appear shorter are a result of directional differences.



173

Table 25: Correlations for major environmental vectors in three dimensional space defined by the vegetation ordination.

VECTOR	CORRELATION
SI	0.9736
P10	0.5805
NIT10	0.6829
P_NIT10	0.6126
NIT30	0.5938
FE30	0.5658
SILT50	0.4802
RAD	0.5000
RAD_DQ	0.5169
RAIN_LM	0.6102
RAIN_DQ	0.5617

A strong relationship exists between site index and a vector going from the top left hand corner to the lower right hand corner of Figure 19. The size of the circles in Figure 19 are proportional to site index and clearly increase in size in the same general direction as the site index vector. Site index is significantly correlated with all three SSHV axis, with Pearson's correlation coefficients for the relationships being SSHV1 r=-0.9566 (p>0.0001), SSHV2 r=0.3830 (p>0.0036) and SSHV3 r=-0.5816 (p>0.0001).

With the exception of the vector for FE30 other prominent vectors are oblique to the three SSHV axis. Only the SILT50 vector extends in the same general direction of that of site index, while other environmental vectors are oblique to these two.

Individual environmental variables were correlated with the SSHV axes to extract significant linear gradients. Strong gradients were evident for edaphic chemical variables along SSHV1, with weaker gradients evident for climatic variables. Only the FE30 gradient was significant for SSHV2, while edaphic physical variables, particularly from 50 cm depth were significant for SSHV3 (Table 26).

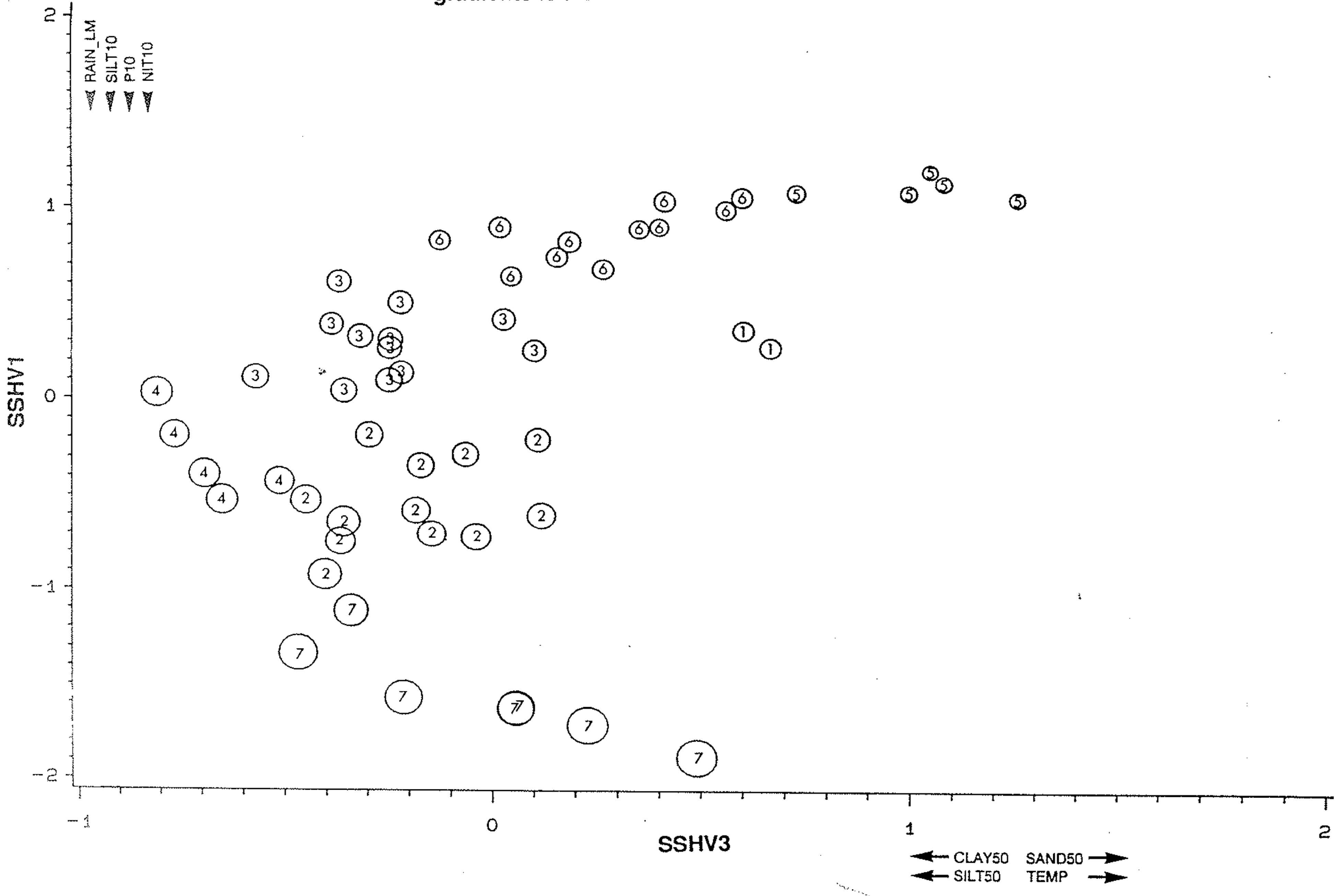
The nature of the relationships between individual environmental variables and the individual SSHV axes was examined graphically. Where a relationship was evident it was usually of a linear nature however, some variables such as the climatic variables did demonstrate curvilinear or nonlinear relationships. As such the gradients and vectors of these attributes will be deemed less prominent by the linear interpretation techniques such as Pearson's correlation and rotational correlation.

Cluster groups defined in Chapter 3 are clearly separated along SSHV1 and SSHV3. SSHV1 separates group 7 from all other groups, groups 2 and 4 from all other groups, groups 1 and 3 from all other groups and groups 5 and 6 from all other groups. SSHV3 separates groups 2 and 4, groups 3 and 1 and groups 5 and 6 (Figure 20).

Table 26: Pearson's correlation coefficient between environmental attributes and the three axes resulting from the vegetation ordination.

VARIABLE	SSHV1	(P>X)	SSHV2	(P>X)	SSHV3	(P>X)
SI	-0.9567	0.0001	0.3830	0.0036	-0.5816	0.0001
P10	-0.5210	0.0001	-	-	-	-
NIT10	-0.6056	0.0001	-	-	-	-
P_NIT10	-0.4938	0.0001	-	-	-	-
FE30	-	-	-0.3880	0.0031	-	-
P_NIT10	-0.3354	0.011	-	-	-	-
BD10	0.4318	0.0012	-	-	-	-
SAND10	0.3484	0.009	-	-	0.3424	0.01
SILT10	-0.3849	0.0037	-	-	-0.3757	0.0047
SIBD10	-0.4330	0.0012	-	-	-	-
GRBD10	0.3521	0.009	-	-	-	-
BD30	0.3912	0.003	-	-	-	-
SAND30	0.3383	0.011	-	-	0.3155	0.019
SILT30	-0.4563	0.0005	-	-	-	-
SIBD30	-0.5620	0.0001	-	-	-	-
GRBD30	0.3194	0.02	-	-	-	-
BD50	0.3732	0.005	-	-	-	-
CLAY50	-	-	-	-	-0.3892	0.004
SAND50	-	-	-	-	0.4343	0.001
SILT50	-0.4625	0.0005	-	-	-0.3471	0.01
CBD50	-	-	-	-	-0.3307	0.01
SABD50	-	-	-	-	0.4292	0.001
SIBD50	-0.4835	0.0003	-	-	-0.3171	0.02
RAD	0.4614	0.0003	-	-	-	-
RAD_HM	0.4331	0.0009	-	-	-	-
RAD_R	0.4346	0.0008	-	-	-	-
RAD_S	-0.3282	0.01	-	-	-	-
RAD_WQ	0.3585	0.006	-	-	-	-
RAD_DQ	0.4804	0.0002	-	-	-	-
TEMP	0.3445	0.009	-	-	0.4039	0.002
TEMP_HM	0.3580	0.007	-	-	-	-
TEMP_DQ	0.4511	0.005	-	-	-	-
RAIN_LM	-0.5577	0.001	-	-	-	-
RAIN_S	0.3895	0.003	-	-	-	-
RAIN_DQ	-0.5251	0.0001	-	-	-	-

Figure 20: **The position of cluster groups in two dimensions defined by SSHV1 and SSHV3. The direction of significant environmental gradients is also shown.**



4.2.4 DISCUSSION AND CONCLUSION

The application of numerical multivariate methods to ecological data has been criticised (Levin and Lewontin 1980; Salt 1983) and in some situations the assumptions inherent in these methods are violated when applied to ecological data (Austin 1976b, 1980, 1987). As such the problems encountered during the data exploration phase of this study should be noted.

Firstly, it became obvious that the shapes of some vectors and gradients were not linear. As such the linear techniques employed to interpret ordinate space, rotational correlation and Pearson's correlation, will not represent the importance of such variables fully.

Secondly, the choice of metric used to define the dissimilarity matrix will have an influence upon which gradients and vectors are extracted (Gauch 1973; Clymo 1980; Beals 1984; Faith *et al.* 1987). A number of different metrics were examined for their suitability to both data sets. For the environmental ordination the Gower metric (Gower 1971) was used, as the application of this metric assumes data to be interval (see Belbin 1989) in nature, the scale recommended for such data (Belbin pers. comm).

The Bray-Curtis metric (Bray and Curtis 1957) was applied to the unstandardised top height data after the recommendations of Faith *et al.* (1987). Although criticised by Orloci (1974) the metric was found to yield realistic results. Standardisation of the top height data had little effect upon the ordination's outcome. Although an uncommon result (Noy-Meir *et al.* 1975; Faith *et al.* 1987), it is not surprising given that the data is of the same units and from one species only. Contrary to their robust nature the Kulczynski (Hajdu 1981) and the relativised Manhattan (Sokal and Michener 1957) metrics were found to be influenced by missing data to an unacceptable degree.

As the top height data has a time series associated, the two dimensional profile algorithm (Faith *et al.* 1985) was applied to derive a dissimilarity matrix in the same manner as was done in Chapter 3. However, when the SSH algorithms were applied to the matrix, the final configuration of plots were considered less realistic than other outcomes. This is possibly due to the influence of missing data and a consequence of the small dissimilarities yielded via this algorithm.

For the purposes of this discussion nomenclature will follow Austin (1980) and Austin and Cunningham (1981) who divided environmental gradients into three types, indirect, direct and resource gradients. Indirect gradients are those whose influence on productivity is indirect such as SLOPE_PC. Direct gradients are those that exert a direct physiological effect upon productivity e.g. PH. A resource gradient is one where the factor is directly used as a resource for productivity e.g. NIT10. Austin *et al.* (1985) proposes the hypothesis that if environmental gradients are expressed as resource gradients improved predictions will result.

4.2.4.1 ENVIRONMENTAL ORDINATION

The relationship which exists between the ordinate space, defined via the environmental ordination, and site index should not be considered optimal for extracting the maximum correlation between site index and its vector. Such an ordination merely produces a configuration of plots with no reference to the productivity attributes, with the final configuration of plots in ordinate space influenced by variables which may have little or no effect upon productivity. The relationship which exists between the environmental ordinate space and productivity should be viewed as a holistic relationship rather than any optima. Although a relationship between the environment and productivity was evident, the relationship between the environment and polymorphism was not clear.

It is concluded that site index is significantly related to the ordinate space defined by environmental attributes, while cluster groups do not exhibit clear separation in the same space. Site index would seem to be a more generalist measure while cluster groups, representing both productivity and top height development patterns are less general, and specific relationships with environmental attributes may be required to yield clear separation. Attributes from all classes of variable i.e., climatic, edaphic physical and chemical were prominent when the ordinate space was interpreted, suggesting a reduced data set should contain members from each of these groups. Most gradients of significance were either direct or resource gradients.

4.2.4.2 VEGETATION ORDINATION

The site index vector has a very high correlation with the ordinate space defined from top height. This suggests that site index is a reasonable representation of top height development as a whole. Likewise, the cluster groups were clearly separated, suggesting that the polymorphism encapsulated within the clusters is not an artifact of any one numerical analysis technique.

The vegetation ordination, unlike the environmental ordination, produced a space which differed in both productivity and top height development pattern. As such, any environmental attribute whose vector has a sufficiently large correlation coefficient could reasonably be expected to exert either direct or indirect influence on productivity and/or polymorphism. Of the ten most prominent vectors, five are resource gradients. Again, no one class of variable dominates with climatic, edaphic physical and chemical attributes showing significant relationships.

The P10, NIT10, P_NIT10, NIT30 and FE30 vectors represent the directions of major fertility gradients. Both P and N are well recognised as essential elements which affect productivity (Summer and Farina 1986). The P_NIT10 vector is a reflection of the synergism which exists between these two elements. Terman *et al.* (1977) suggests that increased levels of N will increase the ability of the plant to absorb P, however, the mechanism of this interaction is not clearly understood (Summer and Farina 1986). Examination of the direction of vectors, in relation to the site index vector, and the gradients along the three SSHV axes shows that as the value of the NIT10, P10, P_NIT10 and NIT30 variables increase, so too does site index.

The FE30 variable is unlikely to be exerting direct influence upon productivity. Although Fe is classed as an essential element, the FE30 variable is more likely to be acting as a surrogate for the P fixing capabilities of the site (Ballard and Fiskell 1974; Lewis *et al.* 1981; Lewis *et al.* 1987a,b; Schwab and Kulyingyong 1989). The direction of the FE30 vector is approximately

180° to that of the P10 vector suggesting that the influence of FE30 on site index and top height development is opposite to that of P10.

These edaphic chemical vectors may be reflecting the differences in previous land use of the plots. Throughout the study area it is frequently observed that plots which are located on land which was previously pasture are more productive than plots located on land which was previously unimproved. This phenomenon is termed the pasture effect in Australia and the old-field effect in North America (c.f. Haines *et al.* 1973; Skinner and Attiwill 1981). Under a pasture regime, conditions for tree growth become more favourable as nutrients are accumulated in the soil with the addition of fertiliser (Lewis *et al.* 1987 a,b; Schwab and Kulyingyong 1989). Edaphic physical properties may also be improved through the actions of pasture crop roots (Martin 1944; Skinner and Attiwill 1981). In a comprehensive study of the pasture effect, Skinner and Attiwill (1981) conclude that the effect was not due to the changes in the microflora and fauna composition or the modification of edaphic physical characteristics but "... the effect is associated with a significant increase in the availability of soil phosphorus." Whether site index and top height development are influenced by the pasture effect to the same degree as total volume, is yet to be ascertained.

Of the edaphic physical attributes, only the SILT50 vector had a relatively high correlation coefficient. This vector probably represents the direction of the soil water storage capability gradient and reflects the degree to which the profile will dry out during the summer period. South west Western Australia experiences a Mediterranean type climate and, as such, some areas within the study area will suffer water deficits during the dry season. Fine textured soils have low hydraulic conductivities and therefore resist drying out. However, the absence of other edaphic physical vectors, particularly those expressed as bulk densities was surprising.

Although edaphic physical vectors were not as prominent as other vectors, their gradients along SSHV axes were significant. These gradients suggest that as the quantity of silt and clay increase in the profile so to does site index.

The remaining four vectors describe the direction of climatic gradients. Because of the correlation which exists between the climatic variables, discerning causal relationships is impossible, however, broad description is warranted. The direction of the RAD and RAD_DQ vectors may be viewed as the direction in which sites become progressively hotter and drier, particularly during the summer months. Site index decreases in the same direction. The directions corresponding to the RAIN_DQ and RAIN_LM vectors may be interpretated as the direction in which sites suffer less from water deficits during summer. Site index increases in this direction.

Both radiation and water deficits have been shown to influence *E. globulus* productivity (Beadle and Inions 1990). In their comparison of *E. globulus* productivity at various sites, Beadle and Inions (1990) found that increased radiation was associated with increased above ground biomass production. However, in south west Western Australia, increased radiation is also associated with decreased rainfall and higher temperatures leading to water deficits in summer.

It may be concluded from the vegetation ordination that;

(i) both site index and the polymorphism encapsulated within the cluster groups, are strongly related to the ordinate space;

(ii) that environmental variables from the climatic and edaphic physical and chemical groups are significantly related to this space and, therefore;

(iii) that both site index and polymorphism are related to environmental variables;

(iv) that half of the more prominent environmental vectors are resource gradients which supports Austin's *et al.* (1985) hypothesis.

A disadvantage associated with the data exploration techniques reported here is that no mechanism exists whereby the synergistic relationship which may exist between classes of variable, can be explored. For example, evidence suggests that there may be a synergistic relationship between the N status of a stand and water-use efficiency, due to improved stomatal control as N status increases (Brix 1972; Brix and Mitchell 1986). As a result of the synergistic and interactive relationships which may exist between variables it is prudent to use a formal modelling approach to express the relationships between productivity and environmental variables. Nonetheless, the use of such multivariate techniques has identified the nature and pattern of the relationships which exist between individual variables and site productivity. This knowledge is useful for constructing formal empirical models to predict site productivity.

4.3 RELATIONSHIPS BETWEEN SITE PRODUCTIVITY AND ENVIRONMENTAL ATTRIBUTES FOR *E.GLOBULUS* IN SOUTH WEST WESTERN AUSTRALIA. II. EMPIRICAL MODELS TO PREDICT PRODUCTIVITY.

4.3.1 INTRODUCTION

When evaluating the productive capabilities of a tract of land, using environmental attributes, the usual approach is to develop a linear regression equation to predict a measure of site productivity, such as site index. Although some success has been attained with this approach many equations are limited in:

(i) their success in accounting for a significant proportion of the productivity variation (Corns and Pluth 1984; Schmoldt *et al.* 1985; Monserud *et al.* 1990);

(ii) their applicability being restricted by the need to derive variables which require laboratory analysis (Daubenmire 1976); and

(iii) the synergistic and nonlinear nature of many of the environmental attributes are often not accounted for by linear combinations of single environmental attributes (McQuilkin 1976).

Some studies have sought to overcome some of these difficulties by including interactive terms in their equations (Jackson and Gifford 1974; Corns 1983), while Czarnowski *et al.* (1971) used nonlinear parameter estimation techniques to overcome some of the problems associated with nonlinearity. However, few alternatives to the standard approach appear in the literature. One exception is the technique of deriving a linear discriminant function to predict class membership. Classes may be site index classes (Gasana and Loewenstein 1984) or any other grouping which reflect differences in productivity (Harding *et al.* 1985).

A slightly different approach is presented by Verbyla and Fisher (1989) who argue that regression equations developed from randomly selected plots underrepresent prime sites and it is these sites only where intensive silviculture is feasible. As a result they used what they termed classification - tree analysis to derive an allocation procedure to nominate sites to prime or nonprime classes on the basis of percentage sand and soil pH.

The multivariate analysis presented in Section 4.2 has shown, that the productivity of *E.globulus* in south west Western Australian can be related to environmental variables. The most prominent variables are those which reflect the sites' soil moisture holding capabilities, such as SIBD50, the soil nutrient status, such as NIT10 and P10 and the degree to which the site experiences water stress over the summer months such as RAD, RAD_DQ, RAIN_DQ and TEMP_HM (see Table 18 for a list of variable codes). Likewise the polymorphism encapsulated within the cluster groups, defined in Section 3.2.3.3.4, may also be related to some of these site attributes. Therefore it is the aim of this study:

(i) to develop equations which predict site index from the environmental variables shown to be related to productivity in Section 4.2. Due regard is to be given to the synergistic and nonlinear relationships which may be encountered; and

(ii) develop a method to allocate plots to a top height development cluster group on the basis of environmental attributes.

4.3.2 METHODS

Details of plot selection, the productivity criteria used and the methods for deriving the environmental attributes are given in section 4.2.2.

4.3.2.1 NUMERICAL ANALYSIS

4.3.2.1.1 REGRESSION ANALYSIS

Regression analysis was used to relate site index to environmental attributes. Prior to the development of any equation the multicolinearity between attributes was examined by correlating all individual attributes with each other. Plots of site index with individual environmental attributes were also examined for nonlinear behaviour. Estimation of the parameters for the linear and nonlinear equations was by the standard criterion of ordinary least squares under the assumptions and methodologies previously detailed (see section 3.2.2.4.1). Where parameter estimates were required for nonlinear functional forms, the derivative free secant method of Ralston and Jennrich (1979) was used.

A large number of equations were developed during data analysis. The equations presented and discussed were judged to be best on the basis of the proportion of the variation in site index explained, their significance and the nature of their residual statistics. Those equations displaying heteroscedasticity among residuals were discarded. The functional forms of candidate equations were derived using the information gathered in Section 4.2, knowledge of the multicolinearity among attributes and the nature of their relationship with site index.

4.3.2.1.2 DISCRIMINANT ANALYSIS

Discriminant analysis enables the assignment of an individual or a group of objects to one of several known alternative populations, on the basis of several measurements taken from the objects (Cacoullos 1973). As such, the technique should enable the allocation of plots to one of the seven cluster groups, defined in Section 3.2.3.3.4, on the basis of environmental attributes. Application of discriminant analysis assumes that the data used to derive the discriminant functions consist of a random sample from a population mixture of multivariate normal populations and that the covariances are homogeneous across all populations (Williams 1981).

Since the seminal work on the technique in 1936 (Fisher 1936), many applications have been found (Cacoullos and Styan 1973; Williams 1981). However, with the exception of forest ecology (eg. Kuusipalo 1985; Callaway *et al.* 1987) it is only recently that the technique has been used frequently in the field of forest science. This is particularly so for the literature concerned with site evaluation. Some examples include Harding *et al.* (1985) who used the technique to predict productivity classes on the basis of environmental attributes. Turvey *et al.* (1986) also used the technique to predict volume classes from edaphic attributes, as did

Harrington (1986), who predicted site index classes from environmental attributes. In all three studies no mention is made as to whether the assumptions inherent in discriminant analysis are violated or of the consequences of any such violation. However, Gasana and Loewenstein (1984) tested the assumption that the covariances of the populations, in their case site index classes, were homogeneous and found the assumptions to be invalid. Gasana and Loewenstein (1984) then used individual population covariances when deriving the canonical function. When individual population covariances are used in this manner the technique is termed quadratic discriminant analysis. When the pooled covariance is used the technique is termed linear discriminant analysis.

Application of linear discriminant analysis to data with heterogeneous population covariances will distort the final configuration of plots in multidimensional space (Williams 1981). In such cases, quadratic discriminant analysis yields a more pronounced separation of populations (Michaelis 1973). However, the degrees of freedom required to correctly test the hypothesis of homogeneous covariances between cluster groups, and to calculate and use individual population covariances, should the hypothesis be rejected, were too few in this study.

The assumption of multidimensional normality, unlikely when dealing with ecological data (Williams 1981), is likely to be violated in this study. As Anderson and Bahadur (1962) point out, deviations from multidimensional normality may affect the results of quadratic discriminant analysis much more than those from linear discriminant analysis. Given the problems of insufficient degrees of freedom and the probable violation of the multidimensional normality assumption, linear discriminant analysis is preferred to the quadratic methodology for allocating plots to cluster groups on the basis of environmental attributes. In this case some misclassifications may occur as a result of any hetereogeneity between covariance matrices.

An alternative to discriminant analysis is logistic discrimination, where the probability of group membership is modelled. For two populations logistic discrimination was first suggested by Cox (1966) and Day and Kerridge (1967). The technique was then extended to incorporate more than two populations (Anderson 1972). The advantage of logistic discrimination is that it may be used with equal facility whether the variables are discrete or continuous and that, unlike discriminant analysis, the estimation procedure is efficient under many different assumptions about the underlying distributions (Anderson 1973).

The parameters of the logistic function were estimated with an iterative procedure. One difficulty which can arise is that non-unique maxima of the likelihood estimates occur if the plots from each population can be separated by hyperplanes. The implications of such a situation are discussed further by Anderson (1972). In this study some cluster groups may be separated by hyperplanes with the net result of nonconvergence of the iterative procedure. Therefore, heuristic allocation rules were developed, based on the identification of subgroups of clusters. Subgroups were then split into individual clusters using environmental attributes and logistic discrimination via the methods of Cox (1966) and Day and Kerridge (1967).

4.3.3 RESULTS

4.3.3.1 EQUATIONS FOR PREDICTING SITE INDEX FROM EACH ATTRIBUTE TYPE

During the data exploration phase of this study (Section 4.2) it was concluded that representatives of all data types (i.e. climatic, edaphic physical and chemical) were related to productivity. This suggests that any regression model which predicts site index would require representatives from each attribute type to explain the maximum amount of variation in site index. To examine this hypothesis further regression equations were derived for each attribute type separately.

Examination of plots of soil chemical variables with site index showed that site index was asymptotically related to P10 and NIT10. As such a log/reciprocal transformation of the data was applied. Other soil chemical variables displayed no relationship with site index. The best model (Eq. [21]) for predicting site index from soil chemical variables, along with model statistics is given in Table 27.

Plots of soil physical attributes with site index showed that where a relationship was evident it was linear in form. As such, no data transformation or nonlinear estimation techniques were required. The best model (Eq. [22]) for predicting site index from soil physical attributes along with model statistics are given in Table 27.

Plots of climatic attributes with site index also showed that where a relationship was evident, it was linear. The best model (Eq. [23]) to predict site index from climatic attributes along with model statistics is given in Table 27. When estimating the parameters of Eq. [23] plot number 60 was removed from the data set. This plot, situated in Albany (Figure 1) has climatic attribute values which are vastly different from other plots. As such the plot was prominent as an outlier during residual analysis.

Although significant ($p>0.006$) the equation to predict site index from topographic attributes only explained 21% of the variation encountered. The equation (Eq.[24]) and its statistics are given in Table 27.

4.3.3.2 EQUATIONS FOR PREDICTING SITE INDEX FROM MOISTURE, NUTRIENT AND LIGHT REGIMES

Productivity has long been assumed to be a function of the light, temperature, soil moisture and soil nutrient regimes of a site (Major 1963; Krajina 1969; Kozlowski 1982; Klinka and Carter 1990). Many expressions for soil nutrient regime and soil moisture regime have been derived (c.f. Kabzems and Klinka 1987a,b; Carter and Klinka 1990; Robertson *et al.* 1990). Soil nutrient regime may be defined as the amount and balance of essential nutrients that are available to vascular plants through root uptake over a period of several years (Klinka *et al.* 1984). Soil moisture regime is defined as the amount of water available to vascular plants through root uptake over an extended period of time (Kabzems and Klinka 1987a).

In this study the data were too limited to derive formal quantitative expressions for soil nutrient and moisture regimes, particularly if temporal variations is to be accounted for. As such equations were derived comprised of those variables which were considered to influence the availability of moisture. Soil texture reflecting ability of the soil to hold water, rainfall and rainfall distribution, reflecting the recharge capacity and temperature, reflecting the speed that which a site will dry out, were used in the equation's construction. Many combinations of these variables were found to be significant, however, the best equation (Eq. [25]) and its statistics are given in Table 27.

Equations, derived of variables which were considered to influence the availability of nutrients, were also constructed. Soil chemical and texture variables were used. Soil chemical variables reflect the fertility of the site and soil texture in the top 10cm of the profile, reflects the amount of exchange sites for nutrient absorption by the roots. Again, many combinations of these variables were significant, however, the best equation (Eq. [26]) and its statistics are given in Table 27.

Table 27: **Equations to predict site index from environmental attributes.**

Equation Identification	Equation	r^2	adj r^2	n	Standard Error	F Value
	Soil Chemical Attributes Only					
Eq. [21]	$lnS = 3.09 - 1.62/NIT10 + 1.03 \left(\frac{1}{NIT10}\right)^2 - 0.24/P10$	0.4338	0.4011	56	0.02	13.28
	Soil Physical Attributes Only					
Eq. [22]	$S = 14.59 - 3.78\ (SIBD50 + CBD50) + 23.06\ SIBD50 - 2.59\ BD10$	0.3878	0.3470	49	0.33	9.50
	Climatic Attributes Only					
Eq. [23]	$S = 79.02 + 1.02\ RAIN\_LM + 2.77\ RAD\_HM$	0.4994	0.4801	55	0.28	25.93
	Site Attributes Only					
Eq. [24]	$S = 11.85 + 0.024\ TOP\_POS - 0.023\ VEG\_CV + 13.78\ SHAPE\_S$	0.2093	0.1637	56	0.35	4.59
	Soil Moisture Attributes					
Eq. [25]	$S = 7.41 - 3.12\ (SIBD50.\ RAIN\_DQ) + 0.12\ (SIBD50.\ RAIN\_DQ.\ TEMP\_HM) - 0.0041\ RAIN + 0.59\ RAIN\_LM - 2.73\ (SIBD50 + CBD50)$	0.6406	0.6015	52	0.25	16.40
	Soil Nutrient Attributes					
Eq. [26]	$lnS = 2.88 - 1.46/NIT10 + 1.0062 \left(\frac{1}{NIT10}\right)^2 - 0.32/P10 + 0.0092\ SILT10$	0.5530	0.5172	55	0.02	15.46
	Temperature and Light					
Eq. [27]	$S = 65.75 - 1.40\ TEMP - 1.19\ RAD\_DQ$	0.2388	0.2100	56	0.35	8.31
	All Attribute Types					
Eq. [28]	$S = 7.26 + 0.00031\ (S1.S2)^2 - 5.99 \times 10^{-7}\ (S1..S2)^3$	0.7875	0.7786	51	0.19	88.93
	Combined Variables via Stepwise Algorithm					
Eq. [29]	$S = -51.50 + 0.37\ NIT10 - 0.0029\ FE30 + 0.88\ RAIN\_LM + 1.10\ ORG\_C10 + 2.75\ RAD + 0.25\ CLAY10 + 6.07\ SIBD50 - 2.58\ CBD50$	0.8378	0.8060	51	0.17	27.11
	Old Pasture sites					
Eq. [35]	$lnS = 2.84 - 1.51/NIT10 + 1.09\left(\frac{1}{NIT10}\right)^2 + 0.011\ SILT10$	0.5827	0.5305	28	0.027	11.17

Cont.

Table 27 (Continued)

Equation Identification	Equation	r^2	adj r^2	n	Standard Error	F Value
	Old Pasture Sites					
Eq. [36]	S = -83.30 + 1.25 RAIN_LM + 2.68 TEMP_HM + 42.83 SIBD50 - 2.30 CBD50 - 0.18 (SIBD50. RAIN_DQ. TEMP_HM)	0.8892	0.8601	25	0.19	30.5
Eq. [37]	S = 3.78 + 0.48 (S3) + 0.00015 $(S3.S4)^2$ - $2.83x10^{-7}$ $(S3.S4)^3$	0.9187	0.9070	25	0.16	79.06
	Old Forest Sites					
Eq. [38]	S = 10.04 + 0.12 SILT10 + 0.18 CLAY10 - 0.0026 FE30	0.3538	0.2695	27	0.36	4.20
Eq. [39]	S = 6.34 + 0.087 SILT50 + 0.29 RAIN_LM	0.3174	0.2605	27	0.36	5.58
Eq. [40]	S = -4.35 + 0.74 S5 + 0.66 S6	0.4935	0.4495	26	0.32	11.25

Where;

S1 = 7.41 - 3.12 (SIBD50.RAIN_DQ) + 0.12 (SIBD50.RAIN_DQ.TEMP_HM) - 0.0041 RAIN + 0.59 RAIN_LM - 2.73 (SIBD50 + CBD50)

S2 = e (2.88 - 1.46/NIT10 + 1.0062 / $(NIT10)^2$ - 0.32/P10 + 0.0092 SILT10)

S3 = - 83.30 + 1.25 RAIN_LM + 2.68 TEMP_HM + 42.83 SIBD50 - 2.30 CBD50 - 0.18 (SIBD50.RAIN_DQ.TEMP_HM)

S4 = e (2.84 - 1.51/NIT10 + 1.09 / $(NIT10)^2$ + 0.011 SILT10

S5 = 6.34 + 0.087 SILT50 + 0.29 RAIN_LM

S6 = 10.04 + 0.02 SILT10 + 0.18 CLAY10 - 0.0026 FE30

Variables expressing the temperature and light were combined and the best equation (Eq. [27]) is given in Table 27. If the hypothesis, that productivity is a function of the soil nutrient and moisture regime is accepted then an equation comprised from a selection of all attribute types should display model statistics which are more desirable that the individual equations Eq. [21] to Eq. [27]. Four such equations are presented. Equation [28] and its statistics is presented in Table 27. Three nonlinear equations (Eqs. [30], [31], [32]) are shown in Table 28. Equations [28] and Eqs. [30], [31], and [32] are not directly comparable due to the different parameter estimation techniques employed in their construction.

Table 28: **Nonlinear equations to predict site index from environmental attributes.**

Equation Identification	Equation	Asymptotic Confidence Interval (95%) B_0			Asymptotic Confidence Interval (95%) B_1			n	STD ERR	Residual variance	Residual mean
Eq [30]	S = S1.S2/0.03.S1.S2 + 7.57	0.023	-	0.036	6.43	-	8.71	51	0.20	1.99	0.004
Eq [31]	S = $24.77\ e^{-100.25/S1.S2}$	22.18	-	27.37	-116.72	-	-83.78	51	0.21	2.18	0.014
Eq [32]	S = $0.73\ (S1.S2)^{0.56}$	0.41	-	1.045	0.48	-	0.65	51	0.20	2.05	-0.005
Eq [33]	S = S1.S2/0.03.S1.S2 + 7.70	0.025	-	0.036	6.52	-	8.89	25	0.19	0.91	-0.0034
Eq [34]	S = $26.08\ e^{-114.91/S1.S2}$	23.74	-	28.42	-132.57	-	97.25	25	0.19	0.92	0.0034

Where;

S1 = 7.41 - 3.12 (SIBD50.RAIN_DQ) + 0.12 (SIBD50.RAIN_DQ.TEMP_HM) - 0.0041 RAIN + 0.59 RAIN_LM - 2.73 (SIBD50 + CBD50)

S2 = e (2.88 - 1.46/NIT10 + 1.0062 / $(NIT10)^2$ - 0.32/P10 + 0.0092 SILT10)

S3 = - 83.30 + 1.25 RAIN_LM + 2.68 TEMP_HM + 42.83 SIBD50 - 2.30 CBD50 - 0.18 (SIBD50.RAIN_DQ.TEMP_HM)

S4 = e (2.84 - 1.51/NIT10 + 1.09 / $(NIT10)^2$ + 0.011 SILT10

4.3.3.3 AN EQUATION FOR PREDICTING SITE INDEX USING A STEPWISE ALGORITHM

The practice of submitting all environmental variables into a variable selection algorithm such as stepwise (Hocking 1976) has become common place in the literature concerned with predicting productivity from environmental variables (eg. Hamilton and Krause 1985; Schmidt and Carmean 1988; Verbyla and Fisher 1989). For comparative purposes a stepwise variable selection technique was used to derive Eq. [29] (Table 27).

4.3.3.4 EQUATIONS FOR PREDICTING SITE INDEX WHEN PREVIOUS LAND USE IS CONSIDERED

To examine the effect of the previous land use on the parameter estimates and variables selected, the data were divided into plots established on unimproved land and plots established on previously agricultural pasture. For both data sets equations were derived from variables which were considered to influence the availability of (a) nutrients and (b) moisture. Again equations derived from all attribute types were constructed for each data set. For plots established on what was agricultural lands, the equations comprised of variables reflecting nutrient availability (Eq. [35]) moisture availability (Eq. [36]) and all attribute types (Eq. [37]) and their statistics are presented in Table 27. Nonlinear forms are given in Table 28. For plots established on unimproved lands, the equations comprised of variables reflecting nutrient availability (Eq. [38]), moisture availability (Eq. [39]) and all attribute types (Eq. [40]) and their statistics are given in Table 28.

4.3.3.5 PREDICTION OF CLUSTER GROUP MEMBERSHIP USING ENVIRONMENTAL ATTRIBUTES AND LINEAR DISCRIMINANT ANALYSIS

To further identify which variables possess discriminatory powers, a one-way analysis of variance was employed to test for differences between the mean values of each environmental attribute, for each top height development cluster group defined in section 3.2.2.4.2. The means and their standard errors, by cluster group, are given in Table 29. Only variables whose F values were significant at probability levels greater than p=0.01 are included in the table. As such soil physical attributes sampled from 50 cm depth are noticeably absent. However, the means of these variables were significantly different ($p>0.02$).

Using environmental attributes only, the most successful discriminant function derived, correctly classified 78.85% of plots. This discriminant function and its classification table are given in Table 30. A discriminant function comprised of RAIN_LM, BD10, P10, RAD and site index correctly classified 86.79% of plots and is presented in Table 31, along with its classification table. Linear discriminant analysis using site index only, correctly classified 69.64% of plots while a discriminant function comprised of P10, BD10, RAIN_LM and RAD only, correctly classified only 52.83% of plots. This demonstrates the discriminating power of site index. However, site index in isolation does not provide an adequate discrimination of cluster groups. For maximum separation a combination of site index and environmental attributes is required (Table 31).

Table 29: Univariate statistics for environmental attributes for each cluster group (Mean ± Standard Error) .Only variables with significant differences between means (P>0.01) shown. Significance level given in parenthesis.

Variable		Cluster group 1	2	3	4	5	6	7
		n = 2	n = 13	n = 13	n = 5	n = 5	n = 11	n = 7
BD10	(P>0.001)	0.97 ± 0.21	1.10 ± 0.06	1.41 ± 0.07	1.00 ± 0.02	1.08 ± 0.12	1.41 ± 0.12	0.97 ± 0.06
GRBD10	(P>0.0008)	0.16 ± 0.14	0.26 ± 0.08	0.87 ± 0.10	0.24 ± 0.07	0.27 ± 0.21	0.67 ± 0.20	0.19 ± 0.07
BD30	(P>0.002)	1.24 ± 0.12	1.30 ± 0.09	1.67 ± 0.09	1.39 ± 0.04	1.24 ± 0.22	1.69 ± 0.08	1.22 ± 0.06
SIBD30	(P>0.005)	0.09 ± 0.08	0.11 ± 0.02	0.07 ± 0.01	0.13 ± 0.02	0.07 ± 0.03	0.06 ± 0.06	0.17 ± 0.03
GRBD30	(P>0.001)	0.22 ± 0.20	0.30 ± 0.10	1.09 ± 0.11	0.52 ± 0.10	0.38 ± 0.29	0.80 ± 0.19	0.34 ± 0.11
P10	(P>0.001)	6.00 ± 2.00	7.25 ± 2.25	9.05 ± 2.32	16.12 ± 3.90	2.94 ± 0.66	4.68 ± 0.96	41.52 ± 16.69
NIT10	(P>0.0001)	1.35 ± 0.35	1.67 ± 0.18	1.56 ± 0.14	3.20 ± 1.04	1.96 ± 0.29	1.48 ± 0.13	7.17 ± 1.18
C_NIT10	(P>0.002)	1.58 ± 0.79	1.76 ± 0.22	2.01 ± 0.20	1.44 ± 0.44	1.04 ± 0.13	1.90 ± 0.30	0.48 ± 0.06
P_NIT10	(P>0.001)	7.35 ± 0.65	15.08 ± 6.03	15.45 ± 4.79	57.14 ± 28.76	6.20 ± 2.30	6.94 ± 1.69	340.38 ± 169.51
P_AMM10	(P>0.002)	24.65 ± 7.35	56.41 ± 17.73	58.4 ± 15.67	147.46 ± 71.49	20.12 ± 5.21	28.40 ± 7.53	343.86 ± 149.01
RAD	(P>0.0021)	17.42 ± 0.32	16.84 ± 0.21	17.65 ± 0.15	17.37 ± 0.30	17.39 ± 0.43	17.33 ± 0.13	16.39 ± 0.13
RAD_HM	(P>0.0044)	27.85 ± 0.47	27.39 ± 0.24	28.18 ± 0.18	27.84 ± 0.34	27.77 ± 0.58	27.81 ± 0.18	26.69 ± 0.14
RAD_DQ	(P>0.0016)	26.86 ± 0.44	26.06 ± 0.25	26.96 ± 0.16	26.62 ± 0.32	26.59 ± 0.55	26.70 ± 0.16	25.47 ± 0.15
RAIN_LM	(P>0.0001)	11.21 ± 0.48	16.34 ± 1.17	12.63 ± 0.76	13.68 ± 1.41	13.99 ± 2.63	13.41 ± 0.52	20.87 ± 0.56
RAIN_DQ	(P>0.0003)	44.57 ± 0.90	62.39 ± 4.78	46.43 ± 2.49	50.87 ± 5.42	49.85 ± 9.37	50.59 ± 2.55	75.59 ± 3.96
S	(P>0.0001)	9.95 ± 0.15	13.82 ± 0.34	11.85 ± 0.17	14.58 ± 0.24	7.8 ± 0.22	9.88 ± 0.17	18.07 ± 0.39

Table 30: Discriminant function and its classification table for allocating plots to cluster groups on the basis of environmental attributes.

Actual Group	Predicted Group						
	1	2	3	4	5	6	7
1	2 (100.00)						
2	1 (7.69)	9 (69.23)		2 (15.38)		1 (7.69)	
3			9 (81.82)	1 (9.09)		1 (9.09)	
4		1 (20.00)		4 (80.00)			
5					3 (100.00)		
6	1 (9.09)	1 (9.09)	2 (18.18)			7 (63.64)	
7							7 (100.00)

Variable	Discriminant function						
Constant	-5761	-6030	-6096	-6072	-5982	-5983	-6228
NIT10	-15.59	-16.20	-15.90	-15.46	-15.33	-15.83	-13.37
GRBD30	-8.65	-15.08	-6.58	-12.02	-3.69	-11.11	-11.11
RAIN_LM	112.35	115.51	115.72	115.61	114.22	114.82	117.60
FE30	0.007	0.007	0.007	0.005	0.017	0.009	0.011
ORG_C10	62.56	66.67	66.80	68.02	60.53	65.71	66.02
BD30	-28.86	-16.48	-24.76	-19.60	-41.09	-18.09	-19.61
RAD	586.37	598.23	602.25	600.44	598.45	596.22	607.18
SILT50	-1.70	-1.58	-1.69	-1.59	-2.02	-1.75	-1.71

Table 31: Discriminant function and its classification table for allocating plots to cluster groups on the basis of environmental attributes and site index.

Actual Group	Predicted Group						
	1	2	3	4	5	6	7
1	2 (100.00)						
2		9 (69.23)	2 (15.38)	2 (15.38)			
3			13 (100.00)				
4		1 (25.00)		3 (75.00)			
5					5 (100.00)		
6	2 (22.22)					7 (77.78)	
7							7 (100.00)

Variable	Discriminant Function						
constant	-4291	-4436	-4494	-4482	-4503	-4466	-4602
P10	-0.78	-0.79	-0.79	-0.78	-0.78	-0.78	-0.69
S	-1.24	4.33	1.63	5.01	-4.76	-1.21	9.89
BD10	-24.25	-15.27	-14.36	-18.08	-24.48	-16.09	-10.55
RAIN_LM	78.08	79.19	79.82	79.26	80.58	79.62	80.22
RAD	444.64	447.76	452.56	449.99	456.81	451.09	450.82

4.3.3.6 SEPARATION OF CLUSTER GROUPS VIA HEURISTIC RULES AND LOGISTIC REGRESSION

As initial attempts to fit logistic discrimination functions failed to converge, alternative hierarchic methods were employed. The high discriminatory power of site index was utilized such that a 100% classification of the data was achieved by simply splitting the site index range appropriately, such that:

$S > 16.4m$	=	cluster group 7
$10.9m < S < 16.3m$	=	cluster groups 2, 3 or 4
$8.7m < S < 10.8m$	=	cluster groups 1 or 6
$S < 8.6m$	=	cluster group 5

No heuristic rule could be derived to separate cluster groups 6 and 1. Within cluster subgroup (2,3,4), cluster group 4 has site index values which range from 14.0 m to 16.3 m, while cluster group 3 has values between 11.0 m and 12.7 m. Thus, complete discrimination of cluster groups 3 and 4 can be attained if they can be separated from cluster group 2.

A logistic regression was carried out to separate cluster group 2 and subgroup (3,4). The only significant variables were S, NIT10 and GRBD50. The estimated probability of belonging to cluster group 2 is given by,

$$p = \frac{e^{-7.356 + 0.892S - 2.053NIT10 - 2.952GRBD50}}{1 + e^{-7.356 + 0.892S - 2.053NIT10 - 2.952GRBD50}} \qquad [41]$$

If the estimated p exceeds 0.5, then the plot is estimated to be more likely to belong to cluster group 2. Using this rule on the 29 plots in subgroup (2,3,4), for which there was no missing data, the function correctly allocates 8 of the 12 plots in cluster group 2, and 16 of the 17 in subgroup (3,4). Since all of cluster group 3 will be discriminated from all of cluster group 4 using the site index range, this rule misallocates seven of the 56 (87.50%) plots. Cluster group 1 is assumed to be misallocated.

4.3.4 DISCUSSION

4.3.4.1 EQUATIONS FOR THE PREDICTION OF SITE INDEX

When only those equations which were derived from individual attribute types are considered, the equation comprised of climatic attributes accounted for the most variation in site index (48%). The two attributes comprising this equation RAIN_LM and RAD_HM both reflect the degree to which the stand will suffer water stress during the summer months. Given the very seasonal nature of south west Western Australia's mediterranean climate this is not an unexpected result. Equations comprised of soil chemical and soil physical variables only account for 40% and 35% of the variation in site index respectively. No significant interactive terms were identified for any of the three equations suggesting that multiplicative interaction occurs between attribute types rather than within. This result is unexpected given the well documented interactions which occur between nutrients (Terman *et al.* 1977; Summer and Farina 1986). Formal experimentation is required to identify the nature and form of such interaction prior to submitting such a term to a linear regression equation.

It is unlikely that Eqs. [21], [22], [23], or [24] possess enough precision of estimation to be useful for predicting site index. The aim of deriving such equations was to examine the hypothesis that productivity is not merely due to one attribute type alone. This is clearly not the case as the equations which predict site index from variables reflecting soil moisture and soil nutrients account for 60% and 52%, respectively, of the variation in site index. Eq. [25] also contains multiplicative interactive terms. Again the equation which reflects the influence of moisture stress explains the greater proportion of the site index variability. The equation which represents the temperature and light regime Eq. [27] only accounts for 21% of the variation in site index. It is unlikely that temperature and radiation are as limiting as nutrient and moisture in south west Western Australia (Beadle and Inions 1990). The significance of

TEMP and RAD_DQ in this equation is probably due to their correlation with other climatic variables rather than a meaningful result.

Equations comprised of attributes which reflect both the available nutrients and moisture account for larger proportions of the variability in site index (78% - 91%) than single attribute equations, or the equations which reflect soil moisture or nutrient regimes alone. With the exception of Eq. [29] all such equations require a multiplicative interaction term for maximum explanation of the variation.

The equation derived via the stepwise procedure (Eq. [29]) accounts for 81% of the variation in site index. However, its use is subject to concern. It is the most parameterised model and as such may be subject to prediction bias (Verbyla 1986). Also, the linear combination of attributes, with no provision for their synergistic nature, casts doubt on its extrapolative qualities as the selection of attributes, via a stepwise procedure, may include biologically insignificant attributes (Verbyla and Fisher 1989).

On the basis of model statistics and biological realism it is recommended that Eqs. [28], [30], [31] or [32] are the most appropriate for use when estimating site index from environmental variables. The proportion of the variability in site index accounted for by these equations is high given the size of the study area and the range of environmental attributes spanned (c.f. Hunter and Gibson 1984; Saunders *et al.* 1984; Buckley 1988; Monserud *et al.* 1990) however, these results are not unique (c.f. Page 1976; Brown and Loewenstein 1978). The only other study of this nature which used *E. globulus* as its subject was undertaken by Gasana and Loewenstein (1984) in Rwana. They produced an equation which accounted for 71% of the variation in site index.

Given the size and ecological complexity of the study area it is not surprising to find that the most prominent attributes are those which reflect the moisture and nutrient regimes and that these factors are interdependent (McQuilkin 1976). The most prominent attributes include NIT10, P10, SIBD50, CBD50, BD10, RAIN_LM, RAIN_DQ, TEMP_HM and RAD_HM.

Four of these variables (NIT10, P10, TEMP_HM and RAD_HM) may be viewed as resource gradients. Again the hypothesis as proposed by Austin *et al.* (1985), that using environmental gradients which are classed as resource gradients improved predictions, holds true.

Moisture is obviously critical to tree growth (Kozlowski 1982) however, identifying a variable to express moisture availability to the tree is extremely difficult (Broadfoot 1969). It is particularly difficult to express such an attribute as a resource gradient. The SIBD50 and CBD50 are variables which effect moisture availability through their influence upon the soil's ability to resist drying. The significance of moisture, as expressed by SIBD50 and CBD50 is common in such studies (Page 1976) but not universal (Daubenmire 1968; Monserud *et al.* 1990). The climatic variables are merely reflecting the degree to which soil water stores will be recharged and the severity of the summer water stress.

The attributes discussed here were also prominent during the data exploration phase of this study (see section 4.2).

4.3.4.2 DISCRIMINATION OF CLUSTER GROUPS

For allocating plots to a cluster group it is recommended that the discriminant function comprised of P10, BD10, RAIN_LM, RAD and site index or the heuristic rules detailed in section 4.3.3.6 are the more suitable. The discriminant function comprised of environmental attributes only contains eight attributes and as such, may be subject to prediction bias (Verbyla 1986).

Although the use of discriminant functions, for allocating plots to productivity classes, is becoming common (Gasana and Loewenstein 1984; Harding *et al.* 1985; Harrington 1986; Turvey *et al.* 1986) this study is the first to use the technique as a means of dealing with polymorphism. The technique, as applied in this study, truncates the requirement for two height measurements when applying polymorphic nondisjoint top height development and site index equations. A philosophically similar approach is presented by Monserud (1984) who incorporated habitat type into top height development equations such that the top height development pattern varied with each habitat type. In this study the use of vegetation, or any other classification unit was not available to employ such an approach. As such an allocation methodology was required based upon the causes, or at least variables related to the causes, of polymorphism so that the appropriate top height development and site index equations could be selected (see Section 3.2).

The discriminant analysis and logistic regression exercise emphasises the utility of site index. It is shown to be a major discriminant of cluster groups and the polymorphism contained therein. However, for maximum separation of cluster groups site index and environmental attributes are required.

Whether the use of the discriminant procedures, for separating plots into cluster groups prior to predicting top height development, yields greater accuracy and precision over standard techniques such as the use of Eq.[11] will be the topic of Chapter 5.

4.3.4.3 EFFECT OF PREVIOUS LAND USE

When the dataset is split, on the basis of the previous land use, marked differences in model statistics become obvious. The equation fitted to data derived from land that was previously pasture (Eq. [37]) accounts for 91% of the variability in site index, compared to only 45% for the equation fitted to data derived from land that was previously unimproved (Eq. [40]). The differences in productivity associated with the history of land use are well documented, with pasture sites the more productive (Haines *et al.* 1973; Skinner and Attiwill 1981). Likewise, the differences in soil properties between land use histories are also well documented (Adejuwon and Ekanade 1988).

The differences in productivity between land use history are usually attributed to nutrient accumulation through fertilisation (Lewis *et al.* 1987a,b), nutrient enrichment through the activity of grasses and legumes (Williams 1962; Donald 1970) and the physical structure of the soil becoming enhanced by the activities of pasture species roots (Martin 1944). However, this does not explain the disparity in model statistics between the two land use histories. One possible explanation is that soil properties are more homogeneous under pasture and therefore the error incurred when estimating nutrient levels will be smaller. However, it is unlikely that such a disparity is due to this alone. Another more feasible explanation is that, due to the modified soil environment under a pasture regime (Dalal and Mayer 1986a,b,c; Lewis *et al.* 1987a,b), many of the limiting factors, which may be influential under forest sites, are removed. The nett effect would be that less noise would exist in the pasture data set and productivity becomes more predictable.

4.3.4.4 SOURCES OF ERROR AND BIAS

The equations derived are subject to errors of varying magnitude and therefore warrant discussion. Firstly, the linear and nonlinear equations are not directly comparable due to the different parameter estimation procedures employed. Also, where nonlinear parameter estimation techniques were used, estimates of the variance and therefore confidence intervals may be biased. A further area of concern when using regression is prediction bias (Verbyla 1986). Where many independent variables are screened for relationships with site index the probability of chance correlations is ($1-p^n$), where P is the probability level of significance and n is the number of variables screened. In studies such as this, many attributes are screened with the very real chance that erroneous correlations will occur. The effect of such bias is well documented with equations exhibiting desirable model statistics but not performing at all well when validation occurs (McQuilkin 1976; Verbyla 1986). This situation is of particular concern when a stepwise procedure is used (Brandt 1970). The strategy of undertaking multivariate data exploration prior to model formulation assisted in identifying the nature and pattern of relationships, prior to model formulation. The chance of constructing models with prediction bias is reduced using this strategy.

Secondly, the mathematical expression for the synergistic relationships between attributes was necessarily coarse. The intercorrelations and interactions among environmental attributes made the development of complex interactive expressions impossible. Formal experimentation would be required to define the nature and form of such interaction.

A third weakness stems from the inability to measure attributes which directly influence productivity. Instead, most factors are inferred indirectly from secondary expressions such as SIBD50 etc. This is particularly so when deriving expressions for soil nutrients where the method of extracting the element from the soil sample may influence whether or not it is related to productivity (Keeney 1980; Powers 1980). Likewise, the variables expressing climate used in this study were derived from a set of equations. As such these attributes should be viewed as indices rather than the climate actually experienced at the site.

A fourth source of error stems from the spatial and temporal variation of the soil medium. The temporal variation in soil nutrients was not accounted for in this study. With the exception of Page (1976) this is standard practice in such studies. However, soil chemistry is known to change overtime under forest canopy (France *et al.* 1989; Billet *et al.* 1990) introducing a sampling error of undefined magnitude. In this study, where the rotation is short (10 years) this error is assumed to be small. A further error is incurred as a result of the spatial heterogeniety of the soil (Usher 1970; Blyth and MacLeod 1978; Keeney 1980). Again, the sampling error incurred as a result of the spatial heterogeniety is assumed to be small in this study (see section 4.2.3.1).

A fifth cause of unexplained variation in site index may result from not accounting for all the factors which exert an influence upon productivity. For example, no measure of genetic variability was used in this study, yet genotype has been shown to be about a third more important than the environment in determining phenotypic variation in dominant height of Douglas-fir in the U.S.A. (Monserud *et al.* 1990; Monserud and Rehfeldt 1990). However, in this study the subject species is an exotic to south west Western Australia, and the genetic complexity between environment and genotype evident in Monserud and Rehfeldt's (1990) study is unlikely to be as influential.

A final source of error may result from the small number of sample plots. Fifty six plots is towards the lower end of range used in such studies.

Although the error sources listed above are of real concern, the variation in site index accounted for by the equations is towards the upper end of the published range for such studies. However, the utility, accuracy and precision of such equations is best assessed by validation on independent data, which is the topic of Chapter 5.

4.3.5 CONCLUSION

It may be concluded from this study that:

(i) both site index and the polymorphism in top height development are related to environmental attributes;

(ii) the maximum explanation of the variation in site index is achieved in functions comprised of all variable types i.e., climatic, edaphic chemical and edaphic physical, and interactive terms are required;

(iii) large amounts of the variation in site index were explained by Eqs. [28], [29], [30], [31] and [32];

(iv) cluster groups, representing the polymorphic nature of the top height development patterns, may be separated by heuristic rules and logistic regression or discriminant functions. In either case the procedures require the input of both site index and some environmental attributes;

(v) the use of multivariate data exploration techniques was useful for identifying the nature and patterns which exist between productivity and environmental attributes. Such knowledge is useful when formulating the functional forms of predictive equations or allocation rules; and,

(vi) no recommendations will be made as to which predictive equation or allocation procedure is the most appropriate until after the equations have been validated, which is the subject of Chapter Five.

4.3.6. REFERENCES

Abrams, M.C. 1985. Age-diameter relationships of Quercus species in relation to edaphic factors in gallery forests in north east Kansas. *For. Ecol. Manage.*, 13:181-193.

Adejuwon, J.O., and O. Ekanade. 1988. A comparison of soil properties under different landuse types in a part of the Nigerian cocoa belt. *Catena*, 15:319-331.

Albert, D.A., and B.V. Barnes. 1987. Effects of clearcutting on the vegetation and soil of a sugar maple-dominated ecosystem, western upper Michigan. *For. Ecol. Manage.*, 18:283-298.

Anderson, J.A. 1972. Separate sample logistic discrimination. *Biometrika*, 59:19-35.

Anderson, J.A. 1973. Logistic discrimination with medical applications. *In*: Discriminant analysis and applications. T. Cacoullos (ed.). Academic press, New York.

Anderson, T.W., and R.R. Bahadur. 1962. Classification into two multivariate normal distributions with different covariance matrices. *Ann. Math. Stat.*, 33:420-431.

Auchmoody, L.R., and H. Clay-Smith. 1979. Oak soil-site relationships in north western West Virginia. *USDA. Res. Pap.* NE-434.

Austin, M.P. 1976a. Performance of four ordination techniques assuming three different non-linear species response models. *Vegetatio*, 33:43-49.

Austin, M.P. 1976b. On non-linear species response models in ordination. *Vegetatio*, 33:43-49.

Austin, M.P. 1980. Searching for a model for use in vegetation analysis. *Vegetatio*, 42:11-21.

Austin, M.P. 1985. Continuun concept, ordination methods, and niche theory. *Ann. Rev. Ecol. Syst.*, 16:39-61.

Austin, M.P. 1987. Models for the analysis of species' response to environmental gradients. *Vegetatio*, 69:35-45.

Austin, M.P., and I. Noy-Meir. 1971. The problem of non-linearity in ordination: experiments with two gradient models. *J. Ecol.*, 59:762-773.

Austin, M.P., and P. Greig-Smith. 1968. The application of quantitative methods to vegetation survey. II. Some methodological problems of data from rainforest. *J. Ecol.*, 56:827-844.

Austin, M.P., and R.B. Cunningham. 1981. Observational analysis of environmental gradients. *Proc. Ecol. Soc. Aust.*, 11:109-111

Austin, M.P., R.B. Cunningham, and P.M. Flemming. 1985. New approaches to direct gradient analysis using environmental scalors and statistical curve-fitting procedures. *In*: Plant community ecology: Papers in honour of Robert H. Whittaker. R.K. Peet (Ed.). *Advances in vegetation science* No. 7.

Auten, J.T. 1945a. Some soil factors associated with site quality for planted black locust and black walnut. *J.For.*, 43:592-598.

Auten, J.T. 1945b. Prediction of site index for yellow poplar from soil and topography. *J.For.*, 43:662-668.

Ballard, R., and J.G.A. Fiskell. 1974. Phosphorus retention in coastal plain forest soils. I. Relationship to soil properties. *Proc. Soil. Sci. Soc. Am.*, 38:250-255.

Beadle, C.L., and G.B. Inions. 1990. Limits to growth of Eucalypts and their biology of production. *In*: Prospects for Australian plantations. J. Dargavel, and N. Semple (eds.). ANU press. Canberra, A.C.T., Australia.

Beals, E.W. 1984. Bray-Curtis ordination: an effective strategy for analysis of multivariate ecological data. *Adv. Ecol. Res.*, 14:1-55.

Belbin, L. 1989. PATN: Pattern analysis package. CSIRO Div. Wildl. Ecol. Canberra. Australia.

Belbin, L. *In press*. SSH: semi-strong hybrid scaling. *J. Classification*.

Billett, M.F., F. Parker-Jervis, E.A. Fitzpatrick, and M.S. Cresser. 1990. Forest soil chemical changes between 1949/50 and 1987. *J.Soil. Sci.*, 41:133-145.

Blake, G.R., and K.H. Hartge. 1986. Bulk density. *In*: Methods of soil analysis, Part 1. Physical and mineralogical methods. *Agron. Mono.*, No. 9.

Blyth, J.F., and D.A. MacLeod. 1978. The significance of soil variability for forest soil studies in north-east Scotland. *J. Soil Sci.*, 29:410-430.

Blyth, J.F., and D.A. MacLeod. 1981. Sitka spruce (*Picea sitchensis*) in North East Scotland II. Yield prediction by regression analysis. *Forestry*, 54:63-73.

Booth, T.H. 1978. Numerical classification techniques applied to forest tree distribution data. I. A comparison of methods. *Aust. J. Ecol.*, 3:297-306.

Booth, T.H., H.A. Nix, M.F. Hutchinson, and T. Jovanovic. 1988. Niche analysis and tree species introduction. *For. Ecol. Manage.*, 23:47-59.

Bowman, D.M.J.S., and P.R. Minchin, 1987. Environmental relationships of woody vegetation patterns in the Australian monsoon tropics. *Aust. J. Bot.*, 35:151-169.

Brand, D.G., and P.S. Janas. 1988. Growth and acclimation of planted white pine and white spruce seedlings in response to environmental conditions. *Can. J. For. Res.*, 18:320-329.

Brandt, E.B. 1970. The abuse of stepwise regression or: experiments in fishing. *Fand. Corp. publ.* AD 701804.

Bray, J.R., and J.T. Curtis. 1957. Ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.*, 27:325-34978.

Brix, H. 1972. Nitrogen fertilization and water effects on photosynthesis and earlywood-latewood production in Douglas-fir. *Can. J. For. Res.*, 2:467-478.

Brix, H., and A.K. Mitchell, 1986. Thinning and nitrogen fertilization effects on soil and tree water stress in a Douglas-fir stand. *Can. J. For. Res.*, 16:1334-1338.

Broadfoot, W.M. 1969. Problems in relating soil to site index for southern hardwoods. *For. Sci.*, 15:354-364.

Brown, H.G., and H. Loewenstein. 1978. Predicting site productivity of mixed conifer stands in northern Idaho from soil and topographic variables. *Soil Sci. Soc. Am. J.*, 42:967-971.

Buckley, G.P. 1988. Soil factors influencing yields of *Eucalyptus camaldulensis* on former tin-mining land in the Jos Plateau region, Nigeria. *For. Ecol. Manage.*, 23:1-17.

Busby, J.R. 1986. A bioclimatic analysis of *Nothofagus cunninghamii* (Hook.) Oerst. in south-eastern Australia. *Aust. J. Ecol.*, 11:1-7.

Cacoullos, T. 1973. Introductory address. *In*: Discriminant analysis and applications. T. Cacoullos (ed.). Academic press, New York.

Cacoullos, T., and G.P.H. Styan. 1973. A bibliography of discriminant analysis. *In*: Discriminant analysis and applications. T. Cacoullos (ed.) Academic press, New York.

Callaway, R.M., E.E.C. Clebsch, and P.S. White. 1987. A multivariate analysis of forest communities in the western Great Smokey Mountains National Park. *Am. Mid. Nat.*, 118:107-120.

Carmean, W.H. 1967. Soil survey refinements for predicting black oak site quality in south eastern Ohio. *Soil Sci. Soc. Am. Proc.*, 31:805-810.

Carmean, W.H. 1975. Forest site quality evaluation in the United States. *Adv. Agron.*, 27:209-269.

Carter, R.E., and K. Klinka. 1990. Relationship between growing-season soil water-deficit, mineralizable soil nitrogen and site index in coastal Douglas fir. *For. Ecol. Manage.*, 30:301-311.

Chang, M., and D.G. Boyer. 1977. Fitting soil temperature by a periodic regression model. *Tex. J. Sci.*, 29:169-174.

Chang, M., and A.J.R. Aguilar. 1980. Effects of climate and soil on the radial growth of loblolly pine (*Pinus taeda* L.) in a humid environment of south eastern U.S.A. *For. Ecol. Manage.*, 3:141-150.

Charley, J. 1981. Soil as a nutrient reservoir and buffer system. *In*: Productivity in perpetuity. Proceedings of Australian forest nutrition workshop, Canberra 10-14, August.

Christie, J.M., and R. Lines. 1979. A comparison of forest productivity in Britain and Europe in relation to climatic factors. *For. Ecol. Manage.*, 2:75-102.

Churchill, D.M. 1968. The distribution and prehistory of *Eucalyptus diversicolor* F. Muell., *E. marginata* Donn ex Sm., and *E. calophylla* R. Br. in relation to rainfall. *Aust. J. Bot.*, 16:125-151.

Clutter, J.L., J.C.Forston, L.V. Pienaar, G.H. Brister, and R.L. Bailey. 1983. Timber management: A quantitative approach. John Wiley and son. Brisbane.

Clymo, R.S. 1980. Preliminary survey of the peat-bog Hummell Knowe Moss using various numerical methods. *Vegetatio*, 42:129-148.

Cook, A., M.M. Court, and D.A. MacLeod. 1977. The prediction of Scots pine growth in north east Scotland using readily assessable site characteristics. *Scottish Forestry*, 31:251-264.

Corns, I.G.W. 1983. Forest community types of west-central Alberta in relation to selected environmental factors. *Can. J. For. Res.*, 13:995-1010.

Corns, I.G.W., and D.J. Pluth. 1984. Vegetational indicators as independent variables in forest growth prediction in west-central Alberta, Canada. *For. Ecol. Manage.*, 9:13-25.

Cox, D.R. 1966. Some procedures associated with the logistic qualitative response curve. *In*: Research papers in statistics. F.N. David (ed.) Festschriftfor J. Neyman, Wiley, London.

Cromer, R.N. 1973. Perenial weeds in Australian forests. 2nd Vic. Weeds Conf. Weeds Sci. Soc. Vic. Aust.

Czarnowski, M.S., F.R. Humphreys, and S.W. Gentle. 1971. Quantitative expression of site index in terms of certain soil and climate characteristics of *Pinus radiata* D. Don in Australia and New Zealand. *Ekologia Polska.*, 19:295-309.

Dalal, R.C., and R.J. Mayer. 1986a. Long-term trends in fertility of soils under continuous cultivation and cereal cropping in southern Queensland. I. Overall changes in soil properties and trends in winter cereal yields. *Aust. J. Soil. Res.*, 24:265-280.

Dalal, R.C., and R.J. Mayer. 1986b. Long-term trends in fertility of soils under continuous cultivation of cereal cropping in southern Queensland. II. Total organic carbon and its rate of loss from the soil profile. *Aust. J. Soil. Res.*, 24:281-292.

Dalal, R.C., and Mayer, R.J. 1986c. Long-term trends in fertility of soils under continuous cultivation of cereal cropping in southern Queensland. III. Distribution and kinetics of soil organic carbon in particle-size fractions. *Aust. J. Soil. Res.*, 24:293-300.

Dargie, T.C.D. 1984. On the integrated interpretation of indirect site ordinations: a case study using semi-arid vegetation in southeastern Spain. *Vegetatio,* 55:37-55.

Daubenmire, R. 1968. Soil moisture in relation to vegetation distribution in the mountains of northern Idaho. *Ecology*, 49:431-438.

Daubenmire, R. 1976. The use of vegetation in assessing the productivity of forest lands. *Bot. Rev.*, 42:115-143.

Daubenmire, R., and J.B. Daubenmire. 1968. Forest vegetation of eastern Washington and northern Idaho. *Washington Agricultural Experiment Station, Technical Bulletin.* No 60.

Day, N.E., and D.F.Kerridge. 1967. A general maximum likelihood discriminant. *Biometrics*, 23:313-323.

Doley, D. 1982. Photosynthetic productivity of forest canopies in relation to solar radiation and nitrogen cycling. *Aust. For. Res.*, 12:245-261.

Donald, C.M. 1970. The pastures of southern Australia. *In*: The Australian environment. G.W. Leeper (ed.) CSIRO. Melbourne. University Press.

Edwards, A.P., and J.M. Bremmer. 1967. Dispersion of soil particles by sonic vibrations. *J. Soil. Sci.*, 18:47-63.

Evans, J. 1974. Some aspects of growth of *Pinus patula* in Swaziland. *Comm. For. Rev.*, 53:57-62.

Faith, D.P., and R.H. Norris. 1989. Correlation of environmental variables with patterns of distribution and abundance of common and rare freshwater macroinvertebrates. *Biol. Conserv.*, 50:77-98

Faith, D.P., M.P. Austin, L. Belbin, and C.R. Margules. 1985. Numerical classification of profile attributes in environmental studies. *J. Env. Manage.*, 20:73-85.

Faith, D.P., P.R. Minchin, and L. Belbin. 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio*, 69:57-68.

Farr, W.A., and A.S. Harris. 1979. Site index of Sitka spruce along the Pacific coast related to latitude and temperatures. *For. Sci.*, 25:145-152.

Farr, W.A., and J.S. Hard. 1987. Multivariate analysis of climate along the southern coast of Alaska - some forestry implications. *USDA Res. Pap.* PNW-372.

Fasham, M.J.R. 1977. A comparison of nonmetric multidimensional scaling, principal components and reciprocal averaging for the ordination of simulated coemoclines and coenoplanes. *Ecology*, 58:551-561.

Fisher, R.A. 1936. The utilization of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7:179-188.

Fourt, D.F., D.G.M. Donald, J.N.R. Jeffers, and W.O. Binns. 1971. Corsican pine (*Pinus nigra* var. *maritima* (Ait.) Melville) in southern Britain: a study of growth and site factors. *Forestry*, 44:189-207.

Fox, L., J.A. Brockhouse, and N.D. Tosta. 1985. Classification of timberland productivity in north western California using Landsat, topographic, and ecological data. *Photogrammetric Engineering and Remote Sensing*, 51:1745-1752.

Fralish, J.S., and O.L. Loucks. 1975. Site quality evaluation models for aspen (*Populus tremuloides*) in Wisconsin. *Can. J. For. Res.*, 5:523-528.

France, E.A., D. Binkley, and D. Valentine. 1990. Soil chemistry changes after 27 years under four tree species in southern Ontario. *Can. J. For. Res.*, 19:1648-1650.

Gaiser, R.N. 1951. Relation between topography, soil characteristics, and the site index of white oak in south eastern Ohio. *USDA Tech. Pap. CS-121*.

Gasana, J.K., and H. Loewenstein. 1984. Site classification for maiden's gum, *(Eucalyptus globulus* subsp. *maidenii*), in Rwanda. *For. Ecol. Manage.*, 8:107-116.

Gauch, H.G. 1973. The relationship between sample similarity and ecological distance. *Ecology*, 54:618-622.

Gauch, H.G., and R.H. Whittaker. 1981. Hierarchial classification of community data. *J. Ecol.*, 69:537-557.

Gauch, H.G., and T.R. Wentworth. 1976. Canonical correlation analysis as an ordination technique. *Vegetatio*, 33:17-22.

Gauch, H.G., R.H. Whittaker, and S.B. Singer. 1981. A comparative study of nonmetric ordinations. *J. Ecol.*, 69:135-152.

Gauch, H.G., R.H. Whittaker, and T.R. Wentworth. 1977. A comparative study of reciprocal averaging and other ordination techniques. *J. Ecol.*, 65:157-174.

Gee, G.W., and J.W. Bauder. 1986. Particle-size analysis. *In*: Methods of soil analysis, Part 1. Physical and mineralogical methods. *Agron. Mono.*, No. 9.

Gessel, S.P., D.W. Cole, and E.C.Steinbrenner. 1973. Nitrogen balances in forest ecosystems of the Pacific Northwest. *Soil Biol. Biochem.*, 5:19-34.

Goodall, D.W. 1954. Objective methods for the classification of vegetation. III. An essay in the use of factor analysis. *Aust. J. Bot.*, 2:304-324.

Gower, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53:325-338.

Gower, J.C. 1971. A general coefficient of similarity and some of its properties. *Biometrics*, 23:623-637.

Green, R.N., P.L. Marshall, and K. Klinka. 1989. Estimating site index of Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) from ecological variables in southwestern British Columbia. *For. Sci.*, 35:50-63.

Grey, D.C. 1979. Site quality prediction for *Pinus patula* in the Glengarry area, Transkei. *S.A. For. J.*, 111:44-48.

Grigor'ev, Yu. 1986. [Water tables and forest types]. *Lesnoe Khozyaistvo*, 5:32-36.

Guttman, L. 1968. A general non-metric technique for finding the smallest co-ordinate space for configuration of points. *Psychometrika*, 33:469-506.

Hagglund, B. 1981. Evaluation of forest site productivity. *Forestry Abstracts*, 42:517-527.

Haines, L.W., T.W. Maki, and S.G. Sanderford. 1973. The effect of mechanical site preparation treatments on soil productivity and tree (*Pinus taeda* L. and *P. elliottii* Engelm. var. *elliottii*) growth. *In*: Forest soils and forest land management. B. Bernier and C.H. Winget (Eds.). Les Presses de l'Universite Laval, Quebec.

Haines, S.G., and G. Cleveland. 1981. Seasonal variation in properties of five forest soils in south west Georgia. *Soil Sci. Soc. Am. J.*, 45:139-143.

Hajdu, L.V. 1981. Graphical comparison of resemblence measures in phytosociology. *Vegetatio*, 48:47-59.

Hamilton, W.N., and H.H. Krause. 1985. Relationship between jack pine growth and site variables in New Brunswick plantations. *Can. J. For. Res.*, 15:922-926.

Harding, R.B., D.F. Grigal, and E.H. White. 1985. Site quality evaluation for white spruce plantations using discriminant analysis. *Soil. Sci. Soc. Am. J.*, 49:229-232.

Harrington, C.A. 1986. A method of site quality evaluation for Red Alder. USDA *Gen. Tech. Rep.* PNW 192.

Havel, J.J. 1968. The potential of the northern swan coastal plain for *Pinus pinaster* Ait. plantations. *Bull. For. Dept. West. Aust.*, No 76.

Havel, J.J. 1975. Site vegetation mapping in the northern jarrah forest (Darling Range). I. definition of site vegetation types. *Bull. For. Dept. West. Aust.*, No 86.

Havel, J.J. 1981a. Application of fundamental synecological knowledge to practical problems in forest management. I. Theory and methods. *For. Ecol. Manage.*, 3:1-29.

Havel, J.J. 1981b. Application of fundamental synecological knowledge to practical problems in forest management. II. Applications. *For. Ecol. Manage.*, 3:81-111.

Heinselman. M.L. 1963. Forest sites, bog processes and peatland types in the Glacial Lake Agassz region, Minnesota. *Ecol. Monogr.*, 33:327-374.

Hellmers, H., and D.A. Rook. 1973. Air Temperature and growth of *P.radiata* seedlings. *N.Z.J. For. Sci.*, 3:271-285.

Hill, M.O. 1973. Reciprocal averaging: an eigenvector method of ordination. *J. Ecol.*, 61:237-249.

Hill, M.O., and H.G. Gauch. 1980. Detrended correspondence analysis, an improved ordination technique. *Vegetatio,* 42:47-58.

Hocking, R.R. 1976. The analysis and selection of variables in linear regression. *Biometrics*, 32:1-50.

Hunter, I.R., and A.R. Gibson. 1984. Predicting *Pinus radiata* site index from environmental variables. *N.Z. J. For. Sci.*, 14:53-64.

Hutchinson, M.F. 1987. A new objective method for spline interpolation of meteorological variables from irregular networks applied to the estimation of monthly mean solar radiation, temperature, precipitation and windrun. *In*: Proc. United Nations University Workshop on need for climatic and hydrological data in agriculture of South-East Asia. E.A. Fitzpatrick (Ed) A.C.T. 12-15 December 1983. United Nations University, Tokyo.

Hutchinson, M.F., and R.J. Bischof. 1983. A new method for estimating the spatial distribution of mean seasonal and annual rainfall applied to the Hunter Valley. *Aust. Meteorol. Mag.*, 31:179-184.

Inions, G.B. 1990. Classification and evaluation of sites in karri regeneration. I. Edaphic and climatic attributes. *For Ecol. Manage.*, 32:117-134.

Inions, G.B., G. Wardell-Johnson, and A.A. Annels. 1990. Classification and evaluation of sites in karri regeneration. II. Floristic attributes. *For. Ecol. Manage.*, 32:135-154.

Jackson, D.S., and H.H. Gifford. 1974. Environmental variables influencing the increment of radiata pine. I. periodic volume increment. *N.Z.J. For. Sci.*, 4:3-26.

Kabzens, R.D., and K. Klinka. 1987a. Initial quantitative characterization of soil nutrient regimes. I. Soil properties. *Can. J. For. Res.*, 17:1557-1564.

Kabzens, R.D., and K. Klinka. 1987b. Initial quantitative characterization of soil nutrient regimes. II. Relationships among soils, vegetation, and site index. *Can. J. For. Res.*, 17:1565-1571.

Keenan, R.J., and S. Candy. 1983. Growth of young *Eucalyptus delegatensis* in relation to variation in site factors. *Aust. For. Res.*, 13:197-205.

Keeney, D.R. 1980. Prediction of soil nitrogen availability in forest ecosystems: A literature review. *For. Sci.*, 26:159-171.

Kendell, D.G. 1970. Seriation from abundance matrices. *In*: Anglo-Romanian conference on mathematics in the archeological and historical sciences. R.F. Hodson. D.G. Kendell and P. Tautu. (Eds.) Edinburgh University Press. Edinburgh.

Klinka, K., and R.E. Carter. 1990. Relationships between site index and synoptic environmental factors in immature coastal Douglas-fir stands. *For. Sci.*, 36:815-830.

Klinka, K., P.J. Courtin, and R.N. Green. 1984. An approach to quantitative classification of hygrotopes and trophotopes of forest soils. *In*: First IUFRO workshop on qualitative and quantitative assessment of forest sites with special reference to soils. Swiss federal institute of forest research. Birmensdorf, Switzerland.

Kozlowski, T.T. 1982. Water supply and tree growth. Part 1. Water deficits. *For. Abstr.*, 43:57-95.

Krajina, V.J. 1969. Ecology of forest trees in British Columbia. *Ecol. West. N. Am.*, 2:1-146.

Kruskal, J.B. 1964. Non-metric multidimensionsal scaling : A numerical method. *Psychometrika*, 29:115-129.

Kubota, T. 1972. Aggregate-formation of allophonic: effects of drying on the dispersion of the soils. *Soil. Sci. Plant. Nutr.*, 18:79-87.

Kuusipalo, J. 1985. An ecological study of upland forest site classification in southern Finland. *Acta Forestalia Fennica*, No. 192.

Layser, E.F. 1974. Vegetative classification : Its application to forestry in the northern Rocky Mountains. *J. For.*, 72:354-357.

Lea, R., and R. Ballard. 1982. Predicting loblolly pine growth response from N fertilizer, using soil-N availability indices. *Soil Sci. Soc. Am. J.*, 46:1096-1099.

Levins, R., and R. Lewontin. 1980. Dialectics and reductionism in ecology. *In*: Conceptual issues in ecology, E. Saarinen (ed.). Reidel, Dordrecht.

Lewis, D.C., A.L. Clarke, and W.B. Hall. 1981. Factors affecting the retention of phosphorus applied as superphosphate to the sandy soils in south-eastern South Australia. *Aust. J. Soil. Res.*, 19:167-174.

Lewis, D.C., A.L., Clarke, and W.B. Hall. 1987a. Accumulation of plant nutrients and changes in soil properties of sandy soils under fertilized pasture in south-eastern South Australia. I. Phosphorus. *Aust. J. Soil. Res.*, 25:193-202.

Lewis, D.C., A.L. Clarke, and W.B. Hall. 1987b. Accumulation of plant nutrients and changes in soil properties of sandy soils under fertilized pasture in south-eastern South Australia. II. Total sulfur and nitrogen, organic carbon and pH. *Aust. J. Soil. Res.*, 25:203-210.

Lloyd, W.J., and P.E. Lemmon. 1970. Rectifying azimuth (of aspect) in studies of soil-site index relationships. *In* : Tree growth and forest soils. C.T. Youngberg, and C.B. Davey (eds.), Oreg. State. Univ. Press, Corvallis, Oreg.

MacLean, C.D., and C.L. Bolsinger, 1973. Estimating Dunning's site index from plant indicators. *USDA. Res. Note* PNW-197.

Mader, D.L. 1976. Soil-site productivity for natural stands of white pine in Massachusetts. *Soil Sci. Soc. Am. J.*, 40:112-115.

Mader, D.L., and D.F. Owen. 1961. Relationships between soil properties and red pine growth in Massachusetts. *Soil Sci. Soc. Am. Proc.*, 25:62-65.

Major, J. 1963. A climatic index to vascular plant activity. *Ecology*, 44: 485-498.

Malan'in, A.N. 1985. [Effect of ground water depth on forest growth conditions on sands in the arid zone]. *Biologicheskie Nauki*, 8:93-101.

Malcolm, D.C. 1970. Site factors and the growth of Sitka spruce. Unpublished Ph.D. Thesis, Uni. Edinburgh.

Martin, W.S. 1944. Grass covers in their relation to soil structure. *Emp. J. Exp. Agric.*, 12:21-32.

McLeod, S.D., and S.W. Running 1988. Comparing site quality indices and productivity in ponderosa pine stands of western Montana. *Can. J. For. Res.*, 18:346-353.

McNab, W.H. 1985. Geomorphic factors associated with quality of yellow-poplar sites in the southern Appalachians. *In*:.Proc. third biennial southern silvicultural research conference. E. Shoulders (Ed.). *USDA Gen. Tech. Rep.* S0-54.

McNabb, D.H., K. Cromack, and R.L. Fredriksen. 1986. Variability of nitrogen and carbon in surface soils of six forest types in the Oregon Cascades. *Soil Sci. Soc. Am. J.*, 50:1037-1041.

McQuilkin, R.A. 1976. The necessity of independent testing of soil-site equations. *Soil. Sci. Soc. Am. J.*, 40:783-785.

Meeuwig, R.O., and S.V. Cooper. 1981. Site quality and growth of pinyon-juniper stands in Nevada. *For. Sci.*, 27:593-601.

Michaelis, J. 1973. Simulation experiments with multiple group linear and quadratic discriminant analysis. *In:* Discriminant analysis and applications. T. Cacoullos (ed.). Academic Press. New York.

Mikhail, E.H., and G.P. Briner. 1978. Routine particle size analysis of soils using sodium hypochlorite and ultrasonic dispersion. *Aust. J. Soil. Res.*, 16:241-244.

Miller, W.S., and A.N. Auclair. 1974. Factor analytic models of bioclimate for Canadian forest regions. *Can. J. For. Res.*, 4:536-548.

Minchin, P.R. 1987a. An evaluation of the relative robustness of technqiques for ecological ordination. *Vegetatio*, 71:145-156.

Minchin, P.R. 1987b. Simulation of multidimensional community patterns: towards a comprehensive model. *Vegetatio*, 71:145-156.

Moir, W.H., and E.P. Bachelard. 1969. Distribution of fine roots in three Pinus radiata plantations near Canberra, Australia. *Ecology*, 50:658-662.

Monserud, R.A. 1984. Height growth and site index curves of inland Douglas-fir based on stem analysis data and forest habitat type. *For. Sci.*, 30: 943-965.

Monserud, R.A., and G.E. Rehfeldt. 1990. Genetic and environmental components of variation of site index in inland Douglas-fir. *For. Sci.*, 36:1-9.

Monserud, R.A., V. Moody, and D.W. Brever. 1990. A soil-site study for inland Douglas-fir. *Can. J. For. Res.*, 20: 686-695.

Munn, L.C., and J.P. Vimmerstedt. 1980. Predicting height growth of Yellow-poplar from soils and topography in south eastern Ohio. *Soil Sci. Soc. Am. J.*, 44:384-387.

Murphy, J., and T. Riley. 1962. A modified single solution method for the determination of phosphate in natural waters. *Anal. Chimica. Acta.*, 27:31-36.

Newnham, R.M. 1968. A classification of climate by principal component analysis and its relationship to tree species distribution. *For. Sci.*, 14:254-264.

Nix, H.A. 1986. A biogeographic analysis of Australian Elapid snakes. *In:* Atlas of Australian Elapid Snakes. R. Longmore. (ed.).Bureau of Flora and Fauna, Canberra, A.C.T. Aust.

Northcote, K.H. 1971. A factual key for the recognition of Australian soils. Rellim Technical Publications. Glenside South Australia.

Noy-Meir, I., D. Walker, and W.T. Williams. 1975. Data transformations in ecological ordination. II. On the meaning of data standardization. *J. Ecol.*, 63:779-800.

Okansen, J. 1983. Ordination of boreal heath-like vegetation with principal components analysis, correspondence analysis and multidimensional scaling. *Vegetatio*, 52:181-189.

Oliver, C.D. 1978. Subsurface geologic formations and site variation in upper sand hills of South Carolina. *J. For.*, 76:352-354.

Orloci, L. 1966. Geometric models in ecology. I. The theory and application of some ordination techniques. *J. Ecol.*, 54:193-215.

Orloci, L. 1974. Revisions for the Bray and Curtis ordination. *Can. J. Bot.* 52:1773-1776.

Orloci, L. 1988. Community organisation: recent advances in numerical methods. *Can. J. Bot.*, 66:2626-2633.

Page, G. 1974. Effects of forest cover on the properties of some Newfoundland forest soils. Can. Dep. Environ., *For. Serv. Publ.*, ND 1332.

Page, G. 1976. Quantitative evaluation of site potential for spruce and fir in Newfoundland. *For. Sci.*, 2:131-143.

Peterson, D.L., and G.L. Rolfe. 1985. Temporal variation in nutrient status of a floodplain forest soil. *For. Ecol. Manage.*, 12:73-83.

Pfister, R.D., and S.F. Anro. 1980. Classifying forest habitat types based on potential climax vegetation. *For. Sci.*, 26:52-70.

Powers, R.F. 1980. Mineralisable soil nitrogen as an index of nitrogen availability to forest trees. *Soil Sci. Soc. Am. J.*, 44:1314-1320.

Prentice, I.C. 1977. Non-metric ordination methods in ecology. *J. Ecol.*, 65:85-94.

Prentice, I.C. 1980. Vegetation analysis and order invariant gradient models. *Vegetatio*, 42:27-34.

Ralston, M.L., and R.I. Jennrich. 1979. DUD a derivative-free algorithm for nonlinear least squares. *Technometrics*, 1:7-14.

Roberts, J. 1976. A study of root distribution and growth in a *Pinus sylvestris* L. (Scots pine) plantation in East Anglia. *Plant Soil*, 44:607-621.

Robertson, E.O., L.A. Jozsa, and D.L. Spittlehouse. 1990. Estimating Douglas-fir wood production from soil and climate data. *Can. J. For. Res.*, 20:357-364.

Running, S.W., R.R. Nemani, and R.D. Hungerford. 1987. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. *Can. J. For. Res.*, 17:472-483.

Ryan, P.J. 1986. Characterization of soil and productivity of *Pinus radiata* (D. Don) in New South Wales. II. Pedogenesis on a range of parent materials. *Aust. J. Soil Res.*, 24:95-102.

Salt, G.W. 1983. Roles: their limits and responsibilities in ecological and evolutionary research. *Am. Nat.*, 122:697-705.

Saunders, J.C., T.H. Booth, J. Turner, and E.R. Williams. 1984. Some effects of site, particularly upper horizon soil, on the growth of *Pinus radiata*. *Aust. For. Res.*, 14:27-36.

Schiller, G. 1982. Significance of bedrock as a site factor for aleppo pine. *For. Ecol. Manage.*, 4:213-223.

Schmidt, M.G., and W.H. Carmean. 1988. Jack pine site quality in relation to soil and topography in north central Ontario. *Can. J. For. Res.*, 18:297-305.

Schwab, A.P., and S. Kulyingyong. 1989. Changes in phosphate availability indexes with depth after 40 years of fertilization. *Soil. Sci.*, 147:179-186.

Shepard, R.N. 1974. Representation of structure in similarity data - problems and prospects. *Psychometrika*, 39:373-421.

Shoulders, E., and A.E. Tiarks. 1980. Predicting height and relative performance of major southern pines from rainfall, slope, and available soil moisture. *For. Sci.*, 26:437-447.

Shumway, J., and W.A. Atkinson. 1978. Predicting nitrogen fertilizer response in unthinned stands of Douglas fir. *Commun. Soil Sci. Plant Anal.*, 9:529-539.

Skinner, M.F., and P.M. Attiwill. 1981. The productivity of pine plantations in relation to previous land use. I. growth responses in agricultural and forest soils. *Plant and soil*, 60:161-176.

Smalley, G.W. 1986. Classification and evaluation of forest sites on the northern Cumberland Plateau. *USDA. Gen. Tech. Rep.* S0-50.

Sokal, R.R., and C.D. Michener. 1957. The effect of different numerical techniques on the phenetic classification of bees of the *Hoplitis* complex (Megachilidae). *Proc. Linn. Soc. London*, 178:59-74.

Specht, R.L. 1981. Growth indices - their role in understanding the growth, structure and distribution of Australian vegetation. *Oecologia*, 50:347-356.

Squire, R.O., G.C. Marks, and F.G. Craig. 1978. Root development in a *Pinus radiata* D. Don plantation in relation to site index, fertilizing and soil bulk density. *Aust. For. Res.*, 8:103-114.

Stage, A.R. 1976. An expression for the effect of aspect, slope, and habitat type on tree growth. *For. Sci.*, 22:457-460.

Stahr, K. 1979. [Soil parameters in relation to site response in the southern Black Forest.] Bodenkundliche Parameter fur die Standortsansprache in Sudschwarzwald. *Allgemeine Forstzeitschrift*, 47:1284.

Summer, M.E., and M.P.W. Farina. 1986. Phosphorus interactions with other nutrients and lime in field cropping systems. *Adv. Soil. Sci.*, 5:201-230.

Swift, L.W. 1976. Algorithm for solar radiation on mountain slopes. *Water Resour. Res.*, 12:108-112.

Tajchman, S.J. 1984. Distribution of the radiative index of dryness and forest site quality in a mountainous watershed. *Can. J. For. Res.*, 14:717-721.

Tajchman, S.J., and C.J. Lacey. 1986. Bioclimatic factors in forest site potential. *For. Ecol. Manage.*, 14:211-218.

Terman, G.L., J.C. Noggle, and C.M. Hunt. 1977. Growth rate - nutrient concentration relationships during early growth of corn as affected by applied N, P and K. *Soil. Sci. Soc. Am. J.*, 41:363-368.

Thrower, J.S. 1987. Growth intercepts for estimating site quality of young white spruce plantations in north central Ontario. *Can. J. For. Res.*, 17:1385-1389.

Torgerson, W.S. 1952. Multidimensional scaling. I. Theory and method. *Psychometrika*, 17:401-419.

Truman, R., F.R. Humphreys, and M.J. Lambert. 1983. Prediction of site index for *Pinus radiata* at Mullions Range State Forest, New South Wales. *Aust. For. Res.*, 13:207-215.

Turvey, N.D., A.B. Rudra, and J. Turner. 1986. Characteristics of soil and productivity of *P. radiata* (D. Don) in New South Wales. I. Relative importance of soil physical and chemical parameters. *Aust. J. Soil Res.*, 24:95-102.

Usher, M.B. 1970. Pattern and seasonal variability in the environment of Scots pine forest soil. *J. Ecol.*, 58:669-680.

Van Groenewoud, H. 1984. The climatic regions of New Brunswick: a multivariate analysis of meteorological data. *Can. J. For. Res.*, 14:389-394.

Vance, E.D., and G.S. Henderson. 1984. Soil nitrogen availability following long-term burning in an oak-hickory forest *Soil Sci. Soc. Am. J.*, 48:184-190.

Verbyla D.L., and R.F. Fisher. 1989. Effect of aspect on ponderosa pine height and diameter growth. *For. Ecol. Manage.*, 27:93-98.

Verbyla, D.A. 1986. Potential prediction bias in regression and discriminant analysis. *Can. J. For. Res.*, 16:1255-1257.

Wahba, G., and J. Wendelberger. 1980. Some new mathematical methods for variational objective analysis using splines and cross validation. *Mon. Weather. Rev.*, 108:1122-1143.

Walkley, A., and I.A. Black. 1934. An examination of the egtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil. Sci.*, 37:29-38.

Watanable, F.S. and Olsen, S.R. 1965. Test of an ascorbic acid method for determining phosphorus in water and Na HCO_3 extracts from soil. *Soil Sci. Soc. Am. Proc.*, 29:677-678.

Watts, R.F., and M.L. Heinselman. 1965. Foliar nitrogen and phosphorous levels related to site quality in a northern Minnesota spruce bog. *Ecology*, 46:357-361.

Weaver, T., and F. Forcella. 1979. Seasonal variations in soil nutrients under six Rocky Mountain vegetation types. *Soil Sci. Soc. Am. J.*, 43:589-593.

West, P.W. 1981. Simulation of diameter growth and mortality in regrowth Eucalyptus forests of southern Tasmania. *For. Sci.* 27:603-616.

White, E.J. 1979. The prediction and selection of climatological data for ecological purposes in Great Britain. *J. Appl. Ecol.*, 16:141-160.

White, E.J. 1982a. Relationships between height growth of Scots pine (*Pinus sylvestris* L.) and site factors in Great Britain. *For. Ecol. Manage.*, 4:225-245.

White, E.J. 1982b. Relationship between height growth of stand and open-grown single trees of Scots pine (*Pinus sylvestris* L.) and site factors in Great Britain. *For. Ecol. Manage.*, 4:297-259.

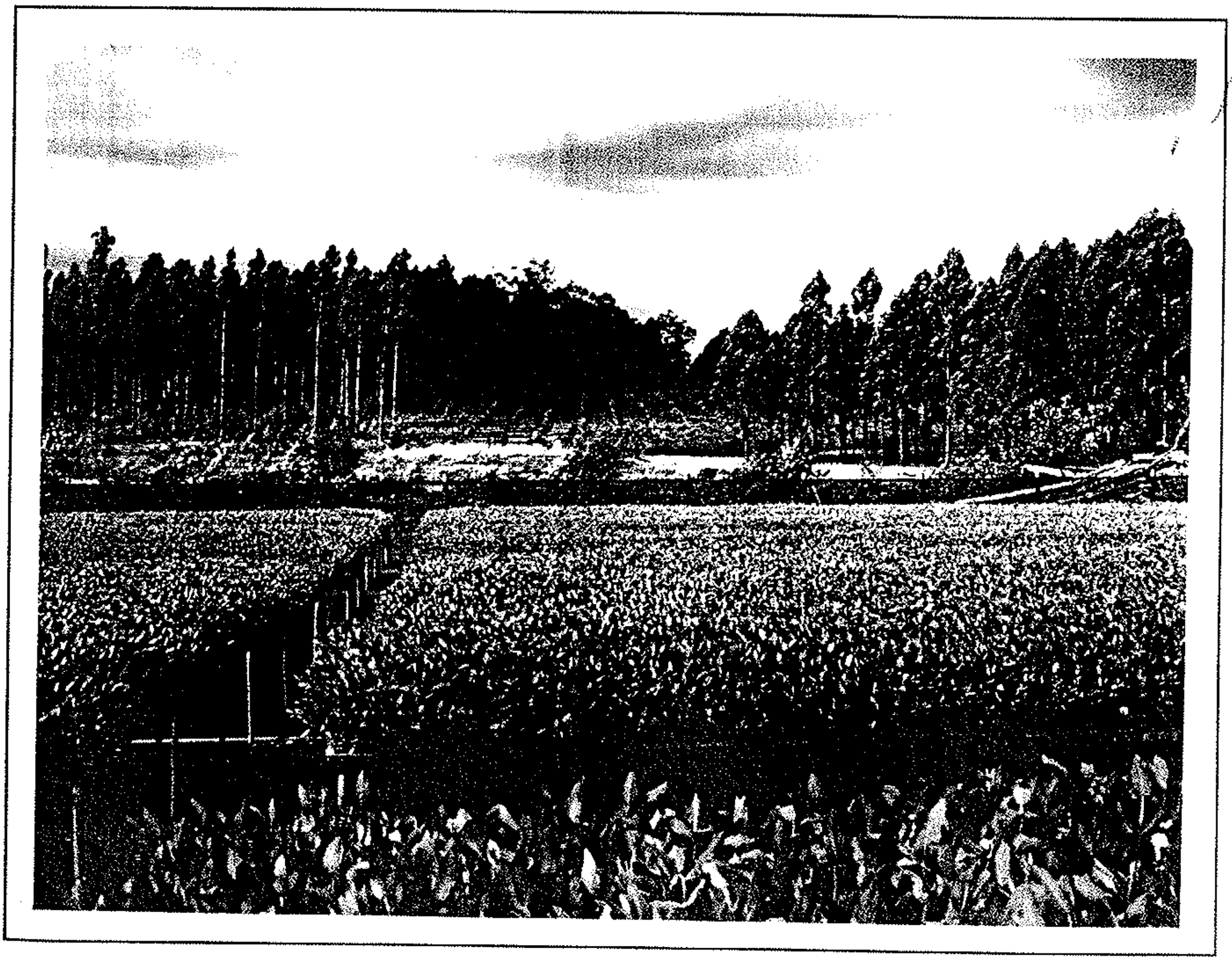
Will, G.M., and R. Ballard. 1976. Radiata pine-soil degrader or improver? *N.Z. J. For.*, 21:248-252.

Williams, B.K. 1981. Discriminant analysis in wildlife research: theory and applications. *In*: The use of multivariate statistics in studies of wildlife habitat. D.E. Capen. (ed.). USDA *Gen. Tech. Rep.* RM-87.

Williams, C.H. 1962. Changes in nutrient availability in Australian soils as a result of biological activity. *J. Aust. Inst. Agric. Sci.*, 28:196-205.

CHAPTER FIVE

Validation of Predictive Equations



5.1 LITERATURE REVIEW: VALIDATION OF PREDICTIVE EQUATIONS

5.1.1 INTRODUCTION

Equations which predict an outcome, such as those detailed in Chapters 3 and 4, are merely mathematical abstractions. As such they cannot be expected to make predictions which agree exactly with reality. Prior to the use of any system of equations, collectively termed a model. it should be examined to compare its agreement with the system it is intended to represent (Goodall 1972). Such an examination is termed validation (Schaeffer 1980). Validation involves the comparison of model output with a data set which is independent from that used to calibrate the model.

There are two philosophical approaches to model validation. The first approach uses a framework of statistical hypothesis testing to formally assess whether a model meets the specified accuracy requirement of the user. The second approach uses no specified standard of accuracy and the objective is simply to give the user some estimate of how far estimated values will be from the truth.

A third approach, which is not validation under the above definition but is worthy of mention. is cross-validation. Here the i^{th} sample is deleted during model calibration. The model is then tested on the excluded case and an error derived. The excluded case is returned to the data set and the $i^{th} + 1$ case is excluded and the procedure is repeated. The procedure continues until all cases have been removed. The accuracy of the model is then the mean of all errors. This procedure has proven popular particularly where small sample sizes prevents segregation of the data into calibration and validation datasets (Frank *et al.* 1984; Harding *et al.* 1985). Modifications of this cross-validation procedure are available (Efron 1983) and a comparison of such methodologies is given by Gong (1986).

A slightly different approach is given by Gertner (1986) and Gertner (1987) who examined the random errors in the independent variables and their effect on model precision. For the purposes of this review these procedures are not classed as validation methodologies and will not be discussed further.

5.1.2 THE STATISTICAL HYPOTHESIS TESTING APPROACH

The first of such methodologies was proposed by Freese (1960). The approach requires the user to define an acceptable level of accuracy. The accuracy attained by the model is then estimated, and finally a statistical test is applied to decide whether the model meets the accuracy required. The method is prominent in the forest science literature (Moser and Hall 1969; Pearson and Sternitzke 1974; Evert 1981) and modifications of the procedure have also arisen (Rennie and Wiant 1978; Ek and Monserud 1979). In such modifications a maximum anticipated error or critical error is calculated and defined as the smallest value of the error, which will lead to the acceptance of the null hypothesis. The critical error may also be used as a validation statistic (Ek and Monserud 1979).

It is argued that the assumptions underlying Freese's procedure have not been explicitly stated by Freese (1960). Reynolds (1984) clarified the assumptions inherent in Freese's methodology and proposed a more powerful test of accuracy by relating the critical error bounds to an interval estimator for a particular quantile of the distribution of the errors. These procedures were modified further by Gregoire and Reynolds (1988) who also discussed the robustness of such tests to departures from error normality.

Although examples of such tests appear in the literature (Ek and Monserud 1979; West 1981; Dolph 1989; Borders and Patterson 1990), the methodology has been criticized for its requirement for a user specified level of accuracy (Holdaway and Brand 1983). Under such a situation it is possible that a model may be deemed acceptable by one user and rejected by another. As such most models are validated using independent data to yield indications of accuracy and precision in the absence of formal hypothesis testing.

5.1.3 MEASURES OF ACCURACY AND PRECISION

With this approach two or more models may be compared, or a single model tested for its predictive capabilities. The measure of success, or otherwise, is the residual, defined as the observed minus the predicted value. Some authors have defined the residual as the predicted minus the observed, however this definition violates regression theory (Zuuring *et al.* 1988). The accuracy and precision of the model under examination is given by the mean and standard deviation of the residuals. However, no set of rules exist which stipulate the appropriateness of the model. As pointed out by Zuuring *et al.* (1988), final acceptance of the model is subjective, even though rigorous statistical analysis is employed to derive the model. Whether the model is accepted, given its accuracy and precision, will largely depend upon the end use of the model predictions (Newberry and Stage 1988).

This approach is employed more often than the formal hypothesis testing methodologies. It has been used to examine the accuracy and precision of a single model (Holdaway and Brand 1983), to compare a number of models (Lenhart 1988; Patterson and Stiff 1988) and to compare different modelling procedures (Stage and Renner 1988). The mean and standard deviation of the residuals are the most commonly used statistics (Devan and Burkhart 1982; Holdaway and Brand 1983; Stage and Renner 1988), although slight variations may occur. For example, the percentage bias of the residuals, defined as the ratio of the mean error to the sample mean, was used by Dolph (1989) for validating height-diameter equations. Other variations include that proposed by West (1983) who developed regression equations which examined the relationships between residuals and various independent attributes such as stand age, site index and time period of the simulation projection. Another approach used a Student's t-test to examine whether the mean of the residuals was equal to zero. Rejection of the null hypothesis would imply bias (Dyer and Bailey 1987). Zuuring *et al.* (1988) describes a method for graphically displaying residuals with various independent variables. A simple linear regression relating observed to predicted estimates has been used by Ek and Monserud (1979) and West (1981).

Some authors have adopted the strategy of using both validation philosophies when testing or comparing models, presenting both formal hypothesis tests as well as descriptive validation statistics (Ek and Monserud 1979; West 1981).

5.1.4 VALIDATION OF TOP HEIGHT DEVELOPMENT, SITE INDEX AND SOIL-SITE EQUATIONS

Although Carmean and Lenthall (1989) used both Freese's (1960) and Reynold's (1984) accuracy tests, most studies use descriptive statistics when validating top height development and site index equations. For example, Devan and Burkhart (1982) and Lappi and Bailey (1988) use the mean and standard deviation of the residuals when validating their top height development curves. Newnham (1988) employed the strategy of examining the mean residual for each site index class when validating site index curves in Canada. Some authors employ a range of residual statistics such as the mean, standard deviation, minimum, maximum and range (Cieszewski and Bella 1989). Smith and Watts (1987) compare seven top height development equations using the maximum absolute residual, the residual standard deviation and the frequencies of the absolute residuals by residual classes.

In those studies concerned with predicting productivity from environmental attributes, validation is usually presented as a description of residuals. Cross validation has also been used, but infrequently (White 1982; Harding *et al.* 1985; Verbyla and Fisher 1989). The correlation coefficient or the r-squared between observed and predicted values is a common validation statistic in such studies (Broadfoot 1969; Blyth and MacLeod 1981). However, as with site index and top height development equations, most studies describe the nature of the residuals with such statistics as the mean, standard deviation or standard error (Page 1976; Shoulders and Tiaks 1980; Schmidt and Carmean 1988).

5.2 VALIDATION OF TOP HEIGHT DEVELOPMENT EQUATIONS, SITE INDEX EQUATIONS AND EQUATIONS WHICH PREDICT SITE INDEX FROM ENVIRONMENTAL ATTRIBUTES

5.2.1 INTRODUCTION

Much has been written about the need for validating models (House 1974; Caswell 1977; Shaeffer 1980; Holdaway and Brand 1983), particularly models which predict productivity from environmental attributes (Broadfoot 1969; McQuilkin 1976; Verbyla 1986).

The top height development and site index equations presented in Chapter 3 and the equations which predict site index from environmental attributes presented in Chapter 4, have been compared on the basis of the residual statistics, derived from the statistical analysis of the calibration data sets only. However, these equations will form the bases for land acquisition and managerial decisions. As such some indication or test of their ability to predict is obligatory.

It is the aim of this study to:

(i) define the accuracy and precision of the predictions made by the individual candidate equations detailed in Chapters 3 and 4;

(ii) make recommendations as to which equations are the more appropriate and under which circumstances; and

(iii) define the accuracy and precision of predictions when equations are used in concert.

5.2.2 METHODS

5.2.2.1 VALIDATION DATA SETS

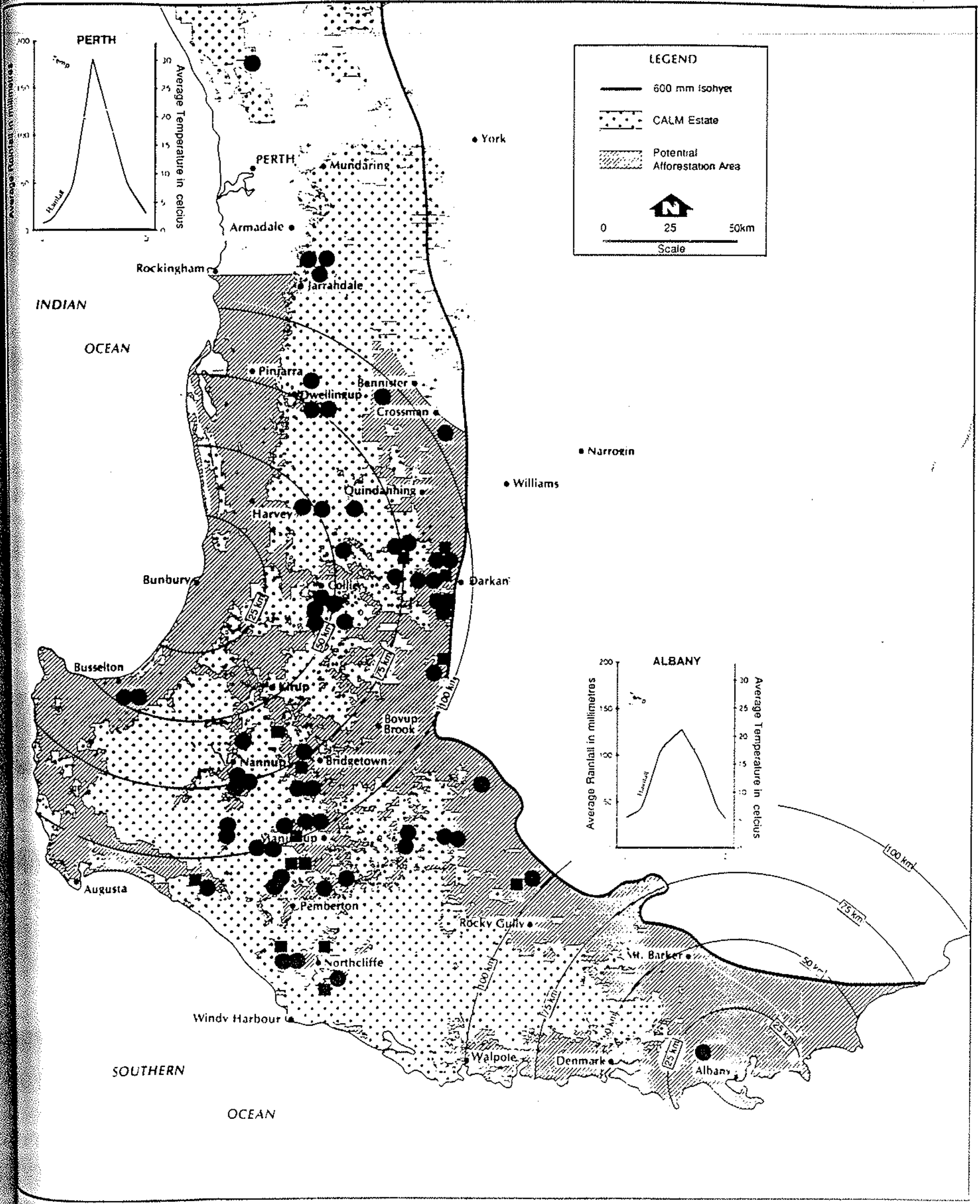
To derive the data set used for the validation of the top height development and site index curves, fourteen plots were established across the geographic range of the study area (Figure 21).A single site tree was felled in each plot with the selection of the site trees and the reconstruction of their growth pattern, via the methods previously described in Section 3.2.2. Of the 14 plots four were aged five years, five were aged six years, three were aged seven years, one was aged eight years and one was aged ten years, yielding 89 age top height data points.

At each of the 14 validation plots environmental attributes were derived via the assumptions and methodologies detailed in Section (4.2.2.3). A further four plots existed where top height development and site index data were previously derived and used in the construction of the equations reported in Chapter 3. However, these plots had incomplete collections of environmental data and therefore were not used in the construction of the equations reported in Chapter 4. These four plots were resampled for environmental attributes, bringing the total plots available for the validation of the equations which predict site index from environmental attributes to 18.

5.2.2.2 VALIDATION CRITERIA

The validation philosophy proposed by Freese (1960) and Reynolds (1984) is not deemed appropriate for this study as it requires a user specified accuracy level. The models under validation in this study will be used for different reasons by different people, each with a different accuracy requirement. As such it is considered more appropriate to provide such users with an estimate of the accuracy and precision of the equation in a descriptive manner. Such descriptive statistics are also useful for model comparisons.

Figure 21: South west Western Australia showing the location of plots used for model calibration and model validation.



● calibration plots
■ validation plots

Models were compared on the basis of residuals, defined as the observed minus the predicted value. Candidate models were initially compared using the mean of the residuals $(\overline{D})$ and the standard deviation of the residuals (D_{SD}). $\overline{D}$ represents the accuracy of the equation while D_{SD} represent the precision. The r-squared statistic from a linear regression analysis using the observed value as the dependent variable and the predicted value as the independent is also given. The most accurate and precise models were validated further by graphically examining the relationship between the residuals and the independent variables.

5.2.3 RESULTS

5.2.3.1 VALIDATION OF TOP HEIGHT DEVELOPMENT EQUATIONS

Residual statistics derived from the validation of equations which predict top height development are given in Table 32. Of the equations validated, Ek's (1971) and Payendeh's (1974) modification of the Chapman-Richards functional form (Eq. [11]) (see section 3.2.3.3.1) displays the best residual statistics. It has the smallest values of $\overline{D}$, indicating it is the most accurate equation tested. Eq. [11] and the algebraic difference equation Eq. [18] (see section 3.2.3.3.3) have similar levels of precision, as indicated by their D_{SD} values.

The allocation of plots to cluster groups, via heuristic rules and logistic regression, yielded better residual statistics than when discriminant analysis was used as the allocation criteria. However, neither strategy yields residual statistics which were an improvement over those yielded by the use of Eqs. [11] or [18].

On the basis of residual statistics it is recommended that Eq. [11] be used to model top height development. Examination of plots of residuals with S, H and A revealed no patterned bias. However, when a linear regression, using the observed top height as the dependent, and the predicted top height as the independent variable, was undertaken the intercept term was found to be significantly ($p<0.0373$) different from zero. No explanation can be offered as such bias is not evident in plots of observed versus predicted top heights (Figure 22).

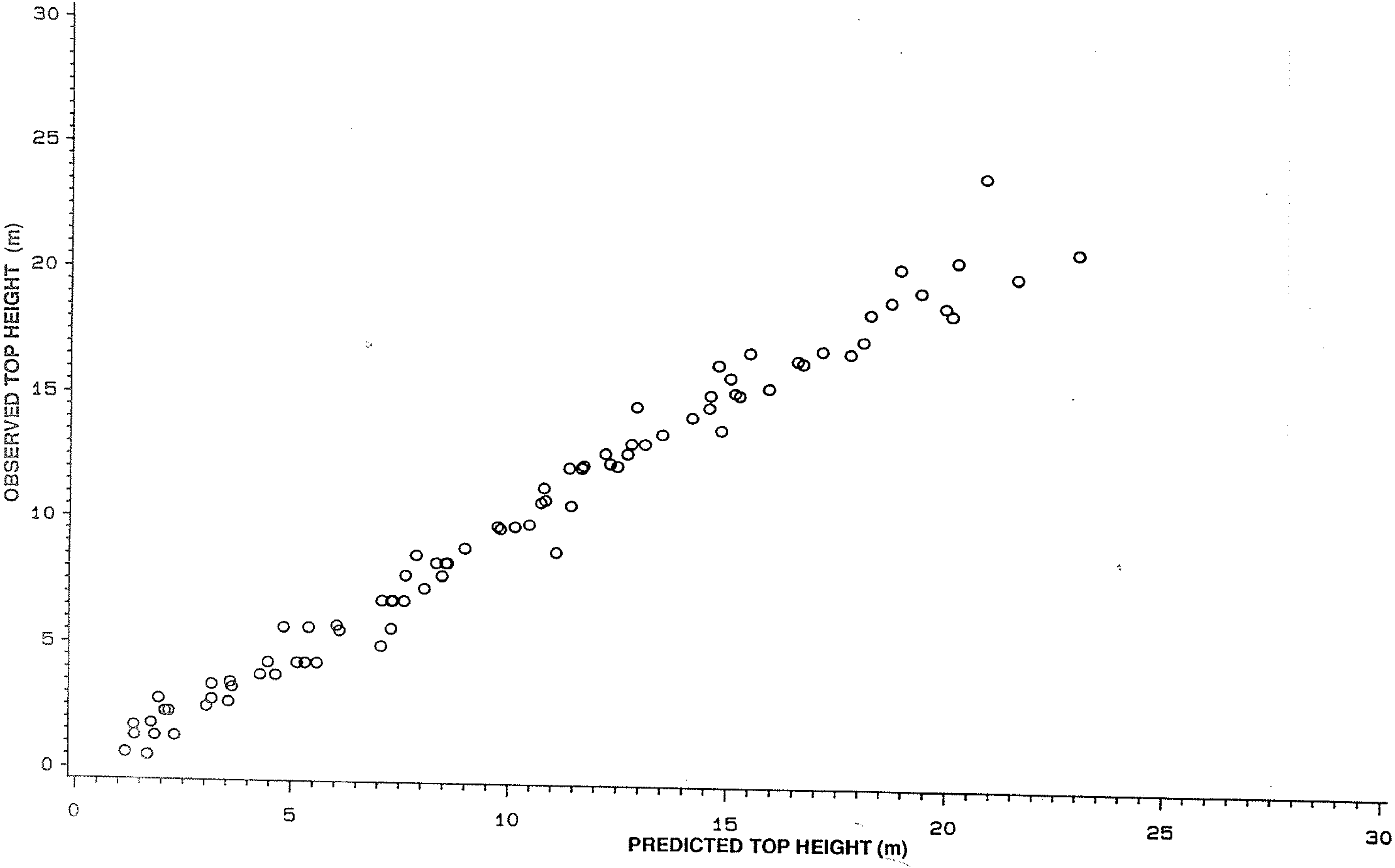
Table 32: Residual statistics generated from the validation of top height development equations(n=89).

Equation	$\overline{D}$	D_{SD}	r^2
Eq.[11]	-0.13	0.84	0.9811
Eq.[18]	-0.23	0.83	0.9811
Eq.[20]*	-0.15	1.01	0.9741
Eq.[20]**	-0.29	1.45	0.9469
Eq.[18]*	-0.16	0.90	0.9787
Eq.[18]**	-0.32	1.08	0.9684

* equation applied to each cluster group where the plot was allocated to a cluster group via heuristic rules and logistic regression.

** equation applied to each cluster group where the plot was allocated to a cluster group via discriminant functions.

Figure 22: **Observed top height by predicted top height derived from the validation of the top height development equation Eq. [11].**



5.2.3.2 VALIDATION OF SITE INDEX EQUATIONS

Residual statistics derived from the validation of equations which predict site index from H and A, are given in Table 33. The alternative form of Ek's (1971) and Payendeh's (1974) modification of the Chapman-Richards functional form (Eq. [13]) (see section 3.2.3.3.1) was both inaccurate and imprecise. Bias was detected with increasing values of S for this equation and it is therefore rejected. The algebraic difference site index equation (Eq. [19]) (see section 3.2.3.3.3) displayed poor accuracy. However, if data derived from years one and two were deleted from the validation data set, the equation yields acceptable accuracy and the highest level of precision of the site index equations validated.

Unlike top height development equations, separation of the data into cluster groups and applying cluster specific versions of Eq. [19], improved accuracy but not the level of precision. The allocation of plots to cluster groups via heuristic rules and logistic regression yields greater accuracy than the allocation of plots via discriminant functions. Little effect upon precision was detected between the two allocation procedures (Table 33).

In all cases the removal of the first two years data improved both accuracy and precision. It is recommended that Eq. [19] is used to predict S from H and A and the equation should not be applied until the stand is at least three years of age. If environmental attributes are available, the allocation of plots to cluster groups via heuristic rules and logistic regression and the application of the algebraic difference site index equation will yield increased accuracy. Again the application of these equations should only occur when the stand is three or more years of age. No patterned bias was detected with increasing values of A, H or S for either equation. The intercept terms from linear regressions using observed site index as the dependent variable, were not significant ($p>0.1076$; $p>0.3153$) for Eq. [19] or the algebraic difference site index equation applied to each cluster group respectively. Plots of predicted site index with observed site index is given in Figure 23 for both methods.

Table 33: Residual statistics generated from the validation of site index equations (n=89).

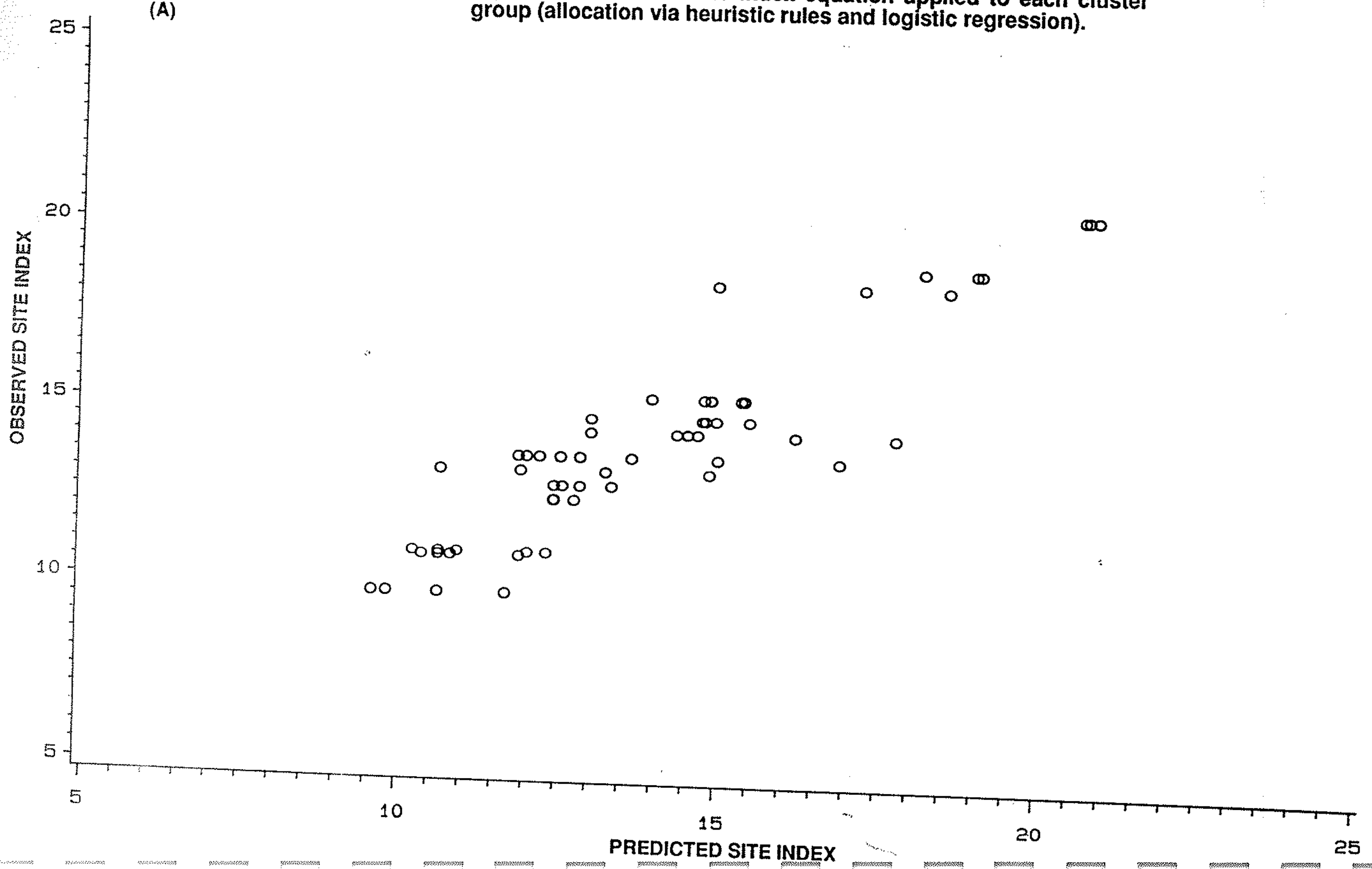
Equation	$\overline{D}$	D_{SD}	r^2
Eq.[13]	0.93	1.72	0.7021
Eq.[13]ˡ	0.76	1.41	0.9132
Eq.[19]	0.65	1.93	0.7215
Eq.[19]ˡ	0.18	0.87	0.9175
Eq.[19]*	-0.19	3.52	0.3681
Eq.[19]*ˡ	0.03	1.06	0.8744
Eq.[19]**	0.24	3.55	0.3430
Eq.[19]**ˡ	0.09	1.17	0.8494

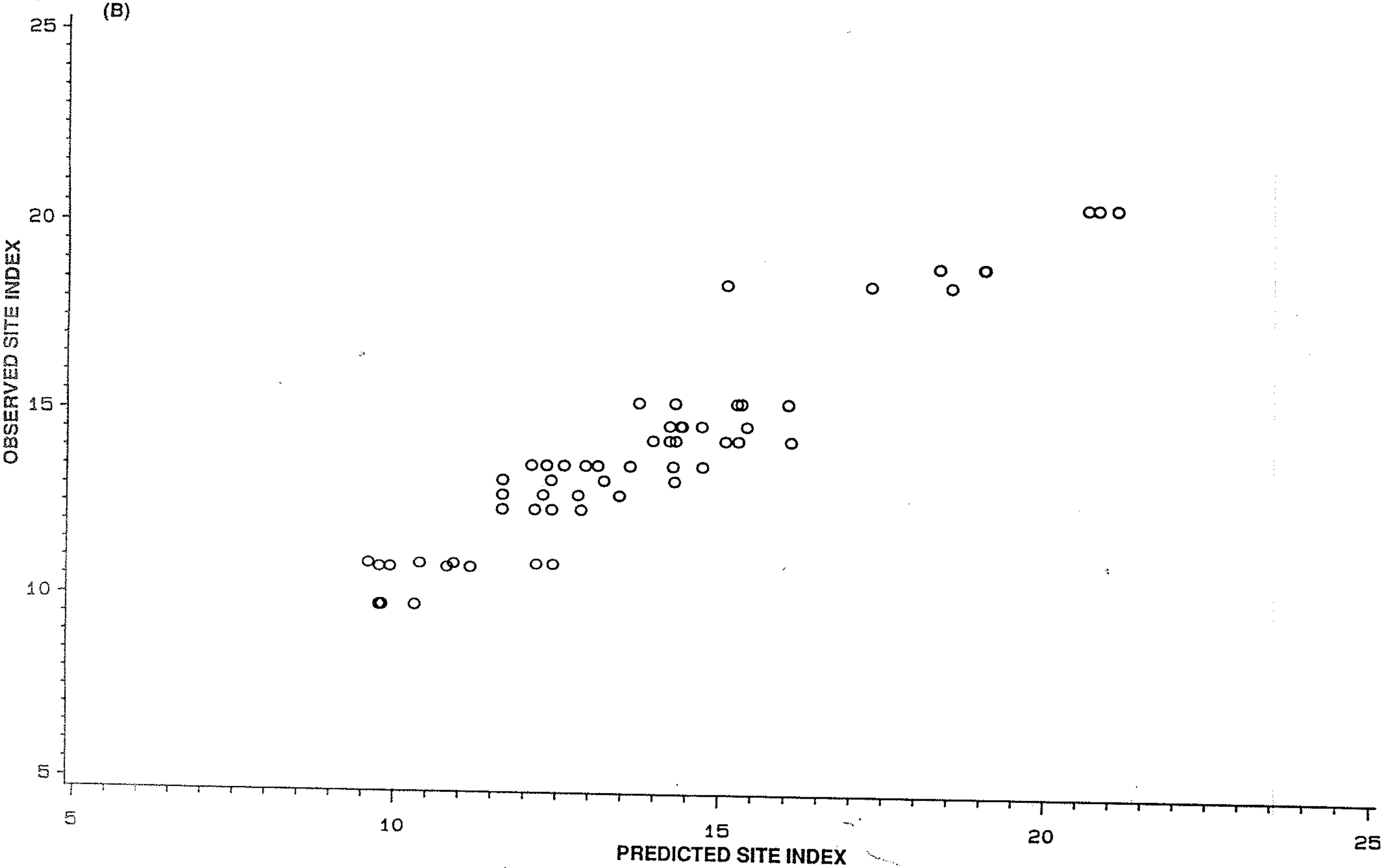
* equation applied to each cluster group where the plot was allocated to a cluster group via heuristic rules and logistic regression.

** equation applied to each cluster group where the plot was allocated to a cluster group via discriminant functions.

ˡ data from years one and two removed from the data set.

Figure 23: Observed site index by predicted site index derived from the validation of (a) the site index equation Eq. [19] and (b) the algebraic difference site index equation applied to each cluster group (allocation via heuristic rules and logistic regression).




(B)
OBSERVED SITE INDEX
PREDICTED SITE INDEX
25
20
15
10
5
5
10
15
20
25

5.2.3.3 VALIDATION OF EQUATIONS WHICH PREDICT SITE INDEX FROM ENVIRONMENTAL ATTRIBUTES

Residual statistics derived from the validation of equations which predict site index from environmental attributes are given in Table 34. The most accurate and precise results were obtained from equations comprised of all attribute types (see Section 4.3.3.2) (i.e. Eqs. [28], [30], [31], [32]). Equations comprised of single attribute types, or combinations of attribute types, and the equation derived via a stepwise variable selection algorithn, were less accurate and less precise than these equations.The equation displaying the least desirable residual statistics is that derived via the stepwise algorithm (Eq. [29]) (see section 4.3.3.3).

Although displaying desirable residual statistics Eq. [30] and Eq. [31] both exhibit patterned bias with site index, such that site index is underestimated for larger values of S. This problem does not occur when Eq. [28] or Eq. [32] are used. Eq. [32] is less accurate than Eq. [28] although their precision is similar. Therefore, it is recommended that Eq. [28] be used to predict site index from environmental attributes. When a linear regression, using predicted values as the independent variable and the observed values as the dependent, was fitted the intercept term was not significant ($p>0.7565$) indicating a lack of bias. Plots of observed versus predicted site index for this equation is given in Figure 24.

No improvement in residual statistics was gained by using equations or data specific to a previous land use (Table 34).

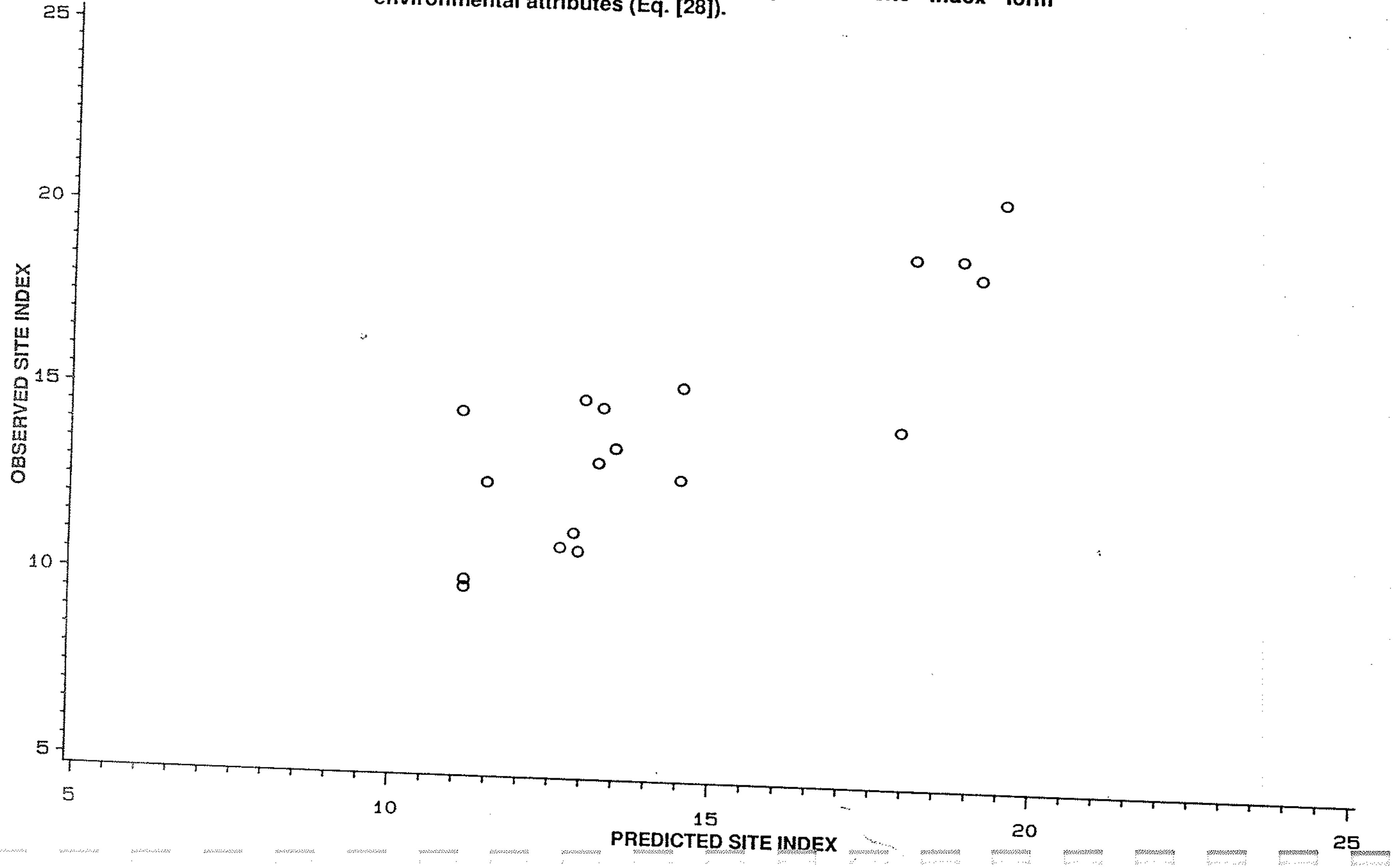
Table 34: (A) Residual statistics generated from the validation of equations which predict site index from environmental attributes (n=18).

Equation	$\overline{D}$	D_{SD}	r^2
Eq.[21]	0.66	2.02	0.7007
Eq.[22]	1.44	2.70	0.3250
Eq.[23]	-0.003	2.18	0.5543
Eq.[25]	-0.13	2.18	0.5551
Eq.[26]	0.87	1.87	0.6883
Eq.[27]	1.04	2.68	0.4095
Eq.[28]	-0.12	1.73	0.7233
Eq.[29]	0.64	3.42	0.1183
Eq.[30]	0.01	1.74	0.7213
Eq.[31]	-0.08	1.84	0.6981
Eq.[32]	0.38	1.71	0.7334

(B) Residual statistics generated from the validation of equations which predict site index from environmental attributes. Pasture sites only (n=16).

Equation	$\overline{D}$	D_{SD}	r^2
Eq.[35]	0.52	1.89	0.6652
Eq.[36]	-0.39	1.68	0.7341
Eq.[37]	-0.27	1.63	0.7468
Eq.[33]	-0.24	1.56	0.7595
Eq.[34]	-0.10	1.59	0.7507

Figure 24: **Observed site index by predicted site index derived from the validation of the equation which predicts site index form environmental attributes (Eq. [28]).**



5.2.3.4 VALIDATION OF TOP HEIGHT DEVELOPMENT EQUATIONS IN COMBINATION WITH EQUATIONS WHICH PREDICT SITE INDEX FROM ENVIRONMENTAL ATTRIBUTES.

Not only must equations perform well individually they must also perform well in unison. Residual statistics were examined when equations were used in concert, where site index was predicted from environmental attributes and used as input into top height development equations. Residual statistics for some of the combinations examined are given in Table 35.

The most accurate and precise combination of equations, where cluster groups were not defined, was Eq. [32] to predict site index and Eq. [11] to predict top height development. If the predicted site index (from Eq.[32]) was substituted into Eq. [20], which is specific to cluster groups, a marked improvement in accuracy is evident. Of the four combinations (i.e. Eq. [28] & Eq. [20]; Eq. [30] & Eq. [20]; Eq. [31] & Eq. [20]; Eq. [32] & Eq. [20]), Eq. [30] & Eq. [20] is the most accurate. However, bias is evident with site index when this combination is used.

On the basis of residual statistics from the validation of the equations used in combination and the performance of equations when used individually, it is recommended that Eq. [32] is used in combination with Eq. [11] when predicting top height development from environmental attributes.

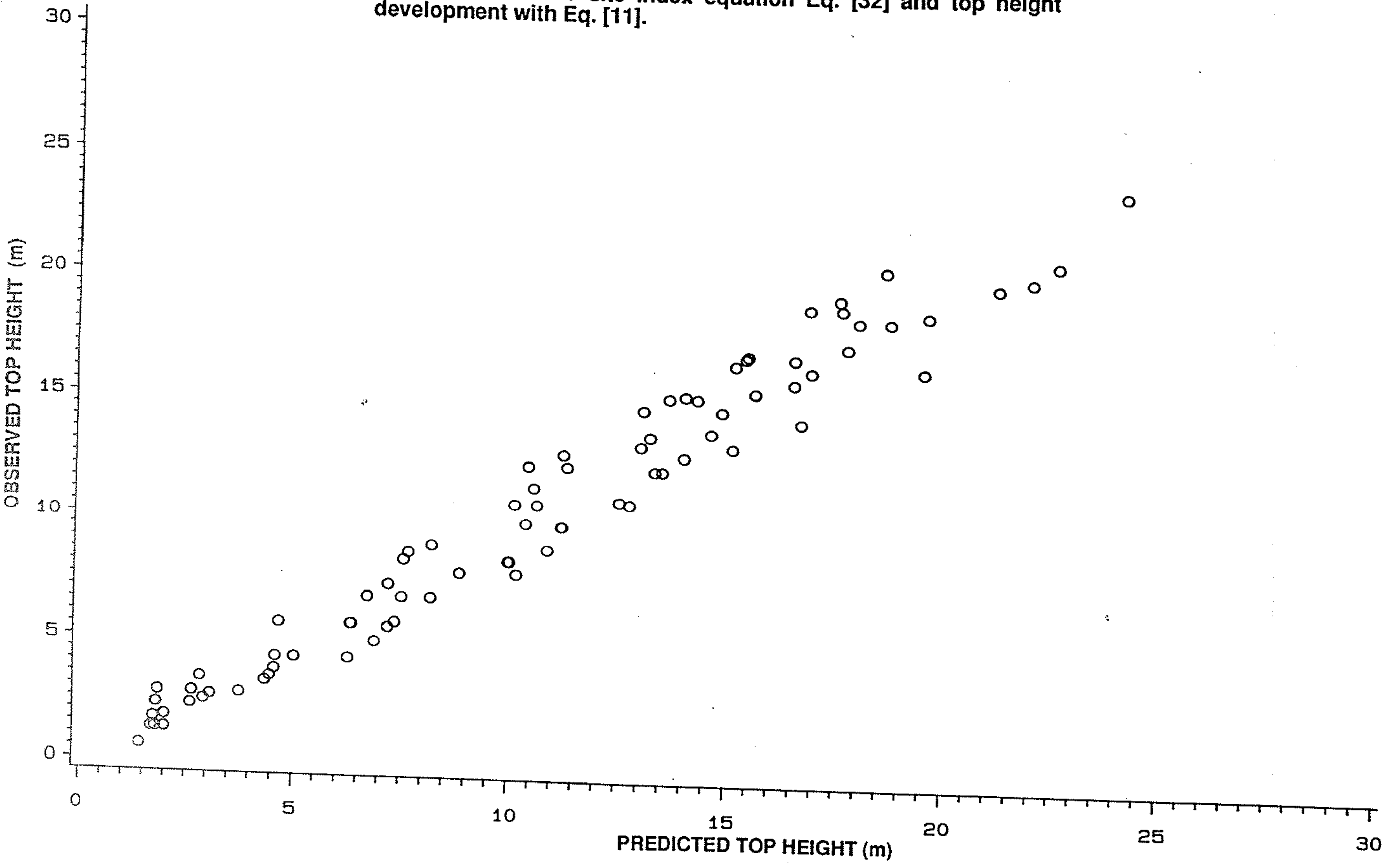
Linear regression, using the observed top height as the dependent variable and predicted top height as the independent variable, yielded intercepts which were not significant for the Eq. [32] & Eq. [11] ($p>0.1370$) combination. A plot of observed top height by predicted top height is given in Figure 25.

Table 35: Residual statistics generated from the validation of top height development equations and equations which predict site index from environmental attributes (n = 83).

Equation	$\overline{D}$	D_{SD}	r^2
Eq.[28] + Eq.[11]	-0.54	1.22	-0.9598
Eq.[30] + Eq.[11]	-0.41	1.21	0.9606
Eq.[31] + Eq.[11]	-0.50	1.27	0.9567
Eq.[32] + Eq.[11]	-0.10	1.21	0.9614
Eq.[28] + Eq.[11]*	-0.08	1.21	0.9616
Eq.[30] + Eq.[11]*	0.01	1.25	0.9587
Eq.[31] + Eq.[11]*	-0.05	1.33	0.9538
Eq.[32] + Eq.[11]*	0.23	1.51	0.9601
Eq.[28] + Eq.[18]	-0.64	1.29	0.9569
Eq.[30] + Eq.[18]	-0.58	1.23	0.9594
Eq.[31] + Eq.[18]	-0.61	1.33	0.9532
Eq.[32] + Eq.[18]	-0.20	1.24	0.9587

* equation applied to each cluster group where the plot was allocated to a cluster group via heuristic rules and logistic regression.

Figure 25: **Observed top height by predicted top height derived by predicting site index with the site index equation Eq. [32] and top height development with Eq. [11].**



5.2.4 DISCUSSION

5.2.4.1 TOP HEIGHT DEVELOPMENT EQUATIONS

It is surprising that the validation statistics pertaining to Eq. [11] were better than other candidate top height development equations. Based upon the residuals generated from calibrating the equations on the calibration data set (see Section 3.2.3.4) Eq. [18], Eq. [20] and Eq. [18] applied to each cluster group, displayed the most desirable residuals. Likewise, Eq. [18] has qualities as a top height development equation, not possessed by Eq. [11] (see Section 3.2.4.3). Nonetheless, Eq. [11] has been used with success by other authors (Carmean and Hahn 1981; Carmean and Lenthall 1989). Smith and Watts (1987) found the equation gave the best validation statistics of the seven equations they tested for their ability to predict top height development.

The allocation of plots to cluster groups and applying cluster specific top height equations to each cluster did not improve the validation statistics. This is a surprising result given the polymorphic nature of the top height data set and Monserud's (1984) conclusion, that the Ek-Payendah functional form lacked the flexibility to track the polymorphism in his data set. A possible explanation is that the two allocation procedures, used in this study, were unable to consistently allocate independent plots to appropriate cluster groups, thus incurring errors.

Two concerns exist when using the Ek-Payendeh equation, firstly, predicted top height does not equal site index at the reference age and secondly, the equation is not reference age invariant. In this study both problems do not seem to incur an error which is of any practical importance.

5.2.4.2 SITE INDEX EQUATIONS

The failure of the alternative form of Ek's (1971) and Payendeh's (1974) modification of the Chapman-Richard's functional form Eq. [13] to yield appropriate validation statistics is not uncommon. Smith and Watts (1987) found that this functional form yielded the worst validation statistics of the seven site index equations they validated and concluded that the algebraic difference site index equation yielded the best. Monserud (1984) concludes that this functional form forces on the data an equation that could not represent the expected model behaviour and that where the equation has been used successfully it has been fitted to the inverse of published height growth curves (Payendeh 1974; Monserud and Ek 1976) rather than to actual site index data.

In this study the algebraic difference equation yields the best validation statistics. Similar results are reported by Smith and Watts (1987). However, unlike the top height development equation, the allocation of plots to cluster groups and application of cluster specific versions of Eq. [19] did improve the accuracy of the validation statistics. However, increased accuracy was at the cost of decreased precision. As the allocation of plots to cluster groups by heuristic rules and logistic regression is strongly influenced by site index, an increased accuracy is not surprising. However, the decrease in the level of precision is unexpected and possibly due to the failure of the allocation procedures.

Improvements in both accuracy and precision occurred if data from years one and two were excluded from the validation data set. This may be due to the influence of establishment practice on height growth during the early stages of stand development. Three years seem to be required for the stand to reflect the influences of the site.

5.2.4.3 THE PREDICTION OF SITE INDEX FROM ENVIRONMENTAL ATTRIBUTES

The most accurate and precise equations were those comprised of all attribute types and containing multiplicative interactive terms. This was also the case if model calibration statistics are examined (see Section 4.3.3). This is not surprising given the scale of the study and the synergistic nature of the influences upon *E. globulus* growth across the study area.

The equation derived via the stepwise procedure (Eq. [29]) yielded validation statistics which are unacceptable. With the use of stepwise procedures variables are selected on the basis of individual contributions to the model sums of squares. As such the equation may be subject to prediction bias (Verbyla 1986) and include attributes which are biologically insignificant (Verbyla and Fisher 1989). Although this equation yields the most desirable model statistics (Table 27) it yields the worst validation statistics, casting doubt upon the common practice of selecting such attributes by stepwise procedures.

Segregating the data set into land use histories improved the model statistics during model parameterisation (Table 27). However, no improvement was evident in the validation statistics after segregating the validation data set. A practical disadvantage of segregating the data in this manner is that it is often quite difficult to ascertain whether a property is freshly cleared and sown to pasture or whether it has been pasture for some time, thus casting doubt as to which equation is applicable.

5.2.4.4 TOP HEIGHT DEVELOPMENT EQUATIONS IN COMBINATION WITH EQUATIONS WHICH PREDICT SITE INDEX FROM ENVIRONMENTAL ATTRIBUTES

On the basis of the validation statistics derived for individual equations, one would expect that Eq. [28] combined with Eq. [11] would yield the best validation statistics. However, where cluster groups were not accounted for, this was not the case with Eq. [32] and Eq. [11] yielding a much improved level of accuracy over other equations (Table 35). Unlike the validation of individual equations, the allocation of plots to cluster groups via heuristic rules and logistic regression and the application of the cluster specific version of Eq. [11] in concert with Eqs. [28], [30], [31] and [32] markedly improved accuracy without a loss in precision. No explanation can be offered for this phenomenon given that it was not the case when equations were validated individually. However, this does serve to illustrate the importance of validating equations in concert as well as individually.

5.2.4.5 POLYMORPHIC NONDISJOINT TOP HEIGHT DEVELOPMENT

Combining plots of similar top height development patterns into cluster groups, using the two-dimensional profile algorithm and the UPGMA fusion strategy, then fitting top height development and site index equations to each cluster was a strategy which confined polymorphism to that which exists between clusters. As such the calibration model statistics applied to these clusters were more desirable than those derived via standard methodologies. This approach is philosophically similar to that taken by Monserud (1984) who built habitat type into his top height and site index equations, to account for such polymorphism. However, habitat type was quite assessable in Monserud's study (Pfister and Arno 1980) unlike cluster groups used here.

The strategy of allocating plots to cluster groups via discriminant functions or heuristic rules and logistic regression did not generally improve the validation statistics as expected. This may be because:

(i) the polymorphism encapsulated in clusters, derived via cluster analysis, is a mathematical artifact and bears little resemblance to reality; or,

(ii) the techniques employed to allocate independent plots to cluster groups did so without the level of accuracy required to use such specific equations.

On the basis of the multivariate data exploration (see chapter 4) I do not believe that the polymorphism is a mathematical artifact. Rather, the failure of this strategy is probably due to the allocation procedures. To derive multivariate logistic equations and discriminant functions based upon heterogeneous covariance matrices, many degrees of freedom are required (see Section 4.2.2.4). As such these techniques could not be fully explored in this study. I therefore conclude that algorithms specific to individual clusters are not appropriate until further research is undertaken to derive and test more appropriate allocation techniques.

5.2.5 CONCLUSION

On the basis of validation statistics it is recommended that Eq. [11] be used for the prediction of top height given stand age and site index and Eq. [19] be used for the prediction of stand site index given top height and age. Eq. [19] should only be applied to stands of age three or older. When predicting site index from environmental attributes Eq. [28] is the most appropriate. When site index is estimated from environmental attributes and used as input into top height equations Eq. [32] and Eq. [11] are recommended.

5.2.6 REFERENCES

Borders, B.E., and W.D. Patterson. 1990. Projecting stand tables: A comparison of the Weibull diameter method, a percentile-based projection method, and a basal area growth projection method. *For. Sci.*, 36:413-424.

Blyth, J.F., and D.A. MacLeod. 1981. Sitka spruce (*Picea sitchensis*) in north-east Scotland. II. Yield prediction by regression analysis. *Forestry*, 54:63-73.

Broadfoot, W.M. 1969. Problems in relating soil to site index for southern hardwoods. *For. Sci.*, 15:354-364.

Carmean, W.H., and J.T. Hahn. 1981. Revised site index curves for balsam fir and white spruce in the Lake States. *USDA Res. Note NC-269.*

Carmean, W.H., and D. J. Lenthall, 1989. Height-growth and site-index curves for jack pine in north central Ontario. *Can. J. For. Res.*, 19:215-224.

Caswell, H. 1977. The validation problem. *In*: Systems analysis and simulation in ecology, B.C. Pattem (ed.). Vol. IV. Academic Press, New York.

Cieszeweski, C.J., and I.E. Bella. 1989. Polymorphic height and site index curves for lodgepole pine in Alberta. *Can. J. For. Res.*, 19:1151-1160.

Devan, J.S., and H.E. Burkhart. 1982. Polymorphic site index equations for loblolly pine based on a segmented polynomial differential model. *For. Sci.*, 28:544-555.

Dolph, K.L. 1989. Height-diameter equations for young-growth red fir in California and Southern Oregon. USDA *Res. Note*. PSW-408.

Dyer, M.E., and R.L. Bailey. 1987. A test of six methods for estimating true heights from stem analysis data. *For. Sci.*, 33:3-13.

Efron, B. 1983. Estimating the error rate of a prediction rule: improvement on cross-validation. *J. Am. Stat. Assoc.*, 78:316-331.

Ek, A.R. 1971. A formula for white spruce site index curves. *Univ. Wis. For. Res. Note*, No 161.

Ek, A.R., and R.A. Monserud. 1979. Performance and comparison of stand growth models based on individual tree and diameter-class growth. *Can. J. For. Res.*, 9:231-244.

Evert, F. 1981. A model for regular mortality in unthinned white spruce plantations. *For. Chron.*, 57:77-79.

Frank, P.S., R.R. Hicks, and E.J. Harner. 1984. Biomass predicted by soil-site factors: a case study in north central West Virginia. *Can. J. For. Res.*, 14:137-140.

Freese, F. 1960. Testing accuracy. *For. Sci.*, 6:139-145.

Gertner, G. 1986. Postcalibration sensitivity procedure for regressor variable errors. *Can. J. For. Res.*, 16:1120-1123.

Gertner, G. 1987. Approximating precision in simulation projections: an efficient alternative to Monte Carlo methods. *For. Sci.*, 33:230-239.

Goodall, D.W. 1972. Building and testing ecosystem models. *In*: Mathematical models in ecology, J.N.R. Jeffers (ed.). Blackwell Sci. Publ. Oxford.

Gong, G. 1986. Cross-validation, the jacknife, and the bootstrap: excess error estimation in forward logistic regression. *J. Am. Stat. Assoc.*, 81:108-113.

Gregoire, T.G., and M.R. Reynolds. 1988. Accuracy testing and estimation alternatives. *For. Sci.*, 34:302-320.

Harding, R.B., D.F. Grigal, and E.H. White. 1985. Site quality evaluation for white spruce plantations using discriminant analysis. *Soil. Sci. Soc. Am. J.*, 49:229-232.

Holdaway, M.R., and G.J. Brand. 1983. An evaluation of the STEMS tree growth projection system. *USDA Res. Pap.* NC-234.

House, P.B. 1974. Diogenes revisited - the search for a valid model. *Simulation*, 22:117-125.

Lappi, L., and R.L. Bailey. 1988. A height prediction model with random stand and tree parameters: an alternative to traditional site index methods. *For. Sci.*, 34:907-927.

Lenhart, J.D. 1988. Evaluation of explicit and implicit yield prediction in loblolly and slash pine plantations in east Texas. *In*: Forest growth modelling and prediction, A.R. Ek, *et al.* (eds.). USDA. *Gen. Tech. Rep.* NC-120.

McQuilkin, R.A. 1976. The necessity of independent testing of soil-site equations. *Soil. Sci. Soc. Am. J.*, 40:783-785.

Monserud, R.A., 1984. Height growth and site index curves for inland Douglas-fir based on stem analysis data and forest habitat type. *For. Sci.*, 30:943-965.

Monserud, R.A., and A.R. Ek. 1976. Site curves and equations for several northern hardwood forest species. *Coll. Agric. and Life. Sci. Res. Bull.*, R-2772. Univ. Wis., Madison.

Moser, J.W., and O.F. Hall. 1969. Deriving growth and yield functions for uneven-aged forest stands. *For. Sci.*, 15:183-188.

Newberry, J.D., and A.R. Stage. 1988. Validating forest growth models: procedures defined by resource decisions. *In*: Forest growth modelling and prediction, A.R. Ek, *et al.* (eds.). USDA *Gen. Tech. Rep.* NC-120.

Newnham, R.M. 1988. A modification of the Ek-Payendah nonlinear regression model for site index curves. *Can. J. For. Res.*, 18:115-120.

Patterson, D.A., and C.T. Stiff. 1988. An evaluation and comparison of two distance-independent forest projection models in the Inland Northwest. *In*: Forest growth modelling and prediction, A.R. Ek, *et al.* (eds.). USDA *Gen. Tech. Rep.* NC-120.

Page, G. 1976. Quantitative evaluation of site potential for spruce and fir in Newfoundland. *For. Sci.*, 22:131-143.

Payendeh, B. 1974. Nonlinear site index equations for several major Canadian timber species. *For. Chron.*, 50:194-196.

Pearson, H.A., and H.S. Sternitzke. 1974. Forest-range inventory: a multiple-use survey. *J. Range. Manage.*, 27:404-407.

Pfister, R.D., and S.F. Arno. 1980. Classifying forest habitat types based on potential climax vegetation. *For. Sci.*, 26:52-70.

Rennie, J.C., and H.V. Wiant. 1978. Modification of Freese's chi-square test of accuracy. USDA Bur. Land Manage. Res. Inventory, Note BLM 14.

Reynolds, M.R. 1984. Estimating the error in model predictions. *For. Sci.*, 30:454-469.

Schaeffer, D.L. 1980. A model evaluation methology applicable to environmental assessment models. *Ecol. Modelling*, 8:275-295.

Schmidt, M.G., and W.H. Carmean. 1988. Jack pine site quality in relation to soil and topography in north central Ontario. *Can. J. For. Res.*, 18:297-305.

Shoulders, E., and A.E. Tiarks. 1980. Predicting height and relative performance of major southern pine from rainfall, slope and available soil moisture. *For. Sci.*, 26:437-447.

Smith, V.G., and M. Watts. 1987. An assessment of the structural method of deriving a black spruce site equation.*Can. J. For. Res.*, 17:1181-1189.

Stage, A.R., and D.L. Renner. 1988. Comparison of yield-forecasting techniques using long-term stand histories. *In*: Forest growth modelling and prediction, A.R. Ek, *et al.* (eds.). Vol. 2. USDA *Gen. Tech. Rep.* NC-120.

Verbyla, D.L. 1986. Potential prediction bias in regression and discriminant analysis. *Can. J. For. Res.*, 16:1255-1257.

Verbyla, D.L., and R.F. Fisher. 1989. Effect of aspect on ponderosa pine height and diameter growth. *For. Ecol. Manage.*, 27:93-98.

West, P.W. 1981. Simulation of diameter growth and mortality in regrowth eucalypt forests of southern Tasmania. *For. Sci.*, 27:603-616.

West, P.W. 1983. Accuracy and precision of volume estimates from a simulation model for regrowth eucalypt forest.*Aust. For. Res.*, 13:183-188.

White, E.J. 1982. Relationships between height growth of Scots pine (*Pinus sylvestris* L.) and site factors in Great Britain.*For. Ecol. Manage.*, 4:225-245.

Zuuring, H.R., J.D. Arney, and K.S. Milner. 1988. Generic graphical analysis for tree growth model invalidation. *In*: Forest growth modelling and prediction, A.R. Ek, *et al.* (eds.). USDA. *Gen. Tech. Rep.* NC-120.

CHAPTER SIX

Construction of Empirical Equations for Estimating Plantation Yield

6.1 LITERATURE REVIEW

6.1.1 INTRODUCTION

The need to estimate forest growth and yield has long been recognised. Prior to the advent of computer technology the most common method of obtaining such estimates was from yield tables and alignment charts (Schlich 1925; Tesch 1981). Today more sophisticated, computer based projection systems are in use and are commonly termed *growth models*. Due to the concurrent development of many kinds of growth models in various parts of the world, a confused set of terminology is evident in the literature. For consistency, the terminology of Bruce and Wensel (1988) will be adopted for this study. As such a growth model is defined as a mathematical function, or system of functions, used to relate actual growth rates to measured tree, stand, and site variables.

A review of the historical development of growth and yield modelling is given by Tesch (1981), while a categorisation of the historical development stages has been delivered by Moser (1980). Moser (1980) suggests the following historical chapters in the development of growth and yield models:

- (i) yield tables;
- (ii) yield functions;
- (iii) compatible growth and yield functions;
- (iv) diameter distribution approaches; and
- (v) simulation approaches.

6.1.2 A CLASSIFICATION OF GROWTH MODELS

A vast variety of growth models appear in the literature, each reflecting different silvicultural practices, modelling philosophies, applications and mathematical complexity. Any classification of growth models can therefore be based on a variety of characteristics. For example, Burkhart

(1977a) classified growth models as either stochastic or deterministic, where stochastic growth models include random variables whose values are generated, whereas deterministic models do not. Other categories can be defined by classes of predictive functions and/or statistical techniques used in their definition (Lowell and Mitchell 1987). Bruce and Wensel (1988) argue that growth models should be categorised according to the condition of the forest to which they apply and the purpose of the model.

A more detailed classification is offered by Clutter *et al.* (1983) who suggests that models be classified on the basis of the target populations to which the predictions will apply. Therefore models are classified as those pertaining to:

(i) natural forests;

  (a) uneven-aged,

  (b) even-aged,

(ii) plantation forests;

  (a) thinned,

  (b) unthinned.

Alternatively, Clutter *et al.* (1983) suggests models can be classified by the complexity of the mathematical approach involved. thus models may be classified as:

(i) models in tabular form:

(ii) models as equations and systems of equations:

  (a) direct prediction of unit area values,

  (b) unit area values obtained by summation.

    (1) equations for classes of trees,

    (2) equations for individual trees.

Clutter *et al.* (1983) also makes the important distinction between explicit and implicit growth models. Where the solution of the model provides estimates of volume per unit area, the model is explicit. With implicit models the solution produces basic information on stand structure. Yield estimates implied by the predicted structure are then calculated from further computations based on the stand structure information.

The first and most frequently utilized classification of models is that of Munro (1974). Munro presents three classes of models. The first, stand growth models, are those where the independent variables are stand characteristics such as age, site index and basal area. Secondly, single tree distance dependent models require single tree input as well as some measure of inter-tree distance. Finally, single tree distance independent models require single tree input but no measure of inter-tree distance is required. As Munro's classification has gained acceptance and is frequently used in the literature it will be adopted for the purposes of this review. However, when describing a model Clutter's *et al.* (1983) definition of implicit and explicit methods will accompany Munro's (1974) classification.

6.1.3 STAND GROWTH MODELS

The approaches to modelling stand level growth and yield are many and varied and include the use of differential equations (Clutter 1963), generalised least squares (Ferguson and Leech 1978), logistic regression analysis (Lowell and Mitchell 1987), diameter distribution projections using probability density functions (Bailey 1980) or Markov chains (Peden *et al.* 1973) or simple linear regressions (Payendeh 1990). It should be noted that some stand level models, such as diameter distribution models may produce some tree level output, such as the frequency of trees per dbh class. Although, Garcia (1988) classed such models as distant independent individual tree growth models, they are still classified as stand level models in the majority of cases because the independent variables are stand level statistics (Burkhart 1977a; Clutter *et al.* 1983).

6.1.3.1 EXPLICIT PREDICTION OF GROWTH AND YIELD

Originally normal yield tables were used for the estimation of current and future yields. These tables were then generalised by including a measure of stand density as a third independent variable in addition to site index and age (Schumacher 1939). Nowadays, such equations estimate current yield from functions of variables such as site index, age, stocking, density and dominant height (Bennett 1970; Sullivan and Clutter 1972; Chambers 1980; Sadiq and Beckwith 1986; Payandeh 1990). Such equations usually estimate the yield in the units of interest, for example, volume per acre, outside bark of all trees of 4.6 inches dbh and larger to a top diameter of 4 inches outside bark (Bennett 1970). This approach is adequate where the resource is used for a single product with defined merchantable limits. However, where the resource is used for a variety of products the yield will vary depending upon which product is being estimated. Many studies simply predict total volume which is subsequently divided into product classes (Burkhart *et al.* 1972). However, such a strategy may yield illogical estimates for certain combinations of independent attributes. An alternative approach utilizes the strategy employed in the single tree volume equation derived by Burkhart (1977b) and Van Deusen *et al.* (1981). A ratio equation was developed where the merchantable yield is a function of total yield, quadratic mean diameter, stocking, top diameter merchantability limit and threshold dbh limit (Amateis *et al.* 1986; Matney *et al.* 1988).

The explicit prediction of future yield requires the independent variables of an equation, which explicitly predicts current yield, to be projected to some point into the future. For example, Schumacher's variable density yield equation as given by Pienaar and Shiver (1986);

$$ln(Y) = \beta_0 + \beta_1\left(\frac{1}{A}\right) + \beta_2 f(S) + \beta_3 g(D)$$

Where;

Y	=	yield/unit area
A	=	stand age
f(S)	=	some function of site quality
g(D)	=	some function of stand density
ln	=	natural logarithm

As presented this equation may be viewed as one which explicitly predicts current yield. However, if S is expressed as top height and D as basal area, future values of S and D may be obtained from top height and basal area development functions and substituted into the above equation. As such a future value of Y is obtained.

The nature and functional forms of the equations which describe the development of the independent stand attributes are many and varied. The most frequently projected attribute is basal area. Usually projected basal area is a function of age and the basal area at the beginning of the projection period (Clutter and Jones 1980). Other equations may include site index (Borders and Bailey 1986), stocking (Chojnacky 1988) and dominant height (Varmola 1988) in various combinations (Pienaar and Shiver 1986). Bailey and Ware (1983) developed a basal area development equation in which they included a variable to reflect the type of thinning applied to stands.

It is argued that stand models which predict future yield must take into account future mortality (Hamilton and Edwards 1976). Where the yield equation or basal area projection equations use stocking as an independent variable, changes in stems per unit area over time must be accounted for. Many mortality functions have been constructed (Lee 1971; Ek 1974; Stiell 1974; Monserud 1976; Evert 1981; Buchman *et al.* 1983; Rennolls and Peace 1986) however, the majority predict stocking at some future time as a function of present time, and current stocking (Clutter and Jones 1980; Pienaar and Shiver 1981). An alternative approach involves the prediction of the probability of mortality (Hamilton 1974) but is more applicable to diameter class or single tree equations. Equations to predict the future values of top heights have been covered in Chapter 3.

Another approach to predicting future yields involves the prediction of volume increment and summing the increments to obtain yield (Oliver 1979).

6.1.3.1.1 PARAMETER ESTIMATION

Growth and yield are viewed as functions of site quality, age and density as well as the interactions between these variables. Stand density is also taken to be a function of site quality, age and initial measurements of stand density. Site quality is usually a reference to height and age. Each of these equations describes a different relationship, and all are assumed to hold simultaneously (Borders and Bailey 1986). However, when deriving such models the parameters are estimated using ordinary least squares on individual equations in isolation. As such the sum of individual increments may give different results to the solution of a yield equation (Clutter 1963).

As a result of such inconsistency Clutter (1963) proposed the idea of compatibility among growth and yield equations such that when the growth function is integrated over a time interval, the yield is obtained for the end of the projection interval. The benefits of such compatibility, as listed by Clutter (1963), are logical consistency and that the knowledge of existing yield models will suggest appropriate growth models upon differentiation. Clutter's (1963) proposal of compatibility has been well accepted and is now common in stand level growth models (Pienaar and Turnbull 1973; Borders and Bailey 1986; Pienaar and Shiver 1986; Borders 1989).

Estimations of the parameters of stand level growth models are subject to a number of concerns. Firstly, most parameters are estimated from repeated measures of permanent inventory plots. Such time series data proposes statistical problems as discussed in Section 3.1.7. Secondly, equations which explicitly predict future yields are usually a system of interrelated equations such as those proposed by Clutter (1963). Parameters of such equations are usually estimated by ordinary least squares applied separately to each equation comprising the system. Such systems are classed as recursive or simultaneous equations under the definition of Pindyck and Rubinfeld (1981) and Borders (1989). Where ordinary least squares are used to estimate interdependent multi-equation systems, resulting estimates may not be the most efficient.

Attempts to solve this problem are many. For example, Sullivan and Clutter (1972) estimated the parameters of the equations, defined by Clutter (1963), simultaneously by substituting the predicted values of basal area, in place of actual basal area, into the yield prediction equation prior to estimating its parameters. The final equation was fitted using ordinary least squares and maximum likelihood. A similar approach was employed by Burkhart and Sprinz (1984). Furnival and Wilson (1971) used econometric techniques (Fomby *et al.* 1984) to fit multi-equation models. Recently, the application of such techniques, usually termed two or three stage least squares, has become common when addressing such problems (Murphy and Sternitzke 1979; Murphy and Beltz 1981; Amateis *et al.* 1984; Borders and Bailey 1986; Gregoire 1987; Borders 1989).

Although the problem of parameter estimation of interdependent models is receiving more attention in the recent literature, the error structure associated with such models is poorly understood. With the exception of Bailey (1981) and Borders and Bailey (1986) this topic has received scant attention thus inhibiting the use of confidence intervals around yield predictions.

The problems inherent within parameter estimations of interrelated equations are at least of theoretical concern. However, as Borders (1989) points out, parameter estimation is the final step in the modelling process. Prior to this, much time and effort must be put forth for data acquisition and development of theory upon which mathematical models are based. Any gains in predictive ability are much more likely to be attributable to appropriate data bases and sound growth theory than to theoretically sound parameter estimation procedures.

6.1.3.2 IMPLICIT PREDICTION OF STAND LEVEL GROWTH AND YIELD

On a stand level the implicit prediction of current and future yields involves the use of diameter distribution methods. The frequency of trees per unit area by diameter class is represented by some probability density function. Diameter distribution methods provide more information on stand structure than the explicit methods previously discussed. Such information is used for determining the value of raw material, harvesting costs, product mixes and for deriving management plans (Hyink and Moser 1983). These techniques are generally applied to even-aged plantation yields (Bennett and Clutter 1968; Burkhart and Strub 1974; Smalley and Bailey 1974; Alder 1979; Bailey 1980; Pienaar and Harrison 1988) however, application to uneven-aged stands have also been reported (Schreuder *et al.* 1979; Hyink and Moser 1983).

Basically, the method represents the diameter distribution of a stand with some density function, such as the Weibull probability density function (Bailey and Dell 1973). The frequency of trees in each diameter class is then derived and their volumes estimated from single tree volume equations, such as those discussed in Chapter 2. Prior to estimating the tree volumes, tree heights are estimated using height-diameter equations, such as those presented by West (1982a), Zakrzewski and Bella (1988), Dolph (1989a) and Borders and Patterson (1990). The parameters of the probability density function are usually represented as functions of stand attributes such as site index, age, basal area, etc. Therefore, the implicit prediction of future yields merely involves the projection of those stand attributes used to estimate the parameters, as was the case for the explicit prediction of future yields (Smalley and Bailey 1974).

The choice of probability density function to represent diameter distribution is arbitrary. There is no biological basis on which to argue for or against a particular functional form (Borders *et al.* 1987). Early work in diameter distribution models used the exponential distribution (Leak 1965; Schmelz and Lindsay 1965). This function represents negative J-shaped distributions. Functions which have been used to represent mound shaped distributions include the gamma distribution (Nelson 1964; Bailey 1980), the three parameter lognormal distribution (Bliss and Reinker 1964), the beta distribution (Lenhart and Clutter 1971; Goodwin and Candy 1986) and Johnson's S_B distribution (Hafley and Schreuder 1977; Tham 1988). However, since Bailey and Dell's (1973) paper, the Weibull probability density function (Weibull 1951) has been utilized more than any other function for even-aged diameter distributions (Schreuder *et al.* 1979; Matney and Sullivan 1982; Zutter *et al.* 1986; Shiver 1988). Its popularity stems from its flexibility and the fact that a closed form cumulative density function can be derived.

The three parameter function has a probability density of the form;

$$f(x) = \frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1} . \exp . \left(-\left(\frac{x-a}{b}\right)^{c}\right)$$

for $x \geq = a$, $a \geq = 0$
$b > 0$
$c > 0$

Where;

x = dbh
a = location parameter
b = scale parameter
c = shape parameter

To derive a model using this methodology a probability density function is fitted to each plot such that a matrix of parameter estimates is obtained. These parameters are then expressed as functions of stand characteristics (Smalley and Bailey 1974; Burk and Burkhart 1984). This approach is termed the parameter prediction method (Hyink and Moser 1983). However, such parameters are usually highly correlated and vary independently to stand characteristics, making it difficult to develop predictive equations that explain a high percentage of the variation in the parameters. Consequently, stand table statistics such as basal area, quadratic mean diameter and diameter percentiles are related to analytical equivalents for a probability density function (Dubey 1967). This technique is termed the parameter recovery method (Hyink and Moser 1983) and usually involves the use of percentiles for obtaining the probability density function's parameter estimates. The technique is based upon the fact that, if three sample percentiles are known, each can be equated to its corresponding Weibull cumulative distribution function and the three equations solved interactively (Clutter *et al.* 1983). The problems caused by correlations between parameter estimates are overcome with this method and the copious calculations involved with fitting a probability density function to each plot are unnecessary. Borders *et al.* (1987) argue that the stand characteristics used in the parameter recovery method really are the variables of interest and that predicting highly variable probability density function parameters is an unnecessary and unsatisfactory step in estimating a stand table.

Borders *et al.* (1987) developed a method which does not rely on a predefined probability density function. To characterise the diameter distribution, they define an empirical probability density function based on 12 percentiles. The 55th percentile is a function of the quadratic mean diameter. The other percentiles are functions of adjacent percentiles and in some cases quadratic mean diameter. This method assumes a uniformed distribution of tree frequency between adjacent percentiles and is termed the percentile method by Borders and Patterson (1990).

Another approach to implicit stand level modelling, which does not require a predetermined probability density function is that proposed by Pienaar and Harrison (1988). This procedure projects an individual tree diameter list into the future based on a hypothesis concerning the expected change in relative tree size over time. Relative size is defined as the ratio of tree basal area to the mean tree basal area and is an extension of the earlier work of Clutter and Jones (1980). Mortality and basal area growth equations are also required. Deriving such a system requires time series data, where each tree is uniquely identified.

In a comparison of the parameter recovery method, percentile method and the method of Pienaar and Harrison (1988), it was shown that the basal area growth projection method of Pienaar and Harrison (1988) was superior in accuracy and precision (Borders and Patterson 1990).

Another method of projecting diameter class frequencies is the use of Markov chains (Bruner and Moser 1973; Peden *et al.* 1973). With this methodology the probability of a tree progressing from one diameter class to the next is derived. The probabilities are assembled into a transition matrix. Mortality and harvest removals may also be built into these probabilities (Bruner and Moser 1973).

6.1.3.2.1 PARAMETER ESTIMATION

As the estimation of the parameters of a probability density function requires iterative computations some earlier works chose to transform the functions such that the estimators were then derived as for linear regression (Bain and Antle 1967). However, this procedure is absent from the recent literature.

Other parameter estimation technique include the percentile estimation procedures put forward by Zanakis (1979). Zarnoch and Dell (1985) compared the percentile estimation procedure to that of maximum likelihood estimation and found the percentile estimators to be biased but with smaller variances. Moment estimation has also been used to estimate the parameters of a Weibull function (Garcia 1981). In a comparison of maximum likelihood, the percentile estimation and moment estimation procedures, Shiver (1988) found that maximum likelihood estimation provided the best estimates of the parameters of known distributions.

In a noteworthy approach, the parameters for the equations which comprise the percentile method of Borders *et al.* (1987) were first estimated with ordinary least squares. However, the errors were found to be contemporaneously correlated. Therefore, the parameters were re-estimated using seemingly unrelated regression (Zellner 1962).

Although there are numerous techniques of estimating the parameters of a probability density function it is generally accepted that maximum likelihood estimators are best (Bailey and Dell 1973; Shiver 1988). When deriving equations which project driving variables, such as basal area etc, into the future the parameter estimation procedures suffer from the same problems as those encountered with explicit predictions, as discussed in Section 6.1.3.1.1.

6.1.4 SINGLE TREE GROWTH MODELS

A vast variety of single tree growth models have appeared in the literature. All have the requirement of time series data for model parametrization. Therefore, single tree growth models are not an option for this study. The copious literature concerned with the topic will, therefore, not be fully reviewed.

6.1.4.1 SINGLE TREE DISTANCE DEPENDENT MODELS

As the output of single tree distance dependent models are the sum of individual tree growth and yield, these models are implicit. The models require a tree list, individual tree data such as dbh, height etc. and information on between tree distances, usually inputted as a set of x-y coordinates, as major components of input data. Single tree distance dependent growth models are based on the postulation that if the competitive influences on an individual tree, induced by its neighbours, can be quantified greater detail on tree development is provided. The amount of competition to which a tree is subject is often expressed as being proportional to the amount that the "competitive circle" of the subject tree is overlapped by the competitive circles of neighbouring trees.

There has developed an abundance of techniques for determining the radius of the competitive circle and for calculating overlap areas and weighing these by the relative sizes of competitors. These procedures may be based on the mean distance to various neighbours (Thompson 1956), angles (Lin 1974) and units of area (Ek and Monserud 1979) to name a few (c.f. Opie 1968; Bella 1971; Ek and Monserud 1974; Daniels 1976).

The application of single tree distance dependent growth models proceeds in the following order, as listed by Clutter *et al.* (1983):

(i) A competitive index value is calculated for each tree in the tree list;

(ii) Mortality is derived as functions of the competitive index;

(iii) Predicted periodic growth for each tree is calculated and added to the current size;

(iv) Steps (i), (ii) and (iii) are repeated interactively until the end of the projection period is arrived at;

(v) Individual tree volumes are calculated from the final tree size measurements and summed.

Single tree distance dependent growth models were favoured during the 1960's and 1970's because of their assumed ability to characterize spatial variation in planting, mortality, thinning and competition. The disadvantages of this approach are that it is computationally more expensive than other classes of models and the difficulties involved with obtaining spatial coordinates of tree positions, variables not often included in standard forest inventory. The fundamental assumption underlaying this approach involves the assertion that individual tree growth can be predicted more precisely if the sizes and locations of neighbouring competitors are known. However, various studies have suggested that the use of competition indices have contributed little or no improvement in growth prediction over that obtained using basal area as the measure of competition (Beck 1974; Daniels and Burkhart 1975).

6.1.4.2 SINGLE TREE DISTANCE INDEPENDENT MODELS

Single tree distance independent growth models use stand table data as input, in addition to some stand level data. Trees are grown individually or in groups of similar dbh classes to generate stand table data at some future point in time. No set of x-y coordinates of tree positions is used to calculate competitive effects, rather functions involving basal area are frequently used to modify the potential growth response (Alder 1979; Belcher *et al.* 1982). Other modifiers include measures of green crown size (West *et al.* 1982b) and crown competition factor (Krajicek *et al.* 1961; Arney 1985). The range of density measures available as growth modifiers are reviewed by Curtis (1970), West (1982b) and Payendeh and Ek (1986). Other extrinsic modifiers may also be used (Wykoff 1990).

This class of growth model provides for all the mixed size class, age class and species mixture capabilities of distance dependent approaches, without the need for the input of tree positions. As such they may utilize information from standard tagged tree plot inventories. Where the projection of the future stand table does not require iterative calculations little more computer time is required than for stand level models (Clutter and Allison 1974; Alder 1979; Clutter and Jones 1980). Other models of this class, designed to handle mixed species and/or uneven-aged stands, do require iterative computations and are more demanding of computer resources

(Stage 1973; Belcher *et al.* 1982). However, with the increasing advances in computer technology these considerations have become secondary.

The utility of the single tree distance independent approach is that much detail on stand composition and product size distributions are available. Also various silvicultural options may be incorporated into the algorithms (Lowell 1988; Knowe and Foster 1989).

6.2 CONSTRUCTION OF EMPIRICAL EQUATIONS FOR ESTIMATING THE YIELD OF PLANTATION GROWN *E. GLOBULUS*.

6.2.1 INTRODUCTION

The ability to estimate current and future yields, under various conditions of the stand, is critical to any plantation scheme. In south west Western Australia the State requires estimates of plantation yields so that annuity calculations can be made and uneconomic sites avoided (see Chapter 1).

Published growth and yield studies of *E. globulus* have generally used the single tree distance independent approaches. For example, Tomé (1988) developed a single tree distance independent growth model for *E. globulus* plantations in Portugal. Tomé (1988) also presents a stand level yield equation.

In Australia, West (1981) presents a single tree distance independent approach for estimating growth and yield of mixed species, even-aged regrowth forests in Tasmania. In these forests *E. globulus* is only one component. In West's study diameter increments of individual trees were expressed as functions of various tree and stand parameters. A mortality function which predicts the death of individual trees as a stochastic process is also presented. Other functions useful in West's simulator are presented in West (1979) and West (1982a).

In a less elaborate study, also in the regrowth forests of Tasmania, Goodwin and Candy (1986) use both a single tree distance independent and a diameter distribution approach. Their study was based on a limited data set from a single stand comprised of *E. globulus* only. Equations for top height development and individual tree diameter increments are presented. The diameter distribution was modelled via the parameter prediction methodology using a beta distribution. Their mortality function was based on the 3/2 self thinning rule because of a lack of data to develop empirical functions.

For short rotation *E. globulus* plantations in Rwana, Gasana (1983) has derived a set of explicit equations for the prediction of current yield. After testing many functional forms, the final equation predicted stand volume as a function of stand age, site index, stocking and density percentage. Density percentage was defined as the average spacing between trees as a percentage of site index. Stand basal area was also expressed as a function of these parameters.

No particular modelling philosophy can be argued to be the most appropriate for *E. globulus* with the final choice being influenced by the purpose of growth and yield predictions, resources and quantity and quality of the data.

It is the aim of this study to define a method for the prediction of current and future yields for *E. globulus* plantations in south west Western Australia.

6.2.2 METHODS

6.2.2.1 RATIONALE BEHIND THE CHOICE OF MODELLING PHILOSOPHY

After much consideration of the topic it was decided to pursue a stand level, explicit modelling philosophy for the following reasons. Firstly, with the exception of the top height development data, presented in Chapter 3, all other data stem from static point measurements. To apply the single tree modelling philosophies time series data is obligatory. Although time series data will eventually become available, this study was restricted to a single measurement.

Secondly, only pulpwood will be harvested from the plantation estate. As such the piece size distribution and product differentiation capabilities of implicit modelling philosophies are not essential. For planning harvesting operations the knowledge of piece size distributions would be of advantage. However, such planning will not occur for about eight years, by then time series data will be available and more elaborate models may be constructed.

6.2.2.2 DATASETS

Two data sets were used to develop the yield functions. The first resulted from the measurement of 100 temporary inventory plots established in the plantation estates owned by Bunnings Tree Farms Ltd. This dataset will be referred to as the Bunnings data set. Each tree was measured for diameter at breast height over bark (dbhob) with a standard diameter tape, to the nearest 0.1 cm. The heights of all trees were estimated with a sunnto clinometer to the nearest 0.1 m. 3795 trees were measured in the 100 plots.

A second data set was available as a result of establishing 58 permanent inventory plots for the purposes of this study. The plots were established by the Department of Conservation and Land Management (CALM) and will be referred to as the CALM data set. These plots were measured in winter 1987 and 55 plots were remeasured in winter 1988. Each tree was measured for dbhob with a standard diameter tape to the nearest 0.1 m. Height was measured to the nearest 0.1 m with a sunnto clinometer and bark thickness was measured at two points, at right angles at breast height, with a Swedish bark gauge. 3336 trees were measured in this fashion.

Although the plots were measured in two consecutive years the data were not treated as time series. As this increment period may be atypical the consecutive measurements were treated as independent.

6.2.2.3 BARK THICKNESS AND STAND VOLUME ESTIMATION

To estimate the volume of each plot individual tree volumes were estimated using Eq.[6]$_{(CF)}$ presented in Section 2.2. Plot volume is then the sum of all the individual tree volumes and is expressed as total merchantable volume under bark ha^{-1}. However, to apply Eq.[6]$_{(CF)}$ dbhub is required. The Bunnings data set lacked this measurement and therefore required estimates of bark thickness prior to estimating plot volumes.

6.2.2.3.1 ESTIMATION OF BARK THICKNESS

To develop a bark thickness equation the CALM dataset was used. Univariate statistics for dbhob, dbhub and two times bark thickness is given in Table 36.

To estimate dbhub two approaches are possible. The first is to develop an equation which predicts bark thickness from dbhob and/or tree and stand attributes (West 1979; West 1982a; Gordon 1983) or develop equations to estimate dbhub directly from dbhob (Monserud 1979; Dolph 1989b). After plotting bark thickness against dbhob and dbhub against dbhob the strategy of predicting dbhub directly from dbhob was pursued (Figure 26).

The literature was examined for candidate functional forms which may be appropriate for the estimation of dbhub and those selected are listed in Table 37. The CALM data set was then split into a calibration and validation data set by setting aside about one randomly selected tree per plot (n = 113 trees). After parametrization the final candidate equations were validated via the assumptions and methods presented in Chapter 5.

Parameter estimation was via the standard criteria of ordinary least squares. For the nonlinear equations the parameters were estimated via nonlinear ordinary least squares using the derivative free secant method of Ralston and Jennrich (1979). Most candidate equations displayed heterogeneity among residuals upon parametrization. As such the parameters were re-estimated via weighted least squares. Statistics derived via weighted least squares were identified by the subscript (WT) with the equation identification number. The variance assumption for weighted least squares was $\sigma^2 \propto 1/dbhob^2$ and was found to be a reasonable one during the data exploration phase of this study.

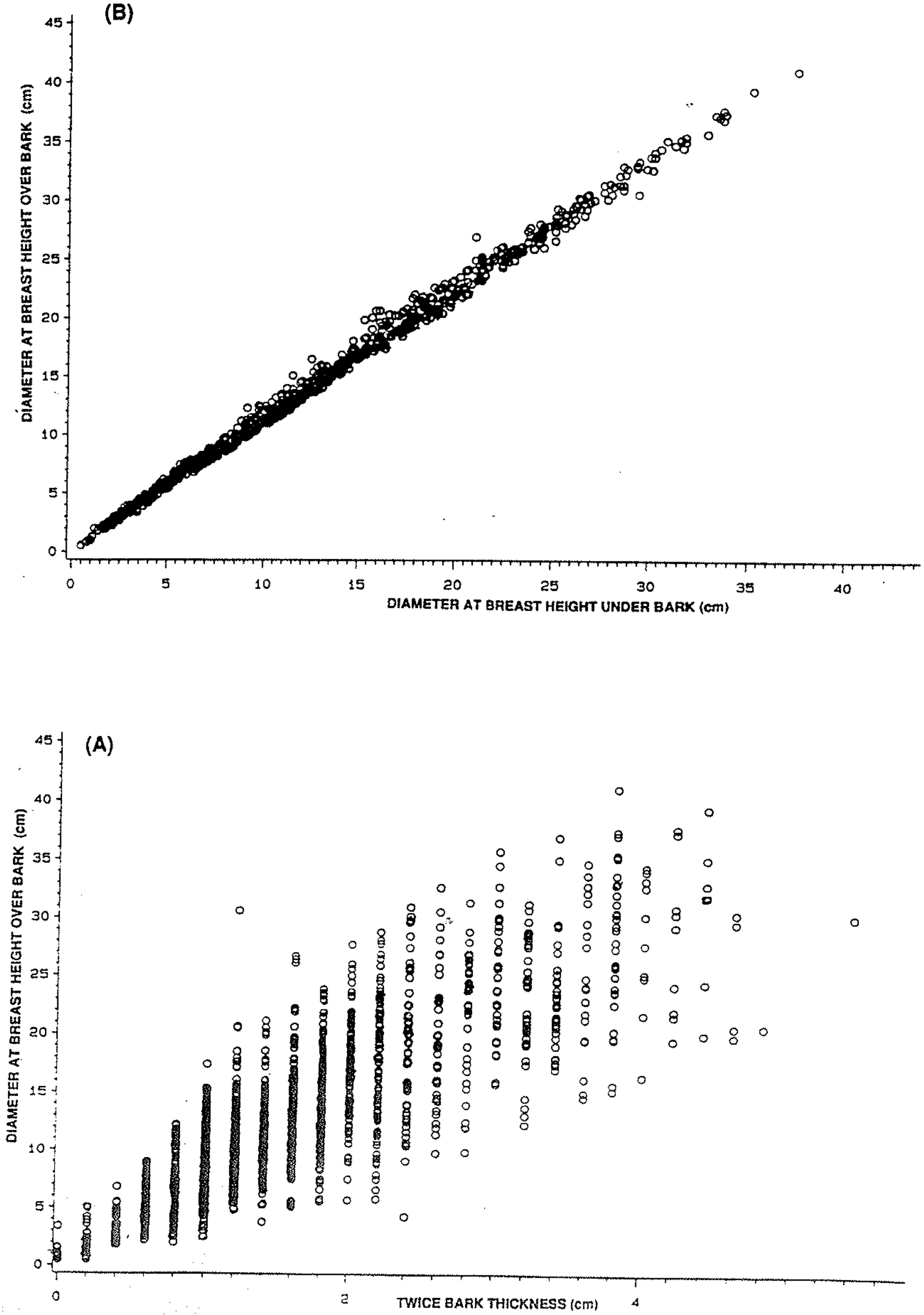
Table 36: Univariate statistics for variables used to develop a bark thickness equation

Variable (cm)	$\overline{X}$	St. Dev.	Range	n
DBHOB	14.1	7.1	0.5 - 41.9	3342
DBHUB	12.6	6.5	0.3 - 39.6	3336
2 x Bark thickness	1.5	0.8	0.0 - 5.4	3336

Table 37: Functional forms of candidate bark equations

Functional form	Citation	
dbhub $= \beta_1 Dbhob$	Powers	(1969)
dbhub $= \beta_0 + \beta_1 Dbhob$	Dolph	(1989)
dbhub $= Dbhob / (\beta_0 + \beta_1 Dbhob)$	West	(1979)
dbhub $= \beta_0 Dbhob^{\beta_1}$	Dolph	(1989)
dbhub $= \beta_0 + \beta_1 Dbhob + \beta_2 Dbhob^2 + ... + \beta_n Dbhob^n$	-	

Figure 26: (a) Twice bark thickness against diameter at breast height overbark and (b) diameter at breast height underbark by diameter at breast height overbark for the 3336 trees in the CALM data set.

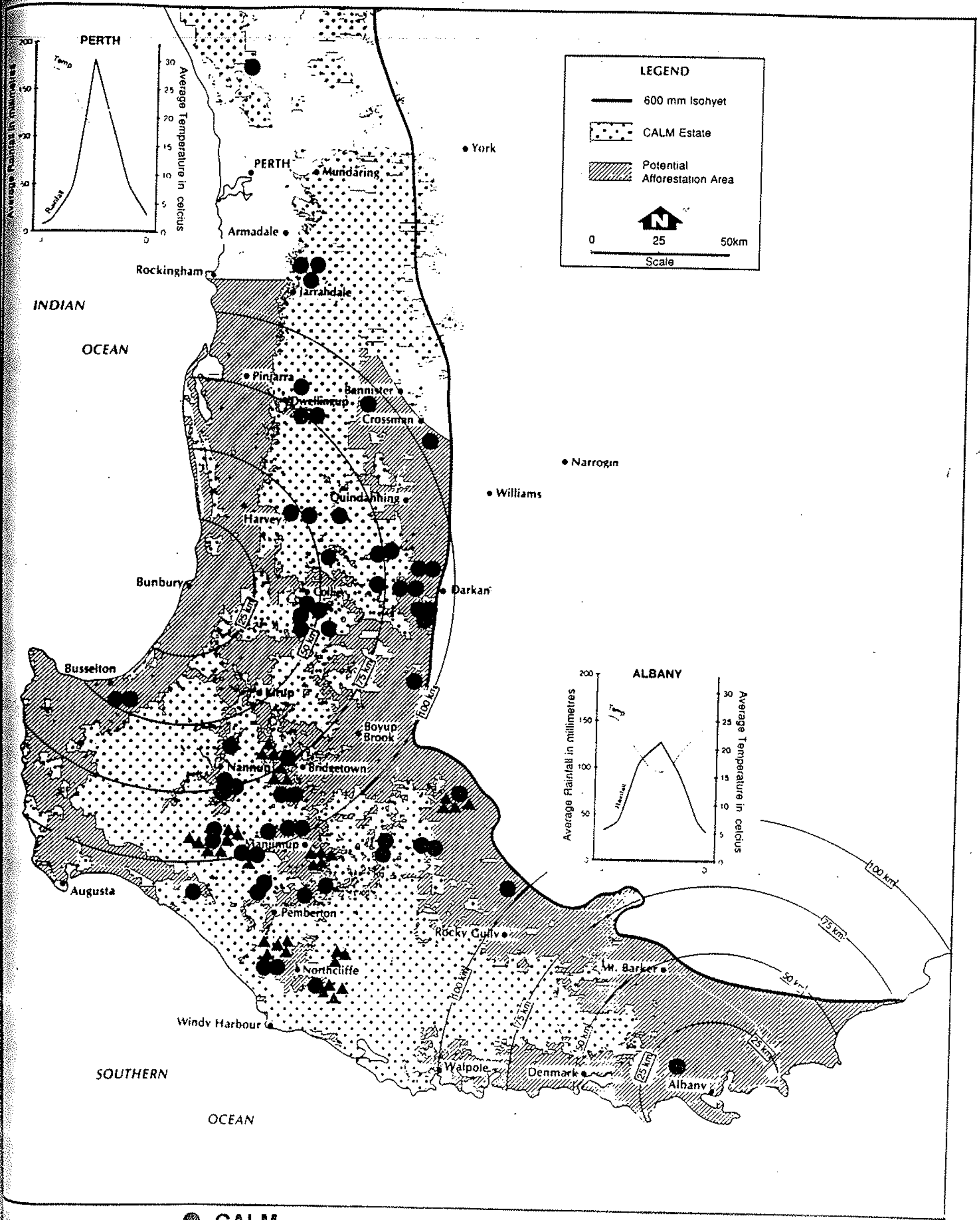


6.2.2.3.2 CALCULATION OF PLOT VOLUMES AND STAND VARIABLES

Using the derived dbhub equation (Eq.[42]) the dbhub of each tree in the Bunnings data set was estimated. The merchantable volumes under bark of each tree in both data sets were estimated using the single tree volume equation Eq.$[6]_{(CF)}$. Plot volumes were calculated as the sum of individual trees in the plots. Plot volume, basal area and stocking were converted to per hectare figures. The site index of each plot in the CALM data set was known from stem analysis (see Section 3.2.2.2). The site index of the plots in the Bunnings data set was estimated using the site index equation (Eq.[19]) (see Section 3.2.3.3.3). The top height of each plot in both datasets was defined as the average of the tallest one or two trees per plot depending upon plot size. In this study top height is defined as the average height of the tallest 40 stems ha^{-1}.

Univariate statistics for the combined data sets is given in Table 38, while the location of plots comprising the datasets are shown in Figure 27.

Figure 27: South west Western Australia showing the location of plots which comprise the Bunnings and CALM data sets.



● CALM

▲ Bunnings

Table 38: Univariate statistics for various stand attributes for the combined data set, data split by landuse history and for the calibration and validation data sets.

	TOTAL DATA SET			
	$\bar{X}$	SD	N	Range
V	84.8	60.0	201	6.5 - 296.0
BA	11.7	5.4	201	1.8 - 27.6
S	14.7	3.8	201	7.5 - 23.9
H	17.2	4.8	201	7.5 - 34.3
A	7.0	3.8	201	3.0 - 17.0
N	863.1	309.1	201	154.0 - 2000.0

	LAND HISTORY							
	PASTURE				UNIMPROVED			
	$\bar{X}$	SD	N	Range	$\bar{X}$	SD	N	Range
V	76.5	46.7	132	7.5 - 240.1	99.6	77.4	69	6.5 - 296.0
BA	11.6	4.7	132	2.2 - 23.5	12.2	6.4	69	1.7 - 27.6
S	16.3	3.4	132	10.0 - 23.9	11.5	2.3	69	7.5 - 16.2
H	16.2	4.1	132	7.5 - 27.6	19.0	5.5	69	9.1 - 34.3
A	5.2	1.9	132	3.0 - 13.0	10.3	4.2	69	4.0 - 17.0
N	899.1	254.3	132	154.0 - 1488.0	792.5	383.6	69	18.0 - 2000.0

	CALIBRATION				VALIDATION			
	$\bar{X}$	SD	N	Range	$\bar{X}$	SD	N	Range
V	85.5	61.9	171	6.5 - 296.0	81.2	48.7	30	14.9 - 210.5
BA	11.9	5.4	171	1.8 - 27.5	11.8	5.0	30	4.1 - 27.6
S	14.7	3.8	171	7.5 - 23.9	14.4	4.0	30	7.5 - 21.4
H	17.2	3.9	171	7.5 - 34.3	17.1	4.6	30	10.3 - 27.6
A	7.0	3.8	171	3.0 - 17.0	6.8	3.6	30	3.5 - 17.0
N	858.3	308.3	171	180.0 - 2000.0	901.0	315.9	30	154.0 - 1355.0

v = stand merchantable volume underbark per hectare, to a top diameter of 4cm (m^3ha^{-1})

BA = stand basal area at breast height under bark (m^2ha^{-1})

S = site index (m)

H = top height (m)

A = age (yrs)

N = stems per hectare

6.2.2.4 EXPLICIT PREDICTION OF CURRENT STAND YIELD AND CURRENT BASAL AREA

To develop the equations which estimate stand yield and basal area, the CALM and Bunnings data sets were combined. A set of candidate functional forms were selected after extensive correlation and graphical analysis of the data. The literature was also searched for functional forms which may be appropriate. The initial set of candidate functional forms were parametrized and equations compared on the basis of model statistics and the behaviour of the residuals. A reduced set of candidate equations were thus selected for further analysis.

The history of land use prior to the establishment of the plantation has been found to influence site productivity (Chapman 1938; Skinner and Attiwill 1981). As such the combined data set was split into two data sets reflecting the land use histories. Those plots in plantations established on pastured lands (PASTURE data set) were separated from those in plantations established on previously unimproved lands (UNIMPROVED data set). For each candidate equation the hypothesis, that the equation fits both data sets equally well and that no significant loss of explanatory power is experienced by combining the datasets, was tested.

Having determined whether the functional forms were applicable only to the individual land use data sets or the combined data set, the combined data set was separated into a calibration data set and a validation data set by randomly selecting 29 plots to be withheld from model parametrization. The integrity of land use histories was maintained in both the parametrization and validation data sets (Table 38).

6.2.2.5 PARAMETER ESTIMATION AND VALIDATION CRITERIA

Parameter estimation was via the standard criteria of ordinary least squares. Where logarithmic transformations were employed, Sprugel's (1983) correction factor was applied after transformation of the logarithmic terms to standard units (see Section 2.2.2.3). For nonlinear equations the parameters were estimated via nonlinear ordinary least squares using the derivative free secant method of Ralston and Jennrich (1979). The final candidate equations were validated via the assumptions and methods presented in Chapter 5.

6.2.3 RESULTS

6.2.3.1 ESTIMATION OF DIAMETER INSIDE BARK AT BREAST HEIGHT

An initial set of candidate equations were parametrized and their model statistics and residual patterns examined. A final set of three candidate functional forms were selected in this manner for further examination. Table 39 presents the model coefficients for the three equations after their parameters were estimated using the calibration dataset. Table 40 presents the validation statistics of the final equations after they were applied to the validation dataset.

On the basis of parametrization and validation statistics the fourth order polynomical (Eq.[42]) was found to be the more appropriate. This equation was the least biased and most precise of the final candidate equations. It is noteworthy that weighting the equations did not improve the model statistics or validation statistics of any of the equations.

Table 39: Equations for estimating the diameter of breast height underbark from diameter at breast height overbark

Equation	Model			r^2	n
Eq. [41]	Dbhub	=	0.89 Dbhob	0.9987	3228
Eq. [41]wt	Dbhub	=	0.89 Dbhob	0.9974	3228
Eq. [42]	Dbhub	=	0.83 Dbhob + 0.0078 $Dbhob^2$ -0.00032 $Dbhob^3$ + 0.0000044 $Dbhob^4$	0.9988	3228
Eq. [42]wt	Dbhub	=	0.84 Dbhob + 0.0066 $Dbhob^2$ -0.00026 $Dbhob^3$ + 0.0000035 $Dbhob^4$	0.9976	3228
Eq. [43]	Dbhub	=	0.85 $Dbhob^{1.017}$	0.9941[a]	3228
Eq. [43]wt	Dbhub	=	0.84 $Dbhob^{1.022}$	0.9955[a]	3228

[a] r^2 values estimated by (1 - a/b)

where a = residual sums of squares
b = corrected total sums of squares

Table 40: Validation statistics derived from the validation of three bark equations (n = 113)

Equation	$\overline{D}$	D_{SD}
Eq. [41]	-0.08	0.55
Eq. [41]wt	0.04	0.56
Eq. [42]	0.04	0.54
Eq. [42]wt	-0.04	0.55
Eq. [43]	-0.05	0.55
Eq. [43]wt	-0.06	0.56

6.2.3.2 EXPLICIT PREDICTION OF CURRENT YIELD

Univariate statistics for stand level attributes are given in Table 38. After extensive graphical and correlation analysis and a review of the literature, a number of functional forms were either derived or selected for further analysis. These functional forms were parametrized using the total data set and are presented along with their adjusted r^2 values in Table 41. After examination of the model statistics and the patterns of the residuals for each candidate equation, the log/log transformation and Schumacher's variable density yield equation (Pienaar and Shiver 1986) were selected for further examination.

Table 41: Functional forms and adjusted r^2 for candidate equations used for the explicit prediction of current yield

Equation			Adj r^2
v	=	$\beta_0 + \beta_1 B + \beta_2 H + \beta_3 BH + \beta_4 A^{\beta 5}$	0.9777
v	=	$\beta_0 SB (1 - e^{-\beta 1 A})^{\beta 2}$	0.9779
Inv	=	$\beta_0 + \beta_1 A^{-1} + \beta_2 S + \beta_3 \mathit{In}\, B$	0.9894
Inv	=	$\beta_0 + \beta_1 H + \beta_3\ \mathit{In}\, B$	0.9873
v	=	$\beta_0 + \beta_1 BH$	0.9718
v	=	$\beta_0 + \beta_1 B^2H$	0.9446
Inv	=	$\beta_0 + \beta_1 \mathit{In}B + \beta_2 \mathit{In}\, H$	0.9903
v	=	$\beta_0 + \beta_1 B + \beta_2 H + \beta_3 BH$	0.9772
Inv	=	$\beta_0 + \beta_1 \mathit{In}\, B$	0.9473
Inv	=	$\beta_0 + \beta_1 \mathit{In}\, H$	0.7636

$\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ = paramaters to be estimated
In = natural logarithm
B = basal area (m^2ha^{-1})
H = top height (m)
S = site index (m)

Prior to final model formulation the log/log transformation and Schumacher's equations were parametrized using the PASTURE, UNIMPROVED and TOTAL datasets, to test the hypothesis that the functional forms and their parameters are equally applicable to both sources of data. In both cases the hypothesis that the equations are equally applicable to either data source was accepted ($p < 0.001$). Consequently, the equations were parametrized using the calibration data set (Table 38) which is comprised of data from both land use histories. The final equations are as follows:

$$\ln V = -0.99 + 0.93 \ln H + 1.092 \ln B \qquad \text{Eq. [44]}$$

Adj r^2 = 0.9920 n = 171 RSS = 0.867 F value = 10539

$$\ln V = 1.44 - 3.97A^{-1} + 0.044S + 1.19 \ln B \qquad \text{Eq. [45]}$$

Adj r^2 = 0.9852 n = 171 RSS = 1.57 F value = 3781

All terms in both equations were significant ($p < 0.0001$) and no heteroscedasticity among residuals was detected.

The use of Eq.[44] or Eq.[45] requires the basal area of the stand to be known or estimated. This will not be possible for some applications of the results of this study. As such an equation which explicitly predicts current yield from stand attributes other than basal area was sought. Again candidate equations were tested for their applicability to both the PASTURE and UNIMPROVED data sets. In all cases the hypothesis that the candidate equations were equally applicable to both datasets was rejected ($p > 0.001$). Consequently, separate equations were derived and parametrized for the PASTURE and UNIMPROVED data using the calibration data set (Table 38).

The two best equations are as follows:

If the stand is established on pastured lands -

$$lnV = 4.25 + 0.00073N - 29.58H^{-1} + 0.043S + 0.087A \quad \text{Eq. [46]}$$

Adj r^2 = 0.8610 n = 112 RSS = 7.90 F value = 172

If the stand is established on unimproved lands -

$$lnV = -4.35 + 2.88\ln H + 0.001N - 493A^{-1} \quad \text{Eq. [47]}$$

Adj r^2 = 0.8405 n = 59 RSS = 7.26 F value = 103

It is noteworthy that not only do the parameters of any particular functional form require re-estimation for both sources of data, but in this case the best fits were obtained via the application of different functional forms to both sources of data.

6.2.3.3 EXPLICIT PREDICTION OF CURRENT BASAL AREA

After graphical and correlation analysis and a review of the literature, a number of functional forms were either derived or selected for further analysis. The candidate functional forms were parametrized using the PASTURE, UNIMPROVED and TOTAL data sets, to test the hypothesis that the functional forms and their parameters are equally applicable to both sources of data. In all cases the hypothesis that a single parametrized equation is equally applicable to either data source was rejected ($p > 0.001$). Consequently, individual equations were fitted to both the PASTURE and UNIMPROVED components of the calibration data set.

The final equations for the explicit prediction of current basal area are as follows;

If the stand was established as pastured lands -

$$lnB = 1.85 - 15.H^{-1} + 0.00068N + 0.0385S + 0.06A \quad \text{Eq. [48]}$$

Adj r^2 = 0.7555 n = 112 RSS = 7.11 F value = 129

If the stand was established on unimproved lands -

$$lnB = 3.84 - 28.46H^{-1} + 0.00085N - 4.73A^{-1} \qquad \text{Eq. [49]}$$

Adj r^2 = 0.7367 n = 59 RSS = 6.35 F value = 64.43

All terms in both equations were significant ($p < 0.005$) and no heteroscedasticity among residuals was detected.

6.2.3.4 VALIDATION OF PREDICTIVE EQUATIONS

Equations Eq.[44], Eq.[46], Eq.[47], Eq.[48], and Eq.[49] were validated by predicting stand volume or the basal area of the plots comprising the validation data set (Table 38). The mean residual, the standard deviation of the residuals, and the r-squared statistic, from a linear regression analysis using the observed value as the dependent variable for each of the above equations, are given in Table 42. Eq.[44] was also validated using a predicted basal area as input. The basal area was predicted using Eqs.[48] and [49]. The validation statistics stemming from this approach are included in Table 42.

Validation residuals were found to be random and showed no pattern with any stand attribute. On the basis of the validation statistics it is recommended that for the prediction of current stand yield Eq.[44] is used. However, where basal area is not available Eqs.[46] and [47] should be used. Figures 28 and 29 show the predicted stand yields with the actual stand yields for both approaches. For the prediction of stand basal area equations Eqs. [48] and [49] yield acceptable results. All equations yield validation statistics which are within practically acceptable limits.

Table 42: Validation statistics for equations which explicitly predict stand volume or basal area (n = 29).

Equation	Statistic		
	$\overline{D}$	D_{SD}	r^2
Yield equations			
Eq. [44]	-0.01	6.78	0.9750
Eq. [46] + Eq. [47][1]	-0.75	17.42	0.8374
Eq. [44][2]	-1.42	17.38	0.8358
Basal area equations			
Eq. [48] + Eq. [49]	0.04	2.35	0.6781

1 : prediction of current yield using equations specific to land use histories

2 : prediction of current yield using Eq. [44] where basal area is predicted using Eq. [48] and Eq. [49]

$\overline{D}$ = mean of residuals

D_{SD} = standard deviation of residuals

r^2 = the r-squared statistic from a linear regression, using the observed value as the dependent variable

Figure 28: **Observed stand volume by predicted stand volume derived from the validation of the stand volume equation Eq. [44].**

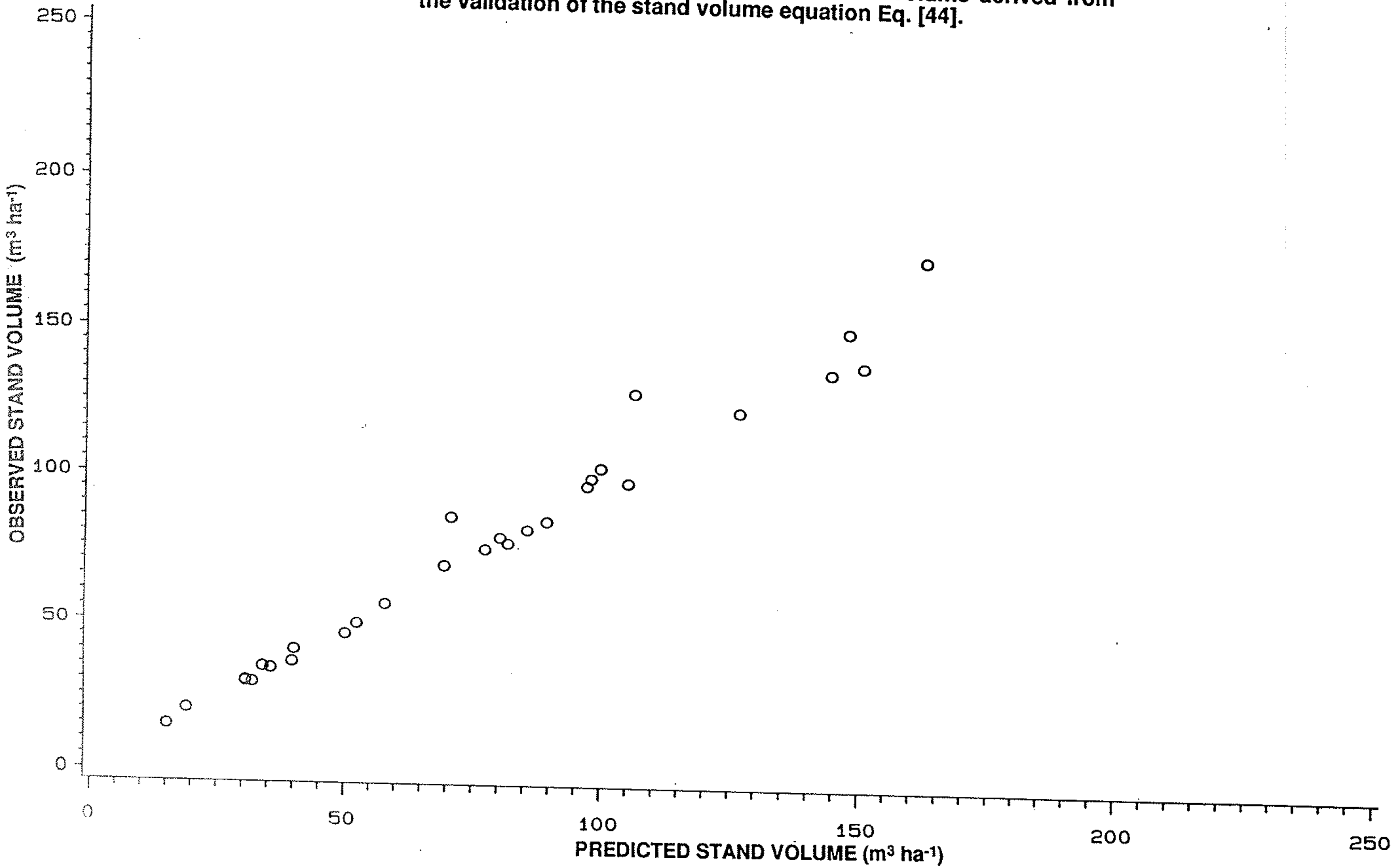
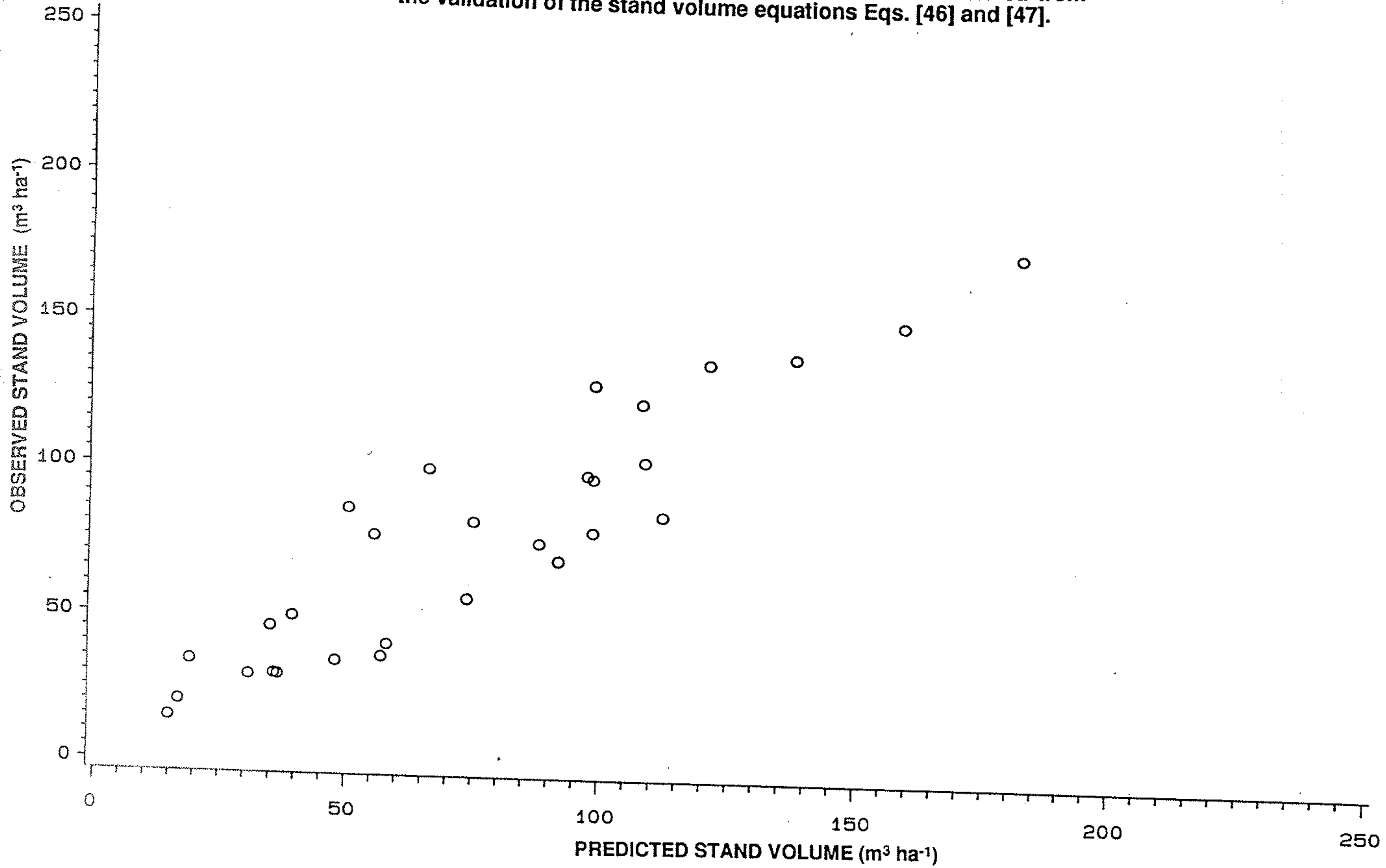


Figure 29: **Observed stand volume by predicted stand volume derived from the validation of the stand volume equations Eqs. [46] and [47].**



6.2.4 DISCUSSION

In the recent literature most growth and yield simulation studies have tended to follow the single tree distance independent philosophies or a biology of production approach (Botin *et al.* 1972). Such studies are most applicable to complex multi-species, uneven-aged dynamic forests, such as those which occur in the Pacific North-west of the United States. Under such conditions growth and yield simulation methods need to be consistent and robust to a wide range of forest conditions and have the capacity to predict the yield of the range of products harvested from such forests (Stage 1973; Wykoff *et al.* 1982; Wykoff 1986; Wykoff 1990).

Alternatively, where the forest of interest is an even-aged monoculture and particularly where a single product is harvested, there is no evidence to suggest that a stand level approach is less applicable. Such situations occur in the southern states of the United States and in the forests on which this study is based (Murphy and Farrar 1983; Burkhart *et al.* 1985; Matney *et al.* 1988; Pienaar *et al.* 1990; Walters *et al.* 1990). The stand level equations used in this study validate with levels of accuracy and precision acceptable for practical purposes.

The functional forms presented are similar to other studies which seek to explicitly predict current yield (c.f. Clutter *et al.* 1983; Burkhart *et al.* 1985). However, this study differs by accounting for the effect of land use history. Land use history is not a concern when using Eq.[44] to estimate yield. Here the effect is accounted for by using both basal area and top height as independent variables. However, when basal area is not available the common practise of estimating basal area, with equations such as Eq.[48] and Eq.[49] and substituting the value into a yield equation such as Eq.[44], did not validate as well as using the equations specific to a land use history (i.e. Eq.[46] and Eq.[47]) (Table 42).

The pasture or old-field effect has received some attention in the literature (see Section 4.3.4.3). However, little effort has been directed towards quantifying or incorporating the effect into growth and yield studies. For example, Clutter *et al.* (1976) and Burkhart *et al.* (1985) found no evidence that land use history influenced yield and concluded that, 'a single yield equation is adequate for predicting total cubic-foot volume yield across all site preparation methods and physiographic regions sampled'. It is noteworthy that the equation of Burkhart *et al.* (1985) did not include basal area as an independent variable.

On the other hand Amateis and Burkhart (1987), while studying the influence of land use history upon individual tree volumes, concluded that 'trees grown in old-field plantations are more conical in shape than trees from cutover-site plantations. This means that for trees of the same stump diameter and total height, old-field-plantation-grown trees have more volume'.

In this study it was observed that for two stands, of different land use histories, but of equivalent stocking, age and top heights, the stand established on pasture will have a larger basal area. Under such a situation it is obligatory to account for such differences when basal area and top height are not included in the same equation. In this study the strategy of applying separate equations to each source of data was adopted. The alternative strategy of incorporating dummy variables into a single equation, was not examined (e.g. Monserud 1984).

The effect of land use history on basal area is assumed to be nutritionally induced (Skinner and Attiwill 1981) as a result of nutrient accumulation under an agricultural regime (Lewis *et al.* 1987a,b). No evidence was found to suggest that top height is influenced by land use history. As such the use of site index and top heights as a measure of the potential of a site to yield produce is assumed to be sound. However, whether the site actually achieves the potential will be a function of many things, including the previous land use and nutritional status.

6.2.4.1 PREDICTION OF FUTURE YIELD

The explicit prediction of future yield requires the projection of the independent variables of the equations presented in this study, to some future point in time. A number of strategies are available to achieve this aim. Firstly, substitution of future values of H and B into Eq.[44] will give an estimate of future yield. However, such an approach requires a method to predict future values of H and B. Future values of H may be estimated using the equations derived in Chapter 3. However, because of the lack of basal area growth data, no method of predicting future values of stand basal area has been derived. An example of a stand level basal area growth function is given in Pienaar *et al.* (1990).

An alternative strategy for predicting future values of B is to substitute future values of H, N, S and A into Eqs.[48] and [49]. The predicted future value of B may then be substituted into Eq.[44]. However, this approach proved less accurate and less precise, when estimating current yield, than using Eqs.[46] and [47] directly. There is no reason to suspect that this result would not apply equally to the prediction of future yields. Therefore, the strategy of substituting future values of H, S, N and A into Eqs.[46] and [47] is recommended for the explicit prediction of future yields.

Future values of H may be estimated using Eq.[11] presented in Chapter 3. S is invariant over time and may be either measured directly when the stand is five years of age, estimated from measurements of stand top height and age using Eq.[19], also described in Chapter 3, or estimated using environmental variables using Eq.[28] presented in Chapter 4. A is simply nominated as the age of the stand at which the future estimate of yield is required. Future estimates of N are usually obtained from mortality functions (c.f. West 1981; Burkhart *et al.* 1985; Matney *et al.* 1987; Pienaar *et al.* 1990). However, the data available to this study are too limited to derive such functions. Also faced with the problem of limited data on which to base a mortality function, Goodwin and Candy (1986) employed the so called 3/2 self thinning rule (Drew and Flewelling 1979; White 1981) to their growth and yield model for *E. globulus*. In their study natural mortality was observed to have commenced at age 12 years. Similar patterns were observed in this study. Few, if any mortality was observed between the stand ages of two and ten years. As such, a mortality function is not critical for the prediction of future yields when dealing with such short rotation, even-aged monocultures.

The variation in stocking observed in the dataset (Table 38) is attributed to differences in planting densities and the quality of site preparation employed at planting. The quality of site preparation is observed to have a marked influence on the survival over the critical first summer, where most mortality occurs. It is therefore assumed that N will remain constant over the duration of the rotation once the stand has reached two years of age. The management practises of the Department of Conservation and Land Management stipulate that a minimum of 95% survival is required at a stand age of two years. Therefore, for the prediction of future yields, future values of N should be the current value if the stand is older than two years or 95% of the planting density for younger stands.

It is recognised that the lack of an empirical mortality function remains a weakness of this study. However, mortality is observed to have less influence on these fast growing, short rotation plantations, than on the longer term forests to which most growth simulators are applied. It is also recognized that a lack of data prevents the formal validation of the equations ability to predict future yields and that this also remains a weakness of this study and a priority for future research.

6.2.5 CONCLUSION

For the population of plantations sampled in this study it was concluded that:

(i) The current yield of plantation grown *E. globulus* is best estimated using Eq.[44] and that this single yield equation is adequate for estimating the total merchantable volume under bark, across all land use histories sampled;

(ii) Where basal area is not available as an independent variable, the current yield of plantation grown *E. globulus* is best estimated using Eqs.[46] and [47] which are specific to the two land use histories sampled;

(iii) Both methods of estimating current yield validate with levels of accuracy and precision within practically acceptable standards; and,

(iv) The explicit prediction of future yield is obtained by estimating the future values of the independent variables comprising Eqs.[44], [46] and [47], using the equations developed in Chapters 3 and 4. However, the accuracy and precision of such a course of action has not been determined formally.

6.2.6 REFERENCES

Alder, D. 1979. A distance-independent tree model for exotic conifer plantations in East Africa. *For. Sci.*, 25:59-71.

Amateis, R.L., and H.E. Burkhart. 1987. Tree volume and taper of loblolly pine varies by stand origion. *South. J. Appl. For.*, 11:185-189.

Amateis, R.L., H.E. Burkhart, and T.E. Burk. 1986. A ratio approach to predict merchantable yields of unthinned loblolly pine plantations. *For. Sci.*, 32:287-296.

Amateis, R.L., H.E. Burkhart, B.J. Greber, and E.E. Watson. 1984. A comparison of approaches for predicting multiple-product yields from weight scaling. *For. Sci.*, 30:991-998.

Arney, J.D. 1985. A modelling strategy for the growth projection of managed stands. *Can. J. For. Res.*, 15:511-518.

Bailey, R.L. 1980. Individual tree growth derived from diameter distribution models. *For. Sci.*, 26:216-236.

Bailey, R.L. 1981. Consideration of error structure in regressions to predict biomass. *In:* (C.A. Gresham (Ed.). Proc southern forest biomass workshop. Georgetown, S.C.

Bailey, R.L., and K.D. Ware. 1983. Compatible basal area growth and yield model for thinned and unthinned stands. *Can. J. For. Res.*, 13:563-571.

Bailey, R.L., and T.R. Dell. 1973. Quantifying diameter distributions with the Weilbull function. *For. Sci.*, 19:97-104.

Bain, L.J., and C.E. Antle. 1967. Estimation of parameters in the Weibull distribution. *Technometrics*, 9:621-627.

Beck, D.E. 1974. Predicting growth of individual trees in thinned stands of yellow-poplar. *In:* J. Fries(Ed). Growth models for tree and stand simulation. *Royal Coll. Res. Notes.*, No. 30. Stockholm.

Belcher, D.W., M.R. Holdaway, and G.J. Brand. 1982. A description of STEMS: the stand and tree evaluation and modelling system. *USDA Gen. Tech. Rep.* NC-79.

Bella, I.E. 1971. A new competition model for individual trees. *For. Sci.*, 17:364-372.

Bennett, F.A. 1970. Variable-density yield tables for managed stands of natural slash pine. *USDA Res. Note.* SE-141.

Bennett, F.A., and J.L. Clutter. 1968. Multiple-product yield estimates for unthinned slash pine plantations - pulpwood, sawtimber, gum. *USDA Res. Pap.* SE-35.

Bliss, C.I., and K.A. Reinker. 1964. A log-normal approach to diameter distributions in even-aged stands. *For. Sci.*, 10:350-360.

Borders, B.E., 1989. Systems of equations in forest stand modelling. *For. Sci.*, 35:548-556.

Borders, B.E., and R.L. Bailey. 1986. A compatible system of growth and yield equations for slash pine fitted with restricted three-stage least squares. *For. Sci.*, 32:185-201.

Borders, B.E., and W.D. Patterson. 1990. Projecting stand tables: A comparison of the Weibull diameter distribution method, a percentile-based projection method, and a basal area growth projection method. *For. Sci.*, 36:413-424.

Borders, B.E., R.A. Souter, R.H.Bailey, and K.D. Ware. 1987. Percentile-based distributions characterize forest stand tables. *For. Sci.*, 33:570-576.

Botkin, D.B., J.E. Janak, and J.R. Wallis. 1972. Some ecological consequences of a computer model of forest growth. *J. Ecol.*, 60:849-872.

Bruce, D., and L.C. Wensel. 1988. Modelling forest growth: approaches, definitions, and problems. *In:* A.R.Ek, *et al.* (Eds.). Forest growth modelling and prediction. *USDA Gen. Tech. Rep.*, NC-120.

Bruner, H.D., and J.W. Moser. 1973. A Markov chain approach to the prediction of diameter distributions in uneven-aged forest stands. *Can. J. For. Res.*, 3:409-417.

Buchman, R.G., S.P.Pederson, and N.R. Walters. 1983. A tree survival model with application to species of the Great Lakes Region. *Can. J. For. Res.*, 13:610-608.

Burk, T.E., and H.E. Burkhart. 1984. Diameter distributions and yields of natural stands of loblolly pine. *Virginia Politech Inst and State Univ.* FWS-1-84.

Burkhart, H.E. 1977a. Stand modelling for radiata pine in New Zealand. *N.Z.J. For.*, 9:354-371.

Burkhart, H.E. 1977b. Cubic-foot volume of loblolly pine to any merchantable top limit. *South. J. Appl. For.*, 1:7-9.

Burkhart, H.E., and M.R. Strub. 1974. A model for simulation of planted loblolly pine stands. In: J. Fries (Ed.). Growth models for tree and stand simulation. *Royal. Coll. For. Res. Note.*, No. 30. Stockholm.

Burkhart, H.E., and P.T. Sprinz. 1984. Compatible cubic volume and basal area projection equations for thinned old-field loblolly pine plantations. *For. Sci.*, 30:86-93.

Burkhart, H.E., D.C. Cloerem, and R.L. Amateis. 1985. Yield relationships in unthinned loblolly pine plantations on cutover, site-prepared lands. *South. J. Appl. For.*, 9:84-91.

Burkhart, H.E., R.C. Parker, M.R. Strub, and R.G. Octerwald. 1972. Yields of old-field loblolly pine plantations.
Va. Polytech. Inst. and State Univ., Div. For. and Wildl. Resour., FWS-3-72.

Chambers, C.J. 1980. Empirical growth and yield tables for the Douglas-fir zone.
State of Washington, Dept. of Nat. Res., Rep No. 41.

Chapman, R.A. 1938. The effect of origion of stand on the site index of longleaf pine.
J. For., 36:75-77.

Chojnacky, D.C. 1988. Modeling volume growth for Arizona's pinyon-juniper forest.
In: A.R. Ek, *et al.* (Eds.). Forest growth modelling and prediction.
USDA Gen. Tech. Rep. NC-120.

Clutter, J.L. 1963. Compatible growth and yield models for Loblolly pine.
For. Sci., 9:354-371.

Clutter, J.L., and E.P. Jones. 1980. Prediction of growth after thinning in old-field slash pine plantations. *USDA Res. Pap.* SE-217.

Clutter, J.L., and B.J. Allison. 1974. A growth and yield model for *Pinus radiata* in New Zealand. *In:* J. Fries (Ed.). Growth models for tree and stand simulation.
Royal Coll. For. Res. Notes. No. 30. Stockholm.

Clutter, J.L., J.C. Fortson, and B.D. Shiver. 1976. Some relationship between soils and site preparation in flatwoods slash pine plantations. *In*: Proc., Sixth South. For. Soils Workshop, Charleston, Sc.

Clutter, J.L., J.C. Fortson,L.V. Pienaar, G.H. Brister, and R.L. Bailey. 1983. Timber management: A quantitative approach. John Wiley and Son. New York.

Curtis, R.O. 1970. Stand density measures: an interpretation. *For. Sci.*, 16:403-414.

Curtis, R.O. 1973. A method of estimation of gross yield of Douglas-fir. *For. Sci .Monogr*, 13.

Daniels, R.F. 1976. Simple competition indices and their correlation with annual loblolly pine tree growth. *For. Sci.*, 22:454-456.

Daniels, R.F., and H.E. Burkhart. 1975. Simulation of individual tree growth and stand development in managed loblolly pine plantations.
Div. For. Wildl. Res, Va Polytechnic Inst. and State Univ. FWS-5-75.

Dolph, K.L. 1989a. Height-diameter equations for young-growth red fir in California and southern Oregon. *USDA Res. Note* PSW -408.

Dolph, K.L. 1989b. Nonlinear equations for predicting diameter inside bark at breast height for young-growth red fir in California and southern Oregon. *USDA Res. Note* PSW-409.

Drew, T.J., and J.W. Flewelling. 1979. Stand density management: an alternative approach and its application to Douglas fir plantations. *For. Sci.*, 25:518-532.

Dubey, S.D. 1967. Some percentile estimators for Weibull parameters.
Technometrics, 9:119-129.

Ek, A.R. 1974. Nonlinear models for stand table projection in northern hardwood stands.
Can. J. For. Res. 4:23-27.

Ek, A.R., and R.A. Monerud. 1974. FOREST: A computer model for simulating the growth and reproduction of mixed species forest-stands.
Univ. Wisc. Sch. Nat. Resources Res. Rep. RZ635.

Ek, A.R. and R.A. Monerud. 1979. Performance and comparison of stand growth models based on individual tree and diameter-class growth. *Can. J. For. Res.*, 9:231-244.

Evert, F. 1981. A model for regular mortality in unthinned white spruce plantations.
For. Chron., 81:77-79.

Ferguson, I.S., and J.W. Leech. 1978. Generalized least squares estimation of yield functions.
For. Sci., 24:27-42.

Fomby, J.B., Hill, R.C., and S.R. Johnson. 1984. Advanced econometric methods. Springer-Verlag, New York.

Furnival, G.M., and R.W. Wilson. 1971. Systems of equations for predicting forest growth and yield. *Statist. Ecol.*, 3:43-55.

Garcia, O. 1981. Simplified method of moments for the Weibull distribution.
N.Z. J. For. Sci., 11:304-306.

Garcia, O. 1988. Growth modelling - a (re)view. *N.Z. For.*, 33:14-17.

Gasana, J.K. 1983. Site factors affecting growth of Maiden's gum, *Eucalyptus globulus* Labill. subsp. *Maidenii* (F. Muell) kirkpatrick in Rwana. Ph.D dissertation, University of Idaho.

Goodwin, A.N., and S.G. Candy. 1986. Single-tree and stand growth models for a plantation of *Eucalyptus globulus* Labill. in Northern Tasmania. *Aust. For. Res.*, 16:131-144.

Gordon, A. 1983. Estimating bark thickness of *Pinus radiata*. *N.Z. J. For. Sci.*, 13:340-353.

Gregoire, T.G. 1987. Generalized error structure for forestry yield models. *For. Sci.*, 33:423-444.

Hafley, W.L., and H.T. Schreuder. 1977. Statistical distributions for fitting diameter and height data in even-aged stands. *Can. J. For. Res.*, 7:481-487.

Hamilton, D.A. 1974. Event probabilities estimated by regression. *USDA Res. Pap.* INT-152.

Hamilton, D.A., and B.M. Edwards. 1976. Modelling the probability of individual tree mortality. *USDA Res. Pap.* INT -185.

Hyink D.M. and J.W. Moser. 1983. A generalized framework for projecting forest yield and stand structure using diameter distributions. *For. Sci.*, 29:85-95.

Knowe, S.A., and G.S. Foster. 1989. Application of growth models for simulating genetic gain of loblolly pine. *For. Sci.*, 35:211-228.

Krajicek, J.E., K.A. Brinkham, and S.F. Gingrich. 1961. Crown competition as a measure of density. *For. Sci.*, 7:35-42.

Leak, W.B. 1965. The J-shaped probability distribution. *For. Sci.*, 11:405-409.

Lee, Y.J. 1971. Predicting mortality for even-aged stands of loblolly pine. *For. Chron.*, 47:29-32.

Lenhart, J.B., and J.L. Clutter. 1971. Cubic-foot yield tables for old-field loblolly pine plantations in the Georgia Piedmont. *Ga. For. Res. Counc. Rep.* 22, series 3.

Lewis, D.C., A.L. Clarke, and W.B. Hall. 1987a. Accumulation of plant nutrients and changes in soil properties of sandy soils under fertilized pasture in south-eastern South Australia. I. Phosphorus. *Aust. J. Soil. Res.*, 25:193-202.

Lewis, D.C.,A.L. Clarke, and W.B. Hall. 1987b. Accumulation of plant nutrients and changes in soil properties of sandy soils under fertilized pasture in south-eastern South Australia. II. Total sulfer and nitrogen, organic carbon and pH. *Aust. J. Soil. Res.*, 25:203-210.

Lin, J.Y. 1974. Stand growth simulation models for Douglas-fir and Western Henlock in the North-western United States. *In:* J. Fries (Ed.) Growth models for tree and stand simulation. *Royal. Coll. For. Res. Notes,* No. 30. Stockholm.

Lowell, K.E. 1988. Modelling the growth and yield effects of forest fertilization on radiata pine plantations in New Zealand. *In:* A.R. Ek *et al.* (Eds.) Forest growth modelling and prediction. *USDA Gen. Tech. Rep.* NC-120.

Lowell, K.E., and R.J. Mitchell. 1987. Stand growth projection: simultaneous estimation of growth and mortality using a single probabilistic function. *Can. J. For. Res.*, 17:1466-1470.

Matney, T.G., and A.D. Sullivan. 1982. Compatible stand and stock tables for thinned and unthinned loblolly stands. *For. Sci.*, 28:161-171.

Matney, T.G., J.R. Ledbetter, and A.D. Sullivan. 1987. Diameter distribution yield systems for unthinned cutover site-prepared slash pine plantations in southern Mississippi. *South. J. Appl. For.*, 11:32-36.

Matney, T.G., A.D. Sullivan, and J.R. Ledbetter. 1988. Stand level cubic-foot volume ratio equations for planted loblolly pine on site-prepared land in the mid-south Gulf Coastal Plain. *South. J. Appl. For.*, 12:7-11.

Monserud, R.A. 1976. Simulation of forest tree mortality. *For. Sci.*, 22:438-444.

Monserud, R.A. 1979. Relations between inside and outside bark diameter at breast height for Douglas-fir in northern Idaho and northwestern Montana. *USDA Res. Note.* INT-266.

Monserud, R.A. 1984. Height growth and site index curves for inland Douglas-fir based on stem analysis and forest habitat type. *For. Sci.*, 30:943-965.

Moser, J.W. 1980. Historical chapters in the development of modern forest growth and yield theory. *In:* K.M. Brown and F.R. Clarke (Eds.). Forecasting forest stand dynamics. Proceedings of the workshop held at the school of forestry, Lakeshead University, Thunder Bay, Ontario. June 24-25th 1980.

Munro, D.D. 1974. Forest growth models - a prognosis. *In:* J. Fries (Ed.). Growth models for tree and stand simulation. *Royal Coll. For. Res. Note* No. 30. Stockholm.

Murphy, P.A., and H.S. Sternitzke. 1979. Growth and yield estimation for loblolly pine in the West Gulf. *USDA Res. Pap.* S0-154.

Murphy, P.A., and R.C. Beltz. 1981. Growth and yield of shortleaf pine in the West Gulf region. *USDA Res. Pap.* S0-169.

Murphy, P.A,. and R.M. Farrar. 1983. Sawtimber volume predictions for uneven-aged loblolly - shortleaf pine stands on average sites. *South. J. Appl. For.*, 7:45-50.

Nelson, T.C. 1964. Diameter distribution and growth of loblolly pine. *For. Sci.*, 10:105-115.

Oliver, W.W. 1979. Fifteen-year growth patterns after thinning a ponderosa-Jeffrey pine plantation in north-eastern California. *USDA Res. Pap.* PSW-141.

Opie, J.E. 1968. Predictability of individual tree growth using definitions of competing basal area. *For. Sci.*, 14:314-323.

Payandeh, B. 1990. Equations describing attributes of major tree species in north central Ontario. *For. Ecol. Manage.*, 36:245-252.

Payandeh, B., and A.R. Ek. 1986. Distance methods and density estimators. *Can. J. For. Res.*, 16:918-924.

Peden, L.M., J.S. Williams, and W.E. Frayer. 1973. A Markov model for stand projection. *For. Sci.*, 19:303-314.

Pienaar L.V., and W.H. Harrison. 1988. A stand table projection approach to yield prediction in unthinned even-aged stands. *For. Sci.*, 34:804-808.

Pienaar, L.V., and B.D. Shiver. 1981. Survival functions for site prepared slash pine plantations in the flatwoods of Georgia and northern Florida. *S. J. App. For.*, 5:59-62.

Pienaar, L.V., and B.D. Shiver. 1986. Basal area prediction and projection equations for pine plantations. *For. Sci.*, 32:626-633.

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. *For. Sci.*, 19:2-22.

Pienaar, L.V., H.H. Page, and J.W. Rheney. 1990. Yield prediction for mechanically site-prepared slash pine plantations. *South. J. Appl. For.*, 14:104-109.

Pindyck, R.S., and D.L. Rubinfeld. 1981. Econometric models and economic forecasts. Ed. 2. McGraw-Hill, New York.

Powers, R.F. 1969. Estimating past diameters of ponderosa pine in northern California. *USDA Res. Note* PSW-194.

Ralston, M.L., and R.I. Jennrich. 1979. DUD a derivative-free algorithm for nonlinear least squares. *Technometrics*, 1:7-14.

Rennolls, K., and A. Pearce. 1986. Flow models of mortality and yield for unthinned forest stands. *Forestry*, 59:47-58.

Sadig, R.A., and A.F. Beckwith. 1986. Modified Australian stand growth formula for managed plantations. *Aust. For. Res.*, 16:301-312.

Schlich, W. 1925. Schlich's manual of forest. Vol. III. Forest management. 5th ed. Bradbury, Agnew and Co. London.

Schmelz, D.V., and A.A. Lindsey. 1965. Size-class structure of old-growth forests in Indiana. *For. Sci.*, 11:258-264.

Schreuder, H.T., W.L. Hafley, and F.A. Bennett. 1979. Yield prediction for unthinned natural slash pine stands. *For. Sci.*, 25:25-30.

Schumacher, F.X. 1939. A new growth curve and its application to timber yield studies. *J. For.*, 37:819-820.

Shiver, B.D. 1988. Sample size and estimation methods for the Weibull distribution for unthinned slash pine plantation diameter distributions. *For. Sci.*, 34:809-814.

Skinner, M.F., and P.M. Attiwill. 1981. The productivity of pine plantations in relation to previous land use. I. Growth responses in agricultural and forest soils. *Plant and Soil*, 60:161-176.

Smalley, G.W., and R.L. Bailey. 1974. Yield tables and stand structure for shortleaf pine plantations in Tennessee, Alabama, and Georgia Highlands. *USDA Res. Pap.* S0-97.

Sprugel, D.G. 1983. Correcting for bias in log-transformed allometric equations. *Ecology*, 64:209-1210.

Stage, A.R. 1973. Prognosis model for stand development. *USDA Res. Pap.* INT-137.

Stiell, W.M. 1974. An application of stand volume and tree mortality models in yield table construction. *In:* J. Fries (Ed.) Growth models for tree and stand simulations. *Roy. Coll. For. Res. Note*, No. 30.

Sullivan, A.D., and J.L. Clutter. 1972. A simultaneous growth and yield model for loblolly pine. *For. Sci.*, 18:76-86.

Tesch, S.D. 1981. The evolution of forest yield determination and site classification. *For. Ecol. Manage.*, 3:169-182.

Tham, A. 1988. Estimate and test frequency distributions with the Johnson S_B function from stand parameters in young mixed stands after different thinning treatments. *In:* A.R. Ek *et al.* (Eds.) Forest growth modelling and prediction. *USDA Gen. Tech. Rep.* NC-120.

Thompson, H.R. 1956. Distribution of distance to n^{th} neighbour in a population of randomly distributed individuals. *Ecology*, 37:391-394.

Tomé, M. 1988. A growth and yield model for eucalypt plantations in Portugal. *In:* A.R. Ek *et al.* (Eds.). Forest growth modelling and prediction. *USDA Gen. Tech. Rep.* NC-120.

Van Deusen, P.C., A.D. Sullivan, and T.G. Matney. 1981. A prediction system for cubic foot volume of loblolly pine applicable through much of its range. *South. J. Appl. For.*, 5:186-189.

Varmola, M. 1988. A stand model for simulating the early development of Scots pine cultures in Finland. *In:* A.R. Ek *et al.* (Eds.). Forest growth modelling and prediction. *USDA Gen. Tech. Rep.* NC-120.

Walters, D.K., A.R. Ek, and D. Czysz. 1990. Construction of variable-density empirical yield equations from forest management inventory data. *North. J. Appl. For.*, 7:110-113.

Weibull, W. 1951. A statistical distribution function of wide applicability. *J. Appl. Mech.*, 18:293-297.

West, G.G., R.L. Knowles, and A.R. Koehler. 1982. A model to predict the effects of pruning and early thinning on the growth of radiata pine. *N.Z. For. Serv., F.R.I. Bull.*, No. 5

West, P.W. 1979. Estimation of height, bark thickness and plot volume in regrowth eucalypt forest. *Aust. For. Res.*, 9:295-308.

West, P.W. 1981. Simulation of diameter growth and mortality in regrowth eucalypt forest of southern Tasmania. *For. Sci.*, 27: 603-616.

West, P.W. 1982a. Functions to estimate tree height and bark thickness of Tasmanian regrowth eucalypts. *Aust. For. Res.*. 12:183-190.

West, P.W. 1982b. Comparison of stand density measures in even-aged regrowth eucalypt forest of southern Tasmania. *Can. J. For. Res.*, 13:22-31.

White, J. 1981. The allometric interpretation of the self-thinning rule. *J. Theor. Biol.*, 89:475-500.

Wykoff, W.R. 1986. Supplement to the user's guide for the stand prognosis model: Version 5.0. *USDA. Gen. Tech. Rep.* INT 218.

Wykoff, W.R. 1990. A basal area increment model for individual conifers in the northern Rocky mountains. *For. Sci.*, 36:1077-1104.

Wykoff, W.R., N.L. Crookston, and A.R. Stage. 1982. User's guide to the stand prognosis model. *USDA Gen. Tech. Rep.* INT-133.

Zakrzewski, W.T. and I.E. Bella. 1988. Two new height models for volume estimation of lodgepole pine stands. *Can. J. For. Res.*, 18:195-201.

Zanakis, S.H. 1979. A simulation study of some simple estimators for the three parameter Weibull distribution. *J. Stat. Comput. Simul.*, 9:101-116.

Zarnoch, S.J., and T.R. Dell. 1985. An evaluation of percentile and maximum likelihood estimators of Weibull parameters. *For. Sci.*, 31:260-268.

Zellner, A. 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *J. Am. Statist. Assoc.*, 57:348-368.

Zutter, B.R., R.G. Oderwald, P.A. Murphy, and R.M. Farrar. 1986. Characterizing diameter distributions with modified data types and forms of the Weibull distribution. *For. Sci.*, 32: 37-48.